



HAL
open science

SASH1, a new potential link between smoking and atherosclerosis

Henri Weidmann

► **To cite this version:**

Henri Weidmann. SASH1, a new potential link between smoking and atherosclerosis. Cardiology and cardiovascular system. Université Pierre et Marie Curie - Paris VI; Universität Hamburg, 2015. English. NNT : 2015PA066267 . tel-01242976

HAL Id: tel-01242976

<https://theses.hal.science/tel-01242976>

Submitted on 14 Dec 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Pierre et Marie Curie

Ecole Doctorale 394: Physiology and physiopathology

UMRS 1166 ICAN Institute, Insitute of Cardiometabolism And Nutrition

Equipe 1: Genomics and physiopathology of cardiovascular diseases

SASH1, a new potential link between smoking and atherosclerosis

By Henri Weidmann

To obtain the Degree of Doctor of physiology and physiopathology of the University
Pierre et Marie Curie

Directed by Dr Ewa Ninio

Co-directed by Pr Tanja Zeller

Publicly presented on the 23th of September 2015

Jury:

Dr Mustapha Rouis, President of the Jury

Dr Marie-Paul Jacob-Lenet, Reporter

Dr Alain-Pierre Gadeau, Reporter

Dr Klaus-Peter Janssen, Examiner

Dr Fabienne Foufelle, Examiner

To my family

To my friends

“Science is a way of thinking, much more than it is a body of knowledge”

Carl Sagan

Acknowledgment

Having reached the end of my thesis, I would like to thank all the people that made this work possible and thus allowed me to attain my fondest dream, working in biological research.

First of all I would like to thank Dr Ewa Ninio who supervised all my work on this project, whose determination and tenacity allowed us to overcome many difficulties, both scientific and administrative that paved the way along my thesis project.

I also thank Pr Blankenberg and Pr Tanja Zeller, who gave me the opportunity to do half of my thesis in her team at the Universitätsklinikum Eppendorf and without whom this project wouldn't have been possible. Furthermore, I would like to also give a special thanks to Laurence Tiret who made this collaboration possible in addition to contributing to the microarrays analyses and François Cambien who participated in the initiation of the project and microarrays analyses.

I would also like to thank the members of my jury Dr Mustapha Rouis, Dr Marie-Paul Jacob-Lenet, Dr Alain-Pierre Gadeau, Dr Klaus-Peter Janssen and Dr Fabienne Fougère, for giving me the honor of evaluating my scientific work.

My thanks also goes to all the people from the core facilities who contributed to the experiments during my thesis: Solenne Chardonnet Cedric Pionneau who were extremely helpful for the protein part of this project, Dr Frederic Charlotte and Annette Lesot who performed and contributed to the analysis of the histological staining, Dominique Langui, Aurélien Dauphin and Perrine Frère who taught me how to use confocal microscopy.

Nathalie Abiola also has all my thanks for helping to overcome all the administrative paperwork that goes along the thesis and international collaboration.

Of course I would like to thank my lab mates from Paris, Hervé Durand, Carole Proust, Claire Perret, Dominique Stengel and Zahia Touat to whom the hour 11h30 should ring a pavlovian bell. And the bioinformaticians from our team, Nicolas who worked with me in the late hours be it on GHS database analyses or plane engineering, Dylan who helped with the methylation data and who's fascination for Spanish salami sometimes go over the edge, Guillaume who invented the double deck human blood shaker, Vinh with whom I had a lot of interesting talk regarding mRNA sequencing and who initiated me to rock climbing deep in the scary woods of Fontainebleau, Ares who is now working on the hypothesis that catalan and spanish could be two different languages, Ulli who supervised part of my intense training before leaving for Germany, Bathilde who is probably drinking beer at the time I write these lines, Jessica who scientifically demonstrated that there is no need for us to put on a helmet while riding a bike, Lyamine who gave me the trick to always get nice food for lunch, Nadjim who's dedication to washing its Tupperware is now legendary, Romain who definitely "knows about this theory", Alexandre who picked up most of our bad habits, Maguelonne who managed my communication with the team brilliantly, Veronica who always had a nice vaff' for me, Yasmine "ptb", François who's hears leave no one indifferent,

Bastien who will always be forever unemployed in my heart. I also thank Rajai my predecessor for his introduction to the lab, Badreddine who's calm and patience are to be commanded, and finally all the interns that came along the way and gave me the occasion to learn how to supervise students in the lab, H  l  ne "un, deux, trois, j'ai gagn  " who tricked Nicolas into getting croissant for everyone, Chlo   who I guess is the only person now working in pedagogy being able to perform quantitative PCR, Saakib who learned that experiments in the wet lab have their part of mystery, Auxanne who fought for the right of the interns to have their own personalized thanks in this part and Rima who charitably always smile at my jokes, even when they miss the spot.

My German lab mates are not to be forgotten as well, as they gave me a warm welcome in Hamburg and tried to kidnap me away from the lab from time to time! So thank you, Tina who took care of me when I first arrive in Hamburg, Diana who took time to give me good advice for the redaction of my manuscript, especially for the figures, Simon "mister Z" my personal body coach and ping pong partner with whom I wrote a song after extracting RNA at 1am, Christian "herr Mueller" who must now fear the sound of my steps moving toward his office, Mereile "mip mip" who always brought a lot of fun in the lab, Tim who's gattung ist zu selten and loves water-soluble bear, Caro who thinks french are so much into bread that they call wand "baguette magique", Svenja who holds a picture of me proving my med' cred', fossil Svenja who now have a -152 million dead fish at home because of me, Vivian who I only met briefly but who's good mood is contagious, taller Vivian who always took time to practice English with me, Jule who is also an honorary member of the "spreading the good mood team", Frauke who is into boxing so I shall not make any further comments here, Maerit who still owes me a severe correction at tennis, Sebastian with whom I had very interesting discussions, Hartwig who helped me a lot with the logistic, Denise who would make a fortune if she had a penny every time someone came to her for a technical advice, Rose who everyone think is german and a special thought for Arius "herr jami" with whom I spent long nights/week-ends working in the lab and who learned the most important French sentence to be used in a lab "je n'ai pas de resultats". I also thank the german interns I supervised: Tobias who discovered the syndrome of the ghost drop and Sabrina whose performances were very impressive.

Finally I would like to thank my friends and family who were by my side all the way.

Table of content

| | |
|---|-----------|
| <i>I Introduction/Scientific background</i> | 10 |
| I.1. Atherosclerosis | 10 |
| I.1.a Physiopathology of atherosclerosis..... | 10 |
| I.1.b cell types involved in the plaque progression | 14 |
| I.2 High-Throughput technologies and atherosclerosis | 16 |
| I.3 Smoking and atherosclerosis..... | 17 |
| I.3.a The pathophysiology linking smoking and atherosclerosis | 17 |
| I.3.b Molecular pathways linking smoking to atherosclerosis | 18 |
| I.3.c Modeling cigarette smoke exposure <i>in vitro</i> | 22 |
| I.4 SASH1..... | 25 |
| I.4.a <i>SASH1</i> gene..... | 25 |
| I.4.b SASH1 protein..... | 25 |
| I.4.c SASH1 in the literature, the quest for SASH1 function..... | 26 |
| I.5 Thesis project: from Genome Wide Analysis in the Gutenberg Health Study to <i>in vitro</i> experiments.... | 30 |
| I.6 Aim of the thesis and experimental strategy | 32 |
| II Article: SASH1 a new link between cigarette smoke and atherosclerosis | 33 |
| II.3 SASH1 ongoing experiments and unpublished results..... | 65 |
| II.3.a Materials and methods..... | 65 |
| II.3.b Detecting the SASH1 protein in HAECs..... | 68 |
| II.3.c SASH1 expression in vascular cells..... | 69 |
| II.3.d Testing the effect <i>SASH1</i> knockdown on the SASH1 network..... | 71 |
| II.3.e Effect of HAECs stimulation with inflammation related compounds on SASH expression and its sub cellular localization | 71 |
| II.3.f Transcriptomic analysis of cigarette smoke condensate stimulated HAECs..... | 75 |
| II.3.g <i>In silico</i> study of SASH1 methylation | 75 |
| II.4 Searching for SASH1 partner proteins..... | 77 |

| | |
|---|------------|
| II.4.a Yeast two hybrid (Y2H) | 77 |
| II.4.b Co-immunoprecipitation | 78 |
| II.4.c Mass spectrometry | 81 |
| III Discussion | 83 |
| References | 92 |
| ANNEXES | 109 |
| Annex 1: SLC39A8 (ZIP8) and cigarette smoking, genesis of a new project..... | 110 |
| I Introduction..... | 110 |
| I.1 <i>SLC39A8</i> gene organisation | 111 |
| I.2 SLC39A8 protein (ZIP8)..... | 111 |
| I.3 Zip8 in the literature..... | 112 |
| I.4 Material and methods..... | 114 |
| II Results | 115 |
| II.1 ZIP8 expression in vascular cells and in the vascular wall..... | 115 |
| II.2 ZIP8 expression and the plaque formation | 117 |
| II.3 ZIP8 expression in human carotid plaque according to the smoking status..... | 118 |
| II.4 Investigating the interdependence of ZIP8 and Gas6 expression in human macrophages | 119 |
| II.5 Methallothionein and ZIP8 expression..... | 120 |
| III Discussion | 123 |
| Références..... | 125 |
| Annex 2. Highlights for the ZIP8 project..... | 128 |
| Annex 3. Hightlights of the SASH1 project | 129 |
| Annex 4. Verdugo et al 2013..... | 130 |
| Annex 5. Touat et al (publication in preparation)..... | 145 |

Abbreviations

ABCA1 : ATP-binding cassette sub-family A member 1
ABCG1 : ATP-binding cassette sub-family G member 1
AFF1 : AF4/FMR2 family member 1
AHRR : Aryl-Hydrocarbon Receptor Repressor
ANGPT2 : Angiopoietin 2
ANLN : Anillin
C1D : Nuclear receptor co-repressor
CCL2 : Chemokine ligand 2
CCND1 : Cyclin D1
CCND3 : Cyclin D3
CRP : C-reactive protein
CSC : Cigarette smoke condensate
CTNNA1 : Catenin alpha 1
CTNNBIP1 : Catenin, Beta Interacting Protein 1
CTTN : Cortactin
CVD : Cardiovascular disease
CXCR4 : Chemokine (C-X-C Motif) Receptor 4
CYP1A1 : Cytochrome P450, family 1, subfamily A, polypeptide 1
DMSO : Diméthylsulfoxyde
DUH : Dyschromatosis universalis heredita
EDTA : Ethylenediaminetetraacetic acid
eNOS : Endothelial NO synthase
ESM1 : Endothelial Cell-Specific Molecule 1
FADD : FAS-Associating Death Domain-Containing Protein
FOXM1 : Forkhead Box M1
FOXO1 : Forkhead Box O1
GAPDH : Glyceraldehyde-3-Phosphate Dehydrogenase
Gas6 : growth arrest 6
GHS : Guntenberg health study
GM-CSF : Granulocyte macrophage colony stimulating factor
GPR15 : G protein receptor 15
GWAS : Genome wide association study
GWE : Genome wide expression
HAECs : Human aortic endothelial cell
HCASMC : Human coronary artery smooth muscle cell
HEK : Human embryonal kidney
HDL : High density lipoprotein
HMEC : Human microvascular endothelial cell
HMOX1 : Heme Oxygenase 1
HUVEC : Human umbilical vein cells
ICAM1 : Intercellular adhesion molecule 1
IkbkA : Inhibitor Of Kappa Light Polypeptide Gene Enhancer In B-Cells, Kinase alpha
IkbkB : Inhibitor Of Kappa Light Polypeptide Gene Enhancer In B-Cells, Kinase Beta
IL-4 : Interleukin 4
IL-6 : Interleukin 6

IL-8 : Interleukin 8
INF γ : Interferon gamma
IQGAP1 : IQ Motif Containing GTPase Activating Protein 1
KPNA2 : Karyopherin alpha 2
LDL : Low density lipoprotein
LPS : liposaccharide
LRP5 : Low Density Lipoprotein Receptor-Related Protein 5
LXR : Liver X receptor
MAP2K2 : Mitogen-Activated Protein Kinase Kinase 2
M-CSF : Macrophages colony stimulating factor
MDM2 : Mouse Double Minute 2, Human Homolog Of
MMP : Matrix metalloproteinase
MYH9 : Myosin heavy chain 9
NES : Nuclear export signal
NF κ B : Nuclear Factor Of Kappa Light Polypeptide Gene Enhancer In B-Cells
NLS : Nuclear localization signal
NO : Nitric oxide
OAS2 : 2'-5'-Oligoadenylate synthase 2
PAF : Platelet activating factor
PBS : Phosphate buffer saline
PDGF : Platelet derived growth factor
PLA2G4A : Phospholipase A2, Group IVA
PTGS2 : Prostaglandin-Endoperoxide Synthase 2
RAC1 : Rho Family small GTP binding protein 1
RPS25 : 40S ribosomal protein S25
SASH1 : SAM and SH3 domain containing 1
SLC39A8 : Solute carrier family 39, subfamily A, member 8
SLY1 : SH3 Protein Expressed In Lymphocytes
SLY2 : Src Homology Domain 3 (SH3)-Containing Adapter Protein SH3 Lymphocyte Protein 2
SMC : Smooth muscle cells
SNP : Single nucleotide polymorphism
SR-A1 : Scavenger receptor A1
TAK1 : TGF-Beta Activated Kinase 1
TGF β : Transforming growth factor beta
TLR4 : Toll like receptor 4
TNF α : Tumor necrosis factor alpha
TOP2A : Topoisomerase (DNA) II Alpha
TP53 : Tumor Protein P53
TPM1 : Tropomyosin 1
TRAF6 : TNF Receptor-Associated Factor 6, E3 Ubiquitin Protein Ligase
VCAM1 : Vascular cell adhesion protein 1
VEGF : Vascular endothelial growth factor
Y2H : Yeast two hybrid
ZIP8 : Zrt- And Irt-Like Protein 8

I Introduction/Scientific background

I.1. Atherosclerosis

I.1.a Physiopathology of atherosclerosis

Atherosclerosis is a chronic physiopathological condition that starts early in the childhood and evolves asymptotically throughout aging, with atherosclerotic plaques accumulating in the arteries. These lipid laden plaques can rupture in late stages of the pathology, leading to thrombosis. Depending on the location of the thrombus, this can cause a variety of complications, including events like stroke or myocardial infarction.

The vascular wall

To understand how the pathology progresses, one should first consider the organization of the vascular wall. The vascular wall is composed of three different layers (also called tunica), from the inner- to the outermost layer: the tunica Intima, the tunica media and the tunica externa (adventitia) (Figure 1).

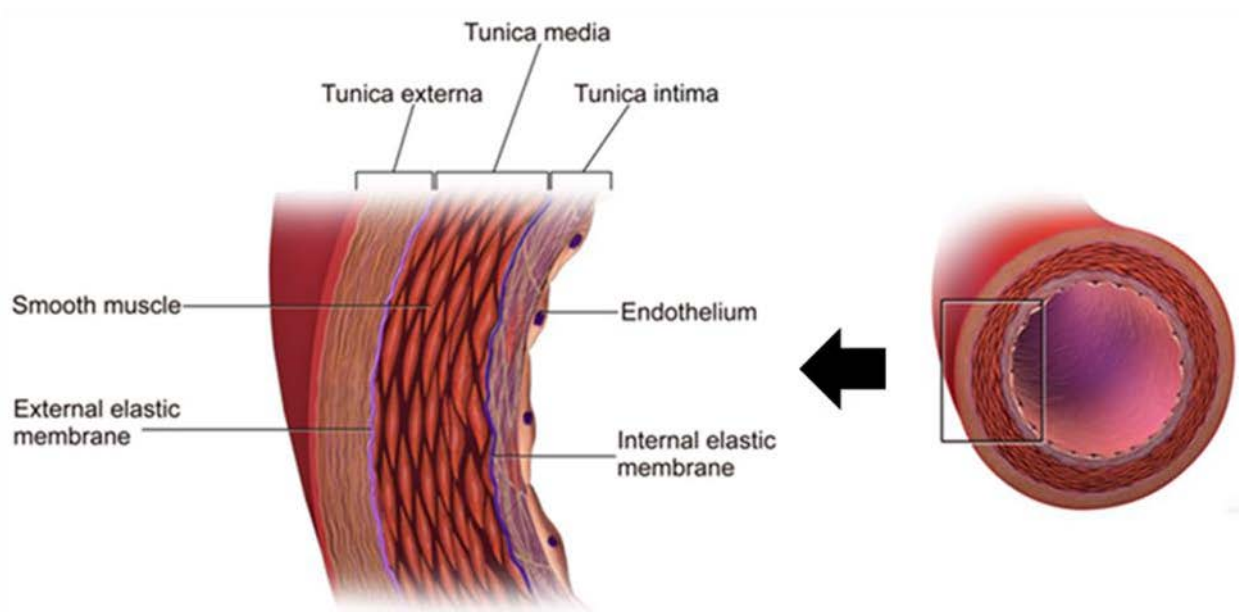


Figure 1. Structure of the vascular wall (Adapted from Wikipedia). Disposition of the three vascular layers, tunica intima, tunica media and tunica adventitia (tunica externa) that form the vascular wall.

Each of these layers is composed of different cells and plays a unique role in the vessel. Those layers communicate with one another through various small molecules transiting between them *via* gap junctions¹. Consequently, vascular layers should not be considered as isolated players, as they cooperate together to allow the vascular wall to fulfill its functions.

The *tunica Intima*, the innermost tunica of the vascular wall, is composed exclusively of endothelial cells organized as a monolayer, lying on a fibro-elastic connective tissue. The cells are in direct contact with the blood flow and as a consequence, regulate the permeability of the vascular wall. They are also regulating the vascular tone and prevent smooth muscle cells from activating.

The *tunica media*, located between the intima and adventitia, is the thickest layer of the vascular wall. It is composed of smooth muscle cells organized in concentric layers wrapped by an extracellular matrix composed of fibrous and elastic proteins as well as mucopolysaccharides. The media produces the mechanical force that allows vasoconstriction and vasodilatation.

The tunica externa (also known as adventitia), the outmost layer of the vascular wall, is composed mostly of connective tissue (collagen and elastic fibers) in which you can find few fibroblasts and adipocytes. The cells are irrigated by a network of vessels called the *vasa vasorum*. The role of the adventitia is to limit the expansion of the vascular wall and to allow the anchoring of the vessels to nearby organs or structures.

The lipoproteins as triggering factor of the pathology

Lipoproteins are considered as the most robust risk factors for the pathology. Plasma lipoproteins are particles containing proteins, cholesterol and other lipids which allow their distribution through the body until their metabolization or degradation. The plasma lipoproteins are traditionally separated between high density lipoproteins (also known as “good cholesterol”) and the low density lipoproteins (LDL) (the infamous “bad cholesterol”).

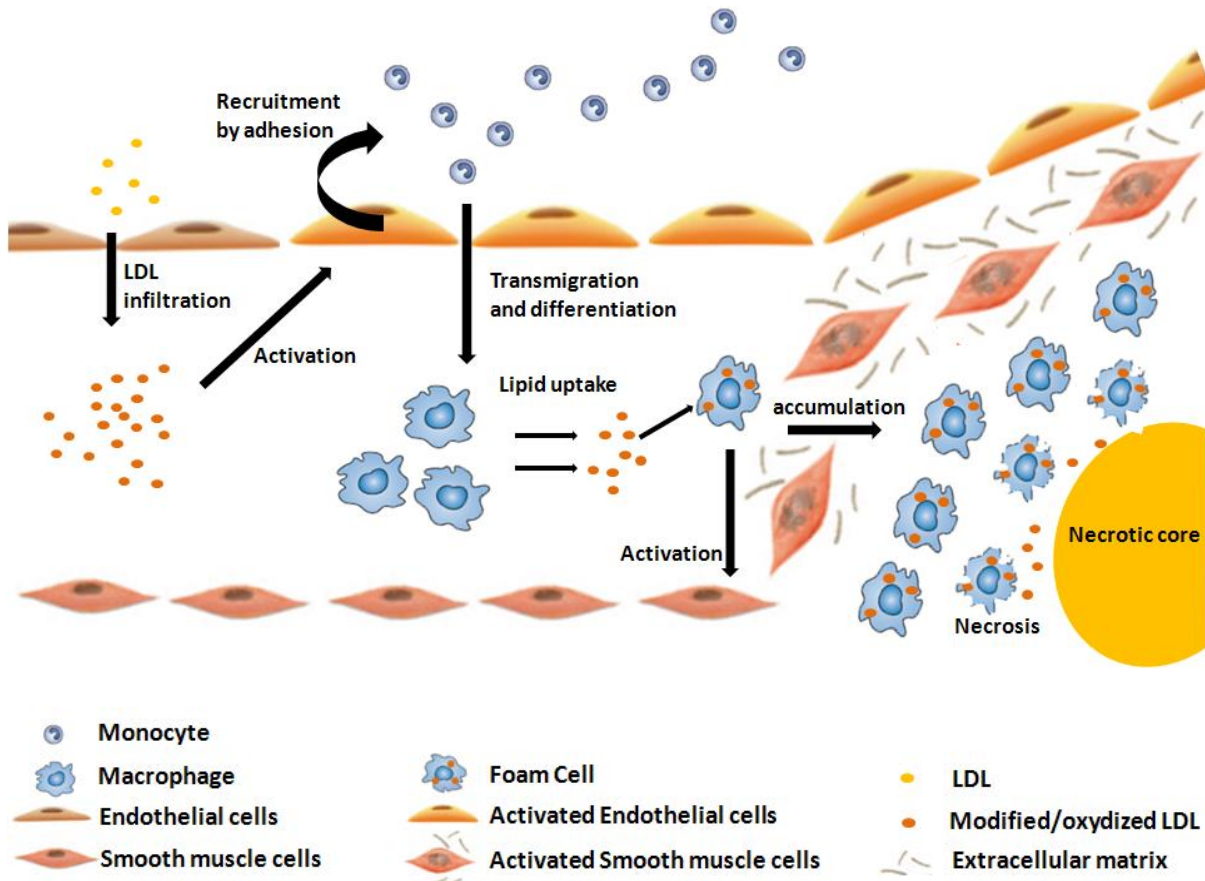


Figure 2. Simplified overview of the atherosclerotic plaque formation. From left to right, LDL infiltrate in the intimal space where they are modified. Modified LDL activate endothelial cells which allow recruitment of monocytes. After infiltration, monocytes differentiate into macrophages and capture modified LDL, hence becoming lipid laden foam cells. Foam cells activate smooth muscle cells that switch from a quiescent state to dedifferentiated phenotype. Foam cells undergo necrosis, freeing lipids and cholesterol into the intimal space. Free lipids and cholesterol crystallize and form the necrotic core. Activated smooth muscle cells migrate to the neo intimal space where they proliferate and secrete extracellular matrix that leads to the formation of a fibrous cap between the intima and intimal space.

The onset of atherosclerosis (Figure 2) starts when LDL transmute through the endothelial monolayer into the intimal space (a space between the intima and the media). Those LDL are trapped in the intimal space and undergo oxidative or enzymatic modifications. These altered LDL cause an inflammatory reaction that activates the endothelial cells. The activated endothelial cells express vascular cell adhesion protein 1 (VCAM1), Intercellular adhesion molecule 1 (ICAM1) and

selectin (SELE) adhesion molecules which participate in the recruitment of the monocytes from the circulation into the intimal space and their differentiation into macrophages. Those macrophages express at their plasma membrane, among others, a panel of receptors called the scavenger receptors (SRB1, CD36, CD68...). Those scavenger receptors recognize modified LDL and allow their uptake by the macrophages. The macrophages accumulate the LDL in their cytoplasm, transforming them into foam cells. Small conglomerates of foam cells are considered as the first stage of atherosclerosis as described first by Stary in 1994² and 1995³ (all stages of plaque formation are presented in Figure 3).

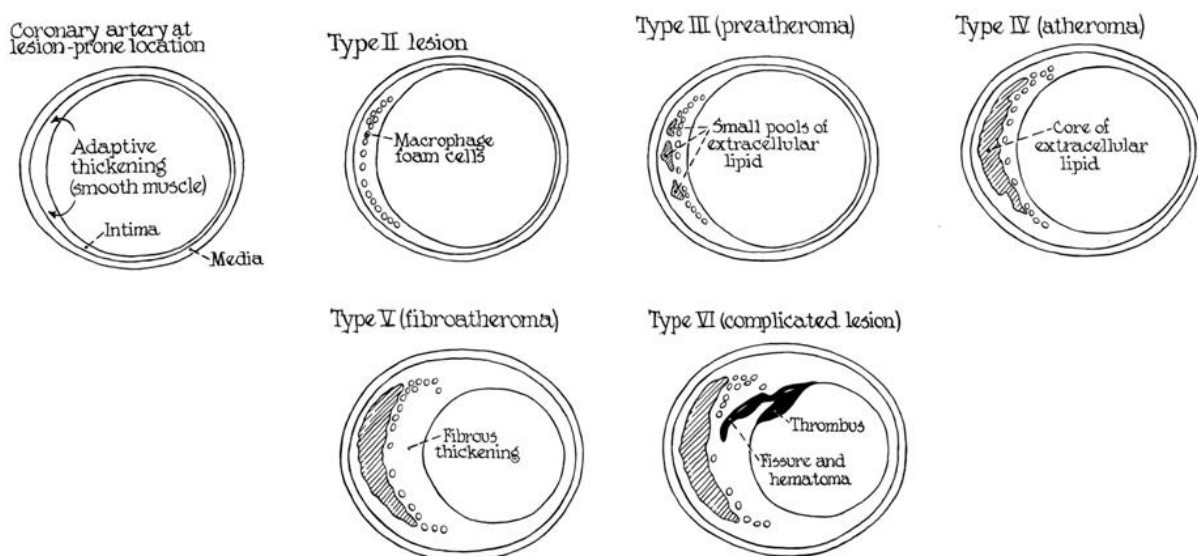


Figure 3: Original representation of the various stages of atherosclerosis from Stary et al, 1994. Description of the different type of atherosclerotic lesion throughout atherosclerosis development. Early atherosclerotic events (type II and III) precedes the formation of the plaque (type IV and V) that can then become complicated plaque (type VI) prone to rupture.

Foam cells secrete various inflammatory cytokines that further activate endothelium permitting increased recruitment of monocytes and thus leading to the formation of additional foam cells which form the fatty streaks, visible as a yellow linear protrusion in the intima (stage II). At this stage, the inflammation triggered by the foam cells also modifies the phenotype of the smooth muscle cells (SMC) in the intima. These SMC are quiescent and contractile under normal physiological condition, but after activation by PDGF (Platelet Derived Growth Factor) released by the foam cells they start to migrate into the intimal space, where they proliferate and secrete cytokines. The fatty streaks are mostly localized at the arterial bifurcations, where the shear stress is lower and oscillatory, and can in turn become pre-atherosclerotic lesions, in which small pools of extracellular lipids accumulate in the neointimal space (stage III). The accumulation of monocyte-derived macrophages is irreversible as they cannot return to circulation, thus they die directly in the plaque either by necrosis or apoptosis (both have been observed within the plaque) and ultimately liberate their lipid content in the intimal space. This leads to the formation of the necrotic core, an acellular accumulation of crystallized cholesterol and dead cells debris, surrounded by numerous

living foam cells (stage IV). Dedifferentiated smooth muscle cells secrete an important quantity of fibrous proteins that form the fibrous cap, a layer of connective tissue that separate the necrotic core from the intima. The result is a fibro-lipid plaque (stage V). At this stage the intima is still present but the initial connective tissue on which it was lying is already degraded. Stage V plaques can have disruptions at their surface, such as hematoma or hemorrhage, in that case they are considered complicated lesions (stage VI).

Some of the plaques may calcify (stage VII) which impairs the contractile function of the vessels. Finally there are plaques which are highly fibrotic with very low lipid content and are then classified as stage VIII.

I.1.b cell types involved in the plaque progression

The endothelial cells

The endothelial cells that form the Intima are organized as a monolayer, forming the innermost tunica in direct contact with the blood flow. The endothelial cells are considered the guardians of the vascular homeostasis, maintaining the balance between vasodilatation and vasoconstriction. These effects are mainly regulated through the nitric oxide (NO) production, a gas produced by the endothelial NO synthase (eNOS) which expression is regulated by prostacyclin and bradykinin⁴. NO diffuses easily within the vascular wall, its actions includes inhibiting platelet activation⁵, leukocytes transmigration, smooth muscle cells proliferation⁶ and inhibits LDL oxydative modifications⁷. All of these mechanisms can be considered as anti-atherogenic, nevertheless, endothelial cells also expresses vasoconstrictive hormones such as endothelin⁸ and angiotensin II⁹ which promote various proatherogenic pathways, including smooth muscle cells proliferation¹⁰. In atherosclerosis, endothelial cells are activated, switching from a quiescent NO producing phenotype, to an activated phenotype producing reactive oxygen species (ROS) and recruiting monocytes. The mechanisms of endothelial activation are similar to a normal immune response to infection, where ROS production aim at killing the pathogens. However, a steady production of ROS also damages the nearby tissues and increases inflammation¹¹. In addition, activated endothelial cells express adhesion proteins such as ICAM1, VCAM1 and various members of the selectin family¹² along with leukocytes chemoattractant such as monocyte chemoattractant protein 1 (MCP-1 also known as CCL2). This results in a constant influx of leukocytes into the vascular wall where they are trapped and ultimately participate to plaque formation¹³. Endothelial cells are also known to have anti-thrombotic properties *via* the production of prostaglandin I₂ (PGI₂), anti-thrombin III and the plasminogene activator (t-PA) which either directly inhibits thrombin or degrades thrombotic proteins like fibrin¹⁴.

In summary endothelial cells are considered as key regulators of plaque formation and their activation (also referred to as “endothelial dysfunction”) is considered to be the first step of plaque formation. Indeed correlations have been found between endothelial activation and cardiovascular disease and more specifically with cardiac events linked to atherosclerosis. Alteration of the endothelial function has been observed in patients with coronary artery risk factors, however

ultrasound measurement showed no evidence of the presence of plaques in their arteries^{15,16}, indicating that endothelial activation is indeed the prequel to the plaque formation.

The smooth muscle cells

Smooth muscle cells provide the required force for the arterial wall contraction, allowing the vasoconstriction. Under normal physiological condition, they are quiescent and contractile however they may switch to a pathological, dedifferentiated phenotype characterized by increased migration and proliferation as observed in atherosclerosis¹⁷. Smooth muscle cells are the main producers of extracellular matrix within the vascular wall. In the normal vascular wall, they mainly express Type I and Type III fibrillar collagen, while under atherosclerotic conditions they produce scattered type I collagen fibrils and fibronectine¹⁸. This new composition of the extracellular matrix can trap more modified LDL particles within the wall¹⁹. Like macrophages, smooth muscle cells also express the scavenger receptors at their surface and are thus able to internalize oxidized LDL and become foam cells²⁰. Smooth muscle cells in the atherosclerotic lesions also produce monocytes chemoattractants, such as CCL2, and express adhesion proteins VCAM1 and ICAM1²¹ that allow them to directly attract and bind to monocytes and to macrophages. This interaction is thought to regulate apoptosis in macrophages, which implies that smooth muscle cells play a role in accumulation of foam cells within the plaque²². The smooth muscle cells only express those specific adhesion proteins in atherosclerosis and not in the normal vascular tissue^{23,24,25}. Finally smooth muscle cells also secrete a number of cytokines involved in inflammation, including platelet derived growth factor (PDGF), transforming growth factor beta (TGF β) and interferon γ (IFN γ)²⁶. While the role of smooth muscle cells is well established in advanced atherosclerotic lesions, many recent studies aimed to determine whether or not they are also major actors in the early onset of the pathology. For instance it was recently demonstrated that SMC express VCAM1 and ICAM1 prior or concomitantly to monocytes transmigration²⁷ indicating that they could also participate in the recruitment of monocytes, further into the plaque. In summary SMC are the main proliferative cell type in atherosclerosis, the main producers of extracellular matrix component and share the LDL internalization processes with macrophages.

Monocytes and macrophages

Monocytes and macrophages are normally absent from the vascular wall, their presence in this tissue always imply ongoing local inflammation that amplifies the recruitment of monocytes from the blood stream into the vascular wall. Monocytes which transmigrate into the inflamed tissue ultimately differentiate into macrophages. Once in the intima, macrophages scavenge the modified LDL and the cellular debris from the extracellular compartment. Under normal physiological conditions, macrophages may leave the tissue (phenomenon known as egress) when inflammation is resolved, however in atherosclerosis, inflammation is chronic and macrophages stay within the vascular wall²⁸. Under normal conditions, after their uptake by the LDL receptor (LDL-R), the LDL are eliminated via the LXR nuclear receptor pathway that triggers the expression of transporters such as ABCA1 and ABCG1 allowing cholesterol efflux from the cells^{29,30}. Modified LDL follow the same efflux pathways but are internalized through different receptors, the scavenger receptors SR-A1

and SR-A2 . The difference between scavenger receptors and LDLR is that LDLR is not recycled to the membrane in case of overaccumulation of intracellular lipids, whereas the scavenger receptors keep importing the oxidized lipids. This ultimately leads to macrophages apoptosis and the liberation of apoptotic bodies within the plaque. The mechanisms by which macrophages enter apoptosis involve excessive intra cellular lipid accumulation leading to the stiffening the ER membrane³¹ leading to unfolded protein response³², a highly conserved mechanism that aim at restoring the endoplasmic reticulum function. When this mechanism fails, and the unfolded protein response remains active for a prolonged period, apoptosis is activated. As macrophages are overwhelmed they cannot clear the apoptotic bodies which leak their content in a processes known as post apoptotic necrosis³³. Macrophages and foam cells secrete an important number of pro-inflammatory and anti-inflammatory cytokines. Macrophages are phenotypically diverse, they are typically distributed in a wide range between two extreme phenotypes referred to as M1 (pro inflammatory macrophages) and M2 (anti inflammatory macrophages)³⁴. The M1 macrophages, refer to the classically activated monocyte usually by either interferon gamma (IFN γ) or granulocyte macrophage colony stimulating factor (GM-CSF)³⁵.The M1 macrophages secrete pro inflammatory cytokines such as IL-6 and IL-8 and are thus considered as pro-inflammatory. The M2 macrophages, refer to an alternative activation of the monocytes, usually achieved by either IL-4 or macrophages colony stimulating factor (M-CSF) stimulation. They secrete TGF β and IL-10 and are therefore considered as anti-inflammatory.

Overall, monocytes and macrophages are by far the most studied cell types in atherosclerosis as they hold the key to the lipid accumulation in the plaques and are among the main cell types that modulate inflammation within the vascular wall. Current studies on those cells include restoring the cholesterol efflux pathways in foam cells and studying macrophages egress from the plaques to understand why the macrophages are unable to leave the lesions^{36,37}.

While this part focused on the “main” cell types involved in atherosclerosis, it is important to underline the fact that other cell types are also involved in plaque formation, including lymphocytes and neutrophils. However, those cells where not a part of our study thus their contribution to atherosclerosis will not be described here.

I.2 High-Throughput technologies and atherosclerosis

Atherosclerosis is a multifaceted pathology that involves a wide and diverse range of pathological mechanisms. Identification of relevant therapeutic targets therefore requires an integrated approach. Most of biological studies focus on a single candidate or a pathway, while most pathologies are usually multi-factorial and results from complex cross talk between different pathways. High throughput technologies (also known as –OMICs) allow to investigate all the pathways involved in the pathology in a single experiment. They can be divided into several categories: the genome wide association studies (GWAS) focusing on the correlations between DNA mutations (such as single nucleotide polymorphisms, SNP) and a given phenotype (e.g. disease, body mass index, age...), genome wide expression studies (GWE) establishing the relations between

genes expression (RNA levels) and a given phenotype, finally proteomics approach is used to determine the link between protein expressions and a phenotype. There are also a number of other –omic technologies such as metabolomics, studying the metabolites or epigenomics that focuses on DNA non sequential modification such as methylation. As the omics are less accurate than conventional, targeted studies, high throughput methods are often used in combination, for example performing a GWE after a GWAS allows to assess whether the detected SNP in regulatory sequences in the DNA affects the RNA expression of the gene. Candidates selected by these methods are then further studied by more specific methods (real time PCR, Western blot...). The aim of those new global approaches is usually to find new therapeutically important targets or biomarkers, thus improving prognosis and patient care. This thesis study was done in the continuity of a GWE³⁸, performed in the frame of a population base study (see §1.5).

I.3 Smoking and atherosclerosis

I.3.a The pathophysiology linking smoking and atherosclerosis

The link between smoking, both active and passive, and atherosclerosis is well established. Cigarette smoking affects all stages of atherosclerosis contributing to endothelial dysfunction³⁹, increasing low density lipoproteins (LDL) oxidation⁴⁰, monocyte recruitment⁴¹ and thrombus formation following plaque rupture⁴². In line with these findings, smoking aggravates many cardiovascular diseases including coronary artery disease leading to myocardial infarction and is implicated in sudden death^{43,44}. Cigarette smoke is usually divided in two phases, first phase is called the tar-phase and regroups the particles with a size superior to 0.1 µm, and the second one is the gas-phase that contains the smaller particles. Both phases contain a high amount of free radicals, however the half-life of these free radicals are different between the two phases. Free radicals from the tar-phase are long lived and can last hours to month in the body, while the gas-phase contains free radicals that usually persist for only a few seconds after inhalation^{45,46}. Furthermore, there is a difference between the mainstream smoke that is inhaled by smokers and the sidestream smoke, that comes from the burning end of the cigarette⁴⁷. The later contains much more toxic compounds than the first one. Environmental tobacco smoke, to which both passive and active smokers are exposed is composed of 85% sidestream smoke and only 15% of mainstream smoke⁴⁸.

It is still a matter of debate whether or not cigarette smoke affects the CVD risk in a dose dependent fashion. Indeed, Passive smoking increases the risk of CVD of 30%, while active smoking increases it by 80%^{49,50}. Passive smokers are exposed roughly to an equivalent of 1% of what an active smoker, smoking 20 cigarette per day, would inhale⁵⁰, however their CVD risk increases by 30%, thus the dose-response to cigarette smoke is not linear. In the case of atherosclerosis, it has been demonstrated that intimal-medial thickness, a routine measurement to determine the advancement of atherosclerosis, is increased in both passive and active smokers⁵¹. Furthermore, both active and passive smoking are associated with decreased vasodilatation in various vascular

beds (in different sites in the arteries)^{52,53,54}. Finally, cigarette smoking significantly increases, LDL levels⁵⁵, lipid oxidation, leukocyte counts and inflammatory markers in humans^{56,57,58} (see section I.3.b), all of these processes being known to favor plaque growth.

I.3.b Molecular pathways linking smoking to atherosclerosis

Cigarette smoke contains thousands of active components, most of which have unknown impact on the human body. Since atherosclerosis itself is a complex, multifactorial disease, the impact of cigarette smoke is likely to be a combination of various factors, most of which are still unknown. Listed below are the main current research topics linking smoking to atherosclerosis as illustrated in Figure 4.

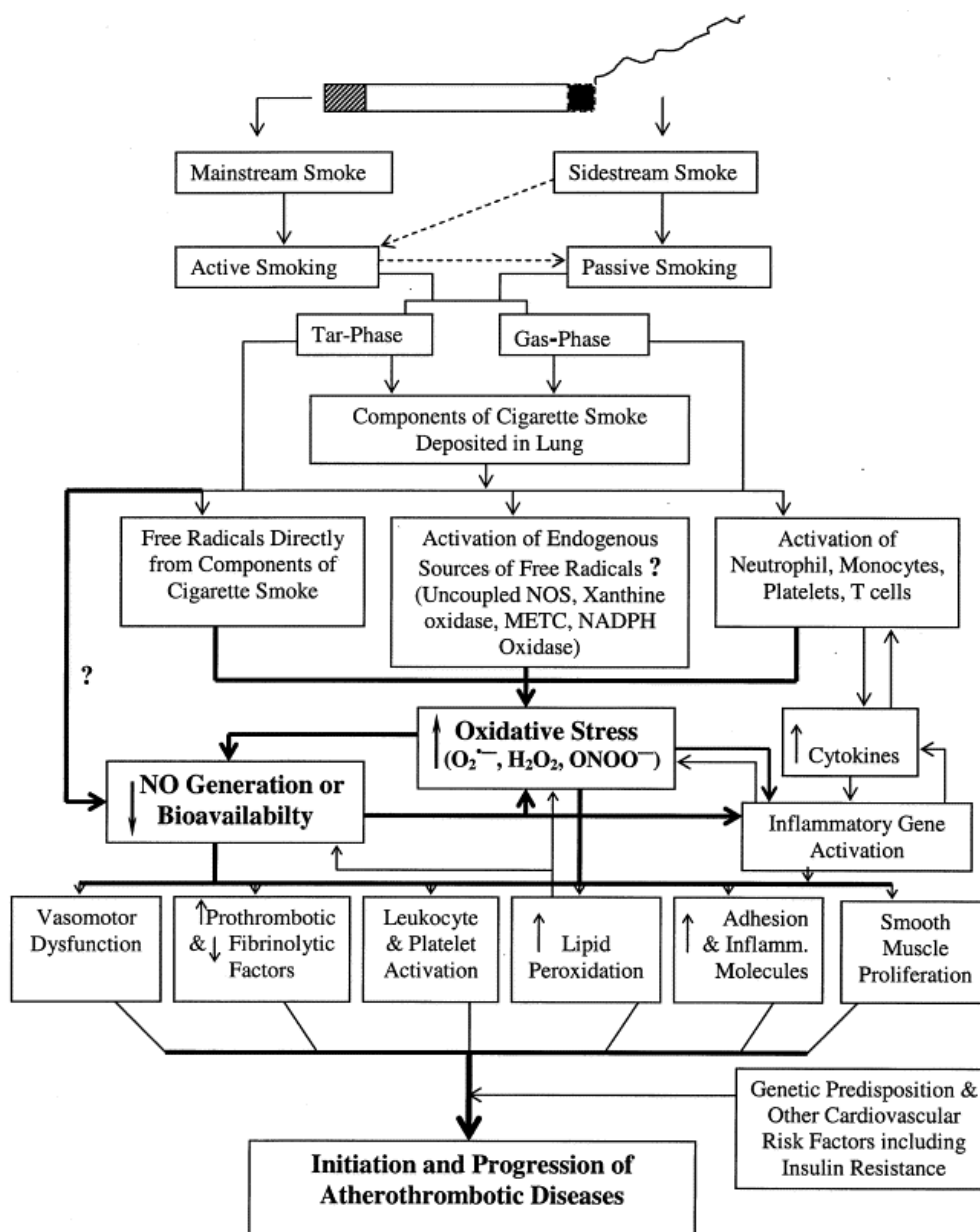


Figure 4. Pathways linking smoking to atherosclerosis. (from Ambrose et al 2004)

Effect of smoking on inflammation

Inflammation plays a key role in atherosclerosis as it promotes plaque growth. It increases the recruitment of monocytes into the intimal space of the arteries and thus the number of macrophages that, in turn become foam cells and accumulate in the lesions⁵⁹. Cigarette smoke increases inflammation, however the cellular and molecular pathways by which it aggravates atherosclerosis are still under scrutiny.

Smoking increases the blood leukocytes counts by 20%⁵⁶, consequently more monocytes are available for recruitment to the plaques. Additionally, it was shown that cigarette smoke increases the expression of adhesion proteins known to be involved in atherosclerosis by increasing monocyte recruitment, such as VCAM1, ICAM1 and E-selectin^{60,61}. Monocytes accordingly show a 70 to 90% increased adherence to human endothelial cells⁶², as well as a 200% increased transendothelial migration when they are stimulated with cigarette smoke extracts⁶³ (this in vitro model will be described further in the next part). Similar results were found when exposing human endothelial cells and monocytes to smokers' serum⁶⁴. Markers of inflammation (C-reactive protein (CRP), interleukin-6 (IL-6), tumor necrosis factor alpha (TNF α) and fibrinogen) were also significantly increased in the serum from smokers when compared to the non-smokers, regardless of the gender^{57,58,65}. The main pathway by which cigarette smoke is believed to exacerbate inflammation is oxidative stress due to its high content of free radicals which saturate the defensive antioxidant mechanisms leading to cell stress. The latter may explain the increased adhesion molecules expression at the monocyte and endothelial cell surface and thus result in a higher number of inflammatory cells recruited to the plaques. Several studies have focused on the effects of individual component of the cigarette smoke on inflammation. For example, nicotine triggers inflammatory response in human monocyte derived dendritic cells⁶⁶, including the secretion of the pro-inflammatory cytokine IL-12. In mice models stimulation by nicotine increases the monocyte-endothelial cell interactions resulting in increased rolling and adhesiveness of monocytes in the cerebral microcirculation⁶⁷. In the same study the authors showed, using a common mice model of atherosclerosis CD57/BL6 ApoE knock-out mice, that intravenous injection of nicotine increased the number of dendritic cells recruited to the atherosclerotic plaques.

Effect of cigarette smoking on the endothelium

Endothelial activation (also referred to as endothelial dysfunction) is the early step of atherosclerosis. It is characterized by the increased expression of surface adhesion proteins, but also by a decreased production of nitric oxide (NO). NO is the main stimulus for vasodilatation, it also inhibits platelet aggregation, leukocyte adhesion and thrombosis⁶⁸. It is therefore considered to be atheroprotective. Both active and passive smokers show decreased vasodilatation⁵²⁻⁵⁴, which corroborates many studies that have shown that cigarette smoke is associated with decreased NO availability^{54,69}. Several studies investigated the molecular pathways leading to decreased NO availability in smokers. It was demonstrated that exposure to smokers serums, increased endothelial nitric oxide synthase (eNOS) expression but decreased its enzymatic activity in human endothelial cells⁷⁰. In this context, the increase in eNOS expression could be seen as a

compensation mechanism for decreased eNOS activity. Furthermore, the free radicals contained in cigarette smoke react with NO, which also leads to a lowering of its availability. It is noteworthy that the expression of other molecules with known vasodilatation effects, such as tissue plasminogen activator (tPA) and plasminogen activator inhibitor-1 (PAI-1) have also been shown to be altered by cigarette smoke⁷¹. Interestingly, impairment of vasodilatation in smokers has been demonstrated to be partially reversed by administration of vitamin C (known to have anti oxidative properties), which suggests that it is mainly due to the oxidative stress⁷².

These modifications of endothelial functions leads to differential response to vascular injury between smokers and non-smokers, with visible morphologic differences. Indeed, the intimal/medial thickness of carotids measured by echo Doppler was consistently increased in both passive and active smokers, as compared to non smokers⁷³. More surprisingly, nicotine alone increased neointimal thickening in a rat model of balloon-induced endothelium injury⁷⁴. This corroborates a study on ApoE knockout mice, where a treatment with nicotine stimulated angiogenesis and plaque formation⁷⁵.

Effect of cigarette smoking on lipid levels and oxidation

Lipid composition and their modifications are important issues in atherosclerosis. This is due to the fact that LDL and modified lipids are considered the major triggering factors of the pathology. Interestingly cigarette smoke affects the balance between HDL considered as atheroprotective and LDL considered as pro atherogenic, by decreasing the blood concentration of the former and increasing the blood concentration of the latter⁵⁵. This switch in LDL/HDL ratio is clearly in favor of an atherogenic lipoprotein profile in smokers. Interestingly, this profile has been shown to be reversible, with increased HDL concentration within two weeks after smoking cessation⁷⁶. The pathway by which cigarette smoke switches the lipoprotein balance in favor of LDL is unknown, but could be due to insulin resistance. Indeed cigarette smoking is considered to be a risk factor for type 2 diabetes and intravenous infusion of nicotine in type 2 diabetic patients significantly reduces the insulin secretion^{77,78}.

Not only does cigarette smoke increases LDL concentration but it also increases lipid modification through increased oxidation⁷⁹, mainly because of its high content in free radicals. Those oxidized lipids activate the endothelium and thus increase inflammation. Furthermore, among those oxidized lipids, are PAF-like lipids that bind the PAF receptor (Platelet Activating Factor receptor) which triggers leukocytes adhesion and platelet activation and aggregation⁸⁰. The latter effect was reversed by the use of vitamin C pointing out the implication of the oxidative stress in these processes. However, when using serum from smokers to trigger leukocyte adhesion to human endothelial cells, another study could only reverse the process with L-arginine but not vitamin C⁸¹ suggesting that NO levels play a central role in this interaction. Therefore further studies are required to characterize the pathways involved in leukocyte-endothelial interactions.

Effect of cigarette smoking on thrombosis

Thrombosis is the consequence of the plaque rupture. When the integrity of the plaque is compromised, it causes the exposure of the highly pro-thrombotic lipid core to the blood stream triggering platelet activation/aggregation and the formation of a blood clot that obstructs the arteries leading to various pathological consequences (MI, stroke...) depending on the localization of the thrombus in the vascular tree. Cigarette smoke increases thrombosis by several synergic actions that alter the balance between antithrombotic and prothrombotic factors. Fibrinogen, a thrombotic agent converted by thrombin into insoluble fibrin strands, is increased in smokers as compared to non smokers^{82,83}, this increase is reversed in the ex smokers to the level of non smokers. Tissue plasminogen activator (tPA), a protein involved in the breakdown of the blood clot, is present at lower levels in serum from smokers⁸⁴, suggesting an impaired reaction to blood clot formation. Accordingly tPA inhibitor (PAI-1) is present in higher levels in the serum from smokers and has been shown to be upregulated by nicotine in human endothelial cell from the brain vasculature⁸⁵. Furthermore the tissue factor (TF) and the TF pathway inhibitor (TFPI-1), two major proteins involved in the formation of thrombin from pro thrombin show the same pattern of altered expression in both human⁸⁶ and mice⁸⁷. Moreover, cigarette smoke contains carbon monoxide (CO), a gas which competes with oxygen to bind hemoglobin and thus decreases the oxygen supply. This creates hypoxemia (abnormally low amount of oxygen transported by the blood) in both passive and active smokers and leads to an increased number of circulating red blood cell⁵⁶ that also are a major components of the blood clot formation. Finally, cigarette smoke induces platelet dysfunction as platelets from smokers aggregate spontaneously and in greater proportion following oxidative stress⁸⁸, suggesting that they are hyper activated even at basal conditions. This was further demonstrated by exposing platelets from the non smokers to a serum from the smokers, resulting in an increased aggregation⁸⁹, that was explained by the decreased NO level. To summarize, the cigarette smoke creates hypercoagulation.

Concomitant action of smoking and genetic polymorphisms

Genetics plays an important role in the susceptibility to cigarette smoke exposure. Indeed, it is very surprising that despite the large number of deleterious compounds present in the cigarette smoke, about 50% of the smokers do not die prematurely of their addiction. This suggests a strong implication of the genetic background in the susceptibility of smokers to develop cardiovascular diseases (CVD). For example, in the CYP1A1 gene, coding for a member of the cytochrome P450 protein family known to be involved in the metabolism of xenobiotics, the presence of the rare C allele in the 3'-flanking region of the CYP1A1 gene in smokers may enhance predisposition to severe CAD and type 2 diabetes⁹⁰. Surprisingly, this was observed only in light smokers but not in heavy smokers, highlighting the fact that there is no obvious evidence for a dose response linking cigarette smoke and CVD. Another polymorphism in the NOS3 gene coding for the nitric oxide synthase (eNOS) has been shown to be associated with increased coronary stenosis in smokers and ex smokers as compared to non smokers^{91,92}. A few other mutations were also characterized in proteins involved in blood coagulation, including factor V in which the allele FV:Q506 allele was

found to increase the risk of myocardial infarction in smokers⁹³. In summary, there are some genetic evidences linking smoking and atherosclerosis, suggesting that combination of genetic background and cigarette smoke can impact the risk for CVD.

I.3.c Modeling cigarette smoke exposure *in vitro*

As previously described, cigarette smoke contains over 5000 thousand active compounds unevenly distributed between mainstream smoke and sidestream smoke. Those compounds enter the body via the lungs, then diffuse to the blood stream which allow them to spread throughout the whole body. Several of these compounds like the benzo(a)pyren, are not toxic until they are processed in the liver or kidney⁹⁴. Because of that, modeling cigarette smoke *in vitro* can be a challenge. The most common approach, and the one used in this study, is to filter the cigarette smoke through a Cambridge filter and then resuspend the particles trapped in the filter by soaking it with an organic solvent (typically DMSO)⁹⁵. The resulting solution, called cigarette smoke extract (CSC), is then applied to the cells in culture. This method has the advantage of allowing reproducible testing, as one can aliquot, freeze the solution and use the same extract for many experiments. However, the vapor phase molecules are lost by this procedure. To solve this issue, another method consists of bubbling the whole cigarette smoke into the culture media or PBS, using a system (usually custom made) of pumps and tubing or simple syringe⁹⁶. This approach has the advantage of containing both compounds from the vapor phase and the particulate phase, however it also comes with a number of drawbacks. First the reproducibility may be poor, as the extract can only be used within 24h, and must be prepared freshly for each experiment. The main issue to obtain each time a similar extract is that cigarettes from different suppliers or even from different batches of the same supplier may differ in composition. Thus, using this mode of smoke extraction requires the use of standardized research cigarettes such as the 3R4F cigarette produced by the University of Kentucky (USA), and specially designed for research purpose. However, those cigarettes are not easy to obtain, especially for a research group outside of the USA. Another issue concerning the reproducibility of this model is that bubbling of the smoke through the medium gives uneven concentrations of the compounds contained in the smoke. Thus adjustment on a reference (typically by absorbance measurement at 320 nm to quantify turbidity) is required to ascertain that the amount of compounds is similar in all preparations. Several commercial systems have been created over the years to ensure reproducible extraction (reviewed by David Thorne and Jason Adamson⁹⁷). Those commercial systems all follow the same pattern: a first device burns the cigarettes, then in a second compartment the resulting smoke is bubbled in the culture medium, and finally the cells are exposed to the medium in a third compartment, the culture chamber. These systems are expensive, thus only few laboratories are equipped with them.

All previously mentioned models lack a major component of the cigarette smoke toxicity, the metabolization of the compounds by the human body. As previously stated, many compounds from cigarette smoke only acquire their toxicity after metabolization by the organism which mostly occurs in the liver and kidney, but also in the lungs. A number of solutions have been suggested to circumvent this issue. The first is to incubate the cigarette smoke extract with microsomes, prior to

the incubation with the cells. Microsomes are fragments of endoplasmic reticulum obtained by cell fractionation which contain several enzymes involved in the detoxification of xenobiotics. By adding them with their co-factors in the media with cigarette smoke extracts, it allows a partial metabolization of the extract⁹⁸. However this method can only be used for a short lasting stimulation, as the microsomal preparations are highly cytotoxic. Another possibility is to use conditioned media obtained by the exposure of cells with high metabolic activity, such as hepatocytes, to cigarette smoke extract and then to use their media to stimulate the cells of interest⁹⁹. Another version of this method is to co-culture directly those two cell types in the same compartment. Nevertheless, cultivating both cell types together usually requires physical separation of the cells in the culture compartment (using a culture insert to create a sub-compartment) as they have different growth rates¹⁰⁰. For example, cultivating hepatocytes with smooth muscle cells would result in the hepatocyte occupying the whole compartment before the smooth muscle cells had a chance to grow. However, both these approaches only give a fraction of the metabolites derived from cigarette smoke. To increase the number of different metabolites present in the smoke extract, another model consists of stimulating the cells with either plasma from smokers or non smokers. The main drawback of this model is that it requires an access to a constant source of well characterized plasma with known smoking records. All the described *in vitro* methods are summarized in Figure 5.

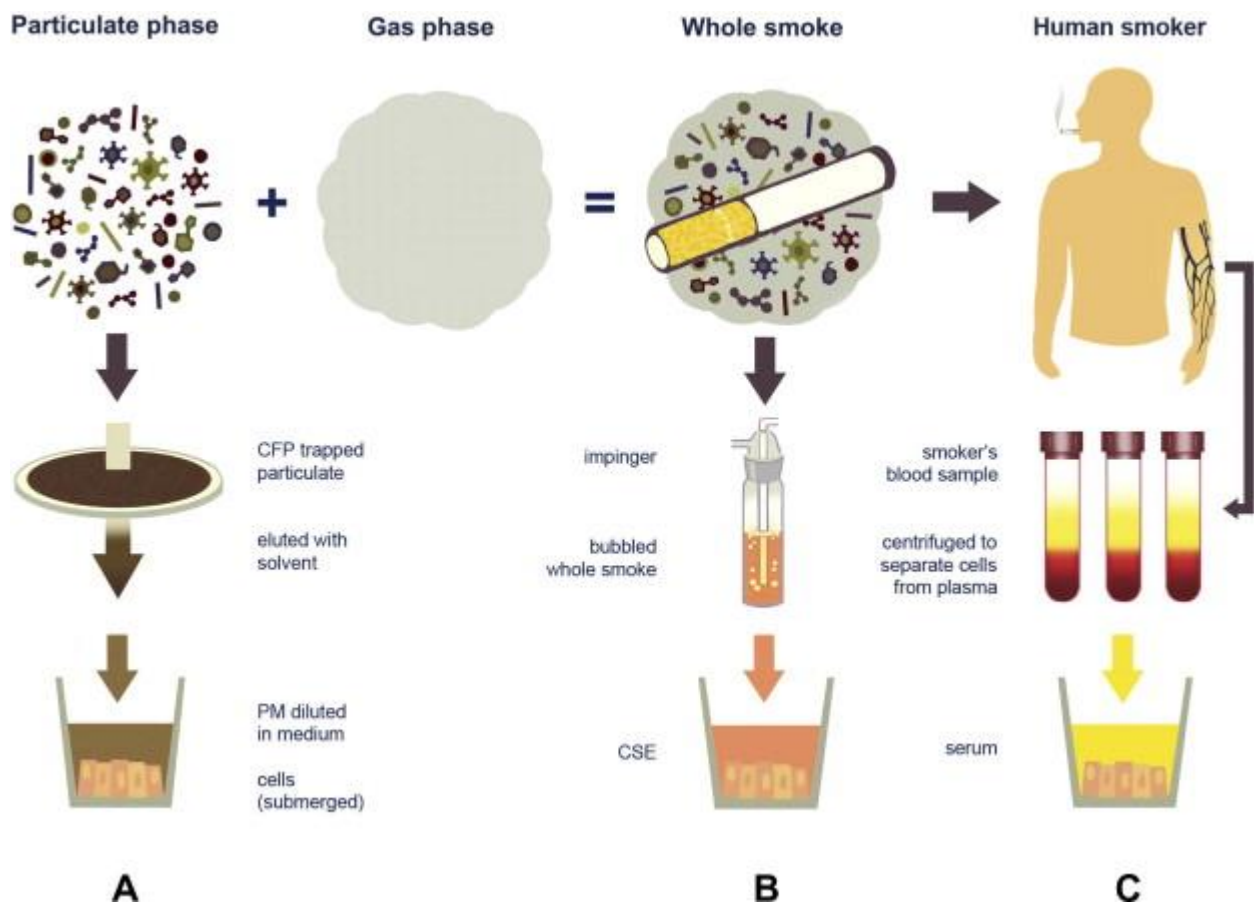


Figure 5. *in vitro* models for cigarette smoke exposure. (from Fearon et al, *Toxicology in vitro*, 2012) (A) Cigarette smoke condensate (CSC) stimulation, CFP=cambridge filter pad. (B) Cigarette smoke extract stimulation. (C). Stimulation with serum from smokers.

Regardless of the model, *in vitro* models of cigarette smoke exposure all lack at least one critical parameter, which is the chronic exposure to cigarette smoke. Indeed, *in vitro* settings limit the exposure from few seconds to 96h to a single dose of cigarette smoke extract. While those models provide a valuable insight into the cigarette smoke effects, *in vivo* confirmation of the results they provide is required. *In vivo* model will not be detailed here, as they are outside the frame of this study, but they usually involve exposing mice to a massive amount of environmental smoke either by using direct exposure to cigarette smoke in a close compartment or by forcing them to inhale smoke through a small gas mask. It is surprising to note the paucity of the literature on cigarette smoke models. Indeed very few reviews are available on this subject, underlying the lack of harmonization in modeling of cigarette smoke exposure. This greatly impairs the ability of newcomers to choose a pertinent *in vitro* model.

I.4 SASH1

I.4.a SASH1 gene

SASH1 gene, standing for SAM and SH3 domain containing 1 gene, was first described by Zeller et al in 2003¹⁰¹. It is located on the chromosome 6 at the position 6q24.3, has a size of 210kb and contains 20 exons (Figure 6). The *SASH1* gene sequence shows 85% homology between human and mouse, suggesting a high conservation of the sequence between species. Despite its big size, only two polyadenylation signals were found, both in the 3'UTR region, at positions 4407 and 7685 respectively, suggesting that alternative splicing would only affect the length of the 3'UTR region. The 3'UTR region is known to contain a variety of regulatory regions that affects stability, localization, and translation efficiency of the transcript^{102,103}. Consequently, while the two transcripts should result in the same protein sequence, their level of translation may differ. Indeed, Northern blot experiments on a collection of human tissues showed the presence of two ubiquitous transcripts of approximately 4.4 and 7.5kb. The smaller transcript was expressed in higher levels in lung, placenta, thymus and spleen. The longer transcript showed the same tissue repartition for high expression with the exception of the thymus.

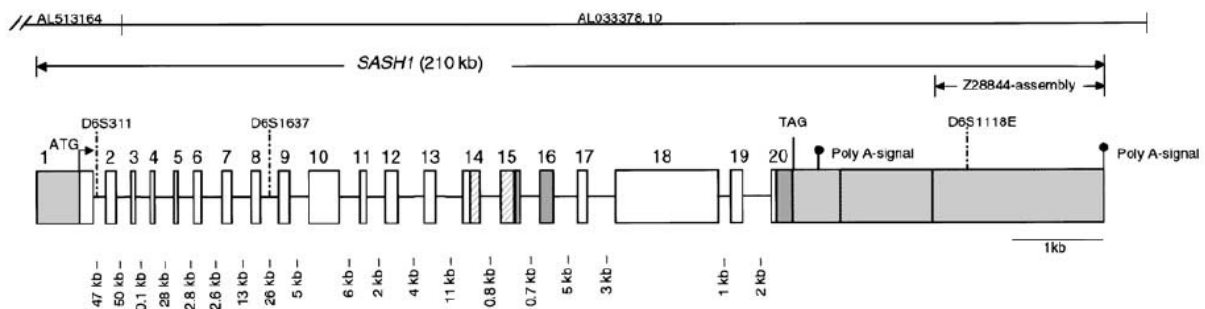


Figure 6. Schematic view of the genomic organization of SASH1. (from Zeller et al 2003). *SASH1* gene have 20 exon and two poly A signals, both localized in the 3'UTR region of the gene, suggesting that both isoforms codes for the same protein and differ only by post transcriptional regulations.

I.4.b SASH1 protein

The *SASH1* gene codes for a protein of an expected molecular weight of 137kDa (1247 amino acid). The protein contains 4 different protein interaction domains: one coiled-coil, two SAM domain (sterile alpha motif) and one SH3 domain. This pattern is typical of signal transduction proteins, and thus *SASH1* was added to this category by default. The repartition of *SASH1* has been shown in the literature¹⁰⁴ and in our experiments to be both cytoplasmic and nuclear. However, while *SASH1* possesses indeed two nuclear localization signals (NLS) in its N-terminus along with one nuclear export signal (NES) in its sequence, it lacks a DNA binding domain.

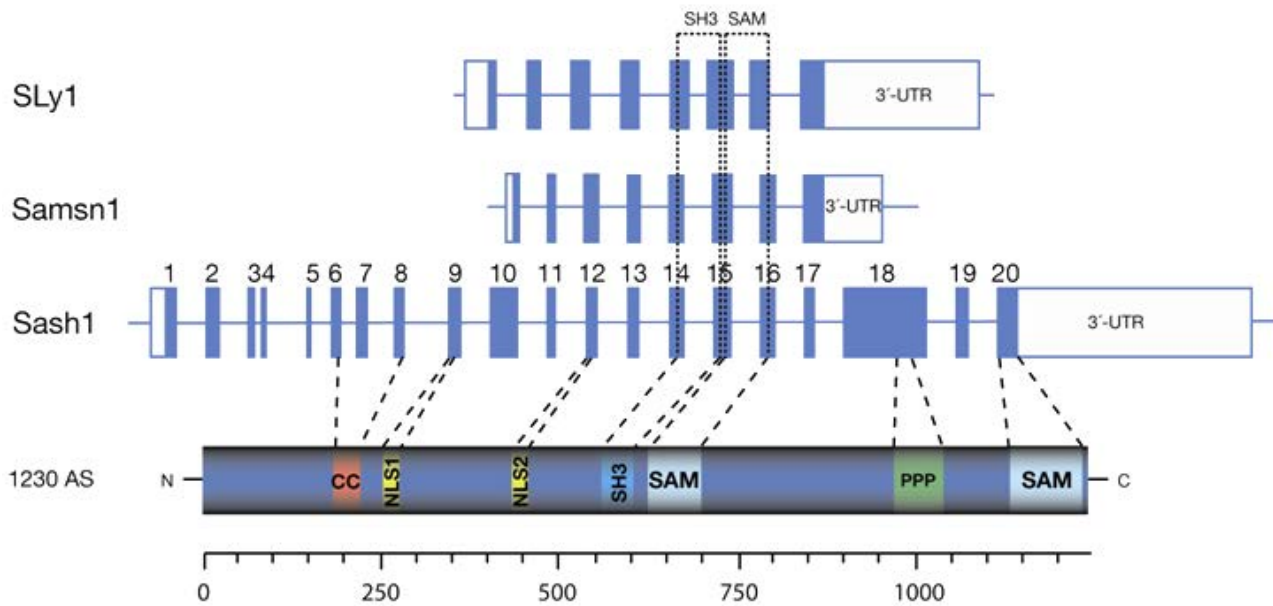


Figure 7. SLY family mRNA transcripts and SASH1 protein schematic representation. (from the Atlas of Genetics and Cytogenetics in Oncology and Hematology). SASH1, SASH2 (Samsn1) and SASH3 (SLy1) transcript and the SASH1 protein structure. The conserved SLY domain that encompasses the SH3 domain and one Sam domain (coded by exon 14 to 16) is between position 500 and 750 in the protein.

SASH1 is a member of the SLY protein family along with two other genes, SLY (also known as SASH3) and Sly2 (also known as SASH2). The latter 2 genes are mainly expressed in hematopoietic tissue, whereas SASH1 is ubiquitous. The three proteins share a common domain (sometimes called SLY domain) that encompasses the SH3 domain and one of the SAM domain (Figure 7). A study using a SASH1 protein construct containing only this conserved domain was sufficient to change the morphology of human cancer cells to similar extent as the full SASH1 protein¹⁰⁴, indicating that this conserved domain is of critical importance for SASH1 function. Little is known regarding potential post translational modification of the SASH1 protein, however, a study on the effect of ionizing radiation performed in human skin fibroblasts revealed that SASH1 was differentially phosphorylated following exposure to ionizing radiation¹⁰⁵. Thus phosphorylation is likely to play a role in SASH1 function.

I.4.c SASH1 in the literature, the quest for SASH1 function

Since its formal characterization in 2003, SASH1 has been the main focus of a small dozen of publications and also made a few appearances in more general papers. The main field in which SASH1 was studied is cancer, but it was also shown as an important actor of the skin function. As SASH1 bibliography is not abundant, this part will be organized by theme, each of them being organized chronologically to give a brief overview at the evolution in the SASH1 field.

SASH1 in cancer

SASH1 was first shown to be down-regulated in breast cancer in a first study in 2003. In this study Zeller et al studied the 6q23-25 region, a known site for allelic loss in breast Cancer¹⁰¹ and examined the expression of the genes present in this locus. They found that the mRNA expression of SASH1 was reduced in primary breast cancer tumors. Using a cancer expression profiling array, which contained cDNA from normal and tumor tissues from different organs and donors, the authors also showed that SASH1 expression was also down regulated in a number of other cancers, including colon cancer, thyroid cancer and lung cancer. These findings gave the first evidence for SASH1 implication in cancer.

In 2006, Rimkus et al¹⁰⁶ published additional evidence of SASH1 down regulation in colon cancer. They analyzed by RT-qPCR samples of normal and tumor tissues from patients (113 cases compared to 15 controls) and showed that SASH1 expression was lower in advanced and late stages of colon cancer and in liver metastases, but not in early stages and pre-cancerous tissues. The authors thus speculated that SASH1 plays a protective role against tumor invasion and dissemination, rather than in tumor initiation. Since the correlation of SASH1 expression and cancer was clearly established, the following studies focused on SASH1 function in order to determine by which mechanisms SASH1 affect the cell cycle regulation.

The first study on the subject was an extensive work on the SASH1 protein performed by Martini et al in 2011¹⁰⁴ using cancerous cell lines. In this publication the authors first showed that SASH1 was present in both nucleus and cytoplasm with an enrichment in the peripheral membrane protrusions. Using different constructs derived from SASH1 coding region they also showed that the N-terminal part of SASH1 (SASH1 protein construct with the deletion of its C-terminus) that contains the NLS, accumulated in the nucleus, while the C-terminal part (construct of SASH1 protein with the deletion of its N-terminus) accumulated in the cytosol. They also showed that SASH1 over-expression affected cell morphology by increasing the overall membrane ruffling. While N-terminally and C-terminally truncated SASH1 constructs did not affect the cell morphology, the SLY domain SASH1 construct (conserved domain of the SASH1 family proteins containing one SAM and the SH3 domain of SASH1) over-expression affected the cell morphology in comparable levels to that of the full SASH1 protein. This study also showed that SASH1 was enriched in the lamellipoda and co-precipitated with cortactin (CTTN), a regulator of the actin cytoskeleton. Thus the authors tested the effect of SASH1 expression on HeLa cells and found that SASH1 over-expression inhibited cell migration. Finally they showed that SASH1 over-expression in HEK cells increased adhesion to fibronectin and laminin but not to gelatin coated plates. This study is still the most complete to date on the SASH1 protein and showed for the first time that SASH1 affected cell migration and adhesion, two mechanisms critical for cancer invasion and dissemination.

Recent studies from 2012 focusing on SASH1 function in cancer, ^{107,108,109,110} showed that SASH1 over-expression in different cancer cell lines lead to accumulation of cells in G0/G1 phase, indicating an inhibitory effect of SASH1 on cell proliferation. Accordingly, cyclin D1 levels were

found lowered in all but one cell line and cell invasiveness *in vitro* assays showed that SASH1 over-expression lead to decreased invasiveness of the cancerous cells.

The same year SASH1 was used for the first time in a clinical study focusing on the risk assessment of metastasis by biological markers in patients with stage II colon cancer¹¹¹. The study, performed by Nitsche et al on 232 patients, showed that SASH1 expression was decreased in tumor tissues and that low level of SASH1 correlated with decreased survival rate.

As SASH1 implication in cancer was now clearly established, the question of its regulation in cancer tissue was beginning to emerge. Two similar studies focused on the methylation of the SASH1 promoter in breast cancer¹¹² and in the hepatocellular carcinoma¹¹³. Using methylation microarrays they found several methylation sites in the SASH1 promoter to be increasingly methylated in tumors than in normal tissues. Additionally treatment with 5-aza deoxy cytosine, that results in demethylation of DNA, increased SASH1 expression in cell lines from breast and hepatic cancer.

Overall, cancer studies on SASH1 revealed that its expression was reduced in several human cancers and was correlated with poor survival of the patients. The *in vitro* experiments showed that SASH1 expression affected the migration and proliferation of the cells, however the pathways involved in those processes are yet to be uncovered. Decrease of the SASH1 expression in tumors appears to be regulated through the differential methylation of the promoter region. All those discoveries potentially point out to SASH1 as a new tumor suppressor gene.

SASH1 in dermatology

The first occurrence of SASH1 in the dermatology field in 2010 was in a study focusing on the effect of ionizing radiation on protein phosphorylation in human skin fibroblasts by Yang et al¹⁰⁵. Indeed, the SASH1 protein possesses three phosphorylation sites on serine residues S407, S837 and S839, the residue S407 being specific to the low levels of ionizing radiation. Most importantly the authors showed that the exposure of fibroblasts to low ionizing radiation induced phosphorylation of 2566 phosphoproteins with SASH1 being the most phosphorylated at S407. This result already suggested an implication for SASH1 in the cell response to UV exposure.

In 2013 a Chinese group working on Dyschromatosis universalis heredita (DUH), a pathology that results in abnormal repartition of the melanocytes in the skin, discovered a mutation in *SASH1* in three non-consanguineous DUH families¹¹⁴. The mutant protein had higher half life (8h) than the wild-type protein (5h). SASH1 mutation also enhanced melanocytes mobility both *in vivo* and *in vitro*. As these effect could account for the abnormal distribution of the melanocytes in the DUH patients, the authors decided to investigate the molecular mechanisms by which SASH1 could affect this cellular process. As *in vivo* cell migration usually requires extracellular matrix degradation, they investigated the expression and the secretion of metalloproteinases and showed that SASH1 over-expression reduced the secretion of MMP1, MMP2 and MMP8. Furthermore, a GST-pulldown was also performed in a melanoma cell line over-expressing wild type SASH1, and showed that IQGAP1, calmodulin 1 and MAP2K2 co-precipitated with SASH1. All those proteins being involved in actin

cytoskeleton polymerisation and cell-cell adhesion, the authors focused their attention on E-cadherin, an adhesion molecule that is required for the adhesion of Langerhans cells to keratinocytes (a cell type involved in melanocyte repartition throughout the skin). They showed that SASH1 expression indeed down-regulated E-cadherin in a calmodulin 1 dependent fashion. Taken together, the results from this study established the first link between SASH1 mutations and a human pathology of the skin, suggesting a strong involvement of SASH1 in cell dissemination and invasiveness through the regulation of their actin cytoskeleton and adhesiveness.

Lastly, in 2014, Courcet et al¹¹⁵ published a study on a consanguineous family in which individuals, with severe skin dysfunction, shared a homozygous variant (c.1849G>A; p.Glu617Lys) for SASH1¹¹⁵. The major symptoms included recurrent skin cancers, loss of hair, abnormal skin pigmentation and palmoplantar keratoderma (excessive keratin production under the foot). Although this publication is mostly descriptive and focused on the patient phenotypes, the authors also performed *in vitro* studies and demonstrated that fibroblasts with mutated SASH1 migrated faster than fibroblasts with the wild type SASH1. Surprisingly, proliferation levels of the fibroblasts with the SASH1 mutation did not differ from those of control fibroblasts. Furthermore, survival assays performed using normal and SASH1 mutated fibroblasts showed no difference for exposure to UV or ionizing radiation. This study confirmed that SASH1 has an important role in skin function, besides its role in cancer.

In summary, studies in the field of dermatology revealed that SASH1 was implicated in the repartition of the melanocytes in the human skin. While *in vitro* studies on human fibroblasts showed that SASH1 phosphorylation was affected by ionizing radiation, however subsequent studies failed to find a difference in the fibroblasts survival between wild type and mutated SASH1 exposed to UV. SASH1 was also found to regulate human melanocytes invasiveness and human fibroblasts migration which corroborates the findings made by cancer studies. Furthermore these works found several other partners for SASH1, including IQGAP1, calmodulin 1 and MAP2K2. The latter three proteins are involved in the actin cytoskeleton and adhesiveness regulation in cells, comforting a role for SASH1 in those cellular processes.

Other significant findings regarding SASH1 in the literature

While the previously described studies on SASH1 are quite consistently pointing out SASH1 as a regulator of the actin cytoskeleton, an extensive study, by Daupinee et al¹¹⁶, regarding a potential role for SASH1 in the immune response, shed a new light on SASH1. This study focused on identifying the downstream signal proteins of toll like receptor 4 (TLR4) in the lipid rafts from mouse embryonic fibroblasts after lipopolysaccharide (LPS) stimulation. Comparing FADD (a negative regulator of TLR4) knockdown mice to wild type they found SASH1 to be present following LPS stimulation in the lipid rafts of the FADD KO mice but not in the lipid rafts from wild type. Using Human Microvascular Endothelial Cells (HMEC) they showed that SASH1 positively regulated LPS signal transduction. Using SASH1 over-expression followed by immune-precipitation in HMEC they showed that SASH1 co precipitated with several members of an NFkB activating complex, TRAF6, IkbkB, IkbkA and TAK1. Thus they hypothesized that SASH1 was a scaffold protein for this

transduction signal pathway and was a positive regulator of the TLR4 mediated NFkB activation. This work has opened several interesting perspective on SASH1 function, but so far stands rather isolated within the SASH1 literature.

Other studies showed correlation between SASH1 expression and various pathologies. SASH1 was found up-regulated in the placenta from women with severe pre-eclampsia¹¹⁷, and down-regulated in the medial temporal lobe in patients with Alzheimer disease¹¹⁸. Regarding single nucleotide polymorphism (SNP), it was shown that SNP rs6930576, located in an intronic region of SASH1, was strongly associated to diabetic nephropathy in a GWAS study in an Afro-American population¹¹⁹. Still in the field of diabetes, a study by Huang et al¹²⁰, found through a genome wide screening that SASH1 was a negative regulator of insulin-mediated inhibition of FOXO1A. All these results place SASH1 as an actor in diverse pathologies, however those results are mostly correlative and require further studies to determine the exact implication of SASH1 in human pathologies.

I.5 Thesis project: from Genome Wide Analysis in the Gutenberg Health Study to *in vitro* experiments

The present study is in the continuity of the collaboration between our laboratory (Dr François Cambien/Dr Laurence Tiret and Dr Ricardo Verdugo) and Pr Tanja Zeller and Pr Stefan Blankenberg (from the Molecular cardiology group of the UKE in Hamburg, Germany). This study is designed as a community-based, prospective, observational single-center cohort study in the Rhein-Main region in Germany. Its main aim is to identify new genetic and non-genetic risk factors contributing to cardiovascular disease, with a particular focus on atherosclerosis, thus allowing better individual cardiovascular risk prediction. Inclusion criteria to participate in this cohort were an age ranging from 35 to 74 years old, and giving a written consent. Exclusion criteria were restricted to insufficient knowledge of the German language or mental disability that would impair examination, instructions and explanations at the study center. Individuals received clinical examinations where known CVD risk factors were acquired and blood samples collected. Circulating monocytes were then negatively selected from the blood samples and processed for DNA/RNA extractions. Among the subjects, 1500 had both RNA and DNA available, allowing transcriptomic measurements. Among the cardiovascular risk factors were smoking status and the number of carotid plaques (measured by echo Doppler). Zeller et al³⁸ studied by microarrays the correlations between gene expression and cardiovascular risk factors and showed that in circulating monocytes the expression of several genes and especially that of SASH1, was positively correlated to the smoking status. In the same study, the authors also found SASH1 expression was positively correlated to the number of plaques. SASH1 was still associated with plaques after adjustment for smoking, suggesting that it may affect atherosclerosis independently of smoking. Following this study, Verdugo et al¹²¹ performed new statistical analyses using the GHS transcriptome data to predict the pathways linking cigarette smoking to atherosclerosis (Figure 8).

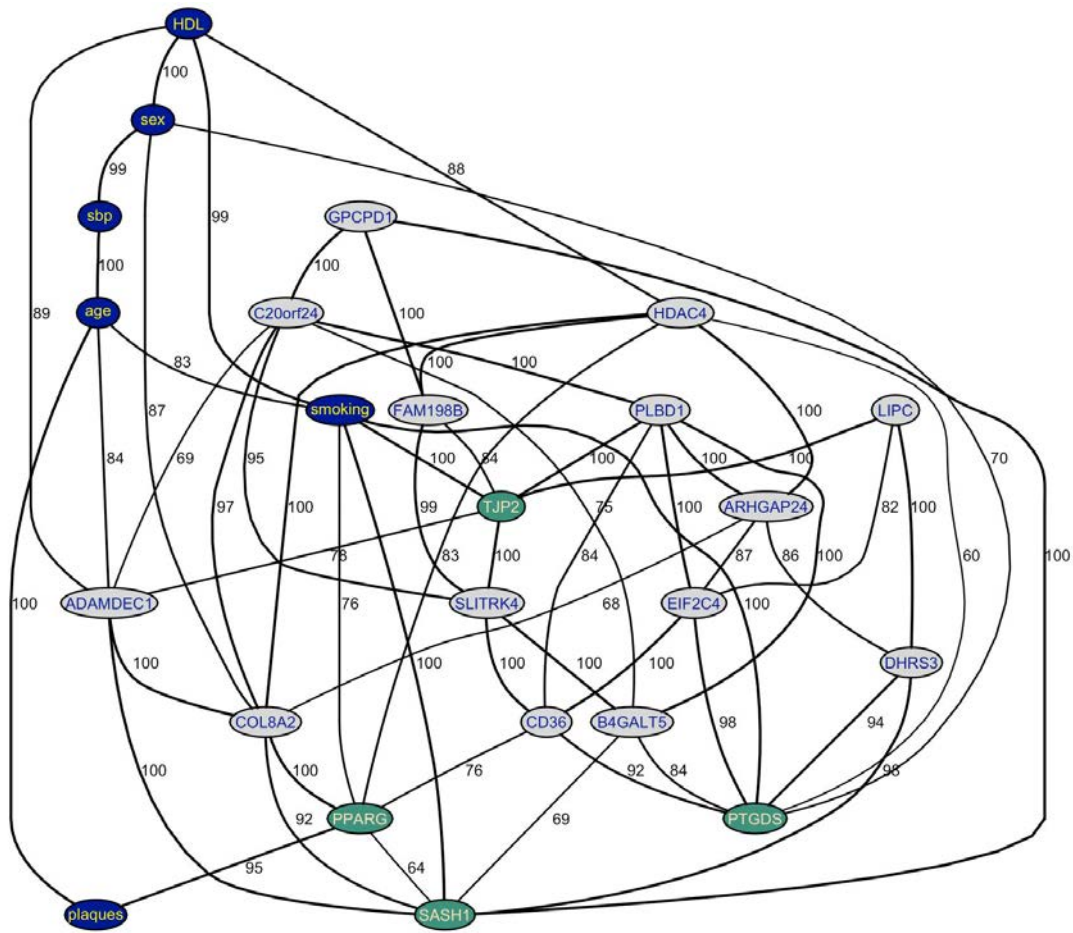


Figure 8: SASH1 predicted network linking smoking to atherosclerosis (from Verdugo et al 2013). This graph represents a consensus network from 1000 bootstraps. Edges among variables are drawn if detected in at least 60% of bootstrapped samples. The recovery percentages are indicated to the right of the medial section of each edge. Line thickness is proportional to the edge's partial correlation. Plaques and risk factors are in blue. Genes directly connected to smoking are in green. Other genes are in gray.

In this latter study, in which I am also a co-author, *SASH1* was found to be the most correlated gene to smoking. The largest reduction in co-variation between smoking and plaques was also observed when adjusting for *SASH1* expression, suggesting that *SASH1* plays an important role in cigarette smoking-driven aggravation of atherosclerosis. Another gene *SLC39A8* (also known as *ZIP8*), a divalent metal ion transporter, also came up in this study as the most significantly correlated with the number of plaques (Figure 16).

This thesis project aimed at investigating whether or not these *in silico* predicted candidates and their related pathways could be confirmed in cellular models and to further characterize the mechanisms linking these candidate genes to smoking and to atherosclerosis.

I.6 Aim of the thesis and experimental strategy

This study aimed to assess the link between SASH1, smoking and atherosclerosis. For this, several questions were raised at the beginning of the study:

What is SASH1 vascular expression?

As *SASH1* expression was to be characterized in vascular tissue at the beginning of this study, we used RNA extraction, quantitative RT-PCR, immunohistochemistry and immunocytochemistry to determine *SASH1* mRNA and protein expression in the vascular wall

Is the increased SASH1 expression previously observed in human monocytes from smokers also present in the vascular wall?

After confirmation of *SASH1* expression in vascular tissue, quantitative RT-PCR was used to test, whether increased *SASH1* expression observed in human monocytes was also present in the vascular wall.

Which cellular and molecular pathways link SASH1 expression and atherosclerosis?

Subsequently it was plan to determine by which cellular and molecular processes *SASH1* could be linked to atherosclerosis. As we wanted to investigate the effect of *SASH1* on proliferation we choose to focus our study on human aortic endothelial cells (HAECs), a cell type that is able to proliferate and that is in direct contact with the blood stream. This makes endothelial cells a direct target for cigarette smoke derived metabolites. We first performed immunocytochemistry and fractionated lysis of cells followed by western blot to determine the subcellular localization of *SASH1* in HAECs.

To determine in which molecular pathways *SASH1* is involved different strategy at the RNA and protein level were used:

-We performed a transcriptomics analyses by using a series of microarrays on *SASH1* silenced HAECs, to identify the genes and pathway affected by the disappearance of the *SASH1* protein. siRNA silencing was confirmed by the diminution of the protein after 48h by western blot analyses.

-After identification of potential *SASH1* interacting protein by the yeast two hybrid technique, *SASH1* co-immunoprecipitations were performed in HAECs to confirm the candidate interacting proteins. As the available antibodies targeting these proteins were quite unspecific and therefore did not allow us to confirm any candidate, we went for a different approach and used mass spectrometry to identify candidate proteins that where precipitated with *SASH1* in HAECs. In order to ensure specificity of the proteins we compare *SASH1* silenced HAECs to HAECs transfected with control siRNA.

- Finally, we tested the in vitro effect of *SASH1* knockdown on proliferation, migration and angiogenesis in HAECs using tetrazolium salt (MTT), wound healing and matrigel assay respectively. As those processes are involved in plaque formation and rupture, these results could form the first trail of evidence to explain *SASH1* implication in atherosclerosis.

II Article: SASH1 a new link between cigarette smoke and atherosclerosis (accepted in Atherosclerosis)

Abstract

Objective: We have previously reported that SASH1 expression is increased in circulating human monocytes from smokers and was positively correlated with the number of carotid atherosclerotic plaques. The aim of this study was to further validate the link between smoking, SASH1 and atherosclerosis within the vascular wall and to assess the impact of SASH1 expression on endothelial cell functions.

Method: Human carotids with atherosclerotic plaques were obtained from 58 patients (45 of them with known smoking status: smoker, non-smoker, ex-smokers) , and were processed for gene expression analyses and immunostaining. To investigate its function, SASH1 was silenced in human aortic endothelial cells (HAECs) using two different siRNA and subcellular localization of SASH1 was determined by immunostaining and subcellular fractionation. Subsequently the transcriptomic analyses and functional experiments (wound healing, WST-1 proliferation or Matrigel assays) were performed to characterize SASH1 function.

Results: SASH1 was expressed in human vascular cells (HAECs, smooth muscle cells) and in monocytes/macrophages. Its tissue expression was significantly higher in the atherosclerotic carotids of smokers compared to non-smokers ($p < 0.01$). In HAECs, SASH1 was expressed mostly in the cytoplasm and SASH1 knockdown resulted in an increased cell migration, proliferation and angiogenesis. Transcriptomic and pathway analyses showed that SASH1 silencing results in a decreased CYP1A1 expression possibly through the inhibition of TP53 activity.

Conclusion: We showed that SASH1 expression is increased in atherosclerotic carotids in smokers and its silencing affects endothelial angiogenic functions; therefore we provide a potential link between smoking and atherosclerosis through SASH1 expression.

Word count: 245

SASH1, a new potential link between smoking and atherosclerosis.

^{1,2,3}Henri Weidmann, ¹Zahia Touat-Hamici, ¹Herve Durand, ^{2,3}Christian Mueller, ⁴Frédéric Charlotte, ⁵Klaus-Peter Janssen, ⁶Ricardo Verdugo, ¹Francois Cambien, ^{2,3}Stefan Blankenberg, ¹Laurence Tiret, ^{2,3}Tanja Zeller* and ¹Ewa Ninio*§

¹Sorbonne Universités, UPMC, UMR_S 1166-ICAN, Genomics and Pathophysiology of Cardiovascular Diseases, Institute of Cardiometabolism and Nutrition, ICAN, Pitié-Salpêtrière Hospital, F-75013, Paris, France.

²University Heart Center Hamburg, Clinic for general and interventional Cardiology, Hamburg.

³German Center for Cardiovascular Research (DZHK e.V.) Partner Site Hamburg, Lübeck, Kiel, Hamburg, Germany.

⁴Department of Pathology, Hôpital de la Pitié-Salpêtrière, AP-HP, Paris, France

⁵Department of Surgery, Klinikum rechts der Isar, Technische Universität München, IsmaningerStrasse 22, Munich, Germany.

⁶Programa de Genetica Humana ICBM, Facultad de Medicina, Universidad de Chile, Chile.

* Both authors contributed equally to this work

§ Author for correspondence: Ewa Ninio, Sorbonne Universités, UPMC, UMR_S 1166-ICAN, 91, Boulevard de l'Hôpital, 75634 Paris cedex 13, France; email ewa.ninio@upmc.fr; phone: +33140779768.

Figures: 6; Tables: 2

Word count: 4747; **References:** 47

List of Abbreviations :

HAECs: human aortic endothelial cells; **HCASMCs:** human coronary artery smooth muscle cells; **SASH1:** SAM and SH3 domain containing 1; **SAM:** sterile alpha-motif; **SH3:** Src-homology 3; **WST-1:** (4-[3-(4-iodophenyl)-2H-5-tetrazolio]-1-3-benzene disulfonate; **CCND1:** cyclin D1; **CCND3:** cyclin D3.

Key words: atherosclerosis, SASH1, smoking, CCND3, CCND1, TP53.

Introduction

Atherosclerosis is the leading cause of death in developed countries. It is characterized by an excessive accumulation of lipids in the arterial wall, which leads to endothelial dysfunction, the first step towards the early onset of fatty streaks and subsequent plaque formation¹²². Both fatty streak and plaque are composed of macrophages, dedifferentiated smooth muscle cells and foam cells derived from these cell types. Plaque rupture causes stroke, myocardial infarction or ischemia of inferior limbs.

Endothelial cells play an important role in the onset of atherosclerosis as their dysfunction triggers leukocyte transmigration into the intima of the vessels leading to foam cell formation¹²³, a hallmark of fatty streak and plaque¹²⁴. Furthermore by playing an important role in intra-plaque neovascularization, endothelial cells contribute to intra-plaque hemorrhages and plaque instability. Among the factors influencing plaque formation, it has been established that smoking increases lesion size and number and is correlated with earlier cardiovascular events^{125,126,127}. Cigarette smoke is composed of more than 5.000 different compounds, including approximately a 100 known to have carcinogenic, cardiovascular or pulmonary adverse effect¹²⁸. Its action on atherosclerosis is exerted through multiple pathways, implicated in a wide range of cellular processes such as inflammation, cell migration and proliferation^{129,130}.

Previously, we showed in a large population based study that the expression of *SASH1* (SAM and SH3 domain containing 1) in circulating monocytes was markedly increased in smokers when compared to non-smokers and was also positively correlated with the number of carotid plaques, indicating a potential role of *SASH1* in atherosclerosis^{131,132}. *SASH1* is a member of the SLY signal-adaptor protein family and encodes an ubiquitously expressed signal transducing protein¹³³. It contains one SH3 domain, one coiled-coil and two SAM domains, all involved in protein-protein interactions¹³⁴. *SASH1* expression is down-regulated in colorectal and breast cancer^{133,135,136} and *SASH1* overexpression is known to decrease proliferation and invasiveness of various cancerous cell lines while increasing their apoptosis rate^{137,138,139}. At the molecular level, *SASH1* has been suggested to be involved in the regulation of the cytoskeleton polymerization¹³⁴, to be a scaffold protein downstream of the TLR4 receptor leading to an increased NFkB activation after liposaccharide (LPS) stimulation¹⁴⁰ as well as a being part of the Gas-IQGAP1-E-Cadherin pathway¹⁴¹.

The present study was conducted to investigate the potential role of *SASH1* in atherosclerosis. Our data support a possible implication of *SASH1* in the pathology of atherosclerosis. We show that

SASH1 is expressed in the arterial tissues and that SASH1 expression is increased in carotid plaques from smokers. In primary human aortic endothelial cells (HAECs), SASH1 siRNA silencing resulted in an increased proliferation, migration and angiogenesis.

Materials and methods

Human tissue samples

Fragments of fresh human full-thickness carotid resections ($n = 58$) were obtained from the Department of Vascular Surgery at the Pitié-Salpêtrière Hospital in compliance with institutional ethical guidelines and in accordance with the Helsinki Declaration of 1975 as revised in 1983. All patients (age 51 to 91; 19 females and 39 males) signed written informed consent. Briefly, plaques were dissected; and total RNA was immediately extracted using the trizol method (Life Technologies, 15596). Patients were separated in three groups: “non smokers” including individuals that never smoked ($n=25$), “ex-smokers” including individuals that stopped smoking prior to the surgery (from 6 month up to 40 years, $n=9$) and “smokers” including individuals still smoking at the time of surgery ($n=11$).

Cell isolation and culture

Human Aortic Endothelial Cells (HAECs) and Human Coronary Artery Smooth Muscle Cells (HCASMCs) were purchased from PromoCell. Cells were cultured in EGM2 medium (Lonza, CC-3156 & CC-4176) and smooth muscle cells growth medium (PromoCell, C-22162), respectively. All experiments were performed using cell passages 3 to 7. Human monocytes were isolated from human buffy coats (Etablissement Français du Sang, Rungis, France) and differentiated into macrophages as previously described¹⁴². Macrophages were cultured for 6 days in RPMI 1640 medium with, 50 µg/ml gentamycin, 10% of fetal calf serum and 10ng/ml MCSF (Life Technologies, 11875-093

siRNA-mediated silencing of SASH1

Transfection of HAECs was performed using RNAiMAX (Life Technologies, 13778). Briefly, using RNAiMAX transfection reagent (Life technologies, 13778-100), two different siRNA targeting *SASH1* (Ambion, S23572 and S23574) at a concentration of 20 nM were applied separately and compared to control siRNA (Ambion, 4390843). The medium was changed after 24h, total RNA and proteins were harvested 48h after transfection.

Efficiency of siRNA transfection and SASH1 knockdown was assessed by RT-qPCR and western blotting as described below.

Wound healing, proliferation and angiogenesis assays

Transfected HAECs were cultured for 48h to reach 90% confluence and were treated with 20 μ M mitomycin-C (Sigma-Aldrich, M4287) for 90 min. Two linear scratches per well were performed in the cell monolayer using a 200 μ l pipette tip and were incubated at 37°C and 5% CO₂ for next 24h. Images of the gap were obtained at time 0h and 24h with fully motorized inverted microscope (Nikon Eclipse TiS) using bright field and 4X magnification. The wounded area was analysed using ImageJ software (NIH) by quantification of the surface of wounded area at 24h as compared to time 0h.

HAECs proliferation was evaluated using a colorimetric tetrazolium salt WST-1 (4-[3-(4-iodophenyl)-2H-5-tetrazolio]-1-3-benzene disulfonate) assay, based on the conversion of WST-1 into formazine by mitochondrial dehydrogenase enzyme in viable cells. After 48h of transfection, cells (5,000 cells/well) were seeded in 96-well plates and cultured for further 24h in 100 μ l medium. Then 10 μ l of WST-1 solution were added to each well, followed by 3h incubation. Absorbance was measured at 450 nm in a microtiter plate reader.

HAECs transfected with siRNA targeting SASH1 or control siRNA were cultured on Matrigel (BD Bioscience, 354234) and tubular formation network was evaluated at 24h. Briefly, 10 μ l of Matrigel were pipetted into 15 wells slides (Ibidi, 81506) and brought to 37°C for 30 min to induce polymerization of the matrix. Subsequently, cells were seeded at a density of 50,000 cells per well. Tubular structures were quantified by manual counting on imageJ.

RNA extraction, reverse transcription and quantitative polymerase chain reaction (PCR)

Total RNA was extracted using the mirVana kit (Life Technologies, AM1560) and RNA quality was assessed by a 2100 Bioanalyzer (Agilent Technologies). Synthesis of cDNA was carried out using the Super Script II Reverse Transcriptase (Life Technologies, 18064-014). Real-time PCR was performed using Mx3005P QPCR System (Agilent Technologies) and SYBR Green (Thermo Scientific, AB-1158/B). The amplification program was: 95 °C for 15 min, 40 cycles of 95°C for 30 s and one cycle of 95°C for 1 min, 60°C for 30 sec and 95°C for 30 sec. Data were analyzed using MxPro® software using the $\Delta\Delta$ CT method¹⁴³ and normalized to the ubiquitin or GAPDH control genes. The sequence of primers is listed in Table 1.

Genome-wide expression analysis and pre-processing of expression data

Total RNA was extracted from HAECs using the mirVANA kit (Life Technologies, AM1560). Transcriptome analysis of total RNA was performed using the *Illumina* HT-12 v4 BeadChip (<http://www.Illumina.com>). Briefly, RNA samples were processed in batches of 24 samples. 250 ng of total RNA was reverse transcribed, amplified and biotinylated using the IlluminaTotalPrep RNA Amplification Kit (Ambion/Applied Biosystems, AMIL1791). Each biotinylated cRNA (750 ng) was hybridized to a single BeadChip at 58°C for 16–18h. BeadChips were scanned using the *IlluminaHiScan* array.

The summary probe-level data delivered by the *Illumina* scanner (mean and SD computed over all beads for a particular probe) was loaded in *Genomestudio*. The pre-processing was done by the *Illumina* software, at the level of the scanner and by *Genomestudio* included: correction for local background effects, removal of outlier beads, computation of average bead signal and SD for each probe and gene, calculation of detection *P*-values using negative controls present on the array, quantile normalization across arrays, check of outlier samples using a clustering algorithm, check of positive controls. Analyses were carried out on the mean level for all probes in each gene. To stabilise variance across expression levels, we applied an arcsinh transformation to the expression data. Compared to a log transformation, this transformation has the advantage not to discard negative expression values which can occur in *Illumina* data. A gene was declared significantly expressed in the dataset, i.e. expressed above background (as measured by the negative controls present on each array), when the detection *p*-value calculated by genome studio was <0.05 in more than 5% of the samples. Microarray analysis was performed using the *Illumina* HT-12 v4 BeadChip. To limit batch effect, each sample treated with the SASH1 siRNA was processed on the same beadchip as its siRNA control.

Western blotting

Cells were washed 3 times with PBS and lysed on ice with RIPA (Thermo scientific, 89900) supplemented with EDTA and protease inhibitors (Thermo Scientific, 78410). Equal amounts of protein were loaded and separated on an SDS-PAGE gel of 6, 8 or 12%. Proteins were transferred onto nitrocellulose 0.2 µm membrane (Biorad) using wet transfer (Tris/glycine, 20% ethanol) overnight (15h) at 30V. Detection was performed with ECL Plus detection kit (GE-Healthcare) using either the digital camera (GE Healthcare, LAS 4000) or the Agfa Gevaert, CURIX 60 system. Primary antibodies for detection were as follows: anti-SASH1 (Novus, NBP1-26650), anti-CCND3 (Santa Cruz Biotechnologies Sc-182), anti-CCND1 (Cell Signaling, # 2926), anti-GAPDH (Cell

Signaling #21185), anti-tubulin (Invitrogen, 322500). The following secondary antibodies: anti-goat (Santa Cruz Biotechnology, Sc-2020), anti-rabbit (Vector laboratories, PI1000) and anti mouse (Vector Laboratories, PI2000) coupled with horseradish peroxidase were used for detection.

Immunostaining of cells and tissues

HAECs were fixed with paraformaldehyde (4%), permeabilized with Triton-X100 (0.5%) and incubated at room temperature with phalloidin-Alexa 488 (1:1000; Life Technologies, A12379) for 1h at room temperature. Subsequently, SASH1 antibodies, 1:400 (Novus Biological, NBP1-26650 or NBP1-26651), 1:800 (Sigma Aldrich S8073) or 1:50 (Santa Cruz Biotechnology Sc-169253) were added for 2h at room temperature. Secondary anti-rabbit and anti-goat Alexa 594-conjugated antibodies were used for detection (Life Technologies, A11036 and A11037). Nuclei were counterstained with DAPI. Microphotographs were obtained with a confocal microscope (Zeiss, LSM 710).

For SASH1 expression in vascular tissue, the immunostaining was performed on formalin-fixed, deparaffinized 3- μ m thick sections. The antigen retrieval reaction was carried out in a water bath for 20 min at 97°C using EDTA buffer (pH 8.0) for all antibodies. The slides were then incubated with the primary antibodies for 60 min at room temperature. Thereafter, the procedure was performed according to manufacturer instructions by using a biotin-free polymeric visualization system (Ultravision LP ref TL-015-HD, Lab Vision, Fremont, CA) for SASH1 antibodies from Novus Biological (1:300, NBP1-26650).

Statistical analysis

All data are presented as mean \pm standard deviation (SD) from at least 5 independent experiments of different donors. Comparison between groups was performed using Wilcoxon or Mann-Whitney test, depending on the experiment, with a significance threshold at $p \leq 0.05$ using Graphpad Prism 6 (version 6.05).

Statistical analysis of microarray data was performed in the R environment (version 3.0.1) using the *lumi* package of Bioconductor¹⁴⁴. The variance stabilizing transformation¹⁴⁵ was applied to raw data. A variance-based filtering of gene expressions was further performed using the *nsFilter* function from the *genefilter* package with default parameters (interquartile range and variance cutoff of 0.5). Differential expression analysis was performed using the *limma* package of Bioconductor. To account for the pairing of each siRNA sample with its control, the difference between gene

expressions of both samples was considered as the variable to be tested to 0. Analysis was further adjusted for beadchip. The Benjamini-Hochberg procedure was applied to adjust for multiple testing.

Results

SASH1 mRNA is expressed in vascular cells and in human carotid tissues

We first investigated *SASH1* expression in human aortic vascular cells and carotid tissues by RT-qPCR. *SASH1* transcript was present at detectable levels in all samples (Figure 1). The expression levels were similar in all tested cell types (HAECs, HCASMCs, monocytes and macrophages), as well as in healthy tissue (mammary arteries), fatty streaks and plaques. To ensure the specificity of the qPCR amplicon, *SASH1* siRNA silencing was performed on HAECs using two different siRNA. Both siRNA showed a >85% decrease of the detected *SASH1* transcript.

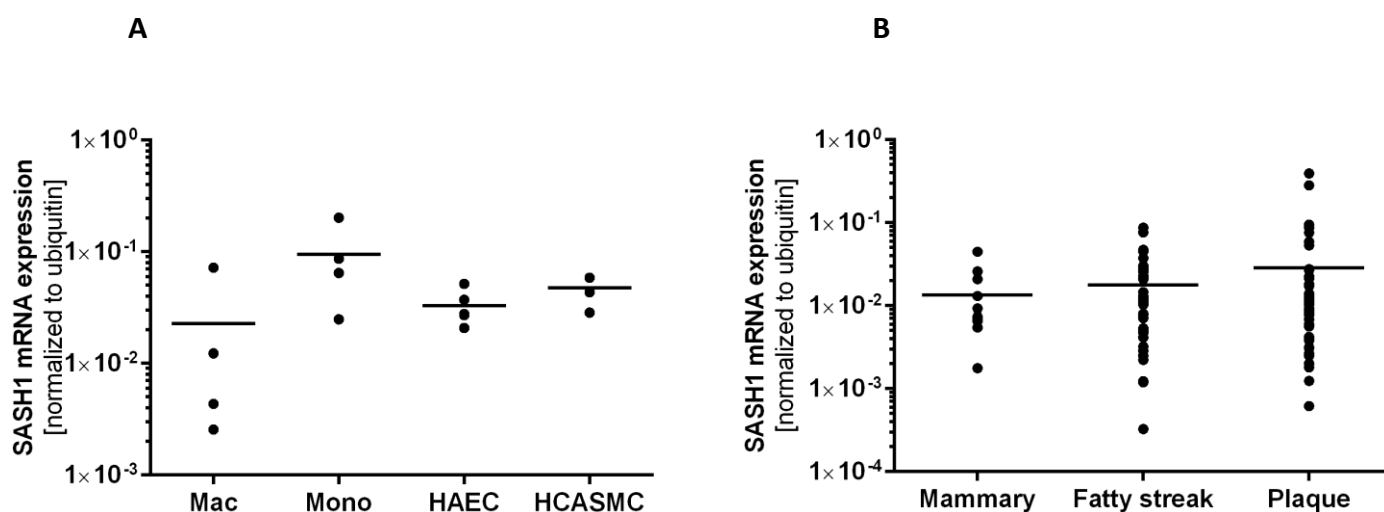


Figure 1A. *SASH1* mRNA expression in human vascular cells (A). *SASH1* mRNA expression in human macrophages (Macro, n=4), monocytes (Mono, n=4), human aortic endothelial cells (HAECs, n=5) and human coronary artery smooth muscle cells (HCASMC, n=4); (B) *SASH1* mRNA expression in the vascular wall, healthy tissue (mammary arteries), n=11, fatty streak, n=36 and plaque, n=53. Ct value for ubiquitin were between 17-22, for *SASH1* 22-27).

To characterize the distribution of *SASH1* in the arterial wall, we performed an immunostaining on human atherosclerotic carotid tissues and showed that *SASH1* was expressed in the media, intima and in macrophages within the vascular wall (Figure 2). *SASH1* staining was also present in the intra plaque capillaries. Thus, *SASH1* is ubiquitously expressed in all cell types present in human arteries and in the atherosclerotic plaque.

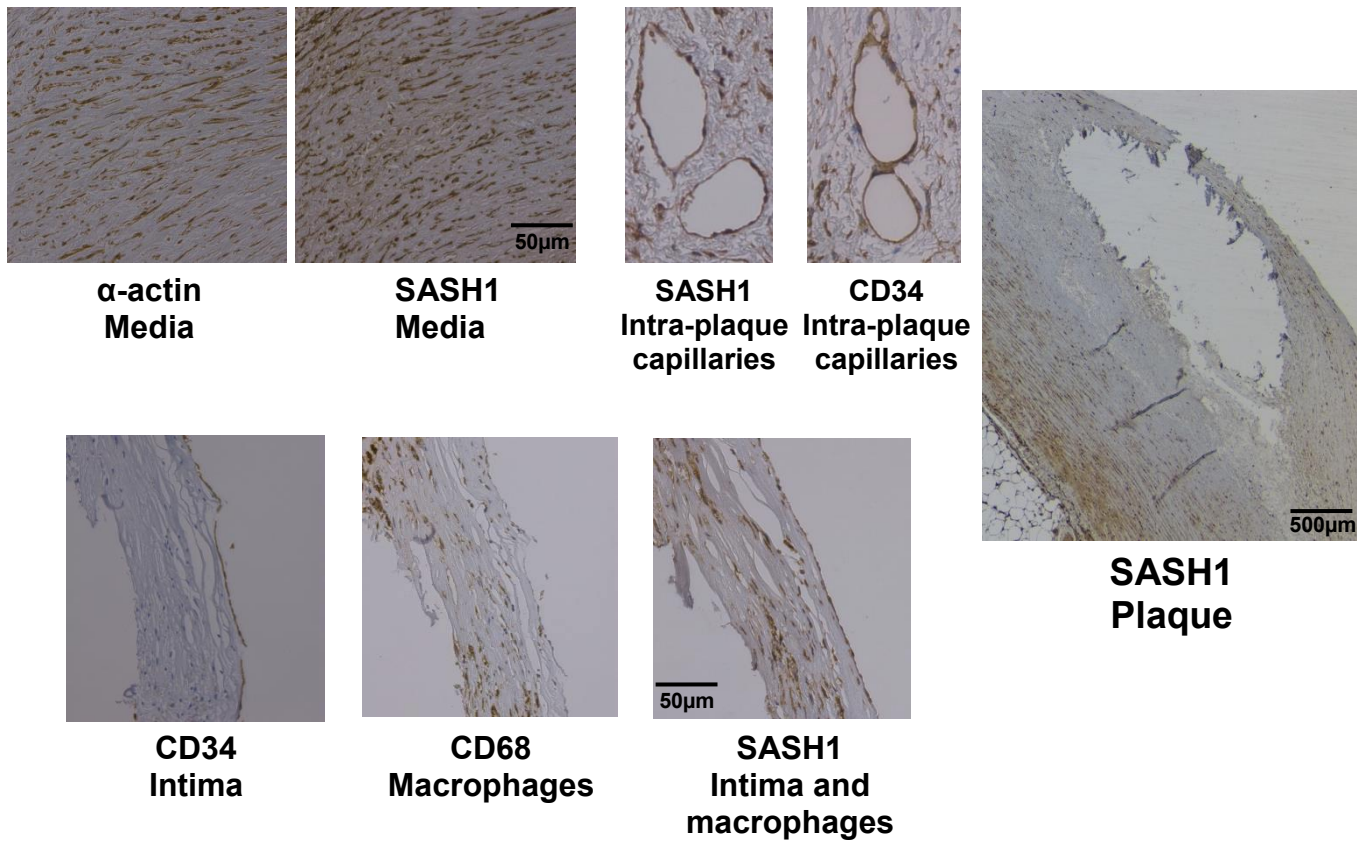


Figure 2. SASH1 protein expression in the human carotid plaque. Human atherosclerotic carotid immunostaining with anti-SASH1 (Novus NBP1-26650), with anti-CD68 (macrophages/foam cells), a-actin (smooth muscle cells) and CD34 (endothelial cells).

SASH1 protein localization in human aortic endothelial cells

The subcellular distribution of SASH1 in HAECs, determined using SASH1 antibody and confocal microscopy analysis, was shown to be mainly cytoplasmic, with a weak nuclear staining (Figure 3A). Immunostaining was performed with four different commercial antibodies that gave similar results (Suppl Fig 1) and was confirmed using fractionated lysis of HAECs followed by western blotting which showed a strong signal at the expected 180kDa size for SASH1 in the cytoplasm and a weaker signal in the nucleus. The specificity of this signal was validated using two different siRNA targeting SASH1. Transfection of both siRNAs reduced the expression of SASH1 protein to a similar extent (Figure 3B).

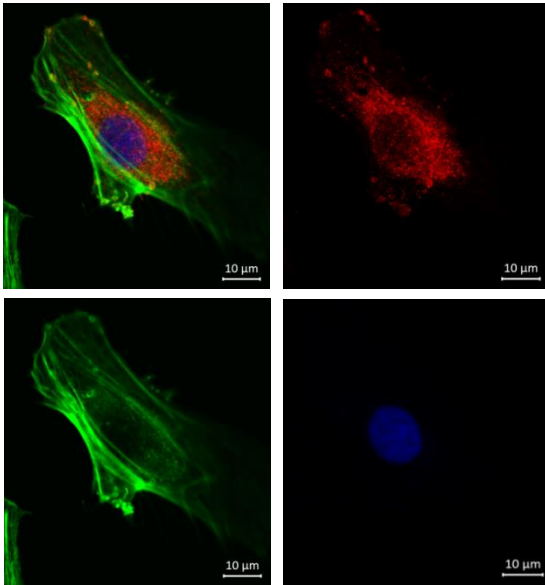
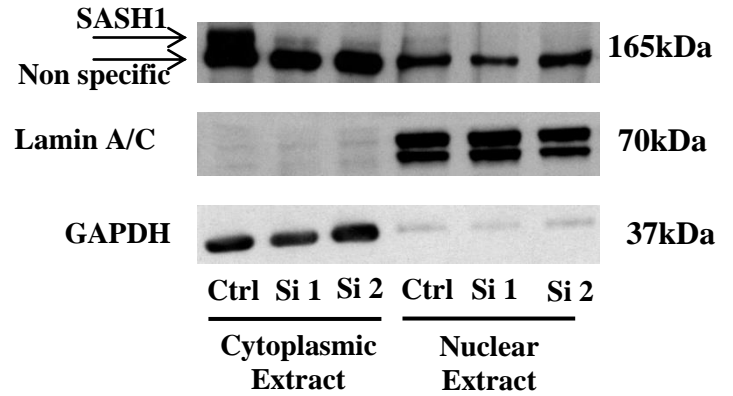
A**B**

Figure 3. Subcellular localization of SASH1 in HAECs. (A) Confocal micrographs of HAECs stained for SASH1 (Novus Biological NBP-1 26650, red), actin cytoskeleton (Phalloidine-488, green) and nucleus (DAPI, blue) and negative control rabbit IgG (Supplementary Figure 1); (B) Western blots of HAECs either transfected for 48h with siSASH1 (si1 and si2) or siCtrl (ctrl) and fractionated into nuclear (Lamin A/C) and cytosolic compartment (GAPDH).

As SASH1 contains a nuclear localization sequence (NLS) and a nuclear export signal (NES), we tested a possible translocation of SASH1 into nucleus using various inflammatory stimuli in HAECs: LPS (100 ng/ml), cigarette smoke condensate (40 μ g/ml), TNF α (10 ng/ml) or VEGF (50 ng/ml) and incubation for 1, 6 and 24h. We did not observe SASH1 translocation in these various experimental settings (data not shown).

SASH1 mRNA expression is increased in carotids from smokers

To determine if the increase of *SASH1* mRNA expression observed in monocytes of smokers in our previous study¹³¹ was also present in the arterial wall, *SASH1* expression was tested in human carotid plaques from smokers (n=11), ex-smokers (n=9) and non-smokers (n=25). *SASH1* mRNA expression was significantly increased in carotids from smokers when compared to non-smokers (p<0.01) and ex-smokers (p<0.001) (Figure 4). No difference was observed between non-smokers and ex-smokers (p=0.38), suggesting that cigarette smoke-induced increase of *SASH1* expression can be reverted.

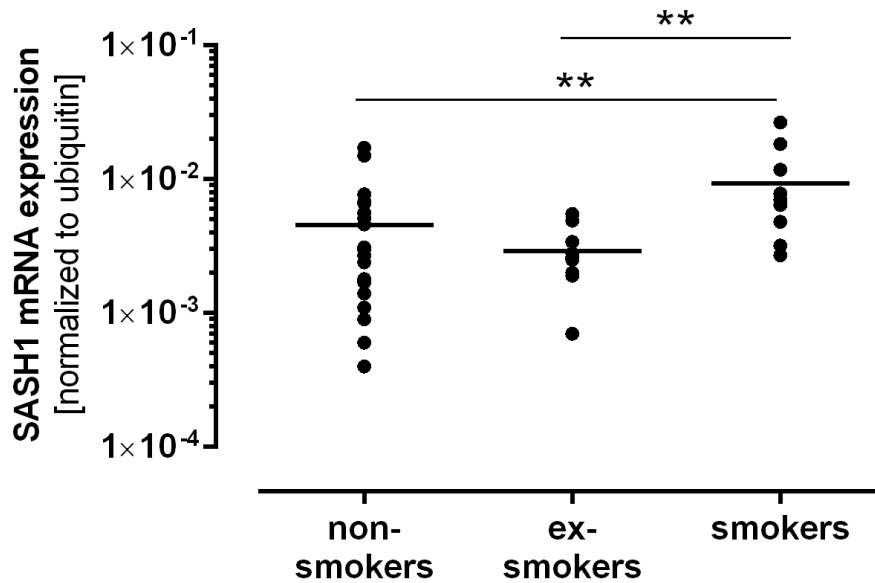


Figure 4. SASH1 expression in human carotids depending on the smoking status.

Expression of SASH1 mRNA in human atherosclerotic carotids from non-smokers, ex-smokers and current smokers.

**indicates a $p < 0.01$

To investigate the acute effect of smoking on *SASH1* mRNA expression, HAECs were incubated with cigarette smoke condensate at 10 and 40 $\mu\text{g/ml}$. Expression *CYP1A1* and *CYP1B1* was used as the positive controls for cigarette smoke condensate exposure as previously described¹⁴⁶. Expression of these control genes was significantly increased; however, the expression of *SASH1* mRNA was not modified by the addition of cigarette smoke condensate, suggesting that *SASH1* expression is not affected by the acute exposure to cigarette smoke (Suppl Fig 2).

SASH1 siRNA-silencing affects cell cycle in human aortic endothelial cells

In order to determine the molecular pathways potentially linking SASH1 to vascular pathology, a transcriptomic analysis was performed using SASH1 siRNA-silenced HAECs from ten different individuals. The cells were transfected either with siSASH1 S23572 (si1), siSASH1 S23572 (si2) or a control siRNA (siCtrl). Analysis of the results revealed that 182 genes were differentially expressed ($\text{FDR} < 0.10$) in both si1- and si2-silenced HAECs as compared to siCtrl-transfected cells (Suppl. data Table 1). Among these genes, *CYP1A1*, a major player in the cigarette smoke metabolism¹⁴⁷, was shown to be down-regulated upon SASH1 knockdown; this decrease was confirmed independently by the RT-qPCR experiments (Figure 5A). Pathway analysis using the Ingenuity Pathway Analysis tool (IPA) performed on the 182 differentially expressed genes, showed that pathways involved in regulation of the cell cycle were over-represented (Table 2). IPA analysis

prediction of established transcription factor activity, based on the list of differentially expressed genes upon SASH1 knockdown, showed a potential inhibition of the tumor suppressor TP53 (Suppl Fig 3), and in parallel an activation of FOXM1; the latter is strongly associated with cell proliferation¹⁴⁸. Taken together, those results suggest an implication of SASH1 in the regulation of the cell cycle.

| Ingenuity Canonical Pathways | BH corrected FDR | Ratio | Nb of genes |
|---|-------------------------|--------------|--------------------|
| Mitotic Roles of polo-Like Kinase | $7.1 \cdot 10^{-5}$ | 1.84E-01 | 6 |
| Cell cycle: G2/M DNA Damage Checkpoint Regulation | $2.1 \cdot 10^{-4}$ | 1.59E-01 | 5 |

Table 2. Pathway analysis of the genes affected by SASH1 silencing. Statistically significant pathways are involved in the cell cycle. BH: Benjamini-Hoechberg, Ratio: number of gene differentially expressed upon SASH1 knockdown divided by total number of genes from the pathway. Nb of genes: number of genes differentially expressed upon SASH1 knockdown involved in the pathway.

Previously, overexpression of SASH1 protein has been shown to increase the expression of cyclin D1¹¹, a major protein involved in the cell cycle control, especially the G₁ phase. While cyclin D1 did not show statistically significant change of expression in our microarrays analysis, the quantification of mRNA by RT-qPCR of SASH1 siRNA-silenced HAECs showed that cyclin D3 (CCND3), usually co-expressed with CCND1, was significantly up regulated by 1.75 fold (p<0.005). Interestingly, Western blot analyses showed an increased protein expression of both CCND1 and CCND3 (Figure 5B). As these cyclins are G₁/S phase specific, an increase of their expression suggested that SASH1 siRNA-silencing may increase the proliferation rate of HAECs.

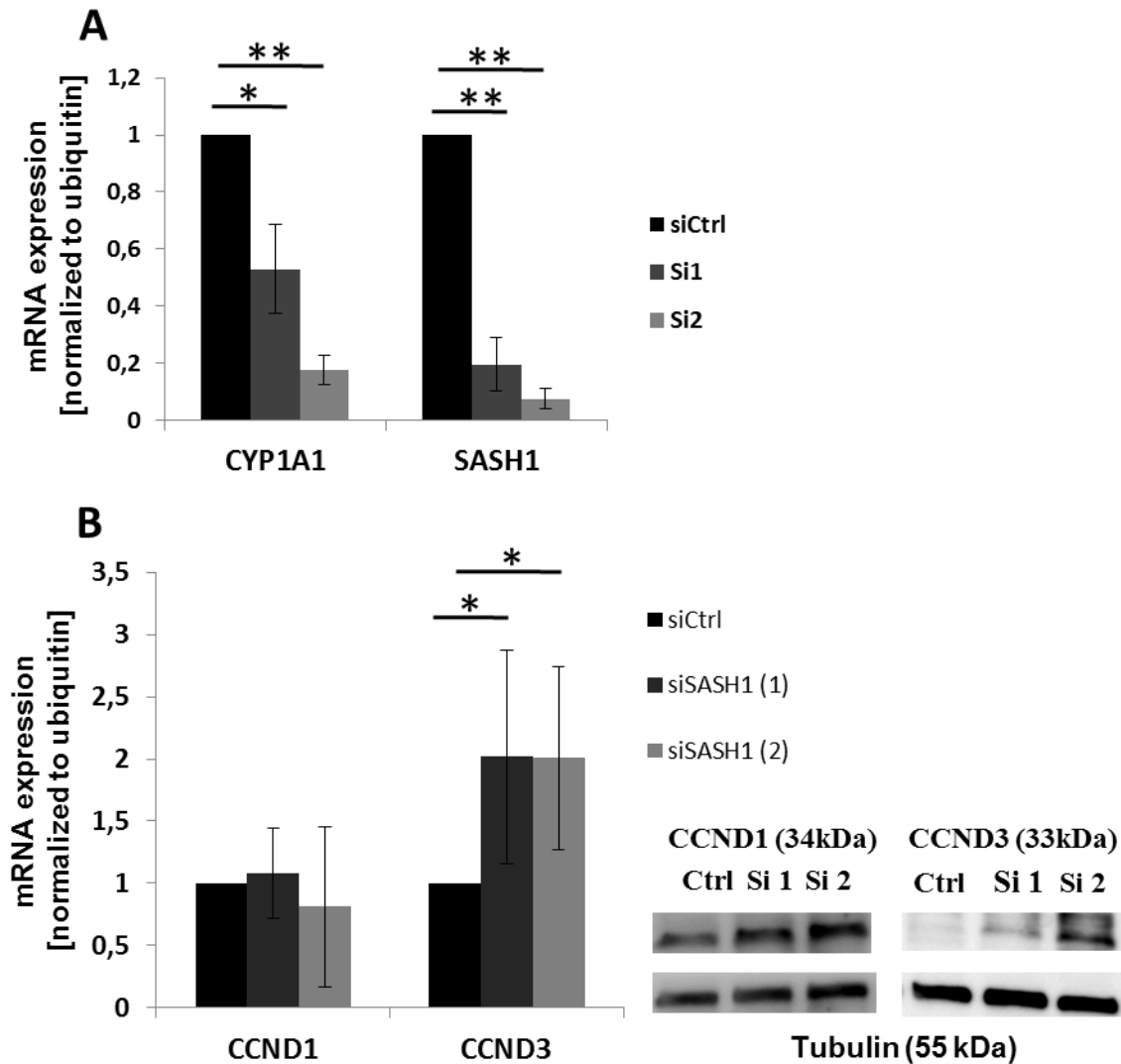


Figure 5. Transcripts affected by SASH1 silencing in HAECs. HAECs (n=6 individuals) were transfected for 48h with either siCtrl (ctrl) or siSASH1 (si1 and si2 used separately); (A) mRNA expression of CYP1A1 and SASH1 (B) mRNA expression and protein expression of CCND3 and CCND1
 * Indicates a $p < 0.05$ ** Indicates a $p < 0.001$

SASH1 siRNA-silencing increases proliferation and migration of human aortic endothelial cells

Our transcriptomic results suggested an implication of SASH1 in the cell cycle regulation. Therefore, we further investigated the effect of SASH1 on proliferation and migration of HAECs by using WST-1 and cell migration assay (wound healing assay, using mitomycin C to block proliferation). Two different siRNAs were used to silence SASH1 in HAECs as described above and the wound area or formazan formation was compared to that of HAECs treated with control siRNA. The closure of the wound and formazan concentration were significantly higher in both SASH1 siRNA-silenced

HAECs than in control HAECs (Figure 6A and 6B), indicating that SASH1 represses migration and proliferation.

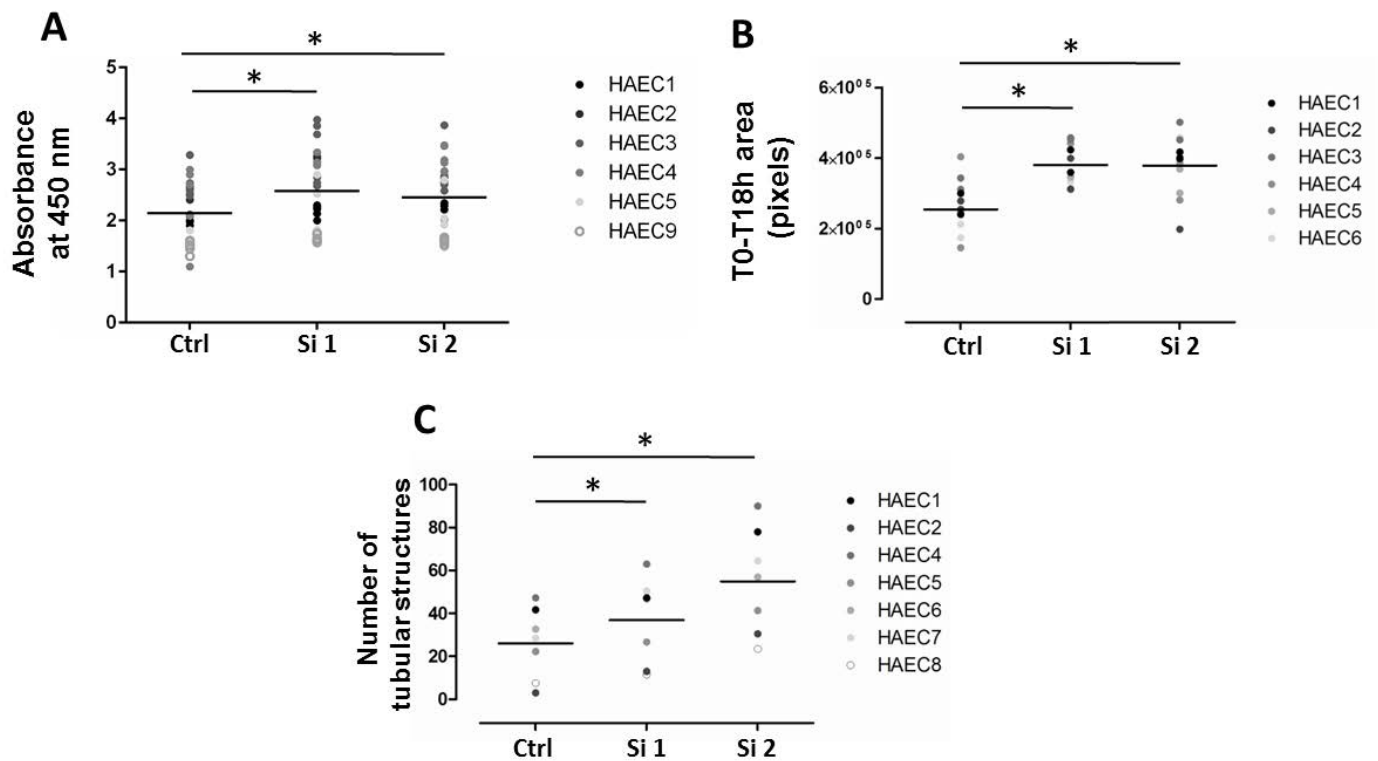


Figure 6. Proliferation, migration and angiogenesis of SASH1 silenced HAECs. HAECs (n=6 to 8 individuals) were transfected for 24h with either control siRNA (ctrl) or siSASH1 (si1 and si2 used separately); (A) the proliferation rate was measured using the WST-1 assay and the absorbance at 450nm was measured after 24h of incubation at 37°C, (B) the wound healing assay was performed on HAECs treated with mitomycin C by comparing the area of the wound at t0 vs t18h of incubation at 37°C; (C) Angiogenesis on Matrigel was performed by comparing the number of tubular structures at t0 and after t18h of incubation at 37°C.

* Indicates a p<0.05

SASH1 siRNA-silencing increases angiogenesis of human aortic endothelial cells

Since both proliferation and migration were increased by SASH1 knockdown, we hypothesized that SASH1 would also affect angiogenesis. To test this hypothesis, the angiogenic capacities of SASH1 silenced HAECs were assessed using a Matrigel assay by comparing the number of tubular structures to control. As expected, an increased number of tubular structures upon SASH1 silencing were formed (Figure 6C), indicating an inhibitory effect of SASH1 on angiogenesis.

Discussion

Recently, *SASH1* expression has been linked to a number of diseases such as pre-eclampsia¹⁴⁹, Alzheimer disease¹⁵⁰, colorectal and breast cancer^{133,135,136}, skin carcinoma¹⁵¹ and Dyschromatosis

universalis hereditaria (DUH)¹⁴¹. While little is known about the role of SASH1 in the vascular physiopathology, an *in vitro* study has reported *SASH1* to be up-regulated following VEGF receptor blockade in human pulmonary microvascular endothelial cells¹⁵². An involvement of SASH1 in such a broad range of diseases could be explained by the fact that *SASH1* expression has an important impact on basic cellular mechanisms, such as proliferation and migration. Nevertheless, these observations are mostly based on correlations, and the molecular mechanism by which SASH1 affect those pathologies is widely unknown.

In the present study we showed for the first time that SASH1 is ubiquitously expressed in human arterial wall, and that *SASH1* mRNA expression was increased in the carotid plaques of smokers, thus extending the results from our previous work showing an increased expression in circulating monocytes from smokers¹³¹. Furthermore, our results indicate that SASH1 acts as an inhibitor of proliferation, migration and angiogenesis in HAECs possibly through a TP53 inhibition mechanism. These results suggest that an up-regulation of SASH1 in smokers could be a cellular defense mechanism, specifically directed against the cigarette smoke-induced migration and proliferation. Accordingly, earlier observations have indicated that SASH1 has tumor-suppressive properties. The fact that smoking increases SASH1 expression in carotids, while SASH1 expression was found to be stable throughout the plaque development, may imply that some of the inflammatory and proliferation pathways activated by smoking through SASH1 expression are distinct from the ones triggered by the lipid accumulation in atherosclerosis and could therefore aggravate the already existing inflammation.

In previous studies on breast and liver cancer, the *SASH1* promoter region was shown to be differentially methylated^{153,154}. In our data, an increased *SASH1* mRNA expression in carotid plaques was observed when comparing smokers to non-smokers. Interestingly, the level of expression in carotid plaques of ex-smokers was similar to that of non-smokers, suggesting that the increase in *SASH1* mRNA expression is transient and may be reversed. One hypothesis to explain this phenomenon could be that the variation of *SASH1* expression in the carotids from smokers might be driven by differential methylation pattern in the *SASH1* promoter region. Cigarette smoke is known to affect methylation¹⁵⁵ in a long-term reversible manner¹⁵⁶. Consistent with this hypothesis, *in vitro* stimulation of HAECs with cigarette smoke condensate to investigate an acute response of *SASH1* to smoking, did not alter *SASH1* expression at the mRNA level, suggesting that *SASH1* expression is only affected by mid-to-long term exposure to cigarette smoke.

Nevertheless, those *in vitro* models for acute cigarette smoke exposure have several limitations, such as the lack of cigarette smoke metabolites or poor mimicking of chronic exposure to cigarette smoke¹⁵⁷. It is likely that testing the impact of acute cigarette smoke exposure on SASH1 requires an *in vivo* model to take into account the chronic exposure and the metabolism of inhaled smoke. Alternatively *in vitro* experiments could also be done using a combination of cigarette smoke extract and microsomes that contain numerous enzymes involved in the metabolism of xenobiotics¹⁵⁸. It is noteworthy that the cigarette smoke condensate used in our study only contained the lipophilic part of the cigarette smoke. Aqueous extract have been used in the literature, but our attempts to stimulate human umbilical vein cells (HUVEC, n=2) with such extracts failed to affect *SASH1* expression (data not shown).

We demonstrated that SASH1 silencing increased cell proliferation, migration and angiogenesis, all processes being implicated in atherosclerosis^{159,160}. Accordingly, SASH1 silencing resulted in an increased CCND1 protein expression, as previously shown in A549 cells¹¹, as well as the increased CCND3 mRNA and protein levels. Both cyclins are part of the cyclin-dependent kinase protein complex required for the G₁/S transition^{161,162} and are known to be co-expressed¹⁶³. However, as SASH1 over-expressing cells have been shown to be arrested in the G₁ phase¹³⁷, it is likely that the increase of CCND1 and CCND3 expression reflects an indirect effect due to SASH1 silencing thus favoring the G₁/S transition rather than a direct effect of SASH1 increasing their expression.

Our transcriptomic results pointed out the implication of SASH1 in the cell cycle regulation, as the two statistically significant pathways (FDR<10), namely “Mitotic Roles of polo-Like Kinase” and “Cell cycle: G2/M DNA Damage Checkpoint Regulation” are both involved in cell replication. Interestingly those results predicted the inhibition of TP53 and activation of FOXM1 transcription factors. Since TP53 is known to negatively regulate the FOXM1 pathway¹⁶⁴, it is possible that the effect of SASH1 on FOXM1 is indirect and mediated by its primary effect on TP53. Little is known about the function of FOXM1 in atherosclerosis, while it could be considered pro-atherogenic because of its mitogenic properties¹⁶⁵, it was also demonstrated to play a critical role in endothelial repair following vascular injury¹⁶⁶ and therefore could also show some atheroprotective function.

CYP1A1, a member of the cytochrome P450 family was found to be down-regulated upon SASH1 silencing in our experiments. CYP1A1 is up regulated in a TP53-dependent fashion by the polycyclic aromatic hydrocarbons¹⁶⁷, present in cigarette smoke, thus the link between the expression of SASH1 and CYP1A1 could be explained by the fact that SASH1 participate to the activation of TP53. The cigarette smoke condensate used in our study to activate HAECs contains high levels of polycyclic

aromatic hydrocarbons and consequently it increased CYP1A1 expression; however SASH1 expression remained unchanged. A possible explanation could be that SASH1 intervenes in the increase of CYP1A1 expression rather at the protein-protein level. However more studies will be required to validate these hypotheses.

As the *SASH1* sequence includes both NLS and nuclear export signal (NES), and as a C-terminus deleted construct of SASH1 accumulates in the nucleus¹³⁴, it is likely that SASH1 also plays a role in the regulation of gene transcription. However there is no evidence or known protein domain in the SASH1 protein for a direct binding to DNA, although SAM domains may bind to RNA¹⁶⁸, suggesting that this action requires SASH1 binding to other proteins which interact with DNA. The nuclear translocation of SASH1 could be regulated either by protein conformation or by chaperone protein(s). Our data from fractionated cell lysates showed that SASH1 is mainly cytoplasmic and to a lesser extent found in the nuclear fraction in quiescent HAECs. When HAECs were stimulated by pro-inflammatory factors or cigarette smoke condensate we could not observe significant translocation of SASH1 to the nucleus, suggesting that SASH1 translocation requires a more specific stimulation, or is not related to inflammation.

In summary, we confirmed the link between expression of SASH1, smoking and atherosclerosis in humans and provide the first evidence of a possible role for SASH1 in endothelial dysfunction leading to atherosclerotic lesions formation.

Acknowledgments

We would like to thank Annette Lesot (Groupe Hospitalier Pitié-Salpêtrière, AP-HP, Paris, France) for her work on the tissue immunostaining, and Dr Diana Lindner (Universitätsklinikum Eppendorf, Hamburg) for her help with the statistical analysis and preparing figures. We also thank the genomic P3S platform for the microarrays, qPCR data acquisition and western blot scanning.

Sources of funding

This work was supported in France: by the Institut National de la Santé et de la Recherche Médicale (INSERM), the Fondation de France (FDF) for the Project R13072DD-Bigot to E. Ninio, the New French Society of Atherosclerosis (NSFA) for the first 12 months of the PhD scholarship of H. Weidmann and by the ECOS Sud-CONICYT cooperation program (C13S01) ; Ewa Ninio is Director of Research in the Centre National de la Recherche Scientifique (CNRS); in Germany: by the German Center for Cardiovascular Research (DZHK e.V.). The present study was also supported by a

joint funding grant from the Federal Ministry of Education and Research, Germany and the Agence Nationale de la Recherche, France (contract BMBF 01KU0908A and ANR 09 GENO 106 01).

Disclosures

None

References

1. Libby P, Aikawa M, Jain MK. Vascular endothelium and atherosclerosis. *Handb Exp Pharmacol.* 2006;(176 Pt 2):285-306.
2. Libby P. Inflammation in atherosclerosis. *Arterioscler Thromb Vasc Biol.* 2012;32(9):2045-2051. doi:10.1161/ATVBAHA.108.179705.
3. Hansson GK, Libby P. The immune response in atherosclerosis: a double-edged sword. *Nat Rev Immunol.* 2006;6(7):508-519. doi:10.1038/nri1882.
4. Auerbach O, Garfinkel L. Atherosclerosis and aneurysm of aorta in relation to smoking habits and age. *Chest.* 1980;78(6):805-809.
5. Redgrave JNE, Lovett JK, Rothwell PM. Histological features of symptomatic carotid plaques in relation to age and smoking: the oxford plaque study. *Stroke J Cereb Circ.* 2010;41(10):2288-2294. doi:10.1161/STROKEAHA.110.587006.
6. McEvoy JW, Blaha MJ, DeFilippis AP, et al. Cigarette smoking and cardiovascular events: role of inflammation and subclinical atherosclerosis from the multiethnic study of atherosclerosis. *Arterioscler Thromb Vasc Biol.* 2015;35(3):700-709. doi:10.1161/ATVBAHA.114.304562.
7. Talhout R, Schulz T, Florek E, van Benthem J, Wester P, Opperhuizen A. Hazardous compounds in tobacco smoke. *Int J Environ Res Public Health.* 2011;8(2):613-628. doi:10.3390/ijerph8020613.
8. Varela-Carver A, Parker H, Kleinert C, Rimoldi O. Adverse effects of cigarette smoke and induction of oxidative stress in cardiomyocytes and vascular endothelium. *Curr Pharm Des.* 2010;16(23):2551-2558.
9. Messner B, Bernhard D. Smoking and cardiovascular disease: mechanisms of endothelial dysfunction and early atherogenesis. *Arterioscler Thromb Vasc Biol.* 2014;34(3):509-515. doi:10.1161/ATVBAHA.113.300156.

10. Zeller T, Wild P, Szymczak S, et al. Genetics and beyond--the transcriptome of human monocytes and disease susceptibility. *PLoS One*. 2010;5(5):e10693. doi:10.1371/journal.pone.0010693.
11. Verdugo RA, Zeller T, Rotival M, et al. Graphical modeling of gene expression in monocytes suggests molecular mechanisms explaining increased atherosclerosis in smokers. *PLoS One*. 2013;8(1):e50888. doi:10.1371/journal.pone.0050888.
12. Zeller C, Hinzmann B, Seitz S, et al. SASH1: a candidate tumor suppressor gene on chromosome 6q24.3 is downregulated in breast cancer. *Oncogene*. 2003;22(19):2972-2983. doi:10.1038/sj.onc.1206474.
13. Martini M, Gnann A, Scheikl D, Holzmann B, Janssen K-P. The candidate tumor suppressor SASH1 interacts with the actin cytoskeleton and stimulates cell-matrix adhesion. *Int J Biochem Cell Biol*. 2011;43(11):1630-1640. doi:10.1016/j.biocel.2011.07.012.
14. Rimkus C, Martini M, Friederichs J, et al. Prognostic significance of downregulated expression of the candidate tumour suppressor gene SASH1 in colon cancer. *Br J Cancer*. 2006;95(10):1419-1423. doi:10.1038/sj.bjc.6603452.
15. Nitsche U, Rosenberg R, Balmert A, et al. Integrative marker analysis allows risk assessment for metastasis in stage II colon cancer. *Ann Surg*. 2012;256(5):763-771; discussion 771. doi:10.1097/SLA.0b013e318272de87.
16. Chen E, Chen Y, Dong L, Zhang J. Effects of SASH1 on lung cancer cell proliferation, apoptosis, and invasion in vitro. *Tumour Biol J Int Soc Oncodevelopmental Biol Med*. 2012;33(5):1393-1401. doi:10.1007/s13277-012-0387-2.
17. Yang L, Liu M, Gu Z, Chen J, Yan Y, Li J. Overexpression of SASH1 related to the decreased invasion ability of human glioma U251 cells. *Tumour Biol J Int Soc Oncodevelopmental Biol Med*. 2012;33(6):2255-2263. doi:10.1007/s13277-012-0487-z.
18. Lin S, Zhang J, Xu J, et al. Effects of SASH1 on melanoma cell proliferation and apoptosis in vitro. *Mol Med Rep*. 2012;6(6):1243-1248. doi:10.3892/mmr.2012.1099.
19. Dauphinee SM, Clayton A, Hussainkhel A, et al. SASH1 is a scaffold molecule in endothelial TLR4 signaling. *J Immunol Baltim Md 1950*. 2013;191(2):892-901. doi:10.4049/jimmunol.1200583.

20. Zhou D, Wei Z, Deng S, et al. SASH1 regulates melanocyte transepithelial migration through a novel Gas-SASH1-IQGAP1-E-Cadherin dependent pathway. *Cell Signal*. 2013;25(6):1526-1538. doi:10.1016/j.cellsig.2012.12.025.
21. Brochériou I, Maouche S, Durand H, et al. Antagonistic regulation of macrophage phenotype by M-CSF and GM-CSF: implication in atherosclerosis. *Atherosclerosis*. 2011;214(2):316-324. doi:10.1016/j.atherosclerosis.2010.11.023.
22. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*. 2001;29(9):e45.
23. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. *Bioinforma Oxf Engl*. 2008;24(13):1547-1548. doi:10.1093/bioinformatics/btn224.
24. Lin SM, Du P, Huber W, Kibbe WA. Model-based variance-stabilizing transformation for Illumina microarray data. *Nucleic Acids Res*. 2008;36(2):e11. doi:10.1093/nar/gkm1075.
25. Nordskog BK, Blixt AD, Morgan WT, Fields WR, Hellmann GM. Matrix-degrading and pro-inflammatory changes in human vascular endothelial cells exposed to cigarette smoke condensate. *Cardiovasc Toxicol*. 2003;3(2):101-117.
26. Slattery ML, Samowitz W, Ma K, et al. CYP1A1, cigarette smoking, and colon and rectal cancer. *Am J Epidemiol*. 2004;160(9):842-852. doi:10.1093/aje/kwh298.
27. Wierstra I, Alves J. FOXM1, a typical proliferation-associated transcription factor. *Biol Chem*. 2007;388(12):1257-1274. doi:10.1515/BC.2007.159.
28. Sitras V, Paulssen RH, Grønåas H, et al. Differential placental gene expression in severe preeclampsia. *Placenta*. 2009;30(5):424-433. doi:10.1016/j.placenta.2009.01.012.
29. Bronner IF, Bochdanovits Z, Rizzu P, et al. Comprehensive mRNA expression profiling distinguishes tauopathies and identifies shared molecular pathways. *PloS One*. 2009;4(8):e6826. doi:10.1371/journal.pone.0006826.
30. Courcet J-B, Elalaoui SC, Duplomb L, et al. Autosomal-recessive SASH1 variants associated with a new genodermatosis with pigmentation defects, palmoplantar keratoderma and skin carcinoma. *Eur J Hum Genet EJHG*. 2014. doi:10.1038/ejhg.2014.213.

31. Sakao S, Taraseviciene-Stewart L, Cool CD, et al. VEGF-R blockade causes endothelial cell apoptosis, expansion of surviving CD34+ precursor cells and transdifferentiation to smooth muscle-like and neuronal-like cells. *FASEB J Off Publ Fed Am Soc Exp Biol.* 2007;21(13):3640-3652. doi:10.1096/fj.07-8432com.
32. Sheyu L, Hui L, Junyu Z, et al. Promoter methylation assay of SASH1 gene in breast cancer. *J BUON Off J Balk Union Oncol.* 2013;18(4):891-898.
33. Peng L, Wei H, Liren L. Promoter methylation assay of SASH1 gene in hepatocellular carcinoma. *J BUON Off J Balk Union Oncol.* 2014;19(4):1041-1047.
34. Lee KWK, Pausova Z. Cigarette smoking and DNA methylation. *Front Genet.* 2013;4:132. doi:10.3389/fgene.2013.00132.
35. Tsaprouni LG, Yang T-P, Bell J, et al. Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. *Epigenetics Off J DNA Methylation Soc.* 2014;9(10):1382-1396. doi:10.4161/15592294.2014.969637.
36. Wan J, Johnson M, Schilz J, Djordjevic MV, Rice JR, Shields PG. Evaluation of In Vitro Assays For Assessing the Toxicity of Cigarette Smoke and Smokeless Tobacco. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol.* 2009;18(12):3263-3304. doi:10.1158/1055-9965.EPI-09-0965.
37. Brandon EFA, Raap CD, Meijerman I, Beijnen JH, Schellens JHM. An update on in vitro test methods in human hepatic drug biotransformation research: pros and cons. *Toxicol Appl Pharmacol.* 2003;189(3):233-246. doi:10.1016/S0041-008X(03)00128-5.
38. Fuster JJ, Fernández P, González-Navarro H, Silvestre C, Nabah YNA, Andrés V. Control of cell proliferation in atherosclerosis: insights from animal models and human studies. *Cardiovasc Res.* 2010;86(2):254-264. doi:10.1093/cvr/cvp363.
39. Moreno PR, Purushothaman K-R, Sirol M, Levy AP, Fuster V. Neovascularization in Human Atherosclerosis. *Circulation.* 2006;113(18):2245-2252. doi:10.1161/CIRCULATIONAHA.105.578955.
40. Zimmet JM, Ladd D, Jackson CW, Stenberg PE, Ravid K. A role for cyclin D3 in the endomitotic cell cycle. *Mol Cell Biol.* 1997;17(12):7248-7259.

41. Weng H-Y, Huang H-L, Zhao P-P, Zhou H, Qu L-H. Translational repression of cyclin D3 by a stable G-quadruplex in its 5' UTR. *RNA Biol.* 2012;9(8):1099-1109. doi:10.4161/rna.21210.
42. Radulovich N, Pham N-A, Strumpf D, et al. Differential roles of cyclin D1 and D3 in pancreatic ductal adenocarcinoma. *Mol Cancer.* 2010;9:24. doi:10.1186/1476-4598-9-24.
43. Barsotti AM, Prives C. Pro-proliferative FoxM1 is a target of p53-mediated repression. *Oncogene.* 2009;28(48):4295-4305. doi:10.1038/onc.2009.282.
44. Wen N, Wang Y, Wen L, et al. Overexpression of FOXM1 predicts poor prognosis and promotes cancer cell proliferation, migration and invasion in epithelial ovarian cancer. *J Transl Med.* 2014;12:134. doi:10.1186/1479-5876-12-134.
45. Zhao Y-Y, Gao X-P, Zhao YD, et al. Endothelial cell-restricted disruption of FoxM1 impairs endothelial repair following LPS-induced vascular injury. *J Clin Invest.* 2006;116(9):2333-2343. doi:10.1172/JCI27154.
46. Wohak LE, Krais AM, Kucab JE, et al. Carcinogenic polycyclic aromatic hydrocarbons induce CYP1A1 in human cells via a p53-dependent mechanism. *Arch Toxicol.* 2014. doi:10.1007/s00204-014-1409-1.
47. Aviv T, Lin Z, Lau S, Rendl LM, Sicheri F, Smibert CA. The RNA-binding SAM domain of Smaug defines a new family of post-transcriptional regulators. *Nat Struct Biol.* 2003;10(8):614-621. doi:10.1038/nsb956.

Supplementary material

SASH1, a new potential link between smoking and atherosclerosis.

^{1,2,3}Henri Weidmann, ¹Zahia Touat-Hamici, ¹Herve Durand, ^{2,3}Christian Mueller, ⁴Frédéric Charlotte, ⁵Klaus-Peter Janssen, ⁶Ricardo Verdugo, ¹Francois Cambien, ^{2,3}Stefan Blankenberg, ¹Laurence Tiret, ^{2,3}Tanja Zeller* and ¹Ewa Ninio*§

¹Sorbonne Universités, UPMC, UMR_S 1166-ICAN, Genomics and Pathophysiology of Cardiovascular Diseases, Institute of Cardiometabolism and Nutrition, ICAN, Pitié-Salpêtrière Hospital, F-75013, Paris, France.

²University Heart Center Hamburg, Clinic for general and interventional Cardiology, Hamburg.

³German Center for Cardiovascular Research (DZHK e.V.) Partner Site Hamburg, Lübeck, Kiel, Hamburg, Germany.

⁴Department of Pathology, Hôpital de la Pitié-Salpêtrière, AP-HP, Paris, France

⁵Department of Surgery, Klinikum rechts der Isar, Technische Universität München, IsmaningerStrasse 22, Munich, Germany.

⁶Programa de Genetica Humana ICBM, Facultad de Medicina, Universidad de Chile, Chile.

* Both authors contributed equally to this work

§ Author for correspondence: Ewa Ninio, Sorbonne Universités, UPMC, UMR_S 1166-ICAN, 91, Boulevard de l'Hôpital, 75634 Paris cedex 13, France; email ewa.ninio@upmc.fr; phone: +33140779768.

| ID | geneSymbol | geneName | FC si1 | qval_si1 | FC si2 | qval_si2 |
|--------------|------------|--|-----------|----------|-----------|----------|
| ILMN_2185984 | SASH1 | SAM and SH3 domain containing 1 | 0,35 | 1,3E-03 | 0,15 | 4,8E-08 |
| ILMN_1800912 | ZNF106 | zinc finger protein 106 | 0,77 | 5,7E-02 | 0,43 | 4,8E-08 |
| ILMN_1697597 | EFCAB14 | EF-hand calcium binding domain 14 | 0,58 | 5,9E-03 | 0,25 | 6,1E-07 |
| ILMN_1803279 | TMED5 | transmembrane emp24 protein transport domain containing 5 | 0,61 | 3,8E-02 | 0,31 | 1,6E-06 |
| ILMN_3236945 | PTPMT1 | protein tyrosine phosphatase, mitochondrial 1 | 2,73 | 4,7E-05 | 3,19 | 1,8E-06 |
| ILMN_1807969 | SNCAIP | synuclein, alpha interacting protein | 0,75 | 3,8E-02 | 0,49 | 5,1E-06 |
| ILMN_1701933 | SNCA | synuclein, alpha (non A4 component of amyloid precursor) | 2,45 | 8,8E-04 | 2,17 | 9,4E-06 |
| ILMN_1777066 | NIF3L1 | NIF3 NGG1 interacting factor 3-like 1 (<i>S. cerevisiae</i>) | 1,25 | 3,9E-02 | 1,65 | 1,2E-05 |
| ILMN_1808789 | MYO5C | myosin VC | 0,75 | 2,8E-02 | 0,44 | 1,9E-05 |
| ILMN_2243308 | ACVR1B | activin A receptor, type IB | 1,42 | 3,4E-02 | 1,65 | 2,0E-05 |
| ILMN_1679796 | TOMM20 | translocase of outer mitochondrial membrane 20 homolog (yeast) | 2,14 | 3,2E-04 | 2,45 | 2,4E-05 |
| ILMN_2093027 | MYO1B | myosin IB | 0,64 | 3,0E-02 | 0,51 | 2,8E-05 |
| ILMN_1716730 | BOD1 | biorientation of chromosomes in cell division 1 | 1,73 | 3,0E-02 | 1,96 | 2,9E-05 |
| ILMN_1812902 | CYYR1 | cysteine/tyrosine-rich 1 | 0,73 | 7,1E-02 | 0,23 | 4,1E-05 |
| ILMN_1724789 | CD59 | CD59 molecule, complement regulatory protein | 2,27 | 1,9E-03 | 2,04 | 4,3E-05 |
| ILMN_1809590 | GINS2 | GINS complex subunit 2 (Psf2 homolog) | 1,65 | 2,1E-02 | 2,43 | 4,4E-05 |
| ILMN_1709750 | SUSD1 | sushi domain containing 1 | 1,76 | 7,8E-03 | 1,62 | 4,9E-05 |
| ILMN_3224204 | PSMG4 | proteasome (prosome, macropain) assembly chaperone 4 | 1,45 | 5,7E-02 | 1,73 | 5,3E-05 |
| ILMN_1654653 | KLC1 | kinesin light chain 1 | 1,56 | 2,9E-03 | 1,60 | 9,6E-05 |
| ILMN_1786065 | UHRF1 | ubiquitin-like with PHD and ring finger domains 1 | 1,39 | 4,1E-02 | 1,98 | 1,0E-04 |
| ILMN_1754272 | GINS3 | GINS complex subunit 3 (Psf3 homolog) | 1,57 | 6,2E-03 | 1,63 | 1,2E-04 |
| ILMN_1691506 | NGRN | neugrin, neurite outgrowth associated | 1,62 | 1,4E-02 | 1,59 | 1,5E-04 |
| ILMN_1688103 | CTNNBIP1 | catenin, beta interacting protein 1 | 0,75 | 1,0E-01 | 0,56 | 1,5E-04 |
| ILMN_1732923 | SIPA1L2 | signal-induced proliferation-associated 1 like 2 | 0,66 | 1,4E-02 | 0,68 | 1,7E-04 |
| ILMN_2077550 | RACGAP1 | Rac GTPase activating protein 1 | 1,30 | 6,6E-02 | 1,49 | 1,7E-04 |
| ILMN_1761808 | MCFD2 | multiple coagulation factor deficiency 2 | 1,75 | 5,8E-02 | 1,78 | 2,0E-04 |
| ILMN_1727043 | COLGALT1 | collagen beta(1-O)galactosyltransferase 1 | 0,68 | 5,7E-02 | 0,45 | 2,2E-04 |
| ILMN_1772821 | KIAA1671 | KIAA1671 | 0,72 | 3,3E-02 | 0,49 | 2,4E-04 |
| ILMN_1757845 | SPIRE1 | spire-type actin nucleation factor 1 | 1,58 | 2,7E-02 | 1,59 | 2,4E-04 |
| ILMN_1733535 | ZNF366 | zinc finger protein 366 | 0,80 | 3,4E-02 | 0,75 | 2,5E-04 |
| ILMN_1670238 | CDC45 | cell division cycle 45 | 1,59 | 3,4E-02 | 1,73 | 2,5E-04 |

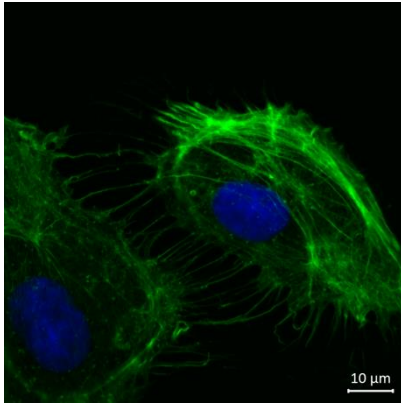
| | | | | | | |
|--------------|----------|--|------|---------|------|---------|
| ILMN_1787345 | FKBP11 | FK506 binding protein 11, 19 kDa | 1,40 | 3,4E-02 | 1,58 | 2,6E-04 |
| ILMN_1694140 | AFAP1L1 | actin filament associated protein 1-like 1 | 0,59 | 2,5E-02 | 0,53 | 2,9E-04 |
| ILMN_1738554 | ADH5 | alcohol dehydrogenase 5 (class III), chi polypeptide | 1,45 | 1,3E-02 | 1,76 | 3,1E-04 |
| ILMN_1743714 | CARD10 | caspase recruitment domain family, member 10 | 0,70 | 4,5E-02 | 0,60 | 3,8E-04 |
| ILMN_2209748 | DERL1 | derlin 1 | 1,32 | 3,9E-02 | 1,42 | 3,8E-04 |
| ILMN_1802669 | PPP3CB | protein phosphatase 3, catalytic subunit, beta isozyme | 1,61 | 1,1E-02 | 1,53 | 4,0E-04 |
| ILMN_1711470 | UBE2T | ubiquitin-conjugating enzyme E2T (putative) | 1,50 | 3,4E-02 | 1,64 | 4,4E-04 |
| ILMN_1720889 | MSMO1 | methylsterol monooxygenase 1 | 1,44 | 3,1E-02 | 1,45 | 4,4E-04 |
| ILMN_1807291 | CYP1A1 | cytochrome P450, family 1, subfamily A, polypeptide 1 | 0,69 | 2,7E-02 | 0,29 | 4,7E-04 |
| ILMN_1738955 | ANAPC16 | anaphase promoting complex subunit 16 | 0,80 | 3,7E-02 | 0,70 | 4,8E-04 |
| ILMN_1708672 | ACAT2 | acetyl-CoA acetyltransferase 2 | 1,55 | 2,8E-02 | 1,89 | 5,0E-04 |
| ILMN_1785284 | ALDH6A1 | aldehyde dehydrogenase 6 family, member A1 | 0,61 | 1,7E-02 | 0,58 | 5,2E-04 |
| ILMN_2124471 | SLC36A1 | solute carrier family 36 (proton/amino acid symporter), member 1 | 1,22 | 8,2E-02 | 1,38 | 6,7E-04 |
| ILMN_1749213 | SDF2L1 | stromal cell-derived factor 2-like 1 | 1,34 | 8,3E-02 | 1,64 | 7,1E-04 |
| ILMN_1668540 | ZNHIT6 | zinc finger, HIT-type containing 6 | 1,54 | 1,3E-02 | 1,39 | 7,2E-04 |
| ILMN_1763907 | CENPW | centromere protein W | 1,46 | 7,2E-02 | 1,79 | 7,2E-04 |
| ILMN_1709683 | RASSF2 | Ras association (RalGDS/AF-6) domain family member 2 | 0,55 | 1,7E-02 | 0,45 | 7,3E-04 |
| ILMN_3248094 | ANO2 | anoctamin 2, calcium activated chloride channel | 0,56 | 3,8E-02 | 0,51 | 7,4E-04 |
| ILMN_1687978 | PHLDA1 | pleckstrin homology-like domain, family A, member 1 | 0,51 | 1,3E-02 | 0,61 | 7,7E-04 |
| ILMN_1777397 | MSX1 | msh homeobox 1 | 0,72 | 7,3E-02 | 0,62 | 7,8E-04 |
| ILMN_1673129 | PCDH12 | protocadherin 12 | 0,50 | 2,9E-02 | 0,55 | 8,2E-04 |
| ILMN_1727194 | CALU | Calumenin | 2,41 | 3,9E-02 | 1,60 | 8,3E-04 |
| ILMN_1737205 | MCM4 | minichromosome maintenance complex component 4 | 1,43 | 2,2E-02 | 1,65 | 8,6E-04 |
| ILMN_1696622 | SLC38A6 | solute carrier family 38, member 6 | 1,33 | 5,6E-02 | 1,34 | 9,1E-04 |
| ILMN_2362549 | ZWINT | ZW10 interacting kinetochore protein | 1,34 | 6,1E-02 | 1,49 | 9,2E-04 |
| ILMN_1790354 | RABGGTB | Rab geranylgeranyltransferase, beta subunit | 2,00 | 6,0E-02 | 2,02 | 9,9E-04 |
| ILMN_1659364 | RFC5 | replication factor C (activator 1) 5, 36.5kDa | 1,31 | 4,0E-02 | 1,37 | 1,1E-03 |
| ILMN_1677719 | CHST1 | carbohydrate (keratan sulfate Gal-6) sulfotransferase 1 | 0,63 | 1,4E-02 | 0,51 | 1,1E-03 |
| ILMN_1715175 | MET | MET proto-oncogene, receptor tyrosine kinase | 1,81 | 5,9E-03 | 1,64 | 1,2E-03 |
| ILMN_2334296 | IL18BP | interleukin 18 binding protein | 0,70 | 9,1E-02 | 0,61 | 1,2E-03 |
| ILMN_1793770 | DNAJB6 | DnaJ (Hsp40) homolog, subfamily B, member 6 | 0,79 | 3,7E-02 | 0,74 | 1,2E-03 |
| ILMN_2401978 | STAT3 | signal transducer and activator of transcription 3 | 0,75 | 8,6E-02 | 0,63 | 1,2E-03 |
| ILMN_2139970 | ALDH1A3 | aldehyde dehydrogenase 1 family, member A3 | 0,61 | 5,8E-02 | 0,61 | 1,2E-03 |
| ILMN_1745954 | CORO1C | coronin, actin binding protein, 1C | 0,72 | 8,2E-02 | 0,69 | 1,3E-03 |
| ILMN_1712803 | CCNB1 | cyclin B1 | 1,41 | 2,8E-02 | 1,48 | 1,3E-03 |
| ILMN_1720124 | RCC2 | regulator of chromosome condensation 2 | 1,77 | 1,7E-02 | 1,60 | 1,3E-03 |
| ILMN_2399893 | RPS24 | ribosomal protein S24 | 1,56 | 1,5E-02 | 1,74 | 1,5E-03 |
| ILMN_1703955 | FBXO32 | F-box protein 32 | 1,53 | 6,1E-02 | 0,70 | 1,5E-03 |
| ILMN_1734229 | SPPL2A | signal peptide peptidase like 2A | 1,34 | 6,7E-02 | 1,50 | 1,5E-03 |
| ILMN_1716382 | C12orf75 | chromosome 12 open reading frame 75 | 1,29 | 7,6E-02 | 1,41 | 1,5E-03 |
| ILMN_1676088 | MSRB3 | methionine sulfoxide reductase B3 | 1,92 | 4,7E-05 | 1,72 | 1,6E-03 |
| ILMN_1753745 | HDCC2 | HD domain containing 2 | 1,46 | 5,9E-03 | 1,67 | 1,6E-03 |

| | | | | | | |
|--------------|----------|--|------|---------|------|---------|
| ILMN_2181432 | PAPL | iron/zinc purple acid phosphatase-like protein | 1,34 | 4,0E-02 | 1,76 | 1,7E-03 |
| ILMN_1668863 | LYPD1 | LY6/PLAUR domain containing 1 | 1,57 | 7,8E-03 | 1,61 | 1,7E-03 |
| ILMN_1713751 | ADAM19 | ADAM metallopeptidase domain 19 | 0,54 | 2,5E-02 | 0,53 | 1,7E-03 |
| ILMN_1747344 | IL3RA | interleukin 3 receptor, alpha (low affinity) | 0,78 | 9,1E-02 | 0,72 | 2,2E-03 |
| ILMN_1800958 | TMEM237 | transmembrane protein 237 | 1,25 | 1,0E-01 | 1,52 | 2,3E-03 |
| ILMN_1719032 | LSM3 | LSM3 homolog, U6 small nuclear RNA associated (<i>S. cerevisiae</i>) | 1,46 | 3,9E-02 | 1,83 | 2,4E-03 |
| ILMN_2051373 | NEK2 | NIMA-related kinase 2 | 1,32 | 5,7E-02 | 1,42 | 2,6E-03 |
| ILMN_1657796 | STMN1 | stathmin 1 | 1,79 | 3,9E-02 | 1,68 | 2,7E-03 |
| ILMN_1714397 | CRYL1 | crystallin, lambda 1 | 0,81 | 3,3E-02 | 0,71 | 2,7E-03 |
| ILMN_2120695 | TSPAN7 | tetraspanin 7 | 0,58 | 3,9E-02 | 0,46 | 2,7E-03 |
| ILMN_1684227 | GPR146 | G protein-coupled receptor 146 | 0,77 | 5,2E-02 | 0,66 | 2,7E-03 |
| ILMN_2397880 | CSTF3 | cleavage stimulation factor, 3' pre-RNA, subunit 3, 77kDa | 1,35 | 1,9E-02 | 1,48 | 2,8E-03 |
| ILMN_1890614 | FAM212B | family with sequence similarity 212, member B | 0,78 | 6,9E-02 | 0,75 | 3,0E-03 |
| ILMN_1747303 | DDX39A | DEAD (Asp-Glu-Ala-Asp) box polypeptide 39A | 1,29 | 7,3E-02 | 1,77 | 3,1E-03 |
| ILMN_3307700 | SPCS3 | signal peptidase complex subunit 3 homolog (<i>S. cerevisiae</i>) | 1,26 | 5,2E-02 | 1,43 | 3,2E-03 |
| ILMN_2202948 | BUB1 | BUB1 mitotic checkpoint serine/threonine kinase | 1,48 | 6,9E-02 | 1,60 | 3,3E-03 |
| ILMN_1691339 | CLEC1A | C-type lectin domain family 1, member A | 0,74 | 4,8E-02 | 0,70 | 3,9E-03 |
| ILMN_1703906 | HJURP | Holliday junction recognition protein | 1,38 | 2,1E-02 | 1,38 | 4,0E-03 |
| ILMN_1721868 | KPNA2 | karyopherin alpha 2 (RAG cohort 1, importin alpha 1) | 1,38 | 4,7E-02 | 1,62 | 4,2E-03 |
| ILMN_1810228 | TTF2 | transcription termination factor, RNA polymerase II | 1,29 | 3,7E-02 | 1,25 | 4,3E-03 |
| ILMN_1786125 | CCNA2 | cyclin A2 | 1,47 | 6,3E-02 | 1,69 | 4,3E-03 |
| ILMN_1765701 | TUBA1B | tubulin, alpha 1b | 1,68 | 9,8E-02 | 1,92 | 4,4E-03 |
| ILMN_1726210 | GPIHBP1 | glycosylphosphatidylinositol anchored high density lipoprotein binding protein 1 | 0,40 | 2,1E-02 | 0,36 | 4,5E-03 |
| ILMN_1798030 | XPR1 | xenotropic and polytropic retrovirus receptor 1 | 0,79 | 4,0E-02 | 0,75 | 4,5E-03 |
| ILMN_2100689 | MAP2K4 | mitogen-activated protein kinase kinase 4 | 1,32 | 4,2E-02 | 1,35 | 4,7E-03 |
| ILMN_1656501 | DUSP5 | dual specificity phosphatase 5 | 0,76 | 5,7E-02 | 1,40 | 4,9E-03 |
| ILMN_1691156 | MT1A | metallothionein 1A | 1,54 | 6,9E-02 | 1,78 | 4,9E-03 |
| ILMN_1789123 | PLK4 | polo-like kinase 4 | 1,34 | 5,6E-02 | 1,42 | 5,0E-03 |
| ILMN_1674243 | TFRC | transferrin receptor | 1,60 | 4,0E-02 | 1,96 | 5,0E-03 |
| ILMN_1670353 | RAD51AP1 | RAD51 associated protein 1 | 1,34 | 3,9E-02 | 1,44 | 5,0E-03 |
| ILMN_1784227 | MCRS1 | microspherule protein 1 | 1,40 | 3,7E-02 | 1,51 | 5,1E-03 |
| ILMN_1732799 | CD34 | CD34 molecule | 0,67 | 2,2E-02 | 0,56 | 5,5E-03 |
| ILMN_1737728 | CDCA3 | cell division cycle associated 3 | 1,29 | 6,9E-02 | 1,31 | 5,7E-03 |
| ILMN_1808110 | TCEB3 | transcription elongation factor B (SIII), polypeptide 3 (110kDa, elongin A) | 1,30 | 1,8E-02 | 1,21 | 5,8E-03 |
| ILMN_1702301 | DOCK10 | dedicator of cytokinesis 10 | 1,34 | 8,8E-02 | 1,28 | 5,9E-03 |
| ILMN_1699980 | TSPAN18 | tetraspanin 18 | 0,61 | 1,2E-02 | 0,51 | 6,1E-03 |
| ILMN_1653828 | CHFR | checkpoint with forkhead and ring finger domains, E3 ubiquitin protein ligase | 1,53 | 6,7E-02 | 1,37 | 6,5E-03 |
| ILMN_1679133 | SERPINB1 | serpin peptidase inhibitor, clade B (ovalbumin), member 1 | 0,48 | 1,2E-02 | 0,39 | 6,7E-03 |
| ILMN_1802257 | PCTP | phosphatidylcholine transfer protein | 0,81 | 2,9E-02 | 1,28 | 6,8E-03 |

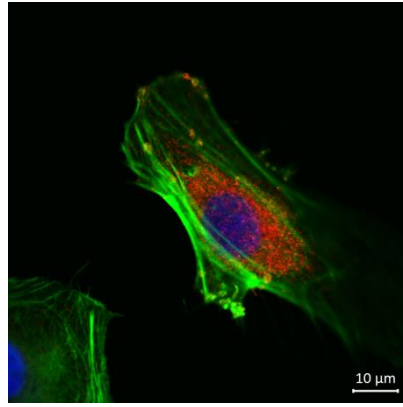
| | | | | | | |
|--------------|----------|---|------|---------|------|---------|
| ILMN_1771051 | RPL29 | ribosomal protein L29 | 1,40 | 7,5E-02 | 1,66 | 7,2E-03 |
| ILMN_1677765 | LRP8 | low density lipoprotein receptor-related protein 8, apolipoprotein e receptor | 1,44 | 3,0E-02 | 1,20 | 7,3E-03 |
| ILMN_1747911 | CDK1 | cyclin-dependent kinase 1 | 1,73 | 1,4E-02 | 1,63 | 7,7E-03 |
| ILMN_1756220 | DDX18 | DEAD (Asp-Glu-Ala-Asp) box polypeptide 18 | 1,46 | 5,4E-02 | 1,26 | 9,0E-03 |
| ILMN_1784005 | RAB13 | RAB13, member RAS oncogene family | 0,72 | 4,4E-02 | 0,68 | 9,0E-03 |
| ILMN_1754660 | ZCCHC24 | zinc finger, CCHC domain containing 24 | 0,81 | 9,9E-02 | 0,71 | 9,2E-03 |
| ILMN_1768271 | SMAP1 | small ArfGAP 1 | 1,28 | 4,3E-02 | 1,23 | 9,4E-03 |
| ILMN_1728934 | PRC1 | protein regulator of cytokinesis 1 | 1,66 | 2,2E-02 | 1,69 | 9,5E-03 |
| ILMN_1663390 | CDC20 | cell division cycle 20 | 2,11 | 7,8E-03 | 1,75 | 9,5E-03 |
| ILMN_1788955 | PDLIM1 | PDZ and LIM domain 1 | 0,57 | 5,3E-02 | 0,60 | 9,7E-03 |
| ILMN_1776490 | C17orf53 | chromosome 17 open reading frame 53 | 1,23 | 8,6E-02 | 1,38 | 1,0E-02 |
| ILMN_1743635 | ALG14 | ALG14, UDP-N-acetylglucosaminyltransferase subunit | 1,27 | 8,8E-02 | 1,24 | 1,0E-02 |
| ILMN_1756326 | CKS2 | CDC28 protein kinase regulatory subunit 2 | 1,58 | 9,3E-03 | 1,70 | 1,0E-02 |
| ILMN_1654563 | EFNB1 | ephrin-B1 | 0,72 | 3,4E-02 | 0,65 | 1,0E-02 |
| ILMN_1719344 | NRBF2 | nuclear receptor binding factor 2 | 1,55 | 1,2E-02 | 1,26 | 1,1E-02 |
| ILMN_1771966 | BCCIP | BRCA2 and CDKN1A interacting protein | 1,69 | 2,1E-02 | 1,45 | 1,1E-02 |
| ILMN_1796216 | VASH1 | vasohibin 1 | 0,68 | 3,7E-02 | 0,55 | 1,1E-02 |
| ILMN_1681591 | PTPN1 | protein tyrosine phosphatase, non-receptor type 1 | 1,84 | 1,1E-02 | 1,33 | 1,2E-02 |
| ILMN_1815261 | PDIA4 | protein disulfide isomerase family A, member 4 | 1,49 | 2,9E-02 | 1,28 | 1,3E-02 |
| ILMN_1728197 | CLDN5 | claudin 5 | 0,52 | 2,9E-02 | 0,53 | 1,3E-02 |
| ILMN_1668125 | MYRIP | myosin VIIA and Rab interacting protein | 0,78 | 3,8E-02 | 0,56 | 1,4E-02 |
| ILMN_1768480 | VGLL4 | vestigial-like family member 4 | 1,45 | 2,7E-02 | 1,27 | 1,4E-02 |
| ILMN_1771026 | GARS | glycyl-tRNA synthetase | 1,45 | 2,9E-02 | 1,48 | 1,4E-02 |
| ILMN_2370624 | COL13A1 | collagen, type XIII, alpha 1 | 1,67 | 3,5E-02 | 1,44 | 1,5E-02 |
| ILMN_1734486 | TSEN15 | TSEN15 tRNA splicing endonuclease subunit | 1,22 | 8,4E-02 | 1,27 | 1,5E-02 |
| ILMN_1673966 | POLR3F | polymerase (RNA) III (DNA directed) polypeptide F, 39 kDa | 1,26 | 3,9E-02 | 1,16 | 1,5E-02 |
| ILMN_1681103 | AQP1 | aquaporin 1 (Colton blood group) | 0,69 | 2,2E-02 | 0,61 | 1,6E-02 |
| ILMN_1768110 | ZAK | sterile alpha motif and leucine zipper containing kinase AZK | 1,34 | 4,3E-02 | 1,26 | 1,7E-02 |
| ILMN_1785158 | HERPUD2 | HERPUD family member 2 | 0,78 | 3,8E-02 | 0,80 | 1,7E-02 |
| ILMN_1749829 | DLGAP5 | discs, large (Drosophila) homolog-associated protein 5 | 1,65 | 3,8E-02 | 1,66 | 1,7E-02 |
| ILMN_1733811 | JUP | junction plakoglobin | 0,60 | 7,3E-02 | 0,64 | 1,8E-02 |
| ILMN_1702487 | SGK1 | serum/glucocorticoid regulated kinase 1 | 1,53 | 9,3E-02 | 1,34 | 1,9E-02 |
| ILMN_1769245 | GLIPR1 | GLI pathogenesis-related 1 | 1,51 | 1,4E-02 | 1,28 | 1,9E-02 |
| ILMN_1809010 | PSMC3 | proteasome (prosome, macropain) 26S subunit, ATPase, 3 | 1,35 | 7,3E-02 | 1,40 | 2,0E-02 |
| ILMN_1801257 | CENPA | centromere protein A | 1,36 | 3,0E-02 | 1,33 | 2,1E-02 |
| ILMN_1746013 | SPOCK1 | sparc/osteonectin, cwcv and kazal-like domains proteoglycan (testican) 1 | 1,52 | 7,5E-02 | 1,29 | 2,1E-02 |
| ILMN_2230892 | IL10RB | interleukin 10 receptor, beta | 1,38 | 2,2E-02 | 1,20 | 2,2E-02 |
| ILMN_1727055 | PARPBP | PARP1 binding protein | 1,38 | 6,0E-02 | 1,28 | 2,3E-02 |
| ILMN_1695658 | KIF20A | kinesin family member 20A | 1,65 | 6,9E-02 | 1,57 | 2,3E-02 |
| ILMN_1673673 | PBK | PDZ binding kinase | 1,47 | 3,4E-02 | 1,37 | 2,4E-02 |
| ILMN_1651719 | MBTPS1 | membrane-bound transcription factor peptidase, site | 1,44 | 2,2E-02 | 1,35 | 2,5E-02 |

| | | | | | | |
|--------------|----------|--|------|---------|------|---------|
| ILMN_1693882 | TAPT1 | transmembrane anterior posterior transformation 1 | 1,20 | 6,9E-02 | 1,13 | 2,5E-02 |
| ILMN_1733931 | PDCD6 | programmed cell death 6 | 1,94 | 1,4E-02 | 1,24 | 2,8E-02 |
| ILMN_1752914 | ARHGAP31 | Rho GTPase activating protein 31 | 0,78 | 9,8E-02 | 0,79 | 3,0E-02 |
| ILMN_1664511 | NDC80 | NDC80 kinetochore complex component | 1,34 | 1,3E-02 | 1,33 | 3,2E-02 |
| ILMN_1747016 | CEP55 | centrosomal protein 55kDa | 1,61 | 9,1E-02 | 1,56 | 3,2E-02 |
| ILMN_1713124 | AKR1C3 | aldo-keto reductase family 1, member C3 | 1,64 | 1,1E-02 | 1,24 | 3,2E-02 |
| ILMN_1777564 | MAD2L1 | MAD2 mitotic arrest deficient-like 1 (yeast) | 1,40 | 3,8E-02 | 1,52 | 3,3E-02 |
| ILMN_1697024 | STYX | serine/threonine/tyrosine interacting protein | 0,67 | 4,2E-02 | 0,79 | 3,4E-02 |
| ILMN_1686097 | TOP2A | topoisomerase (DNA) II alpha 170kDa | 1,68 | 3,9E-02 | 1,77 | 3,6E-02 |
| ILMN_2313672 | IL1RL1 | interleukin 1 receptor-like 1 | 1,77 | 7,3E-02 | 1,42 | 3,8E-02 |
| ILMN_1760201 | DNMT1 | DNA (cytosine-5-)-methyltransferase 1 | 1,42 | 3,7E-02 | 1,47 | 4,0E-02 |
| ILMN_1711408 | ANXA4 | annexin A4 | 0,77 | 4,3E-02 | 0,87 | 4,0E-02 |
| ILMN_1733562 | TFB1M | transcription factor B1, mitochondrial | 1,20 | 5,3E-02 | 1,23 | 4,1E-02 |
| ILMN_1738784 | PPP2R5A | protein phosphatase 2, regulatory subunit B', alpha | 0,79 | 6,1E-02 | 0,86 | 4,9E-02 |
| ILMN_1815705 | LZTFL1 | leucine zipper transcription factor-like 1 | 1,39 | 1,7E-02 | 1,15 | 5,0E-02 |
| ILMN_1764769 | VWA5A | von Willebrand factor A domain containing 5A | 0,75 | 3,4E-02 | 0,81 | 5,2E-02 |
| ILMN_1708537 | RBPJ | recombination signal binding protein for immunoglobulin kappa J region | 1,25 | 2,9E-02 | 1,15 | 5,3E-02 |
| ILMN_2405018 | PPP1CB | protein phosphatase 1, catalytic subunit, beta isozyme | 1,36 | 4,0E-02 | 1,25 | 5,4E-02 |
| ILMN_1801584 | CXCR4 | chemokine (C-X-C motif) receptor 4 | 0,49 | 1,7E-02 | 0,55 | 5,4E-02 |
| ILMN_2209180 | RPRD1A | regulation of nuclear pre-mRNA domain containing 1A | 1,39 | 2,9E-02 | 1,23 | 5,8E-02 |
| ILMN_1779228 | CDH2 | cadherin 2, type 1, N-cadherin (neuronal) | 1,47 | 8,3E-02 | 1,20 | 6,5E-02 |
| ILMN_1684982 | PK4 | pyruvate dehydrogenase kinase, isozyme 4 | 0,78 | 7,3E-02 | 0,81 | 6,6E-02 |
| ILMN_1764043 | TTL | tubulin tyrosine ligase | 1,67 | 5,9E-03 | 1,19 | 7,4E-02 |
| ILMN_1679929 | KLF13 | Kruppel-like factor 13 | 0,81 | 4,3E-02 | 0,84 | 8,0E-02 |
| ILMN_1713482 | CWC15 | CWC15 spliceosome-associated protein | 1,42 | 5,2E-02 | 1,30 | 8,1E-02 |
| ILMN_2169089 | C18orf54 | chromosome 18 open reading frame 54 | 1,26 | 3,4E-02 | 1,15 | 8,2E-02 |
| ILMN_1802888 | ZNF185 | zinc finger protein 185 (LIM domain) | 1,33 | 6,1E-02 | 1,36 | 8,3E-02 |
| ILMN_1757129 | TMEM88 | transmembrane protein 88 | 0,50 | 1,7E-02 | 0,83 | 8,9E-02 |
| ILMN_1712452 | KIF20B | kinesin family member 20B | 1,44 | 3,9E-02 | 1,23 | 9,0E-02 |

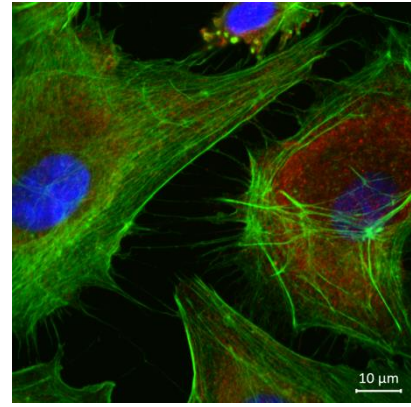
Supplementary Table 1. Genes affected by SASH1 knockdown in HAECs (FDR<10) List of genes differentially expressed after SASH1 silencing in HAECs (n=9)



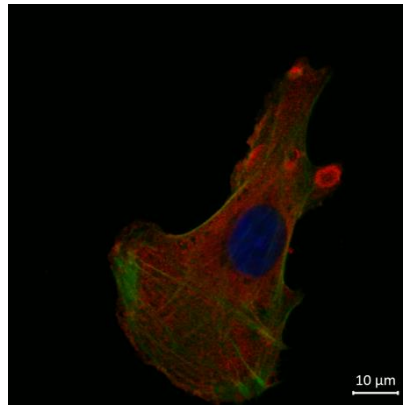
**Cell signalling #3900
rabbit IgG**



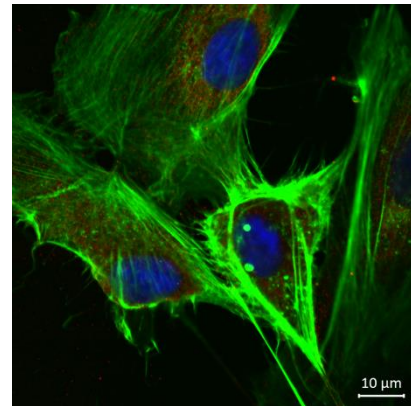
**Novus NBP1 26650
SASH1**



**Novus NBP1 26651
SASH1**

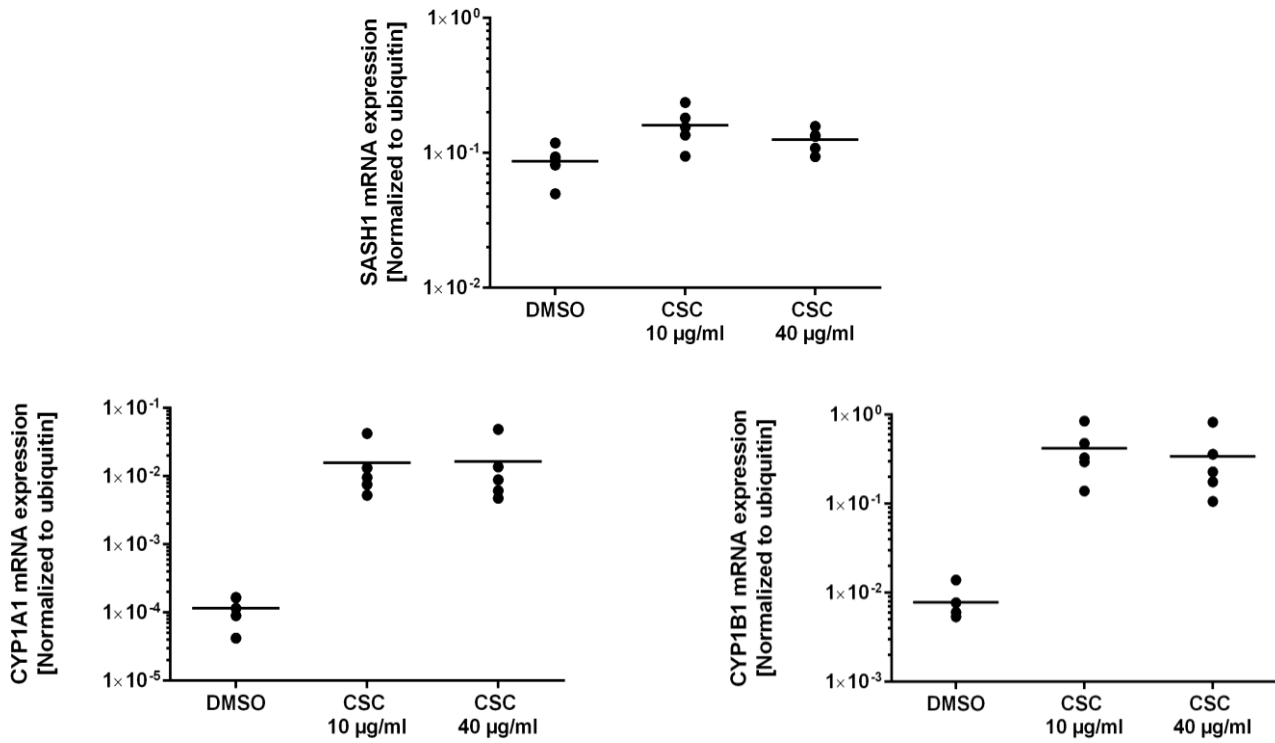


**Sigma S8073
SASH1**



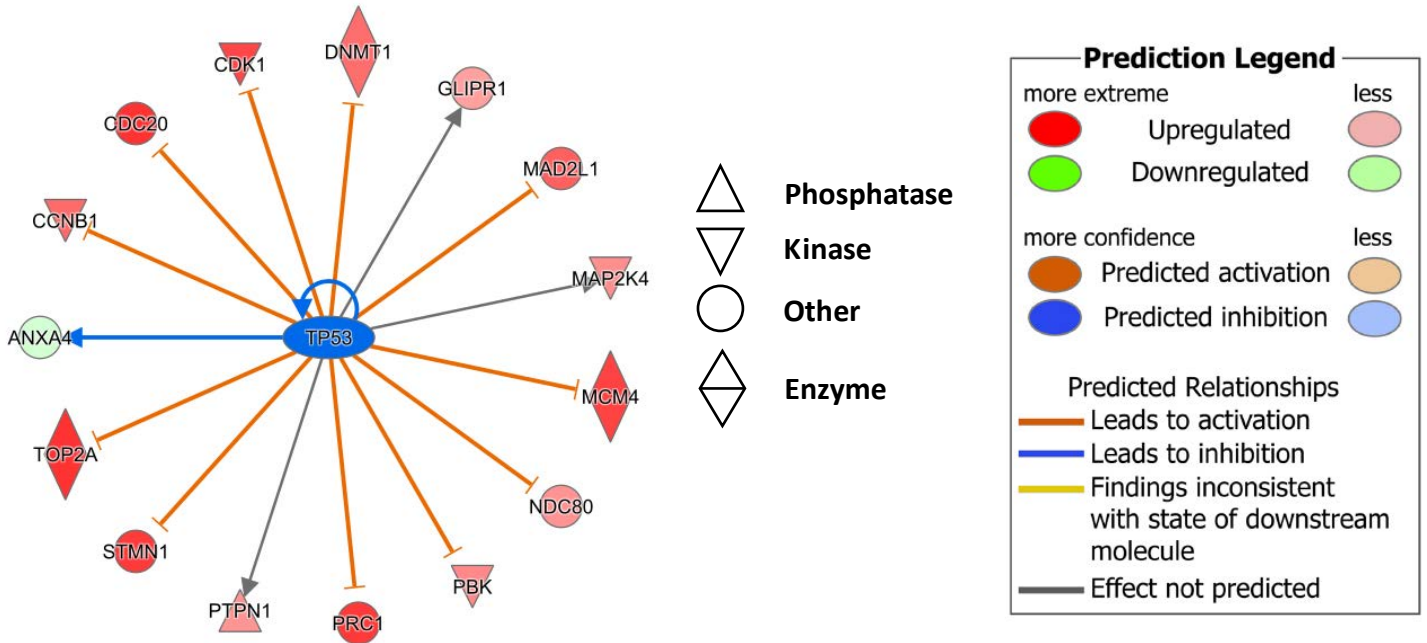
**Santa Cruz Sc169253
SASH1**

Supplementary Figure 1. SASH1 subcellular localisation in HAECs. Confocal microscope microphotographs of HAECs stained with 4 different antibodies targeting human SASH1 (red), actin cytoskeleton (Phalloidine-488, green), negative control for SASH1 with rabbit IgG and nucleus (DAPI, blue).



Supplementary Figure 2. SASH1 expression in HAECs stimulated with cigarette smoke condensate (CSC) for 24h. RT-qPCR expression of SASH1 and CYP1A1 and CYP1B1. Results are expressed as fold change, n=5 different individuals.

| | Predicted Activation State | Activation z-score | p-value of overlap | Target molecules in dataset |
|-------|----------------------------|--------------------|--------------------|--|
| FOXM1 | Activated | 2.157 | 7.60E-12 | CCNA2, CCNB1, CDC20, CDK1, CENPA, KIF20A, NEK2, PLK4, PRC1, STMN1 |
| TP53 | Inhibited | -3.359 | 3.87E-07 | ANXA4, BUB1, CCNA2, CCNB1, CDC20, CDK1, CEP55, DNMT1, DUSP5, GLIPR1, MAD2L1, MAP2K4, MCM4, NDC80, NEK2, PBK, PRC1, PTPN1, STMN1, TOP2A |



Supplementary Figure 3. Predicted activation of transcription factors in microarrays analysis of HAECs silenced for SASH1. IPA prediction of the activated or inhibited state of the transcription factor FOXM1 and TP53 based on the number of target genes affected by SASH1 knockdown.

II.3 SASH1 ongoing experiments and unpublished results

II.3.a Materials and methods

Methylation analysis

The Martha cohort

The MARTHA study is a collection of 1,542 patients with venous thrombosis (VT) recruited from the Thrombophilia centre of La Timone hospital (Marseille, France)^{169,170,171,172}. All subjects had a documented history of VT, were free of chronic diseases, and were free of inherited thrombophilia including: anti-thrombin, protein C and protein S deficiencies and homozygosity for the Factor V Leiden and Factor II G20210A mutations. For the methylation project, 349 MARTHA patients were randomly selected for DNA methylation analysis^{173,174,175,176}.

F5L study

The family study is composed of five extended French-Canadian pedigrees, totaling 255 relatives, ascertained at the Thrombosis Clinic of the Ottawa Hospital through single probands with idiopathic VT and heterozygote for the Factor V Leiden mutation. Probands were free of acquired VT risk factors such as cancer, myeloproliferative disease, pregnancy, puerperium, prolonged immobilization, trauma, surgery and antiphospholipid syndrome, and were free of inherited thrombophilia. A detailed description of this study can be found in previous work from the group¹⁶⁹. Only 218 family members for whom DNA was still available were included in the current work.

The Array

From the 485,577 probes available on the Illumina array, were excluded probes that measured single nucleotide polymorphisms (n = 65), that are either cross-reactive (n = 30,969) or polymorphic at the targeted CpG site (n = 66,877). Of note, 4,464 probes shared the two last features.

Methylation values were corrected for background by use of the Noob method implemented in the "methyumi" package¹⁷⁷, for dye bias following the manufacturer's recommendation http://support.illumina.com/downloads/genomestudio_m_module_v18 Ug%2811319130 b%29.i lmn and normalized for design type bias according to the SWAN method¹⁷⁸ implemented in the minfi R package¹⁷⁹. Probes (n = 4,010) with a detection p-value (as described in the "minfi" package) greater than 0.05 in more than 5% of the total processed samples were then excluded from further analyses. This led to a final selection of 388,120 probes (among which 1,289 tagged for CpH sites) that were tested for association with the homocysteine.

Statistical analysis:

In the MWA study, linear regression analyses were used to assess the influence of the tobacco, as covariates, on each CpG sites, as the outcome. In the F5L pedigree study, a linear mixed-model accounting for the non-independence between family members was used. All analyses were

adjusted for age, sex, and batch and chip effects. The two last variables correspond to microarray plate on which samples were processed and their position on the plate, respectively to account for experimental variability between position and plates¹⁸⁰. Because DNA methylation levels measured in peripheral blood DNA reflect the average level of DNA methylation in different cell types including lymphocytes, monocytes, neutrophils, basophils and eosinophils, all analyses were also adjusted for cell type composition to avoid any contamination bias^{181,182,183}. For MARTHA samples we used available specific biological counts of lymphocytes, monocytes, neutrophils, eosinophils and basophils to characterize their cell type composition (measured with the cellular hematologic analyzer ADVIA©) 120 Hematology System (Siemens Healthcare Diagnostics, Deerfield, IL). In F5L pedigrees the cell type counts were not available, however we handled adjustment for cell type composition using the method described in Koestler et al¹⁸⁴.

Mass spectrometry

In-gel digestion of proteins (colored with silver nitrate)

Each lane of the SDS-PAGE gel was cut in 16 strips. The strips were then discolored with a solution of 15mM potassium ferricyanide and 50 mM thiosulfate and cut down into bits.

After several alternate washing with H₂O and acetonitrile the proteins were reduced and alkylated by immersion in 10mM dithiotreitol (DTT) diluted in 50mM ammonium bicarbonate (30 min at 56°C) followed by immersion in 50mM iodoacetamide diluted in 50mM AMBIC (30 min RT).

After dehydration with acetonitrile and drying under a fume hood, gel strips were re-hydrated for 30 min, on ice, with a solution of trypsin 10 ng/μl diluted in 50mM AMBIC/5% acetonitrile (20ul per strip). Enzymatic digestion was done overnight at 37°C.

Gel strips were then washed 2 times with a 60% acetonitrile/ 0.1% TFA solution, peptides in the supernatant were dried using a speed-vac, resuspended in 3 μl of 30%acetonitrile/0.1% TFA and further diluted to a final volume of 20 μl with 0.1% TFA.

Analysis using LC-MS/MS with ion trap (HTC ultra, Bruker)

The solutions containing the peptides were first concentrated and desalted for 5 min on a RP-C18 pre-column (5 mm, 300 μm i.d., 100 Å) using a mobile phase of 2%acetonitrile/0.1% formic acid, at a rate of 20 μl/min. Peptides were then separated on a RP-C18 analytic column (15 cm, 75μm i.d., 5 μm particles, 100 Å) and eluted with a 2-40% gradient of 95% acetonitrile/0.1% formic acid for 60 min at a rate of 300 nl/min. Total acquisition time was 90 min including C18 columns washing and equilibration.

Peptides were then analyzed by mass spectrometer using positive mode, at a source voltage of 1.9kV. For each MS spectra, the 8 most detected peptides were fragmented in the analyzer by collision induced dissociation.

Protein identification using Mascot software

MS/MS data were processed using the Data analysis software (Bruker) to acquire fragmentation data obtained between 18 and 72 min with a 300'000 (UA) threshold. Peptides were then identified using the Mascot software (version 2.2.07, Matrix Science, London). The parameters for identification were as follow:

- Data bank: UniProt Human (21/11/14)
- Digestion enzyme: trypsin
- 1 missing trypsin cut accepted
- Constant modifications: Cystein carbamidomethylation (alkylation IAM)
- Variable modification: methionines oxydation, N-terminus acetylation
- Masse error tolerance: 0.5 in MS and 0.5 in MS/MS
- Minimum score for peptide validation: 38 (equivalent to a $p < 0.05$)
- Protein validation: at least 2 unique peptides with a score above 38

Yeast two hybrid screening for SASH1 partners

Yeast two-hybrid screening was performed by Hybrigenics Services, S.A.S., Paris, France (<http://www.hybrigenics-services.com>). The coding sequence for Homo sapiens - SASH1 (aa 1-1247) (GenBank accession number gi: 45935384) was PCR-amplified and cloned into pB27 plasmid as a C-terminal fusion to LexA (N-LexA- SASH1-C-Fusion). The construct was verified by sequencing the entire insert and used as a bait to screen a random-primed human placenta cDNA library constructed into pP6. pB27 and pP6 derive from the original pBTM116¹⁸⁵ and pGADGH¹⁸⁶ plasmids, respectively. 22.1 millions (pB27_C) + 42.9 (pB27_D) millions clones (2-fold the complexity of the library and 4 fold the complexity of the library) were screened using a mating approach with YHGX13 (Y187 ade2-101::loxP-kanMX-loxP, mat) and L40ΔGal4 (mata) yeast strains as previously described¹⁸⁷. 175 (pB27_C) + 272 (pB27_D) His⁺ colonies were selected on a medium lacking tryptophan, leucine and histidine, and supplemented with 20mM 3-aminotriazole to handle bait autoactivation. The prey fragments of the positive clones were amplified by PCR and sequenced at their 5' and 3' junctions. The resulting sequences were used to identify the corresponding interacting proteins in the GenBank database (NCBI) using a fully automated procedure. A confidence score (PBS, for Predicted Biological Score) was attributed to each interaction as previously described¹⁸⁸.

Description of the Predicted Biological Score:

The Predicted Biological Score (PBS) relies on two different levels of analysis. Firstly, a local score takes into account the redundancy and independency of prey fragments, as well as the distribution

of reading frames and stop codons in overlapping fragments. Secondly, a global score takes into account the interactions found in all the screens performed at Hybrigenics using the same library. This global score represents the probability of an interaction being nonspecific. For practical use, the scores were divided into four categories, from A (highest confidence) to D (lowest confidence). A fifth category (E) specifically flags interactions involving highly connected prey domains previously found several times in screens performed on libraries derived from the same organism. Finally, several of these highly connected domains have been confirmed as false-positives of the technique and are now tagged as F. The PBS scores have been shown to positively correlate with the biological significance of interactions^{189,190}.

Co-immunoprecipitation

HAECs were transfected at 90 % confluency, with 20 μ M of either SASH1 siRNA or siCtrl using RNAiMAX. After 48h cells were stimulated with 100ng/ml LPS for 20 min and lysed in modified RIPA (Thermo scientific, 87787) compatible with co-immunoprecipitation. Cell lysates were then incubated at a quantity of 300 μ g protein per 50 μ l of magnetic beads (Life Technologies, 10001D) coupled with a DMP crosslinked antibody targeting either SASH1 (Novus Biological NBP1-26650) or targeting SASH1 candidate partner proteins Rac1 (Lsbio, LS-C9852), IKBKB (Cell signaling #8943, Novus Biological NBP1-32258, alpha catenin (Cell signaling #3240, Novus Biological NB110-55563), AFF1 (Novus Biological NBP1-28728), OAS2 (BioCat GmbH TA802773) and C1D (Novus Biological NBP2-15639). Beads were then washed 3 times with PBS-tween 0.05% and eluted with 50 μ l of 50 mM Glycine pH 2.5. Each experiment included co-immunoprecipitation of SASH1 and co-immunoprecipitation of a SASH1 candidate partner proteins.

II.3.b Detecting the SASH1 protein in HAECs

When this project started, little was known about SASH1 and no experiments had been performed at the protein levels. We thus tested several antibodies to detect SASH1. The main difficulty encountered in western blots was that SASH1 had a non-specific signal in its direct vicinity, later identified by mass spectrometry as myosin heavy chain 9 (MYH9). Signal for MYH9 in western blots is stronger than SASH1, with just a few KD of difference. Another isoform of MYH9 with a strong signal also appears about 20-30kDa above SASH1. To separate those two bands we used an acrylamide 6% gel, and performed migration until the 70kDa marker reached the bottom of the gel (hence all protein present in the cell lysate with a molecular weight below 70 are lost). When using the 8 % or gradient gels SASH1 band was in close contact with the non-specific MYH9 band. These conditions allowed us to obtain SASH1 band as an independent signal (Figure 9).

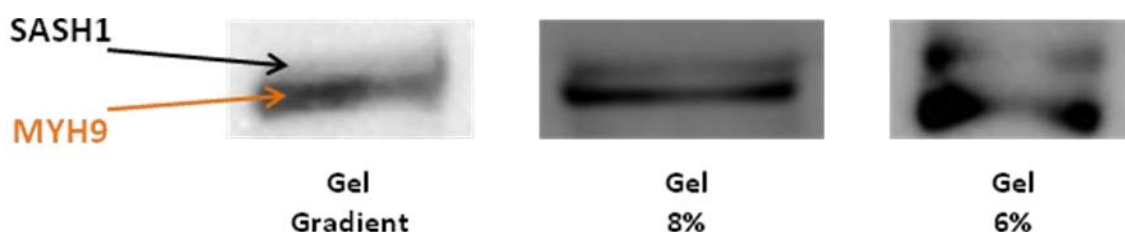


Figure 9. SASH1 detection in HAECs by western blotting. SASH1 (upper band) and the non specific MYH9 (lower band) using either a gradient, 8% or 6% polyacrylamid gel.

The efficiencies of antibodies in the test for each experiment are described in the table 1.

| | clonality | WB | IF | IP |
|------------------------------|------------|-----|-----|-----|
| Novus NBP1-26650 | polyclonal | Yes | Yes | Yes |
| Novus NBP1-26651 | polyclonal | No | Yes | No |
| Santa Cruz 169253 | polyclonal | Yes | Yes | No |
| Abcam ab110776 | polyclonal | No | NA | No |
| Ab Nova H00023328-M01 | monoclonal | No | NA | No |
| Sigma Aldrich S8073 | Polyclonal | No | Yes | No |

Table 1. Testing commercial SASH1 antibodies in western blotting (WB), immunofluorescence (IF) and immunoprecipitation (IP).

While SASH1 has a molecular weight of 137kDa, as calculated *in silico* from its sequence, it was detected at around 170-180kDa (as described in the literature). However migration on a 6% acrylamide gel for a longer period of time resulted in SASH1 signal being detected at around 150 kDa. This suggests that the migration of SASH1 at a higher molecular weight could be due to its structure being resistant to denaturation and thus slowing down the protein migration in the gel, rather than the protein itself undergoing post translational modification. Nevertheless, these observations do not exclude the possibility that both parameters could be involved in the fact that SASH1 migrates at a higher molecular weight than expected.

II.3.c SASH1 expression in vascular cells

We investigated SASH1 protein expression in several human vascular cells: Human Aortic Endothelial Cells (HAECs), Human Coronary Artery Smooth Muscle Cells (HCASMC), monocytes and macrophages (Figure 10). SASH1 expression was found to be mostly cytoplasmic, with enrichment at the leading edge of the lamellipodia observed in endothelial cells and macrophages (orange arrows). SASH1 staining was also present in the nucleus of all cell types except in monocytes, where it was restricted to the cytoplasm.

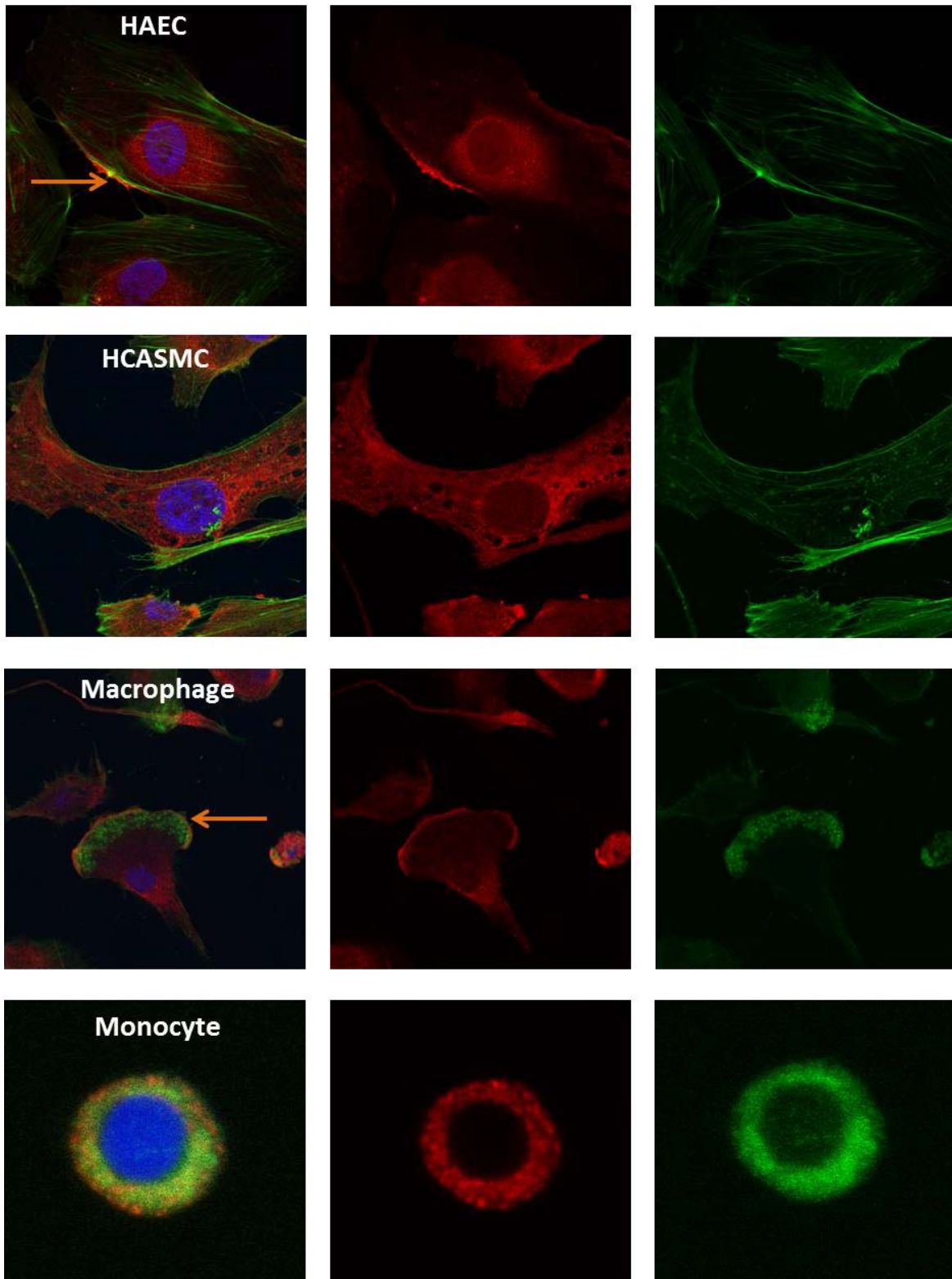


Figure 10. SASH1 immunostaining in vascular cells. Immunostaining for SASH1 (red), actin cytoskeleton (Phalloidine-488, green) and nucleus (DAPI, blue). SASH1 expression was mostly cytoplasmic, with enrichment at the leading edge of the lamellipodia observed in HAECs and macrophages (orange arrows). SASH1 staining was also present in the nucleus of all cell types except in monocytes, where it was restricted to the cytoplasm.

II.3.d Testing the effect *SASH1* knockdown on the *SASH1* network

To investigate whether the gene expression of key candidates from the *SASH1* statistically predicted network (Figure 8) was affected by the knockdown of *SASH1*, we performed a series of knockdown experiments on HAECs. We selected a number of candidate genes from the *SASH1* network, either directly connected to *SASH1* or directly connected to smoking (*PPARg*, *ADAMDEC*, *CD36*, *PTGDS*, *TJP2*). The *SASH1* knockdown led to a decrease of >90% of the *SASH1* transcript as well as a visible reduction of the protein in western blot, however, no difference in expression of candidate genes was observed (Figure 11).

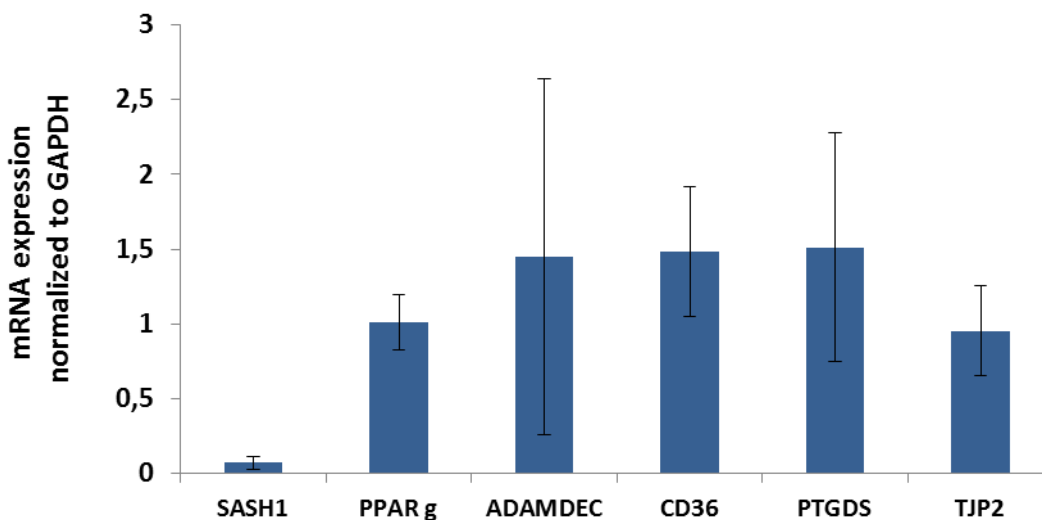


Figure 11: Gene expression of the main genes from the *SASH1* predicted network following *SASH1* knockdown in HAECs. Gene expression measured by RT-qPCR in HAECs after a 48h *SASH1* knockdown. Results are expressed as fold change +/- standard deviation of n=9 experiments. No difference in the expression of these selected genes was observed. *ADAMDEC* expression had a Ct value above 30 and was therefore considered null.

It was concluded that the expression of these candidate genes from the network was independent from the expression of *SASH1*.

II.3.e Effect of HAECs stimulation with inflammation related compounds on *SASH1* expression and its sub cellular localization

As no change of *SASH1* expression was observed in HAECs incubated with cigarette smoke condensate,, we investigated whether other stimuli relevant to the pathology of atherosclerosis could significantly affect *SASH1* expression. We first tested the stimulation of HAECs with VEGF, the main activator of migration, proliferation and angiogenesis of endothelial cells, as *SASH1* had been shown to affect the migration and proliferation of HAECs. Four genes known to be affected by VEGF stimulation were used as positive controls, *CXCR4*, *ESM1*, *TGF α* and *ANGPT2*. *SASH1* expression showed no variation upon VEGF stimulation (50 ng/ml), while the expression of all control genes was significantly affected (Figure 12). Thus, it is likely that the effects of *SASH1* on migration,

proliferation and angiogenesis are not VEGF dependent. However, we cannot exclude the possibility of a regulation only at the protein level, following VEGF stimulation.

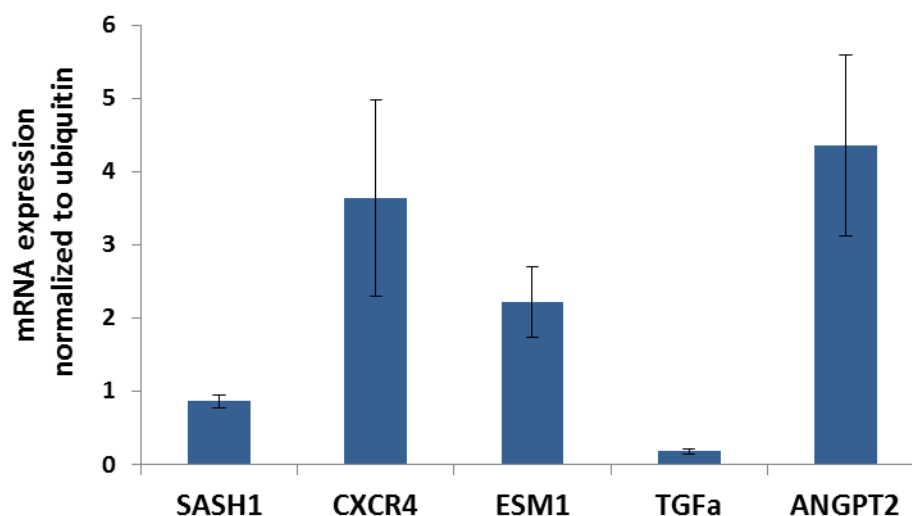


Figure 12: Expression of SASH1 and positive control genes in HAECs stimulated with VEGF. Gene expression measured by RT-qPCR in HAECs stimulated for 24h with VEGF (50ng/ml). Results are expressed as fold change +/- standard deviation of n=5 experiments. SASH1 expression was unchanged following stimulation CXCR4, ESM1 and ANGPT2 were increased and TGFa was decreased, indicating that VEGF stimulation was successful but had no effect on SASH1 expression.

SASH1 has been suggested to be a scaffold protein for the IKBKB/IKBKA complex activated downstream of the TLRA receptor after LPS stimulation¹¹⁶. Since LPS affects several TLR receptors, we searched for a stimulation that could be specific of the TLR4 receptor. The answer came from the publication of Schmidt et al¹⁹¹, which demonstrated that stimulation by nickel in mice induced inflammation solely via TLR4. Hence stimulation of human cells with nickel is likely to trigger an inflammatory response, solely mediated by TLR4, which would confirm the specificity of SASH1 to this receptor. HUVEC and THP1 cells were exposed to a wide range of nickel sulfate (NiSO₄) treatment (15, 30, 75, 100, 300, 1000 ng/ml) for 6 and 24h. All concentrations above 75ng/ml induced detectable levels of cytotoxicity (as confirmed by MTT assay) and were thus discarded from the analysis. No change in SASH1 expression was observed (Figure 13).

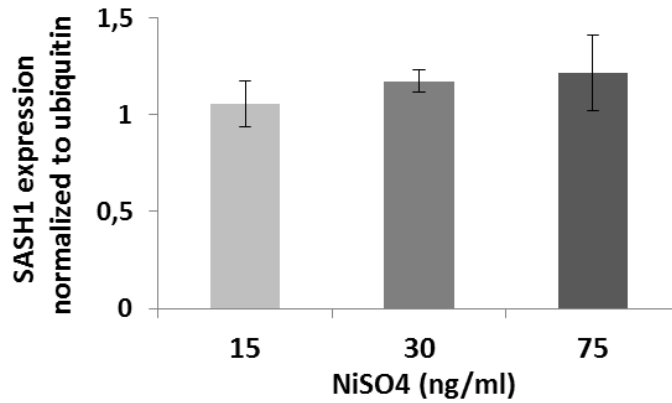


Figure 13: SASH1 expression in THP1 stimulated with NiSO4. SASH1 expression in THP1 after 24h stimulation with NiSO4 was measured by RT-qPCR. Results are expressed as fold change +/- standard deviation of n=4 experiments. SASH1 expression remained unchanged following NiSO4 stimulation.

Since SASH1 expression was not affected by 3 tested compounds (CSC, VEGF and nickel), we explored the possibility that those stimuli could lead to changes in SASH1 subcellular localization. SASH1 protein sequence includes both NLS and NES, consequently we were interested in searching for its possible nuclear translocation. A new set of experiments was performed to study SASH1 translocation using immunocytochemistry. We stimulated HAECs with TNF α (5 ng/ml), CSC (40 ng/ml), LPS (100 ng/ml) or VEGF (50 ng/ml) for 1, 6 and 24h; however no change in SASH1 subcellular location was observed in any of the tested conditions (Figure 14).

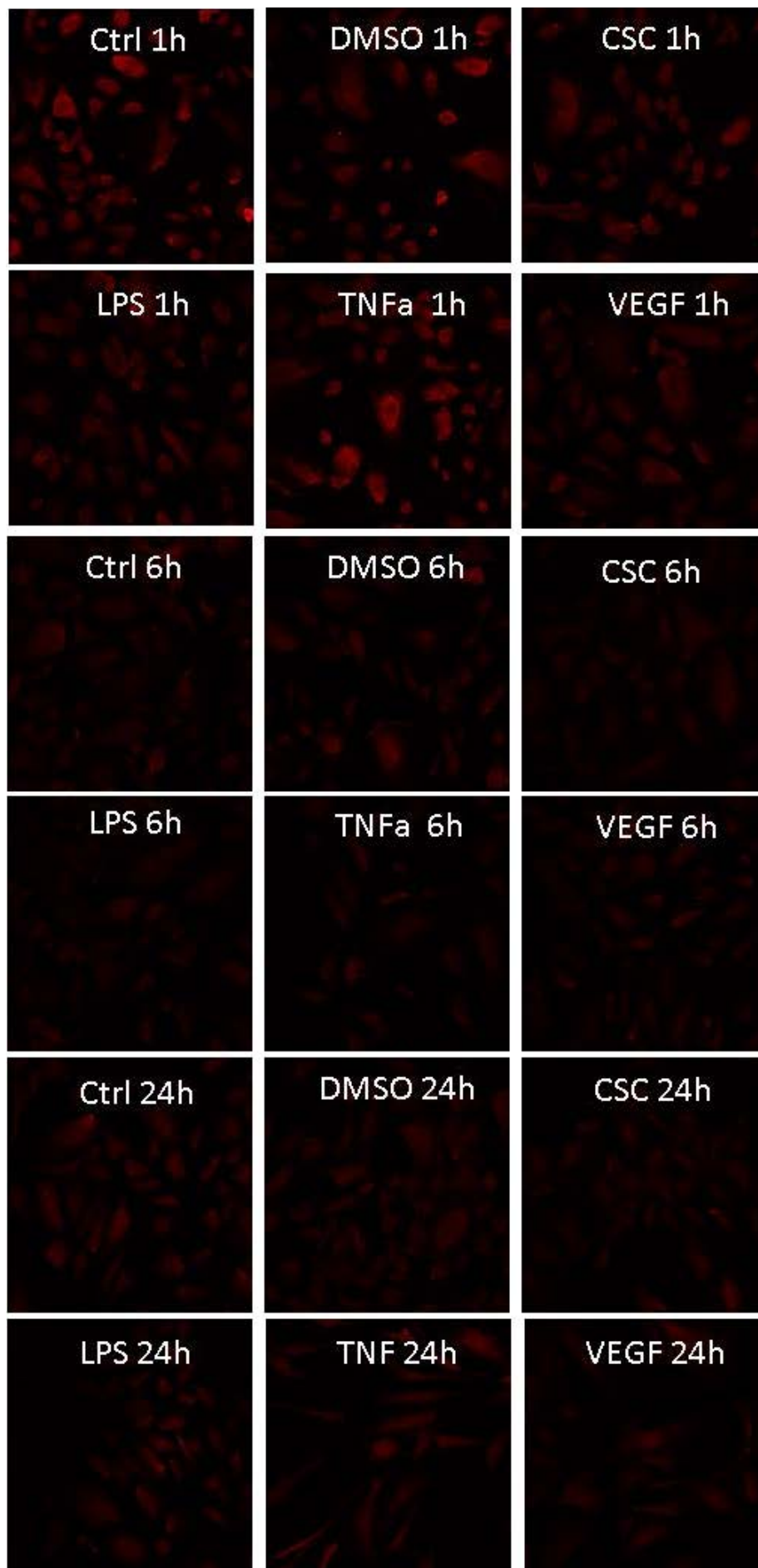


Figure 14. Immunostaining of SASH1 in HAECs following stimulation with various inflammation-related stimuli for 1, 6 and 24h to investigate SASH1 translocation into the nucleus. SASH1 (red) location was unchanged after stimulation with CSC (10 μ g/ml) or TNF α (10ng/ml) or VEGF (50ng/ml) or LPS (10ng/ml).

II.3.f Transcriptomic analysis of cigarette smoke condensate stimulated HAECs

To investigate the effect of cigarette smoke on endothelial cells, we performed transcriptomic analysis of CSC stimulated HAECs. Cells were treated either with 40 ng/ml or 10 ng/ml of CSC, or with the control vehicle (DMSO). Microarrays were performed on HAECs from two different donors. Results showed that 19 and 8 genes were differentially expressed upon 40 ng/ml and 10 ng/ml respectively. Among the differentially expressed genes were the already known members of the P450 cytochrome family, *CYP1A1* and *CYP1B1*, as well as two major inflammation related enzymes *PLA2G4A* and *PTGS2*. Those genes were significantly affected by both concentrations of CSC along with two other genes *TOP2A* (Topoisomerase II alpha) and *ANLN* (Anillin, Actin Binding Protein). Surprisingly *MMP1* and *HMOX1*, two genes which expression is known to be strongly affected by cigarette smoke exposure, were not significantly up-regulated in our microarrays data. However, the RT-qPCR analysis of the same samples revealed that *MMP1* and *HMOX1* were in fact significantly up-regulated by more than 6 and 3 fold respectively, whereas *CYP1A1* and *CYP1B1* were both upregulated by more than 50 fold. These results confirmed that *CYP1A1* and *CYP1B1* are very sensitive and reliable markers for cigarette smoke exposure. *PLA2G4A* and *PTGS2* (also known as *COX2*) being also strongly affected by CSC exposure, it is likely that this stimulation triggers a pro inflammatory reaction. As *SASH1* has been hypothesized to be involved in inflammation¹¹⁶ using CSC in our *in vitro* model to simulate inflammation, seemed relevant. Finally two genes *TOP2A* and *ANLN* came up significantly down-regulated in our experiments, while little is published linking *TOP2A* to smoking, nothing, to our knowledge as been shown linking smoking to *ANLN* in the literature. As *ANLN* is an important actor of the cellular process of cytokinesis, it would be interesting to analyze it further in the context of signaling pathways linking smoking and cellular proliferation.

II.3.g *In silico* study of SASH1 methylation

In collaboration with Dylan Aïssi and Dr David Tregouet from our team, we investigated *SASH1* methylation *in silico*, using data from the MARTHA and F5L cohort, as our data suggested that it could affect *SASH1* expression following chronic cigarette smoke exposure. These data were available to the team as we are also participating to a study linking genome methylation pattern to venous thrombosis. Data on smoking status in Martha study were available for 347 individuals including 201 non-smokers, 98 smokers and 48 ex-smokers. Linear regression was performed to compare each group, using the non smoker group as a reference. No CpG sites were found to be differentially methylated depending on the smoking habit (corrected Bonferonni $p < 8.33 \cdot 10^{-4}$). We also checked for the effect of age on *SASH1* methylation and found several sites which methylation levels significantly changed with aging. (Table 2).

| Probe Name | MARTHA | | | | | | F5L-Pedigrees | | | | | | UCSC RefGene Name | | | | |
|------------|--------------|------------|---------------------------|-------------------------|----------------------------|--------------------------|---------------|------------|----------------------------|--------------------------|--------------|------------|----------------------------|--------------------------|-----|-----------|-------|
| | Age Estimate | Age Pvalue | Smoking Ex vs No Estimate | Smoking Ex vs No Pvalue | Smoking Yes vs No Estimate | Smoking Yes vs No Pvalue | Age Estimate | Age Pvalue | Smoking Yes vs No Estimate | Smoking Yes vs No Pvalue | Age Estimate | Age Pvalue | Smoking Yes vs No Estimate | Smoking Yes vs No Pvalue | CHR | MAPINFO | |
| cg02233217 | -6.35E-05 | 8.82E-02 | -1.13E-04 | 9.37E-01 | -2.61E-04 | 8.22E-01 | 8.02E-05 | 4.92E-01 | -3.84E-04 | 9.08E-01 | 8.02E-05 | 4.92E-01 | -3.84E-04 | 9.08E-01 | 6 | 147891156 | SAMD5 |
| cg18818644 | -1.32E-04 | 9.86E-02 | 4.84E-04 | 8.74E-01 | -3.21E-03 | 1.98E-01 | -1.02E-05 | 9.60E-01 | -9.11E-03 | 1.14E-01 | -1.02E-05 | 9.60E-01 | -9.11E-03 | 1.14E-01 | 6 | 147924672 | |
| cg25105578 | -5.36E-05 | 2.81E-01 | -1.03E-04 | 9.57E-01 | -8.46E-05 | 9.57E-01 | -1.03E-06 | 9.94E-01 | 4.57E-04 | 9.02E-01 | -1.03E-06 | 9.94E-01 | 4.57E-04 | 9.02E-01 | 6 | 147973101 | |
| cg20778199 | -1.70E-03 | 2.98E-10 | -1.44E-02 | 1.49E-01 | -9.39E-03 | 2.51E-01 | -1.47E-03 | 1.17E-05 | -3.62E-02 | 1.13E-04 | -1.47E-03 | 1.17E-05 | -3.62E-02 | 1.13E-04 | 6 | 148020881 | |
| cg23345508 | -1.07E-04 | 6.51E-02 | -8.43E-04 | 7.02E-01 | 2.03E-03 | 2.60E-01 | 4.52E-04 | 8.21E-02 | 7.81E-03 | 2.89E-01 | 4.52E-04 | 8.21E-02 | 7.81E-03 | 2.89E-01 | 6 | 148079619 | |
| cg08992972 | -7.43E-05 | 1.07E-01 | -2.17E-03 | 2.19E-01 | 9.64E-04 | 5.04E-01 | 4.26E-05 | 7.75E-01 | -4.98E-03 | 2.40E-01 | 4.26E-05 | 7.75E-01 | -4.98E-03 | 2.40E-01 | 6 | 148271271 | |
| cg14673315 | -6.54E-05 | 3.78E-01 | -1.38E-03 | 6.26E-01 | 1.57E-03 | 4.98E-01 | 3.04E-05 | 3.05E-01 | 1.40E-03 | 9.77E-02 | 3.04E-05 | 3.05E-01 | 1.40E-03 | 9.77E-02 | 6 | 148336294 | |
| cg07211465 | -7.04E-05 | 3.83E-01 | -3.93E-03 | 2.03E-01 | 5.15E-03 | 4.15E-02 | -3.89E-05 | 6.01E-01 | -3.69E-03 | 8.25E-02 | -3.89E-05 | 6.01E-01 | -3.69E-03 | 8.25E-02 | 6 | 148476471 | |
| cg23489226 | -2.05E-04 | 3.77E-01 | 9.65E-03 | 2.76E-01 | 4.77E-03 | 5.10E-01 | -4.67E-05 | 7.67E-01 | 2.62E-03 | 5.58E-01 | -4.67E-05 | 7.67E-01 | 2.62E-03 | 5.58E-01 | 6 | 148524798 | |
| cg26588612 | -2.19E-04 | 5.09E-03 | -7.97E-03 | 7.70E-03 | -1.96E-03 | 4.19E-01 | -6.52E-04 | 1.20E-05 | -6.43E-03 | 1.18E-01 | -6.52E-04 | 1.20E-05 | -6.43E-03 | 1.18E-01 | 6 | 148562337 | |
| cg12668216 | -6.16E-05 | 4.01E-02 | -1.47E-03 | 2.00E-01 | 3.92E-04 | 6.75E-01 | -7.12E-05 | 2.45E-01 | 1.61E-03 | 3.54E-01 | -7.12E-05 | 2.45E-01 | 1.61E-03 | 3.54E-01 | 6 | 148615459 | |
| cg20459543 | -1.29E-04 | 1.36E-04 | -6.95E-04 | 5.86E-01 | -3.85E-04 | 7.12E-01 | -1.29E-04 | 1.36E-04 | 1.10E-04 | 9.39E-01 | -1.29E-04 | 1.36E-04 | 1.10E-04 | 9.39E-01 | 6 | 148627789 | |
| cg07189366 | -9.35E-05 | 3.60E-04 | 4.65E-04 | 6.39E-01 | -1.77E-04 | 8.27E-01 | 4.02E-05 | 3.67E-01 | -9.00E-04 | 4.74E-01 | 4.02E-05 | 3.67E-01 | -9.00E-04 | 4.74E-01 | 6 | 148634858 | |
| cg22185288 | -8.96E-05 | 1.71E-02 | -3.33E-03 | 2.07E-02 | 1.30E-03 | 2.68E-01 | -8.96E-05 | 3.57E-01 | -3.59E-03 | 3.56E-02 | -8.96E-05 | 3.57E-01 | -3.59E-03 | 3.56E-02 | 6 | 148660242 | |
| cg08365388 | 1.96E-05 | 8.37E-01 | -1.83E-03 | 6.16E-01 | 2.92E-04 | 9.22E-01 | 5.08E-04 | 1.72E-03 | 2.18E-03 | 6.30E-01 | 5.08E-04 | 1.72E-03 | 2.18E-03 | 6.30E-01 | 6 | 148663353 | SASH1 |
| cg22174886 | 3.44E-04 | 8.41E-08 | -1.91E-03 | 4.25E-01 | -3.20E-03 | 1.03E-01 | 5.61E-04 | 1.08E-09 | -2.89E-03 | 2.42E-01 | 5.61E-04 | 1.08E-09 | -2.89E-03 | 2.42E-01 | 6 | 148663426 | SASH1 |

Table 2 . Effect of Age or smoking on SASH1 methylation sites. Estimates and p-values for SASH1 methylation sites preceding SASH1 coding region, comparing smokers to non-smokers (smoking Yes vs No) and ex-smokers to non smokers (smoking ex vs No). No differences in the methylation levels were observed with regard to the smoking status. Age affected several methylation sites (corrected Bonferonni $p < 8.33 \cdot 10^{-4}$).

To ensure that cigarette smoke induced CpG island methylation was detectable in our analysis, we analyzed 3 CpG islands cg05575921, cg19859270, cg21611682, located in the proximity of the of AHRR, GPR15 and LRP5 gene respectively, which methylation pattern are known to be affected by smoking¹⁹². Indeed these CpG island were found to be significantly less methylated in smokers than in non-smokers for all tested CpG island (table 3) in Martha and F5L.

| Gene | CpG site | Comparision | Estimate | P-value | Study |
|-------|------------|---------------------------|------------|----------|---------------|
| GPR15 | cg19859270 | Non-smokers VS smokers | -1.11E-002 | 1.70E-09 | F5L-Pedigrees |
| | | Non-smokers VS smokers | -1.59E-002 | 6.41E-28 | MARTHA |
| | | Non-smokers VS Ex-smokers | -6.96E-003 | 1.83E-05 | |
| AHRR | cg05575921 | Non-smokers VS smokers | -1.76E-001 | 5.11E-26 | F5L-Pedigrees |
| | | Non-smokers VS smokers | -1.72E-001 | 5.94E-40 | MARTHA |
| | | Non-smokers VS Ex-smokers | -4.76E-002 | 5.53E-04 | |
| LRP5 | cg21611682 | Non-smokers VS smokers | -2.99E-002 | 1.16E-06 | F5L-Pedigrees |
| | | Non-smokers VS smokers | -3.39E-002 | 1.81E-11 | MARTHA |
| | | Non-smokers VS Ex-smokers | -5.94E-003 | 3.16E-01 | |

Table 3: Control methylation sites for cigarette smoke exposure in the F5L and Martha cohort.

Estimates and p-values for methylation sites known to be affected by smoking, comparing smokers to non-smokers and ex-smokers to non smokers. All tested candidates showed differential methylation in smokers when compared to non-smokers and ex-smokers (corrected Bonferonni $p < 8.33 \cdot 10^{-4}$).

II.4 Searching for SASH1 partner proteins

II.4.a Yeast two hybrid (Y2H)

To uncover new partner proteins for SASH1, we decided to perform a Yeast two hybrid (Y2H) assay. As Y2H may yield many false positive and requires extensive knowledge of the method for both protocol and analysis, we decided to use the Hybrigenics commercial facility in Paris, a company specialized in this technique. We first amplified and tested a PBUd-SASH1 plasmid (provided by Dr Peter Klaus Janssen) containing the full length SASH1. After sequencing of the SASH1 insert in the plasmid we expressed the recombinant SASH1 protein in HAECs and verified it by western blotting. The plasmid was subsequently sent to Hybrigenics to perform a screening on a placental library. The choice of this library was made based on the fact that it contained most of the transcripts that were significantly affected by SASH1 knockdown in our microarrays analyses. Additionally placenta contains a broad range of cells and is highly vascularized, which make it an interesting organ to study SASH1 expression. The bait containing SASH1 was non-toxic to the yeast, but was self-activating, hence 20mM of aminotriazol was used to inhibit non specific transcriptional activation. Two screenings were performed, yielding 175 and 272 positive clones out of 22 and 42 millions tested interaction, respectively. The positive clones were then sequenced to identify SASH1 potential partners. The best candidates are listed in the Table 4. For each target a Predicted Biological Score (PBS) was computed to assess interaction reliability based on the comparison between the number of independent prey fragments found for an interaction and the chance of

finding them at random (background noise). Results were then classified in 4 different categories: A “very high confidence”, B “high confidence”, C “good confidence”. The D category mixes false positive interactions with hardly detectable specific interactions (low levels of mRNA, prey folding, prey toxicity in yeast) and thus was excluded for further experimentation.

| Gene | significance | size (kDa) |
|---------|--------------|------------|
| AFF1 | A | 131 |
| AKAP10 | C | 73 |
| ARRB1 | A | 47 |
| ASCC2 | C | 86 |
| BAP1 | C | 80 |
| C1D | A | 16 |
| CCNDBP1 | A | 40 |
| CNKSR1 | B | 79 |
| CRK | A | 33 |
| CRKL | C | 33 |
| CTNNA1 | A | 100 |
| EPS8L2 | A | 80 |
| IKBKB | A | 88 |
| JADE2 | A | 87 |

| | | |
|-----------|---|-----|
| JADE3 | A | 93 |
| KIF3B | B | 85 |
| MCRS1 | A | 51 |
| OAS2 | B | 82 |
| PPP4R1 | A | 107 |
| RBM12B | B | 118 |
| SECISBP2L | B | 121 |
| SNX1 | C | 59 |
| SPG7 | C | 88 |
| SPTBN1 | B | 274 |
| SYNE2 | B | 796 |
| TNKS2 | A | 126 |
| UTRN | C | 394 |
| EXOC8 | A | 81 |

Table 4. Candidate SASH1 partners found by Yeast two hybrid. List of the potential SASH1 interacting proteins found in human placenta library screening for SASH1 using a SASH1-pBud plasmid. A: very high confidence. B: high confidence. C: good confidence

II.4.b Co-immunoprecipitation

We selected several proteins among the candidate SASH1 partners from the Y2H for confirmation using co-immunoprecipitation (co-IP) and western blot:

Inhibitor of Kappa Light Polypeptide Gene Enhancer In B-Cells, Kinase (IKBKB): this serine kinase is part of the IKK complex that activates NF-kB by phosphorylating its inhibitor leading to its degradation¹⁹³. This protein has already been shown to co-precipitate with SASH1 in Human Microvascular Endothelial Cells (HMEC) over-expressing SASH1¹¹⁶.

Catenin alpha 1 (CTNNA1): This protein is known to be involved in the actin cytoskeleton dynamics, by competing with the ARP 2/3 complex at the actin branching points¹⁹⁴, thus inhibiting actin polymerization. As SASH1 was previously shown to precipitate with the cortactin, a major actor of the actin cytoskeleton polymerization, and actin-F (polymerized actin)¹⁰⁴, we were interested in studying the interaction of SASH1 with CTNNA1 as it was the only candidate on the list that directly interacts with the actin cytoskeleton.

2'-5'-Oligoadenylate Synthase 2 (OAS2): This protein is known to be involved in the innate immune system response to viral infection and to be part of the interferon gamma (IFN γ) pathway¹⁹⁵. It activates the latent RNase L that degrades viral RNA effectively inhibiting the virus replication.

Rho Family, Small GTP Binding Protein Rac1 (RAC1): This protein belongs to the RAS superfamily of small GTP-binding proteins. It is involved in many cellular processes, including proliferation and migration *via* its effect on the cytoskeleton dynamics. Although RAC1 did not come up in our Y2H test, SASH1 was found in several cDNA library as a prey for Rac1 (Communication personnelle, Dr. Stéphanie Blachon, Hybrigenics Services). Its implication in atherosclerosis has already been demonstrated^{196,197,198} and it is involved in the actin cytoskeleton polymerization, making it a very relevant candidate partner protein for SASH1.

Finally, **C1D Nuclear Receptor co-repressor (C1D)** and **AF4/FMR2 Family, Member 1(AFF1)** were also selected to be tested in co-immunoprecipitation. Little is known about those two transcription factors, AFF1 is believed to be part of the innate immune system response to viral infection¹⁹⁹ and C1D has been shown to be involved in the activation of apoptosis²⁰⁰. SASH1 possesses both NLS and NES signal, indicating that it can be translocated into nucleus, however its nuclear function is still unknown. As SASH1 lacks a DNA binding domain, its action should be exerted through binding to other nuclear proteins. C1D and AFF1 were the only two transcription factors in our Y2H screenings and thus could be potentially involved in SASH1 interaction in the nucleus.

The first step of our experiments was the immunoprecipitation of SASH1 using different anti-SASH1 antibodies. Out of the 6 tested antibodies, only one (Novus Biological NBP1-26650) was able to precipitate SASH1. Specificity of the isolated signal was assessed using SASH1 siRNA to inhibit SASH1 expression. It is noteworthy that, although this antibody did precipitate SASH1, it also precipitated a number of non specific proteins, including two which were very close to the SASH1 signal, in western blots (Figure 14).

Co-immunoprecipitation of SASH1 partner proteins was hindered by the overall bad quality of the commercial antibodies directed against the proteins selected from Y2H screening. Out of the 9 tested antibodies, 5 were potentially able to detect by western blotting the specific proteins. Out of those 5 antibodies only 2, both targeting CTNNA1, had a clear signal. None of the tested antibody was able to precipitate its target. They are two main reasons for this. First, several antibodies did not bind to the protein A magnetic beads, despite being rabbit IgG class antibodies, known to have high affinity for protein A. Second, several antibodies did correctly bind the beads (as confirmed by denaturing elution), but failed to bind their specific target proteins. Additionally, all tested IgG class antibodies had high affinity for several non-specific proteins that were detected by western blotting.

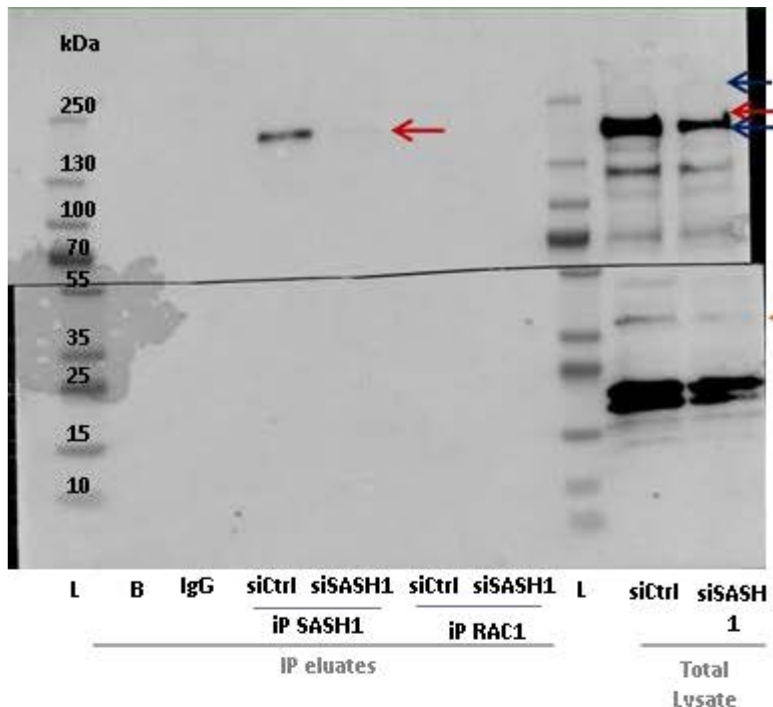
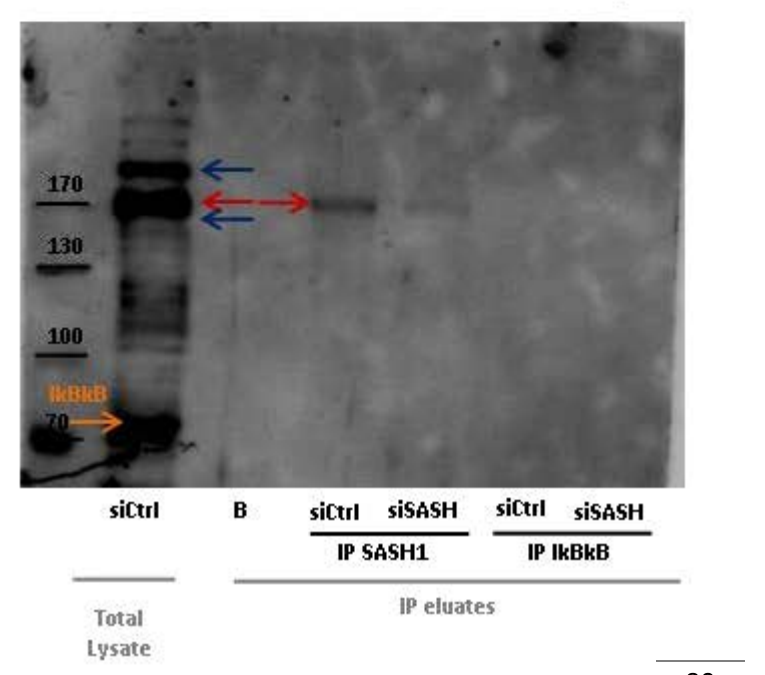
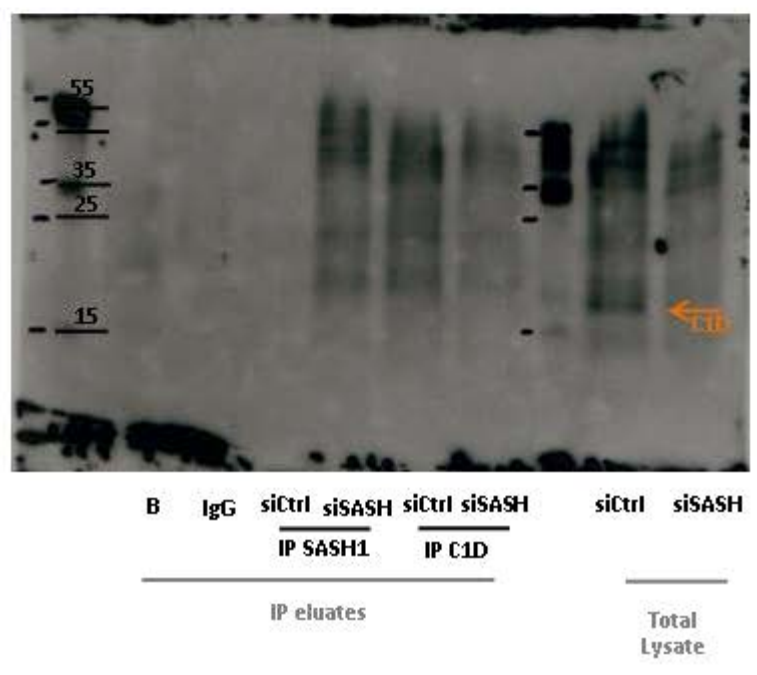
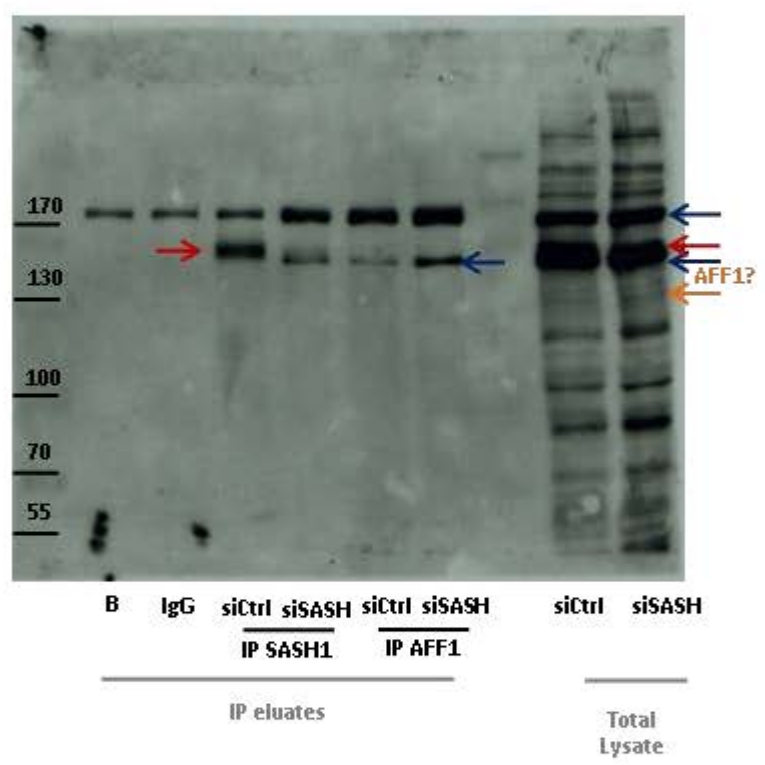
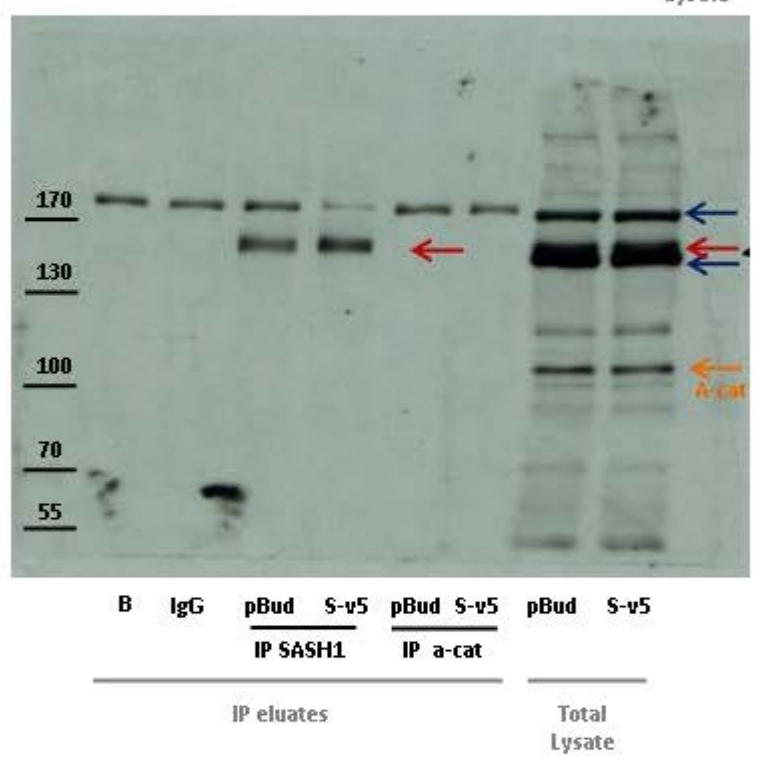


Figure 15 : Western blot of CoIP eluates of SASH1 and candidate partner proteins in HAEC. For each experiment/image, two co-IP were performed, one with SASH1 and one with the partner protein. Each membrane were first stained for SASH1 partner protein and then restained for SASH1. Only the SASH1 co-IP succeeded. Total Lysate have been loaded onto the gel to visualize the proteins.

- ← SASH1
- ← Non specific signal (MYH9)
- ← Candidate partner protein

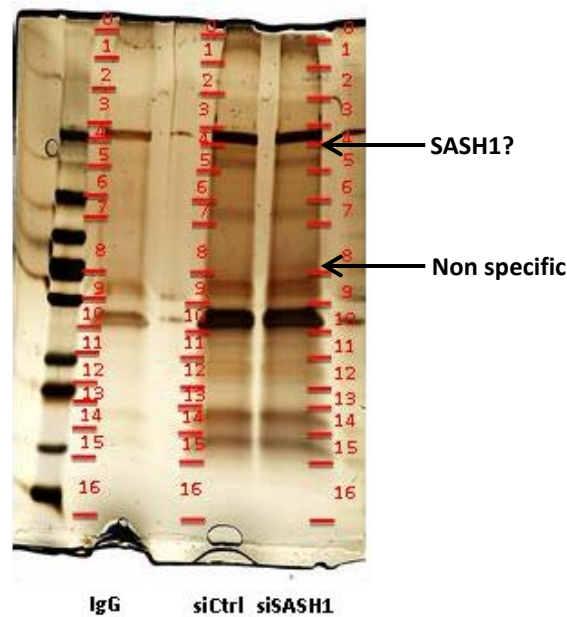
B: Beads alone, IgG: non spécifique IgG, pBud: Ctrl plasmid, S-V5: surexpression SASH1-V5, L: Ladder



II.4.c Mass spectrometry

As SASH1 co-immunoprecipitation experiments were rather inconclusive due to the incapacity of the tested antibodies to detect the specific proteins, it was decided to use 1D electrophoresis, followed by mass spectrometry analysis to identify SASH1 partner proteins in our protein lysates. In each experiment, SASH1 co-immunoprecipitation eluates from HAECs either transfected with siCtrl or siSASH1 were separated on a gradient electrophoresis gels. The gels were then stained with silver nitrate. Out of the 12 clearly observable signals (blurred or very low intensity signal are excluded for this particular observation), only two signals showed different intensity between siCtrl and siSASH1 transfected HAECs. The first one was close to the expected size for SASH1 and showed a higher intensity in the siCtrl compared to the SASH1 siRNA, it was thus considered that this band was probably SASH1. The other band showed a higher intensity in the siSASH1 transfected HAECs than in the siCtrl. It was thus hypothesized that this band was a non-specific protein which bound to the specific site of the SASH1 antibody, this site being then available because of the absence of the SASH1 protein. The other bands were considered as non-specific. As some specific signals could be either too low or hidden by a strong nearby non-specific signal, the gel was cut into slides using a pattern (Figure 16A) that aimed at isolating the non-specific bands from the area with either weak or no visible signal. Results from the mass spectrometry analysis are presented in Figure 16B.

A



B

| Protein | MW [kDa] | pI | Mascot score | | | Peptides number | | | sequence coverage (%) | | |
|---------|----------|------|--------------|--------|---------|-----------------|--------|---------|-----------------------|--------|---------|
| | | | IgG | siCtrl | siSASH1 | IgG | siCtrl | siSASH1 | IgG | siCtrl | siSASH1 |
| MYH9 | 226.4 | 5.4 | 1302 | 3443 | 3518 | 20 | 54 | 53 | 12.4 | 30.8 | 32.4 |
| VIM | 49.6 | 5 | | 490 | 498 | | 9 | 9 | | 21.3 | 21.1 |
| MYL6 | 16.9 | 4.4 | | 388 | 271 | | 7 | 5 | | 45.7 | 39.7 |
| ACTB | 41.7 | 5.2 | | 270 | 487 | | 5 | 9 | | 19.2 | 30.4 |
| MYL12B | 19.8 | 4.5 | | 248 | 226 | | 4 | 4 | | 24.4 | 24.4 |
| TPM1 | 32.7 | 4.5 | | 210 | | | 3 | | | 13.7 | |
| RPLP0 | 15.8 | 10.2 | | 114 | 122 | | 2 | 2 | | 22.5 | 15.3 |
| RPS25 | 13.7 | 10.6 | | 74 | | | 2 | | | 15.2 | |
| CALML5 | 15.9 | 4.1 | | | 108 | | | 2 | | | 25.3 |
| JUP | 81.7 | 5.7 | 424 | | | 7 | | | 10.9 | | |
| RPL12 | 17.8 | 10.3 | | | 112 | | | 2 | | | 14.5 |
| TPM3 | 26.4 | 4.6 | | | 180 | | | 3 | | | 16.3 |
| ACTB | 41.7 | 5.2 | 322 | | | 6 | | | 21.3 | | |
| TPM4 | 17.6 | 4.6 | | | 153 | | | 3 | | | 23.4 |
| S100A8 | 10.8 | 6.6 | 62 | 58 | 58 | 1 | 1 | 1 | 11.8 | 11.8 | 11.8 |

Figure 16. Mass spectrometry identification of SASH1 partner proteins. (A) Gel Cutting pattern for co-immunoprecipitation experiments (red lines). IgG: immunoprecipitation with a non specific IgG; siCtrl SASH1 co-immunoprecipitation in a lysate of HAEC transfected with siCtrl, siSASH1: SASH1 co-immunoprecipitation in a lysate of HAEC transfected with siSASH1. (B) Mass spectrometry identification of the SASH1-interacting proteins. MW: Molecular weight, pI: isoelectric point. Protein only find to precipitates under normal SASH1 expression (siCtrl) are highlighted in green: tropomyosin 1 alpha (TPM1) and ribosomal protein 25 (rps25).

Only two proteins were identified only in the siCtrl condition, where SASH1 expression is at the basal level: tropomyosin 1 (TPM1) and 40S ribosomal protein S25 (RPS25). SASH1 protein itself could not be detected, however, being in very close proximity of MYH9, one of the most abundant non-specific proteins detected in our gel, it is likely that the SASH1 detection was compromised because of the very strong MYH9 signal.

III Discussion

This study on SASH1 showed for the first time functional evidence for a link between smoking and atherosclerosis. SASH1 was described only recently, nevertheless it already appears as an important regulator of crucial cell function. SASH1 expression in the literature is correlated with a wide variety of diseases, including various cancers, preeclampsia, diabetic nephropathy, Alzheimer disease and dyschromatosis universalis^{114,106,119,117}. Considering the diverse pathophysiology of these diseases, it is likely that SASH1 is linked to basic cellular processes that underlie these pathologies. As it was demonstrated in the literature and in this study, SASH1 negatively regulates cell migration and proliferation in a wide spectrum of cell types. Alteration of either of these mechanisms may at least partially account for the development of the above mentioned pathologies. While the role of SASH1 in cancer and dyschromatosis universalis clearly relates to cell migration and proliferation, the role of SASH1 in the other diseases where it was found differentially expressed is still unknown.

The case of a possible involvement of SASH1 in preeclampsia is particularly interesting, as it could also involve the effect of SASH1 on the vascular system. Preeclampsia is a complication during pregnancy characterized by hypertension and proteinuria, leading to vascular and organ damages and sometimes, death²⁰¹. While its exact cause is unknown, an abnormally poor vascularization of the placenta could be the at the origin of this condition²⁰². As we demonstrated in this study, SASH1 knockdown significantly affected *in vitro* angiogenesis, probably because of its impact on proliferation and migration. Therefore the increased SASH1 expression observed in preeclampsia could consequently affect the neo-vascularization in the newly formed placenta by inhibiting angiogenesis. As it was shown in my thesis work that SASH1 expression potentially inhibits FOXM1, a transcription factor involved in endothelium repair¹⁶⁶, another possible effect is that the increased SASH1 expression in preeclampsia impairs the healing of vascular injury induced by hypertension. Consequently, any progress achieved in understanding the precise function of SASH1 in vascular biology will also allow to uncover the mechanisms involved in the pathophysiology of preeclampsia.

In this study, we found SASH1 to be ubiquitously distributed in the vascular wall. This result was indeed expected from what we knew about SASH1 at the beginning of the project from the literature and the available online databases. Indeed, Zeller et al¹⁰¹ found SASH1 to be expressed in all available human tissues, albeit at different levels. Additionally, all databases explored (EMBL-EBI, Human proteome atlas and TIGER) consistently showed a wide distribution of SASH1. Interestingly, all data corroborate with the fact that SASH1 expression is lower in leukocytes, a cell type where SLY, another member of SASH1 protein family is highly expressed¹⁰⁶. Since both proteins share a common domain, which over-expression has been shown to affect cell morphology, there is a possibility that both proteins, along with the third member of their family SLY2, share a common function, but are mutually exclusive. SASH1 is the best characterized member of the SLY family and interestingly the mutations in the SLY protein have been shown to affect cell migration²⁰³, comforting the hypothesis that SASH1 and SLY could be involved in similar pathways. SLY2 also showed strong functional similarities with SASH1, as its expression decreases in multiple myeloma

in humans²⁰⁴. In the same study, the authors showed that this change of expression is likely to be driven by differential methylation of the *SLY2* gene and that re-expression of *SLY2* in a mice model of multiple myeloma completely cured the disease. Furthermore, *SLY2* was also found to regulate the actin cytoskeleton dynamics in both human cancer cell lines and in transgenic mice over-expressing *SLY2*²⁰⁵. In the latter study *SLY2* over-expression increased cell spreading and membrane ruffling in human cancer cell line in a *Rac1* dependent fashion. *SLY2* also co-precipitated with cortactin, a protein also found to co-precipitate with *SASH1*¹⁰⁴. Thus the function of the *SLY* family proteins is probably redundant, however it is cell type specific as different members of this family are differentially expressed depending of the cell type.

Our results from immunostaining and fractionated cell lysis showed that *SASH1* was mostly cytoplasmic, with a weak nuclear staining, confirming earlier studies^{104,114}. Indeed, the *SASH1* protein sequence contains both nuclear localization signal (NLS) and nuclear export signal (NES), indicating that it may play a role in gene expression. Due to its important size, *SASH1* must require a transport protein, most likely an importin, to enter the nucleus. Interestingly we found an upregulation of karyopherin alpha 2 (*KPNA2*) an importin subunit involved in the recognition of the NLS complex in our experiments upon *SASH1* knockdown. *KPNA2* expression affects cell proliferation and migration in a wide range of human cancer cells^{206,207,208}, including breast cancer²⁰⁹, moreover it is over-expressed in cancer tissues and its expression correlates with poor survival²¹⁰. *KPNA2* was also shown to be a target for *p53*²¹¹. Hence, by crossing our results with those from the literature, one can hypothesize that *KPNA2* expression increases in cancerous cells by a *SASH1/p53* dependent pathways. If this hypothesis is confirmed, *SASH1* would regulate the cytoskeleton dynamics both by directly binding to the cortactin and F-actin, and indirectly by inducing transcription of proteins involved in the regulation of the cytoskeleton. Another possibility exists that *KNPA2* participates in the protein complex required for *SASH1* translocation into the nucleus, and its expression is controlled by a feedback mechanism involving *SASH1*.

Martini et al¹⁰⁴ showed that the C-terminus deleted *SASH1* construct, accumulates in the nucleus of the transfected cells¹⁰⁴. Thus it is likely that the NLS signal is hidden at basal conditions either because of the protein folding or because of an interaction with a chaperone protein. However, while several co-immunoprecipitation/GST-pull down experiments^{104,114,116}, including our own, revealed several potent partner proteins for *SASH1*, no chaperone proteins were detected. Protein conformation can be modified by post translational modification such as phosphorylation. *SASH1* is differentially phosphorylated in human fibroblasts exposed to ionizing radiation¹⁰⁵ which damages DNA, and induces DNA repair²¹². Interestingly, in our experiments, one of the two statistically significant pathways following *SASH1* knockdown in HAECs was the "Cell cycle: G2/M DNA Damage Checkpoint Regulation". It is thus possible that *SASH1* translocation may be induced by phosphorylation following DNA damage. This would explain why none of the tested stimuli (*TNF α* , cigarette smoke condensate, VEGF, LPS) elicited *SASH1* translocation, as none of them induces DNA damage in our experimental settings. On the other hand cigarette smoke contains several components with DNA damaging properties such as cadmium²¹³, which participates to the increased oxidative damage to DNA²¹⁴. However, those cigarette smoke components are present at

a relatively low levels and it is likely that a repeated exposure to smoke is required to accumulate enough toxic compounds in the tissue in order to induce DNA damage.

In our study, SASH1 expression was shown to be higher in the carotids from smokers compared to carotids from non-smokers and ex smokers. This result corroborates our previous data, showing an up-regulation of SASH1 in the human circulating monocytes from smokers, as compared to non smokers³⁸. The mechanisms behind this increased expression are yet to be determined, one hypothesis could be that DNA damage induced by cigarette smoke lead to an increased SASH1 expression, as our results suggest that SASH1 is involved in DNA repair. It could also be that chronic inflammation leads to increased SASH1 expression. Indeed, Dauphinee et al¹¹⁶ suggested that SASH1 could be involved in the TLR4/TRAF6/NF- κ B signaling pathway one of the main pathways involved in inflammation. Cigarette smoke have been shown to regulate TLR4 expression in human macrophages²¹⁵ and cigarette smoke also induces pulmonary inflammation via TLR4²¹⁶. As cigarette smoke also participates in vascular inflammation, it is likely to activate the TLR4/TRAF6/NF- κ B pathway in atherosclerosis. However Dauphinee et al did not study SASH1 expression following LPS stimulation in HMEC, additionally our data on NiSO₄ stimulated THP1 cells showed no change of SASH1 expression. Moreover, in our experiments where HAECs were stimulated with LPS prior to SASH1 co-immunoprecipitation, we did not confirm the SASH1 partners found by Dauphinee et al¹¹⁶. However, as the endogenous SASH1 protein is expressed at low levels it is likely insufficient to co-precipitate all its partners. Indeed, only two proteins, TPM1 and RPS25 were found to co-precipitate with SASH1 in our experiments.

Acute stimulation with cigarette smoke condensate (CSC) failed to increase SASH1 expression in our experiments. However, two of the 6 genes most sensitive to cigarette smoke condensate in our microarray experiments, CYP1A1 and TOP2A, were differentially expressed upon SASH1 knockdown. Thus, while CSC do not affect SASH1 expression, SASH1 expression is likely to modulate the cellular response to this CSC. The SASH1 expression was very stable in our *in vitro* experiments, regardless of the stimulation we used. Nevertheless, SASH1 is affected in a wide range of pathologies. Most of these pathologies such as atherosclerosis, cancer, diabetes and Alzheimer disease evolve very progressively over several years. DNA methylation has been demonstrated to play a role in those diseases^{217,217,218,219}, it could thus be hypothesized that SASH1 expression is affected by DNA methylation in those pathologies. Consistent with this hypothesis, the SASH1 promoter region has also been shown to be differentially methylated in breast and hepatic cancerous tissue^{112,113}. Using data from venous thrombosis cohorts Martha and FL5, we investigated whether or not the SASH1 promoter region and the coding region were differentially methylated when stratifying the individuals by smoking habits. Our results did not show any significant differences in methylation levels, however, only 13 probes were present in the SASH1 between the preceding gene and SASH1 coding region that includes SASH1 promoter. Other methylation sites have been characterized in SASH1 promoter region by previous studies, therefore it could be envisioned that the CpG involved in SASH1 regulation were not present in the methylation array used for this analysis.

In our study, the microarray experiments were performed in SASH1 knockdown HAECs with two distinct siRNAs to determine the possible pathways SASH1 could be involved in. A total of 182 genes were differentially expressed in both SASH1 siRNA when compared to siCtrl. Ingenuity pathway analysis (IPA) revealed two statistically significant pathways for those genes: "Mitotic Roles of polo-Like Kinase" and "Cell cycle: G2/M DNA Damage Checkpoint Regulation". The polo-like kinases (Plks), are a family of enzymes involved in the regulation of mitosis (Figure 17). They are necessary for the complete activation of Cdc2 through Cdc25 which is required for the entry into mitosis. During mitosis, they are also required for the correct alignment of the chromosomes, as they are involved in the polymerization of tubulin. During metaphase and anaphase, Plks allow degradation of cyclin B and inactivation of p34 and therefore allow the exit from the M-phase²²⁰.

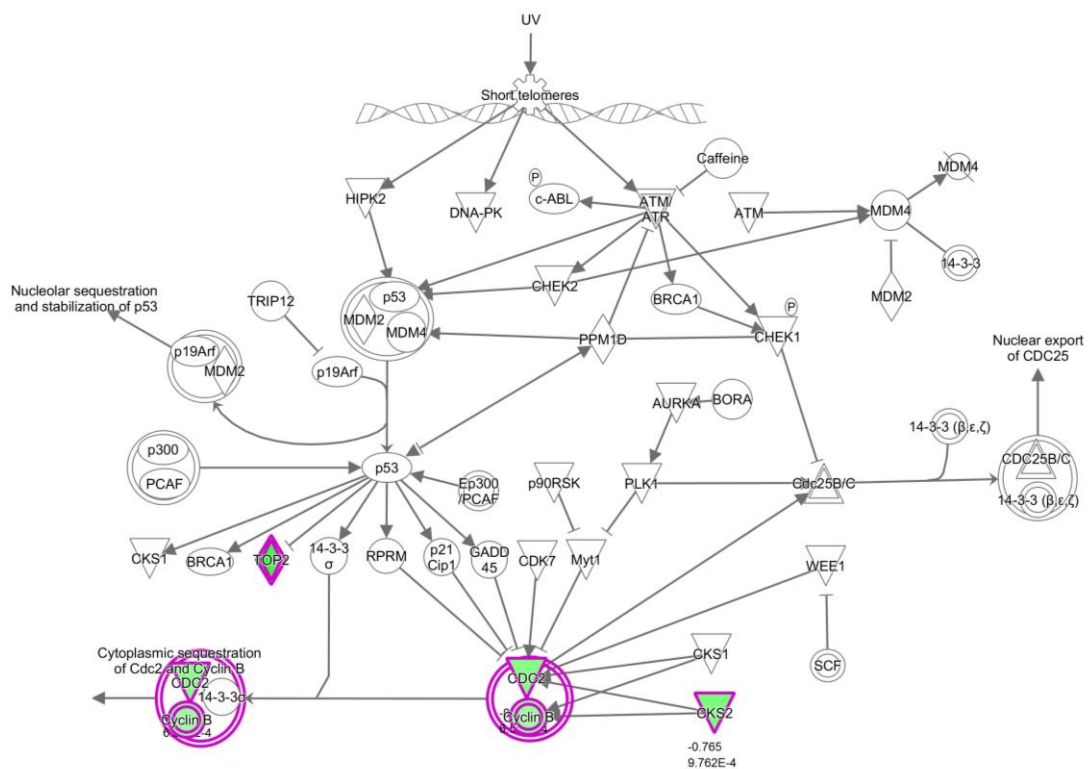


Figure 17: Members of the Cell cycle G2M DNA damage Checkpoint regulation pathway affected by SASH1 knockdown. Genes up-regulated by SASH1 knockdown in HAECs are represented in green.

The "Cell cycle: G2/M DNA Damage Checkpoint Regulation pathway" is a cell cycle checkpoint in eukaryotic cells, conserved from yeast to mammals (Figure 18). During this step, integrity of the DNA is checked and potential damages are repaired prior to cell mitosis. Mutations in proteins crucial for this pathways usually lead to cell death²²¹. Both pathways have strong ties to the cell cycle and thus are connected to atherosclerosis and cancer. The microarrays' results further strengthen the hypothesis for a role of SASH1 in the regulation of the cell cycle.

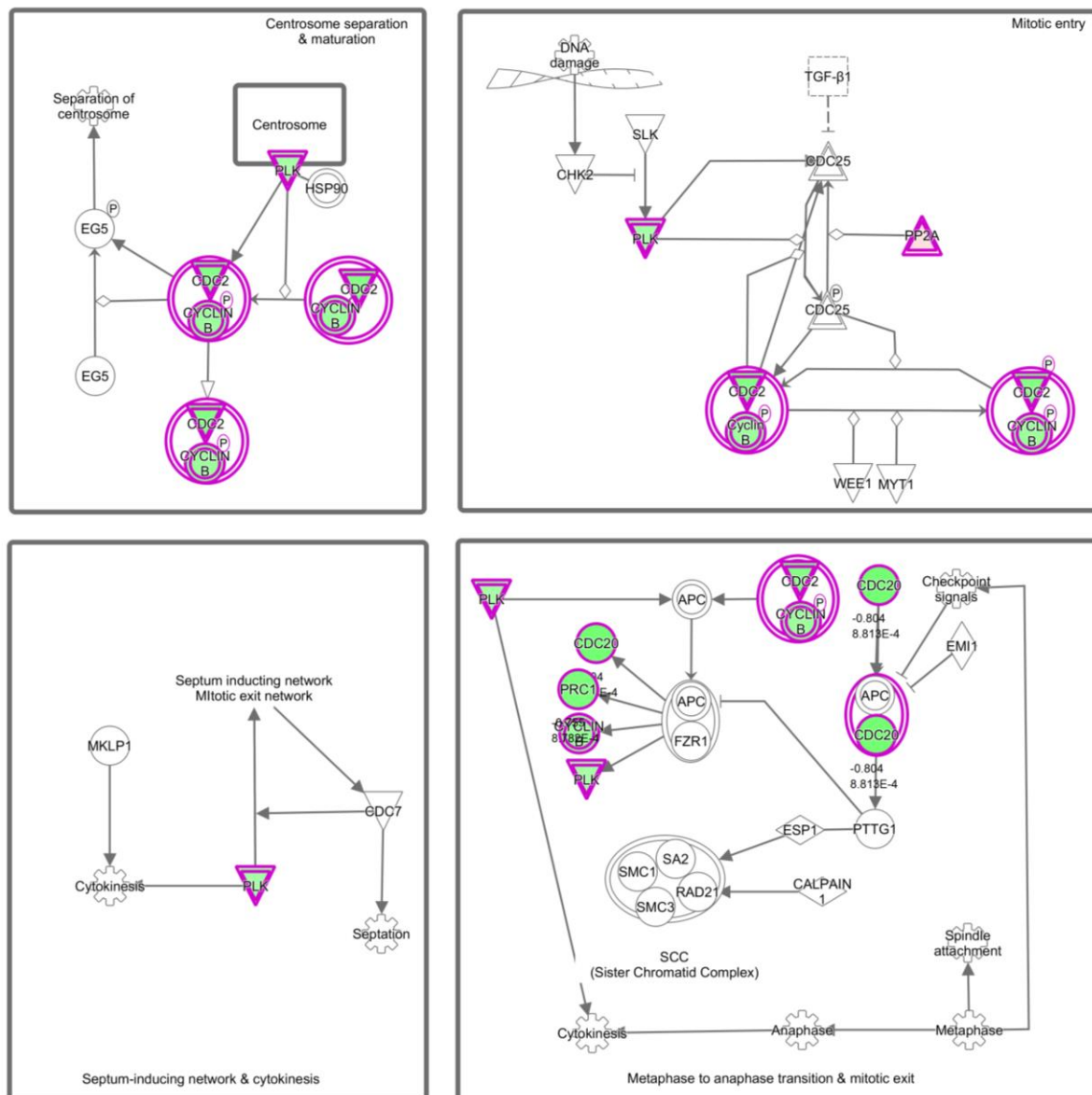


Figure 18: Members of the mitotic role of polo-like kinase pathway affected by SASH1 knockdown. Genes up-regulated by SASH1 knockdown in HAECs are represented in green.

Among the genes that were differentially expressed upon SASH1 knockdown, two caught particularly our attention, as they also were differentially expressed in HAECs exposed to low or high dose of CSC (microarrays): *TOP2A* and *CYP1A1*. Little is known in the literature about a potential link between *TOP2A* and cigarette smoke, as there is only one publication on the subject²²². In this study *TOP2A* is shown to be down-regulated following cigarette smoke condensate exposure, which corroborates with our microarrays' results. *TOP2A* is involved in the process of decatenation, in which its ability to cut double strand DNA allows the separation of the DNA of daughter cells after mitosis. It is over-expressed in most solid tumor cancer such as breast cancer^{223,224} and is therefore used as a biomarker. Since SASH1 knockdown increases proliferation in human aortic endothelial cells, it is not surprising to observe that *TOP2A* expression increases, however the question of whether it is a cause or a consequence requires more studies.

Nevertheless, *TOP2A* is known to be negatively regulated by P53²²⁵, thus it could be hypothesized that the increased *TOP2A* expression in SASH1 knockdown HAECs is a consequence of inhibition of p53. *CYP1A1*, another candidate was also altered by both SASH1 knockdown and cigarette smoke condensate stimulation, although in opposite direction. It is a member of the cytochrome P450 family that is composed of many enzymes involved in the metabolism of xenobiotics. Its expression is up-regulated in a p53 dependent fashion by the polycyclic aromatic hydrocarbons²²⁶ that are present in the particulate phase of the cigarette smoke. It is thus expected to see its expression increased in our experiments where HAECs were stimulated with the CSC (particulate phase diluted in DMSO). The fact that SASH1 is not affected by CSC itself but affects *CYP1A1* expression, suggests that SASH1 exerts its effects at the protein level, possibly through modulation of the p53 activation. Low expression of *CYP1A1* following SASH1 decrease, could be deleterious to the cell cycle, as it degrades many carcinogenic compounds, including the members of the aryl hydrocarbon family such as benzo(a)pyren in even more cancerogenic substances²²⁷. However, some discrepancy exists, as illustrated by the study performed by Uno et al²²⁸ who compared *CYP1A1* knockout to wild type mice and showed that *CYP1A1* promotes survival in mice exposed to orally applied benzo(a)pyren²²⁸. Increased SASH1 expression in smokers could also favor an increase in *CYP1A1* expression, thereby postulating the induction of a defense mechanism against the cigarette smoke toxic compounds.

Among the transcription factors which were predicted to be activated by SASH1 knockdown, *FOXM1*, which is particularly relevant to the vascular field, is negatively regulated by p53²²⁹. *FOXM1* controls positively mitosis and migration, and thus could also be activated following p53 inhibition by SASH1 knockdown. While there are no studies linking directly *FOXM1* to atherosclerosis, its implication in tumorigenesis and poor prognosis has been demonstrated²³⁰. Interestingly, *FOXM1* expression is up-regulated in breast tumor²³¹, in which SASH1 is down-regulated¹⁰¹. In the vasculature, *FOXM1* was shown to promote endothelial repair in mice after inflammatory lung injury²³². In this model, endothelial bone marrow progenitors promote healing of the wound, however in *FOXM1* deficient mice, the protective effect of endothelial bone marrow progenitor cells after LPS induced vascular injury was abrogated. The same study also showed that the human endothelial cell line HMEC required *FOXM1* for cell cycle progression. Moreover, *FOXM1* expression positively regulates the expression of beta-catenin, an adhesion protein required for endothelium regeneration and re-annealing of adherent junction after vascular repair²³³. The restoration of the beta-catenin expression in *FOXM1* deficient mice, rescued the defective endothelial barrier phenotype, indicating that *FOXM1* exerts its vascular effects mainly by controlling the expression of beta-catenin. While beta-catenin was not among the 182 genes affected by SASH1 knockdown in our experiment, *CTNNBIP1* (also known as *ICAT*), a protein which binds to *CTNNB1* was down-regulated. *CTNNBIP1* inhibits the Wnt proliferation pathway, by binding to *CTNNB1* and inhibiting its interaction with the downstream transcription factors of the Wnt pathway (TCF family)²³⁴. Thus decrease in *CTNNBIP1* could also explain partially the increased proliferation observed in SASH1 knockdown HAECs.

We also searched for SASH1 partners using a yeast two hybrid screening and obtained a list of 29 candidate proteins. We tested in HAECs 6 of those proteins, IkbkB, CTNNA1, OAS2, C1D, AFF1 and Rac1 by performing co-immunoprecipitation (co-IP) of either SASH1 or the tested partner proteins followed by western blot identification. None of those tested candidates could be confirmed because of weak antibody efficiency for immunoprecipitation. As only the SASH1 co-IP gave reliable results we thus decided to go for a non-targeted detection method of the co-IP products by mass spectrometry instead of western blotting. By this procedure, we uncovered two proteins: ribosomal protein 25 (RPS25) and tropomyosin 1 (TPM1) as possible partners of SASH1. RPS25 is a nucleolar protein²³⁵ which is involved in p53 regulation by directly binding MDM2 which inhibits its E3 ligase activity, preventing the degradation of p53²³⁶. In addition, RPS25 expression is regulated by p53 through a negative loop, as p53 directly binds to RPS25 promoter²³⁶. Thus, SASH1 directly bind to RPS25, and for this reason SASH1 knockdown is likely to inhibit p53 activity. It may be thus suggested that SASH1 promotes p53 activation by promoting inhibition of MDM2 via RPS25. Apart from the cytoplasm, RPS25 also localizes in the nucleus, and it is thus one the best candidates to explore the nuclear function of SASH1.

The second partner of SASH1 found in our study was TPM1. This protein is involved in the polymerization of the actin cytoskeleton and protects endothelial cell-cell junctions from the oxidative stress triggered by aqueous cigarette smoke extract²³⁷. TPM1 phosphorylation state regulates endothelial permeability to cancerous cells²³⁸, and TPM1 was also demonstrated to be a tumor suppressor in various cancers, including breast cancer^{239,240}, where it resensitizes cancer cells to anoikis (programmed cell death triggered by detachment from the extracellular matrix). As SASH1 binds TPM1, it could be envisioned that it preserves endothelial integrity through the stabilization of the actin cytoskeleton. Consistent with this hypothesis, TPM1 inhibits the activity of the Arp2/3 complex²⁴¹ which is regulated by cortactin (CTTN) another SASH1 partner¹⁰⁴. This study of the SASH1 partner suggests that SASH1 could exert its vascular effects through the control of cell proliferation via both RPS25/P53/FOXM1 and TPM1 dependent actin cytoskeleton polymerization pathways (Figure 19).

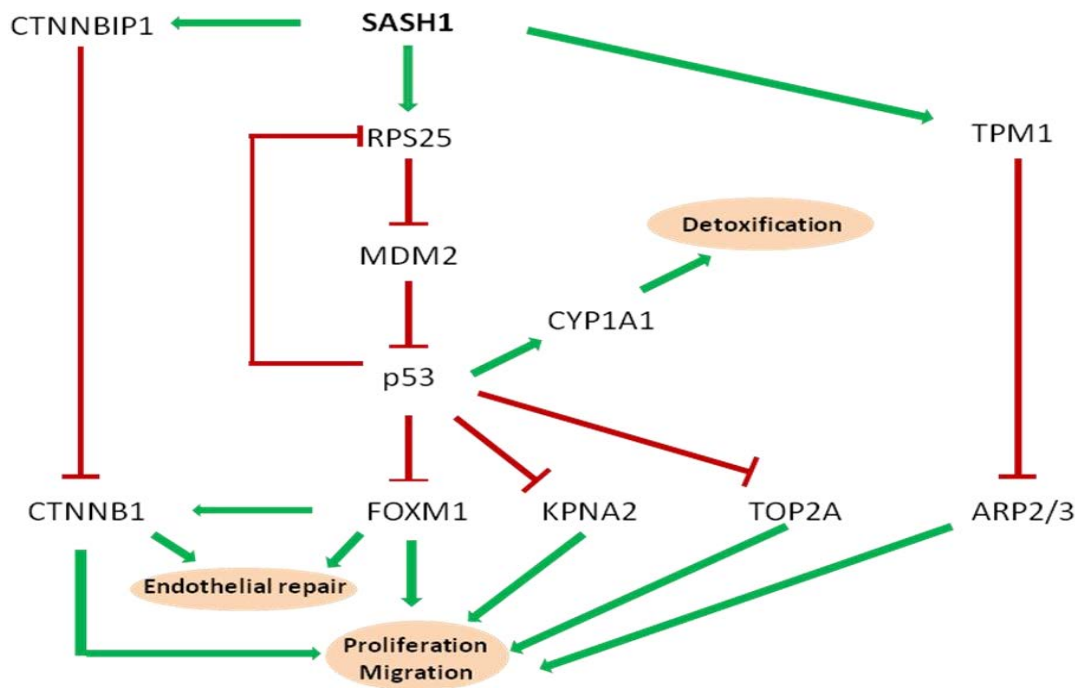


Figure 19. Graphical representation of potential pathways predicted from our co-immunoprecipitation and microarrays experiments performed on SASH1 knockdown HAECs. Green arrows represents activation, red lines represent inhibition, cellular processes are circled in light orange.

After focusing on the molecular aspects of SASH1 knockdown, we investigated the effect of SASH1 on HAECs proliferation and migration. We found that SASH1 knockdown increase both proliferation and migration in HAECs. This result is in accordance with the literature, as SASH1 knockdown increases proliferation and migration in various human cancer cell lines^{104,109,108}, and a SASH1 variant (c.1849G>A; p.Glu617Lys) increases proliferation of human fibroblasts¹¹⁵. In accordance with our proliferation results, CCND1 and CCND3 were found to be up-regulated at the protein level following SASH1 knockdown. While CCND1 had already been shown to be linked to SASH1 expression in the literature^{108,107,110}, this is the first evidence for an impact of SASH1 on cyclin D3 expression. Both cyclins are known to be co-expressed²⁴² and peaks at the G1/S phase²⁴³. Therefore, it is a possibility that the increase of CCND1 and CCND3 expression upon SASH1 knockdown is an indirect effect of SASH1 blocking the cells in the G1 phase rather than a direct effect of SASH1 specifically targeting the transcription of these two cyclins. Furthermore, CCND1 and CCND3 were shown to be up-regulated in human vascular smooth muscle cells, implicated in atherosclerosis upon plaque formation, when stimulated by oxLDL²⁴⁴. Therefore, it may be expected that the expression of those cyclins will be increased in the plaques. The increase of SASH1 expression in the carotids from smokers may lead to a decrease of CCND1 and CCND3 which would point out to an atheroprotective role of SASH1 by inhibiting the proliferation of smooth muscle cells within the plaque. However, this could also have deleterious effects, as smooth muscle cells participate in the formation of fibrous cap and for this reason are involved in plaque stability.

Finally in this study, we showed that SASH1 knockdown increased *in vitro* angiogenesis of HAECs. This result was expected as SASH1 affected both migration and proliferation. The contribution of those two processes to the increase of tubular formation following SASH1 knockdown could be further characterized by performing experiments using both Matrigel and mitomycin-C. Implication of SASH1 in proliferation, migration and angiogenesis of HAECs suggests that SASH1 could play a role in endothelial activation, the first step toward plaque formation, therefore participating in the early onset of the disease. Plaque neo-vascularization, a feature of the late stage of the plaque could also be affected by SASH1 expression. As the newly formed vessels are mostly composed of permeable capillaries, plaque neovascularization is believed to play an important role in intra plaque hemorrhage and therefore in plaque stability^{245,246}. Thus SASH1 could also play a role in the late stage of atherosclerosis by inhibiting intra plaque neo-vascularization and therefore increasing plaque stability.

Taken together our results showed that SASH1 is differentially expressed in smokers and could have an atheroprotective role, inhibiting cellular proliferation and migration of the cells from the vascular wall by regulating p53 activity and the actin cytoskeleton dynamics. However SASH1 also has the potential to impair the healing of vascular injuries by inhibiting FOXM1 in a p53 dependent manner, which would also give it an atherogenic property. Thus, while SASH1 clearly establish itself as a link between smoking and atherosclerosis, its effects remain to be further characterize before it can be decided whether SASH1 is a defense mechanism against cigarette smoke or mediates cigarette smoke deleterious effects on the vascular wall.

References

1. Figueroa XF, Duling BR. Gap Junctions in the Control of Vascular Function. *Antioxid Redox Signal*. 2009;11(2):251-266. doi:10.1089/ars.2008.2117.
2. Stary HC, Chandler AB, Glagov S, 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. *Circulation*. 1994;89(5):2462-2478.
3. Stary HC, Chandler AB, Dinsmore RE, 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. *Circulation*. 1995;92(5):1355-1374. doi:10.1161/01.CIR.92.5.1355.
4. Drexler H. Factors involved in the maintenance of endothelial function. *Am J Cardiol*. 1998;82(10A):3S - 4S.
5. Lüscher TF, Barton M. Biology of the endothelium. *Clin Cardiol*. 1997;20(11 Suppl 2):II - 3-10.
6. Endres M, Laufs U, Huang Z, et al. Stroke protection by 3-hydroxy-3-methylglutaryl (HMG)-CoA reductase inhibitors mediated by endothelial nitric oxide synthase. *Proc Natl Acad Sci U S A*. 1998;95(15):8880-8885.
7. Hogg N, Kalyanaraman B, Joseph J, Struck A, Parthasarathy S. Inhibition of low-density lipoprotein oxidation by nitric oxide Potential role in atherogenesis. *FEBS Lett*. 1993;334(2):170-174. doi:10.1016/0014-5793(93)81706-6.
8. Kinlay S, Behrendt D, Wainstein M, et al. Role of endothelin-1 in the active constriction of human atherosclerotic coronary arteries. *Circulation*. 2001;104(10):1114-1118.
9. Saye JA, Singer HA, Peach MJ. Role of endothelium in conversion of angiotensin I to angiotensin II in rabbit aorta. *Hypertension*. 1984;6(2 Pt 1):216-221.
10. Yaghini FA, Song CY, Lavrentyev EN, et al. Angiotensin II-induced vascular smooth muscle cell migration and growth are mediated by cytochrome P450 1B1-dependent superoxide generation. *Hypertension*. 2010;55(6):1461-1467. doi:10.1161/HYPERTENSIONAHA.110.150029.
11. Victor VM, Rocha M, Solá E, Bañuls C, Garcia-Malpartida K, Hernández-Mijares A. Oxidative stress, endothelial dysfunction and atherosclerosis. *Curr Pharm Des*. 2009;15(26):2988-3002.
12. Davies MJ, Gordon JL, Gearing AJ, et al. The expression of the adhesion molecules ICAM-1, VCAM-1, PECAM, and E-selectin in human atherosclerosis. *J Pathol*. 1993;171(3):223-229. doi:10.1002/path.1711710311.
13. Mestas J, Ley K. Monocyte-Endothelial Cell Interactions in the Development of Atherosclerosis. *Trends Cardiovasc Med*. 2008;18(6):228-232. doi:10.1016/j.tcm.2008.11.004.
14. Deanfield JE, Halcox JP, Rabelink TJ. Endothelial Function and Dysfunction Testing and Clinical Relevance. *Circulation*. 2007;115(10):1285-1295. doi:10.1161/CIRCULATIONAHA.106.652859.
15. Reddy KG, Nair RN, Sheehan HM, Hodgson JM. Evidence that selective endothelial dysfunction may occur in the absence of angiographic or ultrasound atherosclerosis in patients with risk factors for atherosclerosis. *J Am Coll Cardiol*. 1994;23(4):833-843.

16. Celermajer DS, Sorensen KE, Bull C, Robinson J, Deanfield JE. Endothelium-dependent dilation in the systemic arteries of asymptomatic subjects relates to coronary risk factors and their interaction. *J Am Coll Cardiol.* 1994;24(6):1468-1474.
17. Falk E. Pathogenesis of atherosclerosis. *J Am Coll Cardiol.* 2006;47(8 Suppl):C7-C12. doi:10.1016/j.jacc.2005.09.068.
18. Ross R. Atherosclerosis — An Inflammatory Disease. *N Engl J Med.* 1999;340(2):115-126. doi:10.1056/NEJM199901143400207.
19. O'Brien KD, Olin KL, Alpers CE, et al. Comparison of apolipoprotein and proteoglycan deposits in human coronary atherosclerotic plaques: colocalization of biglycan with apolipoproteins. *Circulation.* 1998;98(6):519-527.
20. Ruan XZ, Moorhead JF, Tao JL, et al. Mechanisms of dysregulation of low-density lipoprotein receptor expression in vascular smooth muscle cells by inflammatory cytokines. *Arterioscler Thromb Vasc Biol.* 2006;26(5):1150-1155. doi:10.1161/01.ATV.0000217957.93135.c2.
21. Braun M, Pietsch P, Schrör K, Baumann G, Felix SB. Cellular adhesion molecules on vascular smooth muscle cells. *Cardiovasc Res.* 1999;41(2):395-401.
22. Cai Q, Lanting L, Natarajan R. Interaction of monocytes with vascular smooth muscle cells regulates monocyte survival and differentiation through distinct pathways. *Arterioscler Thromb Vasc Biol.* 2004;24(12):2263-2270. doi:10.1161/01.ATV.0000146552.16943.5e.
23. O'Brien KD, Allen MD, McDonald TO, et al. Vascular cell adhesion molecule-1 is expressed in human coronary atherosclerotic plaques. Implications for the mode of progression of advanced coronary atherosclerosis. *J Clin Invest.* 1993;92(2):945-951. doi:10.1172/JCI116670.
24. Endres M, Laufs U, Merz H, Kaps M. Focal expression of intercellular adhesion molecule-1 in the human carotid bifurcation. *Stroke J Cereb Circ.* 1997;28(1):77-82.
25. Jang Y, Lincoff AM, Plow EF, Topol EJ. Cell adhesion molecules in coronary artery disease. *J Am Coll Cardiol.* 1994;24(7):1591-1601.
26. Amento EP, Ehsani N, Palmer H, Libby P. Cytokines and growth factors positively and negatively regulate interstitial collagen gene expression in human vascular smooth muscle cells. *Arterioscler Thromb J Vasc Biol Am Heart Assoc.* 1991;11(5):1223-1230.
27. Braun M, Pietsch P, Schrör K, Baumann G, Felix SB. Cellular adhesion molecules on vascular smooth muscle cells. *Cardiovasc Res.* 1999;41(2):395-401.
28. Tabas I. Macrophage death and defective inflammation resolution in atherosclerosis. *Nat Rev Immunol.* 2010;10(1):36-46. doi:10.1038/nri2675.
29. Shibata N, Glass CK. Regulation of macrophage function in inflammation and atherosclerosis. *J Lipid Res.* 2009;50(Supplement):S277-S281. doi:10.1194/jlr.R800063-JLR200.
30. Wang X, Collins HL, Ranalletta M, et al. Macrophage ABCA1 and ABCG1, but not SR-BI, promote macrophage reverse cholesterol transport in vivo. *J Clin Invest.* 2007;117(8):2216-2224. doi:10.1172/JCI32057.

31. Li Y, Ge M, Ciani L, et al. Enrichment of endoplasmic reticulum with cholesterol inhibits sarcoplasmic-endoplasmic reticulum calcium ATPase-2b activity in parallel with increased order of membrane lipids: implications for depletion of endoplasmic reticulum calcium stores and apoptosis in cholesterol-loaded macrophages. *J Biol Chem*. 2004;279(35):37030-37039. doi:10.1074/jbc.M405195200.
32. Feng B, Yao PM, Li Y, et al. The endoplasmic reticulum is the site of cholesterol-induced cytotoxicity in macrophages. *Nat Cell Biol*. 2003;5(9):781-792. doi:10.1038/ncb1035.
33. Seimon T, Tabas I. Mechanisms and consequences of macrophage apoptosis in atherosclerosis. *J Lipid Res*. 2009;50 Suppl:S382-S387. doi:10.1194/jlr.R800032-JLR200.
34. Brochériou I, Maouche S, Durand H, et al. Antagonistic regulation of macrophage phenotype by M-CSF and GM-CSF: implication in atherosclerosis. *Atherosclerosis*. 2011;214(2):316-324. doi:10.1016/j.atherosclerosis.2010.11.023.
35. Martinez FO, Gordon S. The M1 and M2 paradigm of macrophage activation: time for reassessment. *F1000Prime Rep*. 2014;6. doi:10.12703/P6-13.
36. Randolph GJ. The fate of monocytes in atherosclerosis. *J Thromb Haemost JTH*. 2009;7 Suppl 1:28-30. doi:10.1111/j.1538-7836.2009.03423.x.
37. Griffin GK, Lichtman AH. Why don't macrophages leave atherosclerotic lesions? *Circ Res*. 2012;110(10):1273-1275. doi:10.1161/CIRCRESAHA.112.268839.
38. Zeller T, Wild P, Szymczak S, et al. Genetics and Beyond – The Transcriptome of Human Monocytes and Disease Susceptibility. *PLoS ONE*. 2010;5(5):e10693. doi:10.1371/journal.pone.0010693.
39. Puranik R, Celermajer DS. Smoking and endothelial function. *Prog Cardiovasc Dis*. 2003;45(6):443-458. doi:10.1053/pcad.2003.YPCAD13.
40. Heitzer T, Ylä-Herttuala S, Luoma J, et al. Cigarette smoking potentiates endothelial dysfunction of forearm resistance vessels in patients with hypercholesterolemia. Role of oxidized LDL. *Circulation*. 1996;93(7):1346-1353.
41. Shen Y, Rattan V, Sultana C, Kalra VK. Cigarette smoke condensate-induced adhesion molecule expression and transendothelial migration of monocytes. *Am J Physiol*. 1996;270(5 Pt 2):H1624-H1633.
42. Fusegawa Y, Goto S, Handa S, Kawada T, Ando Y. Platelet spontaneous aggregation in platelet-rich plasma is increased in habitual smokers. *Thromb Res*. 1999;93(6):271-278.
43. Rosenberg L, Kaufman DW, Helmrich SP, Shapiro S. The risk of myocardial infarction after quitting smoking in men under 55 years of age. *N Engl J Med*. 1985;313(24):1511-1514. doi:10.1056/NEJM198512123132404.
44. Czernin J, Waldherr C. Cigarette smoking and coronary blood flow. *Prog Cardiovasc Dis*. 2003;45(5):395-404. doi:10.1053/pcad.2003.00104.
45. Smith CJ, Fischer TH. Particulate and vapor phase constituents of cigarette mainstream smoke and risk of myocardial infarction. *Atherosclerosis*. 2001;158(2):257-267.
46. Pryor WA, Stone K, Zang LY, Bermúdez E. Fractionation of aqueous cigarette tar extracts: fractions that contain the tar radical cause DNA damage. *Chem Res Toxicol*. 1998;11(5):441-448. doi:10.1021/tx970159y.

47. Pryor WA, Stone K. Oxidants in cigarette smoke. Radicals, hydrogen peroxide, peroxyxynitrate, and peroxyxynitrite. *Ann N Y Acad Sci.* 1993;686:12-27; discussion 27-28.
48. Taylor AE, Johnson DC, Kazemi H. Environmental tobacco smoke and cardiovascular disease. A position paper from the Council on Cardiopulmonary and Critical Care, American Heart Association. *Circulation.* 1992;86(2):699-702.
49. Ohtani H, Tamaki N, Mohiuddin IH, et al. [Minimal redistribution of thallium-201 representing reversible ischemia after coronary bypass surgery: value of quantitative analysis of exercise thallium-201 SPECT]. *J Cardiol.* 1991;21(4):835-846.
50. Law MR, Morris JK, Wald NJ. Environmental tobacco smoke exposure and ischaemic heart disease: an evaluation of the evidence. *BMJ.* 1997;315(7114):973-980.
51. Howard G, Burke GL, Szklo M, et al. Active and passive smoking are associated with increased carotid wall thickness. The Atherosclerosis Risk in Communities Study. *Arch Intern Med.* 1994;154(11):1277-1282.
52. Ijzerman RG, Serne EH, van Weissenbruch MM, de Jongh RT, Stehouwer CDA. Cigarette smoking is associated with an acute impairment of microvascular function in humans. *Clin Sci Lond Engl 1979.* 2003;104(3):247-252. doi:10.1042/CS20020318.
53. Celermajer DS, Adams MR, Clarkson P, et al. Passive smoking and impaired endothelium-dependent arterial dilatation in healthy young adults. *N Engl J Med.* 1996;334(3):150-154. doi:10.1056/NEJM199601183340303.
54. Mayhan WG, Sharpe GM. Effect of cigarette smoke extract on arteriolar dilatation in vivo. *J Appl Physiol Bethesda Md 1985.* 1996;81(5):1996-2003.
55. Craig WY, Palomaki GE, Haddow JE. Cigarette smoking and serum lipid and lipoprotein concentrations: an analysis of published data. *BMJ.* 1989;298(6676):784-788.
56. Smith CJ, Fischer TH. Particulate and vapor phase constituents of cigarette mainstream smoke and risk of myocardial infarction. *Atherosclerosis.* 2001;158(2):257-267.
57. Tracy RP, Psaty BM, Macy E, et al. Lifetime smoking exposure affects the association of C-reactive protein with cardiovascular disease risk factors and subclinical disease in healthy elderly subjects. *Arterioscler Thromb Vasc Biol.* 1997;17(10):2167-2176.
58. Tappia PS, Troughton KL, Langley-Evans SC, Grimble RF. Cigarette smoking influences cytokine production and antioxidant defences. *Clin Sci Lond Engl 1979.* 1995;88(4):485-489.
59. Libby P, Hansson GK. Inflammation and Immunity in Diseases of the Arterial Tree Players and Layers. *Circ Res.* 2015;116(2):307-311. doi:10.1161/CIRCRESAHA.116.301313.
60. Mazzone A, Cusa C, Mazzucchelli I, et al. Cigarette smoking and hypertension influence nitric oxide release and plasma levels of adhesion molecules. *Clin Chem Lab Med CCLM FESCC.* 2001;39(9):822-826. doi:10.1515/CCLM.2001.136.
61. Bermudez EA, Rifai N, Buring JE, Manson JE, Ridker PM. Relation between markers of systemic vascular inflammation and smoking in women. *Am J Cardiol.* 2002;89(9):1117-1119.

62. Kalra VK, Ying Y, Deemer K, Natarajan R, Nadler JL, Coates TD. Mechanism of cigarette smoke condensate induced adhesion of human monocytes to cultured endothelial cells. *J Cell Physiol.* 1994;160(1):154-162. doi:10.1002/jcp.1041600118.
63. Shen Y, Rattan V, Sultana C, Kalra VK. Cigarette smoke condensate-induced adhesion molecule expression and transendothelial migration of monocytes. *Am J Physiol.* 1996;270(5 Pt 2):H1624-H1633.
64. Weber C, Erl W, Weber K, Weber PC. Increased adhesiveness of isolated monocytes to endothelium is prevented by vitamin C intake in smokers. *Circulation.* 1996;93(8):1488-1492.
65. Mendall MA, Patel P, Asante M, et al. Relation of serum cytokine concentrations to cardiovascular risk factors and coronary heart disease. *Heart Br Card Soc.* 1997;78(3):273-277.
66. Aicher A, Heeschen C, Mohaupt M, Cooke JP, Zeiher AM, Dimmeler S. Nicotine strongly activates dendritic cell-mediated adaptive immunity: potential role for progression of atherosclerotic lesions. *Circulation.* 2003;107(4):604-611.
67. Yong T, Zheng MQ, Linthicum DS. Nicotine induces leukocyte rolling and adhesion in the cerebral microcirculation of the mouse. *J Neuroimmunol.* 1997;80(1-2):158-164.
68. Napoli C, Ignarro LJ. Nitric oxide and atherosclerosis. *Nitric Oxide Biol Chem Off J Nitric Oxide Soc.* 2001;5(2):88-97. doi:10.1006/niox.2001.0337.
69. Mayhan WG, Patel KP. Effect of nicotine on endothelium-dependent arteriolar dilatation in vivo. *Am J Physiol.* 1997;272(5 Pt 2):H2337-H2342.
70. Barua RS, Ambrose JA, Srivastava S, DeVoe MC, Eales-Reynolds L-J. Reactive oxygen species are involved in smoking-induced dysfunction of nitric oxide biosynthesis and upregulation of endothelial nitric oxide synthase: an in vitro demonstration in human coronary artery endothelial cells. *Circulation.* 2003;107(18):2342-2347. doi:10.1161/01.CIR.0000066691.52789.BE.
71. Nordenhem A, Leander K, Hallqvist J, de Faire U, Sten-Linder M, Wiman B. The complex between tPA and PAI-1: risk factor for myocardial infarction as studied in the SHEEP project. *Thromb Res.* 2005;116(3):223-232. doi:10.1016/j.thromres.2004.12.007.
72. Heitzer T, Just H, Münzel T. Antioxidant vitamin C improves endothelial dysfunction in chronic smokers. *Circulation.* 1996;94(1):6-9.
73. Howard G, Burke GL, Szklo M, et al. Active and passive smoking are associated with increased carotid wall thickness. The Atherosclerosis Risk in Communities Study. *Arch Intern Med.* 1994;154(11):1277-1282.
74. Krupski WC, Olive GC, Weber CA, Rapp JH. Comparative effects of hypertension and nicotine on injury-induced myointimal thickening. *Surgery.* 1987;102(2):409-415.
75. Heeschen C, Jang JJ, Weis M, et al. Nicotine stimulates angiogenesis and promotes tumor growth and atherosclerosis. *Nat Med.* 2001;7(7):833-839. doi:10.1038/89961.
76. Stubbe I, Eskilsson J, Nilsson-Ehle P. High-density lipoprotein concentrations increase after stopping smoking. *Br Med J Clin Res Ed.* 1982;284(6328):1511-1513.
77. Reaven G, Tsao PS. Insulin resistance and compensatory hyperinsulinemia: the key player between cigarette smoking and cardiovascular disease? *J Am Coll Cardiol.* 2003;41(6):1044-1047.

78. Axelsson T, Jansson PA, Smith U, Eliasson B. Nicotine infusion acutely impairs insulin sensitivity in type 2 diabetic patients but not in healthy subjects. *J Intern Med.* 2001;249(6):539-544.
79. Harats D, Ben-Naim M, Dabach Y, Hollander G, Stein O, Stein Y. Cigarette smoking renders LDL susceptible to peroxidative modification and enhanced metabolism by macrophages. *Atherosclerosis.* 1989;79(2-3):245-252.
80. Lehr HA, Weyrich AS, Saetzler RK, et al. Vitamin C blocks inflammatory platelet-activating factor mimetics created by cigarette smoking. *J Clin Invest.* 1997;99(10):2358-2364. doi:10.1172/JCI119417.
81. Adams MR, Jessup W, Celermajer DS. Cigarette smoking is associated with increased human monocyte adhesion to endothelial cells: reversibility with oral L-arginine but not vitamin C. *J Am Coll Cardiol.* 1997;29(3):491-497.
82. Kannel WB, D'Agostino RB, Belanger AJ. Fibrinogen, cigarette smoking, and risk of cardiovascular disease: insights from the Framingham Study. *Am Heart J.* 1987;113(4):1006-1010.
83. Smith FB, Lee AJ, Fowkes FG, Price JF, Rumley A, Lowe GD. Hemostatic factors as predictors of ischemic heart disease and stroke in the Edinburgh Artery Study. *Arterioscler Thromb Vasc Biol.* 1997;17(11):3321-3325.
84. Eliasson M, Asplund K, Evrin PE, Lundblad D. Relationship of cigarette smoking and snuff dipping to plasma fibrinogen, fibrinolytic variables and serum insulin. The Northern Sweden MONICA Study. *Atherosclerosis.* 1995;113(1):41-53.
85. Zidovetzki R, Chen P, Fisher M, Hofman FM, Faraci FM. Nicotine increases plasminogen activator inhibitor-1 production by human brain endothelial cells via protein kinase C-associated pathway. *Stroke J Cereb Circ.* 1999;30(3):651-655.
86. Sambola A, Osende J, Hathcock J, et al. Role of risk factors in the modulation of tissue factor activity and blood thrombogenicity. *Circulation.* 2003;107(7):973-977.
87. Matetzky S, Tani S, Kangavari S, et al. Smoking increases tissue factor expression in atherosclerotic plaques: implications for plaque thrombogenicity. *Circulation.* 2000;102(6):602-604.
88. Rival J, Riddle JM, Stein PD. Effects of chronic smoking on platelet function. *Thromb Res.* 1987;45(1):75-85.
89. Blache D. Involvement of hydrogen and lipid peroxides in acute tobacco smoking-induced platelet hyperactivity. *Am J Physiol.* 1995;268(2 Pt 2):H679-H685.
90. Wang XL, Greco M, Sim AS, Duarte N, Wang J, Wilcken DEL. Effect of CYP1A1 MspI polymorphism on cigarette smoking related coronary artery disease and diabetes. *Atherosclerosis.* 2002;162(2):391-397.
91. Wang J, Dudley D, Wang XL. Haplotype-specific effects on endothelial NO synthase promoter efficiency: modifiable by cigarette smoking. *Arterioscler Thromb Vasc Biol.* 2002;22(5):e1-e4.
92. Wang XL, Sim AS, Badenhop RF, McCredie RM, Wilcken DE. A smoking-dependent risk of coronary artery disease associated with a polymorphism of the endothelial nitric oxide synthase gene. *Nat Med.* 1996;2(1):41-45.
93. Holm J, Hillarp A, Zöller B, Erhardt L, Berntorp E, Dahlbäck B. Factor V Q506 (resistance to activated protein C) and prognosis after acute coronary syndrome. *Thromb Haemost.* 1999;81(6):857-860.

94. Ma Q, Lu AYH. CYP1A Induction and Human Risk Assessment: An Evolving Tale of in Vitro and in Vivo Studies. *Drug Metab Dispos*. 2007;35(7):1009-1016. doi:10.1124/dmd.107.015826.
95. Anto RJ, Mukhopadhyay A, Shishodia S, Gairola CG, Aggarwal BB. Cigarette smoke condensate activates nuclear transcription factor- κ B through phosphorylation and degradation of I κ B α : correlation with induction of cyclooxygenase-2. *Carcinogenesis*. 2002;23(9):1511-1518. doi:10.1093/carcin/23.9.1511.
96. Wirtz HR, Schmidt M. Acute influence of cigarette smoke on secretion of pulmonary surfactant in rat alveolar type II cells in culture. *Eur Respir J*. 1996;9(1):24-32.
97. Thorne D, Adamson J. A review of in vitro cigarette smoke exposure systems. *Exp Toxicol Pathol*. 2013;65(7-8):1183-1193. doi:10.1016/j.etp.2013.06.001.
98. Palozza P, Serini S, Trombino S, Lauriola L, Ranelletti FO, Calviello G. Dual role of β -carotene in combination with cigarette smoke aqueous extract on the formation of mutagenic lipid peroxidation products in lung membranes: dependence on pO₂. *Carcinogenesis*. 2006;27(12):2383-2391. doi:10.1093/carcin/bgl074.
99. Totlandsdal AI, Refsnes M, Skomedal T, Osnes J-B, Schwarze PE, Låg M. Particle-induced cytokine responses in cardiac cell cultures--the effect of particles versus soluble mediators released by particle-exposed lung cells. *Toxicol Sci Off J Soc Toxicol*. 2008;106(1):233-241. doi:10.1093/toxsci/kfn162.
100. Vozzi F, Heinrich J-M, Bader A, Ahluwalia AD. Connected culture of murine hepatocytes and HUVEC in a multicompartmental bioreactor. *Tissue Eng Part A*. 2009;15(6):1291-1299. doi:10.1089/ten.tea.2008.0066.
101. Zeller C, Hinzmann B, Seitz S, et al. SASH1: a candidate tumor suppressor gene on chromosome 6q24.3 is downregulated in breast cancer. *Oncogene*. 2003;22(19):2972-2983. doi:10.1038/sj.onc.1206474.
102. Pichon X, Wilson LA, Stoneley M, et al. RNA binding protein/RNA element interactions and the control of translation. *Curr Protein Pept Sci*. 2012;13(4):294-304.
103. Barrett LW, Fletcher S, Wilton SD. Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. *Cell Mol Life Sci*. 2012;69(21):3613-3634. doi:10.1007/s00018-012-0990-9.
104. Martini M, Gnann A, Scheikl D, Holzmann B, Janssen K-P. The candidate tumor suppressor SASH1 interacts with the actin cytoskeleton and stimulates cell-matrix adhesion. *Int J Biochem Cell Biol*. 2011;43(11):1630-1640. doi:10.1016/j.biocel.2011.07.012.
105. Yang F, Waters KM, Miller JH, et al. Phosphoproteomics Profiling of Human Skin Fibroblast Cells Reveals Pathways and Proteins Affected by Low Doses of Ionizing Radiation. *PLoS ONE*. 2010;5(11). doi:10.1371/journal.pone.0014152.
106. Rimkus C, Martini M, Friederichs J, et al. Prognostic significance of downregulated expression of the candidate tumour suppressor gene SASH1 in colon cancer. *Br J Cancer*. 2006;95(10):1419-1423. doi:10.1038/sj.bjc.6603452.
107. Yang L, Liu M, Gu Z, Chen J, Yan Y, Li J. Overexpression of SASH1 related to the decreased invasion ability of human glioma U251 cells. *Tumour Biol J Int Soc Oncodevelopmental Biol Med*. 2012;33(6):2255-2263. doi:10.1007/s13277-012-0487-z.

108. Meng Q, Zheng M, Liu H, et al. SASH1 regulates proliferation, apoptosis, and invasion of osteosarcoma cell. *Mol Cell Biochem.* 2013;373(1-2):201-210. doi:10.1007/s11010-012-1491-8.
109. Lin S, Zhang J, Xu J, et al. Effects of SASH1 on melanoma cell proliferation and apoptosis in vitro. *Mol Med Rep.* 2012;6(6):1243-1248. doi:10.3892/mmr.2012.1099.
110. Chen E, Chen Y, Dong L, Zhang J. Effects of SASH1 on lung cancer cell proliferation, apoptosis, and invasion in vitro. *Tumour Biol J Int Soc Oncodevelopmental Biol Med.* 2012;33(5):1393-1401. doi:10.1007/s13277-012-0387-2.
111. Nitsche U, Rosenberg R, Balmert A, et al. Integrative marker analysis allows risk assessment for metastasis in stage II colon cancer. *Ann Surg.* 2012;256(5):763-771; discussion 771. doi:10.1097/SLA.0b013e318272de87.
112. Sheyu L, Hui L, Junyu Z, et al. Promoter methylation assay of SASH1 gene in breast cancer. *J BUON Off J Balk Union Oncol.* 2013;18(4):891-898.
113. Peng L, Wei H, Liren L. Promoter methylation assay of SASH1 gene in hepatocellular carcinoma. *J BUON Off J Balk Union Oncol.* 2014;19(4):1041-1047.
114. Zhou D, Wei Z, Deng S, et al. SASH1 regulates melanocyte transepithelial migration through a novel Gas-SASH1-IQGAP1-E-Cadherin dependent pathway. *Cell Signal.* 2013;25(6):1526-1538. doi:10.1016/j.cellsig.2012.12.025.
115. Courcet J-B, Elalaoui SC, Duplomb L, et al. Autosomal-recessive SASH1 variants associated with a new genodermatosis with pigmentation defects, palmoplantar keratoderma and skin carcinoma. *Eur J Hum Genet EJHG.* October 2014. doi:10.1038/ejhg.2014.213.
116. Dauphinee SM, Clayton A, Hussainkhel A, et al. SASH1 is a scaffold molecule in endothelial TLR4 signaling. *J Immunol Baltim Md 1950.* 2013;191(2):892-901. doi:10.4049/jimmunol.1200583.
117. Rajakumar A, Tianjiao C, Handley DE, et al. Maternal Gene Expression Profiling During Pregnancy And Preeclampsia In Human Peripheral Blood Mononuclear Cells. *Placenta.* 2011;32(1):70-78. doi:10.1016/j.placenta.2010.10.004.
118. Bronner IF, Bochdanovits Z, Rizzu P, et al. Comprehensive mRNA Expression Profiling Distinguishes Tauopathies and Identifies Shared Molecular Pathways. *PLoS ONE.* 2009;4(8):e6826. doi:10.1371/journal.pone.0006826.
119. McDonough CW, Palmer ND, Hicks PJ, et al. A GENOME WIDE ASSOCIATION STUDY FOR DIABETIC NEPHROPATHY GENES IN AFRICAN AMERICANS. *Kidney Int.* 2011;79(5):563-572. doi:10.1038/ki.2010.467.
120. Huang S-MA, Hancock MK, Pitman JL, Orth AP, Gekakis N. Negative Regulators of Insulin Signaling Revealed in a Genome-Wide Functional Screen. *PLoS ONE.* 2009;4(9):e6871. doi:10.1371/journal.pone.0006871.
121. Verdugo RA, Zeller T, Rotival M, et al. Graphical modeling of gene expression in monocytes suggests molecular mechanisms explaining increased atherosclerosis in smokers. *PLoS One.* 2013;8(1):e50888. doi:10.1371/journal.pone.0050888.
122. Libby P, Aikawa M, Jain MK. Vascular endothelium and atherosclerosis. *Handb Exp Pharmacol.* 2006;(176 Pt 2):285-306.

123. Libby P. Inflammation in atherosclerosis. *Arterioscler Thromb Vasc Biol.* 2012;32(9):2045-2051. doi:10.1161/ATVBAHA.108.179705.
124. Hansson GK, Libby P. The immune response in atherosclerosis: a double-edged sword. *Nat Rev Immunol.* 2006;6(7):508-519. doi:10.1038/nri1882.
125. Auerbach O, Garfinkel L. Atherosclerosis and aneurysm of aorta in relation to smoking habits and age. *Chest.* 1980;78(6):805-809.
126. Redgrave JNE, Lovett JK, Rothwell PM. Histological features of symptomatic carotid plaques in relation to age and smoking: the oxford plaque study. *Stroke J Cereb Circ.* 2010;41(10):2288-2294. doi:10.1161/STROKEAHA.110.587006.
127. McEvoy JW, Blaha MJ, DeFilippis AP, et al. Cigarette smoking and cardiovascular events: role of inflammation and subclinical atherosclerosis from the multiethnic study of atherosclerosis. *Arterioscler Thromb Vasc Biol.* 2015;35(3):700-709. doi:10.1161/ATVBAHA.114.304562.
128. Talhout R, Schulz T, Florek E, van Benthem J, Wester P, Opperhuizen A. Hazardous compounds in tobacco smoke. *Int J Environ Res Public Health.* 2011;8(2):613-628. doi:10.3390/ijerph8020613.
129. Varela-Carver A, Parker H, Kleinert C, Rimoldi O. Adverse effects of cigarette smoke and induction of oxidative stress in cardiomyocytes and vascular endothelium. *Curr Pharm Des.* 2010;16(23):2551-2558.
130. Messner B, Bernhard D. Smoking and cardiovascular disease: mechanisms of endothelial dysfunction and early atherogenesis. *Arterioscler Thromb Vasc Biol.* 2014;34(3):509-515. doi:10.1161/ATVBAHA.113.300156.
131. Zeller T, Wild P, Szymczak S, et al. Genetics and beyond--the transcriptome of human monocytes and disease susceptibility. *PLoS One.* 2010;5(5):e10693. doi:10.1371/journal.pone.0010693.
132. Verdugo RA, Zeller T, Rotival M, et al. Graphical modeling of gene expression in monocytes suggests molecular mechanisms explaining increased atherosclerosis in smokers. *PLoS One.* 2013;8(1):e50888. doi:10.1371/journal.pone.0050888.
133. Zeller C, Hinzmann B, Seitz S, et al. SASH1: a candidate tumor suppressor gene on chromosome 6q24.3 is downregulated in breast cancer. *Oncogene.* 2003;22(19):2972-2983. doi:10.1038/sj.onc.1206474.
134. Martini M, Gnann A, Scheikl D, Holzmann B, Janssen K-P. The candidate tumor suppressor SASH1 interacts with the actin cytoskeleton and stimulates cell-matrix adhesion. *Int J Biochem Cell Biol.* 2011;43(11):1630-1640. doi:10.1016/j.biocel.2011.07.012.
135. Rimkus C, Martini M, Friederichs J, et al. Prognostic significance of downregulated expression of the candidate tumour suppressor gene SASH1 in colon cancer. *Br J Cancer.* 2006;95(10):1419-1423. doi:10.1038/sj.bjc.6603452.
136. Nitsche U, Rosenberg R, Balmert A, et al. Integrative marker analysis allows risk assessment for metastasis in stage II colon cancer. *Ann Surg.* 2012;256(5):763-771; discussion 771. doi:10.1097/SLA.0b013e318272de87.
137. Chen E, Chen Y, Dong L, Zhang J. Effects of SASH1 on lung cancer cell proliferation, apoptosis, and invasion in vitro. *Tumour Biol J Int Soc Oncodevelopmental Biol Med.* 2012;33(5):1393-1401. doi:10.1007/s13277-012-0387-2.

138. Yang L, Liu M, Gu Z, Chen J, Yan Y, Li J. Overexpression of SASH1 related to the decreased invasion ability of human glioma U251 cells. *Tumour Biol J Int Soc Oncodevelopmental Biol Med*. 2012;33(6):2255-2263. doi:10.1007/s13277-012-0487-z.
139. Lin S, Zhang J, Xu J, et al. Effects of SASH1 on melanoma cell proliferation and apoptosis in vitro. *Mol Med Rep*. 2012;6(6):1243-1248. doi:10.3892/mmr.2012.1099.
140. Dauphinee SM, Clayton A, Hussainkhel A, et al. SASH1 is a scaffold molecule in endothelial TLR4 signaling. *J Immunol Baltim Md 1950*. 2013;191(2):892-901. doi:10.4049/jimmunol.1200583.
141. Zhou D, Wei Z, Deng S, et al. SASH1 regulates melanocyte transepithelial migration through a novel G α s-SASH1-IQGAP1-E-Cadherin dependent pathway. *Cell Signal*. 2013;25(6):1526-1538. doi:10.1016/j.cellsig.2012.12.025.
142. Brochériou I, Maouche S, Durand H, et al. Antagonistic regulation of macrophage phenotype by M-CSF and GM-CSF: implication in atherosclerosis. *Atherosclerosis*. 2011;214(2):316-324. doi:10.1016/j.atherosclerosis.2010.11.023.
143. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*. 2001;29(9):e45.
144. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. *Bioinforma Oxf Engl*. 2008;24(13):1547-1548. doi:10.1093/bioinformatics/btn224.
145. Lin SM, Du P, Huber W, Kibbe WA. Model-based variance-stabilizing transformation for Illumina microarray data. *Nucleic Acids Res*. 2008;36(2):e11. doi:10.1093/nar/gkm1075.
146. Nordskog BK, Blixt AD, Morgan WT, Fields WR, Hellmann GM. Matrix-degrading and pro-inflammatory changes in human vascular endothelial cells exposed to cigarette smoke condensate. *Cardiovasc Toxicol*. 2003;3(2):101-117.
147. Slattery ML, Samowitz W, Ma K, et al. CYP1A1, cigarette smoking, and colon and rectal cancer. *Am J Epidemiol*. 2004;160(9):842-852. doi:10.1093/aje/kwh298.
148. Wierstra I, Alves J. FOXM1, a typical proliferation-associated transcription factor. *Biol Chem*. 2007;388(12):1257-1274. doi:10.1515/BC.2007.159.
149. Sitras V, Paulssen RH, Grønnaas H, et al. Differential placental gene expression in severe preeclampsia. *Placenta*. 2009;30(5):424-433. doi:10.1016/j.placenta.2009.01.012.
150. Bronner IF, Bochdanovits Z, Rizzu P, et al. Comprehensive mRNA expression profiling distinguishes tauopathies and identifies shared molecular pathways. *PLoS One*. 2009;4(8):e6826. doi:10.1371/journal.pone.0006826.
151. Courcet J-B, Elalaoui SC, Duplomb L, et al. Autosomal-recessive SASH1 variants associated with a new genodermatosis with pigmentation defects, palmoplantar keratoderma and skin carcinoma. *Eur J Hum Genet EJHG*. October 2014. doi:10.1038/ejhg.2014.213.
152. Sakao S, Taraseviciene-Stewart L, Cool CD, et al. VEGF-R blockade causes endothelial cell apoptosis, expansion of surviving CD34+ precursor cells and transdifferentiation to smooth muscle-like and neuronal-like cells. *FASEB J Off Publ Fed Am Soc Exp Biol*. 2007;21(13):3640-3652. doi:10.1096/fj.07-8432com.

153. Sheyu L, Hui L, Junyu Z, et al. Promoter methylation assay of SASH1 gene in breast cancer. *J BUON Off J Balk Union Oncol*. 2013;18(4):891-898.
154. Peng L, Wei H, Liren L. Promoter methylation assay of SASH1 gene in hepatocellular carcinoma. *J BUON Off J Balk Union Oncol*. 2014;19(4):1041-1047.
155. Lee KWK, Pausova Z. Cigarette smoking and DNA methylation. *Front Genet*. 2013;4:132. doi:10.3389/fgene.2013.00132.
156. Tsaprouni LG, Yang T-P, Bell J, et al. Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. *Epigenetics Off J DNA Methylation Soc*. 2014;9(10):1382-1396. doi:10.4161/15592294.2014.969637.
157. Wan J, Johnson M, Schilz J, Djordjevic MV, Rice JR, Shields PG. Evaluation of In Vitro Assays For Assessing the Toxicity of Cigarette Smoke and Smokeless Tobacco. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol*. 2009;18(12):3263-3304. doi:10.1158/1055-9965.EPI-09-0965.
158. Brandon EFA, Raap CD, Meijerman I, Beijnen JH, Schellens JHM. An update on in vitro test methods in human hepatic drug biotransformation research: pros and cons. *Toxicol Appl Pharmacol*. 2003;189(3):233-246. doi:10.1016/S0041-008X(03)00128-5.
159. Fuster JJ, Fernández P, González-Navarro H, Silvestre C, Nabah YNA, Andrés V. Control of cell proliferation in atherosclerosis: insights from animal models and human studies. *Cardiovasc Res*. 2010;86(2):254-264. doi:10.1093/cvr/cvp363.
160. Moreno PR, Purushothaman K-R, Sirol M, Levy AP, Fuster V. Neovascularization in Human Atherosclerosis. *Circulation*. 2006;113(18):2245-2252. doi:10.1161/CIRCULATIONAHA.105.578955.
161. Zimmet JM, Ladd D, Jackson CW, Stenberg PE, Ravid K. A role for cyclin D3 in the endomitotic cell cycle. *Mol Cell Biol*. 1997;17(12):7248-7259.
162. Weng H-Y, Huang H-L, Zhao P-P, Zhou H, Qu L-H. Translational repression of cyclin D3 by a stable G-quadruplex in its 5' UTR. *RNA Biol*. 2012;9(8):1099-1109. doi:10.4161/rna.21210.
163. Radulovich N, Pham N-A, Strumpf D, et al. Differential roles of cyclin D1 and D3 in pancreatic ductal adenocarcinoma. *Mol Cancer*. 2010;9:24. doi:10.1186/1476-4598-9-24.
164. Barsotti AM, Prives C. Pro-proliferative FoxM1 is a target of p53-mediated repression. *Oncogene*. 2009;28(48):4295-4305. doi:10.1038/onc.2009.282.
165. Wen N, Wang Y, Wen L, et al. Overexpression of FOXM1 predicts poor prognosis and promotes cancer cell proliferation, migration and invasion in epithelial ovarian cancer. *J Transl Med*. 2014;12:134. doi:10.1186/1479-5876-12-134.
166. Zhao Y-Y, Gao X-P, Zhao YD, et al. Endothelial cell-restricted disruption of FoxM1 impairs endothelial repair following LPS-induced vascular injury. *J Clin Invest*. 2006;116(9):2333-2343. doi:10.1172/JCI27154.
167. Wohak LE, Krais AM, Kucab JE, et al. Carcinogenic polycyclic aromatic hydrocarbons induce CYP1A1 in human cells via a p53-dependent mechanism. *Arch Toxicol*. November 2014. doi:10.1007/s00204-014-1409-1.

168. Aviv T, Lin Z, Lau S, Rendl LM, Sicheri F, Smibert CA. The RNA-binding SAM domain of Smaug defines a new family of post-transcriptional regulators. *Nat Struct Biol*. 2003;10(8):614-621. doi:10.1038/nsb956.
169. Antoni G, Morange P-E, Luo Y, et al. A multi-stage multi-design strategy provides strong evidence that the BAI3 locus is associated with early-onset venous thromboembolism. *J Thromb Haemost JTH*. 2010;8(12):2671-2679. doi:10.1111/j.1538-7836.2010.04092.x.
170. Antoni G, Oudot-Mellakh T, Dimitromanolakis A, et al. Combined analysis of three genome-wide association studies on vWF and FVIII plasma levels. *BMC Med Genet*. 2011;12:102. doi:10.1186/1471-2350-12-102.
171. Huang J, Sabater-Lleal M, Asselbergs FW, et al. Genome-wide association study for circulating levels of PAI-1 provides novel insights into its regulation. *Blood*. 2012;120(24):4873-4881. doi:10.1182/blood-2012-06-436188.
172. Oudot-Mellakh T, Cohen W, Germain M, et al. Genome wide association study for plasma levels of natural anticoagulant inhibitors and protein C anticoagulant pathway: the MARTHA project. *Br J Haematol*. 2012;157(2):230-239. doi:10.1111/j.1365-2141.2011.09025.x.
173. Dick KJ, Nelson CP, Tsaprouni L, et al. DNA methylation and body-mass index: a genome-wide analysis. *Lancet*. 2014;383(9933):1990-1998. doi:10.1016/S0140-6736(13)62674-4.
174. Gagnon F, Aïssi D, Carrié A, Morange P-E, Trégouët D-A. Robust validation of methylation levels association at CPT1A locus with lipid plasma levels. *J Lipid Res*. 2014;55(7):1189-1191. doi:10.1194/jlr.E051276.
175. Aïssi D, Dennis J, Ladouceur M, et al. Genome-wide investigation of DNA methylation marks associated with FV Leiden mutation. *PLoS One*. 2014;9(9):e108087. doi:10.1371/journal.pone.0108087.
176. Rocañín-Arjó A, Dennis J, Suchon P, et al. Thrombin Generation Potential and Whole-Blood DNA methylation. *Thromb Res*. December 2014. doi:10.1016/j.thromres.2014.12.010.
177. Triche TJ, Weisenberger DJ, Van Den Berg D, Laird PW, Siegmund KD. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. *Nucleic Acids Res*. 2013;41(7):e90. doi:10.1093/nar/gkt090.
178. Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol*. 2012;13(6):R44. doi:10.1186/gb-2012-13-6-r44.
179. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinforma Oxf Engl*. 2014;30(10):1363-1369. doi:10.1093/bioinformatics/btu049.
180. Wilhelm-Benartzi CS, Koestler DC, Karagas MR, et al. Review of processing and analysis methods for DNA methylation array data. *Br J Cancer*. 2013;109(6):1394-1402. doi:10.1038/bjc.2013.496.
181. Liu Y, Aryee MJ, Padyukov L, et al. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. *Nat Biotechnol*. 2013;31(2):142-147. doi:10.1038/nbt.2487.

182. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*. 2012;13:86. doi:10.1186/1471-2105-13-86.
183. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biol*. 2014;15(2):R31. doi:10.1186/gb-2014-15-2-r31.
184. Koestler DC, Christensen B, Karagas MR, et al. Blood-based profiles of DNA methylation predict the underlying distribution of cell types: a validation analysis. *Epigenetics Off J DNA Methylation Soc*. 2013;8(8):816-826. doi:10.4161/epi.25430.
185. Vojtek AB, Hollenberg SM. Ras-Raf interaction: two-hybrid analysis. *Methods Enzymol*. 1995;255:331-342.
186. Bartel P, Chien CT, Sternglanz R, Fields S. Elimination of false positives that arise in using the two-hybrid system. *BioTechniques*. 1993;14(6):920-924.
187. Fromont-Racine M, Rain JC, Legrain P. Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nat Genet*. 1997;16(3):277-282. doi:10.1038/ng0797-277.
188. Formstecher E, Aresta S, Collura V, et al. Protein interaction mapping: a Drosophila case study. *Genome Res*. 2005;15(3):376-384. doi:10.1101/gr.2659105.
189. Rain JC, Selig L, De Reuse H, et al. The protein-protein interaction map of Helicobacter pylori. *Nature*. 2001;409(6817):211-215. doi:10.1038/35051615.
190. Wojcik J, Boneca IG, Legrain P. Prediction, assessment and validation of protein interaction maps in bacteria. *J Mol Biol*. 2002;323(4):763-770.
191. Schmidt M, Raghavan B, Müller V, et al. Crucial role for human Toll-like receptor 4 in the development of contact allergy to nickel. *Nat Immunol*. 2010;11(9):814-819. doi:10.1038/ni.1919.
192. Tsaprouni LG, Yang T-P, Bell J, et al. Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. *Epigenetics Off J DNA Methylation Soc*. 2014;9(10):1382-1396. doi:10.4161/15592294.2014.969637.
193. Schmid JA, Birbach A. IκB kinase beta (IKKβ/IKK2/IKBKB)--a key molecule in signaling to the transcription factor NF-κB. *Cytokine Growth Factor Rev*. 2008;19(2):157-165. doi:10.1016/j.cytogfr.2008.01.006.
194. Hansen SD, Kwiatkowski AV, Ouyang C-Y, et al. αE-catenin actin-binding domain alters actin filament conformation and regulates binding of nucleation and disassembly factors. *Mol Biol Cell*. 2013;24(23):3710-3720. doi:10.1091/mbc.E13-07-0388.
195. Eskildsen S, Justesen J, Schierup MH, Hartmann R. Characterization of the 2'-5'-oligoadenylate synthetase ubiquitin-like family. *Nucleic Acids Res*. 2003;31(12):3166-3173.
196. Marinković G, Heemskerk N, van Buul JD, de Waard V. The ins and outs of small GTPase Rac1 in the vasculature. *J Pharmacol Exp Ther*. June 2015. doi:10.1124/jpet.115.223610.
197. Adorni MP, Ronda N, Bernini F, Favari E. Rac1 and cholesterol metabolism in macrophage. *J Cardiovasc Pharmacol*. 2013;62(5):418-424. doi:10.1097/FJC.0b013e31829dd874.

198. Ferri N, Contini A, Bernini SK, Corsini A. Role of small GTPase protein Rac1 in cardiovascular diseases: development of new selective pharmacological inhibitors. *J Cardiovasc Pharmacol*. 2013;62(5):425-435. doi:10.1097/FJC.0b013e3182a18bcc.
199. Lu H, Li Z, Xue Y, et al. AFF1 is a ubiquitous P-TEFb partner to enable Tat extraction of P-TEFb from 7SK snRNP and formation of SECs for HIV transactivation. *Proc Natl Acad Sci U S A*. 2014;111(1):E15-E24. doi:10.1073/pnas.1318503111.
200. Rothbarth K, Spiess E, Juodka B, et al. Induction of apoptosis by overexpression of the DNA-binding and DNA-PK-activating protein C1D. *J Cell Sci*. 1999;112 (Pt 13):2223-2232.
201. Al-Jameil N, Aziz Khan F, Fareed Khan M, Tabassum H. A brief overview of preeclampsia. *J Clin Med Res*. 2014;6(1):1-7. doi:10.4021/jocmr1682w.
202. Myatt L, Webster RP. Vascular biology of preeclampsia. *J Thromb Haemost JTH*. 2009;7(3):375-384. doi:10.1111/j.1538-7836.2008.03259.x.
203. Erdem-Eraslan L, Heijmans D, de Wit M, et al. Tumor-specific mutations in low-frequency genes affect their functional properties. *J Neurooncol*. February 2015. doi:10.1007/s11060-015-1741-1.
204. Noll JE, Hewett DR, Williams SA, et al. SAMS1 is a tumor suppressor gene in multiple myeloma. *Neoplasia N Y N*. 2014;16(7):572-585. doi:10.1016/j.neo.2014.07.002.
205. Von Holleben M, Gohla A, Janssen K-P, Iritani BM, Beer-Hammer S. Immunoinhibitory adapter protein Src homology domain 3 lymphocyte protein 2 (SLy2) regulates actin dynamics and B cell spreading. *J Biol Chem*. 2011;286(15):13489-13501. doi:10.1074/jbc.M110.155184.
206. Jiang P, Tang Y, He L, et al. Aberrant expression of nuclear KPNA2 is correlated with early recurrence and poor prognosis in patients with small hepatocellular carcinoma after hepatectomy. *Med Oncol Northwood Lond Engl*. 2014;31(8):131. doi:10.1007/s12032-014-0131-4.
207. Gousias K, Niehusmann P, Gielen G, Simon M, Boström J. KPNA2 predicts long term survival in patients with anaplastic oligoastrocytomas. *J Clin Neurosci Off J Neurosurg Soc Australas*. 2014;21(10):1719-1724. doi:10.1016/j.jocn.2014.01.011.
208. Huang L, Wang H-Y, Li J-D, et al. KPNA2 promotes cell proliferation and tumorigenicity in epithelial ovarian carcinoma through upregulation of c-Myc and downregulation of FOXO3a. *Cell Death Dis*. 2013;4:e745. doi:10.1038/cddis.2013.256.
209. Noetzel E, Rose M, Bornemann J, Gajewski M, Knüchel R, Dahl E. Nuclear transport receptor karyopherin- α 2 promotes malignant breast cancer phenotypes in vitro. *Oncogene*. 2012;31(16):2101-2114. doi:10.1038/onc.2011.403.
210. Christiansen A, Dyrskjøt L. The functional role of the novel biomarker karyopherin α 2 (KPNA2) in cancer. *Cancer Lett*. 2013;331(1):18-23. doi:10.1016/j.canlet.2012.12.013.
211. Garritano S, Inga A, Gemignani F, Landi S. More targets, more pathways and more clues for mutant p53. *Oncogenesis*. 2013;2:e54. doi:10.1038/oncsis.2013.15.
212. Ravanat JL, Douki T, Cadet J. Direct and indirect effects of UV radiation on DNA and its components. *J Photochem Photobiol B*. 2001;63(1-3):88-102.

213. Dally H, Hartwig A. Induction and repair inhibition of oxidative DNA damage by nickel(II) and cadmium(II) in mammalian cells. *Carcinogenesis*. 1997;18(5):1021-1026. doi:10.1093/carcin/18.5.1021.
214. Asami S, Manabe H, Miyake J, et al. Cigarette smoking induces an increase in oxidative DNA damage, 8-hydroxydeoxyguanosine, in a central site of the human lung. *Carcinogenesis*. 1997;18(9):1763-1766. doi:10.1093/carcin/18.9.1763.
215. Sarir H, Mortaz E, Karimi K, et al. Cigarette smoke regulates the expression of TLR4 and IL-8 production by human macrophages. *J Inflamm*. 2009;6(1):12. doi:10.1186/1476-9255-6-12.
216. Doz E, Noulin N, Boichot E, et al. Cigarette Smoke-Induced Pulmonary Inflammation Is TLR4/MyD88 and IL-1R1/MyD88 Signaling Dependent. *J Immunol*. 2008;180(2):1169-1178. doi:10.4049/jimmunol.180.2.1169.
217. Nilsson E, Jansson PA, Perfilyev A, et al. Altered DNA Methylation and Differential Expression of Genes Influencing Metabolism and Inflammation in Adipose Tissue From Subjects With Type 2. *Diabetes*. 2014;63(9):2962-2976. doi:10.2337/db13-1459.
218. Dong C, Yoon W, Goldschmidt-Clermont PJ. DNA methylation and atherosclerosis. *J Nutr*. 2002;132(8 Suppl):2406S - 2409S.
219. Baylin SB. DNA methylation and gene silencing in cancer. *Nat Clin Pract Oncol*. 2005;2:S4-S11. doi:10.1038/ncponc0354.
220. Donaldson MM, Tavares AAM, Hagan IM, Nigg EA, Glover DM. The mitotic roles of Polo-like kinase. *J Cell Sci*. 2001;114(13):2357-2358.
221. Cuddihy AR, O'Connell MJ. Cell-cycle responses to DNA damage in G2. In: *International Review of Cytology*. Vol 222. Elsevier; 2003:99-140. <http://linkinghub.elsevier.com/retrieve/pii/S0074769602220136>. Accessed May 11, 2015.
222. Nordskog BK, Blixt AD, Morgan WT, Fields WR, Hellmann GM. Matrix-degrading and pro-inflammatory changes in human vascular endothelial cells exposed to cigarette smoke condensate. *Cardiovasc Toxicol*. 2003;3(2):101-117.
223. Jacot W, Fiche M, Zaman K, Wolfer A, Lamy P-J. The HER2 amplicon in breast cancer: Topoisomerase IIA and beyond. *Biochim Biophys Acta*. 2013;1836(1):146-157. doi:10.1016/j.bbcan.2013.04.004.
224. Bouchalova K, Cizkova M, Cwiertka K, Trojanec R, Hajduch M. Triple negative breast cancer--current status and prospective targeted treatment based on HER1 (EGFR), TOP2A and C-MYC gene assessment. *Biomed Pap Med Fac Univ Palacký Olomouc Czechoslov*. 2009;153(1):13-17.
225. Joshi AA, Wu Z, Reed RF, Suttle DP. Nuclear factor-Y binding to the topoisomerase IIalpha promoter is inhibited by both the p53 tumor suppressor and anticancer drugs. *Mol Pharmacol*. 2003;63(2):359-367.
226. Wohak LE, Kraus AM, Kucab JE, et al. Carcinogenic polycyclic aromatic hydrocarbons induce CYP1A1 in human cells via a p53-dependent mechanism. *Arch Toxicol*. November 2014. doi:10.1007/s00204-014-1409-1.
227. Beresford AP. CYP1A1: friend or foe? *Drug Metab Rev*. 1993;25(4):503-517. doi:10.3109/03602539308993984.

228. Uno S, Dalton TP, Derkenne S, et al. Oral exposure to benzo[a]pyrene in the mouse: detoxication by inducible cytochrome P450 is more important than metabolic activation. *Mol Pharmacol*. 2004;65(5):1225-1237. doi:10.1124/mol.65.5.1225.
229. Barsotti AM, Prives C. Pro-proliferative FoxM1 is a target of p53-mediated repression. *Oncogene*. 2009;28(48):4295-4305. doi:10.1038/onc.2009.282.
230. Wierstra I. FOXM1 (Forkhead box M1) in tumorigenesis: overexpression in human cancer, implication in tumorigenesis, oncogenic functions, tumor-suppressive properties, and target of anticancer therapy. *Adv Cancer Res*. 2013;119:191-419. doi:10.1016/B978-0-12-407190-2.00016-2.
231. Ahmad A, Wang Z, Kong D, et al. FoxM1 down-regulation leads to inhibition of proliferation, migration and invasion of breast cancer cells through the modulation of extra-cellular matrix degrading factors. *Breast Cancer Res Treat*. 2010;122(2):337-346. doi:10.1007/s10549-009-0572-1.
232. Zhao YD, Huang X, Yi F, et al. Endothelial FoxM1 mediates bone marrow progenitor cell-induced vascular repair and resolution of inflammation following inflammatory lung injury. *Stem Cells Dayt Ohio*. 2014;32(7):1855-1864. doi:10.1002/stem.1690.
233. Mirza MK, Sun Y, Zhao YD, et al. FoxM1 regulates re-annealing of endothelial adherens junctions through transcriptional control of beta-catenin expression. *J Exp Med*. 2010;207(8):1675-1685. doi:10.1084/jem.20091857.
234. Tago K, Nakamura T, Nishita M, et al. Inhibition of Wnt signaling by ICAT, a novel beta-catenin-interacting protein. *Genes Dev*. 2000;14(14):1741-1749.
235. Kubota S, Copeland TD, Pomerantz RJ. Nuclear and nucleolar targeting of human ribosomal protein S25: common features shared with HIV-1 regulatory proteins. *Oncogene*. 1999;18(7):1503-1514. doi:10.1038/sj.onc.1202429.
236. Zhang X, Wang W, Wang H, Wang M-H, Xu W, Zhang R. Identification of ribosomal protein S25 (RPS25)-MDM2-p53 regulatory feedback loop. *Oncogene*. 2013;32(22):2782-2791. doi:10.1038/onc.2012.289.
237. Gagat M, Grzanka D, Izdebska M, Sroka WD, Marszał MP, Grzanka A. Tropomyosin-1 protects endothelial cell-cell junctions against cigarette smoke extract through F-actin stabilization in EA.hy926 cell line. *Acta Histochem*. 2014;116(4):606-618. doi:10.1016/j.acthis.2013.11.013.
238. Simoneau B, Houle F, Huot J. Regulation of endothelial permeability and transendothelial migration of cancer cells by tropomyosin-1 phosphorylation. *Vasc Cell*. 2012;4(1):18. doi:10.1186/2045-824X-4-18.
239. Bharadwaj S, Thanawala R, Bon G, Falcioni R, Prasad GL. Resensitization of breast cancer cells to anoikis by tropomyosin-1: role of Rho kinase-dependent cytoskeleton and adhesion. *Oncogene*. 2005;24(56):8291-8303. doi:10.1038/sj.onc.1208993.
240. Mahadev K, Raval G, Bharadwaj S, et al. Suppression of the transformed phenotype of breast cancer by tropomyosin-1. *Exp Cell Res*. 2002;279(1):40-51.
241. Blanchoin L, Pollard TD, Hitchcock-DeGregori SE. Inhibition of the Arp2/3 complex-nucleated actin polymerization and branch formation by tropomyosin. *Curr Biol CB*. 2001;11(16):1300-1304.
242. Radulovich N, Pham N-A, Strumpf D, et al. Differential roles of cyclin D1 and D3 in pancreatic ductal adenocarcinoma. *Mol Cancer*. 2010;9:24. doi:10.1186/1476-4598-9-24.

243. Zimmet JM, Ladd D, Jackson CW, Stenberg PE, Ravid K. A role for cyclin D3 in the endomitotic cell cycle. *Mol Cell Biol*. 1997;17(12):7248-7259.
244. Liu C-S, Chang C-C, Du Y-C, et al. 2-hydroxy-4'-methoxychalcone inhibits proliferation and inflammation of human aortic smooth muscle cells by increasing the expression of peroxisome proliferator-activated receptor gamma. *J Cardiovasc Pharmacol*. 2012;59(4):339-351. doi:10.1097/FJC.0b013e3182440486.
245. Moulton KS, Vakili K, Zurakowski D, et al. Inhibition of plaque neovascularization reduces macrophage accumulation and progression of advanced atherosclerosis. *Proc Natl Acad Sci U S A*. 2003;100(8):4736-4741. doi:10.1073/pnas.0730843100.
246. Moreno PR, Purushothaman KR, Zias E, Sanz J, Fuster V. Neovascularization in human atherosclerosis. *Curr Mol Med*. 2006;6(5):457-477.

ANNEXES

Annex 1: SLC39A8 (ZIP8) and cigarette smoking, genesis of a new project

I Introduction

As part of this thesis work, we also studied a second GHS candidate linking smoking to atherosclerosis, namely *SLC39A8* which was found strongly correlated to the number of plaques and was also correlated with cigarette smoking. The statistically predicted pathway for *SLC39A8*, described by Verdugo et al¹¹⁴, a work to which I have also contributed, linked *SLC39A8* expression to smoking *via* several other genes. In the present work, one of our main interests was the link between *SLC39A8* and growth arrest specific 6 (*Gas6*), as the latter possesses a number of pleiotropic effects both protective and deleterious in atherosclerosis¹. In this network, *Gas6* expression was directly correlated to smoking in our results from GHS (Figure 1)¹¹⁴.

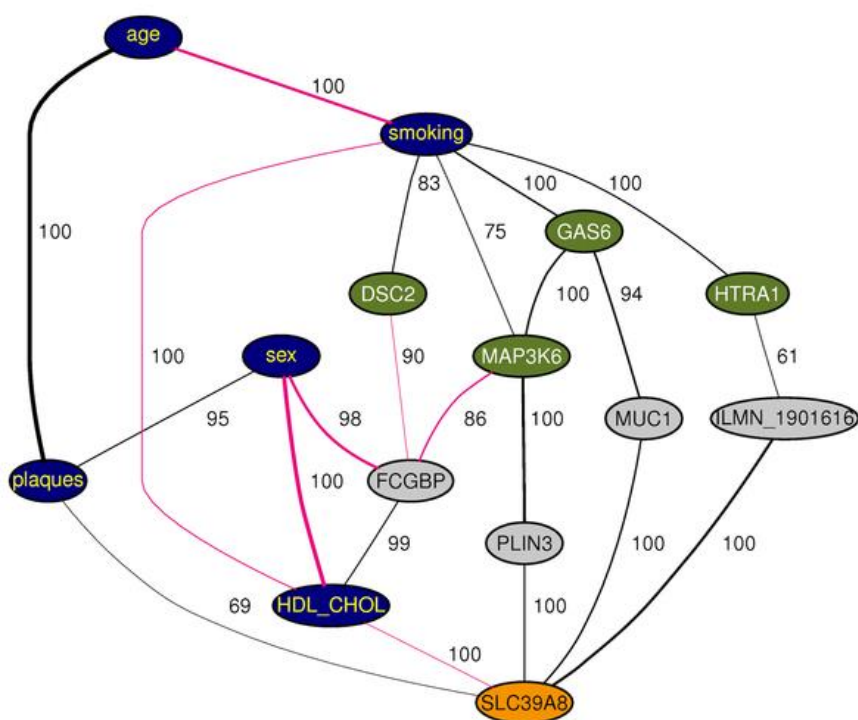


Figure 1: *SLC39A8* predicted network linking smoking to atherosclerosis, from Verdugo et al 2013¹¹⁴. This graph represents a consensus network from 1000 bootstraps. Edges among variables are drawn if detected in at least 60% of bootstrapped samples. The recovery percentages are indicated to the right of the medial section of each edge. Line thickness is proportional to the edge's partial correlation. Black edges denote positive and pink edges negative partial correlations. Plaques and risk factors are in blue. Genes directly connected to smoking are in green and those directly connected to plaques are in orange. Other genes are in gray. Only genes that are involved in the shortest paths connecting smoking to plaques are shown.

We thus decided to study further SLC39A8 expression in human atherosclerotic tissues and vascular cells to better understand its link to atherosclerosis. A secondary objective was also to confirm *in vitro* the interdependence between Gas6 and SLC39A8 expression.

I.1 SLC39A8 gene organisation

Zip8 was first described in 2002 by Begum et al.² in monocytes stimulated with inactivated *Mycobacterium bovis* (a bacteria responsible for the bovine tuberculosis, one of the deadliest cattle disease of the 20th century). The gene coding for Zip8, *SLC39A8*, is located on the chromosome 4 in the region q24, having a length of 1383 bp and divided into 8 exons. This gene produces 4 known alternative transcripts of 2624, 3187, 3309 and 4098 bp (Figure 2).

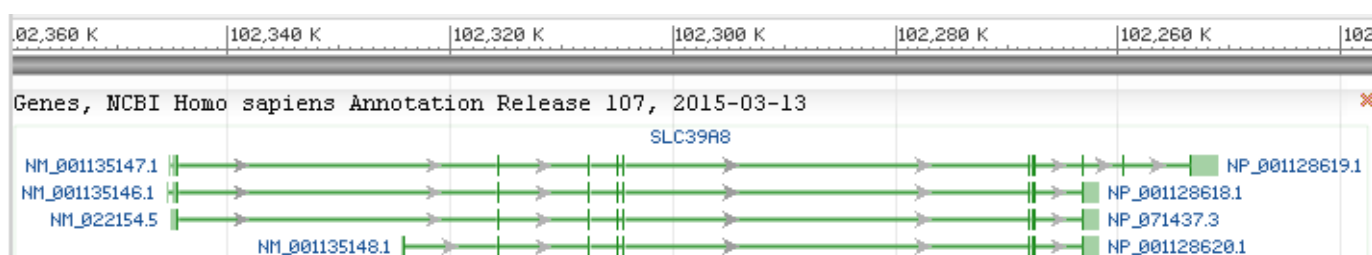


Figure 2. ZIP8 transcripts as described in NCBI database.

I.2 SLC39A8 protein (ZIP8)

The *SLC39A8* gene encodes the ZIP8 (Zrt- and Irt-like Protein 8) protein with 8 transmembrane domains, that acts as a membrane divalent metallic ions / HCO_3^- co-transporter³ (Figure 3).

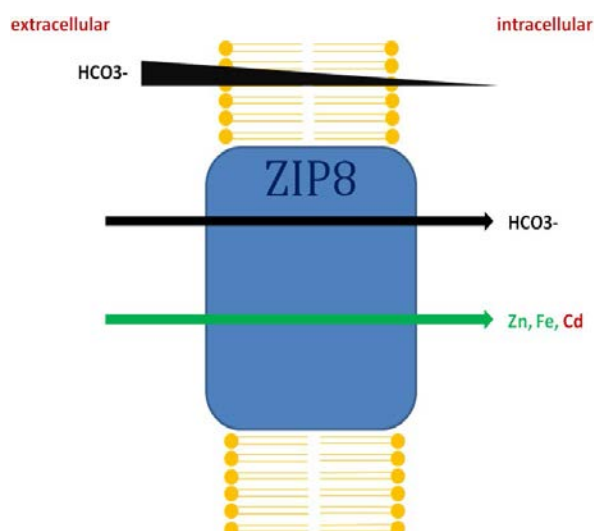


Figure 3. Membrane transport of metal ions through ZIP8.

Zip8 is a symporter that imports simultaneously into the cytoplasm one divalent metal ion (Zn^{2+} , Fe^{2+} , Cd^{2+} ...) and 2 HCO_3^- ions using the HCO_3^- concentration gradient driving force. It is a member of the solute carrier (SLC) family of proteins which is composed of more than 300 members divided

into 52 families. Consequently this group of proteins is involved in the transport of diverse compounds, including a wide range of organic and inorganic molecules. The name SLC39A8 is composed of the root name SLC, followed by the family number, followed by a letter defining the sub-family and ends with the actual number of the member of the family. Hence, SLC39A8 is the 8th member of the sub-family A of the family 39. This family of SLC regroups 14 SLC members, all involved in the transport of metallic ions⁴.

I.3 Zip8 in the literature

Unlike SASH1, Zip 8 function is well characterized, as around 60 publications are found in PubMed as compared to about 20 on SASH1. Zip8 was shown to be a broad-scope metal ion transporter⁵, but most of the research on ZIP8 focused on its ability to transport zinc (Zn) because of the importance of this metal in the activation of the immune system and inflammation^{6,7}. Cadmium (Cd) transport through ZIP8 was also well studied, as Cd is a highly toxic metal that damages DNA and therefore induces cytotoxicity^{8,9}, especially in the vascular tissues. I will focus this brief overview on the ZIP8 literature on those two points as both activation of the immune system and cell survival are involved in atherosclerosis.

ZIP8 activation of the immune system and inflammation

As Zn²⁺ ions are involved in the activation of the immune system through the modulation of various inflammatory pathways in immunocompetent cells⁶, it was expected that ZIP8, that participates in Zn²⁺ transport into the cells, would also be involved in inflammation. Indeed, Aydemir et al¹⁰ showed that activation of human T-cells, using antibody coated beads that mimic antigen presentation, resulted in an increased ZIP8 expression. Similar observations were made in bacterial lysate activated monocytes and macrophages by Begum et al². From a mechanistic point of view, ZIP8 affects both INF γ production and cytolytic activity of perforin. Zip8 is also a target and a regulator for NF κ B, through Zn²⁺ mediated inhibition of IKK^{11,12}.

ZIP8, cadmium transport and vascular system

Another aspect of ZIP8 in the literature focuses on its ability to transport cadmium. Cadmium is a cytotoxic metal that induces DNA damage and inflammation^{9,8}. Lethal when inhaled at high dose, cadmium is present at detectable levels in cigarette smoke¹³ and cadmium accumulation in the blood and vascular wall is significantly higher in smokers when compared to non-smokers. The vascular system is particularly sensitive to chronic cadmium exposure, as it disrupts the endothelial junctions, thus damaging endothelial barrier¹⁴. Cadmium also activates a wide range of mechanisms involved in inflammation and atherosclerosis. At the molecular level, it increases TNF α production¹⁵, plasminogen 1 production¹⁶, PKC activity¹⁷ and extracellular matrix production¹⁸. At the cellular level cadmium increases leukocytes and platelets adhesion to the endothelium¹⁹. Interestingly, at low concentration cadmium increases smooth muscle cells proliferation²⁰, but at high dose it is cytotoxic^{21,22,17}. In the latter cells the expression of metallothioneins, which protect against metal-induced toxicity is weak²³. Additionally cadmium was shown to inhibit several

atheroprotective mechanisms, notably the NO production^{24,25}. In a mice model of atherosclerosis (ApoE knockout mice) cadmium exposure induces expression of the adhesion molecule VCAM1 and the heat shock protein HSP60, both involved in the early steps of inflammation and atherosclerosis²⁶. Exposure to cadmium results in increased atherosclerotic lesions formation in mice when compared to their untreated littermates. In humans, the presence of cadmium in the blood circulation of 195 young females was shown to be an independent risk factor for atherosclerosis²⁷. In the same study, the authors showed that the exposition of human endothelial cells to cadmium resulted in increased endothelial permeability and cell cytotoxicity, along with a reduced proliferation rate.

Taken together, the current knowledge on ZIP8 in the literature points to a role in atherosclerosis through the cadmium transport into the cells and its subsequent cytotoxic effects.

I.4 Material and methods

Cell isolation and culture

As previously described p36

siRNA-mediated silencing of Gas6 and ZIP8

As previously described p36-37

RNA extraction, reverse transcription and quantitative polymerase chain reaction (PCR)

As previously described p37

Western blotting

As previously described p38

Immunostaining of cells and tissues

As previously described p39

II Results

II.1 ZIP8 expression in vascular cells and in the vascular wall

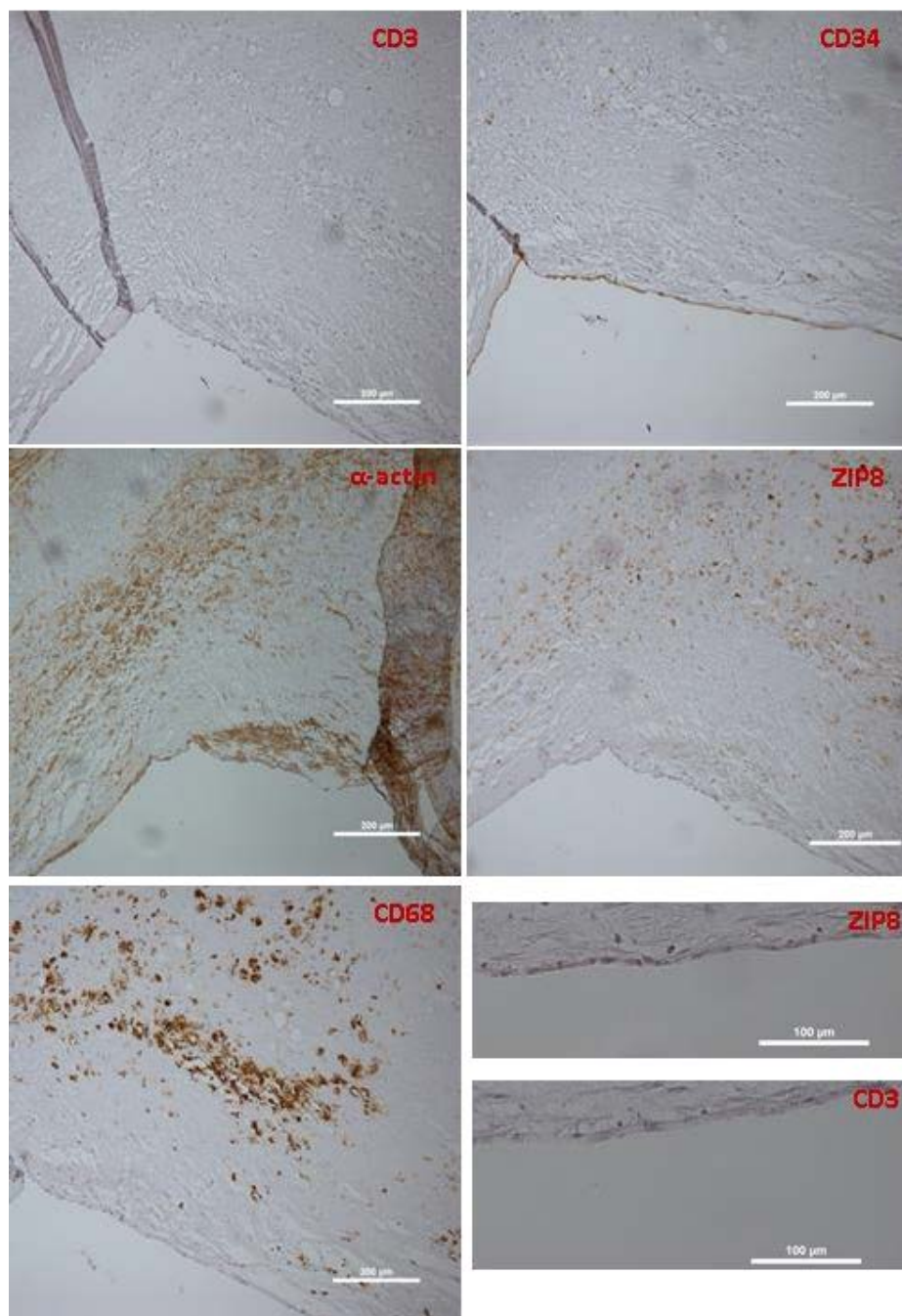


Figure 4. ZIP8 protein expression in human carotid atherosclerotic plaque. Immunostaining of ZIP8 and specific cell type markers: lymphocytes (CD3, no staining), macrophages (CD68), endothelial cells (CD34), smooth muscle cells (α -actin). Zip8 staining co-localized mainly with foam cells and dedifferentiated smooth muscle cells. A weak staining could also be seen at higher magnification in smooth muscle cells of the media and endothelial cells.

As ZIP8 protein expression was not previously investigated within the atherosclerotic plaque, we performed a ZIP8 immunostaining on human carotid plaques (Figure 4). Zip8 staining co-localized mainly with foam cells and dedifferentiated smooth muscle cells. A weak staining could also be seen at higher magnification in smooth muscle cells of the media and endothelial cells, indicating that ZIP8 was also present in those cell types, albeit at lower level.

At the cellular levels, RT-qPCR experiments revealed that ZIP8 was expressed in human monocytes and monocyte-derived macrophages (Figure 5). ZIP8 expression was significantly different between these cell types ($p < 0.05$) with the highest levels being in macrophages and the lowest in HAECs.

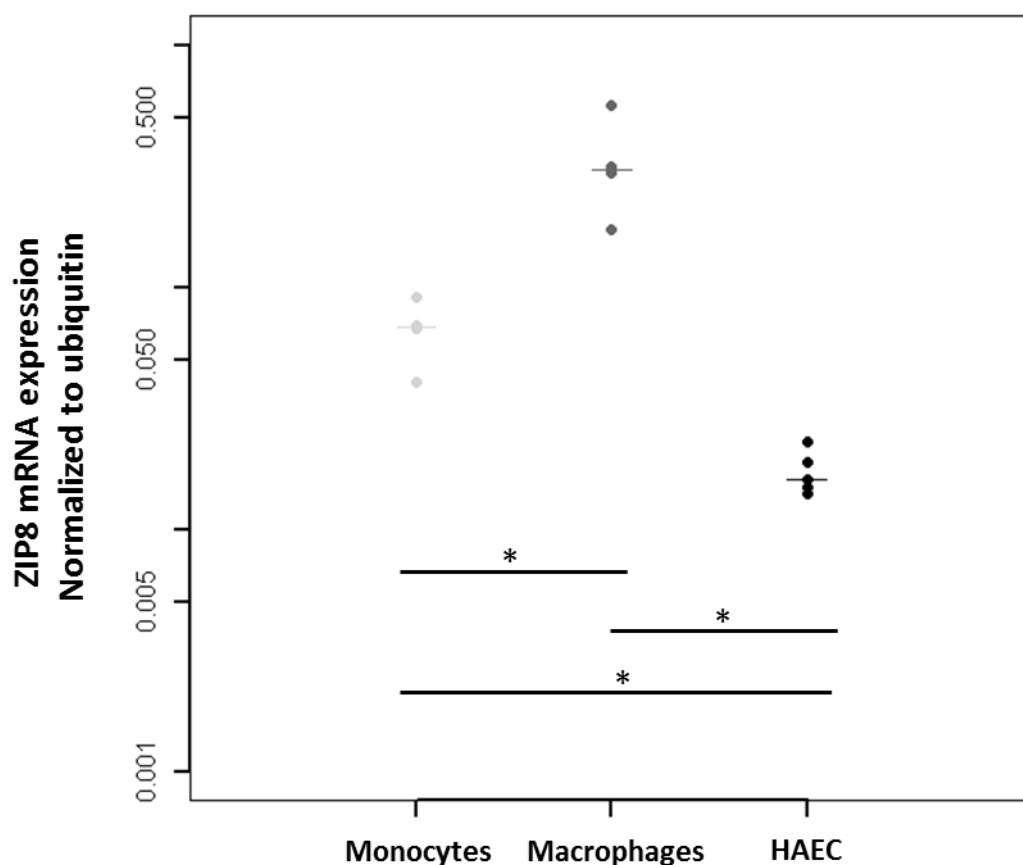


Figure 5. ZIP8 expression in human vascular cells in vitro. Expression of ZIP8 in monocytes (positively selected from buffy coats, $n=4$), monocytes derived macrophages ($n=4$) and HAECs ($n=5$) was measured by RT-qPCR. ZIP8 expression is significantly higher in macrophages and weaker in endothelial cells.

II.2 ZIP8 expression and the plaque formation

As ZIP8 was found directly correlated to the number of plaques, we decided to investigate its expression by RT-qPCR in normal vessels (mammary artery, n=11), carotid fatty streaks,(n=28) and plaques (n=34) (Figure 6). Zip8 expression was significantly higher in fatty streaks and plaques when compared to normal vascular tissue ($p<0.001$). Zip8 expression was also higher in plaques compared to fatty streak ($p<0.001$).

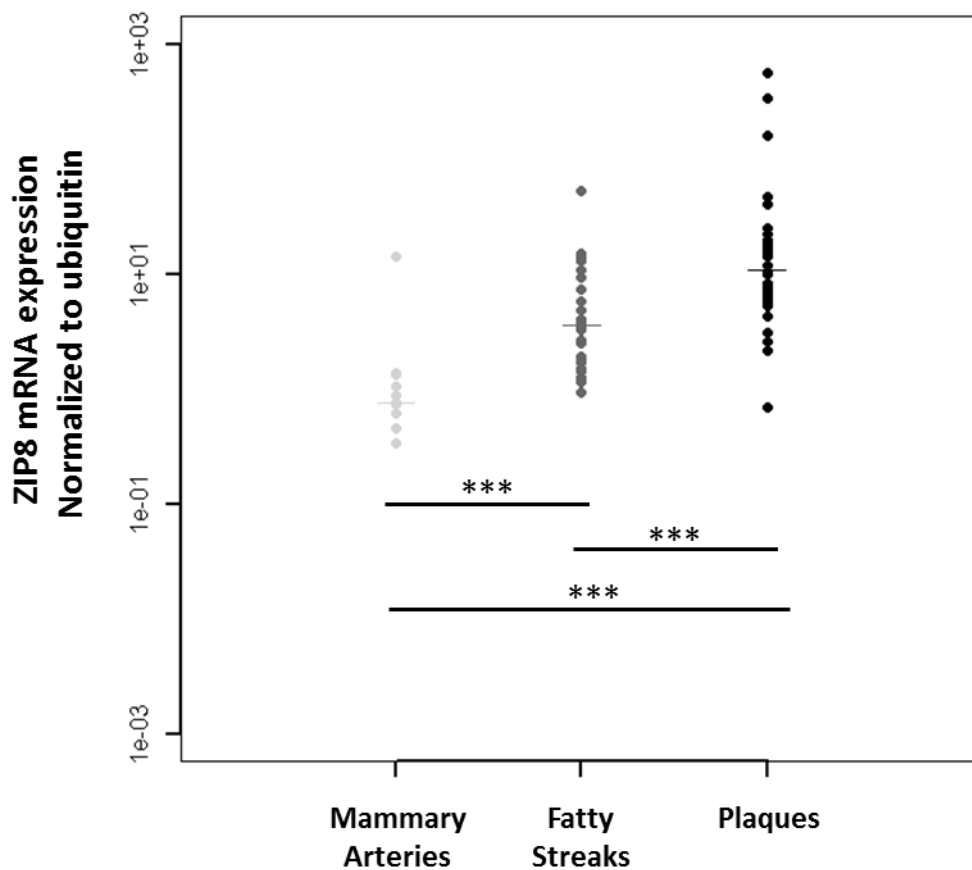


Figure 6. ZIP8 expression in normal vascular tissue (mammary arteries, n=11), fatty streaks (n=27) and plaques from human carotids (n=33). Expression of the ZIP8 measured by RT-qPCR shows that Zip8 expression is significantly higher in plaques and weaker in normal vascular tissues (mammary arteries).

These results indicate that ZIP8 expression increases along with the plaque formation.

II.3 ZIP8 expression in human carotid plaque according to the smoking status

The ZIP8 network (Figure 1) showed an indirect correlation between ZIP8 and smoking, hence we investigated by RT-qPCR the expression of ZIP8 in human atherosclerotic carotid tissues comparing non smokers (n=13) with ex-smokers (n=6) and smokers (n=10). No Statistically significant difference in ZIP8 expression was found between these groups (Figure 7).

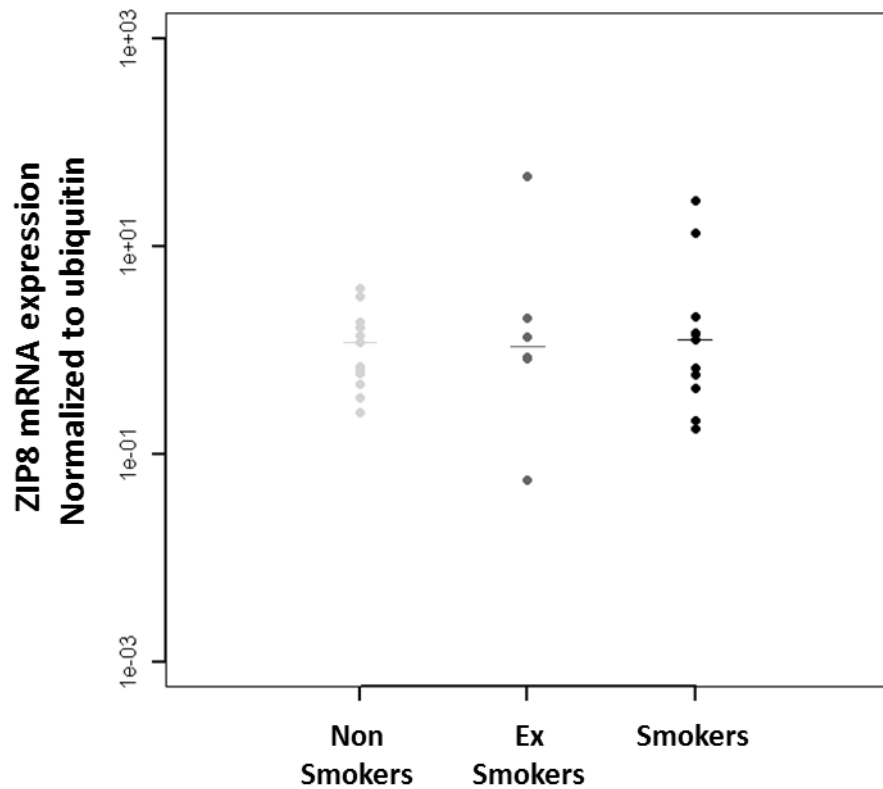


Figure 7. ZIP8 mRNA expression in carotids from non-smokers (NS, n=13), ex-smokers (ES, n=6) and smokers (S, n=10). RT-qPCR of the ZIP8 transcript show no difference in expression regarding the smoking status.

II.4 Investigating the interdependence of ZIP8 and Gas6 expression in human macrophages

The statistically predicted pathway linking ZIP8 to the number of plaques and smoking involved another major candidate, Gas6, which expression was found correlated to smoking. To analyze a possible interaction between Gas6 and ZIP8 expression in human macrophages, we performed siRNA mediated knockdown of either Gas6 or ZIP8 and then explored whether it had an effect on the expression of other members of the pathway. Macrophages silenced for ZIP8 did not show any significant change in Gas6 expression. Conversely, macrophages knocked down for Gas6 did not show any changes of ZIP8 expression (Figure 8).

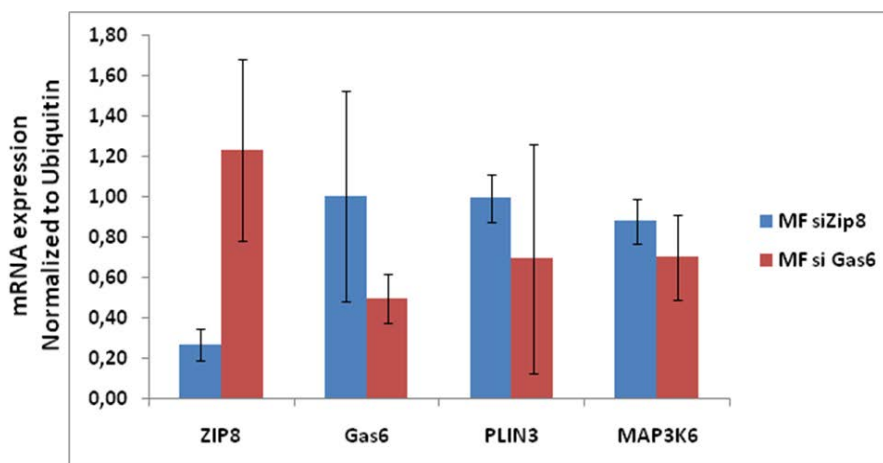


Figure 8. Members of the Gas6/ZIP8 statistically predicted pathways upon either ZIP8 or Gas6 knockdown in human macrophages. Results are expressed as fold change +/- standard deviation of $n=5$ different individuals (delta delta Ct method). Zip8 knockdown did not alter Gas6 expression or other tested genes from the statistically predicted network (PLIN3, MAP3K6).

II.5 Methallothionein and ZIP8 expression

Proteins belonging to the methallothionein family are among the first line of defense against heavy metal exposure, thus we decided to investigate in our GHS data whether the expression of ZIP8 could affect their expression. Out of the 15 metallothionein which expression was detectable, we selected MT1E, MT1A, MT1E, MT1G and MT2A on the following criteria: statistically significant p-value for both ZIP8 probes, and statistically significant p-value for either smoking or the number of plaques. We additionally selected MT1X, as it had the highest correlation to ZIP8 expression, despite the lack of significant p-value for a link to smoking or number of plaques (Table 1).

| | SLC39A8 (ZIP8) | | Plaques | Smoking |
|--------------|-----------------------|----------------|----------------|----------------|
| | Probe1 | Probe2 | | |
| MT1E | 2.2E-11 | 1.1E-04 | 3.2E-03 | 5.6E-01 |
| MT1A | 7.9E-02 | 9.4E-04 | 3.0E-01 | 6.8E-07 |
| MT1B | 5.6E-01 | 5.6E-02 | 8.0E-01 | 9.0E-01 |
| MT1E | 2.5E-02 | 6.3E-02 | 3.6E-04 | 5.1E-01 |
| MT1F | 9.1E-02 | 7.4E-08 | 4.6E-06 | 2.9E-01 |
| MT1G | 1.9E-13 | 4.6E-06 | 3.8E-02 | 5.1E-06 |
| MT1H | 1.1E-01 | 8.8E-01 | 2.1E-03 | 8.9E-01 |
| MT1H | 6.1E-02 | 1.3E-01 | 7.8E-01 | 9.8E-01 |
| MT1M | 3.7E-04 | 3.0E-03 | 6.7E-01 | 8.9E-01 |
| MT1M | 8.9E-03 | 8.9E-04 | 7.5E-02 | 1.3E-01 |
| MT1H | 3.4E-01 | 7.6E-01 | 2.3E-01 | 2.8E-01 |
| MT1X | 2.2E-12 | 6.0E-18 | 9.4E-01 | 1.8E-01 |
| MT2A | 8.5E-03 | 5.8E-02 | 3.0E-01 | 1.5E-11 |
| MT3 | 4.5E-03 | 8.1E-03 | 6.9E-01 | 4.5E-01 |
| MT4 | 1.8E-01 | 1.8E-03 | 1.4E-01 | 9.6E-01 |
| MT1P | 1.0E-04 | 7.4E-01 | 4.8E-02 | 2.8E-04 |
| MTL5 | 3.6E-01 | 1.5E-03 | 3.8E-01 | 1.5E-01 |
| MTL5 | 3.0E-01 | 3.2E-01 | 5.9E-01 | 6.7E-01 |
| MT1DP | 3.8E-05 | 4.3E-01 | 2.9E-01 | 1.8E-01 |

Table 1. Correlations between metallothioneins, ZIP8, number of plaques and smoking status from GHS data (p-value). The selected methallothioneins have been marked in red. Significant p-values are in bold.

Next, we investigated whether or not the cigarette smoke condensate (incubation for 24h; 4, 10, 20, 40, 80 $\mu\text{g/ml}$) was able to affect MT1A, MT1E, MT1G and MTX expression in HAECs. As shown in Figure 9 the expression of these selected metallothioneins did not vary at any studied concentration. The possible explanation could be that the metallic ions, being hydrophilic, were probably absent in the cigarette smoke condensate diluted in DMSO.

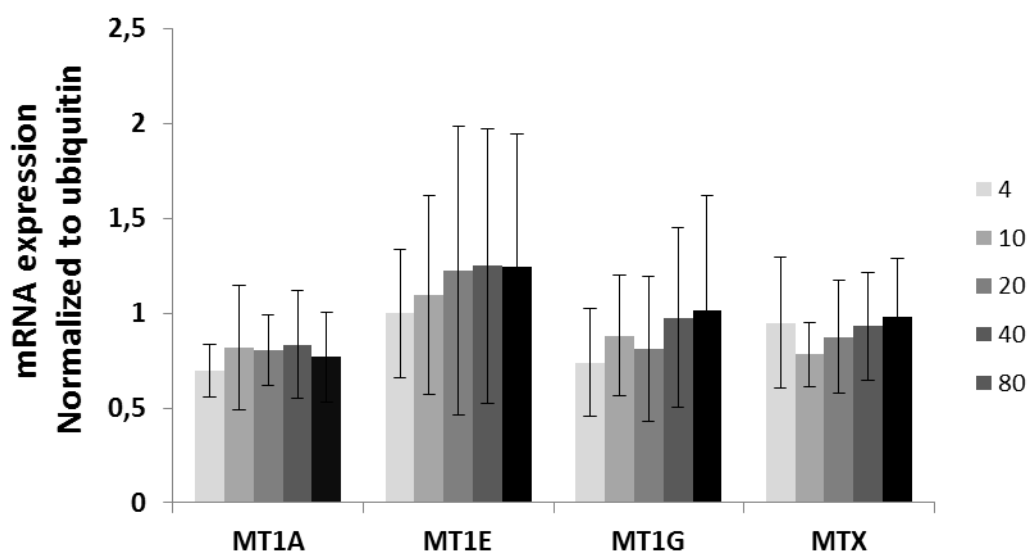


Figure 9. Metallothioneins expression in HAECs following cigarette smoke condensate stimulation. Expression of MT1A, MT1E, MT1G and MTX after cigarette condensate stimulation (incubation for 24h at 4, 10, 20, 40 or 80 $\mu\text{g/ml}$) studied by RT-qPCR. Results are expressed as fold change \pm standard deviation of $n=4$ HAECs preparations. No change of expression of those metallothioneins was observed following stimulation.

During plaque formation, monocytes migrate into the plaque and differentiate into macrophages. As described previously, ZIP8 expression increases in monocytes activated with antibody coated beads (a model of antigen presenting cells)¹⁰. As monocyte activation is a pre-requisite for monocyte migration in the vascular wall and subsequent differentiation into macrophages, we hypothesized that ZIP8 expression could be increased during this process. Furthermore, such increase in ZIP8 expression could affect that of methallothioneins which expressions were correlated to it in GHS data (Table 1). Hence, after isolating monocytes from human blood (buffy coat, n=4) we measured ZIP8 and methallothionein expression before and after differentiation into macrophages (Figure 10).

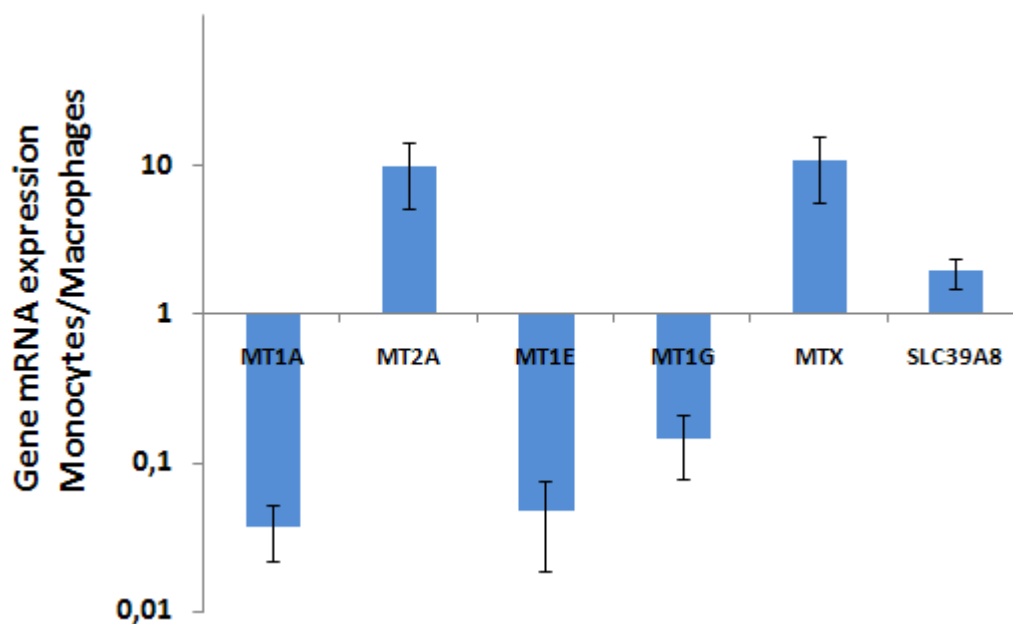


Figure 10. mRNA expression of ZIP8 (SLC39A8) and metallothioneins in monocytes and monocyte-derived macrophages. Expression of SLC39A8 and its correlated metallothioneins was measured by RT-qPCR and the results are log[mean fold change] +/- standard deviation of n=4 monocytes preparations. The results show an increase in SLC39A8, MT2A and MTX expression and a decrease in MT1A, MT1E and MT1G after monocyte differentiation into macrophages.

As expected ZIP8 expression was significantly increased in monocyte-derived macrophages when compared to monocytes (2-fold). All tested methallothioneins also showed a change of expression with MT1A, MT1E and MT1G being downregulated and MT2A and MTX being up-regulated in macrophages as compared to monocytes.

III Discussion

Our results show that ZIP8 is expressed ubiquitously within the atherosclerotic plaque with a higher expression in macrophages/foam cells and dedifferentiated smooth muscle cells. Interestingly ZIP8 expression was higher in plaques as compared to fatty streaks and it was the weakest in normal vascular tissue (mammary artery). The increased expression in plaques and fatty streaks of Zip8, as shown in Figure 10, can be explained by their high macrophages content, which increases along with the plaque formation. Thus ZIP8 expression could be a marker for macrophages accumulation within the vascular wall and probably to the plaque size. Other cell types, such as endothelial cells and smooth muscle cells may also contribute to the accumulation of ZIP8 as its expression is increased by inflammation¹².

Circulating levels of metal ions is correlated with carotid atherosclerosis in elderly²⁸ and many divalent metallic ions, such as iron, zinc, copper and cadmium accumulate within the atherosclerotic plaques²⁹. ZIP8 is one of the main transporters for zinc and cadmium and, as showed above, is increased along atherosclerosis development (Figure 6). Therefore, an increase in ZIP8 expression may, at least partly, explain the increase of heavy metal deposition within the carotid plaques. It has been already demonstrated that smoking aggravates atherosclerosis, as smokers have more plaques which are additionally frequently larger than in non smokers^{30,31}. As cigarette smoke contains cadmium¹³, which is mainly imported in the cells through the ZIP8 transporter⁵, smokers are likely to have substantially more cadmium accumulated within their atherosclerotic plaques. Indeed, such increase in cadmium has already been observed in non atherosclerotic aorta of smokers³². Cadmium accumulates probably mainly in macrophages, as they show the highest Zip 8 expression, as shown by our experiments in Fig 5. Since cadmium is cytotoxic, its accumulation may result in accelerated macrophage death within the plaques. A first consequence of cadmium induced macrophage death could be Cd²⁺ release in the neointimal space leading resulting in death of the nearby cells and thus inducing plaque instability. Another consequence could be an increased lipid release in the neointimal space, accelerating the formation of the necrotic core and thus increasing the plaque size. Taken together, the preliminary results from our studies suggest that the ZIP8 transporter is likely involved in aggravation of atherosclerosis *via* cadmium accumulation.

An appropriate animal model could help to confirm the role of Zip8 in atherosclerosis. Recently, ZIP8 conditional knockout mice were used to study osteoarthritis, which is also an inflammation driven disease³³. One way to confirm a role for ZIP8, as a mediator for cigarette induced cadmium accumulation in the atherosclerotic plaques, could be the use of a double knockout mice ApoE^{-/-}/ZIP8^{-/-} under high fat diet and exposed or not to cigarette smoke. By measuring cadmium accumulation within the aortic plaques of the two groups of mice, one could determine if ZIP8 knockdown reduces cadmium accumulation and plaque size, as hypothesized. As cigarette smoke contains several metallic divalent ions (cobalt, mercury, nickel, lead, cadmium...)¹³, the specificity of the cadmium toxicity could be assessed in these atherosclerosis prone mice exposed to cigarette smoke by feeding them either with zinc supplemented diet or a control diet. As zinc is a direct

competitor for cadmium for the ZIP8 transporter³, complementing food with zinc should partially rescue the phenotype induced by cadmium from cigarette smoke (Figure 11).

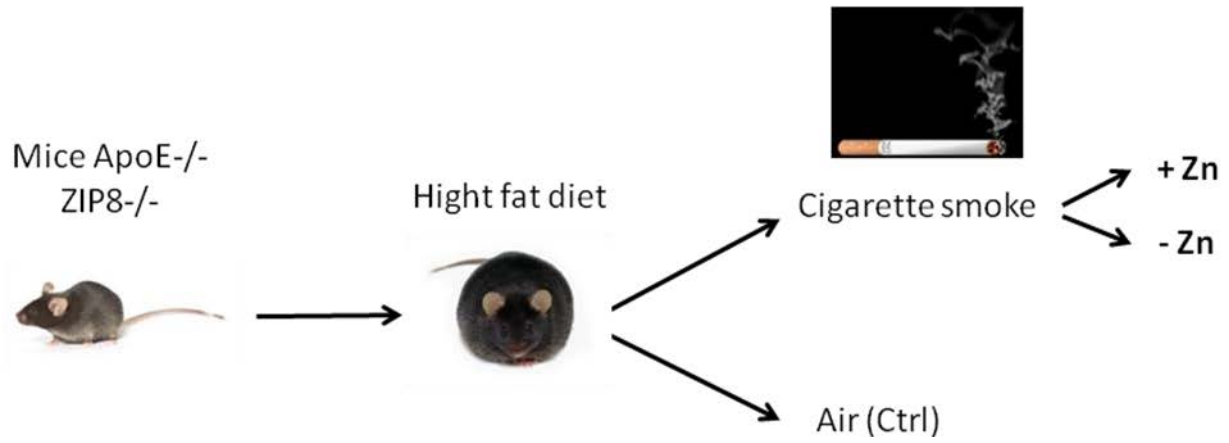


Figure 11. experimental design to test the implication of ZIP8 in cigarette smoke dependent accumulation within the atherosclerotic plaque in mice.

Références

1. Tjwa M, Moons L, Lutgens E. Pleiotropic role of growth arrest-specific gene 6 in atherosclerosis. *Curr Opin Lipidol*. 2009;20(5):386-392. doi:10.1097/MOL.0b013e328330982e.
2. Begum NA, Kobayashi M, Moriwaki Y, Matsumoto M, Toyoshima K, Seya T. Mycobacterium bovis BCG Cell Wall and Lipopolysaccharide Induce a Novel Gene, BIGM103, Encoding a 7-TM Protein: Identification of a New Protein Family Having Zn-Transporter and Zn-Metalloprotease Signatures. *Genomics*. 2002;80(6):630-645. doi:10.1006/geno.2002.7000.
3. He L, Girijashanker K, Dalton TP, et al. ZIP8, member of the solute-carrier-39 (SLC39) metal-transporter family: characterization of transporter properties. *Mol Pharmacol*. 2006;70(1):171-180. doi:10.1124/mol.106.024521.
4. Eide DJ. The SLC39 family of metal ion transporters. *Pflug Arch Eur J Physiol*. 2004;447(5):796-800. doi:10.1007/s00424-003-1074-3.
5. Jenkitkasemwong S, Wang C-Y, Mackenzie B, Knutson MD. Physiologic implications of metal-ion transport by ZIP14 and ZIP8. *Biometals Int J Role Met Ions Biol Biochem Med*. 2012;25(4):643-655. doi:10.1007/s10534-012-9526-x.
6. Bonaventura P, Benedetti G, Albarède F, Miossec P. Zinc and its role in immunity and inflammation. *Autoimmun Rev*. 2015;14(4):277-285. doi:10.1016/j.autrev.2014.11.008.
7. Haase H, Rink L. Multiple impacts of zinc on immune function. *Met Integr Biometal Sci*. 2014;6(7):1175-1180. doi:10.1039/c3mt00353a.
8. Giaginis C, Gatzidou E, Theocharis S. DNA repair systems as targets of cadmium toxicity. *Toxicol Appl Pharmacol*. 2006;213(3):282-290. doi:10.1016/j.taap.2006.03.008.
9. Bertin G, Averbeck D. Cadmium: cellular effects, modifications of biomolecules, modulation of DNA repair and genotoxic consequences (a review). *Biochimie*. 2006;88(11):1549-1559. doi:10.1016/j.biochi.2006.10.001.
10. Aydemir TB, Liuzzi JP, McClellan S, Cousins RJ. Zinc transporter ZIP8 (SLC39A8) and zinc influence IFN-gamma expression in activated human T cells. *J Leukoc Biol*. 2009;86(2):337-348. doi:10.1189/jlb.1208759.
11. Liu M-J, Bao S, Gálvez-Peralta M, et al. ZIP8 regulates host defense through zinc-mediated inhibition of NF- κ B. *Cell Rep*. 2013;3(2):386-400. doi:10.1016/j.celrep.2013.01.009.
12. Napolitano JR, Liu M-J, Bao S, et al. Cadmium-mediated toxicity of lung epithelia is enhanced through NF- κ B-mediated transcriptional activation of the human zinc transporter ZIP8. *Am J Physiol Lung Cell Mol Physiol*. 2012;302(9):L909-L918. doi:10.1152/ajplung.00351.2011.
13. Fresquez MR, Pappas RS, Watson CH. Establishment of toxic metal reference range in tobacco from US cigarettes. *J Anal Toxicol*. 2013;37(5):298-304. doi:10.1093/jat/bkt021.
14. Dong F, Guo F, Li L, et al. Cadmium induces vascular permeability via activation of the p38 MAPK pathway. *Biochem Biophys Res Commun*. 2014;450(1):447-452. doi:10.1016/j.bbrc.2014.05.140.

15. Szuster-Ciesielska A, Lokaj I, Kandefer-Szerszeń M. The influence of cadmium and zinc ions on the interferon and tumor necrosis factor production in bovine aorta endothelial cells. *Toxicology*. 2000;145(2-3):135-145.
16. Yamamoto C, Kaji T, Sakamoto M, Kozuka H. Cadmium stimulation of plasminogen activator inhibitor-1 release from human vascular endothelial cells in culture. *Toxicology*. 1993;83(1-3):215-223. doi:10.1016/0300-483X(93)90103-Y.
17. Washington B, Williams S, Armstrong P, Mtshali C, Robinson JT, Myles EL. Cadmium toxicity on arterioles vascular smooth muscle cells of spontaneously hypertensive rats. *Int J Environ Res Public Health*. 2006;3(4):323-328.
18. Fujiwara Y, Tsumura N, Yamamoto C, Kaji T. Differential effects of cadmium on proteoglycan synthesis of arterial smooth muscle cells: increase in small dermatan sulfate proteoglycans, biglycan and decorin, in the extracellular matrix at low cell density. *Toxicology*. 2002;170(1-2):89-101.
19. Hernández M, Macia M. Free peripheral sulfhydryl groups, CD11/CD18 integrins, and calcium are required in the cadmium and nickel enhancement of human-polymorphonuclear leukocyte adherence. *Arch Environ Contam Toxicol*. 1996;30(4):437-443.
20. Fujiwara Y, Watanabe S, Kaji T. Promotion of cultured vascular smooth muscle cell proliferation by low levels of cadmium. *Toxicol Lett*. 1998;94(3):175-180.
21. Kaji T, Yamamoto C, Miyajima S, et al. Vascular smooth muscle cells in culture are highly sensitive to cadmium cytotoxicity without species-related differences: comparison with Chang liver cells. *Biol Pharm Bull*. 1995;18(10):1392-1395.
22. Kaji T, Suzuki M, Yamamoto C, et al. Sensitive response of cultured vascular smooth-muscle cells to cadmium cytotoxicity: comparison with cultured vascular endothelial cells and kidney epithelial LLC-PK1 cells. *Toxicol Lett*. 1996;89(2):131-137.
23. Kaji T, Suzuki M, Yamamoto C, et al. Sensitive response of cultured vascular smooth-muscle cells to cadmium cytotoxicity: comparison with cultured vascular endothelial cells and kidney epithelial LLC-PK1 cells. *Toxicol Lett*. 1996;89(2):131-137.
24. Skoczynska A, Martynowicz H. The impact of subchronic cadmium poisoning on the vascular effect of nitric oxide in rats. *Hum Exp Toxicol*. 2005;24(7):353-361. doi:10.1191/0960327105ht536oa.
25. Kolluru GK, Tamilarasan KP, Geetha Priya S, Durgha NP, Chatterjee S. Cadmium induced endothelial dysfunction: consequence of defective migratory pattern of endothelial cells in association with poor nitric oxide availability under cadmium challenge. *Cell Biol Int*. 2006;30(5):427-438. doi:10.1016/j.cellbi.2006.02.002.
26. Knoflach M, Messner B, Shen YH, et al. Non-toxic cadmium concentrations induce vascular inflammation and promote atherosclerosis. *Circ J Off J Jpn Circ Soc*. 2011;75(10):2491-2495.
27. Messner B, Knoflach M, Seubert A, et al. Cadmium is a novel and independent risk factor for early atherosclerosis mechanisms and in vivo relevance. *Arterioscler Thromb Vasc Biol*. 2009;29(9):1392-1398. doi:10.1161/ATVBAHA.109.190082.
28. Lind PM, Olsén L, Lind L. Circulating levels of metals are related to carotid atherosclerosis in elderly. *Sci Total Environ*. 2012;416:80-88. doi:10.1016/j.scitotenv.2011.11.064.

29. Stadler N, Lindner RA, Davies MJ. Direct detection and quantification of transition metal ions in human atherosclerotic plaques: evidence for the presence of elevated levels of iron and copper. *Arterioscler Thromb Vasc Biol.* 2004;24(5):949-954. doi:10.1161/01.ATV.0000124892.90999.cb.
30. Andersson J, Sundström J, Kurland L, et al. The carotid artery plaque size and echogenicity are related to different cardiovascular risk factors in the elderly: the Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) study. *Lipids.* 2009;44(5):397-403. doi:10.1007/s11745-009-3281-y.
31. Falk E, Shah PK, Fuster V. Coronary Plaque Disruption. *Circulation.* 1995;92(3):657-671. doi:10.1161/01.CIR.92.3.657.
32. Abu-Hayyeh S, Sian M, Jones KG, Manuel A, Powell JT. Cadmium accumulation in aortas of smokers. *Arterioscler Thromb Vasc Biol.* 2001;21(5):863-867.
33. Kim J-H, Jeon J, Shin M, et al. Regulation of the catabolic cascade in osteoarthritis by the zinc-ZIP8-MTF1 axis. *Cell.* 2014;156(4):730-743. doi:10.1016/j.cell.2014.01.007.

Annex 2. Highlights for the ZIP8 project

Taken together, our results allow us to draw the following conclusions:

- ZIP8 is ubiquitous within the atherosclerotic plaque, with a higher expression in macrophages/foam cells and dedifferentiated smooth muscle cells.
- ZIP8 expression is higher in human monocyte-derived macrophages than in monocytes and HAECs
- ZIP8 expression is increased during differentiation from monocyte to macrophages.
- ZIP8 expression is higher in human carotid plaques than in fatty streaks and normal vascular tissues (mammary arteries)
- ZIP8 expression is not increased by cigarette smoking or cigarette smoke condensate
- ZIP8 expression is correlated to MT1A, MT1G, MT1E, MT1X and MT2A
- ZIP8 related methallothioneins MT1E, MT1G, MT1A are decreased, MT2A and MTX are increased during differentiation from monocyte to macrophages.
- Preliminary results do not show interdependence between ZIP8 and Gas6 expression in human macrophages

Study of SLC39A8 impact on the pathology of atherosclerosis is an ongoing project by the group of Dr Verdugo in collaboration with our group in Paris, as it is the thesis subject of Pamela Andrea Lopez Contreras. This thesis project mainly focus on performing transcriptomic analyses of ZIP8 knockdown in human macrophages after cadmium exposure. It aims at better characterizing how heavy metals, and cadmium in particular, accumulates in atherosclerotic lesions progression.

Annex 3. Highlights of the SASH1 project

The following conclusions can be drawn from this study on SASH1 as a link between smoking and atherosclerosis :

-SASH1 is expressed in human endothelial cells, smooth muscle cells, monocytes and monocytes derived macrophages, and is enriched at the leading edge of the lamellipodia.

-SASH1 is mostly cytoplasmic under normal culture condition in human aortic endothelial cells.

-SASH1 expression is increased in the carotids from smokers when compared to non-smokers, confirming the observations already made in the monocytes.

-SASH1 knockdown affects pathways involved in the cell cycle progression and DNA repair possibly through a p53/FOXM1 dependent fashion.

-SASH1 endogenous protein co-precipitates with RPS25 a protein that regulates p53 activity and TPM1, a protein linked to the polymerization of the actin-cytoskeleton.

-SASH1 affects migration, proliferation and angiogenesis of human aortic endothelial cells

Taken together these data show that SASH1 has the potential to affect the plaque formation and stability in smokers through its impact on the endothelial cell function, and therefore open the perspective to use SASH1 as a biomarker for atherosclerosis.

Annex 4. Verdugo et al 2013

Graphical Modeling of Gene Expression in Monocytes Suggests Molecular Mechanisms Explaining Increased Atherosclerosis in Smokers

Ricardo A. Verdugo^{1‡a}, Tanja Zeller², Maxime Rotival^{1‡b}, Philipp S. Wild^{3,4}, Thomas Münzel³, Karl J. Lackner⁵, Henri Weidmann¹, Ewa Ninio¹, David-Alexandre Trégouët¹, François Cambien¹, Stefan Blankenberg², Laurence Tiret^{1*}

1INSERM UMR_S 937, Pierre and Marie Curie University, Paris, France, **2**University Heart Center Hamburg, Department of General and Interventional Cardiology, Hamburg, Germany, **3**Department of Medicine II, University Medical Center Mainz, Mainz, Germany, **4**Clinical Epidemiology, Center for Thrombosis and Haemostasis, University Medical Center Mainz, Mainz, Germany, **5**Institute for Clinical Chemistry and Laboratory Medicine, University Medical Center Mainz, Mainz, Germany

Abstract

Smoking is a risk factor for atherosclerosis with reported widespread effects on gene expression in circulating blood cells. We hypothesized that a molecular signature mediating the relation between smoking and atherosclerosis may be found in the transcriptome of circulating monocytes. Genome-wide expression profiles and counts of atherosclerotic plaques in carotid arteries were collected in 248 smokers and 688 non-smokers from the general population. Patterns of co-expressed genes were identified by Independent Component Analysis (ICA) and network structure of the pattern-specific gene modules was inferred by the PC-algorithm. A likelihood-based causality test was implemented to select patterns that fit models containing a path “smoking→gene expression→plaques”. Robustness of the causal inference was assessed by bootstrapping. At a FDR ≤ 0.10 , 3,368 genes were associated to smoking or plaques, of which 93% were associated to smoking only. *SASH1* showed the strongest association to smoking and *PPARG* the strongest association to plaques. Twenty-nine gene patterns were identified by ICA. Modules containing *SASH1* and *PPARG* did not show evidence for the “smoking→gene expression→plaques” causality model. Conversely, three modules had good support for causal effects and exhibited a network topology consistent with gene expression mediating the relation between smoking and plaques. The network with the strongest support for causal effects was connected to plaques through *SLC39A8*, a gene with known association to HDL-cholesterol and cellular uptake of cadmium from tobacco, while smoking was directly connected to *GAS6*, a gene reported to have anti-inflammatory effects in atherosclerosis and to be up-regulated in the placenta of women smoking during pregnancy. Our analysis of the transcriptome of monocytes recovered genes relevant for association to smoking and atherosclerosis, and connected genes that before, were only studied in separate contexts. Inspection of correlation structure revealed candidates that would be missed by expression-phenotype association analysis alone.

Citation: Verdugo RA, Zeller T, Rotival M, Wild PS, Münzel T, et al. (2013) Graphical Modeling of Gene Expression in Monocytes Suggests Molecular Mechanisms Explaining Increased Atherosclerosis in Smokers. PLoS ONE 8(1): e50888. doi:10.1371/journal.pone.0050888

Editor: Frank Emmert-Streib, Queen's University Belfast, United Kingdom

Received: June 14, 2012; **Accepted:** October 24, 2012; **Published:** January 23, 2013

Copyright: © 2013 Verdugo et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The Gutenberg Health Study is funded through the government of Rheinland-Pfalz (“Stiftung Rheinland Pfalz für Innovation”, contract number AZ 961-386261/733), the research programs “Wissen schafft Zukunft” and “Schwerpunkt Vasculäre Prävention” of the Johannes Gutenberg-University of Mainz and its contract with Boehringer Ingelheim and PHILIPS Medical Systems including an unrestricted grant for the Gutenberg Health Study. Specifically, the research reported in this article was supported by the National Genome Network “NGFNplus” by the Federal Ministry of Education and Research, Germany (contract number project A3 01GS0833) and by a joint funding from the Federal Ministry of Education and Research, Germany (contract BMBF 01KU0908A) and from the Agence Nationale de la Recherche, France (contract ANR 09 GENO 106 01) for the project CARDomics. Statistical analyses benefit from the C2BIG computing centre funded by the Fondation pour la Recherche Médicale and the Région Ile-de-France. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: laurence.tiret@upmc.fr

‡a Current address: Human Genetics Program, School of Medicine, University of Chile, Santiago, Chile

‡b Current address: Imperial College London, Hammersmith Hospital, London, United Kingdom

Introduction

Smoking is a major risk factor for atherosclerosis and its complications, particularly coronary artery disease (CAD) and peripheral arterial disease [1–3]. Pathophysiological mechanisms by which smoking promotes atherogenesis are relatively well known, in particular through alterations of lipid metabolism [4,5] and endothelial function [6]. However, the molecular mechanisms by which smoking exerts its adverse effects at the cellular level are less documented. The advent of transcriptomic studies allowing

investigation of all the genes expressed in a given type of cell has opened a new window for exploring in a global way the biological mechanisms underlying pathophysiological conditions. Using such transcriptomic approach, widespread perturbation of gene expression by smoking has been recently shown in whole blood [7], circulating lymphocytes [8], and monocytes [9] of humans.

Increasing evidence supports the hypothesis that oxidative stress and activation of the immune system provide a pathophysiological link between cigarette smoking and CAD [10,11]. Monocytes are

Table 1. Characteristics of the Gutenberg Health Study population.

| Characteristics | Variable name | Females | Males | p-value† |
|--------------------------------------|---------------|---------------|---------------|----------|
| Number of individuals | N | 522 (55.77%) | 414 (44.23%) | |
| Age (years) | age | 54.6 (0.49) | 54.2 (0.53) | 0.5825 |
| Body mass index (kg/m ²) | bmi | 26.2 (0.22) | 27.3 (0.2) | 0.0003 |
| Triglycerides (mg/dl) (log) | TRIGLY | 4.63 (0.019) | 4.78 (0.025) | 4.6E-06 |
| HDL cholesterol (mg/dl) | HDL_CHOL | 63.8 (0.69) | 50.6 (0.6) | 2.3E-40 |
| LDL cholesterol (mg/dl) | LDL_CHOL | 144 (1.7) | 142 (1.8) | 0.6472 |
| Systolic blood pressure (mmHg) | sbp | 129 (0.83) | 134 (0.77) | 2.2E-06 |
| Diabetes (type I or II) | diabetes | 30 (5.7%) | 30 (7.2%) | 0.3538 |
| Current smoker | smoking | 115 (22%) | 133 (32.1%) | 0.0005 |
| C-reactive protein (mg/l) (log) | CRP | 0.932 (0.033) | 0.832 (0.038) | 0.0455 |
| Homocysteine (mmol/l) (log) | HCY | 2.25 (0.013) | 2.44 (0.015) | 1.3E-21 |
| Myeloperoxidase (pmol/l) (log) | MPO | 5.76 (0.015) | 5.7 (0.02) | 0.0157 |

† p-values calculated from a χ^2 test for smoking and diabetes (number of subjects), and from an F test for all others. Standard errors or percents of individuals are in parenthesis.

doi:10.1371/journal.pone.0050888.t001

key cells of the immune system involved in the inflammatory response to external agents. We hypothesized that the effect of smoking on atherosclerosis might be reflected by perturbation of gene expression in circulating monocytes and that it might be possible to identify gene networks causally involved in the relationship linking smoking to atherosclerosis.

Questions about causal effects in observational studies can be addressed by statistical methods that can translate statements about correlations and conditional independencies into structural equations [12] or Bayesian Networks [13]. Implementation of both techniques, however, can be difficult when the number of variables is large and inference of the “true” network that generated the data may not be feasible, even with large sample sizes. In addition, current implementations of Bayesian Network inference restrict to systems of Gaussian only, binomial (or multinomial) only or hybrids where binomials can only precede but not be caused by Gaussian variables [14]. A third class of methods based on information theory has been developed for the problem of identification of large networks of direct gene interactions, which does not rely on the correct specification of distribution functions, but where such interactions do not have a causal interpretation (e.g. ARACNE [15], SA-CLR [16], and Parmigene [17], among others). For the problem under study here, an approach that is not limited in the types of variables that can be modeled, i.e. binary, continuous, and counts, or in how they may be associated, that allows inferences about causal relations, and that can deal with large number of variables was needed.

The objective of the present study was to identify groups of genes that may help explain the causal effect of smoking on extent of atherosclerosis. For this purpose, genomewide gene expression in monocytes was modeled as a molecular phenotype potentially linking smoking to carotid atherosclerosis. Data were obtained from the Gutenberg Health Study (GHS), a community-based project primarily aimed at improving cardiovascular risk prediction [9]. We devised a stepwise approach to: 1. Identify patterns of expression associated to smoking and/or atherosclerosis using Independent Component Analysis (ICA); 2. Select expression patterns showing relatively high support for a causal role in the mediation between smoking and atherosclerosis using graphical modeling and Bayesian Network (BN) inference; 3. For patterns

compatible with a potential causal role, infer the network skeleton connecting smoking, genes and atherosclerosis.

Our approach identified three gene networks that were compatible with a causal effect of gene expression mediating the relation between smoking and atherosclerosis where a few genes are candidates for mediating the perturbing effect of smoking in these networks. By performing causal inference on independent patterns of expression instead of the expression of a single gene, we not only dramatically reduced the space of models to be tested, but also applied BN inference in a space that is less prone to the effects of hidden variables. We restricted to detecting classes of best fitting models rather than a single causal model and present arguments for why some causal models may be favored in this study. We also provide cautionary statements to avoid misinterpretation of the reported causal models.

Results

Smoking and extent of atherosclerosis in the GHS cohort

Association between smoking habits and atherosclerosis was investigated in a cohort of subjects of both sexes aged 35 to 74 years who participated in the GHS [9]. Study participants were classified into current (≥ 1 cigarette/day) smokers ($n = 248$) and nonsmokers ($n = 688$). Occasional smokers ($n = 42$) and ex-smokers ($n = 547$) were excluded from the study (Methods). Characteristics of the study population are given in Table 1.

Atherosclerosis extent was defined as the total number of atherosclerotic plaques measured in the two carotid arteries by ultrasound echography (Methods). Carotid intima-media thickness and carotid plaque are well-recognized markers of subclinical atherosclerosis [18] which are influenced by smoking [19]. The number of plaques observed per person ranged from 0 to 11, with a skewed distribution and an average of 0.72 plaques (variance 1.95). The prevalence of atherosclerosis, defined as the presence of at least one plaque, was 31.2% in this middle-aged population. The prevalence and the extent of atherosclerosis were higher in men than in women and in smokers than in non-smokers (Table 2).

In the following, the phenotype considered was the number of atherosclerotic plaques, referred to as “plaques”. Because the distribution showed overdispersion, a negative binomial distribution was used for modeling plaques as a function of covariables (see

Table 2. Mean number of carotid atherosclerotic plaques by sex and smoking status.

| Sex | Smoking | N | Plaques Average [†] | Plaques>0 | |
|---------|------------|-----|------------------------------|----------------|------|
| | | | | n [‡] | % |
| Females | Nonsmokers | 407 | 0.51 | 107 | 26.3 |
| | Smokers | 115 | 0.56 | 30 | 26.1 |
| | All | 522 | 0.52 | 137 | 26.2 |
| Males | Nonsmokers | 281 | 0.74 | 93 | 33.1 |
| | Smokers | 133 | 1.42 | 62 | 46.6 |
| | All | 414 | 0.96 | 155 | 37.4 |
| Both | Nonsmokers | 688 | 0.61 | 200 | 29.1 |
| | Smokers | 248 | 1.02 | 92 | 37.1 |
| | All | 936 | 0.72 | 292 | 31.2 |

[†]average number of plaques per individual;
[‡]number of individuals with at least 1 plaque.
doi:10.1371/journal.pone.0050888.t002

Text S1). The major determinants of plaques were age, sex and smoking, which all together explained 30% of the variability of plaques. The effect of smoking on plaques was stronger in men than in women although the significance of the interaction test was borderline ($p = 0.027$). Additional cardiovascular factors tested for association resulted in a modest increase of the explained variance of plaques (from 30% to 34%) (Table S1).

Gene expression in monocytes is associated to smoking and plaques

The analysis workflow of expression data is outlined in Figure 1. Expression of 18,364 genes was detected in total RNA from circulating monocytes by 23,214 probes in the *Illumina* Human HT-12 BeadChip (Methods). Association of probe expression level with smoking or plaques (log-transformed) was assessed by linear model adjusted for age and sex, as well as for the 6 first singular value decomposition (SVD) components of the expression matrix taken as surrogate variables for technical sources of variability (Methods and Text S1).

In a first step, we identified genes whose expression was associated to either smoking or plaques by univariate analysis. At $FDR \leq 0.1$, we found 3,774 probes (3,062 distinct genes)

associated to smoking. The list of smoking-associated genes was enriched in three “biological processes” in the Gene Ontology (GO) database: platelet activation, interferon-gamma-mediated signaling pathway, and Toll signaling pathway (Table 3). Association between plaques and gene expression was much less prevalent than with smoking, with only 258 probes (236 distinct genes) associated to plaques at $FDR \leq 0.1$. No GO terms were significantly enriched for genes associated to plaques.

Table 4 shows the 10 genes with strongest association to smoking or plaques, respectively. The whole list of associated genes to either phenotype is given in Table S2. There were 2 members common to both top 10 gene lists, *SASH1* and *PTGDS*. In addition to *SASH1* and *PTGDS*, 4 genes of the top 10 list for smoking were ranked among the top 100 genes for plaques: *FUCA1*, *LOC157627*, *MMP25* and *PIDI* (Table S2). Smoking was associated to a much larger variability of gene expression (from 35% to 15% for the top 10 smoking-related genes) than plaques (from 5% to 2% for the top 10 plaques-related genes). The correlation of r_D^2 values for smoking and plaques across all genes was 0.4, indicating a strong association between smoking and plaques effects on gene expression as a result of their confounding effects.

Comparison of smoking-associated gene expressions in monocytes and lymphocytes

We investigated the robustness of the association between gene expression and smoking by comparing the list of smoking-associated genes in monocytes to a list of 323 genes that have been found associated to smoking in lymphocytes [8]. The two studies used different microarray platforms and the 13,707 genes common to both studies were taken as the reference set. Of the 323 genes associated to smoking in lymphocytes, 268 were detected in monocytes of GHS of which 151 were associated to smoking (56.4%). This represented a 2.5-fold enrichment versus the reference ($p = 4.6 \times 10^{-34}$). Using a more stringent FDR threshold of 0.05 rather than 0.10 as in [8] did not significantly affect the results (2,477 unique genes associated to smoking in GHS; 2.6-fold enrichment, $p = 2.4 \times 10^{-34}$). Results from both studies did not only overlap in the list of genes associated to smoking but also in the magnitude and the direction of the effects (Pearson correlation coefficient of 0.72 between the r_D^2 estimated in GHS and the corresponding correlation values reported in [8]) (Figure S1).

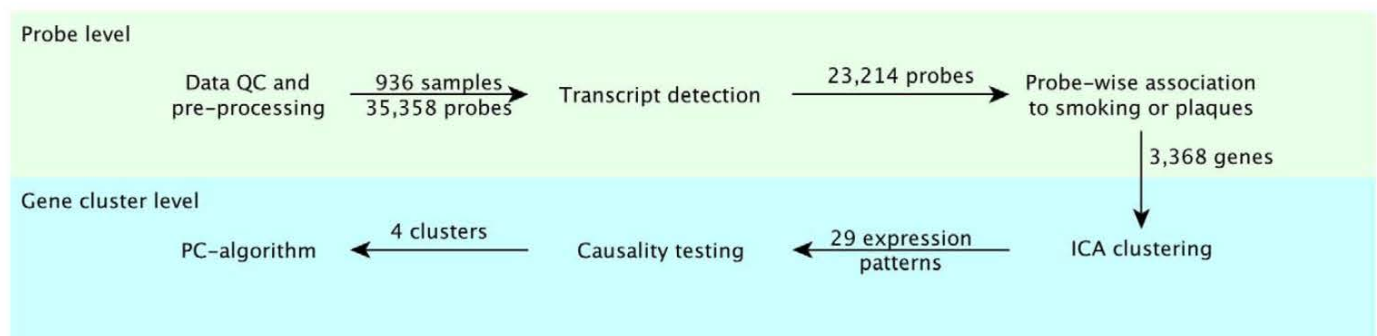


Figure 1. Analysis workflow. Microarray expression data were analyzed at two levels, a probe level (top) and a gene cluster level (bottom). Of the 35,358 probes with a “good” or “perfect” score according to ReMOAT, 23,214 were detected in monocytes of 936 subjects. Of these, 3,960 probes were associated to smoking or atherosclerotic plaques at an $FDR \leq 0.1$, corresponding to 3,368 unique genes that were further clustered in 29 expression patterns by ICA. Causality testing revealed 4 patterns that were compatible with expression mediating the relationship between smoking and plaques. The skeleton of the network connecting genes, smoking, risk factors and plaques was then inferred using the PC algorithm.
doi:10.1371/journal.pone.0050888.g001

Table 3. Gene Ontology categories enriched for smoking-associated gene expressions.

| Term | Genes in GO class | Smoking-associated genes | p-value | Bonferroni corrected* |
|--|-------------------|--------------------------|----------|-----------------------|
| GO: Biological Process (BP) | | | | |
| platelet activation | 224 | 74 | 5.50E-07 | 0.0055 |
| interferon- γ -mediated signaling pathway | 65 | 29 | 4.50E-06 | 0.0451 |
| Toll signaling pathway | 71 | 30 | 4.60E-06 | 0.0461 |
| GO: Cellular Component (CC) | | | | |
| cytosol | 1889 | 469 | 4.30E-15 | 6.02E-12 |
| soluble fraction | 315 | 91 | 2.50E-06 | 0.0035 |
| focal adhesion | 100 | 38 | 2.80E-06 | 0.0039 |
| melanosome | 87 | 33 | 3.40E-06 | 0.0048 |
| cytoplasm | 7509 | 1634 | 3.70E-06 | 0.0052 |
| MHC class II protein complex | 12 | 9 | 3.00E-05 | 0.0420 |
| GO: Molecular Function (MF) | | | | |
| protein binding | 6620 | 1433 | 1.70E-12 | 4.38E-09 |

*Bonferroni correction on the number of GO terms represented in the reference set: 5010 (BP), 700 (CC), 1289 (MF).

doi:10.1371/journal.pone.0050888.t003

Table 4. Top 10 genes associated to smoking or plaques.

| ProbeID | Symbol | beta | r ² _D | p-value |
|---|-----------|-------|-----------------------------|----------|
| Top 10 genes associated to smoking | | | | |
| ILMN_2185984 | SASH1 | 0.58 | 0.35 | 1.4E-88 |
| ILMN_1660031 | P2RY6 | 0.20 | 0.26 | 2.0E-61 |
| ILMN_1717207 | MMP25 | 0.59 | 0.22 | 7.5E-53 |
| ILMN_1752728 | FUCA1 | 0.33 | 0.19 | 7.2E-45 |
| ILMN_1671891 | PID1 | 0.37 | 0.19 | 4.3E-44 |
| ILMN_1664464 | PTGDS | -0.68 | 0.19 | 6.7E-44 |
| ILMN_1656300 | GFRA2 | 0.42 | 0.17 | 1.2E-38 |
| ILMN_1708303 | CYP4F22 | -0.42 | 0.16 | 1.3E-37 |
| ILMN_1655987 | STAB1 | 0.29 | 0.16 | 3.4E-37 |
| ILMN_1818677 | LOC157627 | 0.12 | 0.15 | 1.8E-34 |
| Top 10 genes associated to plaques | | | | |
| ILMN_1800225 | PPARG | 0.09 | 0.05 | 2.26E-11 |
| ILMN_2185984 | SASH1 | 0.16 | 0.04 | 1.77E-09 |
| ILMN_2103107 | ADAMDEC1 | 0.07 | 0.03 | 1.70E-08 |
| ILMN_1664464 | PTGDS | -0.24 | 0.03 | 4.61E-08 |
| ILMN_1706304 | EIF2C4 | 0.05 | 0.03 | 1.31E-07 |
| ILMN_2352633 | ARHGAP24 | 0.05 | 0.03 | 2.40E-07 |
| ILMN_1794914 | UBTD1 | 0.05 | 0.03 | 2.93E-07 |
| ILMN_1802808 | SLC25A37 | 0.12 | 0.03 | 3.17E-07 |
| ILMN_1752478 | DHRS3 | -0.12 | 0.03 | 5.35E-07 |
| ILMN_1687592 | WWC3 | 0.06 | 0.02 | 1.46E-06 |

Association of probe level was tested separately with smoking and ln(plaques+1) by linear model adjusted for age, sex and the 6 first SVD components. The beta regression coefficient is shown. Probes were ranked according to decreasing r²_D for smoking (plaques, respectively). In case of several probes by gene (e.g. *SASH1* and *PPARG*), the probe with the highest r²_D is shown.

doi:10.1371/journal.pone.0050888.t004

Association between plaques and smoking conditional on a single gene expression

In a second step, we investigated whether single gene expressions might mediate the effects of smoking on plaques. For this purpose, we modeled plaques as a function of smoking and each single gene expression. In this one-dimensional scan, no gene could entirely explain the association between plaques and smoking by conditioning on its expression. *SASH1* was the gene that, once accounted for, contributed to the largest reduction in the covariation between plaques and smoking. *PTGDS* and *PPARG* were the second and third genes contributing the most to this reduction (Table S3). These results suggested that, not unexpectedly, the underlying link between smoking and plaques might involve more complex networks, including several genes and/or hidden variables. Next, we devised an approach to explore a more comprehensive set of models on multiple genes and other variables at a time.

Clusters of genes associated to plaques or smoking

Unsupervised gene clustering was used to reduce the dimensionality of the data before testing causal models involving multiple genes. Prior to this, we reduced the set of gene expressions by considering only genes that were significantly associated to smoking or to plaques when tested either separately or jointly (see Methods). This constituted a set of 3,960 probes. The probe with the highest variance in intensity across samples was chosen for each gene, leaving a set of 3,368 distinct gene profiles associated to plaques or smoking.

Independent component analysis (ICA) was used to identify patterns of co-expression in this set of 3,368 genes. ICA is an efficient algorithm that factorizes a matrix of multivariate data into a mixing matrix **A** of *patterns* for independent "hidden" components causing correlation among variables and an **S** matrix of *signatures*, which are vectors of coefficients associating variables (genes) to components (see Methods and [20] for details and [21] for a recent application to gene expression data). A pattern is a linear combination of gene expressions whose level varies among individuals. A signature is a vector of the contributions of a component to each gene expression that can be characterized by the genes that are most affected by that component (see *module*

below). In the following, the terms *pattern*, *signature* or *module* are used for referring to a component according to the context where it applies (i.e. individuals or genes).

The number of components to extract was determined by a permutation test, which indicated that 59 components could be detected in this dataset (see Methods). Components that did not meet pre-specified quality control criteria were discarded, leaving 29 components for analysis (see Methods and Text S1 for details).

Each component was associated to a specific *module* of genes characterizing its *signature*. A module was defined as the subset of genes that were selected on either tail of the signature distribution by controlling local FDR ≤ 0.001 , as done previously [21]. This resulted in 29 modules of 9 to 125 genes (Table S4). Two modules were enriched in GO pathways: module 18 (interferon-gamma-mediated signaling pathway) and module 39 (antigen processing and presentation via MHC class II) (Table S5).

Association of ICA expression patterns with smoking and plaques

As mentioned above, the patterns obtained by ICA factorization are linear combinations of gene expressions whose level can be interpreted as reflecting the “degree of activation” of subsets of co-expressed genes among individuals. Association of expression patterns with smoking or plaques was investigated in a similar fashion to that employed for individual gene expressions, except that we used a Bonferroni-corrected significance threshold ($0.05/29 = 0.0017$) instead of a FDR. At this significance threshold, 14 of the 29 patterns were associated to smoking and 7 to plaques, 5 being common to both (Table 5).

Worthy of note, ICA was able to recover a strong signature of smoking effects on gene expression in monocytes (pattern 43), as reflected by the high proportion of variance of that pattern (56%) explained by smoking (Table 5). The module associated to pattern 43 comprised 34 genes listed in Table S4. At the core of module 43 was *SASH1*, which had the highest coefficient for signature 43 in the **S** matrix and therefore was the most correlated to the overall pattern of expression.

Significant overlap with genes associated to smoking in lymphocytes [8] was tested for each ICA module in the same manner as for the full set of genes associated to smoking (above). Six of the 29 modules were significantly enriched for genes associated to smoking in lymphocytes (Table S6). The most enriched module was module 43, for which 12 of the 34 genes were also observed in lymphocytes ($OR = 27.35$; $p = 9.1 \times 10^{-13}$).

As in the single-gene case, no single pattern was able to entirely account for the association between smoking and plaques. The top-ranking pattern by amount of covariation explained between *P* and *S* was pattern 43 (Table 5). When this pattern was introduced in the model relating plaques to smoking, the proportion of variability of plaque counts explained by smoking (r_D^2) decreased from 8.8% to 2.3%.

Selection of expression patterns with potential causal role in the relationship between smoking and plaques

Though no single pattern could entirely explain the relationship between smoking and plaques, we sought to determine whether some ICA patterns may show evidence for a causal effect partially mediating the link of smoking to plaques. For this analysis, we used graphical modeling (Methods). For each expression pattern, the best-supported causal model involving smoking (*S*), plaques (*P*) and gene expression pattern (*G*) was selected by a likelihood-based model selection approach. All equivalence classes of graphical models among these three variables were enumerated (Figure 2).

Two classes, *f* and *k*, were of primary interest because both comprise models with a path $S \rightarrow G \rightarrow P$, which is the causal relation of interest. Under model class *f*, all the association between smoking and plaques can be explained by the effect of one single pattern. In model *k*, the pattern does not explain all the covariation between smoking and plaques, but it is associated to both (Text S1). Maximum likelihood was used to identify the equivalence class that was best supported by the data for each pattern. The process was repeated for 1000 bootstraps of the data to account for uncertainty in model selection.

The results of causal model inferences are summarized in Figure 3 and Table S4 (spreadsheet “Causality”). The probabilities of the different models across the 1000 bootstraps for each pattern are shown in bottom half of plot in Figure 3. The probability of selecting a model from a *causal class* was defined as the sum of probabilities for the model classes *f* and *k* (top half of plot in Figure 3). According to this criterion, 4 patterns (21, 29, 31 and 51) had a relatively high support for causality (probability ≥ 0.6). Worthy of note, pattern 43, the one the most influenced by smoking, was associated with the model $G-S-P$ with a probability of 0.7, which is incompatible with $S \rightarrow G \rightarrow P$ causality.

Inference of the gene network underlying ICA expression patterns

To further characterize the ICA patterns showing some support for causality, we inferred the topology of the networks underlying patterns. For each pattern, the network was constructed from the subset of genes composing the module specific to that pattern. We applied the PC algorithm 1 to discover the skeleton of conditional independencies (algorithm 1 in [22]; see Methods). The network was represented as an undirected graph. To decrease the possibility of hidden variables for the network, we considered in these analyses all cardiovascular risk factors that were associated to each pattern by stepwise regression (Table S7). This is important to avoid spurious edges resulting from untested confounding variables. To assess uncertainty in the inference, the process was repeated 1000 times by bootstrapping individuals and the proportion of data samples that recovered an edge was recorded. We considered only edges with a recovery probability ≥ 0.6 (see Methods). Graphic representations for all networks are found in Text S2.

Networks selected by causality test

For the 4 modules with suggestive support for causal effects (21, 29, 31 and 51), we estimated the minimum path(s) between smoking and plaques, starting from each gene directly connected to smoking. Pattern 31 did not reveal any path because no gene was connected to plaques in more than 60% of bootstraps.

For pattern 21 (Figure 4), the different paths connecting smoking to plaques involved four genes directly connected to smoking (*MAP3K6*, *GAS6*, *HTRAI* and *DSC2*) and only one gene directly connected to plaques (*SLC39A8*) (Table 6). Therefore, the *SLC39A8* gene funneled all information paths between smoking and plaques in this network. *SLC39A8* expression level was positively correlated to plaques and smoking was positively correlated to all the genes in the cluster, suggesting that an up-regulation of genes of the cluster was associated to increased atherosclerosis extent. Additionally, *SLC39A8* was negatively correlated to HDL-cholesterol levels (~ 1 probability), which is consistent with the protective role of HDL-cholesterol in atherosclerosis.

In pattern 29, there were 7 genes directly connected to smoking (*CXCL16*, *DHRS9*, *FAM20C*, *FPR3*, *PDE4B*, *PTGFRN*, and *TBC1D8*) and only one gene connected to plaques (*RASDI*)

Table 5. Association between ICA gene expression patterns (G), smoking (S) and plaques (P).

| Pattern | n | b | | G~S | | G~P | | P~S G | |
|-----------|-----|---|---|---------------|-----------------------------|--------------|-----------------------------|--------------|-----------------------------|
| | | S | P | p-value | r ² _D | p-value | r ² _D | p-value | r ² _D |
| Pattern43 | 34 | + | + | 4E-166 | 0.56 | 4E-07 | 0.03 | 4E-06 | 0.02 |
| Pattern21 | 36 | - | - | 2E-23 | 0.10 | 7E-06 | 0.02 | 4E-14 | 0.06 |
| Pattern51 | 9 | - | - | 5E-11 | 0.05 | 4E-06 | 0.02 | 2E-16 | 0.07 |
| Pattern39 | 47 | + | + | 8E-10 | 0.04 | 1E-04 | 0.02 | 8E-17 | 0.07 |
| Pattern29 | 98 | + | + | 3E-15 | 0.07 | 0.000 | 0.02 | 5E-17 | 0.07 |
| Pattern18 | 98 | + | | 7E-20 | 0.09 | 3E-02 | 0.01 | 1E-17 | 0.08 |
| Pattern11 | 141 | - | | 3E-23 | 0.10 | 0.005 | 0.01 | 4E-18 | 0.08 |
| Pattern54 | 17 | | - | 2E-02 | 0.01 | 0.000 | 0.02 | 1E-18 | 0.08 |
| Pattern34 | 58 | + | | 0.000 | 0.03 | 1E-02 | 0.01 | 3E-19 | 0.08 |
| Pattern28 | 70 | | | 0.007 | 0.01 | 0.032 | 0.00 | 2E-19 | 0.08 |
| Pattern15 | 14 | + | | 6E-04 | 0.01 | 0.200 | 0.00 | 9E-20 | 0.09 |
| Pattern45 | 51 | | | 7E-02 | 0.00 | 0.413 | 0.00 | 8E-20 | 0.09 |
| Pattern36 | 64 | - | | 0.000 | 0.03 | 0.477 | 0.00 | 7E-20 | 0.09 |
| Pattern19 | 115 | | | 2E-02 | 0.01 | 0.425 | 0.00 | 5E-20 | 0.09 |
| Pattern52 | 49 | + | | 0.001 | 0.01 | 0.127 | 0.00 | 4E-20 | 0.09 |
| Pattern41 | 45 | | | 1E-01 | 0.00 | 0.849 | 0.00 | 4E-20 | 0.09 |
| Pattern14 | 67 | - | | 0.000 | 0.01 | 0.815 | 0.00 | 4E-20 | 0.09 |
| Pattern12 | 103 | | | 0.902 | 0.00 | 0.179 | 0.00 | 4E-20 | 0.09 |
| Pattern49 | 35 | | | 0.941 | 0.00 | 0.585 | 0.00 | 3E-20 | 0.09 |
| Pattern58 | 23 | | | 0.027 | 0.01 | 0.934 | 0.00 | 3E-20 | 0.09 |
| Pattern33 | 43 | | | 0.971 | 0.00 | 0.073 | 0.00 | 2E-20 | 0.09 |
| Pattern17 | 31 | | | 0.791 | 0.00 | 0.651 | 0.00 | 2E-20 | 0.09 |
| Pattern27 | 13 | | | 0.121 | 0.00 | 0.283 | 0.00 | 2E-20 | 0.09 |
| Pattern23 | 40 | | | 9E-01 | 0.00 | 0.360 | 0.00 | 2E-20 | 0.09 |
| Pattern30 | 97 | + | | 0.000 | 0.01 | 0.936 | 0.00 | 2E-20 | 0.09 |
| Pattern42 | 71 | | | 0.138 | 0.00 | 0.038 | 0.00 | 2E-20 | 0.09 |
| Pattern4 | 125 | | | 0.106 | 0.00 | 0.203 | 0.00 | 2E-20 | 0.09 |
| Pattern31 | 85 | | + | 0.072 | 0.00 | 0.006 | 0.01 | 3E-21 | 0.09 |
| Pattern48 | 22 | - | | 6E-06 | 0.02 | 0.071 | 0.00 | 1E-21 | 0.09 |

n: number of genes in the pattern-specific module, b: sign of the regression coefficient of G on S and G on P, respectively. All models included age, sex and the first 6 SVD components. Only the 29 patterns that passed quality control are shown. P-values $\leq 0.05/29 = 0.0017$ are in bold. Genes are ranked by increasing r²_D associated to S in the model P~S|G, which corresponds to decreasing reduction of the amount of covariation between smoking and plaques explained by pattern expression.

doi:10.1371/journal.pone.0050888.t005

(Table 6). *RASD1* was positively correlated to plaques, although this association was recovered with probability of only 0.64. This was a large network with 98 genes (Table S4). However, smoking and plaques were separated by a relatively low number of genes where the 7 paths had between 3 to 5 connecting genes (Table 6).

Pattern 51 was the smallest (15 nodes) and among the most interconnected networks, with an average number of connections per node of 4.3. Among the different paths connecting smoking to plaques, 2 genes were directly connected to smoking (*CYP1B1* and *HOXA10*) and only one gene was directly, negatively, connected to plaques (*TMEM136*) (Table 6).

PPARG network

PPARG was the gene showing the strongest association to plaques and the third in the reduction of the r²_D of smoking to plaques. In addition, this gene is known to be involved in atherosclerosis [23]. This prompted us to examine in more details the network(s) comprising *PPARG*. Actually, *PPARG* was only

present in module 42, which included 71 genes (Table S4). Pattern 42 was not significantly associated to smoking or plaques (Table 5). However, causality testing gave inconclusive results, with two model classes, *h* (*S*→*P*←*G*) and *k* (*S*—*G*—*P*—*S*), alternatively selected by bootstrapping with probabilities 0.5 and 0.31, respectively (Table S4). In model *h* which had the best support, smoking and pattern expression were associated only when conditioned on plaques, meaning that both have causal effects on plaques but independently from one another. Conversely, the topology of the network suggested that the shortest path from smoking to plaques was through *PPARG* only (Text S2). This path, which was recovered in 96% of bootstraps, supported a causal effect of *PPARG*. This discrepancy between *PPARG* and pattern 42 might be explained by the fact that the contribution of pattern 42 to *PPARG* was weak (ranking 67 out of 71 genes of the module) and therefore, the behavior of *PPARG* did not exactly match that of the entire pattern.

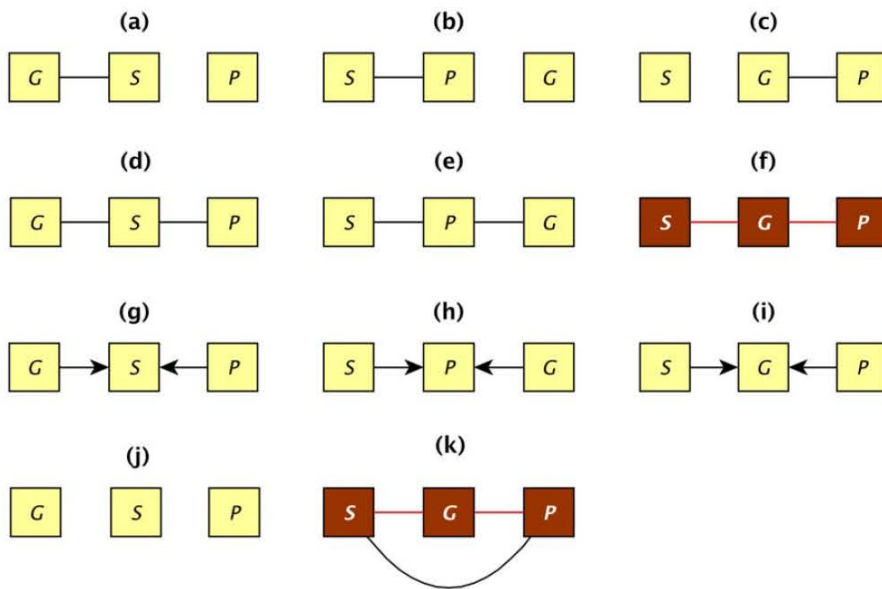


Figure 2. Graphical models for equivalence classes tested among smoking, gene expression and atherosclerotic plaques. Variables are represented by squares and causal associations are indicated by directed edges among nodes. Undirected edges indicate bidirected edges. The two classes colored in brown represent the causal models of interest where gene expression (*G*) mediates the association between smoking (*S*) and plaques (*P*). In class (f), the covariation between *S* and *P* is entirely explained by *G*, whereas in class (k), there is residual covariation between *S* and *P* after conditioning on *G*.

doi:10.1371/journal.pone.0050888.g002

Discussion

We present results from a genome-wide survey of gene expression in monocytes that revealed widespread effects from smoking, with >3000 genes either over- or under- expressed in smokers. Because no information was collected about number of cigarettes smoked per day, we could not test dose-dependence of the effects on gene expression. Due to the non-stringent FDR adopted, the list was rather large because our primary objective was not to miss any gene of potential interest. The list of smoking-associated genes showed significant overlap with those observed in

lymphocytes from a large cohort of Mexican Americans [8] indicating high robustness of smoking effects across different circulating cell types and genetic background. By contrast, the number of gene expressions associated to atherosclerosis was much lower (236 genes), which is probably explained by the fact that atherosclerosis is a complex and distal phenotype with multiple genetic and non genetic determinants.

As expected, a one-dimensional scan across the transcriptome revealed that no single gene could explain the association between smoking and plaques, leading us to search for networks of genes that might be more relevant. Using ICA, we identified 29 patterns

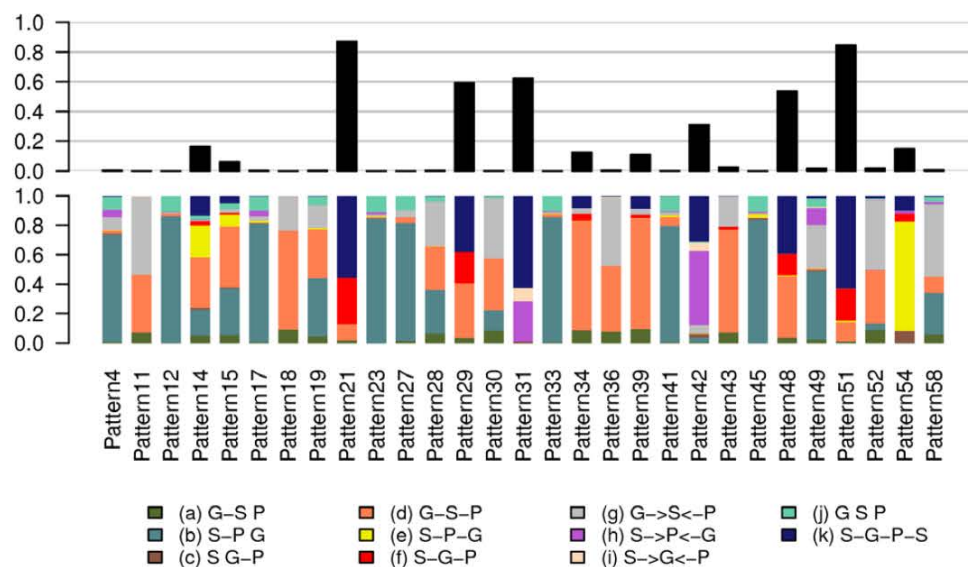


Figure 3. Probability of selection of the different causality models for each ICA expression pattern. Probability was estimated from 1000 bootstraps of the data. The bottom of the graph shows the probabilities of the 11 models described in Figure 2. The top of the graph shows the sum of the probabilities for models (f) and (k) representing the causal classes.

doi:10.1371/journal.pone.0050888.g003

Pattern21 Subnetwork Skeleton

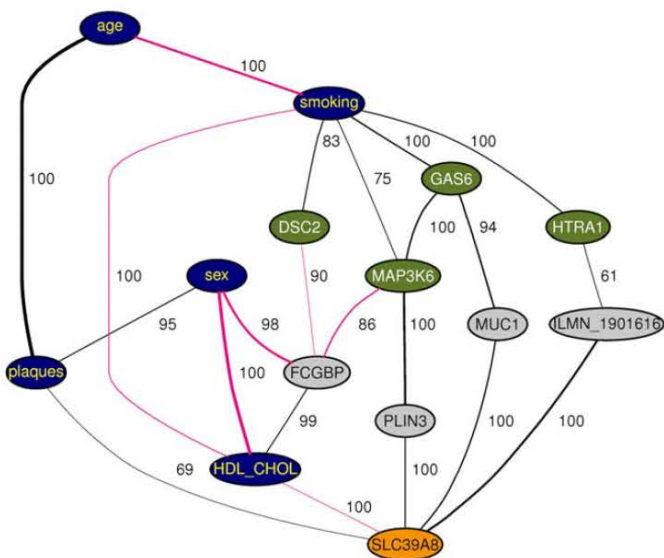


Figure 4. Subnetwork of PC skeleton for Module 21. This graph represents a consensus network from 1000 bootstraps. Edges among variables are drawn if detected in at least 60% of bootstrapped samples. The recovery percentages are indicated to the right of the medial section of each edge. Line thickness is proportional to the edge's partial correlation. Black edges denote positive and pink edges negative partial correlations. Plaques and risk factors are in blue. Genes directly connected to smoking are in green and those directly connected to plaques are in orange. Other genes are in gray. Only genes that are involved in the shortest paths connecting smoking to plaques are shown. The full network for this and other patterns are found in Text S2. doi:10.1371/journal.pone.0050888.g004

of co-expressed genes, 14 of which were strongly associated to smoking. Two patterns were enriched in functional GO categories (interferon-gamma-mediated signaling pathway and MHC class II antigen processing) but none of these two patterns was related to atherosclerosis after conditioning on smoking. Worthy of note, one of the patterns (pattern 43 driven by *SASH1*) could be interpreted as a robust signature of the impact of smoking on the transcriptome of circulating blood cells, as demonstrated by the substantial overlap of smoking-associated genes between monocytes and lymphocytes [8]. Actually, all the genes directly connected to smoking in network 43 (i.e. *SASH1*, *MMP25*, *P2RY6*, *FUCA1*, *PIDI*, *DTNA*, *GFRA2*, *CLEC10A* and *PTGDS*) had been previously identified as the strongest correlates of smoking in GHS [9]. Surprisingly, none of these genes was found differentially expressed in a recent experimental study performed in a human monocytic cell line (THP-1 cells) exposed to cigarette smoke extract [24]. This discrepancy suggests that *in vivo* chronic exposure to cigarette smoke may have a different impact from *in vitro* acute exposure, in particular because of the important role played *in vivo* by the lung, kidney and liver in metabolizing xenobiotics.

We then tested causality by graphical modeling. Actually, two model classes, $S-G-P$ and $S-G-P-S$ contained the causal path of interest $S \rightarrow G \rightarrow P$. Although both of these graphs describe multiple Bayesian networks, *a priori* information can be used to favor only a few of them. Indeed, genes expressed in monocytes are unlikely to affect smoking behavior, eliminating $S \leftarrow G$ edges. Supporting this assumption, only three genes across all networks have been

Table 6. Shortest gene paths connecting smoking to plaques.

| Pattern | Paths |
|-----------|--|
| Pattern21 | smoking - GAS6 - MUC1 - SLC39A8 - plaques |
| | smoking - HTRA1 - ILMN_1901616 - SLC39A8 - plaques |
| | smoking - MAP3K6 - PLIN3 - SLC39A8 - plaques |
| Pattern29 | smoking - DSC2 - GK - FCGR1A - MUC1 - SLC39A8 - plaques |
| | smoking - DHRS9 - ACOX2 - COQ2 - RASD1 - plaques |
| | smoking - FAM20C - NEK6 - ACOX2 - COQ2 - RASD1 - plaques |
| | smoking - FPR3 - AVPI1 - RASD1 - plaques |
| | smoking - CXCL16 - UTP6 - FAR2 - RASD1 - plaques |
| | smoking - TBC1D8 - CMTM4 - FFAR2 - RASD1 - plaques |
| | smoking - PDE4B - PFKFB3 - COQ2 - RASD1 - plaques |
| Pattern51 | smoking - PTGFRN - CDC42EP2 - COQ2 - RASD1 - plaques |
| | smoking - CYP1B1 - TMEM136 - plaques |
| | smoking - HOXA10 - MYB - CYP1B1 - TMEM136 - plaques |

Only networks passing the causality tests are shown.
doi:10.1371/journal.pone.0050888.t006

reported by GWAS to be associated to smoking behavior and none of them were in modules selected for causal effects (Table S8). On the other hand, by excluding ex-smokers, we reduced the possibility that smoking behavior might be modified by the presence of atherosclerosis, making $S \leftarrow P$ edges less likely (see Text S1 for details). Therefore, the networks most likely underlying the two graphical models of interest were $S \rightarrow G \rightarrow P$ for the two-edge case and $S \rightarrow G \rightarrow P \leftarrow S$ or $S \rightarrow G \leftarrow P \leftarrow S$ for the three-edge case.

We identified four patterns of expression (21, 29, 31, and 51) that were compatible with an effect of expression partially mediating the relation between smoking and atherosclerosis. Pattern 31 did not have a gene network topology consistent with any gene mediating smoking to plaques effects. Pattern 29 was associated to a large module (94 genes) with some support for causal effects (0.59 probability). All the paths in this network were connected to plaques through a single gene, *RASD1*. This gene, whose exact function is unknown, encodes a Ras-related protein stimulated by dexamethasone, a drug with anti-inflammatory and immunosuppressive actions. Because of the relatively low probability of causality of pattern 29 and the modest recovery of the *RASD1*-plaques edge (0.64 frequency), caution is needed in the interpretation of this result.

Pattern 51 had a relatively high support for causality (0.85 probability). In this network, smoking was directly connected to 2 genes, *CYP1B1* and *HOXA10*, which both have relevance to atherosclerosis and smoking. *CYP1B1* encodes a member of the cytochrome P450 protein superfamily that localize to the endoplasmic reticulum and metabolize procarcinogens including polycyclic aromatic hydrocarbons in tobacco-smoke [25]. *CYP1B1* expression showed the strongest difference in placenta from smoking and nonsmoking women [26] and was increased in THP-

l cells after *in vitro* exposure to cigarette smoke [24]. In endothelial cells, *CYP1B1* expression is regulated by shear stress and is associated to anti-atherogenic effects [27] and decreased oxidative stress [28]. *HOXA10* encodes a transcription factor associated to cell proliferation in the monocytes cell lineage [29] and repression of PHOX genes involved in oxidative stress [30]. It is expressed in the endothelium in a location-dependent manner, with lower expression in atheroprone than in atheroresistant arteries [31,32]. *HOXA10* expression in endometrium has been shown to be directly affected by cigarette-smoke extract both in humans and mice [33]. Therefore, there is support in the literature linking at least two genes of the network to smoking and plaques. However, *TMEM136*, which was the only gene of the network directly connected to plaques, was expressed at low levels in monocytes, only exceeding detection threshold in 5% of male non-smokers and a few individuals outside this group. *TMEM136* encodes a transmembrane protein of unknown function. Therefore, it is possible that association between plaques and this gene is instead reflecting association to a different unobserved network member.

Pattern 21 which had the highest support for causality (0.87 probability) also appeared to have the highest relevance in the context of smoking-induced atherosclerosis. In this network, connection to plaques was mediated by *SLC39A8* (aka ZIP8), a transmembrane zinc transporter. Genetic variants in *SLC39A8* have previously been associated to several cardiovascular risk factors such as HDL-cholesterol [34,35], blood pressure [36,37], obesity [38], and activation of plasminogen [39]. In addition to being connected to plaques, *SLC39A8* was directly connected to HDL-cholesterol in the network (Figure 4), a result consistent with the association found with *SLC39A8* genetic variants [34,35] and supporting a causal role of *SLC39A8* in atherosclerosis. However, here we provide evidence for an effect on plaques not completely mediated by HDL-cholesterol, since a direct plaques-*SLC39A8* edge was recovered with 0.69 probability. *SLC39A8* is known to have a cytotoxic role by intracellular transport of cadmium, a toxic heavy metal and carcinogen that is abundant in cigarette smoke [40,41]. In lung epithelia, *SLC39A8* expression is increased by TNF α , a pro-inflammatory cytokine that is abundant in smoker's lung [42]. In addition, the cadmium-mediated toxicity induced by cigarette smoke has been shown to be enhanced through NF- κ B-mediated activation of *SLC39A8* expression [42].

The shortest path from smoking to plaques in network 21 was smoking-*GAS6*-*MUC1*-*SLC39A8*-plaques. *MUC1* (mucin 1) encodes a membrane protein involved in cell adhesion and signal transduction, not previously associated to smoking effects or atherosclerosis. By contrast, *GAS6* (growth arrest-specific gene 6) has a strong relevance to atherosclerosis. It belongs to a family of vitamin K-dependent coagulation proteins and has a pleiotropic role in atherosclerosis, with pro- and anti-atherogenic effects [43]. In human atherosclerotic plaques, which are the focus of the present study, *GAS6* has been shown to be expressed mainly by vascular smooth muscle cells and to have an anti-inflammatory action by stimulating the anti-inflammatory cytokine TGF β and inhibiting expression of TNF α [44]. Therefore, TNF α may be a molecular signal underlying the correlation between *SLC39A8* and *GAS6*, but this hypothesis needs to be confirmed. *GAS6* was also among the genes reported to be up-regulated in the placenta of women smoking during pregnancy [26]. To the best of our knowledge, *GAS6* and *SLC39A8* have not been connected before in the context of atherosclerosis and their functional link needs to be experimentally confirmed.

Because ICA was performed on a subset of genes pre-selected by their association with smoking or atherosclerosis, we cannot exclude the possibility that we missed some genes that could be

important nodes in the networks subsequently identified. However, not doing this pre-selection would have led to a larger number of ICA components (91 components were actually detected in our previous ICA application in the complete GHS expression dataset [21]), most of them being irrelevant for the problem under study. Also, we cannot exclude the possibility of spurious edges in networks resulting from untested confounding variables.

Networks discovered in this study may represent diverse mechanisms of gene-by-gene co-expression. For instance, expression of genes may be co-regulated by a signaling pathway in a single cell type or they may represent coordinated variation in the proportion of cell population subclasses of monocytes. Both possibilities are equally interesting since causal effects may be mediated by either mechanism. For instance, HDL levels modulate monocytes proliferation and activation, changing the composition of the myeloid cell lineage, which is thought to explain in part its anti-inflammatory and athero-protective effects [45,46]. Further studies are needed to determine what specific molecular mechanisms underlie the correlation patterns reported here. Additionally, replication of candidate gene networks in independent datasets should be performed before laborious functional studies are undertaken.

A word of caution is in place about inferring causal associations from observational data. Direct associations in causal graphs are appropriate only in the context of the variables that are included in the system, i.e. the variables that were measured. Causal graphs are considered complete only in the sense of common causes between variables but they do not include all causes of variables [47]. Therefore, a direct causal association may represent the net effect of a large number of direct causal associations among variables that were not measured and that mediate the effects among the variables that were observed [48]. A causal graph may change if variables are removed or new ones are included. Furthermore, the real causal associations that can be inferred from graphical models are not in the edges but in the lack of edges between variables, that is, two variables are not causally associated given that we account for the other variables in the system. Instead the presence of an arrow only indicates the possibility of a causal association, which has to be determined from data [49]. The method used here can only identify networks, which in the context of the variables measured, present a correlation structure that is not incompatible with the causal effects of interest $S \rightarrow G \rightarrow P$. An effort was made to include all other variables that may be relevant in the system, but of course there is no guarantee that all relevant variables were included. In addition, although the method does not always allow identifying the best model, providing the best fitting class of models is an honest and useful summary of the information encoded in these data.

In conclusion, we have used a graphical modeling approach to investigate the potential role of gene expression in monocytes in mediating smoking effects on atherosclerosis. The analytic approach implemented here allowed discriminating among competing causal models affecting multiple genes and revealed gene networks that included multiple members with known causal roles in atherosclerosis or mediation of smoke-tobacco effects. To the best of our knowledge, this is the first application of causal inference on gene modules rather than individual genes. Our results put together previously unconnected genes that led to the formulation of new hypotheses about potential molecular mechanisms linking smoking effects to atherosclerosis phenotypes. Therefore, inspection of the correlation structure of risk factors, gene expression and atherosclerosis, revealed candidate genes that would have been missed by looking at strength of gene-phenotype associations alone.

Materials and Methods

More details are provided in Text S1.

Subjects

Study participants of both sexes aged 35–74 yr, were successively enrolled into the GHS, a community-based cohort study in the Rhein-Main region in western mid-Germany. Participants were of European origin. More details about the GHS study are provided in [9]. There were 1,536 individuals with microarray expression data that passed quality control tests. Nine individuals with missing number of plaques were removed. Smoking status was recorded by interview at recruitment. Current smokers (individuals smoking ≥ 1 cigarette per day since at least 6 months before recruitment, $n = 248$) and nonsmokers (individuals who had never been smoking over a period of at least 6 months, $n = 688$) were used for analyses. Occasional smokers ($n = 42$), ex-smokers ($n = 547$) and individuals with missing information on smoking status ($n = 2$) were excluded.

Ethics statement

All subjects gave written informed consent. Ethical approval was given by the local ethics committee and by the local and federal data safety commissioners (Ethik-Kommission der Landesärztekammer Rheinland-Pfalz 22/03/2007 Number 837.020.07 (5555)).

Evaluation of the number of atherosclerotic plaques

Intima-media thickness (IMT) in the common carotid arteries was assessed with an ie33 ultrasound system (Philips, NL) using an 11 to 3 MHz linear array transducer. Measurements were performed by an experienced technologist and evaluated in the QLab software (Philips, NL). The presence of an atherosclerotic plaque was determined by an increment of 1.5 mm or more in IMT when compared to a region without plaque 1 cm before the carotid bulb, averaging from left and right carotids. Plaques were counted in the common, external, and internal carotid arteries on both left and right sides [9]. The phenotype considered was the count of plaques summed over both carotid arteries.

RNA samples processing

Total RNA from circulating monocytes was extracted as previously described [9]. Briefly, monocytes were isolated from blood samples with the RosetteSep Monocyte Enrichment Cocktail (StemCell Technologies, Vancouver, Canada), cells were resuspended in Trizol (Invitrogen, Karlsruhe, Germany), RNA was extracted with the RNeasy Mini kit (Qiagen, Hilden, Germany) and controlled for quality in an Agilent Bioanalyzer 2100 (Agilent Technologies, Boeblingen, Germany).

Microarray data processing

RNA samples were hybridized to the *Illumina* HT-12 BeadChip v3 (*Illumina*, San Diego CA) containing 48,803 50-mer DNA probes. Probe-mapping to the genome was obtained from ReMOAT annotations [50] in the *illuminaHumanv3.db* package from Bioconductor v2.8 (<http://www.bioconductor.org>). There were 13,445 probes with “Bad” scores for genome alignment, which were discarded. Probes were annotated with RefSeq and EntrezGene ids using the *org.Hs.eg.db* package from Bioconductor. Of the remaining 35,358 probes, 28,515 were annotated to Ensembl transcripts and 28,137 to EntrezGene ids. For the purpose of counting number of genes throughout, unannotated (but of good quality) probes were considered as targeting distinct genes. Note that probe annotations from *Illumina* were not used to

discard probes, which is different from our previous analyses on this dataset [9,21]. However, 98% of probes that were removed in our previous analyses were also removed by at least one of the filtering steps applied here.

Bead-level data were processed with BeadStudio (*Illumina*, San Diego CA) to perform quality control and summarization of intensity values at probe level. Data were further processed by quantile normalization. An *arcsinh* transformation was applied to stabilize the variance [51]. A transcript was considered detected when the normalized intensity reported by its targeting probe was significantly above that of negative control probes on the same array (detection p -value < 0.04). Probes with “detected” calls in $\geq 5\%$ of samples within any smoking/sex group were considered for analysis, representing 23,214 probes. All analyses were performed with R v2.13 [52].

Correction for technical sources of variability in gene expression

Major components of variance in the gene expression dataset were identified by singular value decomposition (SVD) using the *La.svd* function in R. The six largest components were not associated to any individual characteristic, and therefore were thought to reflect systematic effects from sample-processing protocols. The first 6 SVD components were then used as surrogate variables of technical effects and were adjusted for in expression analyses [53]. In addition, potential sample contamination with B and T cells, and megakaryocytes was corrected as previously described [21] (Text S1).

Modeling the number of plaques

In analyses modeling the number of plaques as dependent variable, we used a negative binomial (NB-2) distribution [54]. Other distribution functions were considered but not used (see Text S1 for details). The negative binomial model was fitted by the *glm.nb* function with a log link using the R MASS package. A pseudo- R^2 coefficient was defined as $R_D^2 = 1 - \frac{D(\mathbf{y}, \hat{\boldsymbol{\mu}})}{D(\mathbf{y}, \hat{\mathbf{y}})}$, where $D(\mathbf{y}, \hat{\boldsymbol{\mu}}) = 2[\log L(\mathbf{y}) - \log L(\hat{\boldsymbol{\mu}})]$ is the deviance around means predicted by model $\hat{\boldsymbol{\mu}}$, $\hat{\mathbf{y}}$ is the saturated model (one coefficient per observation), $D(\mathbf{y}, \hat{\boldsymbol{\mu}})$ is the deviance in the full model and $D(\mathbf{y}, \hat{\mathbf{y}})$ is the deviance in the model with an intercept only [55]. In GLM models of Gaussian response variables, this reduces to the R^2 coefficient of multiple determination. Similarly, a pseudo-coefficient of partial determination was computed to measure the marginal contribution of one independent variable in predicting a response when contribution from all other variables in the model is accounted for [56]. This coefficient for an independent variable x was defined as $r_D^2 = \frac{D(\mathbf{y}, \hat{\boldsymbol{\mu}}_{-x}) - D(\mathbf{y}, \hat{\boldsymbol{\mu}})}{D(\mathbf{y}, \hat{\boldsymbol{\mu}}_{-x})}$, where $D(\mathbf{y}, \hat{\boldsymbol{\mu}}_{-x})$ is the deviance of a reduced model with x dropped. Model selection was performed by the Bayesian Information Criterion $BIC = -2(\log L(\mathbf{y}|\hat{\boldsymbol{\theta}}) + k \log(n))$, where $L(\mathbf{y}|\hat{\boldsymbol{\theta}})$ is the maximum likelihood of the data given parameters $\hat{\boldsymbol{\theta}}$, k is the number of parameters in $\hat{\boldsymbol{\theta}}$ and n is the number of observations. When comparing different negative binomial models, the dispersion parameter α was held constant and fixed at the estimated value from the model with age, sex and smoking effects ($\alpha = 1.34$). Individuals with missing values for any variables being considered were removed before model fitting.

Single gene expression association analysis

Analysis of expression data was performed at the probe level. Association of a single probe expression with plaques counts and

smoking was performed by linear regression analysis where the dependent variable was the probe expression level and the independent variables were plaques and smoking, tested either individually or simultaneously. In these models, the variable considered for plaques was $\ln(\text{plaques}+1)$. All models were systematically adjusted for age, sex and the 6 first SVD components. Linear models were tested by the *glm* function in R. An FDR ≤ 0.1 was used for selecting the probes associated to plaques or to smoking [57]. See Text S1 for details.

The list of 3,368 distinct gene expressions associated to smoking was compared to the list of 323 gene expressions previously reported to be associated to smoking in lymphocytes [8]. This study used the *Illumina* Human WG-6 v1 microarray platform, which is an earlier version of the HT-12 used in GHS. Updated gene annotations were obtained from Bioconductor. There were 19,614 unique genes in WG-6 of which 17,243 were also queried in HT-12. Of these, 13,707 were detected in monocytes of GHS, which were used as the reference set. Gene set enrichment analysis was performed by the Fisher's exact test implemented in the *fisher.test* function in R.

GO enrichment analysis

Gene set enrichment analysis of gene ontology (GO) terms was performed with the *topGO* (v 2.4.0) package in R. GO data were obtained from the *GO.db* Bioconductor metadata package v 2.5.0 (March 2011). Enrichment was tested by a Fisher's exact test using the set of 18,364 unique known detected genes as a reference. In order to account for the redundancy and hierarchy of GO terms, the *weight01* algorithm was used, which is a hybrid between the *elim* and *weight* algorithms described in [58]. Control for multiple testing was done by a Bonferroni correction on the number of GO terms represented in the reference set.

Covariation between plaques and smoking explained by single genes

To test whether a single gene expression could explain part of the covariation between plaques and smoking, we modeled plaques as a function of smoking and gene expression using the *glm.nb* function. The dispersion parameter θ was held constant across genes to a value of 0.75, which was estimated with the *glm.nb* function on the reduced model without gene effect. The strength of the association of smoking before and after including gene expression was tested by the r_D^2 coefficient as described above. The difference between these values was regarded as covariation explained by a single gene.

Clustering gene expressions by Independent Component Analysis (ICA)

The fastICA algorithm was used to factorize the matrix of 3,368 gene expressions associated to smoking or plaques [20]. A single probe per gene was selected, that was the probe showing the largest variance across samples. This approach avoids the bias that would be introduced by, for instance using the probe with the strongest association to smoking or plaques and favors probes with more information content. Using the mean expression across probes for a gene was not considered because of the added noise that that may result from errors in probe annotation and because the largely uneven number of probes per gene would dramatically affect the distribution of technical error across genes. All unannotated probes were kept.

Normalized data from each probe were centered and standardized. SVD was initially performed to determine the *a*

priori number of components for the ICA algorithm as explained in [21]. This number was found to be 59.

The fastICA algorithm identifies major variance components by iteratively estimating the "mixing" matrix \mathbf{A} that satisfies the equation $\mathbf{X} = \mathbf{S}\mathbf{A}$, where \mathbf{X} is an $m \times n$ data matrix, \mathbf{S} is a $m \times p$ matrix of signatures across genes, \mathbf{A} is $p \times n$ matrix of patterns across samples, n is the number of samples ($n=936$), m is the number of genes ($m=3,368$), and p is the number of ICA components set *a priori* ($p=59$). The iterative algorithm minimizes dependency among signatures (columns of \mathbf{S}), while maximizing non-Gaussianity, i.e. negentropy, of signature distributions [20]. The fastICA function in the R-package of same name was used [59]. The algorithm was run multiple times to avoid trapping in a local maximum. The results were processed as previously described in [21] to remove components that were not consistently detected across random start points. Briefly, fastICA was repeated 500 times and the best run (with the maximal negentropy) was selected. The stability of components over the 500 runs was calculated. Components that did not meet quality control criteria were discarded (see Text S1 and Tables S9, S10, S11).

For each component, a signature-specific module was defined as the subset of genes on either tail of the signature distribution selected by controlling local FDR ≤ 0.001 [21]. The association between ICA expression patterns and smoking or plaques was tested in the same way as single gene expression, except that adjustment was made on all risk factors associated to pattern by stepwise regression analysis and not only on age and sex.

Investigation of causality models

All triplets including smoking (S), plaques (P), and a gene expression pattern (or a single gene expression) (G) were constructed. Potential causal relationships among triplet members were represented as Bayesian Networks, i.e. graphical models, where variables are vertices and causal associations are indicated by directed edges, i.e. arrows, between nodes. The probability function used depended on the distribution of each variable. Smoking was modeled by a binomial, plaques by a negative binomial, and gene expression by a Gaussian distribution function. The inference of network structure and parameters was performed as follows. Each node in the graph was fitted a linear model with its parents as independent variables. Given a graph M with variables $X = \{X_1, \dots, X_N\}$ represented as nodes, a BIC score [60] for M was computed as $BIC(M) = -2 \log(L(X|M)) + k \log(n)$, where n is the sample size, N is the number of nodes, k is the number of parameters of the model, and $L(X|M)$ is the maximum likelihood of the data given graph M , which by the *directed Markov property* of the graph, is given by $L(X|M) = \prod_{i=1}^N \Pr(X_i | Pa^M(X_i))$. The term $\Pr(X_i | Pa^M(X_i))$ denotes the conditional probability of observation X_i given its set of parents $Pa^M(X_i)$ in graph M . The graph with the lowest BIC is selected as the most likely model, which produces identical values for equivalent networks [61]. Here, however, because different density functions were used for different variables, equivalent networks may result in slightly different scores. Therefore, we computed BIC scores for all models in a class, which was selected if any network in the class had the minimum score. The probability of selecting a model class, estimated by 1000 bootstraps of the data was used as a measure of confidence in the Bayesian network inference [62].

Inference of network skeleton

The skeleton of the network connecting genes, smoking, risk factors and plaques was learned by conditional associations using the PC algorithm 1 implemented in the *pcalg* R package v 1.1-4 [22]. Briefly, the algorithm starts with a matrix of marginally

associated variables (nodes connected by edges) and for each pair of connected, i.e. adjacent, nodes it successively tests whether the pair becomes independent after conditioning on any group of adjacent nodes. If for any set of conditioning nodes the pair is independent, the edge is removed. The Pearson correlation with significance level of 0.05 was used to test independence. The process was repeated for 1000 data bootstraps and the proportion of samples where the edge is recovered was recorded. Shortest paths between smoking and plaques were inferred with the Dijkstra's algorithm on a reduced graph after removing nodes corresponding to covariates. Whenever more than one shortest path existed, the algorithm chooses one according to a greedy search. The Dijkstra's algorithm implemented in the *sp.between* function of the RBGL package for R was used [63].

Note that although in its full version the PC algorithm (Algorithm 2 in [22]) can asymptotically infer causal associations in a network by directing some of the edges of the skeleton, when applied to samples of finite size the algorithm may lead to inconsistencies that make such inference impossible [64]. This was the case with the present data. Therefore, we decided to limit to inferring the skeleton of significant partial correlations among variables using a first part of the PC algorithm.

Supporting Information

Figure S1 Comparison of the magnitude of smoking effects on gene expression in two independent cohorts. SAFHS: San Antonio Family Heart Study, (Charlesworth *et al.*, 2010); GHS: Gutenberg Health Study (this manuscript). Dots correspond to the 268 smoking-associated genes in SAFHS that were detected in GHS. The x-axis shows the Pearson correlation between gene expression and smoking in SAFHS and the y-axis the signed-square root of the r_D^2 coefficient for smoking in GHS. Genes associated to smoking in GHS are represented by black dots ($n = 151$), others are in gray.

(TIFF)

Figure S2 Density of the distribution of the edge recovery proportion from bootstraps for all possible node pairs. Patterns 4 to 28.

(TIFF)

Figure S3 Density of the distribution of the edge recovery proportion from bootstraps for all possible node pairs. Patterns 29 to 48.

(TIFF)

Figure S4 Density of the distribution of the edge recovery proportion from bootstraps for all possible node pairs. Patterns 49 to 58.

(TIFF)

Table S1 Risk factors associated to plaques count, selected by stepwise negative binomial regression.

(DOC)

Table S2 Results of association analysis of expression levels with smoking and plaques for all detected probes ($n = 23,214$).

(XLSX)

References

1. McGill HC Jr (1990) Smoking and the pathogenesis of atherosclerosis. *Adv Exp Med Biol* 273: 9–16.
2. Teo KK, Ounpuu S, Hawken S, Pandey MR, Valentin V, et al. (2006) Tobacco use and risk of myocardial infarction in 52 countries in the INTERHEART study: a case-control study. *Lancet* 368: 647–658. doi:10.1016/S0140-6736(06)69249-0
3. Yusuf S, Hawken S, Ounpuu S, Dans T, Avezum A, et al. (2004) Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet* 364: 937–952. doi:10.1016/S0140-6736(04)17018-9

Table S3 Top 20 genes explaining covariation between smoking and plaques.

(DOC)

Table S4 Summary tables and figures for 29 ICA expression patterns that passed quality control.

(XLSX)

Table S5 Gene Ontology categories enriched for gene expressions in specific ICA modules.

(DOC)

Table S6 Enrichment of ICA module gene sets in monocytes for genes associated to smoking in lymphocytes [Charlesworth *et al.* 2010].

(DOC)

Table S7 Cardiovascular risk factors associated to expression patterns by stepwise regression.

(DOC)

Table S8 Genes with reported association to smoking in GWAS.

(XLSX)

Table S9 Quality control statistics for ICA components of variance.

(DOC)

Table S10 Pearson correlation coefficients between ICA patterns and 6 first SVD components of expression matrix.

(DOC)

Table S11 Pearson correlation coefficients between ICA patterns and surrogate variables for cell contamination by non-monocytic cells.

(DOC)

Text S1 Supporting information: further details of Methods.

(DOC)

Text S2 Network plots for all ICA modules.

(PDF)

Acknowledgments

We appreciate the contribution of participants of the Gutenberg Health Study. We gratefully acknowledge the excellent medical and technical assistance of all technicians, study nurses and coworkers involved in the Gutenberg Health Study. We thank Raphaële Castagné for her invaluable assistance in the organization of the many files associated with this project. EN is a Director of Research in Centre National de la Recherche Scientifique.

Author Contributions

Conceived and designed the experiments: RAV FC SB LT. Performed the experiments: TZ PSW HW EN. Analyzed the data: RAV MR DAT FC LT. Contributed reagents/materials/analysis tools: TM KJL. Wrote the paper: RAV LT. Critical review of the manuscript: TZ MR PSW TM KJL HW EN DAT FC SB.

7. Dumeaux V, Olsen KS, Nuel G, Paulssen RH, Borresen-Dale A-L, et al. (2010) Deciphering normal blood gene expression variation—The NOWAC post-genome study. *PLoS Genet* 6: e1000873. doi:10.1371/journal.pgen.1000873
8. Charlesworth JC, Curran JE, Johnson MP, Göring HH, Dyer TD, et al. (2010) Transcriptomic epidemiology of smoking: the effect of smoking on gene expression in lymphocytes. *BMC Med Genomics* 3: 29. doi:10.1186/1755-8794-3-29
9. Zeller T, Wild P, Szymczak S, Rotival M, Schillert A, et al. (2010) Genetics and beyond—the transcriptome of human monocytes and disease susceptibility. *PLoS ONE* 5: e10693. doi:10.1371/journal.pone.0010693
10. Csiszar A, Podlutzky A, Wolin MS, Losonczy G, Pacher P, et al. (2009) Oxidative stress and accelerated vascular aging: implications for cigarette smoking. *Front Biosci* 14: 3128–3144.
11. Sopori M (2002) Effects of cigarette smoke on the immune system. *Nat Rev Immunol* 2: 372–377. doi:10.1038/nri803
12. Shipley B (2002) Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference. 1st ed. Cambridge University Press. 330 p.
13. Spirtes P (2010) Introduction to Causal Inference. *J Mach Learn Res* 99: 1643–1662.
14. Murphy KP (1998) Inference and Learning in Hybrid Bayesian Networks. University of California Berkeley. p. Available: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.32.6529>.
15. Margolin A, Nemenman I, Basso K, Wiggins C, Stolovitzky G, et al. (2006) ARACNE: An Algorithm for the Reconstruction of Gene Regulatory Networks in a Mammalian Cellular Context. *BMC Bioinformatics* 7: S7.
16. Watkinson J, Liang K-C, Wang X, Zheng T, Anastassiou D (2009) Inference of regulatory gene interactions from expression data using three-way mutual information. *Ann N Y Acad Sci* 1158: 302–313. doi:10.1111/j.1749-6632.2008.03757.x
17. Sales G, Romualdi C (2011) parmigene - a parallel R package for mutual information estimation and gene network reconstruction. *Bioinformatics*. Available: <http://www.ncbi.nlm.nih.gov/pubmed/21531770>. Accessed 2011 May 12.
18. Peters SA, den Ruijter HM, Bots ML, Moons KG (2012) Improvement in risk stratification for the occurrence of cardiovascular disease by imaging subclinical atherosclerosis: a systematic review. *Heart* 98: 177–184.
19. Johnson HM, Douglas PS, Srinivasan SR, Bond MG, Tang R, et al. (2007) Predictors of carotid intima-media thickness progression in young adults: the Bogalusa Heart Study. *Stroke* 38:900–5.
20. Hyvärinen A, Oja E (2000) Independent component analysis: algorithms and applications. *Neural Networks* 13: 411–430. doi:10.1016/S0893-6080(00)00026-5
21. Rotival M, Zeller T, Wild PS, Maouche S, Szymczak S, et al. (2011) Integrating genome-wide genetic variations and monocyte expression data reveals trans-regulated gene modules in humans. *PLoS Genet* 7: e1002367. doi:10.1371/journal.pgen.1002367
22. Kalisch M, Bühlmann P (2007) Estimating High-Dimensional Directed Acyclic Graphs with the PC-Algorithm. *J Mach Learn Res* 8: 613–636.
23. Staels B (2005) PPARgamma and atherosclerosis. *Curr Med Res Opin* 21 Suppl 1: S13–20. doi:10.1185/030079905X36440
24. Wright WR, Parzych K, Crawford D, Mein C, Mitchell JA, et al. (2012) Inflammatory Transcriptome Profiling of Human Monocytes Exposed Acutely to Cigarette Smoke. *PLoS One* 7. doi:10.1371/journal.pone.0030120
25. Nebert DW, Dalton TP (2006) The role of cytochrome P450 enzymes in endogenous signalling pathways and environmental carcinogenesis. *Nat Rev Cancer* 6: 947–960. doi:10.1038/nrc2015
26. Bruchova H, Vasikova A, Merkerova M, Milcova A, Topinka J, et al. (2010) Effect of maternal tobacco smoke exposure on the placental transcriptome. *Placenta* 31: 186–191. doi:10.1016/j.placenta.2009.12.016
27. Conway DE, Sakurai Y, Weiss D, Vega JD, Taylor WR, et al. (2009) Expression of CYP1A1 and CYP1B1 in human endothelial cells: regulation by fluid shear stress. *Cardiovascular research* 81: 669.
28. Tang Y, Scheef EA, Gurel Z, Sorenson CM, Jefcoate CR, et al. (2010) CYP1B1 and endothelial nitric oxide synthase combine to sustain proangiogenic functions of endothelial cells under hyperoxic stress. *Am J Physiol, Cell Physiol* 298: C665–678. doi:10.1152/ajpcell.00153.2009
29. Shah CA, Wang H, Bei L, Platanius LC, Eklund EA (2011) HoxA10 regulates transcription of the gene encoding transforming growth factor beta2 (TGFbeta2) in myeloid cells. *J Biol Chem* 286: 3161–3176. doi:10.1074/jbc.M110.183251
30. Lindsey S, Huang W, Wang H, Horvath E, Zhu C, et al. (2007) Activation of SHP2 protein-tyrosine phosphatase increases HoxA10-induced repression of the genes encoding gp91(PHOX) and p67(PHOX). *J Biol Chem* 282:2237–49.
31. Zhang J, Burrige KA, Friedman MH (2008) In vivo differences between endothelial transcriptional profiles of coronary and iliac arteries revealed by microarray analysis. *Am J Physiol Heart Circ Physiol* 295: H1556–1561. doi:10.1152/ajpheart.00540.2008
32. Burrige KA, Friedman MH (2010) Environment and vascular bed origin influence differences in endothelial transcriptional profiles of coronary and iliac arteries. *Am J Physiol Heart Circ Physiol* 299: H837–846. doi:10.1152/ajpheart.00002.2010
33. Zhou Y, Jorgensen EM, Gan Y, Taylor HS (2011) Cigarette smoke increases progesterone receptor and homeobox A10 expression in human endometrium and endometrial cells: a potential role in the decreased prevalence of endometrial pathology in smokers. *Biol Reprod* 84: 1242–1247. doi:10.1095/biolreprod.110.087494
34. Waterworth DM, Ricketts SL, Song K, Chen L, Zhao JH, et al. (2010) Genetic variants influencing circulating lipid levels and risk of coronary artery disease. *Arterioscler Thromb Vasc Biol* 30: 2264–2276. doi:10.1161/ATVBAHA.109.201020
35. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, et al. (2010) Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 466: 707–713. doi:10.1038/nature09270
36. Wain LV, Verwoert GC, O'Reilly PF, Shi G, Johnson T, et al. (2011) Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. *Nat Genet* 43: 1005–1011. doi:10.1038/ng.922
37. Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, et al. (2011) Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* 478: 103–109. doi:10.1038/nature10405
38. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, et al. (2010) Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* 42: 937–948. doi:10.1038/ng.686
39. Yang Q, Kathiresan S, Lin J-P, Tofler GH, O'Donnell CJ (2007) Genome-wide association and linkage analyses of hemostatic factors and hematological phenotypes in the Framingham Heart Study. *BMC Med Genet* 8 Suppl 1: S12. doi:10.1186/1471-2350-8-S1-S12
40. Dalton TP, He L, Wang B, Miller ML, Jin L, et al. (2005) Identification of mouse SLC39A8 as the transporter responsible for cadmium-induced toxicity in the testis. *Proc Natl Acad Sci USA* 102: 3401–3406. doi:10.1073/pnas.0406085102
41. He L, Girijashanker K, Dalton TP, Reed J, Li H, et al. (2006) ZIP8, member of the solute-carrier-39 (SLC39) metal-transporter family: characterization of transporter properties. *Mol Pharmacol* 70: 171–180. doi:10.1124/mol.106.024521
42. Napolitano JR, Liu M-J, Bao S, Crawford M, Nana-Sinkam P, et al. (2012) Cadmium-mediated toxicity of lung epithelia is enhanced through NF-κB-mediated transcriptional activation of the human zinc transporter ZIP8. *American Journal of Physiology Lung Cellular and Molecular Physiology*. Available: <http://www.ncbi.nlm.nih.gov/pubmed/22345571>. Accessed 2012 Feb 24.
43. Tjwa M, Moons L, Lutgens E (2009) Pleiotropic role of growth arrest-specific gene 6 in atherosclerosis. *Curr Opin Lipidol* 20: 386–392. doi:10.1097/MOL.0b013e328330982e
44. Clauser S, Meilhac O, Bièche I, Raynal P, Bruneval P, et al. (2012) Increased secretion of Gas6 by smooth muscle cells in human atherosclerotic carotid plaques. *Thromb Haemost* 107: 140–149. doi:10.1160/TH11-05-0368
45. Murphy AJ, Westerterp M, Yvan-Charvet L, Tall AR (2012) Anti-atherogenic mechanisms of high density lipoprotein: Effects on myeloid cells. *Biochim Biophys Acta* 1821: 513–521. doi:10.1016/j.bbali.2011.08.003
46. Norata GD, Pirillo A, Catapano AL (2011) HDLs, immunity, and atherosclerosis. *Curr Opin Lipidol* 22: 410–416. doi:10.1097/MOL.0b013e328330982e
47. Pearl J (2009) Causality: Models, Reasoning and Inference. 2nd ed. Cambridge University Press.
48. Spirtes P, Glymour C, Scheines R (2001) Causation, Prediction, and Search. Second ed. The MIT Press. 565 p.
49. Pearl J (2010) The foundations of causal inference. *Sociological Methodology* 40: 75–149. doi:10.1111/j.1467-9531.2010.01228.x
50. Barbosa-Morais NL, Dunning MJ, Samarajiva SA, Darot JFJ, Ritchie ME, et al. (2010) A re-annotation pipeline for Illumina BeadArrays: improving the interpretation of gene expression data. *Nucleic Acids Res* 38: e17. doi:10.1093/nar/gkp942
51. Lin SM, Du P, Huber W, Kibbe WA (2008) Model-based variance-stabilizing transformation for Illumina microarray data. *Nucleic Acids Res* 36: e11. doi:10.1093/nar/gkm1075
52. R Development Core Team (2011) R: A Language and Environment for Statistical Computing Vienna, Austria: R Foundation for Statistical Computing. Available: <http://www.R-project.org>.
53. Lunceford JK, Chen G, Hu PH, Mehrotra DV (2010) Evaluating surrogate variables for improving microarray multiple testing inference. *Pharmaceut Statist: n/a–n/a*. doi:10.1002/pst.466
54. Hardin JW, Hilbe JM (2007) Generalized Linear Models and Extensions, Second Edition. 2nd ed. Stata Press. 387 p.
55. Cameron AC, Windmeijer FAG (1996) R-squared measures for count data regression models with applications to health-care utilization. *Journal of Business & Economic Statistics*: 209–220.
56. Neter J, Wasserman W, Kutner MH (1996) Applied linear statistical models. Fourth ed. Irwin Homewood, IL. 1408 p. Available: http://books.google.fr/books/about/Applied_linear_statistical_models.html?id=q2sPAQAAAJ&redir_esc=y.
57. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological)* 57: 289–300.
58. Alexa A, Rahnenführer J, Lengauer T (2006) Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* 22: 1600–1607. doi:10.1093/bioinformatics/btl140
59. Marchini JL, Heaton C, Ripley BD (2010) fastICA: FastICA Algorithms to perform ICA and Projection Pursuit. Available: <http://CRAN.R-project.org/package=fastICA>.

60. Schwarz G (1978) Estimating the Dimension of a Model. *The Annals of Statistics* 6: 461–464. doi:10.2307/2958889
61. Chickering DM (1995) A transformational characterization of equivalent Bayesian network structures. *UAI'95*. pp. 87–98.
62. Friedman N, Linial M, Nachman I, Pe'er D (2000) Using Bayesian networks to analyze expression data. *J Comput Biol* 7: 601–620. doi:10.1089/106652700750050961
63. Carey V, Long L, Gentleman R (2010) RBGL: An interface to the BOOST graph library. R package version 1.
64. Maathuis MH, Kalisch M, Bühlmann P (2009) Estimating high-dimensional intervention effects from observational data. *The Annals of Statistics* 37: 3133–3164.

Annex 5. Touat et al (publication in preparation)

Role of lipid phosphate phosphatase (LPP3) in human aortic endothelial cell function and atherosclerosis.

¹Z. Touat-Hamici, ¹H. Weidmann, ²Y. Blum, ¹C. Proust, ¹H. Durand, ¹F. Iannacci, ^{4,5}P. Gaignard, ^{4,5}P Thérond, ²M. Civelek, ³A.S. Karabina, ²A.J. Lusis, ¹F. Cambien, ¹E. Ninio

¹Sorbonne Universités, UPMC, INSERM UMR_S 1166, ICAN, Genomics and Pathophysiology of Cardiovascular Diseases Team, F-75013, Paris, France

²University of California, Los Angeles, David Geffen School of Medicine, Los Angeles, CA 90095-1679, USA

³Sorbonne Universités, UPMC, INSERM UMR_S 933, Hôpital Armand-Trousseau, Paris, France

⁴APHP, Hôpital de Bicêtre, Service de Biochimie, 94275 Le Kremlin Bicêtre, France

⁵Université Paris Sud, UR Lip(Sys)², UFR de Pharmacie, Châtenay-Malabry, 92296, France

Correspondance:

Z. Touat-Hamici (zahia.touat@upmc.fr) and E. Ninio (ewa.ninio@upmc.fr)

Sorbonne Universités, UPMC, INSERM UMR_S 1166, ICAN, Genomics and Pathophysiology of Cardiovascular Diseases Team, 91 BD de l'Hôpital, F-75013, Paris, France

Abstract

Objective:

Lipid phosphate phosphatase 3 (LPP3), encoded by *PPAP2B* gene, is an integral transmembrane protein which dephosphorylates and thereby terminates signalling of a broad range of lipid substrates including lysophosphatidic acid (LPA) and sphingosine-1-phosphate (S1P). Human LPP3 possesses a cell adhesion motif that allows interaction with integrins. A polymorphism (rs17114036) in *PPAP2B* is associated with coronary artery disease which prompted us to investigate the possible role of LPP3 in human endothelial dysfunction, a condition promoting atherosclerosis.

Approach and results:

To study the role of LPP3 in endothelial cells we used human primary aortic endothelial cells (HAECs) in which LPP3 was silenced or overexpressed using either wild type or mutated cDNA constructs. LPP3 silencing in HAECs enhanced inflammatory cytokines, leukocyte adhesion, cell survival and migration and impaired angiogenesis, whereas wild-type LPP3 overexpression reversed these effects. We also demonstrated that LPP3 expression was negatively correlated with VEGF expression. Mutations of either the catalytic or the RGD domains impaired endothelial cell function and pharmacological inhibition of S1P or LPA restored it. The intra- and extracellular levels of S1P in HAEC under- or overexpressing LPP3 were correlated with LPP3 expression, indicating that S1P is probably degraded by LPP3 as well.

Conclusions:

We demonstrated that LPP3 is a negative regulator of inflammatory cytokines, leukocyte adhesion, cell proliferation and migration in HAECs, suggesting a protective role of LPP3 against endothelial dysfunction in humans. Both the catalytic and RGD functional domains were involved and S1P degradation was the main target of the LPP3 catalytic domain.

Abbreviations

LPP3, Lipid Phosphate Phosphatase 3; HAECs, Human Aortic Endothelial Cells; PPAP2B, Type 2 Phosphatidic Acid Phosphatase β ; LPA, Lysophosphatidic Acid; S1P, Sphingosine-1-Phosphate; C1P, Ceramide 1-Phosphate; PA, Phosphatidic Acid; VEGF, Vascular Endothelial Growth Factor; RGD, Arginine-Glycine-Aspartate; RAD, Arginine-Alanine-Aspartate; PLA2G4A, Phospholipase A2, Group IVA; PTGS2, Prostaglandin-Endoperoxide Synthase 2 (Cyclooxygenase 2); IL1b, IL6 And IL8, Interleukin-1 β , -6, -8

MCP1, Monocyte Chemoattractant Protein 1 (CCL2 Chemokine (C-C Motif) Ligand 2); SELE, Selectin E; VCAM1, Vascular Cell Adhesion Molecule 1; ICAM1, Intercellular Adhesion Molecule 1; DMSO, Dimethylsulfoxide; KDR, Kinase Insert Domain Receptor; LEF1, Lymphoid Enhancer-Binding Factor 1; MAPK, Mitogen Activated Protein Kinase; EDG1, Endothelial Differentiation Gene 1. ANGPT2; Angiopoietin 2.

Introduction

The onset of atherosclerosis is caused by several risk factors(1), altering a proper function of endothelial cells, including their dysfunction which is a marker of its initiation and progression(2). Under quiescent conditions, endothelial monolayer is relatively permeable to macromolecules and maintains an anti-inflammatory, anti-thrombotic and anti-proliferative characteristics(3,4). Therefore, the endothelial cells have a key role in the maintenance of vascular integrity as their dysfunction or activation leads to a breakdown of vascular homeostasis, leading to increased vascular permeability, leukocyte adhesion and inflammation all of which underlie atherosclerosis(5).

LPP3 is a membrane-bound enzyme capable of degrading various lipid phosphate mediators. LPP3 is encoded by the *PPAP2B* gene which is located in a locus associated with the risk of coronary artery disease (CAD) in recent Genome-Wide Association Studies (GWAS)(6). The major allele A (f=0.91) of the rs17114036 variant located in the 5th intron of *PPAP2B* was associated with an increased risk of CAD (odds ratio, 1.17; $P=3.81 \times 10^{-19}$)(6). The major allele of rs6588635, a proxy SNP of rs17114036, has been shown to be associated with lower *PPAP2B* mRNA level in human aortic endothelial cells (HAECs)(7). In the same study, an upregulation of *Ppap2b* expression in aortic endothelial cells and in atherosclerotic lesions isolated from *Apoe*^{-/-} mice was reported(7).

Subsequently *in vitro* and *in vivo* studies performed by Panchatcharam *et al.*(8,9) in mice, suggested a role of LPP3 in atherosclerosis. First, targeted inactivation of *Ppap2b* in smooth muscle cells enhanced intimal hyperplasia and promoted dedifferentiation. After vascular injury LPP3 expression limits cellular response to lysophosphatidic acid (LPA) and impaired smooth muscle cells phenotype modulation(8). Second, targeted *Ppap2b* inactivation in endothelial cells enhanced inflammatory responses and vascular permeability(9).

LPP3 belongs to the phosphatidic acid-phosphatase (PAP) enzymes family. In humans, three PAP genes are known: *PPAP2A*, *PPAP2C* and *PPAP2B*, which encode LPP1, LPP2 and LPP3, respectively(10,11). LPP3 is an integral membrane protein with six transmembrane domains localized both to plasma and intracellular membranes. The third hydrophilic loop contains the catalytic domain that localizes to the extracellular (plasma membrane) or luminal (internal membrane) surface. *Ppap2b* inactivation in mice results in early embryonic lethality due to impaired vasculogenesis, indicating that LPP3 is essential for normal vascular development(12,13) [for review(14)]. LPPs dephosphorylate and thereby terminate the signalling of a broad range of lipid substrates including phosphatidic acid (PA), lysophosphatidic acid (LPA), ceramide 1-phosphate (C1P), sphingosine 1-phosphate (S1P), and diacylglycerol pyrophosphate(15,16). LPP3 dephosphorylates LPA and S1P which are implicated in several signalling pathways during atherosclerosis(17), by binding to their specific receptors (S1PR1 to S1PR5 and LPAR1 to LPAR5) that belong to the G-protein coupled receptors family. S1P contributes to vascular development and endothelial barrier functions through the regulation of cellular proliferation, differentiation, migration and angiogenesis and acts as an intra- and extra-cellular mediator(18–21). LPA is involved in cell migration, proliferation and differentiation, contributing to neovascularization and to the induction and the release of proteases leading to cell invasiveness and suggesting that LPA may also play a role in metastasis(22,23). Human LPP3 possesses a cell adhesion motif of arginine-glycine-aspartate (RGD) that allows interactions with $\alpha\beta3$ and $\alpha5\beta1$ integrins(24,25) promoting cell adhesion and intracellular signalling.

We focused our study on primary human aortic endothelial cells (HAEC). These constitute, the first barrier of the vessel wall and exhibit dysfunction in atherosclerosis. In addition, a recent expression quantitative trait locus (eQTL) of primary human study of HAECs showed that the proxy SNP rs6588635 of the risk allele rs17114036 ($r^2 = 0.83$)(7) in *PPAP2B* was associated with the decreased expression of LPP3.

On the other hand, when using the public eQTL database (GTEx portal), we found that the rs17114036 in the *PPAP2B* locus was not associated with the expression of LPP3 in various other tissues, thus pointing out the specific role of LPP3 in HAECs.

Currently there are no functional studies on the role of LPP3 in human atherosclerosis and for this reason we investigated the role of LPP3 in endothelial function in HAECs, combining LPP3 knockdown, overexpression and targeted mutation strategies. To determine the downstream mechanisms implicating LPP3 in endothelial dysfunction, we focused our study on two LPP3 substrates, S1P and LPA, which are known to be implicated in endothelial cell function, mainly in human umbilical vein endothelial cells(26,27).

Materials and Methods

Cell culture and transfections

Human aortic endothelial cells (HAECs) were either purchased (PromoCell, Germany) or isolated from aortas of donors during heart transplantation at the University of California). Cells were cultured in Endothelial Cell Growth Media (PromoCell) supplemented with growth factors in a humidified 5% CO₂ atmosphere at 37°C. All assays were performed using HAECs from at least 3 different donors.

All protocols involving humans (Human aortic endothelial cells; HAEC) were approved by UCLA Institutional Review Board and NIH, were conform to the Helsinki Declaration of 1975 as revised in 1983 and a written consent was given for the use of the human material.

ELISA assays for cytokines

ELISA assays were performed in cell supernatants to detect soluble forms of IL1 β , IL6 and IL8 using respective human DuoSet ELISA kits.

Human peripheral blood mononuclear cells adhesion assay

Human peripheral blood mononuclear cells (hPBMC) were isolated from a single-donor Buffy coats using Ficoll density gradient. Cells were labelled with calcein-AM and incubated with HAECs monolayers

Cell viability assay

Cell viability was performed using the tetrazolium salt WST-1 assay. Cells were treated with the adequate treatments during the last 24h prior to addition of WST-1 solution for 3h, the absorbance was measured at 450 nm.

Cell proliferation assay

Cell proliferation was assayed using the addition of 5-bromo-2'-deoxyuridine (BrdU) to the plates for 6h and the absorbance was quantified at 450 nm.

Cell migration

Transfected HAECs were cultured to reach confluence. Linear scratches were performed in the cell monolayer using a 1000µl pipette tip and then the cells were incubated for 24h. The wounded area was analysed using ImageJ software by quantification of the surface of wounded area at 16h as compared to 0h.

Angiogenesis tube formation in Matrigel

Angiogenesis assays were performed using 15-well µ-angiogenesis slides (Ibidi, Germany). The slides were coated with Matrigel and the cells were seeded at 10.000 cells/well and incubated for 24h. Images of newly structured tubes were captured and tube formation was quantified by counting the number of tubular and branching point structures.

Genome-Wide Expression analysis and pre-processing of expression data

Transcriptomic analysis of total RNA was performed using the *Illumina* HT-12 v4 BeadChip. Briefly, RNA was extracted from HAECs and 250 ng of total RNA was reverse transcribed, amplified and biotinylated using the *Illumina* TotalPrep RNA Amplification Kit. Each biotinylated cRNA was hybridized to a single BeadChip and scanned using the *Illumina* HiScan array.

Determination of S1P levels

S1P after derivatization was analyzed with a Shimadzu HPLC system using an RP 18 Kromasil column and detection using a Shimadzu spectrofluorometer with an excitation wavelength of 334 nm and an emission wavelength of 440 nm. S1P was quantified by comparison of its fluorescent signal with that of the derivative of the internal standard (coefficients of variation <5%).

Statistical analysis

All data are presented as mean \pm standard deviation (SD) from at least 3 independent donors. Comparison between siRNA or overexpressing plasmids with their respective control was performed using Student test, depending on the experiment, with a significance threshold at $p \leq 0.05$.

Results

LPP3 is expressed in human vascular wall

Immunostaining of human atherosclerotic aortas (tissues microarrays) showed that LPP3 is expressed in endothelial cells, intimal macrophages and the smooth muscle cells of the media (Figure S1). The LPP3 transcript was also present in human atherosclerotic carotid tissue as assessed by RT-qPCR (data not shown).

Transcriptional analysis of HAEC silenced for LPP3

To investigate the function of human LPP3 we proceeded by knockdown or overexpression strategies in HAECs from 5 donors. Two LPP3 siRNAs (siLPP3(1) and siLPP3(2)) or LPP3 overexpression plasmids were transiently transfected into HAECs. mRNA knockdown or overexpression was confirmed by RT-qPCR, and its impact on LPP3 protein level was evaluated by immunoblotting. siLPP3(1), suppressed the expression of LPP3 mRNA by 80% to 90% and the corresponding protein level by 40% compared to the control siRNA (Figure 1A and 1B). siLPP3(2) was less effective as it suppressed the expression of LPP3 mRNA by about 60% (Figure S2) compared to the control siRNA, and for this reason we preferentially used siLPP3(1) for the functional analysis. When LPP3 was overexpressed in HAECs, a 4 to 5-fold increase in mRNA and 2.6-fold in protein levels were observed (Figure 1C and 1D).

We explored the consequences of siRNA-mediated LPP3 silencing by global gene expression profiling. cDNAs were labelled and hybridized to the HumanHT-12 v4 Expression BeadChip which targets more than 47,000 probes. Using an adjusted p-value threshold of 0.05, we found 1941 genes that were differentially expressed in siLPP3 transfected HAECs in comparison to the control siRNA (scrambled). Among the genes differentially expressed, 342 had a fold change higher than 1.5. Statistical enrichment of specific biological processes and molecular functions annotated in Gene Ontology (GO <http://www.geneontology.org/>) was assessed by Gene Set Enrichment Analysis (GSEA) using the Genetrail software. Significant pathways were identified (FDR-corrected p-values <0.05). Pathway's analysis of differentially expressed genes highlighted several GO categories of biological processes. Among them, cell cycle, cell adhesion, cell migration, angiogenesis, blood vessel development, leukocyte activation and differentiation and inflammatory response were found (Table S2). All these signalling pathways are related to either endothelial cell dysfunction or neovessel formation.

LPP3 downregulates VEGF expression

VEGF and several proinflammatory cytokines induce LPP3 expression in mice(28). We showed that both LPP3 mRNA and protein levels were upregulated, 1.5- and 2.7-fold, respectively, in HAECs

after 24h treatment with VEGF165 (Figure 2A). Next, we evaluated the expression of VEGF mRNA in HAEC silenced for LPP3 or overexpressing it. We observed a 1.5-fold increase in VEGF expression in cells silenced for LPP3 (Figure 2B) and a 30% decrease of VEGF expression in LPP3 overexpressing cells (Figure 2C). In addition LPP3 knockdown increased the expression level of the VEGFA receptor (KDR) 1.3-fold in our microarray data (adj. P.value=0.0005) consistent with VEGF expression changes. These observations indicate a retrocontrol mechanism between VEGF and LPP3 in HAECs.

siLPP3 upregulates the expression of pro-inflammatory cytokines

To determine the effect of LPP3 level on inflammatory cytokines in HAECs, we assessed the mRNA levels of 3 major proinflammatory cytokines: IL1 β , IL6 and IL8. We found that siLPP3 significantly increased the mRNA levels of these cytokines, reaching 4-, 2.5- and 15-fold, respectively as compared to control siRNA (Figure 3A), whereas, LPP3 overexpression did not change their expression (Figure S4); suggesting that the presence of basal levels of LPP3 are enough to maintain low level of inflammatory cytokines. Next, in order to assess whether the upregulated cytokine genes led to increased secretion we measured their presence in the culture media using ELISA tests. As shown in Figure 3B, both the levels of both IL6 and IL8 were increased in the media of LPP3 silenced HAECs as compared to control siRNA, however, IL1 β level was not affected by siLPP3 which may suggest that LPP3 effect is independent of the inflammasome(29) or that an additional signal maybe be required for the secretion. The inflammasome is a multiprotein complex implicated in inflammatory disorders, including atherosclerosis and its activation leads to the cleavage of pro-caspase 1 and results in the activation of IL-1 β (30).

Among the genes affected by the LPP3 silencing, PTGS2 is one of those with the highest fold change of expression. PTGS2 in concert with PLA2G4A, synthesize prostaglandin E2 (PGE2) which is involved in vascular inflammation, permeability, cell proliferation and migration(31), in addition both enzymes are implicated in atherosclerosis and cardiovascular disease . We tested the impact of LPP3 on PLA2G4A and PTGS2 (or COX2) expression in HAECs, and observed that siRNA-mediated LPP3 silencing increased mRNA level of both genes, and inversely LPP3 overexpression decreased PTGS2 level, but not PLA2G4A (Figure S5) suggesting that LPP3 exerts an anti-inflammatory effect in HAECs.

We tested whether the expression of the above genes could be affected by the rs17114036 in *PPAP2B*, in a collection of HAECs from 147 donors (119 homozygotes for the major allele AA and 28 heterozygotes AG). We compared the correlations of LPP3 expression with these genes, according to their genotype. Indeed, we found a strong correlation between LPP3 expression and IL6, IL8, PLA2G4A, PTGS2, VCAM1 and SELE only in HAECs bearing the protective AG allele. The AA allele expressing cells showed no correlation between *PPAP2B* and these genes. Correlation values of *PPAP2B* with selected genes are represented in (Figure S6), although in statistical

analysis, only the correlation between *PPAP2B* and *IL8* and *PTGS2* remained significant. These data support the hypothesis of the protective role of LPP3 in inflammatory response.

Role of LPP3 in leukocyte adhesion

As a consequence of vascular inflammation, leukocytes are recruited to endothelial surface and transmigrate into the intimal space of the arteries. Induction of adhesion molecules and chemokine expression in response to inflammatory stimuli plays a critical role in leukocyte adhesion to the endothelium(5). To investigate whether or not LPP3 expression impacts leukocyte adhesion, we first evaluated adhesion molecule and chemokine expression in HAECs. Using RT-qPCR, we measured the expression of 3 major adhesion molecules, SELE, VCAM1 and ICAM1, and the chemokine MCP1. We observed that siLPP3 induced 36-, 2-, 6- and 4-fold increases in SELE, VCAM1, ICAM1 and MCP1 mRNA levels, respectively, as compared with the control siRNA (Figure 4A). LPP3 overexpression significantly decreased SELE and ICAM1 mRNA as compared to the control plasmid, whereas VCAM1 and MCP1 mRNA were not affected by LPP3 overexpression (Figure 4B). We further investigated whether LPP3 expression affects the recruitment of leukocytes, after transfection of HAECs with either siLPP3 or LPP3 plasmids. HAECs were incubated for 1h with fluorescent calcein AM-labelled human blood leukocytes, and as shown in Figure 4C and D, leukocyte attachment was markedly enhanced when the cells were transfected with siLPP3 as compared to control siRNA. Conversely, overexpression of LPP3 significantly decreased leukocyte adhesion to the HAECs monolayer, as compared to control plasmid. These observations are concordant with the increased expression of adhesion molecules in siLPP3 transfected cells.

LPP3 downregulates cell survival

In order to assess the role of LPP3 in proliferation and survival of endothelial cells, we performed the mitochondrial activity WST1 assay to determine cell viability and a BrDU incorporation assay to measure cell proliferation rate. LPP3 overexpression strongly decreased proliferation and survival of HAECs (Figure 5A). In contrast, LPP3 silencing significantly increased cell viability, but did not change the proliferation rate. To better assess the LPP3 function in cell viability, HAECs were transfected with plasmids carrying *PPAP2B* cDNA mutations obtained by site-directed mutagenesis in 2 major functional domains. Two mutations H249P or H251P were introduced in the catalytic domain, and one mutation was introduced into the ligand domain by transforming RGD to RAD. Both H249P and H251P induced a significant reduction in cell survival as compared to the wild-type (Figure 5B), indicating that the catalytic domain is involved in the regulation of cell survival. Next, we observed that the RGD->RAD mutation significantly reduced cell survival compared to the wild-type LPP3. This reduced cell survival was accompanied by a significant decrease in cell proliferation rate

(Figure S3). Overall, these results demonstrate that LPP3 downregulates cell proliferation and survival and that both the catalytic and the RGD domain are required for this activity.

To investigate which LPP3 substrate is involved, we performed the WST1 assay on HAECs under- or overexpressing LPP3. siLPP3 transfected cells were treated with 1 μ M of either PF8380 (an inhibitor of LPA synthesis) or FTY720 (a S1P receptors inhibitor), while LPP3 overexpressing cells were treated with LPP3 substrates, 1 μ M (2S)-OMPT (LPA mimetic resistant to dephosphorylation) or S1P (S1P receptor agonist). As indicated above, siLPP3 increased WST1 labelling (Figure 5C) but treatment of the silenced cells with either PF8380 or FTY720 prevented the siLPP3-induced increase of cell survival which was even lower than that of the control siRNA transfected cells (Figure 5C). Conversely, when HAECs were transfected with an LPP3 plasmid, neither (2S)-OMPT nor S1P compensated the decrease of survival induced by LPP3 overexpression (Figure 5D). We conclude that LPP3 silencing substantially increases the viability of HAECs and supplementation with specific drugs which impede LPA or S1P signalling result in a profound inhibition of cell viability. In contrast, overexpression of LPP3 significantly decreases cell survival and this effect cannot be compensated by adding LPA nor S1P, suggesting that the RGD motif plays a key role in the regulation of proliferation in HAECs induced by LPP3.

LPP3 downregulates cell migration

Migration of endothelial cells is a key initiating event in the formation of new blood vessels. To investigate whether LPP3 is involved in HAECs migration, we performed wound healing assays. To compare the differences in migratory behaviour, the surface of the wounded area was determined after 16 h of various treatments. When HAECs were silenced for LPP3 the migration distance in the dish was increased 1.6-fold as compared to the siRNA control, whereas LPP3 overexpression decreased the wounded area up to 50% as compared to the control plasmid (Figure 6A). To determine whether the catalytic or the ligand domain of LPP3 was involved in cell migration, we used the same approach as described above to study cell survival. When LPP3 was silenced, treatment of HAECs with PF8380 slightly decreased the migration distance of HAECs, but it still remained higher than that in the siRNA control cells, indicating that inhibition of LPA is not sufficient to restore LPP3 silencing effect. When siLPP3 cells were treated with FTY720 the wounded area was strongly reduced as compared to siLPP3 (30%), but also in comparison with the control, thus pointing to an important role for S1P in cell migration in this model (Figure 6B). When (2S)-OMPT was added to LPP3 over-expressing cells, no change in the migration was observed, as compared to the control plasmid (Figure 6C). This observation is in agreement with our results obtained with PF8380, suggesting that LPA production is not required for endothelial cell migration. In contrast, addition of S1P to HAECs overexpressing LPP3 increased the migration distance 3.6-fold as

compared to control plasmid. Together, these results suggest that LPP3 may downregulate HAECs migration through S1P dephosphorylation.

LPP3 regulates angiogenesis

To investigate whether LPP3 participates to the formation of tubes, a hallmark of angiogenesis, we performed a matrigel tube formation assay in HAECs under- or overexpressing LPP3. At early times (3-6h), LPP3 silenced cells started to organize into tube structures more effectively than the siRNA control cells and this was delayed when LPP3 was overexpressed (Figure S7A), consistent with our results obtained with viability and migration. As shown in Figure 7A, after 24 hours, LPP3 silencing strongly decreased tube formation (-63%) as compared to the control siRNA (Figure 7B), suggesting that the angiogenic structures formed at 6h became disorganized. On the contrary, overexpression of LPP3 increased the formation of tubes by 50% as compared to the control plasmid (Figure 7C). Branching point structures number was regulated in the same proportion as the number of tubes (data not shown). In order to determine which LPP3 substrate is involved in the regulation of angiogenesis, we treated the LPP3 silenced cells with the PF8380 and showed that LPA did not affect the number of tube structures as compared to the control (DMSO), whereas FTY720 treatment restored the number of tubes in LPP3 siRNA transfected cells above the level of control siRNA (Figure 7B). Treatment of LPP3 overexpressing cells with either S1P or (2S)-OMPT strongly decreased LPP3-induced angiogenesis (Figure 7C). The tube formation was slightly decreased also in control cells. Next, we investigated which of the LPP3 functions is responsible for this regulation using either catalytic or RAD LPP3 mutants. We observed that both catalytic mutants (H249P or H251P) as well as RAD mutant decreased the number of tubes as compared to the plasmid carrying either the wild-type LPP3 or control plasmid (Figure 7D). Our results suggest that both RGD ligand and the enzymatic domains of LPP3 are involved in LPP3-regulated angiogenesis. Among LPP3 substrates, S1P is likely the one implicated in LPP3-mediated regulation of angiogenesis.

LPP3 degrades both intra- and extracellular S1P

We showed that S1P, but not LPA, was the major substrate of LPP3 involved in cell migration and angiogenesis. To determine whether S1P is degraded by LPP3 under our experimental conditions, and to assess if LPP3 dephosphorylates preferentially extra- or intracellular pool of S1P we measured its level, using reverse phase HPLC, in both culture media and cellular extracts from LPP3-silenced or overexpressing HAECs. The S1P content in intact HAECs was low, ranging from 3.1 to 4.2 pmol/10⁶ cells (intracellular) and 1.1nM (extracellular). Notably, siRNA-mediated LPP3 silencing lead to intra- (Figure 8A) and extracellular increases of S1P levels (Figure 8B), whereas, intracellular levels of S1P were decreased in LPP3 overexpressing cells. The concentration of S1P in supernatants was not detectable. These data indicate that both intracellular and extracellular S1P is dephosphorylated by LPP3.

Discussion

Endothelial cell dysfunction is a major mechanism in the initiation of the inflammatory process of atherosclerosis. In the present study, we demonstrated that LPP3 constitutes an important component of human endothelial cell physiology and when its expression is impaired it may affect endothelial function and likely atherosclerosis. Indeed, silencing LPP3 in HAECs readily enhanced inflammation, cell survival and migration, and impaired angiogenesis, whereas LPP3 overexpression reversed these effects. We demonstrated for the first time that LPP3 expression was negatively correlated with VEGF expression, and its receptor KDR. In various human cells LPP3 expression is induced by growth factors and inflammatory cytokines, including VEGF(32); our findings suggest a negative feedback control loop between VEGF and LPP3, probably via LPA. The latter was shown to induce VEGF expression through activation of the c-Myc and SP1 transcription factors(33).

We demonstrated that LPP3 silencing in HAECs promoted their pro-inflammatory response by increasing IL1 β expression and both expression and secretion of IL6 and IL8. This suggests that LPP3 is a negative regulator of inflammatory cytokine synthesis. In addition we showed that LPP3 silencing significantly increased the expression of adhesion molecules and MCP1, which correlated with increased human blood leukocytes adhesion to HAECs' monolayers. Altogether, these results confirm that LPP3 has a key role in inflammation and vascular permeability and we show for the first time that LPP3 down-regulates leukocyte adhesion to endothelium. In HUVEC, high level of S1P induces the expression of SELE and VCAM1 via NF κ B activation(34). More recently, it was reported that targeted inactivation of Ppap2b in mice increased inflammation and vascular permeability induced by LPS mainly due to accumulation of LPA(9). LPP3 may regulate inflammatory response through dephosphorylation of lipid substrates, such as S1P and LPA, which are both implicated in IL6(35), IL8(36) and IL1 β (37) production, through MAPK and NF κ B activation(38,39). Of note, our LPP3 silencing data in HAECs showed upregulation of several genes involved in the I- κ B kinase/NF- κ B pathway compatible with its protective role in inflammation (Table S2).

In this study, a comparative analysis, in a collection of HAECs from 147 donors (119 homozygotes for the major allele AA and 28 heterozygotes AG), of the two genotypes of rs17114036 showed a strong correlation of the expression level of PPAP2B with the inflammatory genes (IL6, IL8, PLA2G4A and PTGS2) when the protective allele AG is present. However, when the AA allele is expressed, there is no correlation. During endothelial dysfunction, there is an increase of VEGF expression concomitant with the increase of these pro-inflammatory mediators and adhesion molecules. Therefore an increased LPP3 expression could be a defensive mechanism tending to reduce the expression of proinflammatory genes. The lack of correlation with the AA allele might therefore be explained by the fact that the expression of LPP3 is too low in these cells to counteract the inflammatory response.

The role of LPP3 in cell survival and proliferation is controversial. It has been shown that LPP3 promotes proliferation of tumor cells by stabilizing β -catenin and activating LEF1 signalling(40), resulting in the synthesis of cyclin D1(40) in subconfluent cells via *Wnt* pathway(41). At confluence, LPP3 contributes to the formation of cellular junctions and inhibits the activation of *Wnt* pathway(41). In the present study, we demonstrated that LPP3 overexpression strongly decreased cell proliferation and survival; however, loss of LPP3 increased the number of living cells, without any impact on the proliferation rate. In addition, silencing of LPP3 showed a small but significant decrease of expression of several cell cycle genes, including cyclin-D1. The inability of LPP3 to activate LEF1/Cyclin-D1 signalling in confluent cells can be explained by the mechanism of contact inhibition(40).

Our results on HAECs show that the regulation of survival by LPP3 does not fully depend on the regulation of the cell cycle, but rather on the regulation of the balance between apoptosis and survival. Several studies have shown that LPPs, including LPP3, may degrade LPA and S1P(42) and then alter activation of the p42/p44 MAPK pathway and activation of caspases 3/7(43). In addition, S1P inhibits the activation of caspase 3 and the apoptotic pathway downstream of ceramide(44); as postulated by the concept of "sphingolipids rheostat" that governs the cell fate leading either to survival or to apoptosis(45). Both pro-apoptotic ceramide-1 phosphate (C1P) and anti-apoptotic S1P are dephosphorylated by LPP3(17).

The exact mechanism by which LPP3 controls cell survival remains elusive; several studies point to the catalytic function and the leading role of its lipid substrates, but it is clear that this approach cannot explain the role of LPP3 in the β -catenin mediated LEF1/Cyclin-D1 proliferation. Thus, we demonstrated that the loss of either catalytic or adhesion function equally decreased the number of viable cells. Furthermore, both LPA and S1P inhibitors blocked cell survival induced by LPP3 silencing, consistent with the previous studies showing that LPP3 inhibits survival via degradation of its lipid substrates. However, when LPP3 was overexpressed, neither LPA nor S1P was able to reverse the inhibition of cell survival, indicating that the RGD domain was involved. Although, the precise mechanism remains to be defined, we suggest that LPP3 regulates cell survival via both catalytic and RGD domains, possibly by steric hindrance due to the proximity of the two domains in the cellular junctions.

Angiogenesis requires proliferation, migration, adhesion and formation of branching point structures. Here we showed that LPP3 downregulates cell migration as measured by wound healing assay. Since the S1P inhibitor blocked siLPP3-induced migration and conversely, the addition of S1P analogue reversed the inhibition of migration mediated by LPP3 overexpression, we concluded that the anti-migratory effect of LPP3 in HAECs depends solely on the degradation of S1P. This is consistent with former studies which demonstrated that S1P stimulates migration through its

receptor EDG1 in human umbilical endothelial and smooth muscle cells(46,47). However, Panchatcharam *et al.*(8) showed that LPP3 is involved in smooth muscle cell migration in mice through LPA. This discrepancy could be due to different cells and different species.

The first evidence implicating of LPP3 in angiogenesis came from the observation that mouse embryos lacking *Lpp3* die at E7.5 due to abnormal vascular development(13). LPP3 is a key element in cell-cell interaction due to its RGD motif which interacts with integrins(48), and it allows bFGF- and VEGF-induced angiogenesis(28). Our work demonstrated that LPP3 downregulates the number of tubes and branching point structures, consistent with its pro-angiogenic role. We showed that both the catalytic and RGD domains were necessary to sustain angiogenesis. When LPP3 was silenced, inhibition of S1P, but not LPA formation, restored the angiogenesis. These results suggest that LPP3 regulates angiogenesis by controlling the availability of S1P.

Our results showed that despite the fact that LPP3 silencing increases the expression of the pro-angiogenic factors VEGF and ANGPT2 (Figure S7B and transcriptomic data), the angiogenic response was strongly decreased; this is probably due to the excess of accumulated S1P in the absence of LPP3, and the upregulation of S1PR1 (EDG1) expression (as shown in our transcriptomic data). Although *Edg1* knockout mice undergo normal vasculogenesis, the loss of *Edg1* gene is lethal *in utero* due to defect in vessel maturation(49), indicating that S1P is indispensable for vessel stabilization. Several studies demonstrated the role of S1P in vascular stabilization and termination of angiogenesis(50), through the inhibition of VEGFA-induced sprouting and the recruitment of mural cells(50). There is a gradual expression of S1PR1 from the mature sprouting region to the growing vascular front(51). When the blood vessels are formed, S1P activates S1PR1 which in turn inhibits VEGF signalling and suppresses sprouting, promotes junction formation and recruitment of mural cells to prevents excessive sprouting(51,52).

Our model suggests that during early stages of angiogenesis, LPP3, due to its enzymatic activity, promotes sprouting through the downregulation of S1P and S1PR1. This is consistent with our finding that following LPP3 silencing, S1P promotes cell proliferation and migration; but it inhibits angiogenesis, due to its putative accumulation, which blocks sprouting. Later, when vascular network is achieved, LPP3 interacts with integrins through the RGD motif to form adherence junctions(41). When the RGD domain is engaged in cell-cell adhesion, this may inhibit its catalytic activity, leading to an increase of S1P level.

To summarize, our studies show that LPP3 downregulates both inflammation and leukocyte adhesion in human aortic endothelial cells by reducing expression of cytokines and adhesion molecules. LPP3 also downregulates cell growth either by degrading LPA and S1P or via the inhibition of cell-to-cell contacts via its RGD domain. LPP3 inhibits cell migration by controlling the

levels of S1P and the expression of its receptor EDG1, limiting hypersprouting induced by enhanced expression of VEGFA; of note, in atherosclerosis, impaired activation of endothelial cells may promote angiogenesis in order to maintain plaque irrigation and local inflammation, thus participate to plaque rupture. LPP3 may have an important role in the switch between physiological and pathological angiogenesis.

In conclusion, our results indicate that the role of LPP3 in proper endothelial function involves both its RGD ligand domain and catalytic site and a fine regulation of both intracellular and extracellular pools of lipid substrates.

Acknowledgements

We are grateful to Pr. F. Koskas (Vascular Surgery Department, Groupe Hospitalier Pitié-Salpêtrière, Paris) for the human vessel arteriosclerotic samples, A. Lessot and Dr. F. Charlotte (Pathology Department Groupe Hospitalier Pitié-Salpêtrière, Paris) and Pr. I. Brochériou (Hopital Tenon, Paris) for tissue microarrays preparation, immunohistochemistry staining and interpretation. We thank the genomic P3S platform for the microarrays and QPCR data acquisition and western blot scanning.

Dr. V. Baecker, ImageJ User and Developer Conference 2012. Luxembourg: Centre de Recherche Public Henri Tudor; 2012; ISBN: 2-919941-18-6

Funding

Supported by the Institut National de la Santé et de la Recherche Médicale and the Transatlantic Networks of Excellence, Fondation Leducq (12CVD02) (France); HL28481 and: K99HL121172 from the National Institutes of Health (USA).

E. Ninio is Director of Research in Centre National de la Recherche Scientifique.

Legends of figures

Figure 1. Lipid phosphate phosphatase 3 (LPP3) expression in HAEC. HAEC were transfected with either siRNA (siCtrl or siLPP3) (A and B) or with overexpressing plasmids (Ctrl or LPP3) (C and D) for 48 h. The level of LPP3 mRNA was determined by RT-qPCR. The results are shown as relative LPP3 mRNA level over the controls and represented as mean \pm SD from 5 independent donors in duplicate (A, C). Lysates from HAEC transfected either with siLPP3 or siCtrl or LPP3-containing or Ctrl plasmid were analyzed by immunoblotting with rabbit polyclonal anti-LPP3 antibody (B, D) and re-probed with anti-tubulin antibody to ensure equal loading and transfer. Immunoblot from one experiment. *: p<0.05, **: p<0.001, *** p<0.0001.

Figure 2. LPP3 negatively regulates VEGF expression in HAEC. HAEC were transfected with either siRNA (siCtrl or siLPP3) or with overexpressing plasmids (Ctrl or LPP3) for 48 h. Cells were treated with 100nM VEGF-165 for additional 24h. LPP3 mRNA levels were assessed by RT-qPCR (A) and LPP3 protein levels by immunoblotting with rabbit polyclonal anti-LPP3 antibody, subsequently re-probed with anti-tubulin antibody. Immunoblot from one experiment (B). Relative levels of VEGF mRNA under siLPP3 knockdown (C) or LPP3 overexpression with LPP3-containing plasmid (D) were assessed by RT-qPCR. The results are shown as mean \pm SD from 5 independent donors in duplicate; *: $p < 0.05$, **: $p < 0.001$, *** $p < 0.0001$.

Figure 3. siLPP3 upregulates the expression of pro-inflammatory cytokines in HAEC HAEC were transfected with either siRNA (siCtrl or siLPP3) or with overexpressing plasmids (Ctrl or LPP3) for 48 h. mRNA relative levels of IL1 β , IL6 and IL8 were determined by RT-qPCR. The results are expressed as a mean \pm SD from 5 independent donors in duplicate (A). The cytokine concentrations were determined using ELISA in supernatants of HAECs cultured for 24 hours in serum free media. The results are expressed as a mean \pm SD from 3 independent donors in triplicate (B). *: $p < 0.05$, **: $p < 0.001$, *** $p < 0.0001$.

Figure 4. LPP3 decreases leukocyte recruitment. HAEC were transfected with siRNA (siCtrl or siLPP3) for 48 h, and subsequently mRNA relative levels of adhesion molecules SELE, ICAM-1, VCAM-1 (A) and MCP1/CCL2 (B) were determined by RT-qPCR. The results are shown as mean \pm SD from 5 independent donors in duplicate. Leukocyte adhesion was assessed using a PBMC adhesion assay as described in methods. HAEC were transfected with either siRNA (siCtrl or siLPP3) or with overexpressing plasmids (Ctrl or LPP3) for 48 h, and then incubated with calcein-AM labelled hPBMC. Images of fluorescent cells were captured with an epi-fluorescence microscope (C) and the adherent cells were quantified by automated counting using ImageJ and represented as a relative change over the siCtrl. The results are shown as mean \pm SD from 3 independent donors in duplicate. *: $p < 0.05$, **: $p < 0.001$, *** $p < 0.0001$. ns: no significant.

Figure 5. LPP3 affects viability of HAEC. HAEC were transfected with either siRNA (siCtrl or siLPP3) or with overexpressing plasmids (Ctrl or LPP3) for 48 h. Cell proliferation was estimated using the BrDU assay (gray bars) and cell viability was measured using the WST-1 assay (black bars). The results are shown as mean \pm SD from 5 independent donors in triplicate (A). HAECs were transfected with control, WT LPP3 or mutants of the catalytic domain (H249P or H251P) or of the adhesion motif (RGD \rightarrow RAD), then WST-1 assays were performed. The results are shown as mean \pm SD from 3 independent donors in triplicate (B). HAECs transfected with either siLPP3 or siCtrl were treated with inhibitors of LPA (PF8380) or S1P (FTY720) (C). HAECs transfected with LPP3-containing or Ctrl plasmid were treated with agonists of LPA ((2S)-OMPT) or S1P (D). The results are shown as % of respective controls. The results are shown as mean \pm SD from 3 independent donors in triplicate

*: $p < 0.05$, **: $p < 0.001$, *** $p < 0.0001$, ns: non significant.

Figure 6. LPP3 is a negative regulator of HAEC cell migration. . HAEC were transfected with either siRNA (siCtrl or siLPP3) or with overexpressing plasmids (Ctrl or LPP3) for 48 h. Cell monolayers were wounded with 1000 μ l pipette tips (right panel) and incubated for 16h to assess their migration. Wounded areas were imaged at 0h and 16h. Results are expressed in % of the respective controls and are shown as mean +/- SD from 5 independent donors in duplicate (A). siLPP3 and siCtrl transfected HAECs were treated with inhibitors of LPA formation (PF8380) or S1P (FTY720) (B). LPP3-containing plasmid or Ctrl plasmid transfected cells were treated with agonists of LPA ((2S)-OMPT) or S1P. The results are shown as mean \pm SD from 3 independent donors in duplicate (C). *: p<0.05, **: p<0.001, *** p<0.0001, ns: no significant.

Figure7. LPP3 is a positive regulator of angiogenesis. HAEC were transfected with either siRNA (siCtrl or siLPP3) or with overexpressing plasmids (Ctrl or LPP3) for 48 h. Cells were harvested and seeded on Matrigel and incubated for an additional 48h. Microphotographs of tube-like structures were taken at 24h. Angiogenesis was quantified by counting the tube-like and branching point structures at 24h, using an automated ImageJ program (A). siLPP3 and siCtrl transfected HAECs were treated with inhibitors of LPA formation (PF8380) or S1P (FTY720) (B). LPP3-containing plasmid or Ctrl plasmid transfected cells were treated with agonists of LPA ((2S)-OMPT) or S1P (C). HAECs were transfected with control, wild-type LPP3 or mutants of the catalytic domain (H249P or H251P) or the adhesion motif RGD->RAD (D). The results are shown as mean \pm SD and expressed as % of respective controls The results are shown as mean +/- SD from 3 independent donors in duplicate. *, †: p<0.05, ns: no significant.

†: siLPP3/siCtrl and LPP3/Ctrl

*: Treatment/DMSO

Figure 8: S1P levels in HAEC cell extracts and supernatants. HAEC were transfected with either siRNA (siCtrl or siLPP3) or with overexpressing plasmids (Ctrl or LPP3) for 48 h. Cell extracts (A) and supernatants (B) were collected. Lipids were extracted and derivative mixtures were analyzed by HPLC. The results are shown as mean +/- SD from 3 independent donors in duplicate. *: p<0.05; **: p<0.01.

References

1. Greenland P, Alpert JS, Beller GA, Benjamin EJ, Budoff MJ, Fayad ZA, Foster E, Hlatky MA, Hodgson JM, Kushner FG, Lauer MS, Shaw LJ, Smith SC, Taylor AJ, Weintraub WS, Wenger NK, Jacobs AK, American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2010;**122**:e584–e636.
2. Lerman A, Cannan CR, Higano SH, Nishimura RA, Holmes DR. Coronary vascular remodeling in association with endothelial dysfunction. *Am J Cardiol*. 1998;**81**:1105–1109.
3. Chiu J-J, Chien S. Effects of disturbed flow on vascular endothelium: pathophysiological basis and clinical perspectives. *Physiol Rev*. 2011;**91**:327–387.
4. Libby P, Ridker PM, Hansson GK. Progress and challenges in translating the biology of atherosclerosis. *Nature*. 2011;**473**:317–325.
5. Hopkins PN. Molecular biology of atherosclerosis. *Physiol Rev*. 2013;**93**:1317–1542.
6. Schunkert H, König IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, Preuss M, Stewart AFR, Barbalic M, Gieger C, Absher D, Aherrahrou Z, Allayee H, Altshuler D, Anand SS, Andersen K, Anderson JL, Ardisino D, Ball SG, Balmforth AJ, Barnes TA, Becker DM, Becker LC, Berger K, Bis JC, Boekholdt SM, Boerwinkle E, Braund PS, Brown MJ, Burnett MS, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet*. 2011;**43**:333–338.
7. Erbilgin A, Civelek M, Romanoski CE, Pan C, Hagopian R, Berliner JA, Lusis AJ. Identification of CAD candidate genes in GWAS loci and their expression in vascular cells. *J Lipid Res*. 2013;**54**:1894–1905.
8. Panchatcharam M, Miriyala S, Salous A, Wheeler J, Dong A, Mueller P, Sunkara M, Escalante-Alcalde D, Morris AJ, Smyth SS. Lipid phosphate phosphatase 3 negatively regulates smooth muscle cell phenotypic modulation to limit intimal hyperplasia. *Arterioscler Thromb Vasc Biol*. 2013;**33**:52–59.
9. Panchatcharam M, Salous AK, Brandon J, Miriyala S, Wheeler J, Patil P, Sunkara M, Morris AJ, Escalante-Alcalde D, Smyth SS. Mice with targeted inactivation of ppap2b in endothelial and hematopoietic cells display enhanced vascular inflammation and permeability. *Arterioscler Thromb Vasc Biol*. 2014;**34**:837–845.
10. Kai M, Wada I, Imai S i, Sakane F, Kanoh H. Cloning and characterization of two human isozymes of Mg²⁺-independent phosphatidic acid phosphatase. *J Biol Chem*. 1997;**272**:24572–24578.
11. Roberts R, Sciorra VA, Morris AJ. Human type 2 phosphatidic acid phosphohydrolases. Substrate specificity of the type 2a, 2b, and 2c enzymes and cell surface activity of the 2a isoform. *J Biol Chem*. 1998;**273**:22059–22067.
12. Escalante-Alcalde D, Morales SL, Stewart CL. Generation of a reporter-null allele of Ppap2b/Lpp3 and its expression during embryogenesis. *Int J Dev Biol*. 2009;**53**:139–147.

13. Escalante-Alcalde D, Hernandez L, Stunff H Le, Maeda R, Lee H-S, Jr-Gang-Cheng null, Sciorra VA, Daar I, Spiegel S, Morris AJ, Stewart CL. The lipid phosphatase LPP3 regulates extra-embryonic vasculogenesis and axis patterning. *Dev Camb Engl*. 2003;**130**:4623–4637.
14. Ren H, Panchatcharam M, Mueller P, Escalante-Alcalde D, Morris AJ, Smyth SS. Lipid phosphate phosphatase (LPP3) and vascular development. *Biochim Biophys Acta BBA - Mol Cell Biol Lipids*. 2013;**1831**:126–132.
15. Brindley DN, English D, Pilquil C, Buri K, Ling ZC. Lipid phosphate phosphatases regulate signal transduction through glycerolipids and sphingolipids. *Biochim Biophys Acta*. 2002;**1582**:33–44.
16. Sciorra VA, Morris AJ. Roles for lipid phosphate phosphatases in regulation of cellular signaling. *Biochim Biophys Acta*. 2002;**1582**:45–51.
17. Brindley DN, Pilquil C. Lipid phosphate phosphatases and signaling. *J Lipid Res*. 2009;**50 Suppl**:S225–S230.
18. Ahn EH, Schroeder JJ. Induction of apoptosis by sphingosine, sphinganine, and C(2)-ceramide in human colon cancer cells, but not by C(2)-dihydroceramide. *Anticancer Res*. 2010;**30**:2881–2884.
19. Camerer E, Regard JB, Cornelissen I, Srinivasan Y, Duong DN, Palmer D, Pham TH, Wong JS, Pappu R, Coughlin SR. Sphingosine-1-phosphate in the plasma compartment regulates basal and inflammation-induced vascular leak in mice. *J Clin Invest*. 2009;**119**:1871–1879.
20. Huang Y-L, Lin H-S, Chen S-U, Lee H. Tyrosine sulphation of sphingosine 1-phosphate 1 (S1P1) is required for S1P-mediated cell migration in primary cultures of human umbilical vein endothelial cells. *J Biochem (Tokyo)*. 2009;**146**:815–820.
21. Moriue T, Igarashi J, Yoneda K, Nakai K, Kosaka H, Kubota Y. Sphingosine 1-phosphate attenuates H₂O₂-induced apoptosis in endothelial cells. *Biochem Biophys Res Commun*. 2008;**368**:852–857.
22. Brindley DN, Lin F-T, Tigyi GJ. Role of the autotaxin-lysophosphatidate axis in cancer resistance to chemotherapy and radiotherapy. *Biochim Biophys Acta*. 2013;**1831**:74–85.
23. Samadi N, Bekele R, Capatos D, Venkatraman G, Sariahmetoglu M, Brindley DN. Regulation of lysophosphatidate signaling by autotaxin and lipid phosphate phosphatases with respect to tumor progression, angiogenesis, metastasis and chemo-resistance. *Biochimie*. 2011;**93**:61–70.
24. Jia Y-J, Kai M, Wada I, Sakane F, Kanoh H. Differential localization of lipid phosphate phosphatases 1 and 3 to cell surface subdomains in polarized MDCK cells. *FEBS Lett*. 2003;**552**:240–246.
25. Kai M, Wada I, Imai S, Sakane F, Kanoh H. Identification and cDNA cloning of 35-kDa phosphatidic acid phosphatase (type 2) bound to plasma membranes. Polymerase chain reaction amplification of mouse H₂O₂-inducible hic53 clone yielded the cDNA encoding phosphatidic acid phosphatase. *J Biol Chem*. 1996;**271**:18931–18938.
26. Wilkerson BA, Argraves KM. The role of sphingosine-1-phosphate in endothelial barrier function. *Biochim Biophys Acta BBA - Mol Cell Biol Lipids*. 2014;**1841**:1403–1412.
27. Schober A, Siess W. Lysophosphatidic acid in atherosclerotic diseases. *Br J Pharmacol*. 2012;**167**:465–482.

28. Wary KK, Humtsoe JO. Anti-lipid phosphate phosphohydrolase-3 (LPP3) antibody inhibits bFGF- and VEGF-induced capillary morphogenesis of endothelial cells. *Cell Commun Signal CCS*. 2005;**3**:9.
29. Martinon F, Burns K, Tschopp J. The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proIL-beta. *Mol Cell*. 2002;**10**:417–426.
30. Wen H, Ting JP-Y, O’Neill LAJ. A role for the NLRP3 inflammasome in metabolic diseases--did Warburg miss inflammation? *Nat Immunol*. 2012;**13**:352–357.
31. Gomez I, Foudi N, Longrois D, Norel X. The role of prostaglandin E2 in human vascular inflammation. *Prostaglandins Leukot Essent Fat Acids PLEFA*. 2013;**89**:55–63.
32. Humtsoe JO, Feng S, Thakker GD, Yang J, Hong J, Wary KK. Regulation of cell-cell interactions by phosphatidic acid phosphatase 2b/VCIP. *EMBO J*. 2003;**22**:1539–1554.
33. Song Y, Wu J, Oyesanya RA, Lee Z, Mukherjee A, Fang X. Sp-1 and c-Myc mediate lysophosphatidic acid-induced expression of vascular endothelial growth factor in ovarian cancer cells via a hypoxia-inducible factor-1-independent mechanism. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2009;**15**:492–501.
34. Xia P, Gamble JR, Rye KA, Wang L, Hii CS, Cockerill P, Khew-Goodall Y, Bert AG, Barter PJ, Vadas MA. Tumor necrosis factor-alpha induces adhesion molecule expression through the sphingosine kinase pathway. *Proc Natl Acad Sci U S A*. 1998;**95**:14196–14201.
35. Liang J, Nagahashi M, Kim EY, Harikumar KB, Yamada A, Huang W-C, Hait NC, Allegood JC, Price MM, Avni D, Takabe K, Kordula T, Milstien S, Spiegel S. Sphingosine-1-Phosphate Links Persistent STAT3 Activation, Chronic Intestinal Inflammation, and Development of Colitis-Associated Cancer. *Cancer Cell*. 2013;**23**:107–120.
36. Wang L, Cummings R, Usatyuk P, Morris A, Irani K, Natarajan V. Involvement of phospholipases D1 and D2 in sphingosine 1-phosphate-induced ERK (extracellular-signal-regulated kinase) activation and interleukin-8 secretion in human bronchial epithelial cells. *Biochem J*. 2002;**367**:751–760.
37. Liu Q, Rehman H, Shi Y, Krishnasamy Y, Lemasters JJ, Smith CD, Zhong Z. Inhibition of sphingosine kinase-2 suppresses inflammation and attenuates graft injury after liver transplantation in rats. *PLoS One*. 2012;**7**:e41834.
38. Cummings R, Zhao Y, Jacoby D, Spannhake EW, Ohba M, Garcia JGN, Watkins T, He D, Saatian B, Natarajan V. Protein kinase Cdelta mediates lysophosphatidic acid-induced NF-kappaB activation and interleukin-8 secretion in human bronchial epithelial cells. *J Biol Chem*. 2004;**279**:41085–41094.
39. Saatian B, Zhao Y, He D, Georas SN, Watkins T, Spannhake EW, Natarajan V. Transcriptional regulation of lysophosphatidic acid-induced interleukin-8 expression and secretion by p38 MAPK and JNK in human bronchial epithelial cells. *Biochem J*. 2006;**393**:657–668.
40. Chatterjee I, Humtsoe JO, Kohler EE, Sorio C, Wary KK. Lipid phosphate phosphatase-3 regulates tumor growth via β -catenin and CYCLIN-D1 signaling. *Mol Cancer*. 2011;**10**:51.
41. Humtsoe JO, Liu M, Malik AB, Wary KK. Lipid phosphate phosphatase 3 stabilization of beta-catenin induces endothelial cell migration and formation of branching point structures. *Mol Cell Biol*. 2010;**30**:1593–1606.

42. Alderton F, Darroch P, Sambhi B, McKie A, Ahmed IS, Pyne N, Pyne S. G-protein-coupled receptor stimulation of the p42/p44 mitogen-activated protein kinase pathway is attenuated by lipid phosphate phosphatases 1, 1a, and 2 in human embryonic kidney 293 cells. *J Biol Chem*. 2001;**276**:13452–13460.
43. Long J, Darroch P, Wan KF, Kong KC, Ktistakis N, Pyne NJ, Pyne S. Regulation of cell survival by lipid phosphate phosphatases involves the modulation of intracellular phosphatidic acid and sphingosine 1-phosphate pools. *Biochem J*. 2005;**391**:25–32.
44. Cuvillier O, Rosenthal DS, Smulson ME, Spiegel S. Sphingosine 1-phosphate inhibits activation of caspases that cleave poly(ADP-ribose) polymerase and lamins during Fas- and ceramide-mediated apoptosis in Jurkat T lymphocytes. *J Biol Chem*. 1998;**273**:2910–2916.
45. Brocklyn JR Van, Williams JB. The control of the balance between ceramide and sphingosine-1-phosphate by sphingosine kinase: oxidative stress and the seesaw of cell survival and death. *Comp Biochem Physiol B Biochem Mol Biol*. 2012;**163**:26–36.
46. Wang F, Brocklyn JR Van, Hobson JP, Movafagh S, Zukowska-Grojec Z, Milstien S, Spiegel S. Sphingosine 1-phosphate stimulates cell migration through a G(i)-coupled cell surface receptor. Potential involvement in angiogenesis. *J Biol Chem*. 1999;**274**:35343–35350.
47. Kluk MJ, Hla T. Role of the sphingosine 1-phosphate receptor EDG-1 in vascular smooth muscle cell proliferation and migration. *Circ Res*. 2001;**89**:496–502.
48. Humtsoe JO, Feng S, Thakker GD, Yang J, Hong J, Wary KK. Regulation of cell-cell interactions by phosphatidic acid phosphatase 2b/VCIP. *EMBO J*. 2003;**22**:1539–1554.
49. Liu Y, Wada R, Yamashita T, Mi Y, Deng CX, Hobson JP, Rosenfeldt HM, Nava VE, Chae SS, Lee MJ, Liu CH, Hla T, Spiegel S, Proia RL. Edg-1, the G protein-coupled receptor for sphingosine-1-phosphate, is essential for vascular maturation. *J Clin Invest*. 2000;**106**:951–961.
50. Carmeliet P, Jain RK. Molecular mechanisms and clinical applications of angiogenesis. *Nature*. 2011;**473**:298–307.
51. Jung B, Obinata H, Galvani S, Mendelson K, Ding B, Skoura A, Kinzel B, Brinkmann V, Rafii S, Evans T, Hla T. Flow-regulated endothelial S1P receptor-1 signaling sustains vascular development. *Dev Cell*. 2012;**23**:600–610.
52. Shoham A Ben, Malkinson G, Krief S, Shwartz Y, Ely Y, Ferrara N, Yaniv K, Zelzer E. S1P1 inhibits sprouting angiogenesis during vascular development. *Dev Camb Engl*. 2012;**139**:3859–3869.

Figure(s)

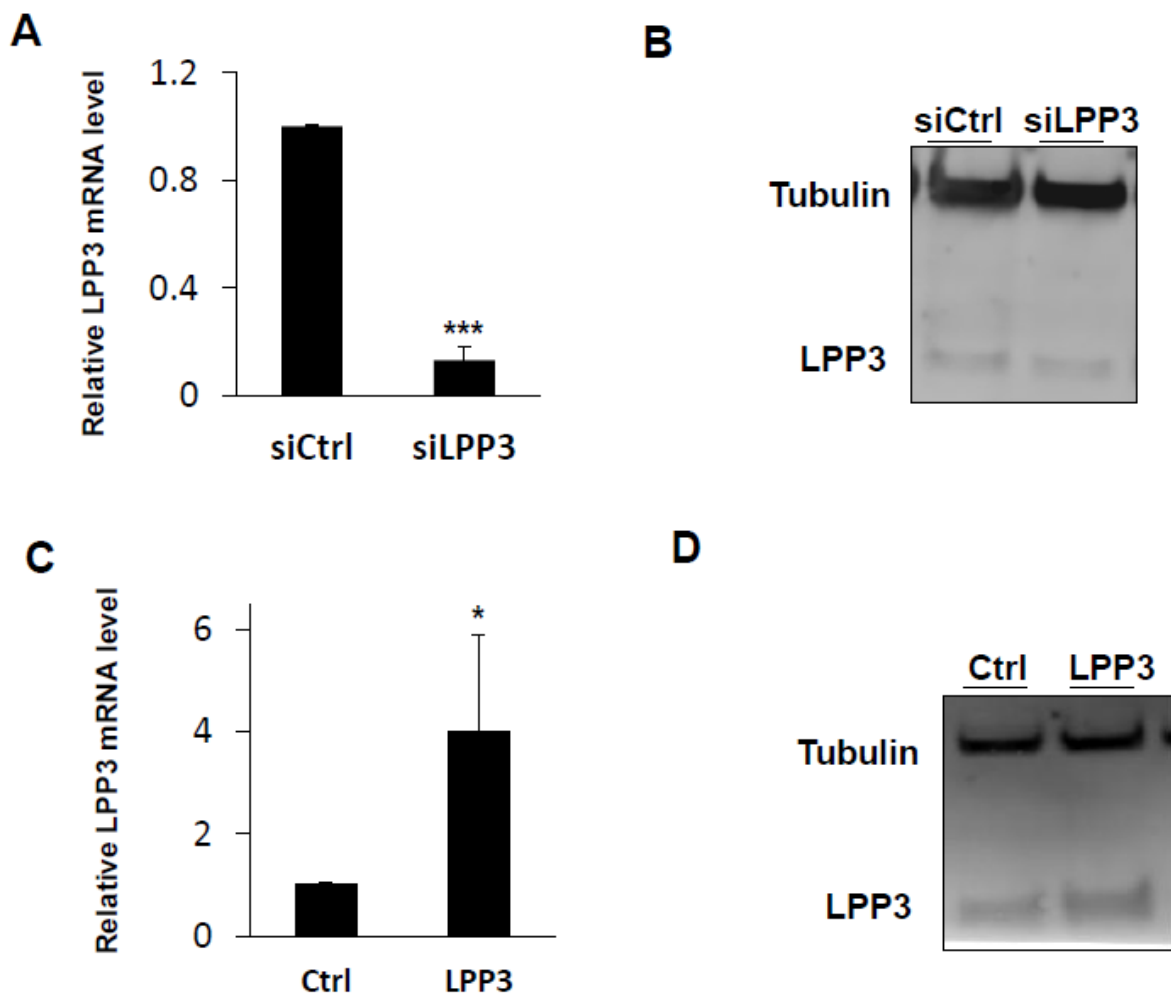


Figure 1. Lipid phosphate phosphatase 3 (LPP3) expression in HAEC. HAEC were transfected with either siRNA (siCtrl or siLPP3) (A and B) or with overexpressing plasmids (Ctrl or LPP3) (C and D) for 48 h. The level of LPP3 mRNA was determined by RT-qPCR. The results are shown as relative LPP3 mRNA level over the controls and represented as mean \pm SD from 5 independent donors in duplicate (A, C). Lysates from HAEC transfected either with siLPP3 or siCtrl or LPP3-containing or Ctrl plasmid were analyzed by immunoblotting with rabbit polyclonal anti-LPP3 antibody (B, D) and re-probed with anti-tubulin antibody to ensure equal loading and transfer. Immunoblot from one experiment. *: $p < 0.05$, **: $p < 0.001$, *** $p < 0.0001$.

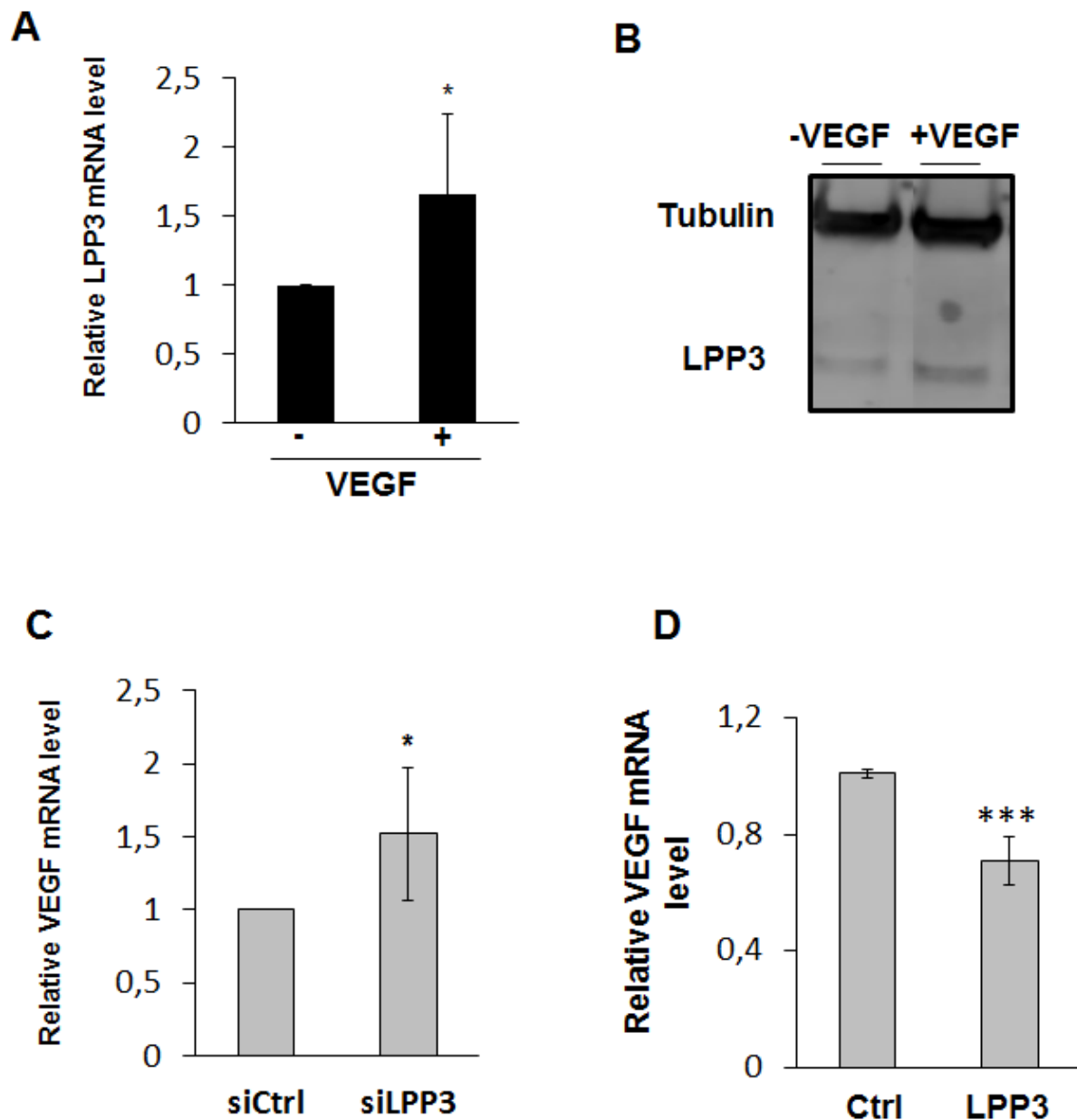


Figure 2. LPP3 negatively regulates VEGF expression in HAEC. HAEC were transfected with either siRNA (siCtrl or siLPP3) or with overexpressing plasmids (Ctrl or LPP3) for 48 h. Cells were treated with 100nM VEGF-165 for additional 24h. LPP3 mRNA levels were assessed by RT-qPCR (A) and LPP3 protein levels by immunoblotting with rabbit polyclonal anti-LPP3 antibody, subsequently re-probed with anti-tubulin antibody. Immunoblot from one experiment (B). Relative levels of VEGF mRNA under siLPP3 knockdown (C) or LPP3 overexpression with LPP3-containing plasmid (D) were assessed by RT-qPCR. The results are shown as mean \pm SD; *, $p < 0.05$, **, $p < 0.001$, *** $p < 0.0001$.

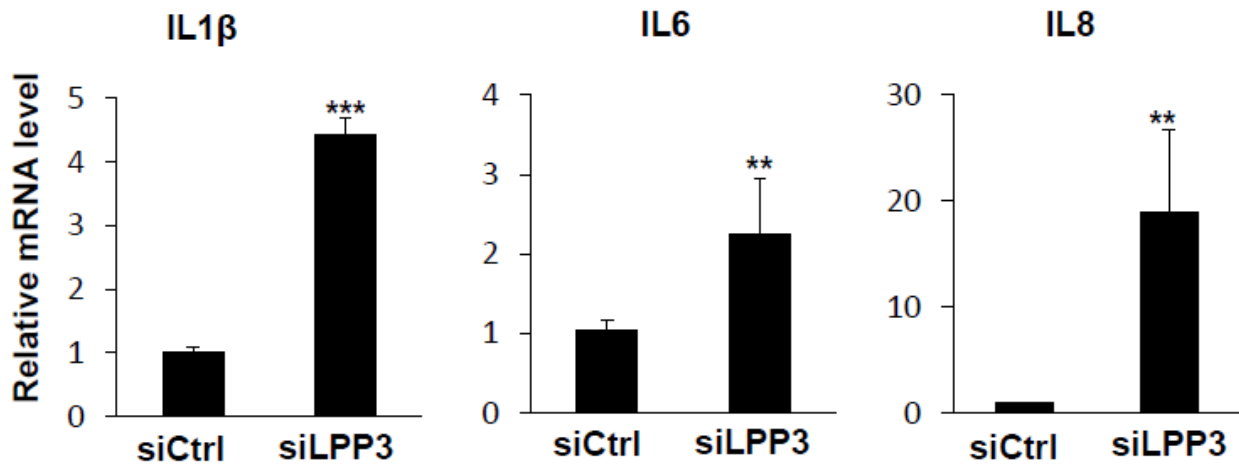
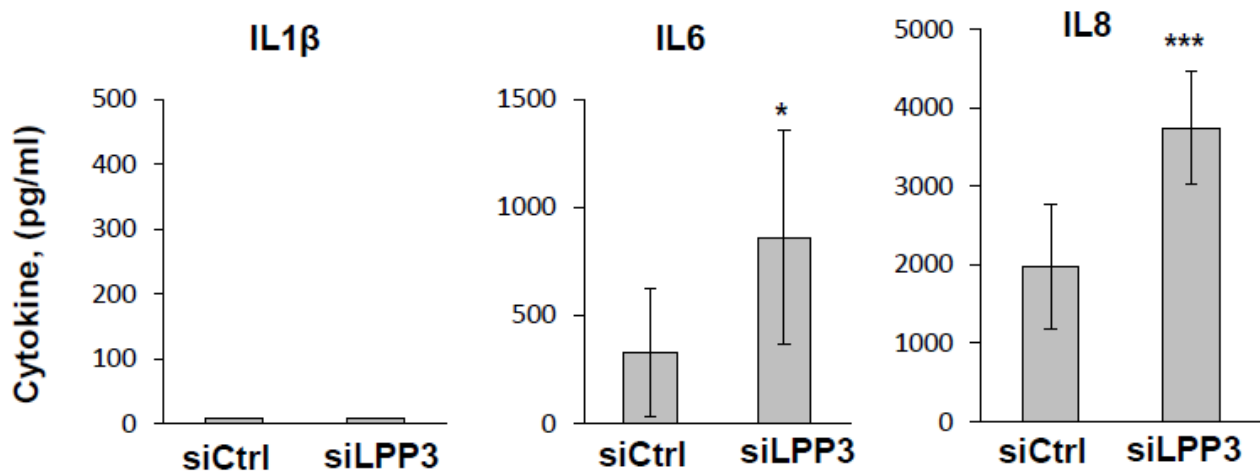
A**B**

Figure 3. siLPP3 upregulates the expression of pro-inflammatory cytokines in HAEC HAEC were transfected with either siRNA (siCtrl or siLPP3) or with overexpressing plasmids (Ctrl or LPP3) for 48 h. mRNA relative levels of IL1 β , IL6 and IL8 were determined by RT-qPCR. The results are expressed as a mean \pm SD from 5 independent donors in duplicate (A). The cytokine concentrations were determined using ELISA in supernatants of HAECs cultured for 24 hours in serum free media. The results are expressed as a mean \pm SD from 3 independent donors in triplicate (B). *: p < 0.05, **: p < 0.001, *** p < 0.0001.

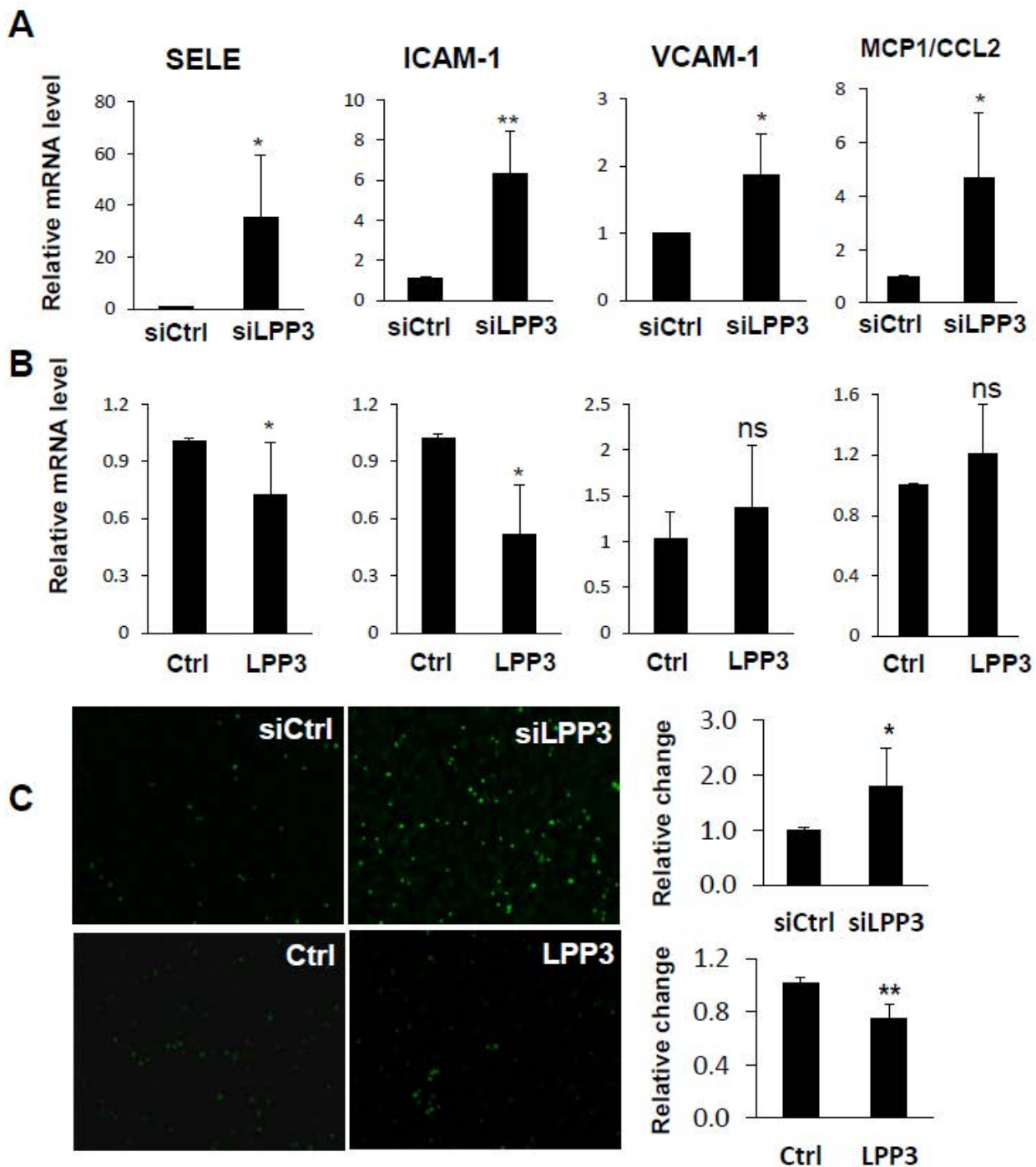


Figure 4. LPP3 decreases leukocyte recruitment. HAEC were transfected with siRNA (siCtrl or siLPP3) for 48 h, and subsequently mRNA relative levels of adhesion molecules SELE, ICAM-1, VCAM-1 (A) and MCP1/CCL2 (B) were determined by RT-qPCR. The results are shown as mean \pm SD from 5 independent donors in duplicate. Leukocyte adhesion was assessed using a PBMC adhesion assay as described in methods. HAEC were transfected with either siRNA (siCtrl or siLPP3) or with overexpressing plasmids (Ctrl or LPP3) for 48 h, and then incubated with calcein-AM labelled hPBMC. Images of fluorescent cells were captured with an epi-fluorescence microscope (C) and the adherent cells were quantified by automated counting using ImageJ and represented as a relative change over the siCtrl. The results are shown as mean \pm SD from 3 independent donors in duplicate. *: $p < 0.05$, **: $p < 0.001$, *** $p < 0.0001$. ns: no significant.

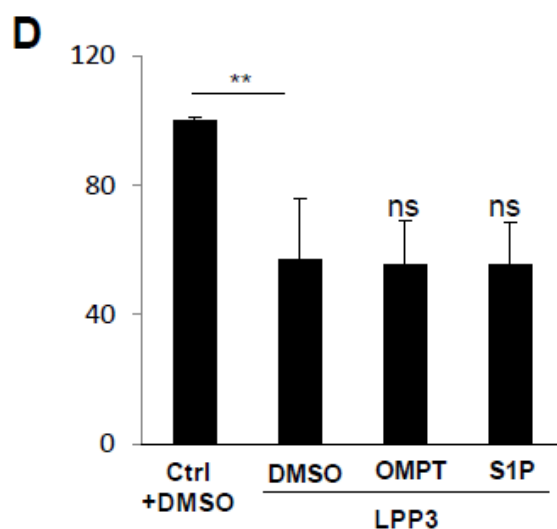
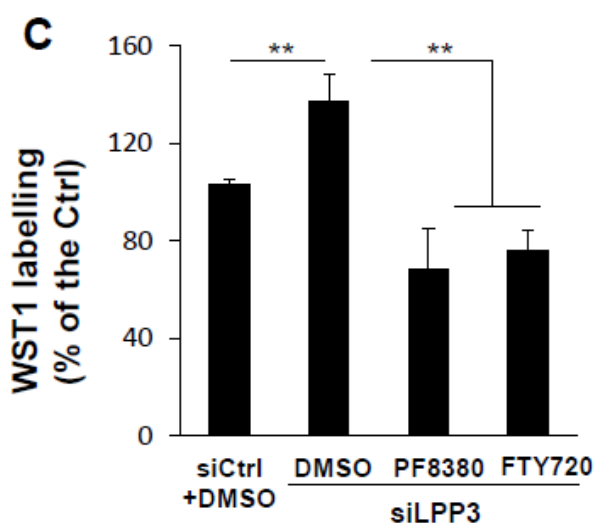
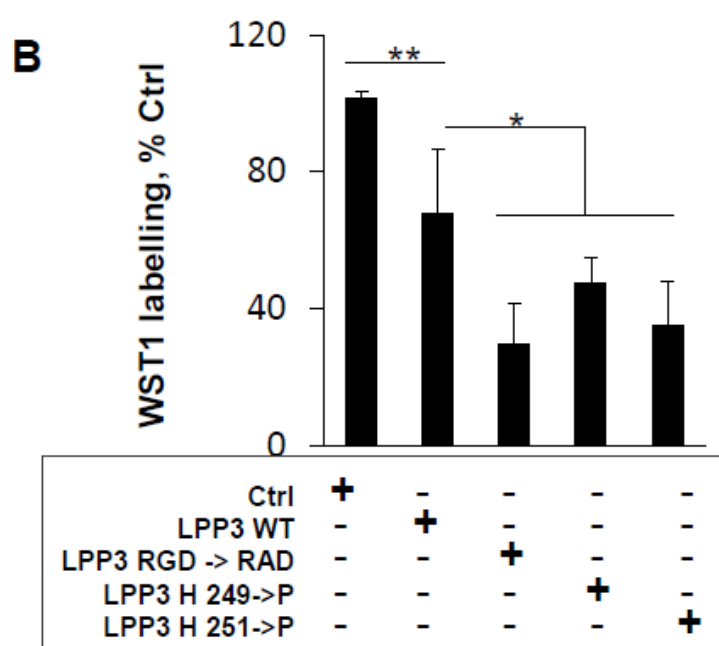
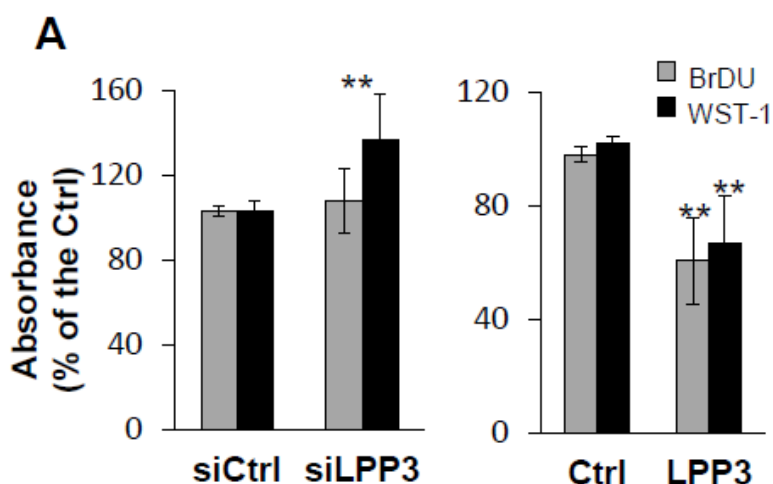


Figure 5. LPP3 affects viability of HAEC. HAEC were transfected with either siRNA (siCtrl or siLPP3) or with overexpressing plasmids (Ctrl or LPP3) for 48 h. Cell proliferation was estimated using the BrDU assay (gray bars) and cell viability was measured using the WST-1 assay (black bars). The results are shown as mean \pm SD from 5 independent donors in triplicate (A). HAECs were transfected with control, WT LPP3 or mutants of the catalytic domain (H249P or H251P) or of the adhesion motif (RGD->RAD), then WST-1 assays was performed. The results are shown as mean \pm SD from 3 independent donors in triplicate (B). HAECs transfected with either siLPP3 or siCtrl were treated with inhibitors of LPA (PF8380) or S1P (FTY720) (C). HAECs transfected with LPP3-containing or Ctrl plasmid were treated with agonists of LPA ((2S)-OMPT) or S1P (D). The results are shown as % of respective controls. The results are shown as mean \pm SD from 3 independent donors in triplicate. *: $p < 0.05$, **: $p < 0.001$, *** $p < 0.0001$, ns: non significant.

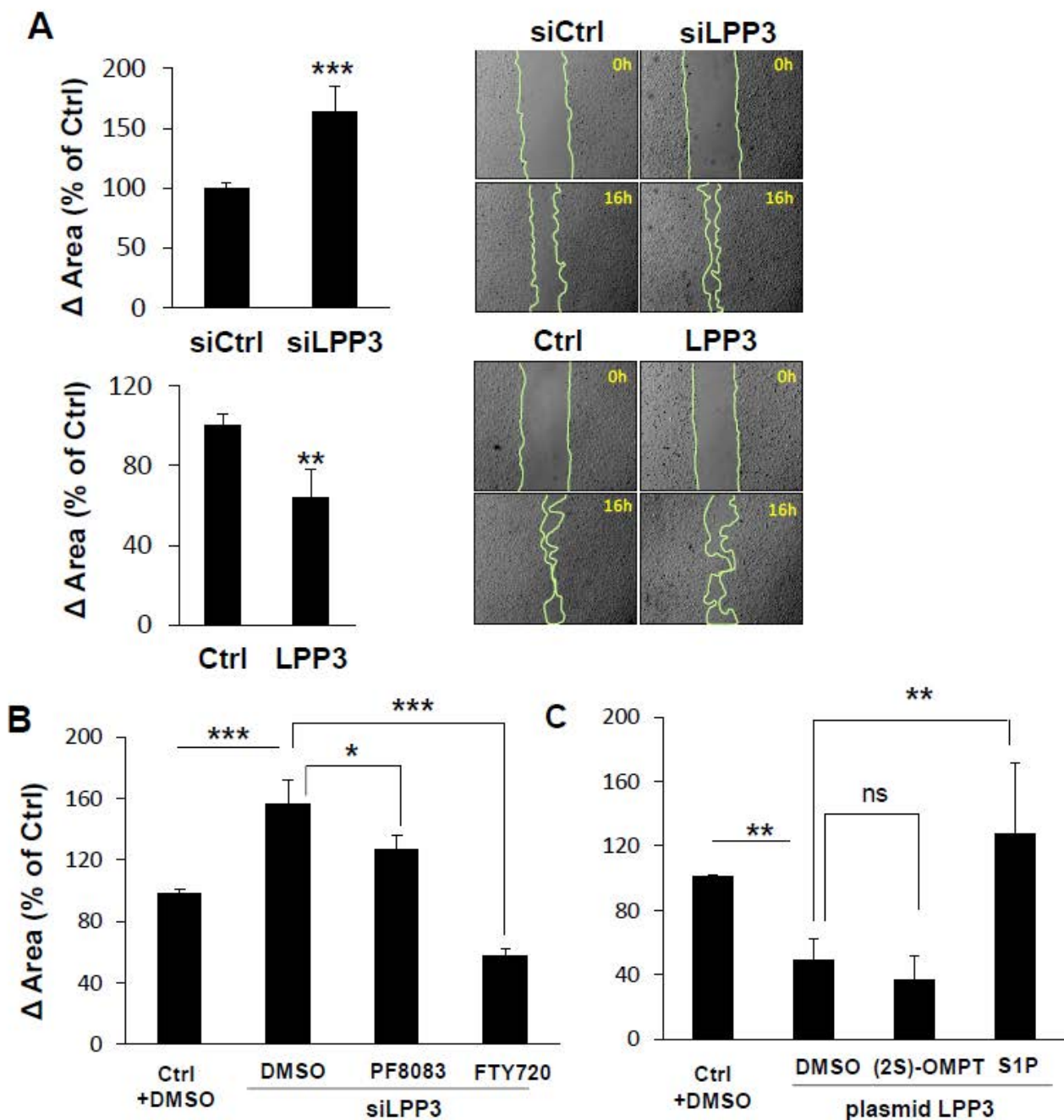


Figure 6. LPP3 is a negative regulator of HAEC cell migration. HAEC were transfected with either siRNA (siCtrl or siLPP3) or with overexpressing plasmids (Ctrl or LPP3) for 48 h. Cell monolayers were wounded with 1000 μ l pipette tips (right panel) and incubated for 16h to assess their migration. Wounded areas were imaged at 0h and 16h. Results are expressed in % of the respective controls and are shown as mean \pm SD from 5 independent donors in duplicate (A). siLPP3 and siCtrl transfected HAECs were treated with inhibitors of LPA formation (PF8380) or S1P (FTY720) (B). LPP3-containing plasmid or Ctrl plasmid transfected cells were treated with agonists of LPA ((2S)-OMPT) or S1P. The results are shown as mean SD from 3 independent donors in duplicate (C). *: $p < 0.05$, **: $p < 0.001$, *** $p < 0.0001$, ns: no significant.

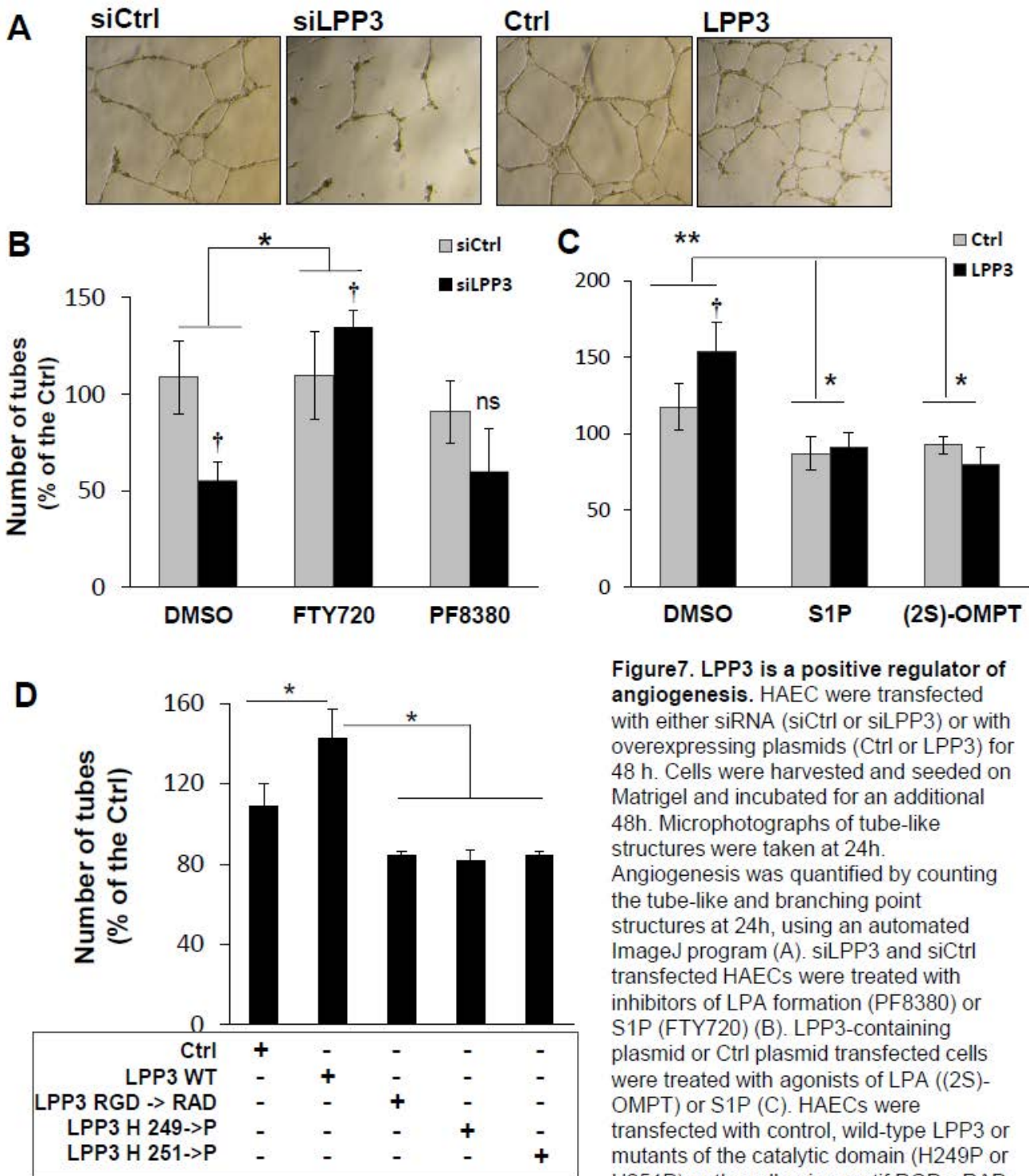
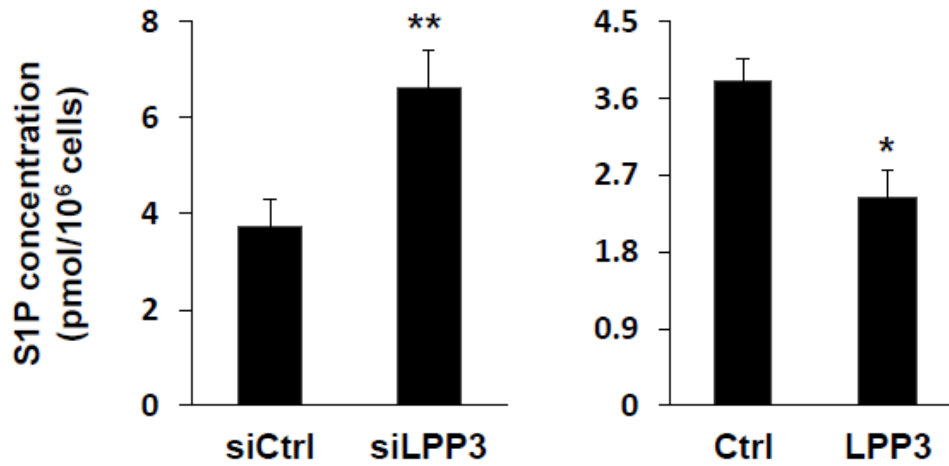


Figure 7. LPP3 is a positive regulator of angiogenesis. HAEC were transfected with either siRNA (siCtrl or siLPP3) or with overexpressing plasmids (Ctrl or LPP3) for 48 h. Cells were harvested and seeded on Matrigel and incubated for an additional 48h. Microphotographs of tube-like structures were taken at 24h. Angiogenesis was quantified by counting the tube-like and branching point structures at 24h, using an automated ImageJ program (A). siLPP3 and siCtrl transfected HAECs were treated with inhibitors of LPA formation (PF8380) or S1P (FTY720) (B). LPP3-containing plasmid or Ctrl plasmid transfected cells were treated with agonists of LPA ((2S)-OMPT) or S1P (C). HAECs were transfected with control, wild-type LPP3 or mutants of the catalytic domain (H249P or H251P) or the adhesion motif RGD->RAD (D). The results are shown as mean \pm SD and expressed as % of respective controls. The results are shown as mean \pm SD from 3 independent donors in duplicate. *: $p < 0.05$; †: siLPP3/siCtrl and LPP3/Ctrl; *: Treatment/DMSO

A Intracellular S1P



B Extracellular S1P

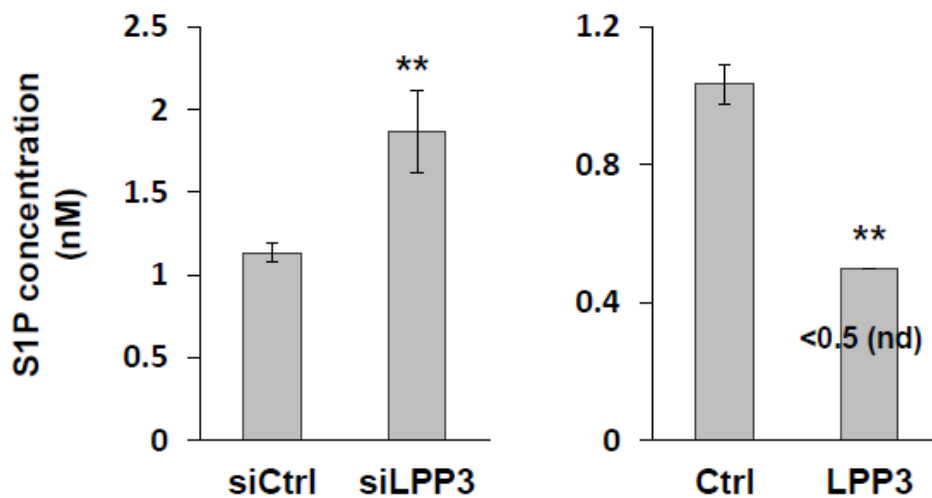


Figure 8: S1P levels in HAEC cell extracts and supernatants. HAEC were transfected with either siRNA (siCtrl or siLPP3) or with overexpressing plasmids (Ctrl or LPP3) for 48 h. Cell extracts (A) and supernatants (B) were collected. Lipids were extracted and derivative mixtures were analyzed by HPLC. The results are shown as mean \pm SD from 3 independent donors in duplicate. *: $p < 0.05$; **: $p < 0.01$.

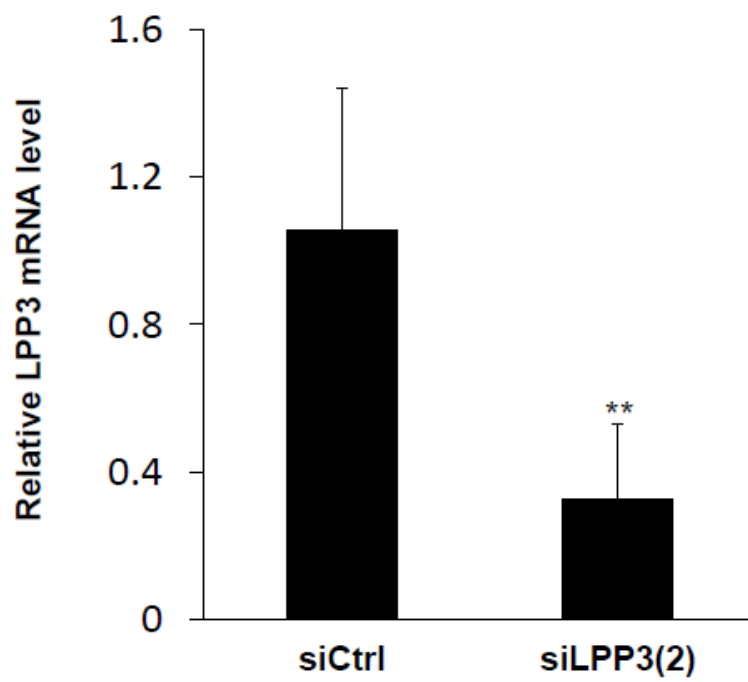


Figure S2: . Silencing of LPP3 with siLPP3 (2). HAECs were transfected for 48h with siLPP3 (siRNA(2): ID s16385) or siCtrl. The level of LPP3 mRNA was determined by RT-qPCR and the results are expressed as a relative LPP3 mRNA level over the control and graphed as mean ± SD from 3 independent donors in duplicate . **: $p < 0.001$.

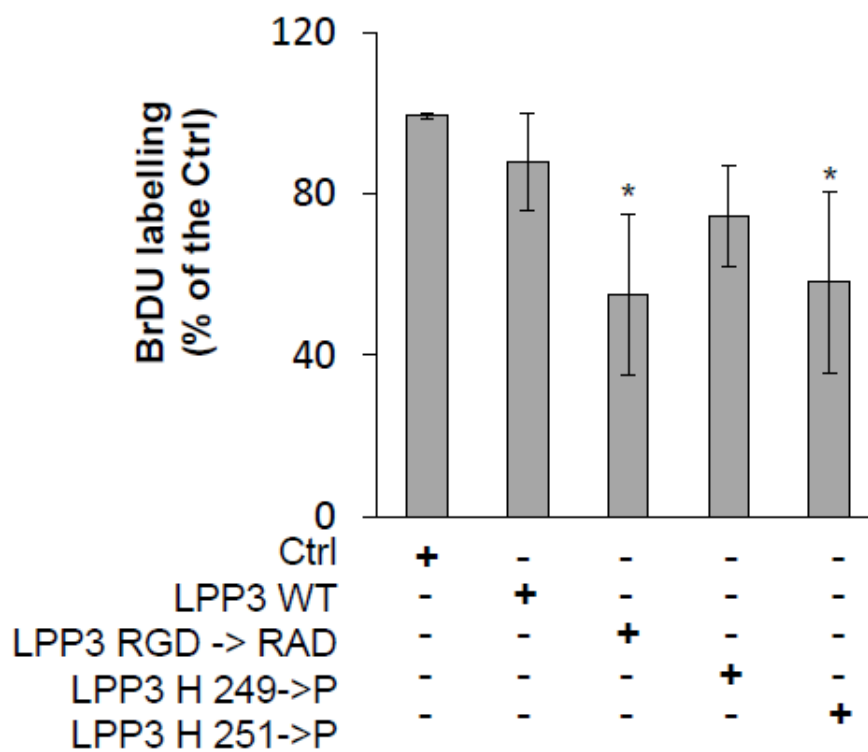


Figure S3. RGD ligand and catalytic domains of LPP3 are implicated in cell proliferation. HAECs were seeded in 96-well plates and transfected for 48h with either control, WT LPP3 or mutants of the catalytic domain (H249P or H251P) or the adhesion motif (RGD->RAD). Cell proliferation was estimated using the BrDU assay. The results are expressed as % of respective controls. The results are shown as mean ± SD from 3 independent donors in triplicate *: p<0.05.

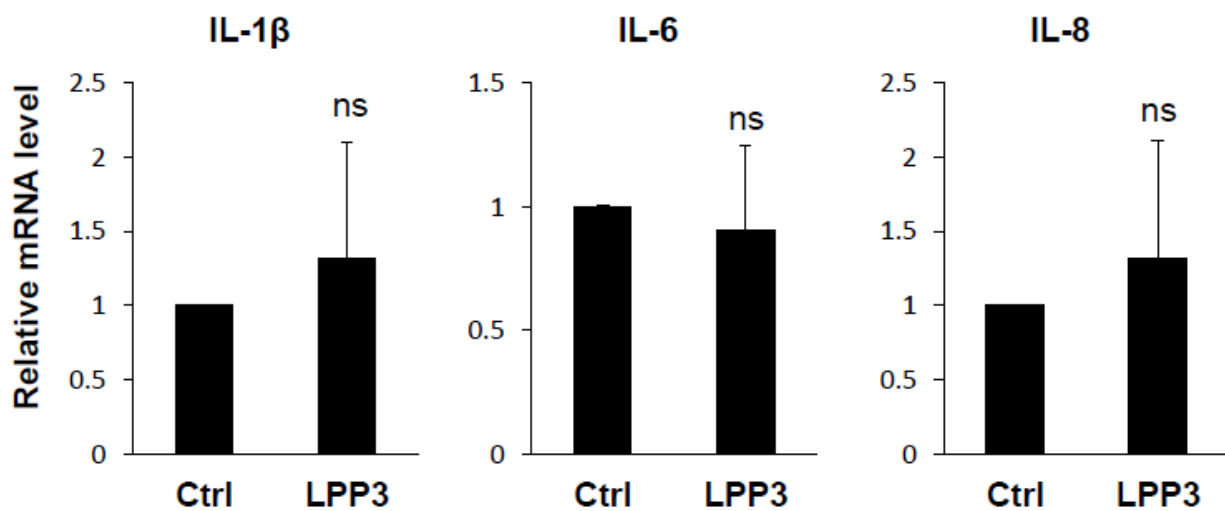


Figure S4. LPP3 does not affect the expression of pro-inflammatory cytokines. HAEC were transfected with overexpressing plasmids (Ctrl or LPP3) for 48 h. mRNA relative levels of IL1 β , IL6 and IL8 were determined by RT-qPCR. ns: non significant. The results are shown as mean \pm SD from 4 independent donors in duplicate

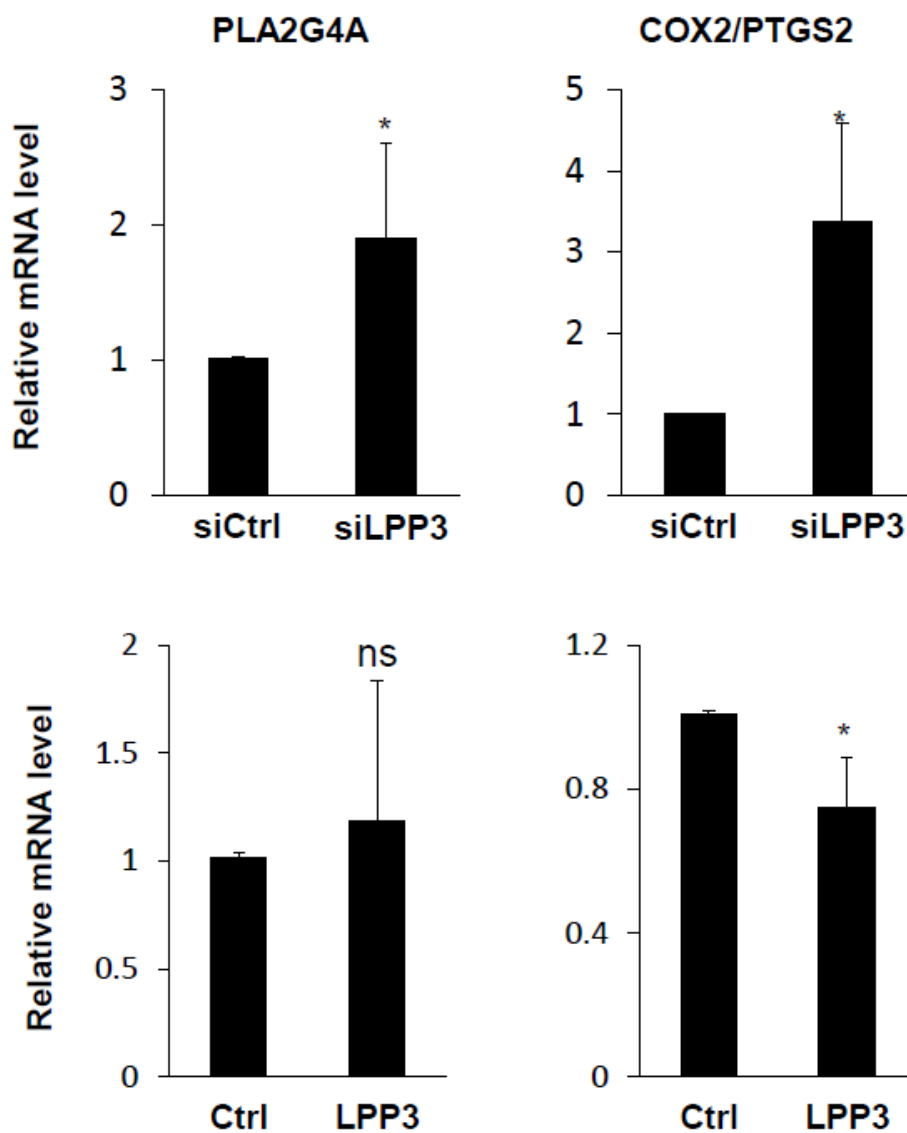
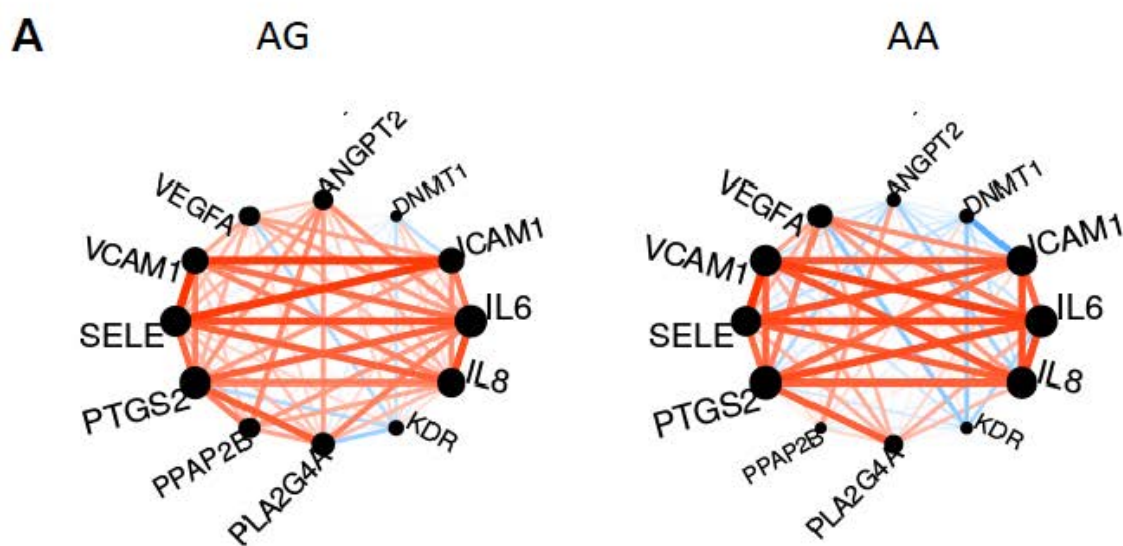


Figure S5. LPP3 negatively regulates PLA2G4A and PTGS2 levels. HAECs were transfected for 48h with either siLPP3 or siCtrl and either LPP3-containing plasmid or Ctrl plasmid. mRNA relative levels of PLA2G4A and PTGS2 (COX2) were determined by RT-qPCR. The results are shown as mean \pm SD from 4 independent donors in duplicate *: $p < 0.05$, **: $p < 0.001$, *** $p < 0.0001$.



B

| | AG (28) | AA (119) | p-value |
|---------|-------------------|-------------------|---------|
| ICAM1 | 0.075 (0.685) | -0.092 (0.330) | 0.4413 |
| IL6 | 0.414 (0.03) | 0.094 (0.330) | 0.1164 |
| IL8 | 0.374 (0.052) | -0.067 (0.449) | 0.0375 |
| KDR | -0.102 (0.612) | -0.097 (0.279) | 1 |
| PLA2G4A | 0.385 (0.040) | 0.175 (0.050) | 0.2983 |
| PTGS2 | 0.538 (0.003) | 0.071 (0.449) | 0.0155 |
| SELE | 0.305 (0.108) | 0.142 (0.128) | 0.4179 |
| VCAM1 | 0.207 (0.283) | 0.095 (0.279) | 0.6101 |
| VEGFA | 0.084 (0.685) | 0.125 (0.158) | 0.8181 |

Figure S6. Correlation between PPAP2B/LPP3 expression and other genes for the AA and AG genotypes. (A) Circle plots of the correlations. Edges are colored in red/blue for positive/negative correlations respectively and are weighted according to the correlation value. **(B)** Statistical comparison of the correlation coefficients between the two genotypes (red; positive and green; negative correlation). Values of the correlations are reported for each genotype, with the p-value into brackets. The p-value related to the comparison tests is given in the last column.

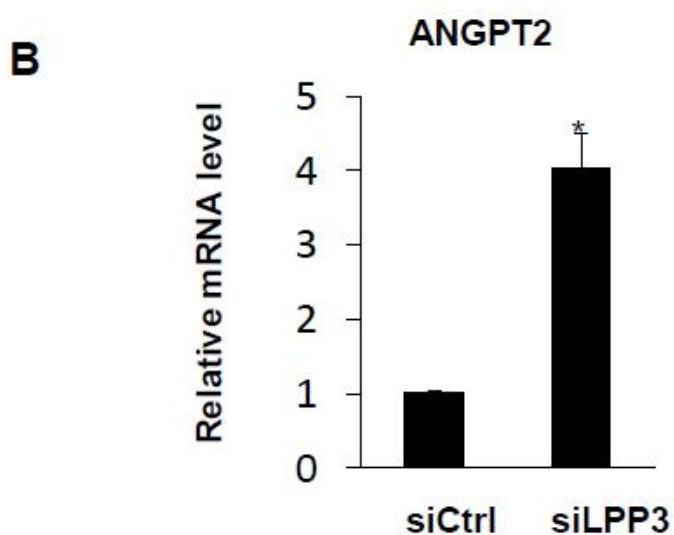
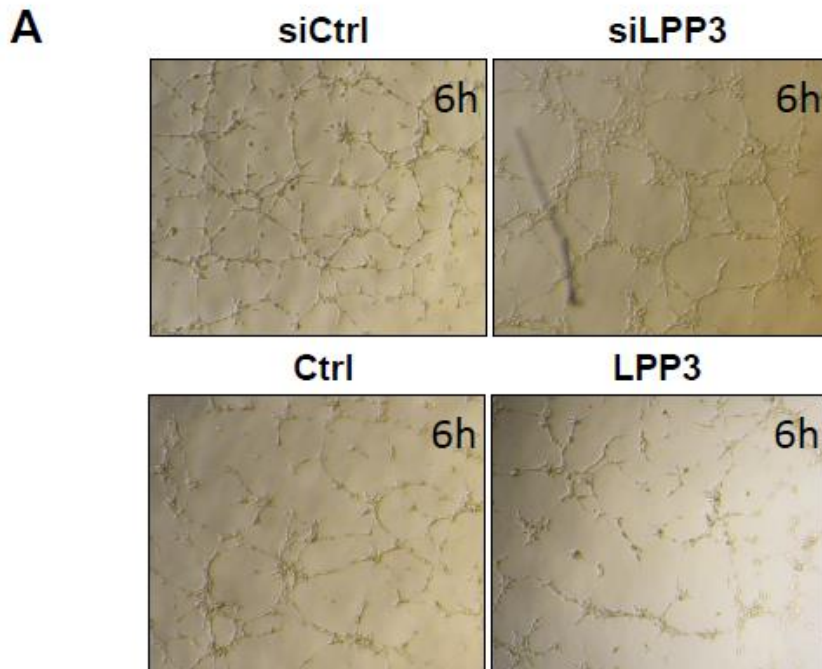


Figure S7. Effect of LPP3 on tube formation and ANGPT2 expression level. (A) HAECs were transfected for 48h with siRNAs or plasmids. Cells were harvested and seeded on Matrigel and incubated further for 48h. Microphotographs of tube-like structures were taken at 6h. (B) Cells were transfected for 48h and mRNA relative levels of ANGPT2 were determined by RT-qPCR. The results are shown as mean \pm SD from 5 independent donors in duplicate *: $p < 0.05$

Materials and Methods

Oligonucleotides and Chemicals

Recombinant human vascular endothelial growth factor-165 (VEGF-165) was purchased from R&D Systems (VEGF 293-VE/CF). Lipid substrates were purchased from Cayman Chemical: (2S)-OMPT (10005707), PF8380 (12018), FTY720 (10006292) and S1P (62570). Human IL1 β /IL-1F2, IL6 and CXCL8/IL8 DuoSet ELISA (enzyme-linked immunosorbent assay) kits were purchased from R&D Systems.

siRNA oligonucleotides were purchased from Ambion/Applied Biosystems: Silencer Select Negative Control siRNA (4390843), PPAP2B (4390824; siRNA(1): ID s16386, siRNA(2): ID s16385). Plasmids were purchased from Origen: *PPAP2B* cDNA (RC203480) and control pCMV-Entry vector (PS100001). *PPAP2B* mutant constructs were made in the laboratory using *PPAP2B* cDNA (Origen).

Mutant primers were: catalytic site (H249->P):

Forward. 5'TCTCGCGTATCAGACCAAAAGCACCATCCCAGTGAT

Reverse. 5'ATCACTGGGA- TGGTGCTTTTGGTCTGATACGCGAGA ;

catalytic site (H251->P):

Forward. 5'TCTCGCGTATCAGACCACAAGCAACATCCCAGTGAT

Reverse. 5'ATCACTGGGATGTTGCTTGTGGTCTGATACGCGAGA ;

adhesion motif (RGD-> RAD):

Forward.5'CTACAGATGCAGAGCTGATGACAGCAAAG

Reverse. 5'CTTTGCTGTCATCAGCTCTGCATCTGTAG.

Cell culture and transfections

Human aortic endothelial cells (HAECs) were either purchased from Promocell (Promocell, Germany) or isolated from aortas of donors during heart transplantation at the University of California, Los Angeles Medical Center(1,2). Cells were cultured in Endothelial Cell Growth Media (Lonza) supplemented with growth factors (EGM-2 bullet kit: 2% FBS, 0.4% hFGF-2, 0.1% VEGF, 0.1% R3-IGF1, 0.1% hEGF, 0.1% ascorbic acid, 0.1% heparin, 0.1% GA-100) in a humidified 5% CO₂ atmosphere at 37°C. All assays were performed using HAECs from at least 3 different donors (passages 4-6). Prior to assays, HAECs were split in 6-wells plates and grown to 70-80% of confluence and then transfected for 48h either with 20nM of siRNA using Lipofectamine RNAiMax (Life Technologies), or with 1 μ g of plasmids using Turbofect (Thermo Scientific). 24 hours post-transfection, the media were replaced with fresh complete media containing either 100ng/ml VEGF or various concentrations of LPA or S1P agonists ((2S)-OMPT, FTY720, PF8380 or S1P). HAECs were maintained in cultures for further 24 hours before assays.

All protocols involving humans (Human aortic endothelial cells; HAEC) were approved by UCLA Institutional Review Board and NIH, were conform to the Helsinki Declaration of 1975 as revised in 1983 and a written consent was given for the use of the human material.

ELISA assays for cytokines

HAECs were cultured in 6-well plates and transfected with either siRNA or plasmids containing the PPAP2B constructs. After transfection cells were cultured 48h in serum-free medium, and cell supernatants were collected and centrifuged at 300 g for 5 minutes to remove cellular debris. ELISA assays were performed in cell supernatants to detect soluble forms of IL1 β , IL6 and IL8 using respective human DuoSet ELISA kits (R&D Systems) following manufacturer's instructions. The experiment was performed in duplicate in supernatants from 3 independent donors in triplicate. Results are expressed as a mean +/- SD, and statistical difference was validated using paired t-test.

Human peripheral blood mononuclear cells adhesion assay

Prior to adhesion assay endothelial cells were seeded into 12-well culture plates, transfected with either siRNA or plasmids for 48 hours and treated by adequate stimuli during 24 hours. Human peripheral blood mononuclear cells (hPBMC) were isolated from a single-donor Buffy coats (Etablissement Français du Sang, Rungis, France) using Ficoll density gradient centrifugation for 20 min at 1200 rpm, next the cells were resuspended in RPMI medium at 3×10^6 cells/ml(3). Cells were labelled with 10 μ M calcein-AM (Life Technologies) for 30 min at 37°C in the dark, and then were sedimented and washed three times with PBS to remove excess dye and resuspended in phenol red-free RPMI (with 10% FBS) at a density of 3×10^6 cells/ml. HAECs monolayers were washed after various treatments with PBS and incubated with calcein-AM labeled hPBMCs (3×10^6 cells/well) at 37°C for 1h. HAECs monolayers were washed at least 3 times with PBS to remove the non-adherent calcein-AM labelled cells and the media was replaced with 1 ml PBS. Images of fluorescent cells were captured with epifluorescence microscope NIKON ECLIPSE Ti with Nikon Intensilight C-HGFIE using FITC filter, and the adherent cells were quantified by automated counting using ImageJ software (NIH). The results were expressed as a mean +/- SD of 3 independent donors in duplicate, and statistical difference was validated using paired t-test.

Cell viability assay

Cell viability assessment was performed according to a colorimetric assay using the tetrazolium salt WST-1 (4-[3-(4-iodophenyl)-2H-5-tetrazolio]-1-3-benzene disulfonate), based on the conversion of WST-1 into formazine by mitochondrial dehydrogenase enzyme in viable cells. 10,000 cells/well were seeded in 96-well plates in 100 μ l medium and

transfected for 48 hours. Cells were treated with the adequate treatments during the last 24 hours. Then 10 μ l of WST-1 solution were added to each well, followed by 3 hours incubation. Absorbance was measured at 450 nm in a Flexstation 3 microtiter plate reader (Molecular devices, CA. USA). Experiments were performed in triplicate from 5 (figure 5A) or 3 (figure 5B; C and D) donors. The results were expressed as a mean \pm SD and statistical difference was validated using paired t-test.

Cell proliferation assay

Cell proliferation test was carried out using the ELISA kit (Calbiochem) containing the thymidine analogue the 5-bromo-2'-deoxyuridine (BrdU), which is incorporated into the DNA of proliferating cells. Briefly, one day prior to the ELISA test, cells were seeded in 96-well plates at 10.000 cells/well in 100 μ l medium, treated with various agents and then cultured at 37°C for 24h to allow adherence. BrdU (10 μ M) was added for 6 hours to the plates, next the cells were fixed and incubated with anti-BrdU antibody conjugated with peroxidase which reacts with its substrate to yield a product quantified at 450 nm using a Flexstation 3 microtiter plate reader (Molecular devices, CA. USA). Experiments were performed in triplicate using five donors. The results were expressed as a mean \pm SD of 5 (figure 5A) or 3 (figure S3) independent donors in triplicate, and statistical difference was validated using paired t-test

Cell migration

Transfected HAECs were cultured for 48h to reach 90% confluence and then treated with mitomycin (Sigma-Aldrich) for 2h. Two linear scratches per well were performed in the cell monolayer using a 1000 μ l pipette tip. Cells were treated with lipid substrates and were incubated at 37°C and 5% CO₂ for the next 16h. Images of the gap were obtained at 0 and 16h with fully motorized inverted microscope (Nikon Eclipse TiS) using bright field at 4X magnification. The wounded area was analysed using ImageJ software (NIH) by quantification of the surface of wounded area at 24h as compared to 0h. Experiments were performed in duplicate from 5 (figure 6A) or 3 (figure 6B) donors. The results were expressed as a mean \pm SD of five independent donors in duplicate, and statistical difference was validated using paired t-test

Angiogenesis tube formation in Matrigel

In vitro angiogenesis assays were performed using 15-well μ -angiogenesis slides (Ibidi, Germany). The slides were coated with 10 μ l of Matrigel (VWR) per well and then left to polymerize for 30 min at 37 °C. Transfected cells were seeded at 10.000 cells/well in triplicates. Cells were treated with various agents and were incubated at 37°C for additional

24h. The images of newly structured tubes were captured at 6 and 24 h using 4x magnification (Nikon). Tube formation was quantified manually by counting the number of tubular and branching point structures. The results were expressed as a mean \pm SD of 3 independent donors in duplicate, and statistical difference was validated using paired t-test

RNA extraction and real-time PCR

Total RNA was extracted 48h post-transfection using the mirVANA-miRNA isolation kit (Life Technologies) and RNA quality was assessed by the 2100 Bioanalyser (Agilent technologies, Canada). RNA integrity (RIN) was between 8 to 10. Synthesis of cDNA was carried out using the Super Script II Reverse Transcriptase (Life technologies). Real-time PCR was performed using Mx3005P QPCR System SYBR Green (Thermo Scientific). The amplification program was: 95 °C for 15 min, 40 cycles of 95 °C for 30 s and one cycle of 95°C for 1 min, 60°C for 30 s and 95°C for 30 s. Data were analyzed using MxPro® software using the $\Delta\Delta$ CT method and normalized to the ubiquitin and GAPDH control genes. Experiments were performed in duplicate using five donors. The results were expressed as a mean \pm SD of 5 independent donors in duplicate, and statistical difference was validated using paired t-test

Protein extraction and western blot analysis

Cells were lysed by the Mammalian Whole Cell Protein Extraction Kit supplemented with the protease inhibitor cocktail (Promocell, Germany) and the protein concentration was measured using Quick Start Bradford reagent (BioRad). Protein extracts (50 μ g) were separated in 4-12% NuPAGE Novex gels (Life Technologies) and transferred onto PVDF membranes (BioRad). The membrane was probed with the rabbit anti-LPP3, kindly provided by Dr Susan S. Smyth, Division of Cardiovascular Medicine, The Gill Heart Institute, Lexington, KY(4) and anti-tubulin mouse monoclonal antibody (Invitrogene). HRP-conjugated anti-rabbit or anti-mouse (DAKO) corresponding secondary antibodies were used for detection. The immunoblot correspond to one experiment. The chemiluminescence signal was detected using Pierce ECL Plus Western blotting substrate (Thermo Scientific) using ImageQuant LAS 4000 (GE Healthcare). Scanned bands were normalized with tubulin.

Plasmid constructs

Site-directed mutagenesis (Agilent QuickChange II XL Site-Directed Mutagenesis kit; Agilent Technologies) was performed using *PPAP2B* wild type cDNA (4390824) (Origen). Mutated primers from Sigma-Genosys (Sigma-Aldrich) were used to generate expression plasmids carrying different mutations. After PCR and Dpn I digestion, XL10-Gold Ultracompetent cells provided with the kit were transformed with the different mutant plasmids. All constructs were sequenced.

Genome-Wide Expression analysis and pre-processing of expression data

Total RNA was extracted from 5 donors of HAECs using mirVANA-miRNA kit (Life Technologies) as described above. Transcriptomic analysis of total RNA was performed using the *Illumina* HT-12 v4 BeadChip (<http://www.Illumina.com>). Briefly, RNA was extracted from HAECs and 250 ng of total RNA was reverse transcribed, amplified and biotinylated using the *Illumina* TotalPrep RNA Amplification Kit (Ambion/Applied Biosystems). Each biotinylated cRNA (750 ng) was hybridized to a single BeadChip at 58°C for 16–18 hours. BeadChips were scanned using the *Illumina* HiScan array.

The summary probe-level data delivered by the *Illumina* scanner (mean and SD computed over all beads for a particular probe) was loaded in *Genome Studio*. The pre-processing was done by the *Illumina* software, at the level of the scanner and by *Genome Studio* included: correction for local background effects, removal of outlier beads, computation of average bead signal and SD for each probe and gene, calculation of detection *P*-values using negative controls present on the array, quantile normalization across arrays, check of outlier samples using a clustering algorithm, check of positive controls. Analyses were carried out on the mean level for all probes in each gene. To stabilise variance across expression levels, *arcsinh* transformation was applied to the expression data. Compared to a log transformation, this transformation has the advantage not to discard negative expression values which can occur in *Illumina* data.

The *Illumina* HT-12 v4 BeadChip targets more than 47,000 probes derived from the National Center for Biotechnology Information Reference Sequence (NCBI) RefSeq Release 38 (November 7, 2009) and other sources. A gene was declared significantly expressed in the dataset, i.e. expressed above background (as measured by the negative controls present on each array), when the detection *P*-value calculated by *Genome Studio* software was <0.05 in

Determination of S1P levels

Methanol (1ml) containing 2.5 μ l concentrated HCl was added to 500 μ L of HAECs suspension or supernatants. C17-S1P (27 pmol) was added as internal standard and lipids were extracted by addition of 1 ml chloroform, 200 μ l NaCl (4 M) and 100 μ l NaOH (3 M). The alkaline aqueous S1P-containing phase devoid of other sphingoid bases and of the majority of hydrophobic PL was transferred to a clean tube. The alkaline aqueous phase was acidified with 100 μ l concentrated HCl and extracted twice with 1.5 ml chloroform. The organic phases were evaporated and the dried lipids were dissolved in 100 μ l of methanol. A derivatization mixture of 5 mg *o*-phthalaldehyde, 100 μ l ethanol, 5 μ l 2-mercaptoethanol and 5 ml boric acid (3% v/w) was prepared and adjusted to pH 10.5 with 10 N KOH. Nine μ l of the derivatization mixture was added to the lipids and the solution incubated for 15 minutes at 37°C. The derivatives were analyzed with a Shimadzu HPLC system using an RP 18 Kromasil column (2.1 mm i.d. x 150 mm) maintained at 45°C. Separation was performed with the isocratic eluent containing methanol: K₂HPO₄ (0.07 M) (78:22 v/v) at a flow rate of 0.25 mL/min. The derivatives were detected selectively using a Shimadzu spectrofluorometer with an excitation wavelength of 334 nm and an emission wavelength of 440 nm. S1P was quantified by comparison of its fluorescent signal with that of the derivative of the internal standard (coefficients of variation <5%). The results were expressed as a mean \pm SD of three independent donors in triplicate, and statistical difference was validated using paired t-test

Statistical analysis

All data are presented as mean \pm standard deviation (SD) from at least 3 independent donors. Comparison between siRNA or overexpressing plasmids with their respective control was performed using Student test, depending on the experiment, with a significance threshold at $p \leq 0.05$.

Statistical analysis of microarray data was performed in the R environment (version 3.0.1) using the *lumi* package of Bioconductor(5). The variance stabilizing transformation(6) was applied to raw data. A variance-based filtering of gene expressions was further performed using the *nsFilter* function from the *genefilter* package with default parameters (interquartile range and variance cutoff of 0.5). Differential expression analysis was performed using the *limma* package of Bioconductor. To account for the pairing of each siRNA sample with its control, the difference between gene expressions of both samples was considered as the

variable to be tested to 0. Analysis was further adjusted for beadchip. The Benjamini-Hochberg procedure was applied to adjust for multiple testing.

Table S2. GO biological processes enrichment. HAEC (n=5) were transfected with LPP3 siRNA(1) or control siRNA. Total RNA was prepared and global gene expression analysis was performed. cDNAs were labelled and hybridized to the Illumina HumanHT-12 v4 Expression BeadChip which targets 47,000 probes. Statistical enrichment of specific biological processes and molecular functions annotated in Gene Ontology (GO <http://www.geneontology.org/>) was assessed by GSEA using Genetrail software. Significant GO biological processes were selected (FDR<0.05).

| Mechanism | Biological processes | category | Genes | FDR | Enrichment |
|----------------|---|------------|-------|-------------|------------|
| Inflammation | Inflammatory response | GO:0006954 | 161 | 0.000888283 | Up |
| | Cytokine activity | GO:0005125 | 76 | 0.00291964 | Up |
| | Cytokine production | GO:0001816 | 124 | 0.00511788 | Up |
| | Chemokine activity | GO:0008009 | 19 | 0.0078495 | Up |
| | Chemotaxis | GO:0006935 | 87 | 0.00845029 | Up |
| | Cell adhesion | GO:0007155 | 382 | 9.36178e-06 | Up |
| | Cell adhesion molecule binding | GO:0050839 | 25 | 0.00324616 | Up |
| | Leukocyte activation | GO:0045321 | 176 | 0.00951024 | Up |
| | Regulation of leukocyte migration | GO:0002685 | 22 | 0.015871 | Up |
| Cell survival | Cell cycle | GO:0007049 | 706 | 0 | Down |
| | Regulation of cell cycle | GO:0051726 | 324 | 0.00117872 | Down |
| | Positive regulation of cell proliferation | GO:0008284 | 231 | 0.00614088 | Up |
| | Caspase inhibitor activity | GO:0043027 | 10 | 0.0131256 | Up |
| | I-kappaB kinase/NF-kappaB cascade | GO:0007249 | 127 | 0.0210446 | Up |
| Cell migration | Response to wounding | GO:0009611 | 289 | 1.51154e-05 | Up |
| | Wound healing | GO:0042060 | 112 | 0.00891874 | Up |
| | Cell migration | GO:0016477 | 260 | 0.000168581 | Up |
| | Regulation of cell migration | GO:0030334 | 127 | 0.00149412 | Up |
| Angiogenesis | Angiogenesis | GO:0001525 | 151 | 0.000576977 | Up |
| | Blood vessel development | GO:0001568 | 220 | 0.00153458 | Up |
| | Blood vessel morphogenesis | GO:0048514 | 184 | 0.00286085 | Up |

Table S3: Primer sequences

| Human genes | Primer sequences |
|------------------------------|--|
| GAPDH | For. 5' -TGCACCACCAACTGCTTAGC Rev. 5' -CAGTCTTCTGGGTGGCAGTGA |
| Ubiquitin | For. 5' -TTTTTGGGAATGCAACAACTT Rev. 5' -CACTTGGTCCTGCGCTTGA |
| PPAP2B (LPP3) | For. 5' -TTTACTATGTGGCCTTCCAAAAT Rev. 5' -ACTAGAACATGAACACTTAAACC |
| VEGF A | For. 5' -TCTACCTCCACCATGCCAAGT Rev. 5' -GATGATTCTGCCCTCCTCCTT |
| IL1β | For. 5' -GTGTAGATCCCAAAAATTACCCA Rev. 5' -CCAGGAAGACGGGCATGTT |
| IL6 | For. 5' -AACCTGAACCTTCCAAAGATG Rev. 5' -TCTGGCTTGTTCTCACTACT |
| IL8 | For. 5' -TCTGGACCCCAAGGAAAACCT Rev. 5' -TCTTGATACCACAGAGAATGAATTT |
| SELE | For. 5' -GATGAGAGGTGCAGCAAGAAG Rev. 5' -CTCACACTTGAGTCCACTGAAG |
| ICAM1 | For. 5' -CAGTGACCATCTACAGCTTTCCGG Rev. 5' -GCTGCTACCACAGTGATGATGATGACAA |
| VCAM1 | For. 5' -TGAGTTTTGAGAACGAACACTCTTATCT Rev. 5' -GCTTACAGTGACAGAGCTCCCATT |
| MCP1/CCL2 | For. 5' -TCTCGCCTCCAGCATGAAAGT Rev. 5' -GCATTGATTGCATCTGGCTGA |
| PLA2G4A | For. 5' -CGACGTCCTTCTCTGGTATTGAAT Rev. 5' -GATGCTGGAAGTGACTATCAAAGTGATA |

| | |
|------------|--|
| COX2/PTGS2 | For. 5'-ATATGTTCTCCTGCCTACTGGAA Rev. 5'-GCCCTTCACGTTATTGCAGCTG |
| ANGPT2 | For. 5'-CAGTTCTGAGGATGTTAATCGTC Rev. 5'-AGACATTCACAGAACTTAACACC |

Legends to supplementary figures

Figure S1. LPP3 expression in vascular cells. IHC was performed on 5- μ m sections of TMA blocks. Deparaffinized and rehydrated sections were incubated with primary antibodies against CD68 (macrophages), alpha-actin (smooth muscle cells), CD34 (endothelial cells), CD3 (lymphocytes) and the polyclonal rabbit anti-LPP3. IHC were revealed using the HRP Polymer DAB Plus Chromogen substrate. Slides were counterstained with aqueous hematoxylin and mounted with Immunomount.

Figure S2. Silencing of LPP3 with siLPP3 (2). HAECs were transfected for 48h with siLPP3 (siRNA(2): ID s16385) or siCtrl. The level of LPP3 mRNA was determined by RT-qPCR and the results are expressed as a relative LPP3 mRNA level over the control and graphed as mean \pm SD from 3 independent donors in duplicate. **: $p < 0.001$.

Figure S3. RGD ligand and catalytic domains of LPP3 are implicated in cell proliferation. HAECs were seeded in 96-well plates and transfected for 48h with either control, WT LPP3 or mutants of the catalytic domain (H249P or H251P) or the adhesion motif (RGD->RAD). Cell proliferation was estimated using the BrDU assay. The results are expressed as % of respective controls. The results are shown as mean \pm SD from 3 independent donors in triplicate *: $p < 0.05$.

Figure S4. LPP3 does not affect the expression of pro-inflammatory cytokines. HAEC were transfected with overexpressing plasmids (Ctrl or LPP3) for 48 h. mRNA relative levels of IL1 β , IL6 and IL8 were determined by RT-qPCR. ns: no significant. The results are shown as mean \pm SD from 4 independent donors in duplicate

Figure S5. LPP3 negatively regulates PLA2G4A and PTGS2 levels. HAECs were transfected for 48h with either siLPP3 or siCtrl and either LPP3-containing plasmid or Ctrl plasmid. mRNA relative levels of PLA2G4A and PTGS2 (COX2) were determined by RT-qPCR. The results are shown as mean \pm SD from 4 independent donors in duplicate *: $p < 0.05$, **: $p < 0.001$, *** $p < 0.0001$.

Figure S6. Correlation between PPAP2B/LPP3 expression and other genes for the AA and AG genotypes. (A) Circle plots of the correlations. Edges are colored in red/blue for positive/ negative correlations respectively and are weighted according to the correlation value. (B) Statistical comparison of the correlation coefficients between the two genotypes (red; positive and green;

negative correlation). Values of the correlations are reported for each genotype, with the p-value into brackets. The p-value related to the comparison tests is given in the last column.

Figure S7. Effect of LPP3 on tube formation and ANGPT2 expression level. (A) HAECs were transfected for 48h with siRNAs or plasmids. Cells were harvested and seeded on Matrigel and incubated further for 48h. Microphotographs of tube-like structures were taken at 6h. (B) Cells were transfected for 48h and mRNA relative levels of ANGPT2 were determined by RT-qPCR. The results are shown as mean \pm SD from 5 independent donors in duplicate *: $p < 0.05$

References

1. Romanoski CE, Lee S, Kim MJ, Ingram-Drake L, Plaisier CL, Yordanova R, Tilford C, Guan B, He A, Gargalovic PS, Kirchgessner TG, Berliner JA, Lusk AJ. Systems genetics analysis of gene-by-environment interactions in human cells. *Am J Hum Genet.* 2010;**86**:399–410.
2. Romanoski CE, Che N, Yin F, Mai N, Poudar D, Civelek M, Pan C, Lee S, Vakili L, Yang W-P, Kayne P, Mungrue IN, Araujo JA, Berliner JA, Lusk AJ. Network for activation of human endothelial cells by oxidized phospholipids: a critical role of heme oxygenase 1. *Circ Res.* 2011;**109**:e27–e41.
3. Pégurier S, Stengel D, Durand H, Croset M, Ninio E. Oxidized phospholipid: POVPC binds to platelet-activating-factor receptor on human macrophages. Implications in atherosclerosis. *Atherosclerosis.* 2006;**188**:433–443.
4. Sciorra VA, Morris AJ. Sequential actions of phospholipase D and phosphatidic acid phosphohydrolase 2b generate diglyceride in mammalian cells. *Mol Biol Cell.* 1999;**10**:3863–3876.
5. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. *Bioinforma Oxf Engl.* 2008;**24**:1547–1548.
6. Lin SM, Du P, Huber W, Kibbe WA. Model-based variance-stabilizing transformation for Illumina microarray data. *Nucleic Acids Res.* 2008;**36**:e11.