Journal of Molecular Medicine: Supplemental Material

Gene expression differences between stroke-associated and asymptomatic carotid plaques

Saksi J*[1], Ijäs P*[1,2], Nuotio K[1,2], Sonninen R[1], Soinne L[2], Salonen O[3], Saimanen E[4], Tuimala J[5], Lehtonen-Smeds E[6], Kaste M[2], Kovanen PT[6], Lindsberg PJ[1,2]

*Equal contribution.

[1]Research Programs Unit, Molecular Neurology, Biomedicum Helsinki, University of Helsinki, Finland; [2]Department of Neurology and [3]Department of Radiology, Helsinki University Central Hospital, Helsinki, Finland; [4]South Karelia Central Hospital, Lappeenranta, Finland; [5]CSC-Scientific Computing Ltd., Espoo, Finland; [6]Wihuri Research Institute, Helsinki, Finland

Supplemental Methods

RNA extraction

RNA was obtained from a longitudinal slice of the CP containing part of the atheroma core. Total RNA was extracted using Trizol®- reagent (Invitrogen Life Technologies) and purified with an RNeasy Total RNA Isolation Kit (Qiagen) according to the manufacturers' recommendations. One CP from an asymptomatic patient originally included in the microarray group did not yield enough RNA for microarray analysis, but this sample was included in the qPCR analysis as part of the extended replication group.

RNA quality was assessed both using conventional spectrophotometry (concentration and A260/A280 ratio $\geq$2) and microcapillary electrophoresis (RNA 6000 Nano LabChip Kit, Agilent Bioanalyzer 2100). Some of the samples were moderately degraded but all still showed clear ribosomal 18S and 28S peaks, only moderately elevated baseline and the 28S:18S rRNA ratios over 1. Examples of representative electropherograms are shown in Supplemental figure 1.

Microarray Analysis

Original data files are available at ArrayExpress ID: E-MEXP-2257 (http://www.ebi.ac.uk/microarray-as/ae/).

The scanned microarray images and Affymetrix quality measures were checked using MASv5.1 software (Affymetrix). Signal intensities were calculated and hybridization data normalized across arrays using the Robust Multi-array Average method (RMA) [1]. Genes expressed at a reliable level and showing differential expression were identified by filtering: 1) probes with a raw signal <100 in >50% of the samples in both groups and 2) probes <1.5-fold difference in the mean signal between SCPs and ACPs. These criteria are based on hybridisation data on a technical replicate (the RNA sample from one plaque was hybridised to two arrays): when these signal and fold-change criteria were applied, 90% of the false positives could be excluded. In addition, the X- and Y-chromosomal genes were removed due to a significant gender difference between the stroke and asymptomatic patients.

Pairwise p-values for probe sets were calculated by nonparametric Mann-Whitney U-test. False discovery rate was estimated by using Significance Analysis of Microarrays (SAM) [2], a permutation-based variance-shrinkage method. We used a value of 0.539 for the tuning parameter delta, which corresponds to a median false discovery rate of 0.05. The following values were used for other adjustable parameters: response format - two class unpaired, minimun fold-change - 1.5, test statistic - Wilcoxon rank-sum test, number of permutations - 1000.

In addition to RMA, the hybridization data was also pre-processed by two other commonly used algorithms, GC-RMA [3] and MAS5 (http://www.affymetrix.com). Pre-processing was followed by otherwise similar filtering steps as for RMA-processed data, but the raw signal cut-offs were adjusted to 100 for GC-RMA and 200 for MAS5 based on hybridisation data on a technical replicate (see methods above). Genes with significantly differing expression levels between ACPs and SCPs at 5% FDR were again identified by SAM using the same values for the adjustable parameters as in the case of RMA. The three probe lists where combined in Venn diagram to get the intersection of the lists.

Hierarchical clustering was performed with GeneSpring v.7.3 (Silicon Genetics) using Pearson uncentered as the distance metric and centroid as the linkage rule. Gene enrichment analysis for Gene Ontology (GO) categories and functional pathways (KEGG and Biocharta) was performed for the probes showing ≥1.5-fold expression change (n=103) using WebGestalt software [4]. We used a prestored gene set (WebGestalt_HG_U133A) as a reference set and applied a hypergeometric test to identify enriched GO categories and pathways. The default significance level of 0.01 was used and a minimum number of genes was set as two. For more information on Gene-Ontology see http://www.geneontology.org/.

Quantitative Real-Time PCR

QPCR was performed using TaqMan® Gene Expression Assays (Supplemental Table 1) and the ABI PRISM® 7000 Sequence Detection System (Applied Biosystems) according to the manufacturer's recommendations. Briefly, mRNA was converted to cDNA using High-Capacity cDNA Archive Kit (Applied Biosystems). Real-time qPCR reactions were performed in a 25 ul mixture of 20 ng of cDNA, 1x TaqMan Universal Master Mix and 1x Assay Mix containing primers and TaqMan probe (all from Applied Biosystems). Each sample was run in triplicate and only Ct values with standard deviation ≤ 0.16 were accepted. Each plate contained negative controls (no-template-controls) and the calibrator sample.

Prior to the analyses, suitable endogenous control genes for qPCR were determined by screening the expression of 11 commonly used endogenous control genes using a TaqMan® Human Endogenous Control Plates (Applied Biosystems). Beta-actin showed the most constant expression levels within $\Delta C_T \leq +/-0.5$ (Supplemental Figure 2).

Gene expression was determined by the comparative C_T- or the relative standard curve method. For each assay a delta-delta-Ct validation experiment was carried out. If the absolute value of the slope of log input amount vs. deltaCt was less than 0.1, data analysis was done by delta-delta-Ct method (Applied Biosystems User Bulletin no. 2 P/N 4303859B). Otherwise, the relative standard curve method was used (see Supplemental Table 1).

Supplemental Table 1. QPCR assays and the data analysis method

	Gene Name	Gene Symbol	Assay ID[a]	Method[b]
1	CD36 molecule	CD36	Hs01567191_m1	Comparative
2	CD163 molecule	CD163	Hs00174705_m1	Comparative
3	Fatty acid binding protein 4	FABP4	Hs00609791_m1	Standard curve
4	Perilipin 2	PLIN2	Hs00605340_m1	Comparative
5	Glutamate-ammonia ligase	GLUL	Hs00374213_m1	Standard curve
6	Chemokine (C-C motif) ligand 18	CCL18	Hs00268113_m1	Comparative
7	Fucosidase, alpha-L-1, tissue	FUCA1	Hs00609173_m1	Comparative
8	Interleukin 1 receptor antagonist	IL1RN	Hs00893626_m1	Standard curve
9	Heme oxygenase (decycling) 1	HMOX1	Hs00157965_m1	Comparative
10	S100 calcium binding protein A8	S100A8	Hs00374263_m1	Comparative
11	Chemokine (C-X-C motif) receptor 4	CXCR4	Hs00976734_m1	Comparative
12	Mannose receptor, C type 1	MRC1	Hs00267207_m1	Comparative
13	Heat shock 70kDa protein 6	HSPA6	Hs00275682_s1	Comparative
14	Oxysterol binding protein-like 8	OSBPL8	Hs00970777_m1	Comparative
15	Apolipoprotein E	APOE	Hs00171168_m1	Comparative
16	ADP-ribosylation factor-like 4C	ARL4C	Hs00255039_s1	Comparative
17	Regulator of G-protein signalling 5	RGS5	Hs01555176_m1	Standard curve
18	Myosin, heavy chain 10, non-muscle	MYH10	Hs00992055_m1	Comparative

[a]TaqMan® Gene Expression Assays, Applied Biosystems (www.appliedbiosystems.com)

[b]All genes were normalized to beta-actin.

Supplemental Table 2. Antibodies.

Antigen	Clonality/clone[a]	Working concentration[b]	Manufacturer
Cell type markers			
HAM56 (macrophage)	PC	0.71 µg/mL	Dako
CD45RO (activated T-cell)	MC/UCHL-1	4.20 µg/mL	Dako
Tryptase (mast cell)	MC/G3	0.12 µg/mL	Chemicon
α-smooth muscle cell actin	MC/1A4	7.6 µg/ml	Dako
Replicated genes[c]			
CD36	PC	2 µg/ml	Chemicon
CD163	PC	10 µg/ml	R&D Systems
FABP4	PC	0.3 µg/ml	Sigma
PLIN2	MC/AP125	1:100	American Research Products
GLUL	PC	5 µg/ml	BD Biosciences Europe
CCL18	PC	25 µg/ml	R&D Systems
IL1RN	PC	2.5 µg/ml	R&D Systems
HMOX1	PC	1:1500	Stressgen
S100A8	PC	2.5 µg/ml	Lifespan Biosciences

[a]PC = polyclonal, MC = monoclonal. [b]Dilution is reported if antibody concentration is unknown. Mouse IgG1 and -3 (Dako), mouse IgG2a-b (Dako) mouse IgM (Dako), goat IgG (Zymed Laboratories Inc.) in equivalent dilutions or PBS were used as negative controls. [c]Antibodies tested against FUCA1 failed to produce a reliable staining pattern.

Supplemental Table 3. Microarray probe sets showing ≥ 1.5-fold change in expression between stroke-associated and asymptomatic CPs (n=103)

Probe ID	Gene Symbol[a]	Gene Name[a]	Fold-change	Mann-Whitney U-test *p**	SAM q (%)*
202206_at	ARL4C	ADP-ribosylation factor-like 4C	1.6	0.001	1.72
213418_at	HSPA6	Heat shock 70kDa protein 6	1.8	0.001	1.72
221760_at	MAN1A1	Mannosidase, alpha, class 1A, member 1	1.5	0.002	1.72
210512_s_at	VEGF	Vascular endothelial growth factor	2.0	0.003	1.72
200921_s_at	BTG1	B-cell translocation gene 1, anti-proliferative	1.6	0.004	1.72
202207_at	ARL4C	ADP-ribosylation factor-like 4C	1.6	0.004	1.72
212192_at	KCTD12	Potassium channel tetramerisation domain containing 12	1.6	0.004	1.72
201670_s_at	MARCKS	Myristoylated alanine-rich protein kinase C substrate	1.6	0.006	1.72
202998_s_at	LOXL2	Lysyl oxidase-like 2	1.7	0.006	1.72
203060_s_at	PAPSS2	3'-phosphoadenosine 5'-phosphosulfate synthase 2	1.5	0.006	
218149_s_at	ZNF395	Zinc finger protein 395	1.5	0.006	1.72
212372_at	MYH10	Myosin, heavy polypeptide 10, non-muscle	-1.7	0.007	
202499_s_at	SLC2A3	Solute carrier family 2 (facilitated glucose transporter), member 3	1.7	0.009	1.72
205099_s_at	CCR1	Chemokine (C-C motif) receptor 1	1.6	0.009	1.72
212154_at	SDC2	Syndecan 2	1.8	0.009	1.72
212334_at	GNS	Glucosamine (N-acetyl)-6-sulfatase	1.5	0.009	
217028_at	CXCR4	Chemokine (C-X-C motif) receptor 4	1.8	0.009	1.72
202310_s_at	COL1A1	Collagen, type I, alpha 1	1.7	0.011	1.72
212820_at	DMXL2	Dmx-like 2	1.7	0.011	1.72
221059_s_at	COTL1	Coactosin-like 1	1.6	0.011	
211404_s_at	APLP2	Amyloid beta (A4) precursor-like protein 2	1.5	0.012	
201438_at	COL6A3	Collagen, type VI, alpha 3	1.7	0.013	1.72
202679_at	NPC1	Niemann-Pick disease, type C1	1.5	0.013	
212959_s_at	GNPTAB	N-acetylglucosamine-1-phosphate transferase, alpha and beta subunits	1.5	0.013	
221210_s_at	NPL	N-acetylneuraminate pyruvate lyase	1.8	0.013	1.72
201193_at	IDH1	Isocitrate dehydrogenase 1 (NADP+), soluble	1.6	0.016	1.72
201744_s_at	LUM	Lumican	1.6	0.016	
202934_at	HK2	Hexokinase 2	1.7	0.016	1.72
213655_at	YWHAE	tyrosine 3-monooxygenase	1.6	0.016	1.72
215049_x_at	CD163	CD163 molecule	1.6	0.016	1.72
217983_s_at	RNASET2	Ribonuclease T2	1.7	0.016	1.72
218353_at	RGS5	regulator of G-protein signalling 5	-1.7	0.016	
32128_at	CCL18	Chemokine (C-C motif) ligand 18	2.2	0.016	1.72
203645_s_at	CD163	CD163 molecule	1.7	0.019	1.72
209924_at	CCL18	Chemokine (C-C motif) ligand 18	2.0	0.019	1.72
212670_at	ELN	Elastin	-1.7	0.019	

Probe ID	Gene Symbol[a]	Gene Name[a]	Fold-change	Mann-Whitney U-test *p**	SAM q (%)*
212671_s_at	HLA-DQA1 / HLA-DQA2 / LOC650946	Major histocompatibility complex, class II, DQ alpha	1.8	0.019	1.72
201645_at	TNC	Tenascin C	1.6	0.023	1.72
202838_at	FUCA1	Fucosidase, alpha-L- 1, tissue	1.9	0.023	1.72
209070_s_at	RGS5	Regulator of G-protein signalling 5	-1.7	0.023	
209071_s_at	RGS5	Regulator of G-protein signalling 5	-1.7	0.023	
212582_at	OSBPL8	Oxysterol binding protein-like 8	1.6	0.023	1.72
213975_s_at	LYZ / LILRB1	Lysozyme	1.7	0.023	
200638_s_at	YWHAZ	Tyrosine 3-monooxygenase	1.5	0.028	
201170_s_at	BHLHB2	Basic helix-loop-helix domain containing, class B, 2	1.5	0.028	
203814_s_at	NQO2	NAD(P)H dehydrogenase, quinone 2	1.5	0.028	1.72
204774_at	EVI2A	Ecotropic viral integration site 2A	1.5	0.028	
214770_at	MSR1	Macrophage scavenger receptor 1	1.5	0.028	
217202_s_at	GLUL	Glutamate-ammonia ligase	1.7	0.028	1.72
200648_s_at	GLUL	Glutamate-ammonia ligase	1.9	0.033	1.72
201200_at	CREG1	Cellular repressor of E1A-stimulated genes 1	1.5	0.033	
202087_s_at	CTSL	Cathepsin L	1.5	0.033	1.72
202436_s_at	CYP1B1	Cytochrome P450, family 1, subfamily B, polypeptide 1	1.6	0.033	1.72
202859_x_at	IL8	Interleukin 8	1.9	0.033	1.72
202902_s_at	CTSS	Cathepsin S	1.5	0.033	1.72
204438_at	MRC1 / MRC1L1	Mannose receptor, C type 1	1.6	0.033	1.72
208146_s_at	CPVL	Carboxypeptidase, vitellogenic-like	1.5	0.036	1.72
201147_s_at	TIMP3	TIMP metallopeptidase inhibitor 3	1.9	0.039	1.72
202266_at	TTRAP	TRAF and TNF receptor associated protein	1.5	0.039	
210889_s_at	FCGR2B	Fc fragment of IgG, low affinity IIb, receptor (CD32)	1.5	0.039	
218559_s_at	MAFB	V-maf musculoaponeurotic fibrosarcoma oncogene homolog B (avian)	1.5	0.039	
202345_s_at	FABP5 / LOC653327	Fatty acid binding protein 5	1.6	0.047	1.72
204137_at	GPR137B	G protein-coupled receptor 137B	1.5	0.047	
206488_s_at	CD36	CD36 molecule	2.2	0.047	1.72
212158_at	SDC2	Syndecan 2	1.5	0.047	
219032_x_at	OPN3	Opsin 3	1.5	0.047	
219607_s_at	MS4A4A	Membrane-spanning 4-domains, subfamily A, member 4	1.8	0.047	1.72
200766_at	CTSD	Cathepsin D	1.5	0.055	
201669_s_at	MARCKS	Myristoylated alanine-rich protein kinase C substrate	1.6	0.055	
202912_at	ADM	Adrenomedullin	1.5	0.055	1.72

Probe ID	Gene Symbol[a]	Gene Name[a]	Fold-change	Mann-Whitney U-test *p**	SAM q (%)*
203665_at	HMOX1	Heme oxygenase (decycling) 1	2.2	0.055	1.72
203980_at	FABP4	Fatty acid binding protein 4, adipocyte	2.6	0.055	1.72
212657_s_at	IL1RN	Interleukin 1 receptor antagonist	1.6	0.055	1.72
214038_at	CCL8	Chemokine (C-C motif) ligand 8	1.6	0.055	1.72
215223_s_at	SOD2	Superoxide dismutase 2, mitochondrial	1.5	0.055	1.72
201212_at	LGMN	Legumain	1.6	0.065	
201963_at	ACSL1	Acyl-CoA synthetase long-chain family member 1	1.5	0.065	
202381_at	ADAM9	ADAM metallopeptidase domain 9 (meltrin gamma)	1.6	0.065	
204580_at	MMP12	Matrix metallopeptidase 12 (macrophage elastase)	2.3	0.065	1.72
209122_at	PLIN2	Perilipin 2	1.8	0.065	1.72
217294_s_at	ENO1	Enolase 1	1.7	0.065	1.72
200832_s_at	SCD	Stearoyl-CoA desaturase	1.6	0.076	1.72
202283_at	SERPINF1	Serpin peptidase inhibitor, clade F	1.5	0.076	
209351_at	KRT14	Keratin 14	1.8	0.076	1.72
209555_s_at	CD36	CD36 molecule	1.9	0.076	1.72
203561_at	FCGR2A	Fc fragment of IgG, low affinity IIa, receptor	1.5	0.088	
204446_s_at	ALOX5	Arachidonate 5-lipoxygenase	1.5	0.088	
203504_s_at	ABCA1	ATP-binding cassette, sub-family A (ABC1), member 1	1.5	0.102	
201785_at	RNASE1	Ribonuclease, RNase A family, 1	1.5	0.118	3.26
202917_s_at	S100A8	S100 calcium binding protein A8	1.6	0.118	3.26
203936_s_at	MMP9	Matrix metallopeptidase 9	1.6	0.118	
221730_at	COL5A2	Collagen, type V, alpha 2	1.5	0.136	3.26
34210_at	CD52	CD52 molecule	1.5	0.155	
201858_s_at	PRG1	Proteoglycan 1, secretory granule	1.5	0.177	
203381_s_at	APOE	Apolipoprotein E	1.6	0.177	4.00
217148_x_at	IGLV2-14	Immunoglobulin lambda variable 2-14	-1.7	0.177	
202075_s_at	PLTP	Phospholipid transfer protein	1.5	0.201	
204259_at	MMP7	Matrix metallopeptidase 7	1.5	0.201	4.72
203290_at	HLA-DQA1	Major histocompatibility complex, class II, DQ alpha 1	-2.0	0.227	
216984_x_at	IGL@	Immunoglobulin lambda locus	-1.7	0.227	
211896_s_at	DCN	Decorin	1.5	0.256	
213831_at	HLA-DQA1	Major histocompatibility complex, class II, DQ alpha 1	2.0	0.320	
217022_s_at	IGHA1 / IGHA2	Immunoglobulin heavy constant alpha 1	1.7	0.943	

[a]Genes are separated by hyphen if the probe detects several homologous genes, gene name is given for the first one. ***** *p* indicates the pairwise unadjusted p-values from nonparametric Mann-Whitney U-test, q the lowest FDR at which the gene is called significant from Significance Analysis of Microarrays (SAM).

Supplemental Table 4. Enrichment of the genes showing ≥ 1.5-fold expression change (n=103) in Gene Ontology categories

GO Category	Obs. # of genes	Exp. # of genes	*p*-value	Gene Symbol
Biological Process				
Catabolism	9	3.62	0.0096	MMP7, MMP9, GNS, RNASET2, LYZ, HMOX1, FUCA1, IDH1, HK2
Cholesterol homeostasis	2	0.06	0.0015	NPC1, APOE
Locomotory behavior	5	0.89	0.0019	CCL8, CCL18, CCR1, IL8, VEGFA
Organ development	12	4.13	0.0007	DCN, SERPINF1, CYP1B1, ELN, COL1A1, COL6A3, IL8, MAFB, VEGFA, PAPSS2, MMP9, BTG1
Physiological defense response	8	2.55	0.0037	ALOX5, CCL8, CCL18, CD163, CCR1, S100A8, IL1RN, LYZ
Response to wounding	9	2.5	0.0008	ALOX5, CCL8, CCL18, CD163, ADM, CCR1, S100A8, IL1RN, CD36
Physiological response to wounding	8	2.42	0.0027	ALOX5, CCL8, CCL18, CD163, CCR1, S100A8, IL1RN, CD36
Chemotaxis	5	0.85	0.0015	CCL8, CCL18, CCR1, IL8, VEGFA
Inflammatory response	7	1.79	0.0020	ALOX5, CCL8, CCL18, CD163, CCR1, S100A8, IL1RN
Regulation of neurogenesis	2	0.09	0.0038	SERPINF1, APOE
Taxis	5	0.85	0.0015	CCL8, CCL18, CCR1, IL8, VEGFA
Carbohydrate metabolism	8	2.79	0.0064	MMP7, MMP9, MMP12, GNS, SLC2A3, FUCA1, IDH1, HK2
Cellular catabolism	9	3.06	0.0033	MMP7, MMP9, GNS, RNASET2, LYZ, HMOX1, FUCA1, IDH1, HK2
Organ morphogenesis	6	1.72	0.0072	DCN, SERPINF1, ELN, IL8, VEGFA, BTG1
Sensory organ development	2	0.14	0.0082	MAFB, CYP1B1
Vasculature development	4	0.59	0.0027	SERPINF1, IL8, VEGFA, BTG1
Lipid transport	4	0.47	0.0012	PLTP, NPC1, OSBPL8, APOE
Positive chemotaxis	2	0.04	0.0007	IL8, VEGFA
Regulation of chemotaxis	2	0.06	0.0015	IL8, VEGFA
Response to reactive oxygen species	2	0.09	0.0032	SOD2, APOE
Blood vessel morphogenesis	4	0.58	0.0026	SERPINF1, IL8, VEGFA, BTG1
Blood vessel development	4	0.58	0.0026	SERPINF1, IL8, VEGFA, BTG1
Cellular carbohydrate metabolism	7	2.16	0.0060	MMP7, MMP9, MMP12, GNS, FUCA1, IDH1, HK2
Positive regulation of chemotaxis	2	0.05	0.0011	IL8, VEGFA
Regulation of positive chemotaxis	2	0.04	0.0007	IL8, VEGFA
Aminoglycan catabolism	2	0.03	0.0004	FUCA1, GNS
Angiogenesis	4	0.56	0.0023	SERPINF1, IL8, VEGFA, BTG1
Peptidoglycan metabolism	3	0.18	0.0007	MMP7, MMP9, MMP12

Category	Obs. # of genes	Exp. # of genes	*p*-value	Gene Symbol
Positive regulation of positive chemotaxis	2	0.04	0.0007	IL8, VEGFA
Glycosaminoglycan catabolism	2	0.03	0.0004	FUCA1, GNS
Phosphate transport	4	0.47	0.0013	COL1A1, COL5A2, COL6A3, MSR1
Induction of positive chemotaxis	2	0.04	0.0007	IL8, VEGFA
Collagen catabolism	2	0.15	0.0092	MMP7, MMP9
Molecular function				
Collagen binding	2	0.13	0.0067	LUM, MMP9
Peptidase activity	9	3.38	0.0063	MMP7, MMP9, MMP12, CPVL, CTSS, CTSD, CTSL, LGMN, ADAM9
Sterol transporter activity	2	0.54	0.0012	NPC1, ABCA1
G-protein-coupled receptor binding	3	0.38	0.0061	CCL8, CCL18, IL8
Endopeptidase activity	8	2.45	0.0030	MMP7, MMP9, MMP12, CTSS, CTSD, CTSL, LGMN, ADAM9
Scavenger receptor activity	3	0.17	0.0006	LOXL2, CD163, MSR1
C-C chemokine binding	2	0.13	0.0077	CXCR4, CCR1
Chemokine activity	3	0.31	0.0037	CCL8, CCL18, IL8
Chemokine receptor binding	3	0.32	0.0040	CCL8, CCL18, IL8
Metalloendopeptidase activity	4	0.66	0.0042	MMP12, MMP7, MMP9, ADAM9
C-C chemokine receptor activity	2	0.13	0.0077	CXCR4, CCR1
Cellular Component				
Extracellular space	11	3.2	0.0003	LOXL2, CCL18, MMP7, MMP9, IL8, ADM, VEGFA, S100A8, APOE, IL1RN, CCL8
Extracellular matrix (sensu Metazoa)	12	1.84	$2x^{10-7}$	LUM, DCN, TIMP3, MMP7, MMP9, MMP12, ELN, VEGFA, COL1A1, COL5A2, COL6A3, TNC
Collagen	4	0.25	0.0001	LUM, COL1A1, COL6A3, COL5A2
Fibrillar collagen	3	0.08	$5x10^{-5}$	LUM, COL1A1, COL5A2
Vacuole	8	1.06	$1x10^{-7}$	GNS, CTSD, CTSS, CTSL, GNPTAB, NPC1, FUCA1, LGMN
Lytic vacuole	8	0.95	$4x10^{-6}$	GNS, CTSD, CTSS, CTSL, GNPTAB, NPC1, FUCA1, LGMN
Lysosome	8	0.95	$4x10^{-6}$	GNS, CTSD, CTSS, CTSL, GNPTAB, NPC1, FUCA1, LGMN

Gene enrichment analysis was performed for the probes showing ≥1.5-fold expression change (n=103) using WebGestalt software [4]. For details, see Supplemental Methods. Categories below level 3, having two or more genes and enriched with a p-value < 0.01 are listed. Categories with significance level <0.001 are highlighted. Obs. # of genes = the number of observed genes in the category. Exp. # of genes = the expected number of genes in the category. For more information on Gene-Ontology see http://www.geneontology.org/.

Supplemental Table 5. Enrichment of the genes showing ≥1.5-fold expression change (n=103) in KEGG and Biocharta (BioC) pathways

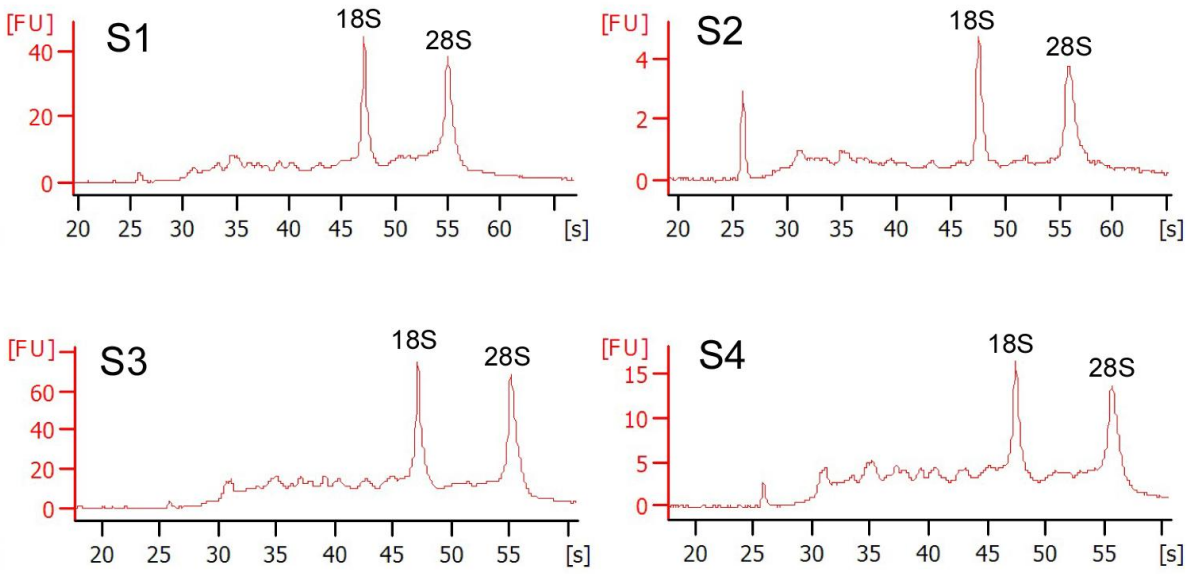
Pathway	Source	Number of genes	*p*-value	Gene Symbol	Gene Name[a]
ECM-receptor interaction (hsa04512)	KEGG	6	$1.9x10^{-5}$	COL1A1, COL5A2, COL6A3, TNC, SDC2, CD36	Collagen types I alpha 1, V alpha 2 and VI alpha 3/ tenascin C/ syndecan 2/ CD36
PPAR signaling pathway (hsa03320)	KEGG	5	$9.3x10^{-5}$	FABP4, ACSL1, PLTP, SCD, CD36	Fatty acid binding protein 4/ acyl-CoA synthetase long-chain family member 1/ phospholipid transfer protein/ stearoyl-CoA desaturase/CD36
Antigen processing and presentation (hs0412)	KEGG	4	$6.6x10^{-4}$	CTSL, CTSS, HLA-DQA1, LGMN	Cathepsins L and S/ major histocompatibility complex class II DQ alpha 1/ legumain
Inhibition of matrix metalloproteinases	BioC	2	0.002	MMP9, TIMP3	Matrix metallopeptidase 9/ TIMP metallopeptidase inhibitor 3
Cytokine-cytokine receptor interaction (hsa04060)	KEGG	6	0.005	CCR1, IL8, CCL8, CCL18, VEGF, CXCR4	Chemokine (C-C motif) receptor 1/ interleukin 8/ chemokine (C-C motif) ligands 8 and 18/ vascular endothelial growth factor/ chemokine (C-X-C motif) receptor 4
Cell communication (hsa01430)	KEGG	4	0.005	COL1A1, COL5A2, COL6A3, TNC	Collagen types I alpha 1, V alpha 2 and VI alpha 3/ tenascin C
Focal adhesion (hsa04510)	KEGG	5	0.009	COL1A1, COL5A2, COL6A3, TNC, VEGF	Collagen types I alpha 1, V alpha 2 and VI alpha 3/ tenascin C/ vascular endothelial growth factor

[a]Hyphen separates genes. Gene enrichment analysis was performed for the probes showing ≥1.5-fold expression change (n=103) using WebGestalt software [4]. We used a prestored gene set (WebGestalt_HG_U133A) as a reference set and applied a hypergeometric test to identify enriched pathways. The default significance level of 0.01 was used and a minimum number of genes was set as two. For more information on KEGG see http://www.genome.ad.jp/kegg/ and on BioCharta see http://www.biocarta.com. Pathways are arranged according to *p*-value.

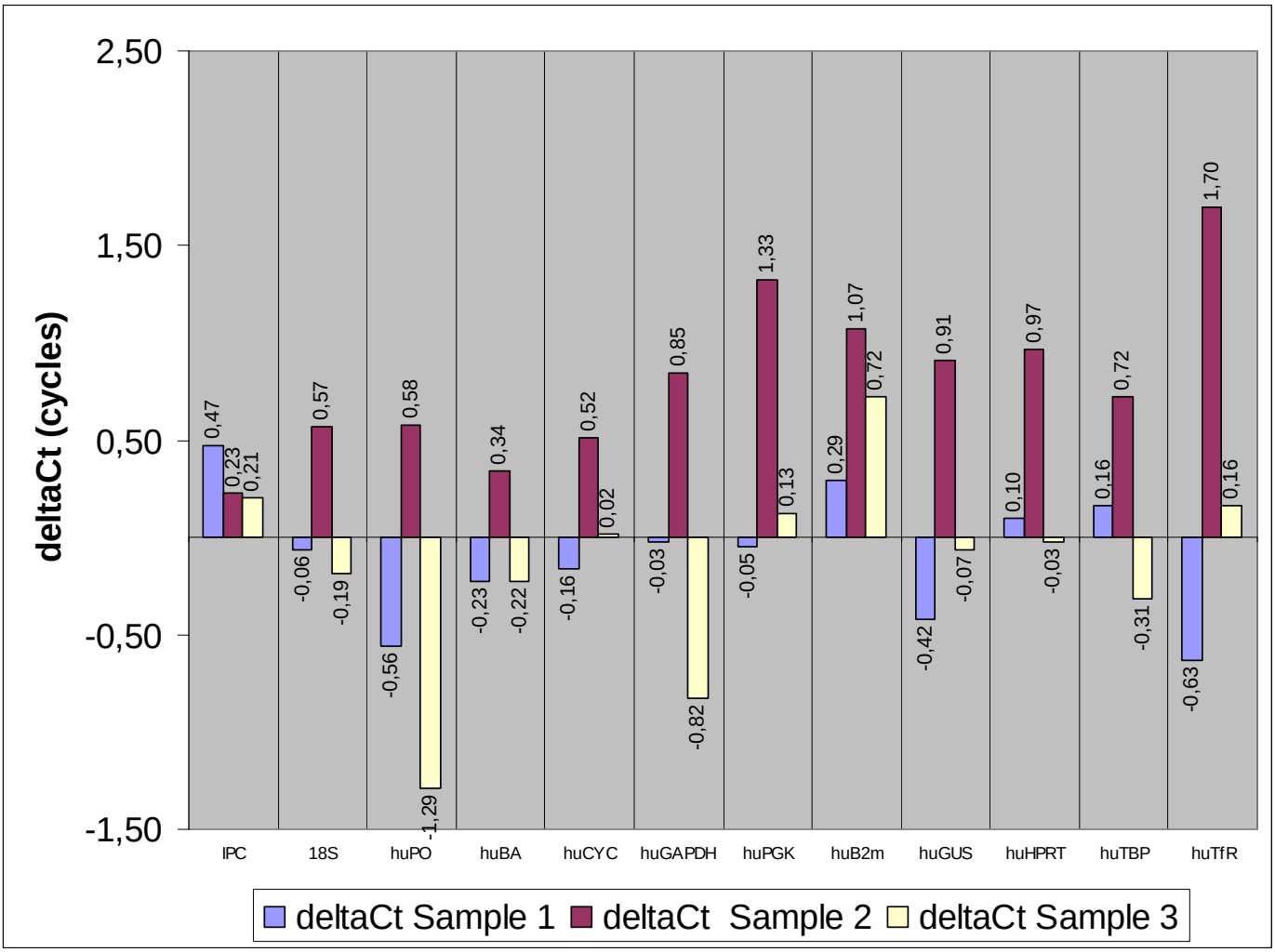
Supplemental Table 6. Correlations of qPCR confirmed expression changes with clinical data and plaque characteristics.

Gene Symbol	CD36	CD163	FABP4	PLIN2	GLUL	CCL18	FUCA1	IL1RN	HMOX1	S100A8
Clinical parameters										
Gender	-0.28	-0.17	-0.14	-0.36*	-0.35*	-0.22	-0.31*	-024	-0.27	-0.10
Degree of ICA stenoses	0.39*	0.38*	0.33*	0.52**	0.44*	0.34*	0.33*	0.42*	0.40*	0.21
Laboratory measurements										
Hematocrit	0.19	0.19	0.11	0.35*	0.23	0.26	0.22	0.15	0.06	0.01
High-sensitivity CRP	-0.12	-0.09	-0.19	-0.23	-0.18	-0.16	-0.06	-0.21	-0.24	-0.14
LDL-cholesterol	0.42*	0.31	0.33	0.35	0.38	0.18	0.31	0.18	0.43*	0.08
HDL-cholesterol	-0.16	-0.18	-0.01	-0.15	-0.03	-0.26	-0.16	-0.00	-0.05	-0.04
Fibrinogen	-0.09	-0.01	-0.03	-0.05	0.02	-0.07	-0.04	0.11	-0.04	-0.30
TPA antigen	0.27	0.29	0.29	0.28	0.25	0.25	0.25	0.44*	0.13	0.26
Macroscopic plaque characteristics										
Ulceration	0.44*	0.45*	0.44*	0.54**	0.46*	0.33*	0.25	0.43*	0.49**	0.13
Intraplaque hemorrhage	0.45*	0.43*	0.36*	0.31	0.40*	0.46*	0.34*	0.27	0.51**	0.21
Intramural thrombus	0.20	0.25	0.28	0.34*	0.28	0.15	0.12	0.34*	0.14	0.21
Loose atheroma	0.11	0.02	0.07	0.04	0.20	0.05	0.18	0.27	0.04	-0.08
Calcification	-0.35*	-0.20	-0.30	-0.22	-0.33*	-0.33*	-0.39*	-0.32*	-0.32*	-0.08
Plaque components[a]										
Macrophage density (HAM 56)(4)	0.19	0.23	0.10	0.32	0.33	0.14	0.13	0.35*	0.35*	0.01
T-cell density (CD45RO)(4)	-0.04	-0.01	-0.10	0.12	0.06	0.08	0.07	0.11	-0.07	-0.22
Mast cell density (anti-tryptase)(4)	0.21	0.18	0.05	0.08	0.13	0.30	0.33	-0.06	0.14	-0.11
Smooth muscle cells (alpha-actin)	-0.10	-0.09	-0.13	-0.06	-0.03	-0.13	0.08	-0.03	-0.11	-0.16
Markers of cell death and proliferation[a]										
Activated caspase 3(5)	0.44*	0.38*	0.31	0.42*	0.51*	0.44*	0.55**	0.37	0.36*	0.24
TUNEL positivity(5)	0.43*	0.35*	0.35*	0.38*	0.44*	0.62**	0.39*	0.38*	0.30	0.21
Ki67 proliferation marker(5)	0.39*	0.41*	0.29	0.51*	0.50*	0.28	0.31	0.44*	0.48*	0.17

The correlation coefficient reported is Spearman's rho. Correlations were analyzed using all available data (n ranging from 24-43). Pairwise unadjusted *p*-values are denoted by * ≤0.05 or ** ≤0.001. Significant correlations ($p<0.05$) after Benjamini and Hochberg multiple testing correction are highlighted with grey; [a]Immunohistochemical analysis, immunostained antigen is specified in parenthesis if not apparent from the heading, the references to the original data are given in parenthesis.

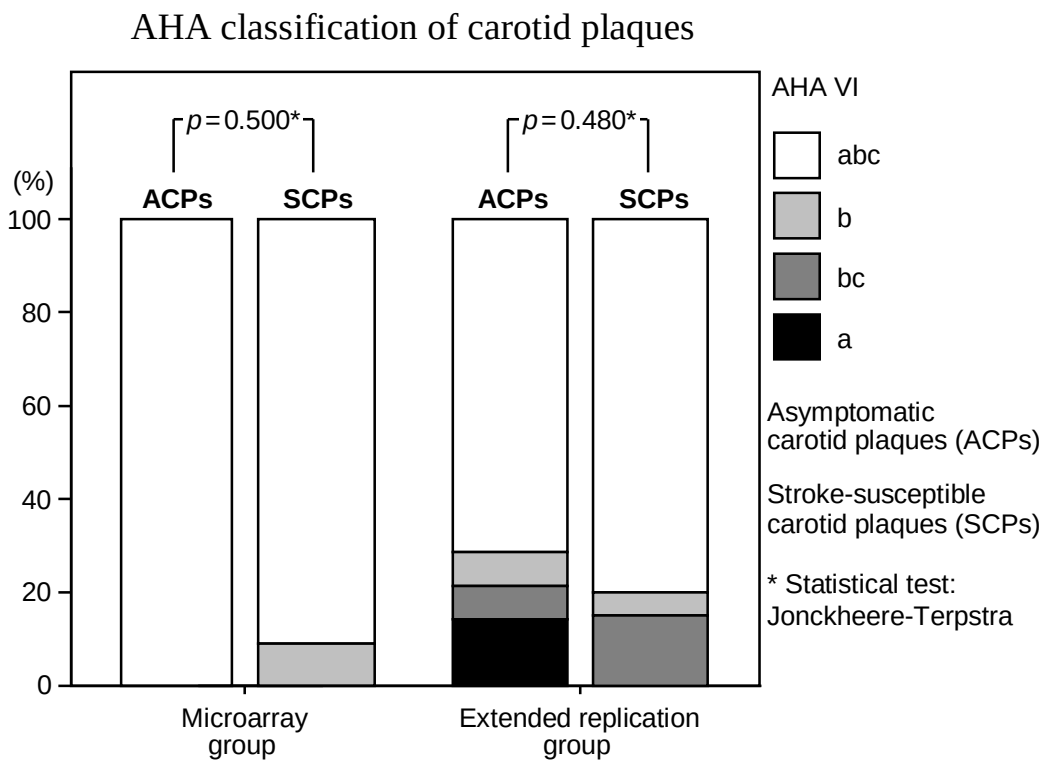


Supplemental Figure 1. Representative electropherograms of mRNA samples from Bioanalyzer. On the y-axis are fluoresence units (FU) and on the X-axis time in seconds (s). S1 to S4 refer to four different mRNA samples extracted from carotid plaques. 18S and 28S refer to corresponding ribosomal RNAs.



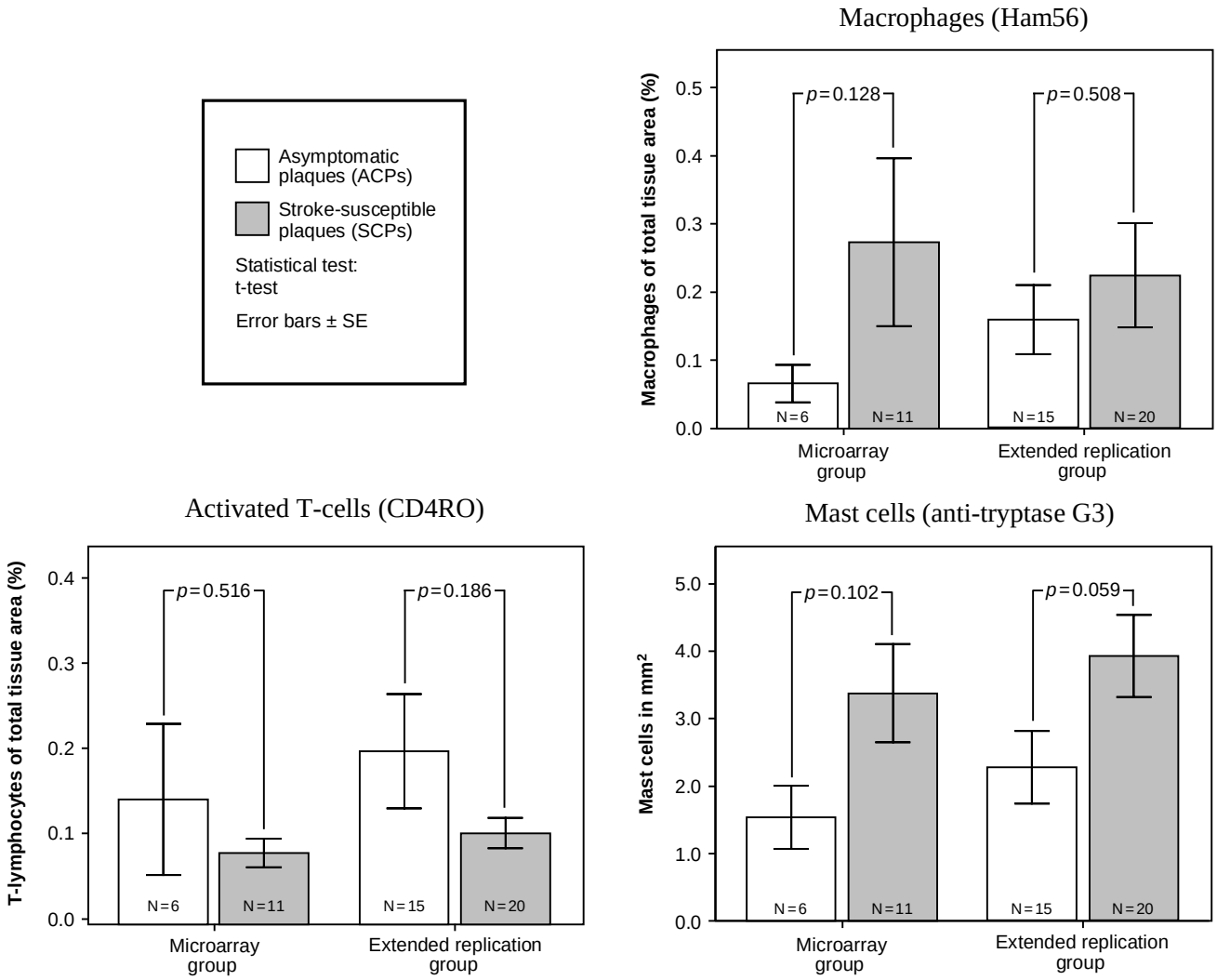
Supplemental Figure 2. Evaluation of 11 endogenous control genes in the carotid plaque cDNA samples by qPCR. The y-axis (deltaCT (cycles)) shows the difference in the threshold cycle between the sample and a calibrator. One deltaCT equals to a twofold difference in initial template concentration. On the X-axis are tested control genes from left to right: IPC = internal plate control; 18S; HuPO (acidic ribosomal protein), huBA (beta-actin), huCYC (cyclophilin), huGAPDH (glyceraldehyde-3-phosphate dehydrogenase), huPGK (phosphoglycerokinase), huB2m (beta2-microglobulin), huGUS (beta-glucuronidase), huHPRT (hypoxanthine ribosyl transferase), huTBP (transcription factor IID, TATA binding protein) and huTfR (transferrin receptor). DeltaCT Sample 1 to Sample 3 refer to three different cDNA samples from carotid plaques. Beta-actin (huBA) was selected for the endogenous control gene because it showed least variability between samples.

Supplemental Figure 3.



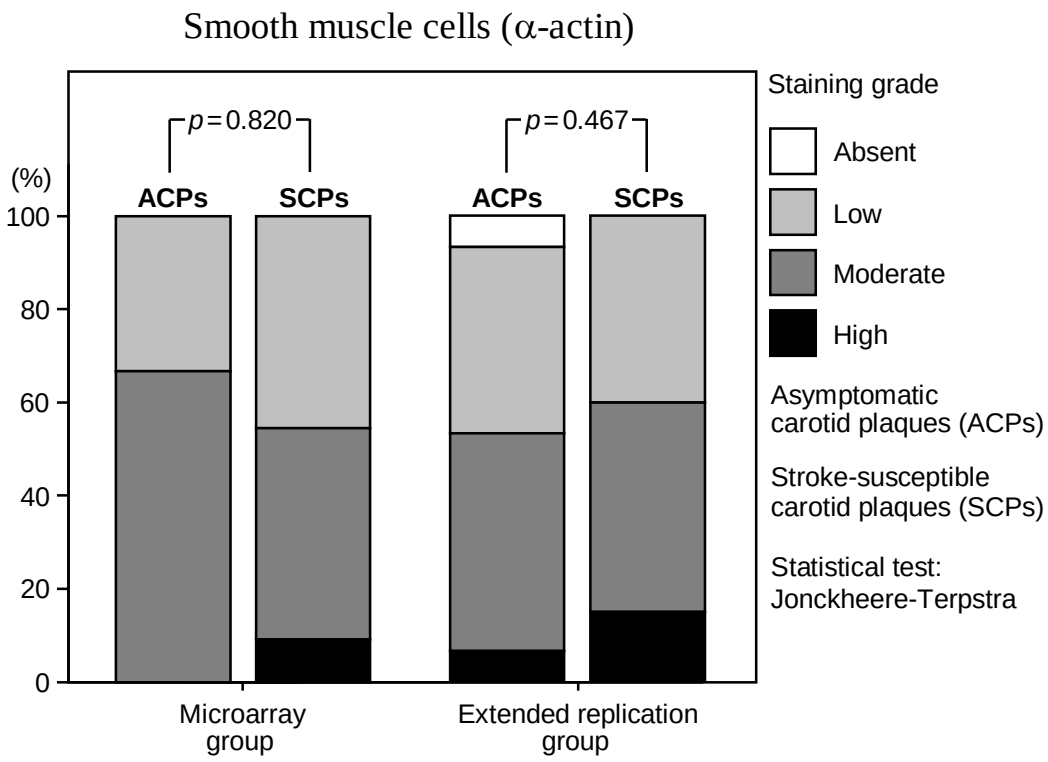
Histological classification of the carotid plaques according to AHA by patient group and symptom. AHA class VIa lesions contain microscopic surface ulceration, VIb hematoma/hemorrhage, and VIc thrombotic deposit [5].

Supplemental Figure 4.

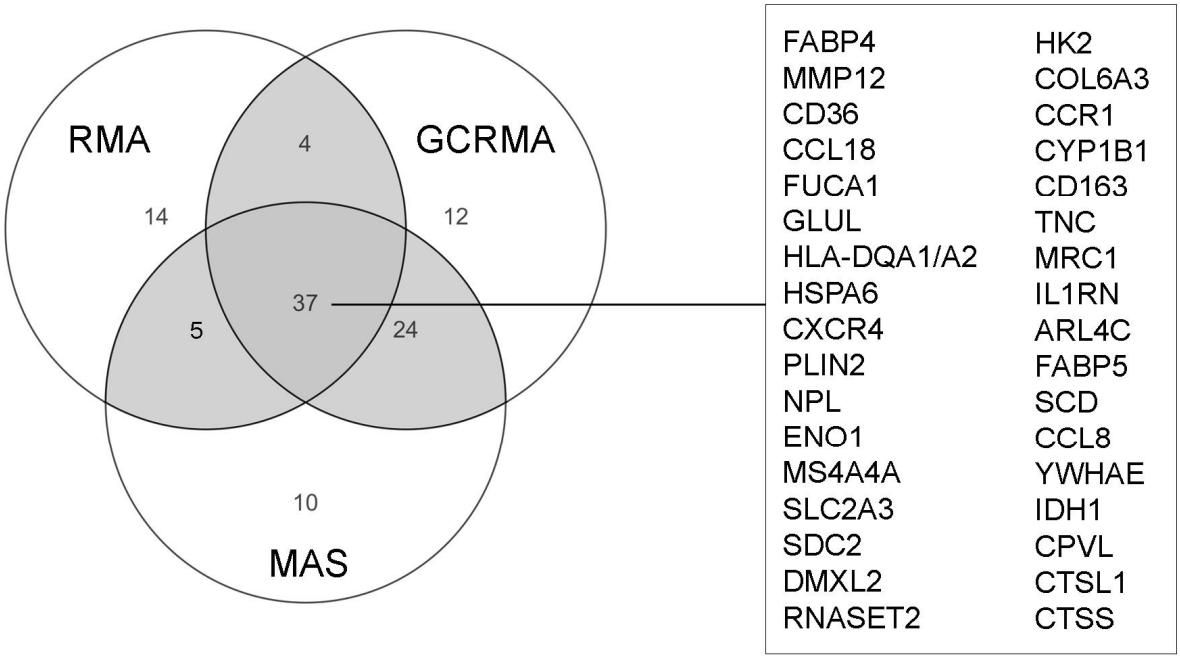


The amounts of macrophages, activated T-cells, and mast cells in ACPs and SCPs in the microarray and extended replication groups. Immunohistochemical analysis was done from the adjacent CP slice used for RNA extraction. The immunostained antigen is given in parenthesis. The results were analyzed using quantitative analysis of immunohistochemistry[6].

Supplemental Figure 5.

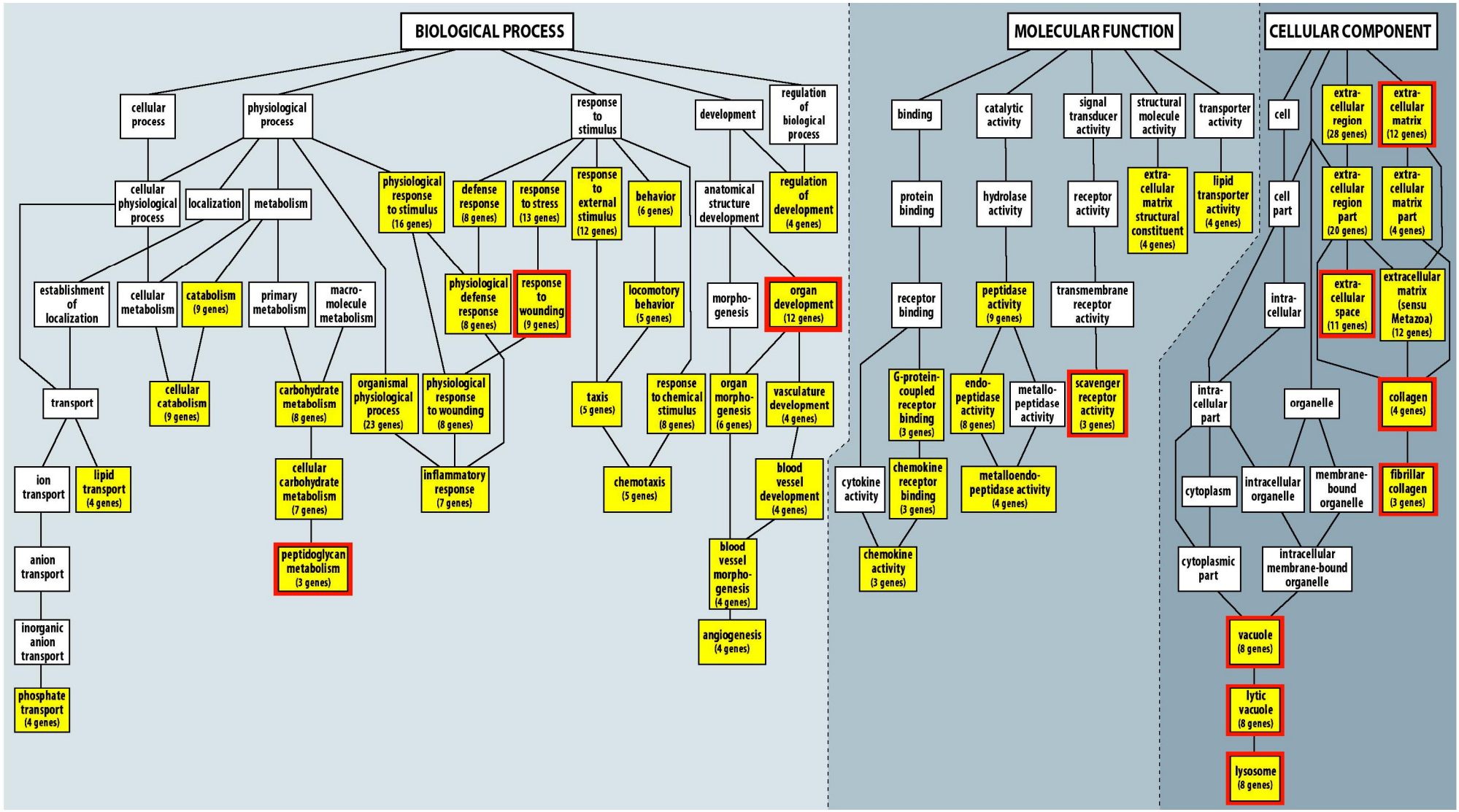


The amount of smooth muscle cells in ACPs and SCPs in the microarray and extended replication groups. Immunohistochemical analysis was done from the adjacent CP slice used for RNA extraction. The immunostained antigen is given in parenthesis.



Supplemental Figure 6. Venn diagram showing the number of significant probes common to different pre-processing algorithms. Hybridization data was pre-processed by either RMA, GC-RMA or MAS5 and filtered by using congruent criteria, after which genes that were differentially expressed at 5% false discovery rate were selected by SAM-analysis. This resulted in 60 probes for RMA, 77 for GC-RMA and 75 probes for MAS5. The three probe lists where combined in the Venn diagram to get the intersection of the lists. The 37 probes corresponding to 35 genes common to all three algorithms are shown in the box on the right. The gene symbols are listed in the descending (and from the leftward panel to the rightward panel) order of RMA fold-change (i.e. FABP4 has the highest fold-change when calculated from the RMA normalized data). The corresponding gene names can be found in Supplemental Table 3.

Supplemental Figure 7.



Directed acyclic graph of enriched gene ontology categories. The 3 main ontologies are shown in the upper row and further discriminated by background colour. Below each ontology, the categories enriched with a *p*-value <0.01 and including ≥3 genes (yellow background) and a *p*-value <0.001 (red outline) and the parental terms necessary for illustration are shown. Genes enriched in each category are listed in Supplemental Table 5. A gene can occur in one or more ontologies.

REFERENCES

1. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4:249-264

2. Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci USA 98:5116-5121

3. Wu Z, Irizarry RA (2004) Preprocessing of oligonucleotide array data. Nat Biotech 22:656-658

4. Zhang B, Kirov S, Snoddy J. WebGestalt (2005) An integrated system for exploring gene sets in various biological contexts. Nucleic Acids Res 33:W741-W748

5. Stary H, Chandler A, Dinsmore R, Fuster V, Glagov S, Insull W, Rosenfeld M, Schwartz C, Wagner W, Wissler R. (1995) 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*. 92:1355-1374.

6. Lehtonen-Smeds EM, Mäyränpää M, Lindsberg PJ, Soinne L, Saimanen E, Jarvinen AA, Salonen O, Carpen O, Lassila R, Sarna S et al (2005) Carotid plaque mast cells associate with atherogenic serum lipids, high grade carotid stenosis and symptomatic carotid artery disease - results from the Helsinki Carotid Endarterectomy Study. *Cerebrovasc Dis*. 19:291-301.