Genes Genet. Syst. (2014) **89**, p. 253–258

# Identification by microarray technology of key genes involved in the progression of carotid atherosclerotic plaque

Jian Wang[1,2#], Bo Wei[1#], Shuwei Cao[3], Fengke Xu[2], Wei Chen[2], Haifeng Lin[2], Chenghua Du[2] and Zhigang Sun[1,2*]

[1]*Department of Neurosurgery, China-Japan Union Hospital of Jilin University, Changchun, Jilin, 130033, China*
[2]*Department of Neurosurgery, The Affiliated Hospital of Inner Mongolia National University, Tongliao, Inner Mongolia, 028000, China*
[3]*Department of Interventional Therapy, Weihai Municipal Hospital, Weihai, Shandong, 264200, China*

*(Received 7 May 2014, accepted 1 November 2014)*

A comparative analysis of gene expression profiles between early and advanced carotid atherosclerotic plaque was performed to identify key genes and pathways involved in the progression of carotid atherosclerotic plaque. Gene expression data set GSE28829 was downloaded from Gene Expression Omnibus, including 13 early and 16 advanced atherosclerotic plaque samples from human carotid. Differentially expressed genes (DEGs) were identified using the package limma of *R*. Principal component analysis was carried out for the DEGs with package rgl of *R*. A gene coexpression network was constructed with information from COXPRESdb and then visualized with Cytoscape. Functional enrichment analysis was performed with DAVID and pathway enrichment analysis was done with KEGG. A total of 319 DEGs were identified in the advanced atherosclerotic plaque samples compared with early atherosclerotic plaque samples, including 267 up-regulated genes and 52 down-regulated genes. In the gene coexpression network, TYRO protein tyrosine kinase binding protein was the hub gene with a degree of 23. Functional enrichment analysis and pathway enrichment analysis suggested that the immune response played a critical role in the progression of carotid atherosclerotic plaque. A number of key genes were revealed in carotid atherosclerotic plaque, and are potential biomarkers for diagnosis or treatment. These findings may also guide future research to better decipher the progression of atherosclerosis.

**Key words:** carotid atherosclerotic plaque, differentially expressed genes, principal component analysis, gene coexpression network, pathway enrichment analysis

## INTRODUCTION

Atherosclerosis is a chronic immunoinflammatory disease elicited by the accumulation of lipids, inflammatory cells, smooth muscle cells, and extracellular matrix in the arterial intima (Galkina and Ley, 2009). Chronically expanding atherosclerotic plaque can occlude arteries completely and thus lead to severe outcomes, such as infarction and stroke (Li et al., 2010; Polak et al., 2011). In addition, carotid atherosclerosis is associated with a high risk of developing coronary heart disease (Hallerstam et al., 2004). These complications of atherosclerosis cause high mortality and a heavy economical burden. Thus, it is necessary to discover biomarkers for early diagnosis or treatment of atherosclerosis.

Cytokines, such as interleukin-10 (Hansson et al., 2002; Pinderski et al., 2002) and transforming growth factor-β (Mallat et al., 2001; Grainger, 2004), within atherosclerotic plaque play a key role in both the development and progression of atherosclerosis. Cytokines provoke monocyte/lymphocyte recruitment and infiltration into the subendothelium (Frieri, 2012). Targeted deletion of genes encoding costimulatory factors and proinflammatory cytokines results in milder disease symptoms in mouse models, whereas interference with regulatory immunity accelerates disease (Hansson and Hermansson, 2011).

Edited by Yoko Satta
* Corresponding author. E-mail: sszhigang@163.com
# Jian Wang and Bo Wei are regarded as co-first author.

Besides, growing evidence supports the critical role of oxidative stress in atherosclerosis (Singh and Jialal, 2006; Bonomini et al., 2008; Victor et al., 2009). A main source of reactive oxygen species in vascular cells is the reduced nicotinamide adenine dinucleotide/nicotinamide adenine dinucleotide phosphate oxidase system (Sorescu et al., 2001; Sheehan et al., 2011). However, the mechanisms contributing to the formation of atherosclerotic plaque remain poorly understood.

To globally identify key genes and pathways associated with atherosclerotic plaque, especially those contributing to its progression, a comparative analysis of gene expression profiles between early and advanced carotid atherosclerotic plaque was performed with various bioinformatic tools. Using the same gene expression profiles, Döring et al. (2012) identified an increase in expression of the plasmacytoid dendritic cell marker LL37 in advanced carotid artery specimens compared with early lesions, but the gene's function and its interaction remain uninvestigated. We aimed to further screen differentially expressed genes (DEGs) between advanced and early carotid artery lesions and explore their underlying function by Gene Ontology (GO) and pathway enrichment analysis. A gene coexpression network was also constructed with information from COXPRESdb. These findings may advance our understanding of the pathogenesis of atherosclerotic plaque and provide potential biomarkers for early diagnosis or treatment of atherosclerotic plaque.

## MATERIALS AND METHODS

**Gene expression data** A gene expression data set (accession number GSE28829) deposited by Döring et al. (2012), based on the GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array platform (Affymetrix, Santa Clara, CA, USA), was downloaded from Gene Expression Omnibus. Annotation files were also acquired. GSE28829 includes 13 early and 16 advanced atherosclerotic plaque samples from human carotid from the Maastricht Pathology Tissue Collection.

**Pretreatment and differential analysis** Probes were mapped to genes and we normalized the microarray data using the Support Vector Regression method (Fujita et al., 2006). Differential analysis between early and advanced atherosclerotic plaque was performed using the package limma (Smyth, 2005) of *R*. Multiple testing correction was performed using the Benjamini & Hochberg method (Benjamini and Hochberg, 1995; Reiner-Benaim, 2007). False discovery rate (FDR) < 0.05 and $|\log_2\text{Fold Change (FC)}| > 1$ were set as the thresholds for selecting DEGs.

**Principal component analysis (PCA)** PCA is a commonly used clustering algorithm that can minimize errors by simplifying the multivariate data matrix (Abdi and Williams, 2010). In the present study, PCA was carried out using the function prcomp in the package rgl (Adler and Murdoch, 2012) of *R*.

**Construction of gene coexpression networks** Coregulation of genes by the same transcription factor is termed gene coexpression. Coexpressed genes may have closely related biological functions. In the present study, DEGs were divided into two groups, up-regulated and down-regulated genes. Gene coexpression networks were constructed with information from COXPRESdb (Obayashi et al., 2008), and were then visualized with Cytoscape (Smoot et al., 2011). Coexpression with score > 0.8 was retained in the networks.

**GO enrichment analysis** Functional enrichment analysis is helpful in revealing altered biological functions. In the present study, functional enrichment analysis was applied to the DEGs using DAVID (Database for Annotation, Visualization and Integration Discovery, http://david.abcc.ncifcrf.gov/) (Huang et al., 2008). FDR < 0.05 was set as the cut-off for selecting significantly enriched functional terms.

**Pathway enrichment analysis** Pathway enrichment analysis was performed using GenMapp from Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2004, 2006). FDR < 0.05 was set as the cut-off for selecting significantly over-represented pathways.

## RESULTS

**Differentially expressed genes** After data pretreatment (Fig. 1), 42,450 genes were obtained from the raw data. In comparison to early atherosclerotic plaque, a total of 319 DEGs (FDR < 0.05 and $|\log_2 FC| > 1$) were identified in advanced atherosclerotic plaque, comprising 267 up-regulated genes and 52 down-regulated genes.

**PCA** The PCA result is shown in Fig. 2. Axes represented functional values, and each column value could be regarded as a factor. The factors were then arranged according to their effects, and the top three factors were used as axes of a three-dimensional diagram. The three factors could well separate early carotid atherosclerotic plaque samples from advanced carotid atherosclerotic plaque samples, suggesting the validity of the DEGs between the two groups of carotid atherosclerotic plaque samples.

**Gene coexpression network of the DEGs** A total of 84 coexpressions (score > 0.8) were revealed in the DEGs by COXPRESdb. They were all up-regulated genes. The network was visualized by Cytoscape (Fig. 3), and

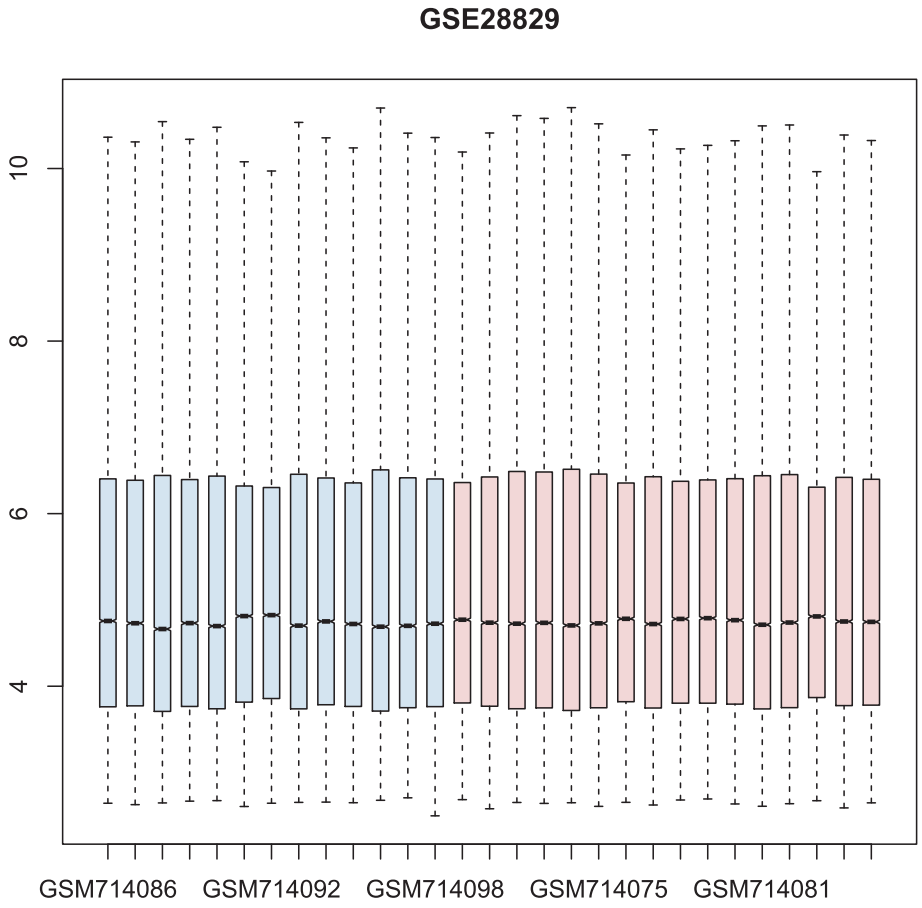


Fig. 1. Box plots for the normalized gene expression data. The x-axis represents 13 early (blue, GSM714086-GSM714098) and 16 advanced (red, GSM714070-GSM714085) atherosclerotic plaque samples, and the y-axis represents the gene expression level in the samples. The black lines in the boxes represent medians and they are almost at the same level, suggesting a good performance of normalization.

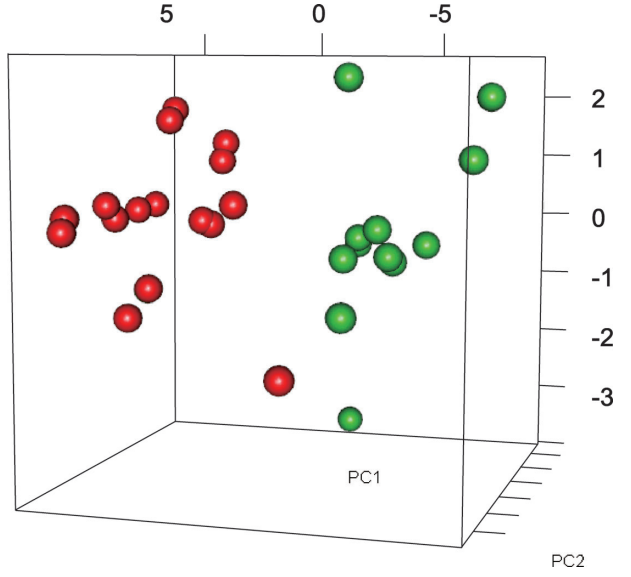


Fig. 2. Result of PCA analysis of the differentially expressed genes. Early and advanced carotid atherosclerotic plaque samples are in red and green, respectively. PC1 and PC2 represent the first two principal components.

consists of 48 nodes including CD14, CD163 and complement component 1, q subcomponent, chains A, B and C (C1QA, C1QB and C1QC, respectively), as well as 84 edges. The hub gene was TYRO protein tyrosine kinase binding protein (TYROBP) with a degree (the number of interactions between genes) value of 23.

**Functional enrichment analysis** Functional enrichment analysis was applied to the genes from the gene coexpression network using DAVID. The top 10 gene ontology (GO) terms are listed in Table 1, including immune response, defense response and adaptive immune response. All three of these GO terms are closely related to immunity.

**Pathway enrichment analysis** Pathway enrichment analysis was performed for the genes from the gene coexpression network using GenMApp from KEGG. Only one significant pathway was revealed: antigen processing and presentation (hsa04612, FDR = 0.019). Six genes were involved in this pathway: interferon gamma-inducible protein 30 (IFI30), cathepsin S (CTSS), major histocom-

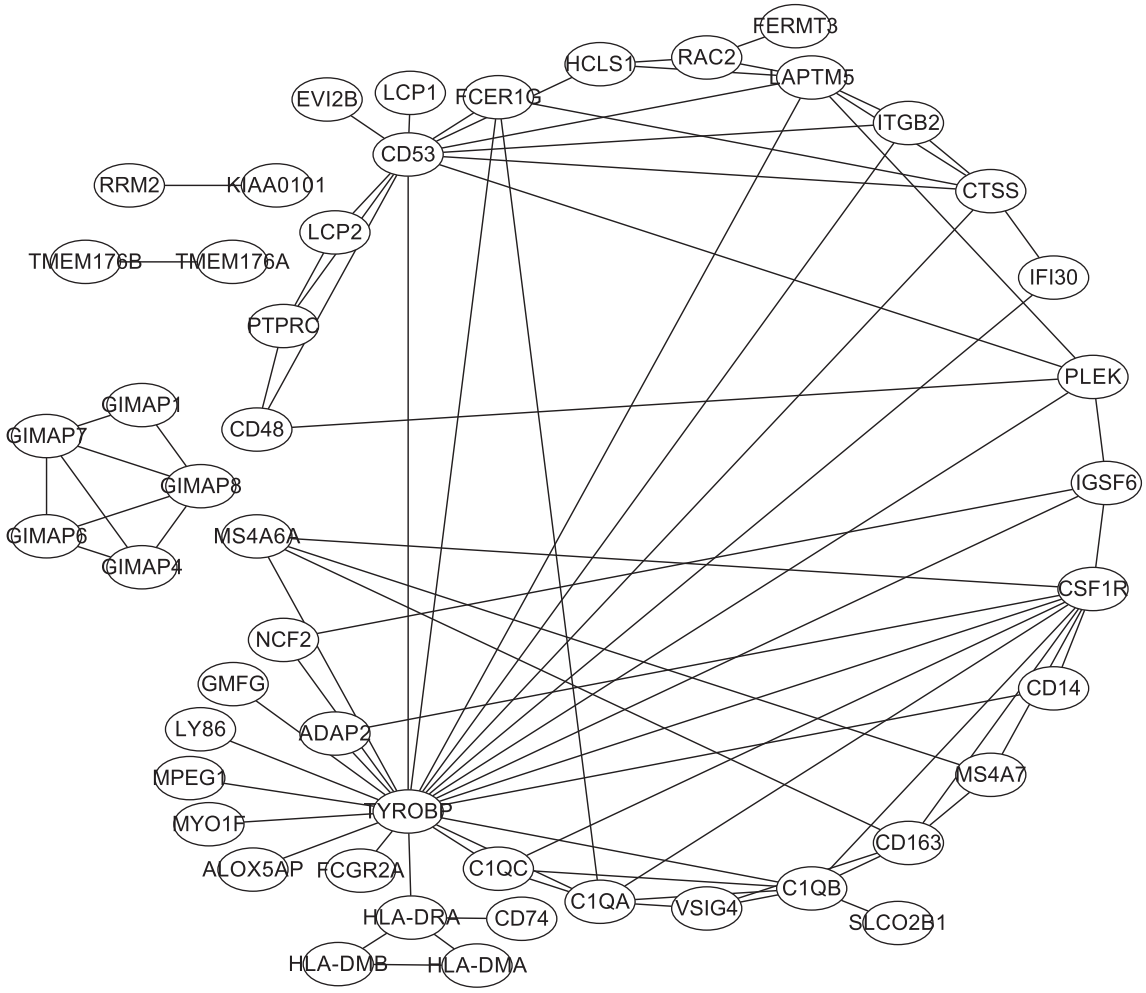


Fig. 3. The gene coexpression networks of the differentially expressed genes.

Table 1. Gene Ontology terms significantly enriched in the genes from the coexpression network

| Term | Count | FDR |
|---|---|---|
| GO:0006955~immune response | 18 | 7.11E-11 |
| GO:0002252~immune effector process | 10 | 3.28E-08 |
| GO:0006952~defense response | 15 | 7.60E-08 |
| GO:0002443~leukocyte mediated immunity | 8 | 1.82E-06 |
| GO:0050778~positive regulation of immune response | 9 | 2.36E-06 |
| GO:0016064~immunoglobulin mediated immune response | 7 | 4.60E-06 |
| GO:0002684~positive regulation of immune system process | 10 | 5.50E-06 |
| GO:0019724~B cell mediated immunity | 7 | 5.76E-06 |
| GO:0002449~lymphocyte mediated immunity | 7 | 2.27E-05 |
| GO:0002250~adaptive immune response | 7 | 4.06E-05 |

GO: Gene Ontology; FDR: false discovery rate.

patibility complex class II DM beta (HLA-DMB), HLA-DMA, CD74, and HLA class II histocompatibility antigen DR alpha chain (HLA-DRA). This pathway is also closely associated with the immune response. Both functional enrichment analysis and pathway enrichment analysis showed that the immune response plays an important role in the development of atherosclerotic plaque.

## DISCUSSION

In the present study, a total of 319 DEGs were identified in advanced carotid atherosclerotic plaque compared with early carotid atherosclerotic plaque, including 267 up-regulated genes and 52 down-regulated genes. The PCA result confirmed that these DEGs could well distinguish advanced atherosclerotic plaque from early atherosclerotic plaque, i.e., these genes exhibited differential expression during the progression of atherosclerotic plaque. Functional enrichment analysis revealed that the immune response was over-represented in the DEGs. Pathway enrichment analysis showed that antigen processing and presentation played an important role in the development of atherosclerotic plaque. This was consistent with previous findings that the immune response plays a critical role in the pathogenesis of atherosclerosis (Hansson and Libby, 2006; Hansson and Hermansson, 2011). Some DEGs have been implicated in atherosclerosis. CTSS is a lysosomal cysteine proteinase that may participate in the degradation of antigenic proteins to peptides for presentation on MHC class II molecules. CTSS has been implicated in the pathogenesis of atherosclerosis and it is overexpressed in atherosclerotic lesions (Sukhova et al., 1998). A study by Cerne et al. (2011) indicates that CTSS-mediated atherogenesis is associated with a CD40-mediated inflammatory and immune response. Inhibition of CTSS decreases atherosclerotic lesions in mice, suggesting a potential therapy for atherosclerotic plaque. CD74 participates in several key processes of the immune system, including antigen presentation, B-cell differentiation and inflammatory signaling (Borghese and Clanchy, 2011). Its level is increased in atherosclerotic plaque (Martín-Ventura et al., 2009), enhancing the inflammatory response and thus promoting the formation of atherosclerotic plaque. Therefore, it may be a novel therapeutic target. Complement C1q is involved in apoptotic cell removal (Ogden et al., 2001) and thus contributes to containing the size and complexity of early lesions in atherosclerosis (Bhatia et al., 2007). All three members of C1q (C1QA, C1QB and C1QC) showed differential expression between early and advanced atherosclerotic plaque. They may therefore be indicators, reflecting the size and complexity of the plaque. Moreover, CD14 polymorphisms have been associated with atherosclerosis (Vainas et al., 2006). CD163 is a plasma marker of coronary atherosclerosis, positively correlated with the extent of coronary atherosclerosis (Aristoteli et al., 2006). Moreno et al. (2009) reported that the CD163-TWEAK (tumor necrosis factor-like weak inducer of apoptosis) plasma level is a potential biomarker of clinical and subclinical atherosclerosis. Therefore, DEGs identified in the present study are worthy of future studies and may reveal more potential biomarkers.

A gene coexpression network was constructed for the DEGs, revealing TYROBP as the hub gene. TYROBP (also known as DAP12) is a transmembrane signaling polypeptide which contains an immunoreceptor tyrosine-based activation motif. It plays a role in the inflammatory response (Turnbull et al., 2005) and the immune response (Lanier, 2009), and can interact with a large family of receptors in hematopoietic cells. It is reported that TYROBP can either activate or inhibit immune responses via different signaling pathways (Lanier, 2009). A study by Bakker et al. (2000) showed that DAP12 signaling is required for optimal antigen-presenting cell function or inflammation. Further research is needed to disclose its role in plaque formation.

Overall, our study further demonstrated the critical role of the immune response in the progression of atherosclerotic plaque. Some DEGs are therapeutic targets, such as CTSS and CD74. Future research on these and other relevant genes (according to the gene coexpression network) should advance our understanding of the development of plaque, and subsequently benefit early diagnosis or treatment.

## REFERENCES


Abdi, H., and Williams, L. J. (2010) Principal component analysis. Wiley Interdisciplinary Reviews: Comp. Stat. **2**, 433–459.

Adler, D., and Murdoch, D. (2012) rgl: 3D visualization device system (OpenGL). R package version 0.92.879.

Aristoteli, L. P., Møller, H. J., Bailey, B., Moestrup, S. K., and Kritharides, L. (2006) The monocytic lineage specific soluble CD163 is a plasma marker of coronary atherosclerosis. Atherosclerosis **184**, 342–347.

Bakker, A. B., Hoek, R. M., Cerwenka, A., Blom, B., Lucian, L., McNeil, T., Murray, R., Phillips, J. H., Sedgwick, J. D., and Lanier, L. L. (2000) DAP12-deficient mice fail to develop autoimmunity due to impaired antigen priming. Immunity **13**, 345–353.

Benjamini, Y., and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. Series B (Methodological) **57**, 289–300.

Bhatia, V. K., Yun, S., Leung, V., Grimsditch, D. C., Benson, G. M., Botto, M. B., Boyle, J. J., and Haskard, D. O. (2007) Complement C1q reduces early atherosclerosis in low-density lipoprotein receptor-deficient mice. Am. J. Pathol. **170**, 416–426.

Bonomini, F., Tengattini, S., Fabiano, A., Bianchi, R., and Rezzani, R. (2008) Atherosclerosis and oxidative stress. Histol. Histopathol. **23**, 381–390.

Borghese, F., and Clanchy, F. I. (2011) CD74: an emerging opportunity as a therapeutic target in cancer and autoimmune disease. Expert Opin. Ther. Targets **15**, 237–251.

Cerne, D., Stern, I., Marc, J., Cerne, A., Zorman, D., Krzisnik-Zorman, S., and Kranjec, I. (2011) *CTSS* activation coexists with *CD40* activation in human atheroma: Evidence from plasma mRNA analysis. Clin. Biochem. **44**, 438–440.

Döring, Y., Manthey, H. D., Drechsler, M., Lievens, D., Megens, R. T. A., Soehnlein, O., Busch, M., Manca, M., Koenen, R. R., Zernecke, A., et al. (2012) Auto-antigenic protein-DNA complexes stimulate plasmacytoid dendritic cells to promote

atherosclerosis. Circulation **125**, 1673–1683.
Frieri, M. (2012) Accelerated atherosclerosis in systemic lupus erythematosus: role of proinflammatory cytokines and therapeutic approaches. Curr. Allergy Asthma Rep. **12**, 25–32.
Fujita, A., Sato, J. R., Rodrigues, L. O., Ferreira, C. E., and Sogayar, M. C. (2006) Evaluating different methods of microarray data normalization. BMC Bioinformatics **7**, 469.
Galkina, E., and Ley, K. (2009) Immune and inflammatory mechanisms of atherosclerosis. Annu. Rev. Immunol. **27**, 165–197.
Grainger, D. J. (2004) Transforming growth factor β and atherosclerosis: so far, so good for the protective cytokine hypothesis. Arterioscler. Thromb. Vasc. Biol. **24**, 399–404.
Hallerstam, S., Larsson, P. T., Zuber, E., and Rosfors, S. (2004) Carotid atherosclerosis is correlated with extent and severity of coronary artery disease evaluated by myocardial perfusion scintigraphy. Angiology **55**, 281–288.
Hansson, G. K., and Hermansson, A. (2011) The immune system in atherosclerosis. Nat. Immunol. **12**, 204–212.
Hansson, G. K., and Libby, P. (2006) The immune response in atherosclerosis: a double-edged sword. Nat. Rev. Immunol. **6**, 508–519.
Hansson, G. K., Libby, P., Schönbeck, U., and Yan, Z. Q. (2002) Innate and adaptive immunity in the pathogenesis of atherosclerosis. Circ. Res. **91**, 281–291.
Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2008) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. **4**, 44–57.
Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y., and Hattori, M. (2004) The KEGG resource for deciphering the genome. Nucleic Acids Res. **32**(suppl 1), D277–D280.
Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K. F., Itoh, M., Kawashima, S., Katayama, T., Araki, K., and Hirakawa, M. (2006) From genomics to chemical genomics: new developments in KEGG. Nucleic Acids Res. **34**(suppl 1), D354–D357.
Lanier, L. L. (2009) DAP10 - and DAP12 - associated receptors in innate immunity. Immunol. Rev. **227**, 150–160.
Li, G. W., Zheng, G. Y., Li, J. G., and Sun, X. D. (2010) Relationship between carotid atherosclerosis and cerebral infarction. Chin. Med. Sci. J. **25**, 32–37.
Mallat, Z., Gojova, A., Marchiol-Fournigault, C., Esposito, B., Kamaté, C., Merval, R., Fradelizi, D., and Tedgui, A. (2001) Inhibition of transforming growth factor-β signaling accelerates atherosclerosis and induces an unstable plaque phenotype in mice. Circ. Res. **89**, 930–934.
Martín-Ventura, J. L., Madrigal-Matute, J., Muñoz-Garcia, B., Blanco-Colio, L. M., Van Oostrom, M., Zalba, G., Fortuño, A., Gomez-Guerrero, C., Ortega, L., Ortiz, A., et al. (2009) Increased CD74 expression in human atherosclerotic plaques: contribution to inflammatory responses in vascular cells. Cardiovasc. Res. **83**, 586–594.
Moreno, J. A., Muñoz-García, B., Martín-Ventura, J. L., Madrigal-Matute, J., Orbe, J., Páramo, J. A., Ortega, L., Egido, G., and Blanco-Colio, L. M. (2009) The CD163-expressing macrophages recognize and internalize TWEAK: potential consequences in atherosclerosis. Atherosclerosis **207**, 103–110.
Obayashi, T., Hayashi, S., Shibaoka, M., Saeki, M., Ohta, H., and Kinoshita, K. (2008) COXPRESdb: a database of coexpressed gene networks in mammals. Nucleic Acids Res. **36**(suppl 1), D77–D82.
Ogden, C. A., Hoffmann, P. R., Bratton, D., Ghebrehiwet, B., Fadok, V. A., and Henson, P. M. (2001) C1q and mannose binding lectin engagement of cell surface calreticulin and Cd91 initiates macropinocytosis and uptake of apoptotic cells. J. Exp. Med. **194**, 781–796.
Pinderski, L. J., Fischbein, M. P., Subbanagounder, G., Fishbein, M. C., Kubo, N., Cheroutre, H., Curtiss L. K., Berliner, J. A., and Boisvert, W. A. (2002) Overexpression of interleukin-10 by activated T lymphocytes inhibits atherosclerosis in LDL receptor–deficient mice by altering lymphocyte and macrophage phenotypes. Circ. Res. **90**, 1064–1071.
Polak, J. F., Pencina, M. J., O'Leary, D. H., and D'Agostino, R. B. (2011) Common carotid artery intima-media thickness progression as a predictor of stroke in multi-ethnic study of atherosclerosis. Stroke **42**, 3017–3021.
Reiner-Benaim, A. (2007) FDR control by the BH procedure for two - sided correlated tests with implications to gene expression data analysis. Biom. J. **49**, 107–126.
Sheehan, A. L., Carrell, S., Johnson, B., Stanic, B., Banfi, B., and Miller, F. J. (2011) Role for Nox1 NADPH oxidase in atherosclerosis. Atherosclerosis **216**, 321–326.
Singh, U., and Jialal, I. (2006) Oxidative stress and atherosclerosis. Pathophysiology **13**, 129–142.
Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L., and Ideker, T. (2011) Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics **27**, 431–432.
Smyth, G. K. (2005) Limma: linear models for microarray data. In: Bioinformatics and Computational Biology Solutions Using R and Bioconductor (eds: R. Gentleman, V. Carey, W. Huber, R. Irizarry and S. Dudoit), pp. 397–420. Springer, New York.
Sorescu, D., Szöcs, K., and Griendling, K. K. (2001) NAD (P) H oxidases and their relevance to atherosclerosis. Trends Cardiovasc. Med. **11**, 124–131.
Sukhova, G. K., Shi, G. P., Simon, D. I., Chapman, H. A., and Libby, P. (1998) Expression of the elastolytic cathepsins S and K in human atheroma and regulation of their production in smooth muscle cells. J. Clin. Invest. **102**, 576–583.
Turnbull, I. R., McDunn, J. E., Takai, T., Townsend, R. R., Cobb, J. P., and Colonna, M. (2005) DAP12 (KARAP) amplifies inflammation and increases mortality from endotoxemia and septic peritonitis. J. Exp. Med. **202**, 363–369.
Vainas, T., Stassen, F. R., Bruggeman, C. A., Welten, R. J. Th. J., van den Akker, L. H., Kitslaar, P. J., Peña, A. S., and Morré, S. A. (2006) Synergistic effect of Toll-like receptor 4 and CD14 polymorphisms on the total atherosclerosis burden in patients with peripheral arterial disease. J. Vasc. Surg. **44**, 326–332.
Victor, V. M., Rocha, M., Sola, E., Banuls, C., Garcia-Malpartida, K., and Hernandez-Mijares, A. (2009) Oxidative stress, endothelial dysfunction and atherosclerosis. Curr. Pharm. Des. **15**, 2988–3002.