PRIMARY RESEARCH　　　　　　　　　　　　　　　　　　　　　**Open Access**

CrossMark

Systematic analysis of the molecular mechanism underlying atherosclerosis using a text mining approach

Dan Xi[1], Jinzhen Zhao[1], Wenyan Lai[2*] and Zhigang Guo[1*]

Abstract

Background: Atherosclerosis is one of the common health threats all over the world. It is a complex heritable disease that affects arterial blood vessels. Chronic inflammatory response plays an important role in atherogenesis. There has been little success in fully identifying functionally important genes in the pathogenesis of atherosclerosis.

Results: In the present study, we performed a systematic analysis of atherosclerosis-related genes using text mining. We identified a total of 1312 genes. Gene ontology (GO) analysis revealed that a total of 35 terms exhibited significance ($p < 0.05$) as overrepresented terms, indicating that atherosclerosis invokes many genes with a wide range of different functions. Pathway analysis demonstrated that the most highly enriched pathway is the Toll-like receptor signaling pathway. Finally, through gene network analysis, we prioritized 48 genes using the hub gene method.

Conclusions: Our study provides a valuable resource for the in-depth understanding of the mechanism underlying atherosclerosis.

Keywords: Atherosclerosis, Pathogenesis, Text mining

Background

Atherosclerosis is a complex heritable disease involving multiple cell types and the interactions of many different molecular pathways [1]. Atherosclerosis is therefore a syndrome affecting arterial blood vessels due to a chronic inflammatory response [2, 3]. Atherosclerosis is at the core of cardiovascular diseases, often leading to myocardial infarctions, stroke, and peripheral vascular diseases.

Recent genome-wide association studies (GWAS), involving hundreds of thousands of individuals, have identified numerous loci contributing to atherosclerotic traits and to risk factors such as blood lipoprotein levels and blood pressure [4]. Plasma lipids are primarily of importance for driving early atherosclerosis development,

consistent with the notion that loci identified by GWAS will be more useful for primary prevention and with the experimental finding that atherosclerosis regression in response to LDL lowering is much greater for early lesions than for mature and advanced lesions [5]. The extensive ongoing studies into the molecular mechanisms of the 153 confirmed GWA-defined CAD loci will shed light on this issue, as these mechanisms will likely be traceable to early versus late events in the pathogenesis of atherosclerosis [6]. Despite a large number of genes are identified, there has been little success in fully identifying functionally important genes in the pathogenesis of atherosclerosis.

Recently, the text mining methodology has been implemented, providing a necessary means to retrieve disease-related genes in an automated way [7]. Here, we reported on a systematic analysis of atherosclerosis-related genes using text mining. Our study provides in-depth insights into the molecular mechanisms underlying atherosclerosis.

* Correspondence: frame007@163.com; guozhigang126@126.com
[2]Laboratory of Department of Cardiology, Nanfang Hospital, Southern Medical University, Guangzhou 510515, Guangdong, People's Republic of China
[1]Division of Cardiology, Huiqiao Medical Center, Nanfang Hospital, Southern Medical University, 1838 North Guangzhou Avenue, Guangzhou 510515, Guangdong, People's Republic of China

Xi *et al. Human Genomics* (2016) 10:14

Page 2 of 8

Results

Identification of atherosclerosis-related genes by using text mining

We ran a key word search in the PubMed database for articles related to atherosclerosis and obtained 45,304 entries as a result (from January 1980 to April 2016). Abstracts of these articles were downloaded and processed through a text mining pipeline shown in Fig. 1a. Cumulative distribution analysis indicated that the number of articles published on atherosclerosis is growing linearly in recent years (Fig. 1b). From these articles, we extracted atherosclerosis-associated genes via text mining. We compiled a list of 1312 atherosclerosis-related genes (Fig. 1c; Additional file 1: Table S1).
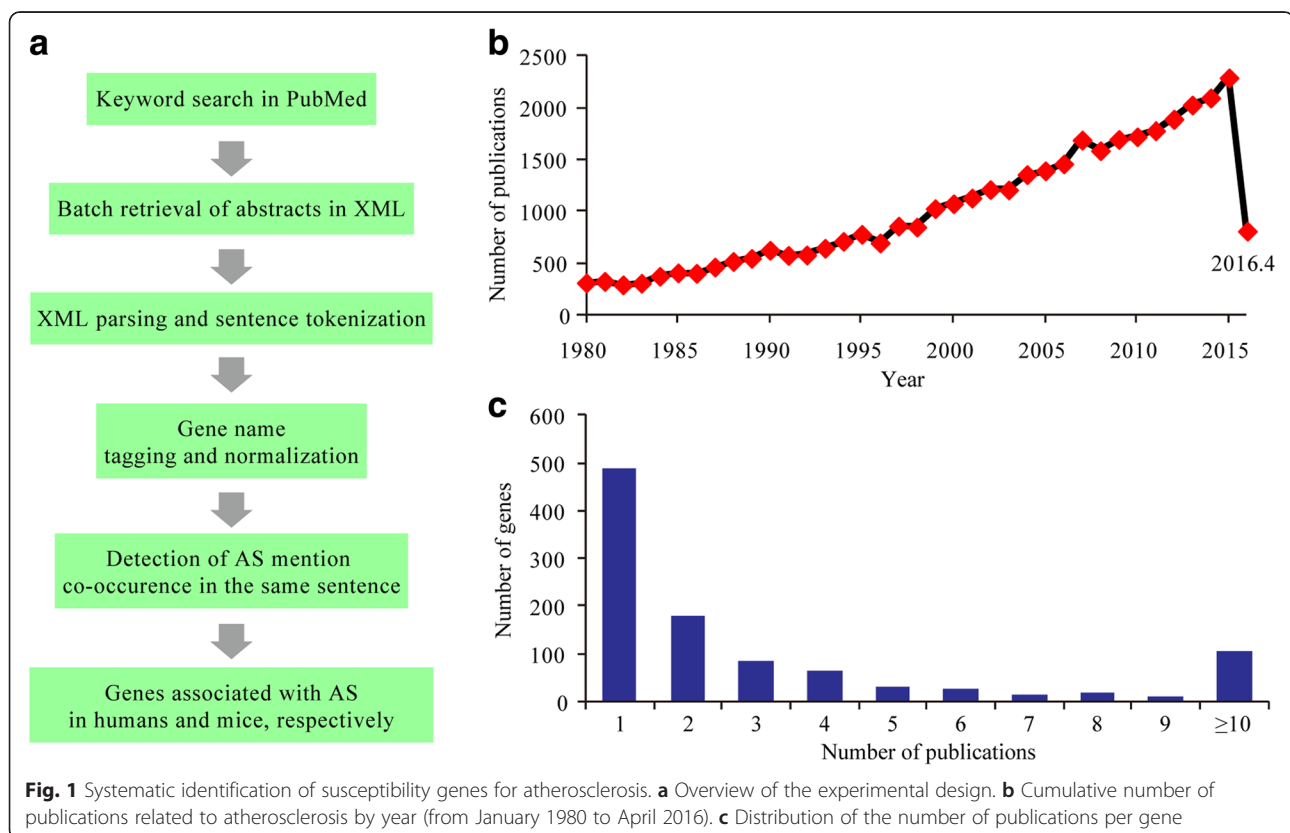
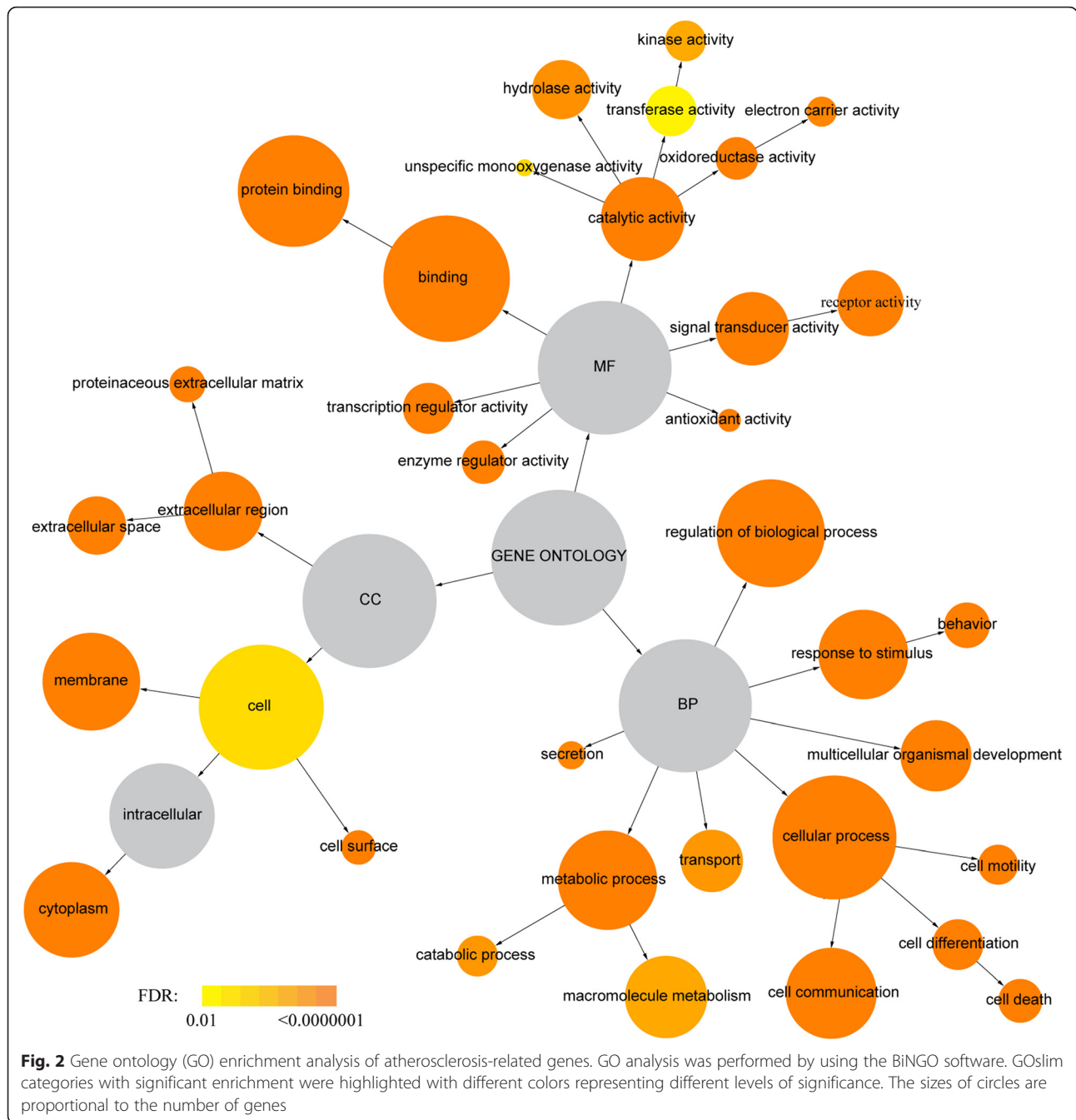### Functional clustering analysis

All 1312 unique genes were functionally categorized based on gene ontology (GO) annotation terms using the BiNGO program package. Enrichment analysis revealed that a total of 35 terms exhibited significance ($p < 0.05$) as overrepresented terms. In the biological process category, response to stimulus, cell communication, regulation of biological process, cellular process, behavior, multicellular organismal development, cell motility, cell death, metabolic process, cell differentiation, enzyme regulator activity, transcription regulator activity, electron carrier activity, secretion, catabolic process, transport, macromolecule metabolic process, and unspecific monooxygenase activity were found to be significantly enriched. GO terms related to extracellular region, extracellular space, cell surface, cytoplasm, membrane, proteinaceous extracellular matrix, and cell were overrepresented under the cellular component category. The overrepresented GO terms in the molecular function category were protein binding, binding, signal transducer activity, receptor activity, antioxidant activity, oxidoreductase activity, catalytic activity, hydrolase activity, kinase activity, and transferase activity (Additional file 2: Table S2). The hierarchical organization of these GO terms is shown in Fig. 2, together with the significance of enrichment indicated by different colors.

Pathway analysis

In addition to the GO analysis, we also performed pathway analysis by using the DAVID tools. Unlike GO, which only contains lists of functional gene groups, the pathway database also stores the information of gene dependencies in each pathway. In the present study, all atherosclerosis-related genes were linked to a total of 50 pathways. Among these pathways, 20 pathways, namely Toll-like receptor signaling pathway, complement and coagulation cascades, hematopoietic cell lineage, NOD-like receptor signaling pathway, adipocytokine signaling pathway, focal adhesion, Jak-STAT signaling pathway,


Fig. 1 Systematic identification of susceptibility genes for atherosclerosis. **a** Overview of the experimental design. **b** Cumulative number of publications related to atherosclerosis by year (from January 1980 to April 2016). **c** Distribution of the number of publications per gene

Xi *et al. Human Genomics* (2016) 10:14

Page 3 of 8

Fig. 2 Gene ontology (GO) enrichment analysis of atherosclerosis-related genes. GO analysis was performed by using the BiNGO software. GOslim categories with significant enrichment were highlighted with different colors representing different levels of significance. The sizes of circles are proportional to the number of genes

apoptosis, T cell receptor signaling pathway, neurotrophin signaling pathway, Fc epsilon RI signaling pathway, PPAR signaling pathway, VEGF signaling pathway, B cell receptor signaling pathway, renin-angiotensin system, leukocyte transendothelial migration, ErbB signaling pathway, TGF-beta signaling pathway, MAPK signaling pathway, and natural killer cell mediated cytotoxicity were significantly enriched ($p < 0.05$) (Fig. 3a; Additional file 3: Table S3). Based on enrichment p value, the most highly

overrepresented pathway went to the Toll-like receptor signaling pathway (Fig. 3b). The Toll-like receptor signaling pathway is known to play an important role during atherosclerosis in both immune and inflammatory response.

Network analysis

In the present study, a genome-wide protein-protein interaction (PPI) network was constructed by merging up-to-date protein-protein interactions available in

Xi *et al. Human Genomics* (2016) 10:14

Page 4 of 8
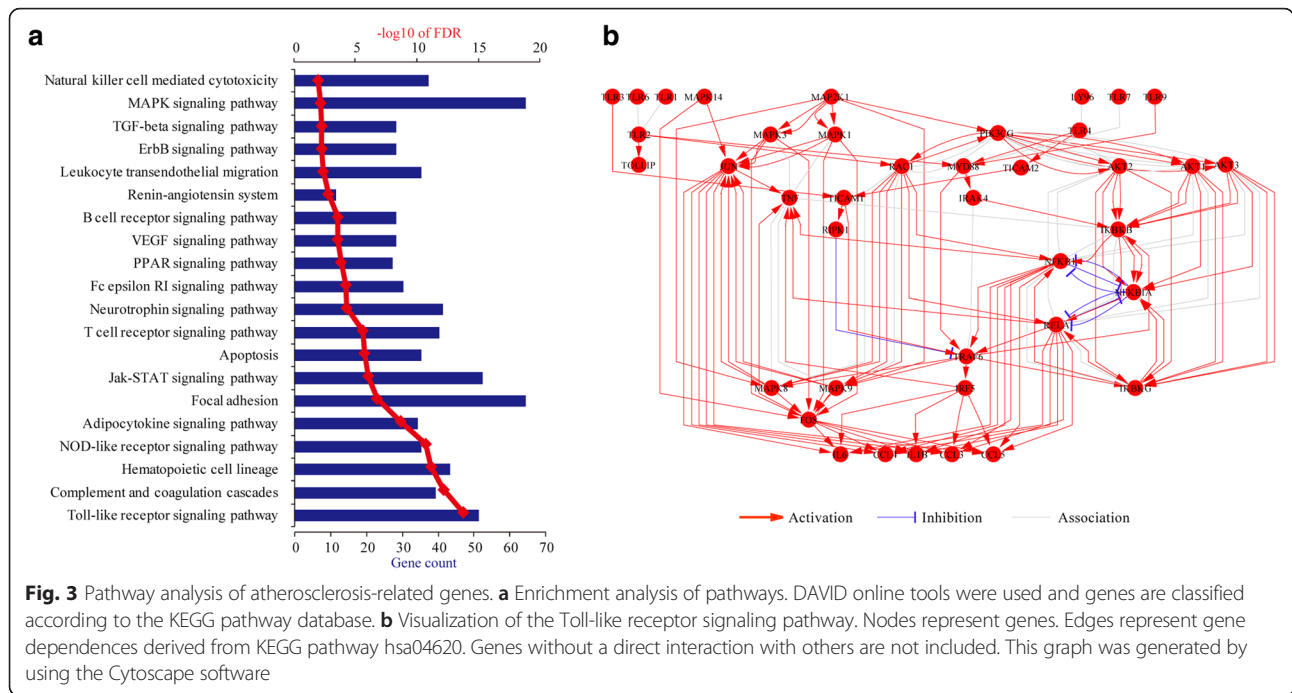
Fig. 3 Pathway analysis of atherosclerosis-related genes. **a** Enrichment analysis of pathways. DAVID online tools were used and genes are classified according to the KEGG pathway database. **b** Visualization of the Toll-like receptor signaling pathway. Nodes represent genes. Edges represent gene dependences derived from KEGG pathway hsa04620. Genes without a direct interaction with others are not included. This graph was generated by using the Cytoscape software

IntAct [8], BioGRID [9], MINT [10], DIP [11], HPRD [12, 13], and MIPS [13]. The network related to atherosclerosis was generated by mapping the atherosclerosis-related genes to the genome-wide PPI network. The atherosclerosis network consisted of 1079 nodes connected via 6089 edges (Fig. 4a). Topological analysis showed that the network follows a power-law distribution (Fig. 4b) and therefore is a scale-free, small-world network [14]. This type of networks has the particular feature that some nodes are highly connected compared with others within the network. These highly connected nodes, also known as hub genes, represent important genes in the network and therefore are treated with special attention. Using a defined cut-off value, we identified 48 hub genes. These hub genes and their connections were extracted from the whole network and rendered as a simplified sub-network (Fig. 4c).
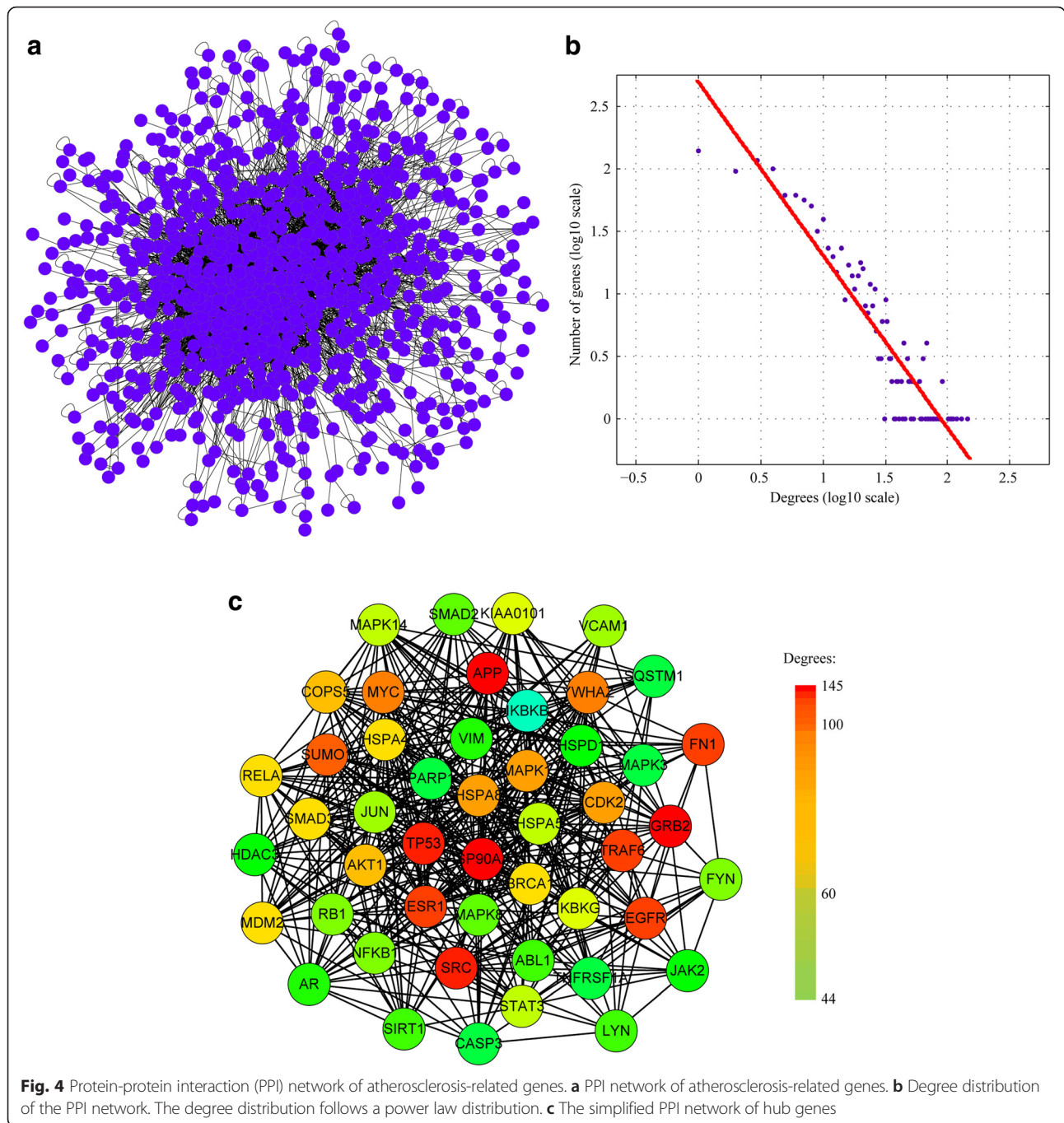
Discussion

In the present study, we attempted to compile a complete list of genes involved in atherosclerosis. In recent years, high-throughput transcriptomic and proteomic approaches make it possible for studying the expression levels of thousands of genes and proteins simultaneously. However, these data suffer from high technical variability and high dimension size [15, 16]. On the contrary, there is a large body of research using conventional gene-by-gene methods. Text mining provides the necessary means to retrieve these data through automated processing of texts [7]. Here, we performed a text mining analysis of atherosclerosis-associated genes. We identified

1312 genes from 45,304 publications. Considering the large body of literature we analyzed, our result may have reasonably good coverage of all atherosclerosis-associated genes.

We found that 1312 genes were associated with atherosclerosis. Based on GO analysis, 35 GO terms were significantly enriched. Additionally, our study also revealed 20 enriched pathways. Based on enrichment *p* value, the most highly overrepresented pathway went to the Toll-like receptor (TLR) signaling pathway. The Toll-like receptor signaling pathway is known to play an important role during atherosclerosis in both immune and inflammatory response. The disruptions of cellular or organismal cholesterol homeostasis that occur as a risk factor of atherosclerosis may lead to an augmentation of inflammatory responses via enhanced TLR signaling or inflammasome activation [17]. TLR activation leads to the expression of pro-inflammatory cytokines and also induces the expression of many negative regulators, acting to limit signal transduction, messenger RNA (mRNA) transcription, or translation [18].

A genome-wide gene network was constructed by using up-to-date interaction data available in the PINA2 database [19]. We obtained a gene network consisting of 1079 nodes connected via 6089 edges. So far, several studies have been conducted to incorporate the topology of gene network in prioritization of disease candidate genes [20–22]. The main concern for these studies is that the incompleteness and noisiness of interaction data may affect the accuracy of prioritization result. By merging up-to-date protein-protein interactions available in

Xi *et al. Human Genomics* (2016) 10:14

Page 5 of 8

Fig. 4 Protein-protein interaction (PPI) network of atherosclerosis-related genes. **a** PPI network of atherosclerosis-related genes. **b** Degree distribution of the PPI network. The degree distribution follows a power law distribution. **c** The simplified PPI network of hub genes

IntAct [8], BioGRID [9], MINT [10], DIP [11], HPRD [12], and MIPS [13], the PINA2 database provides a comprehensive gene network at genome-wide scale. We expected that the use of the PINA2 database may alleviate this problem to a certain extent. Using a defined threshold value for degree, we identified a total of 48 hub genes in this network.

The top 20 hub genes are the following: APP (amyloid beta A4 precursor protein), HSP90AA1 (heat shock

protein 90 kDa alpha class A member 1), GRB2 (growth factor receptor-bound protein 2), SRC (v-src sarcoma viral oncogene homolog), TP53 (tumor protein p53), ESR1 (estrogen receptor 1), FN1 (fibronectin 1), TRAF6 (TNF receptor-associated factor 6), EGFR (epidermal growth factor receptor), SUMO1 (SMT3 suppressor of mif two 3 homolog 1), YWHAZ (14-3-3 zeta), MYC (v-myc myelocytomatosis viral oncogene homolog), CDK2 (cyclin-dependent kinase 2), HSPA8 (heat shock 70-kDa

Xi *et al. Human Genomics* (2016) 10:14

Page 6 of 8

protein 8), MAPK1 (mitogen-activated protein kinase 1), AKT1 (v-akt murine thymoma viral oncogene homolog 1), COPS5 (COP9 constitutive photomorphogenic homolog subunit 5), MDM2 (Mdm2 p53 binding protein homolog), RELA (v-rel reticuloendotheliosis viral oncogene homolog A, NFKB3, p65), and HSPA4 (heat shock 70-kDa protein 4). APP is present in advanced human carotid plaques, in proximity to activated macrophages and platelets [23], and lack of APP attenuates atherogenesis and leads to plaque stability [24]. HSP90 is a candidate autoantigen, target of cellular and humoral immune reactions in patients with carotid atherosclerosis [25]. HSP90 expression is associated with features of plaque instability in advanced human lesions [26]. GRB2 is required for atherosclerotic lesion formation and uptake of oxidized LDL by macrophages [27]. In endothelial cell, SRC contributes to atherosclerotic lesion development by disrupting adherence junction integrity and promoting monocyte transmigration [28]. Increasing P53 activity protects against atherosclerosis by causing proliferation arrest of lesional macrophages [29]. The product of the MDM2 gene is a nuclear protein which forms a complex with P53, thereby inhibiting the negative regulatory effects of wild-type P53 on cell cycle progression. P53 and MDM2 are expressed in human atherosclerotic lesions; P53 and MDM2 may therefore play an important role in regulating cellularity and inflammatory activity in human atherosclerotic plaques [30, 31]. SUMOylation of P53 by SUMO1 contributes to the atherosclerotic plaque formation [32]. ESR1 is expressed in macrophages and other immune cells known to exert dramatic effects on glucose homeostasis. A study suggests that diminished ESR1 expression in hematopoietic/myeloid cells promotes aspects of the metabolic syndrome and accelerates atherosclerosis in female mice [33]. FN is one of the earliest extracellular matrix (ECM) proteins deposited at atherosclerosis-prone sites and was suggested to promote atherosclerotic lesion formation [34]. TRAF6 is expressed in atherosclerotic aortic tissue of low-density lipoprotein-null mice [35]. Endothelial-specific TRAF6 deficiency in females was associated with diminished atherosclerosis and decreased plaque macrophage burden [36]. EGFR mRNA was detected in atherosclerotic plaques but not in morphologically normal aortae and EGFR receptor staining co-localized with macrophage staining in these plaques [37]. Secreted from activated platelets, YWHAZ is present at the atherosclerotic plaques [38]. In cholesterol-fed roosters, MYC was seen in lipid-rich thickened intimal lesions of the entire aorta [39]. CDK2 negatively regulates neointimal thickening in animal models of restenosis and atherosclerosis, and its expression in human neointimal lesions is consistent with a protective role [40]. HSP70 (HSPA4 and HSPA8) is present in human atherosclerotic lesions [41]. Treatment with platelet-

derived growth factor which caused vascular smooth muscle cell migration in an MAPK1 activation-dependent manner suggests a role for MAPK1 in the pathogenesis and/or progression of atherosclerosis [42]. AKT1 expression in vascular smooth muscle cells influences early and late stages of atherosclerosis. The absence of AKT1 in VSMCs induces features of plaque vulnerability including fibrous cap thinning and extensive necrotic core areas [43]. Macrophage migration inhibitory factor (MIF), a cytokine with potent inflammatory functions, was thus considered to be important in atherosclerotic lesion evolution. COPS5 is able to form complexes with MIF and serves critical regulatory functions in atherosclerotic lesion evolution [44]. The transcription factor NF-κB p65 is a key regulator in the regulation of an inflammatory response and in the pathology of atherosclerosis [45].

A limitation for text mining-based strategies is that there is no chance to discover new genes. In order to solve this problem, we enlarged the network by inducing new genes that are not reported to be involved in atherosclerosis. According to the rule of "guilty-by-association," these new genes may be potential susceptibility genes. Finally, we made a list of 50 new genes, all of which have more than 44 connections with known genes (Additional file 4: Table S4).

Conclusions

In summary, we have reported here the first systematic analysis of the molecular mechanism underlying atherosclerosis using a text mining approach. Our study provides a valuable resource for the in-depth understanding of the mechanism underlying atherosclerosis.

Methods

Identification of atherosclerosis-related genes by using text mining

The PubMed database was used as a source of literature for text mining. We conducted a search with the following combinations of query key words: "atherosclerosis" OR "atherogenesis" OR "atheroma" OR "atherosclerotic." The search tag "[Title/Abstract]" was added after each keyword. The relevant articles were retrieved in XML format, which makes information extraction more precise due to the presence of content enclosed within XML tag pairs. For each article, titles and abstract texts were fetched using the dom4j XML parser class in JAVA. Abstract texts were further divided into sentences through sentence tokenizer implemented in LingPipe (Alias-I, Inc.). Text mining was performed at sentence level.

Gene mention recognition was performed using two different gene mention taggers, the hidden Markov model (HMM) tagger implemented in LingPipe and the ABNER tagger based on a machine learning system of conditional random fields (CRF) [46]. Gene mentions

Xi *et al. Human Genomics* (2016) 10:14

Page 7 of 8

from both taggers were merged. Because researchers mention genes in a highly variable manner, we built a gene synonym dictionary from entrez gene database [47]. The dictionary was used for the gene name normalization process during which gene mentions were linked to entrez genes using exact string match. If multiple entrez genes share the same gene mention, the ambiguity was resolved manually. In order to minimize the false positive rate, we required the co-occurrence of atherosclerosis mention and gene mention within a single sentence. Finally, we compiled a complete list of atherosclerosis-related genes.

Enrichment test of gene ontology (GO) terms

GO enrichment analysis was performed by using BiNGO 2.3 with GOslim dataset [48]. To test for enrichment, a hypergeometric test was conducted followed by Benjamini and Hochberg multiple test correction. The adjusted p value <0.01 was used as significance threshold to identify enriched categories.

Pathway analysis

To rank overall importance of pathways involved in atherosclerosis, we calculated Fisher's exact test p values and Benjamini-Hochberg adjusted p values through the DAVID bioinformatics resource 6.7 [49]. The significance threshold was set at 0.01. After enrichment tests, gene sets were collected for each pathway. Gene dependencies in a certain pathway were determined using the R package KEGGSOAP and visualized in Cytoscape [50, 51].

Construction of protein-protein interaction (PPI) network

The genes associated with atherosclerosis were cross-referenced with the PINA2 database to create the PPI network [19]. The PINA2 database provides integrated and up-to-date protein-protein interactions available in IntAct, BioGRID, MINT, DIP, HPRD, and MIPS, which simplifies the task of inter-database mapping [8–13]. To query the PINA2 database, interaction was restricted to human and mouse and all kinds of experimental procedures were included. Cytoscape software was applied for visualization and analysis of PPI network. The topological parameters of PPI network were analyzed by NetworkAnalyzer [52]. The edges in the network were treated as undirected. The degree of a node was the number of its directly connecting neighbors in the network. The threshold degree value for hub genes was the mean plus two standard deviations. As a result, genes with a degree value of larger than 26 were considered hub genes. Hub genes and their connections were extracted from the whole network and rendered as a simplified sub-network.

Additional files

Additional file 1: Table S1. A complete list of genes identified by text mining. (XLSX 148 kb)

Additional file 2: Table S2. Gene ontology analysis. (XLSX 37 kb)

Additional file 3: Table S3. Pathway enrichment analysis. (XLSX 14 kb)

Additional file 4: Table S4. Potential susceptibility genes for atherosclerosis. (XLSX 12 kb)

Authors' contributions
GZG conceived and designed the study. LWY and ZJZ collected the data. XD completed the study and preformed the statistical tests. XD wrote the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

References
1. Stylianou IM, Bauer RC, Reilly MP, Rader DJ. Genetic basis of atherosclerosis: insights from mice and humans. Circ Res. 2012;110:337–55.
2. Glass CK, Witztum JL. Atherosclerosis. The road ahead. Cell. 2001;104:503–16.
3. Andersson J, Libby P, Hansson GK. Adaptive immunity and atherosclerosis. Clin Immunol. 2010;134:33–46.
4. Bennett BJ, Davis RC, Civelek M, Orozco L, Wu J, Qi H, et al. Genetic architecture of atherosclerosis in mice: a systems genetics analysis of common inbred strains. PLoS Genet. 2015;11, e1005711.
5. Bjorkegren JL, Hagg S, Talukdar HA, Foroughi Asl H, Jain RK, Cedergren C, et al. Plasma cholesterol-induced lesion networks activated before regression of early, mature, and advanced atherosclerosis. PLoS Genet. 2014;10, e1004201.
6. Bjorkegren JL, Kovacic JC, Dudley JT, Schadt EE. Genome-wide significant loci: how important are they? Systems genetics to understand heritability of coronary artery disease and other common complex disorders. J Am Coll Cardiol. 2015;65:830–45.
7. Rebholz-Schuhmann D, Oellrich A, Hoehndorf R. Text-mining solutions for biomedical research: enabling integrative biology. Nat Rev Genet. 2012;13:829–39.
8. Aranda B, Achuthan P, Alam-Faruque Y, Armean I, Bridge A, Derow C, et al. The IntAct molecular interaction database in 2010. Nucleic Acids Res. 2010;38:D525–31.
9. Stark C, Breitkreutz BJ, Chatr-Aryamontri A, Boucher L, Oughtred R, Livstone MS, et al. The BioGRID interaction database: 2011 update. Nucleic Acids Res. 2011;39:D698–704.
10. Ceol A, Chatr Aryamontri A, Licata L, Peluso D, Briganti L, Perfetto L, et al. MINT, the molecular interaction database: 2009 update. Nucleic Acids Res. 2010;38:D532–9.
11. Salwinski L, Miller CS, Smith AJ, Pettit FK, Bowie JU, Eisenberg D. The database of interacting proteins: 2004 update. Nucleic Acids Res. 2004;32:D449–51.
12. Keshava Prasad TS, Goel R, Kandasamy K, Keerthikumar S, Kumar S, Mathivanan S, et al. Human protein reference database—2009 update. Nucleic Acids Res. 2009;37:D767–72.
13. Mewes HW, Dietmann S, Frishman D, Gregory R, Mannhaupt G, Mayer KF, et al. MIPS: analysis and annotation of genome information in 2007. Nucleic Acids Res. 2008;36:D196–201.
14. Barabasi AL, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet. 2004;5:101–13.
15. Frantz S. An array of problems. Nat Rev Drug Discov. 2005;4:362–3.

Xi et al. Human Genomics (2016) 10:14

Page 8 of 8

16. Chandramouli K, Qian PY. Proteomics: challenges, techniques and possibilities to overcome biological sample complexity. Hum Genomics Proteomics. 2009;2009.

17. Tall AR, Yvan-Charvet L. Cholesterol, inflammation and innate immunity. Nat Rev Immunol. 2015;15:104–16.

18. Medzhitov R, Horng T. Transcriptional control of the inflammatory response. Nat Rev Immunol. 2009;9:692–703.

19. Cowley MJ, Pinese M, Kassahn KS, Waddell N, Pearson JV, Grimmond SM, et al. PINA v2.0: mining interactome modules. Nucleic Acids Res. 2012;40:D862–5.

20. Chen J, Aronow BJ, Jegga AG. Disease candidate gene identification and prioritization using protein interaction networks. BMC Bioinformatics. 2009;10:73.

21. Guo H, Dong J, Hu S, Cai X, Tang G, Dou J, et al. Biased random walk model for the prioritization of drug resistance associated proteins. Sci Rep. 2015;5:10857.

22. Morrison JL, Breitling R, Higham DJ, Gilbert DR. GeneRank: using search engine technology for the analysis of microarray experiments. BMC Bioinformatics. 2005;6:233.

23. De Meyer GR, De Cleen DM, Cooper S, Knaapen MW, Jans DM, Martinet W, et al. Platelet phagocytosis and processing of beta-amyloid precursor protein as a mechanism of macrophage activation in atherosclerosis. Circ Res. 2002;90:1197–204.

24. Van De Parre TJ, Guns PJ, Fransen P, Martinet W, Bult H, Herman AG, et al. Attenuated atherogenesis in apolipoprotein E-deficient mice lacking amyloid precursor protein. Atherosclerosis. 2011;216:54–8.

25. Businaro R, Profumo E, Tagliani A, Buttari B, Leone S, D'Amati G, et al. Heat-shock protein 90: a novel autoantigen in human carotid atherosclerosis. Atherosclerosis. 2009;207:74–83.

26. Madrigal-Matute J, Lopez-Franco O, Blanco-Colio LM, Munoz-Garcia B, Ramos-Mozo P, Ortega L, et al. Heat shock protein 90 inhibitors attenuate inflammatory responses in atherosclerosis. Cardiovasc Res. 2010;86:330–7.

27. Proctor BM, Ren J, Chen Z, Schneider JG, Coleman T, Lupu TS, et al. Grb2 is required for atherosclerotic lesion formation. Arterioscler Thromb Vasc Biol. 2007;27:1361–7.

28. Sun C, Wu MH, Lee ES, Yuan SY. A disintegrin and metalloproteinase 15 contributes to atherosclerosis by mediating endothelial barrier dysfunction via Src family kinase activity. Arterioscler Thromb Vasc Biol. 2012;32:2444–51.

29. Sayin VI, Khan OM, Pehlivanoglu LE, Staffas A, Ibrahim MX, Asplund A, et al. Loss of one copy of Zfp148 reduces lesional macrophage proliferation and atherosclerosis in mice by activating p53. Circ Res. 2014;115:781–9.

30. Ihling C, Haendeler J, Menzel G, Hess RD, Fraedrich G, Schaefer HE, et al. Co-expression of p53 and MDM2 in human atherosclerosis: implications for the regulation of cellularity of atherosclerotic lesions. J Pathol. 1998;185:303–12.

31. Barillari G, Iovane A, Bonuglia M, Albonici L, Garofano P, Di Campli E, et al. Fibroblast growth factor-2 transiently activates the p53 oncosuppressor protein in human primary vascular smooth muscle cells: implications for atherogenesis. Atherosclerosis. 2010;210:400–6.

32. Heo KS, Chang E, Le NT, Cushman H, Yeh ET, Fujiwara K, et al. De-SUMOylation enzyme of sentrin/SUMO-specific protease 2 regulates disturbed flow-induced SUMOylation of ERK5 and p53 that leads to endothelial dysfunction and atherosclerosis. Circ Res. 2013;112:911–23.

33. Ribas V, Drew BG, Le JA, Soleymani T, Daraei P, Sitz D, et al. Myeloid-specific estrogen receptor alpha deficiency impairs metabolic homeostasis and accelerates atherosclerotic lesion development. Proc Natl Acad Sci U S A. 2011;108:16457–62.

34. Rohwedder I, Montanez E, Beckmann K, Bengtsson E, Duner P, Nilsson J, et al. Plasma fibronectin deficiency impedes atherosclerosis progression and fibrous cap formation. EMBO Mol Med. 2012;4:564–76.

35. Zirlik A, Bavendiek U, Libby P, MacFarlane L, Gerdes N, Jagielska J, et al. TRAF-1, -2, -3, -5, and -6 are induced in atherosclerotic plaques and differentially mediate proinflammatory functions of CD40L in endothelial cells. Arterioscler Thromb Vasc Biol. 2007;27:1101–7.

36. Polykratis A, van Loo G, Xanthoulea S, Hellmich M, Pasparakis M. Conditional targeting of tumor necrosis factor receptor-associated factor 6 reveals opposing functions of Toll-like receptor signaling in endothelial and myeloid cells in a mouse model of atherosclerosis. Circulation. 2012;126:1739–51.

37. Lamb DJ, Modjtahedi H, Plant NJ, Ferns GA. EGF mediates monocyte chemotaxis and macrophage proliferation and EGF receptor is expressed in atherosclerotic plaques. Atherosclerosis. 2004;176:21–6.

38. Hernandez-Ruiz L, Valverde F, Jimenez-Nunez MD, Ocana E, Saez-Benito A, Rodriguez-Martorell J, et al. Organellar proteomics of human platelet dense granules reveals that 14-3-3zeta is a granule protein related to atherosclerosis. J Proteome Res. 2007;6:4449–57.

39. Toda T, Tamamoto T, Shimajiri S, Sadi AM, Nakashima Y, Takei H. Expression of PDGF and C-myc in atherosclerotic lesions in cholesterol-fed chicken. Immunohistochemical and in situ hybridization study. Ann N Y Acad Sci. 1995;748:514–6.

40. Sanz-Gonzalez SM, Melero-Fernandez de Mera R, Malek NP, Andres V. Atheroma development in apolipoprotein E-null mice is not regulated by phosphorylation of p27(Kip1) on threonine 187. J Cell Biochem. 2006;97:735–43.

41. Johnson AD, Berberian PA, Tytell M, Bond MG. Atherosclerosis alters the localization of HSP70 in human and macaque aortas. Exp Mol Pathol. 1993;58:155–68.

42. Yoshizumi M, Kyotani Y, Zhao J, Nagayama K, Ito S, Tsuji Y, et al. Role of big mitogen-activated protein kinase 1 (BMK1)/extracellular signal-regulated kinase 5 (ERK5) in the pathogenesis and progression of atherosclerosis. J Pharmacol Sci. 2012;120:259–63.

43. Rotllan N, Wanschel AC, Fernandez-Hernando A, Salerno AG, Offermanns S, Sessa WC, et al. Genetic evidence supports a major role for Akt1 in VSMCs during atherogenesis. Circ Res. 2015;116:1744–52.

44. Burger-Kentischer A, Goebel H, Seiler R, Fraedrich G, Schaefer HE, Dimmeler S, et al. Expression of macrophage migration inhibitory factor in different stages of human atherosclerosis. Circulation. 2002;105:1561–6.

45. Ye X, Jiang X, Guo W, Clark K, Gao Z. Overexpression of NF-kappaB p65 in macrophages ameliorates atherosclerosis in apoE-knockout mice. Am J Physiol Endocrinol Metab. 2013;305:E1375–83.

46. Settles B. ABNER: an open source tool for automatically tagging genes, proteins and other entity names in text. Bioinformatics. 2005;21:3191–2.

47. Brown GR, Hem V, Katz KS, Ovetsky M, Wallin C, Ermolaeva O, et al. Gene: a gene-centered information resource at NCBI. Nucleic Acids Res. 2015;43:D36–42.

48. Maere S, Heymans K, Kuiper M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics. 2005;21:3448–9.

49. Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, et al. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res. 2007;35:W169–75.

50. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. 2004;5:R80.

51. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498–504.

52. Assenov Y, Ramirez F, Schelhorn SE, Lengauer T, Albrecht M. Computing topological parameters of biological networks. Bioinformatics. 2008;24:282–4.