



## OPEN Transcriptome-wide analysis reveals the coregulation of RNA-binding proteins and alternative splicing genes in the development of atherosclerosis

Runqing Wang<sup>1,2,3</sup>, Jin Xu<sup>1,2,3</sup>, Yuning Tang<sup>1,2,3,4</sup>, Yongxiang Wang<sup>1,2,3,5</sup>, Jing Zhao<sup>1,2,3,5</sup>, Liqiong Ding<sup>2,3,5</sup>, Yu Peng<sup>2,3,5</sup> & Zheng Zhang<sup>1,2,3,5</sup>✉

RNA-binding proteins (RBPs) are involved in the regulation of RNA splicing, stability, and localization. How RBPs control the development of atherosclerosis, is not fully understood. To explore the relevant RNA-binding proteins (RBPs) and alternative splicing events (ASEs) in atherosclerosis. We made a comprehensive work to integrate analyses of differentially expressed genes, including differential RBPs, and variable splicing characteristics related to different stages of atherosclerosis in dataset GSE104140. A total of 3712 differentially expressed genes (DEGs) were identified, including 2921 upregulated genes and 791 downregulated genes. Further analysis screened out 54 RBP genes, and 434 AS genes overlapped DEGs. We selected high expression ten RBP genes (SAMHD1, DDX60 L, TLR7, RBM47, MYEF2, RNASE6, PARP12, APOBEC3G, SMAD9, and RNASE1) for co-expression analysis. Meanwhile, we found seven regulated alternative splicing genes (RASGs) (ABI1, FXR1, CHID1, PLEC, PRKACB, BNIP2, PPP3CB) that could be regulated by RBPs. The co-expression network was used to further elucidate the regulatory and interaction relationship between RBPs and AS genes. Apoptotic process and innate immune response, revealed by the functional enrichment analysis of RASGs regulated by RBPs were closely related to atherosclerosis. In addition, 26 of the 344 alternative splicing genes regulated by the above 10 RBPs were transcription factors (TFs), We selected high expression nine TFs (TFDP1, RBBP7, STAT2, CREB5, ERG, ELF1, HMG3, BCLAF1, and ZEB2) for co-expression analysis. The target genes of these TFs were mainly enriched in inflammatory and immune response pathways that were associated with atherosclerosis. indicating that AS abnormalities of these TFs may have a function in atherosclerosis. Furthermore, the expression of differentially expressed RBPs and the alternative splicing events of AS genes was validated by qRT-PCR in umbilical vein endothelial cells (HUVEC). The results showed that RBM47 were remarkably difference in HUVEC treated with ox-LDL and the splicing ratio of AS in BCLAF1 which is regulated by RBM47 significantly changed. In conclusion, the differentially expressed RBPs identified in our analysis may play important roles in the development of atherosclerosis by regulating the AS of these TF genes.

Atherosclerosis is the accumulation of lipids in the intima of arteries and the formation of the core of atherosclerotic plaques. As the disease progresses, atherosclerotic plaques develop fibrous caps and gradually accumulate calcium minerals<sup>1</sup>. Late unstable atherosclerotic plaques rupture and cause thrombi, occluding blood vessels and

<sup>1</sup>The First School of Clinical Medicine, Lanzhou University, Lanzhou, Gansu, China. <sup>2</sup>Gansu Key Laboratory of Cardiovascular Diseases, The First Hospital of Lanzhou University, Lanzhou, Gansu, China. <sup>3</sup>Gansu Clinical Medical Research Center for Cardiovascular Diseases, The First Hospital of Lanzhou University, Lanzhou, Gansu, China. <sup>4</sup>Department of Cardiology, Lanzhou University Second Hospital, Lanzhou, China. <sup>5</sup>Heart Center, The First Hospital of Lanzhou University, Lanzhou, Gansu, China. ✉email: zhangccu@163.com

causing acute ischemia of organs, while stable atherosclerotic plaques invade the lumen volume and cause chronic tissue ischemia<sup>2</sup>. Atherosclerosis can lead to a variety of cardiovascular diseases, such as myocardial infarction, ischemic cardiomyopathy, stroke, and peripheral arterial disease, of which heart disease and stroke are the two leading causes of death worldwide<sup>3</sup>. In 2016, more than 17 million people died of cardiovascular disease, accounting for 31% of global deaths<sup>4</sup>. At present, the treatment of atherosclerosis focuses on lipid-lowering, anti-platelet and late-stage vascular recanalization, and there are few measures related to how to prevent plaque progression<sup>1</sup>. Therefore, understanding the key physiological mechanisms of atherosclerosis progression is of great significance for preventing patients from developing corresponding clinical symptoms and improving patient prognosis.

Approximately 95% of human multiexon genes generate different transcripts through alternative splicing (AS), which leads to diversity in genetic information transmission<sup>5</sup>. AS processes play important roles in tissue recognition and many key biological processes<sup>6</sup>. To date, functional roles mediated by AS have been discovered in the fields of cardiac development, myocardial infarction, and structural cardiac diseases<sup>7</sup>. Through these studies, research has shown that AS plays an important role in atherosclerosis. For example, AS of FOXP3 can activate regulatory T cells (Treg cells) and inhibit the progression of atherosclerotic disease<sup>8</sup>. In ApoE (−/−) mice fed a high-fat diet, SRPK1-mediated splicing of VEGF-A to proangiogenic VEGF165 was found to contribute to the development of AS<sup>9</sup>. The binding of RNA-binding proteins (RBPs) through specific target sites can drive AS, but the changes in RBPs and AS that mediate atherosclerosis are currently unclear.

We explored RBPs and AS events (ASEs) in atherosclerosis using bioinformatics and expression analysis methods based on previously reported gene expression data of human fibroatheroma samples (GSE104140 dataset). The data were divided into the very early stage of vascular disease (diffuse intimal thickening) and two advanced stages of atherosclerosis (calcified and noncalcified fibroatheroma formation). In the first step, we identified differentially expressed genes (DEGs) in the dataset and then annotated and analyzed their functions. Moreover, we examined differentially prevalent ASEs and differentially expressed RBPs and found changes in AS at different stages and their potential regulatory functions in atherosclerosis, laying a foundation for the in-depth study of the molecular mechanism underlying the occurrence and development of atherosclerosis.

## Materials and methods

**Retrieving and processing data.** We used “atherosclerosis” as the key word to retrieve and select appropriate datasets from the Gene Expression Omnibus (GEO) database and ultimately identified the GSE104140 dataset for use in further analyses. The GSE104140 dataset contains transcriptomic high-throughput sequencing data retrieved using a GPL16791 platform (Illumina HiSeq 2500 Homo sapiens). Total RNA was extracted from sets of thawed plaque cryosections. The expression profiles were converted and standardized into log<sub>2</sub> data to create a series matrix file. GSE104140 is a transcriptome sequencing dataset of human fibroatheromas, which contains 32 samples; The 32 samples were all from carotid endarterectomy plaque sections with various pathological types. 4 samples with inconsistent pathological types were eliminated, including intimal xanthoma, pathologic intimal thickening (with superficial macrophages), and pathologic intimal thickening (macrophage poor). And the remaining 28 samples were allocated into 3 groups: diffuse intimal thickening (DIT) includes 8 biological replicates, calcified fibroatheromas include 12 biological replicates, and noncalcified fibroatheromas include 8 biological replicates.

**Read alignment and DEG analysis.** We utilized TopHat<sup>10</sup> to align the reads to the human genome (GRch38) by setting the mismatch parameter to 4. Each gene was evaluated based on the number of reads and fragments per kilobase of exon per million mapped fragments (FPKM) using unique mapped reads. We utilized DESeq2<sup>11</sup> to analyze all the DEGs in the dataset. We identified differentially expressed genes on the basis of fold change ( $FC \geq 2$  or  $\leq 0.5$ ) and false discovery rate ( $FDR \leq 0.05$ ) criteria.

**Differential AS analysis.** Alternative splicing events (ASEs) and regulated ASEs (RASEs) between different groups were predicted and quantified based on previously reported methods<sup>12,13</sup>. Based on splicing junction reads, we finally detected ten types of ASEs, including alternative 5' splice site (A5SS), alternative 3' splice site (A3SS), cassette exon, exon skipping (ES), intron retention (IR), mutually exclusive exons (MXE), mutually exclusive 5' untranslated regions (UTRs; 5pMXE), mutually exclusive 3'UTRs (3pMXE), A3SS&ES and A5SS&ES data. We calculated the ratio of alternatively spliced to constitutively spliced reads as the RASE ratio between compared samples. We set a *p* value  $\leq 0.05$  for RASE discovery. Student's *t* test was performed to assess an altered ASE ratio in a repetition comparison. The differences in ASEs at the *p*-value cutoff of 0.05 were considered RASEs.

**Functional enrichment analysis.** To distinguish enriched functional pathways of DEGs, we performed a Gene Ontology functional enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis. Functional enrichment statistics of selected gene sets were predicted using GO and KEGG pathway annotation with the KOBAS 2.0 server<sup>14</sup>. For significant pathways, we used a hypergeometric test and Benjamini–Hochberg FDR control to determine the statistical enrichment of each pathway.

**Cell culture and treatment.** We purchased human umbilical vein cells (HUVECs) from the American Type Culture Collection (ATCC; Manassas, VA, USA) and cultured them in endothelial cell medium (ScienCell, California, USA) with 5% (v/v) fetal bovine serum (FBS) (ScienCell, California, USA), 100 U/ml penicillin, 100 mg/ml streptomycin (ScienCell, California, USA), and 1% ECGF (ScienCell, California, USA). The cells were incubated in a 5% CO<sub>2</sub> incubator at 37 °C. To mimic atherosclerosis in vitro, HUVECs were stimulated with 100 µg/ml ox-LDL (UnionBiol, Beijing, China) for 24 h.

**Quantitative real-time polymerase chain reaction (qRT-PCR).** Total RNA was extracted using an M5 HiPer Universal RNA Mini kit (Mei5bio, Beijing, China). The cDNA was reverse transcribed from total RNA by an M5 Super Plus qPCR RT kit with gDNA remover (Mei5bio, Beijing, China), and the reverse transcription reactions were performed on a Quant Studio™ Dx instrument (Thermo Fisher Scientific, USA) using Hieff™ qPCR SYBR® Green (Yeasen, Shanghai, China). The qPCR reactions were run under the following conditions: 5 min at 95 °C for 1 cycle, 15 s at 95 °C and 30 s at 60 °C for 40 cycles, finally, 15 s at 95 °C, 1 min at 60 °C and 15 s at 95 °C for 1 cycle. All groups underwent three technical replicates and three biological replicates. The relative expression levels of target genes were obtained by the  $2^{-\Delta\Delta C_t}$  method, with glyceraldehyde-3-phosphate dehydrogenase (GAPDH) as an internal reference. The primer sequences are listed in Table 1.

**Other statistical analyses.** Principal component analysis (PCA) was performed using the R package *factoextra* (<https://cloud.r-project.org/package=factoextra>) to display the clustering pattern of samples based on the top two components. The sequencing data and genomic annotations were visualized with a script we developed in-house (*sogen*) by normalizing the reads by the tags per million (TPM) of each gene. The *heatmap* package in R was used for clustering based on Euclidean distance. Student's *t* tests were performed to determine the significance of differences between groups.

## Results

**Identification of differentially expressed genes (DEGs) at different stages of atherosclerosis as determined through functional analysis.** IN this study, the RNA-seq data of 32 carotid endarterectomy plaque samples from GSE104140 were downloaded from the GEO database. The data were divided into the very early stage of vascular disease (diffuse intimal thickening) and two advanced stages of atherosclerosis (calcified and noncalcified fibroatheroma formation). We first analyzed samples in each of three different disease stages, and the results showed that the calcified and noncalcified samples contained very few DEGs (Supplementary Fig. 1). Therefore, we combined the calcified and noncalcified samples into one group (SAMP\_advanced, advanced stage), and this combined group and the diffuse intimal thickening samples (SAMP\_DIT, early stage) were analyzed. The sample correlation analysis also showed that there were significant differences in gene expression between SAMP\_DIT and SAMP\_advanced. However, the difference between the calcified group and the noncalcified group was smaller (Fig. 1A). A total of 3712 DEGs between the early stage and the advanced stage of atherosclerosis were identified, including 2921 genes with upregulated expression and 791 genes with downregulated expression (Fig. 1B). The PCA with 18 SAMP\_advanced and 7 SAMP\_DIT samples after normalizing all known DEG expression levels is shown in Fig. 1C. The expression level of all DEGs clearly distinguished the early and advanced atherosclerosis groups, and the samples in the two groups were clustered.

According to GO term enrichment analysis, 2921 upregulated DEGs were significantly enriched in immune-related biological processes, such as the innate immune response, immune response, B-cell receptor signaling pathway and T-cell costimulation pathways. Atherosclerosis is mainly driven by the innate immune response involving myeloid cells, followed by adaptive immunity involving T cells and B cells to modulate the inflammatory state within the plaque, which is the same as the results<sup>15</sup>. Additionally, upregulated DEGs were enriched in “inflammatory response, signal transduction, cell surface receptor signaling pathway, blood agglutination” (Fig. 1D), while “synaptic transmission, ion transmembrane transport, muscle contraction, vasoconstriction regulation” was significantly enriched in 791 downregulated DEGs (Fig. 1E). We also conducted a KEGG pathway analysis, and the five pathways most enriched with DEGs with upregulated expression were cytokine–cytokine receptor interaction, cell adhesion molecules (CAMs) and chemokine signaling pathway, while “cAMP signaling pathway, adrenergic signaling in cardiomyocytes, insulin secretion, long-term potentiation and calcium signaling pathway” were most enriched with DEGs with downregulated expression (Supplementary Fig. 2).

## Regulated alternative splicing events (RASEs) and genes expressed at different stages of atherosclerosis.

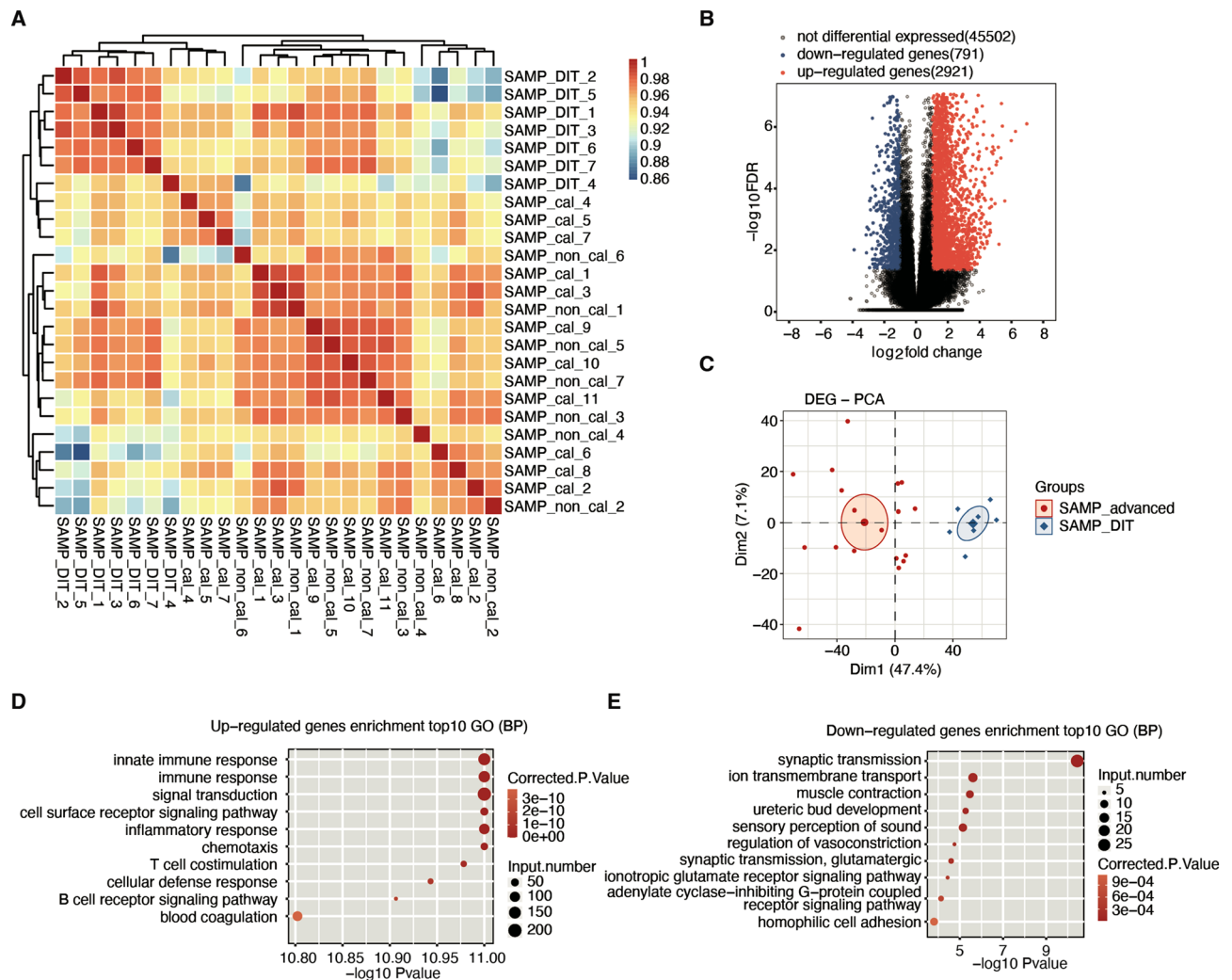
We analyzed all RASE events in all samples (Supplementary Fig. 3). Among 5468 RASEs found between the early stage and advanced stage of atherosclerosis, A5SS (alternative 5' splice site) and A3SS (alternative 3' splice site) were the most frequently reported (Fig. 2A). A PCA of the 18 advanced samples and 7 DIT samples based on the percent spliced in (PSI) value of all differential nonintron retention (NIR) events shows that the DIT group can be separated from the advanced group (Fig. 2B). A PSI heatmap showing all NIR-RAS events among the SAMP\_advanced and SAMP\_DIT samples showed obvious differences (Fig. 2C). Compared with the DIT, GO analysis demonstrated that all regulated alternative splicing genes (RASGs) were mainly enriched in atherosclerosis-related pathways like small GTPase-mediated signal transduction, signal transduction, actin cytoskeleton organization, negative regulation of the I- $\kappa$ B kinase/NF- $\kappa$ B cascade, regulation of cell shape, transcription regulation and other functional pathways (Fig. 2D). Small GTPases (ras) activate a variety of downstream signaling pathways, including phosphatidylinositol 3-kinase (PI3K) and Rac and Rho proteins, associated with the regulation of the cytoskeleton. Through RAS, other signals may be activated such as p38 MAPK, stress-activated protein kinase pathway, and c-Jun N-terminal [JNK] pathway<sup>16</sup>. MAPK, JNK pathways and cell migration caused by cytoskeleton changes is closely related to plaque stability<sup>17–19</sup>. In addition, 434 of the 3698 DEGs were involved in nonintron-retained alternative splicing events, suggesting that alternative splicing of genes may play a role in the progression of atherosclerosis. (Fig. 2E).

## Differential expression of RNA-binding proteins (RBPs) and regulated alternative splicing genes (RASGs).

RBPs play important roles in RASEs. Therefore, we explored the relationship between regulated alternative splicing genes (RASGs) and RBPs. We identified 54 differentially expressed RBPs between the two atherosclerotic sample groups (Fig. 3A). Then, a heatmap of these RBPs whose expression levels was greater

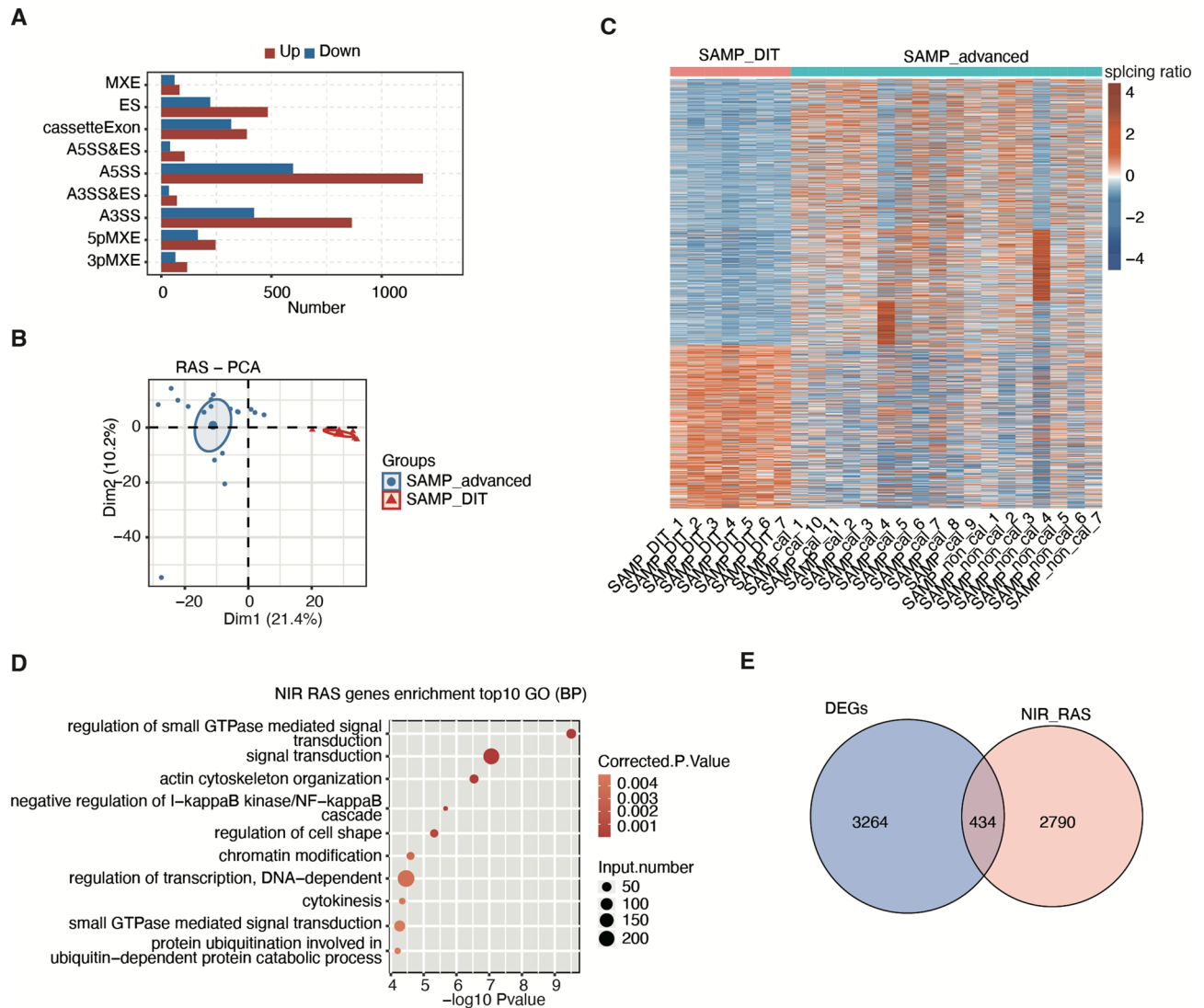
Gene name	Primer sequence (5' to 3')
GAPDH-F	GGTCGGAGTCAACGGATTG
GAPDH-R	GGAAGATGGTGATGGGATTTC
THEMIS2-F	CTGTCTTCCTCATCTCTGTAG
THEMIS2-R	AAGTCCTAAGGTTGTGTTTC
TLR1-F	CCTCTAACACTTCACTTGATAC
TLR1-R	TTAAGTACCTTCCTGGATG
RBM47-F	ACACAGTGGGAATGACAG
RBM47-R	ACAGTGGAGCACATGATC
DDX60L -F	AGTGGTAGTGAGCAGAGAT
DDX60L -R	TGGTATTAGGAGATGATAGGTC
PARP12-F	CTTGGAGGAGGTGGTGTA
PARP12-R	GAGCAACGCCTTCTATGA
IL18-F	GCTAGAAAGTATCCTTCGTATG
IL18-R	GAATCCTCCTGATAACATCAAG
SMAD9-F	TACACATCTGGTCAGTTCA
SMAD9-R	ACTATCAACACGGCTTCC
CIITA -F	CGGAATGAACCACATCTTG
CIITA -R	CGCAATGTCCTTCAGAGA
SAMHD1-F	AGGCTGGTCTCAACTCTT
SAMHD1-R	TTGTATCATTTGTGCTGTCTG
ITGB2-F	TAACCTCACCAACCTCAAG
ITGB2-R	AGGATGTCACCAATTAACCA
TLR7-F	GGCTCTGATTCTCCTGTAAT
TLR7-R	AACTCCTGACCTCGTGAT
PRKACB-M-F	GTGGAGAGCGTGAAAGAGTT
PRKACB-AS-F	GACAGATCAATGAAAGAGTT
PRKACB-M/AS-R	TGAGTTGGATTCTCCCATTT
ABII-M-F	TGAGCCCGGGGCTGGTCCAAT
ABII-AS-F	GCAATAGAACTGGTCCAAT
ABII-M/AS-R	AAGTAGTGGAGGAAGTGGAAG
PPP3CB -M-F	GCAGCTGAACCATCAAACCTG
PPP3CB -AS-F	GCTGAACCTACATCAAACCTG
PPP3CB -M/AS-R	TCTGAGTATTTGCTCTGATG
CHID1-M/AS-F	CTTCTCCAGCAGCGTCTTTG
CHID1-AS-R	AGCCTGGCCAGCCCTCTGGAC
CHID1-M-R	AACAACCAGGTCCTCTGGAC
ELF1-M/AS-F	CAGTTTCCCTGGTTCATTGA
ELF1-AS-R	ATACTTGAAGCCAAGAAGC
ELF1-M-R	TGCTTGTCAGGCCAAGAAGC
FXR1-M/AS-F	CCAGCGAATCTCATCACAGT
FXR1-AS-R	CATCTTTTGCTAGCCCATT
FXR1-M-R	GCAACTGTGACTAGCCCATT
BCLAF1-M-F	TTTATGTTTACTGTCATCTT
BCLAF1-AS-F	TACATTTTGTGTCATCTT
BCLAF1-M/AS-R	AAGTAAAGAACGGGAGATT
PLEC -M/AS-F	TGTCGCATCCACTGAAGCAG
PLEC -AS-R	TGCAGGATGGGAGCTGCAGCT
PLEC -M-R	TGAGGGCCAACGAGCTGCAGCT
BNIP2-M-F	CTTGATCAACTTGTTTTATGC
BNIP2-AS-F	CTTCTTCATACTGTTTATGC
BNIP2-M/AS-R	CGTGTTTAATTTGGCAGAAC
ERG -M/AS-F	CTTTAGTTGCCCTTGGTTCT
ERG -AS-R	TGTGCATGGGGTTATCCAGG
ERG -M-R	CGGCGCTCAGGTTATCCAGG

**Table 1.** The sequences of specific primer.



**Figure 1.** Transcriptome analysis of DEGs in atherosclerotic plaques in the early disease stage (SAMP\_DIT) and advanced disease stages (SAMP\_advanced). (A) Sample correlation between SAMP\_DIT and SAMP\_advanced (SAMP\_cal and SAMP\_non\_cal) samples. The heatmap scale represent correlation coefficient between different samples. (B) Volcano plot displaying DEGs in each group compared to the SAMP\_DIT samples. Red indicates genes with upregulated expression ( $FC \geq 2$  and  $FDR \leq 0.05$ ). Blue indicates genes with downregulated expression ( $FC \leq 0.5$  and  $FDR \leq 0.05$ ). (C) Principal component analysis (PCA) of 18 SAMP\_advanced and 7 SAMP\_DIT samples after normalizing all known DEG expression levels. The samples were grouped by left and right atrial appendages and atherosclerotic stage, and the ellipse for each group is a confidence ellipse. (D, E) GO analysis of DE mRNAs, divided into genes with up- (D) and down-regulated (E) expression.

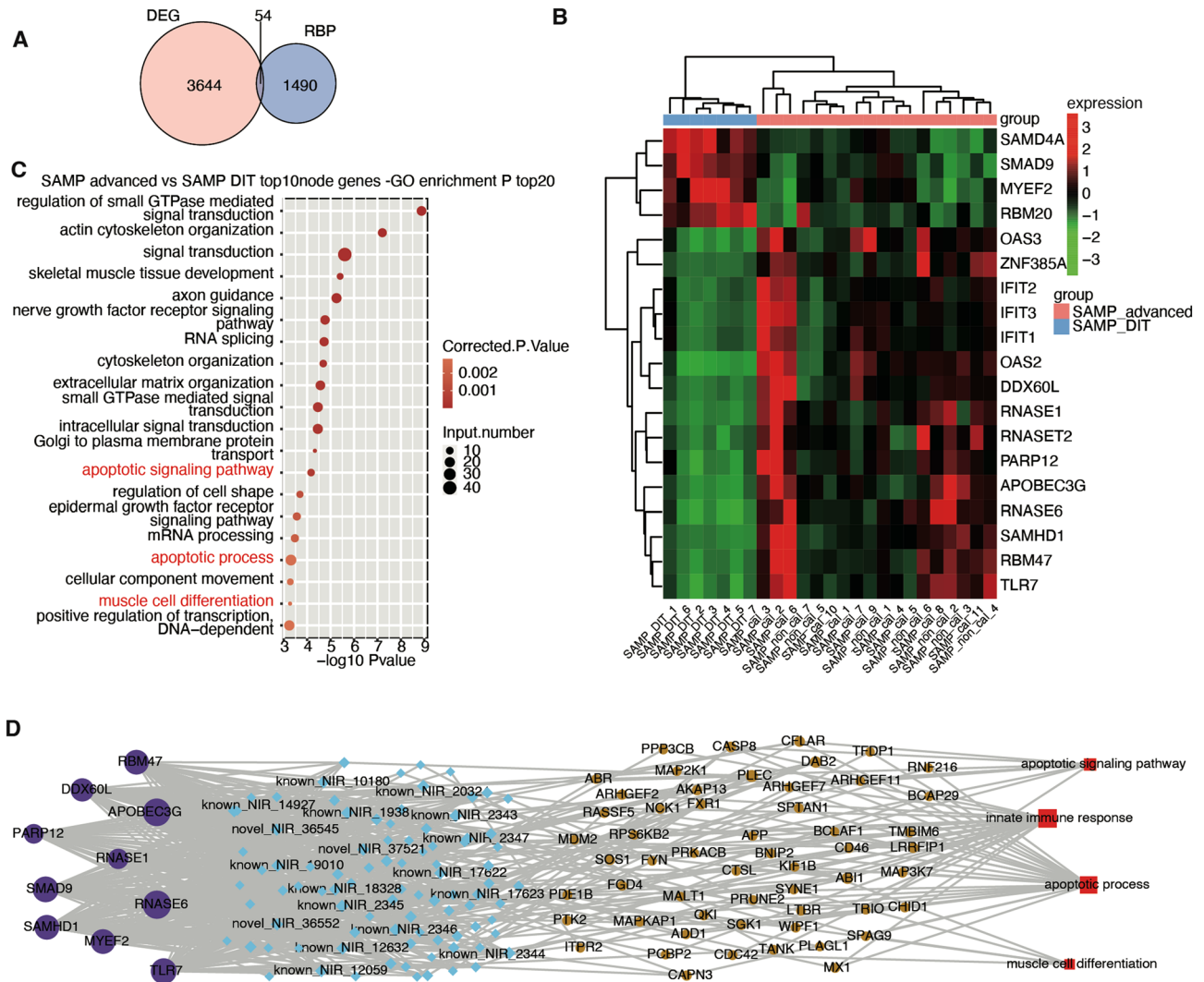
than 1 in 80% of the samples showed differential expression for the DIT and advanced groups (Fig. 3B). To further demonstrate the regulatory relationship of RBPs on RASGs. We established the regulatory relationship between RBPs and RASGs through correlation analysis and speculated that there may be a potential regulatory relationship. The GO analysis of the highest correlation values ten RASGs that co-regulated with these RBPs was enriched in the regulation of small GTP-mediated signal transduction, RNA splicing, apoptotic signaling pathways, Golgi-to-plasma membrane protein transport, intracellular signal transduction, apoptosis, and muscle cell differentiation categories (Fig. 3C). The KEGG analysis indicated that RBP-regulated ASGs between the two stages of atherosclerosis were enriched in focal adhesion, regulation of actin cytoskeleton, vascular smooth muscle contraction, viral carcinogenesis, mTOR signaling pathway, T-cell receptor signaling pathway, and VEGF signaling pathway (Supplementary Fig. 4). The destruction of focal adhesion signaling leads to the loss of adhesion and apoptosis of endothelial cells<sup>20</sup>. Endothelial cells stimulated by long-term lipids and inflammation secrete cytokines such as VEGF and ICAM<sup>15</sup>, which make smooth muscle migrate actively to the intima of blood vessels, all of which affect the stability of plaque fibrous cap<sup>21</sup>. We also selected high expression and consistency ten RBP genes (TLR7, RBM47, SAMHD1, DDX60L, PARP12, SMAD9, APOBEC3G, RNASE1, RNASE6, MYEF2) for another coexpression analysis to further elucidate the regulatory and interaction relationship between these RBPs and AS. (Fig. 3D). We noted that 55 variable splicing genes including ERG, ELF1, BCLAF1, ABI1, FXR1, CHID1, PLEC, PRKACB, BNIP2, and PPP3CB occurred in co-variation with



**Figure 2.** AS deregulation in atherosclerotic plaques in the early disease stage (SAMP\_DIT) and advanced disease stages (SAMP\_advanced). **(A)** Classification of all RAS events except IR events. X-axis: RAS event number. **(B)** Principal component analysis (PCA) of 18 SAMP\_advanced and SAMP\_DIT samples based on the percent spliced in (PSI) value of all differential nonintron retention (NIR) events. The samples were grouped by left and right atrial appendages and atherosclerotic stage, and the ellipse for each group is the confidence ellipse. **(C)** PSI heatmap showing all NIR-regulated alternative splicing events (RASEs) among SAMP\_advanced and SAMP\_DIT samples. NIR RASEs were clustered on the basis of K means. AS filtering was performed to detect splice junctions. AS is indicated if 80% or more of the samples contain 10-splice junction reads. **(D)** A GO analysis of RASGs in the SAMP\_advanced compared with the SAMP\_DIT. **(E)** Venn diagram showing the RASGs and differentially expressed genes (DEGs),  $p$ -value = 1.

these ten RBPs, suggesting a potential regulatory relationship between them. In addition, we validated the genes with the highest rate of variable splicing in vitro (Supplementary Fig. 7). The apoptotic signaling pathway, innate immune response, apoptotic process and muscle cell differentiation were revealed by GO analysis of RASGs regulated by RBPs (Fig. 3D). In addition, the coregulation of the AS network demonstrated the regulatory relationship between the 10 most highly expressed RBPs and RASEs (Fig. 3D).

**RNA-binding proteins (RBPs) regulate the alternative splicing (AS) of transcription factor (TF).** TFs precisely regulate the transcription process of genes, and many studies have shown that abnormalities in transcription factors are closely related to cardiovascular diseases<sup>22,23</sup>. Therefore, we next explored whether transcription factors were involved in the abovementioned RBP-related alternative splicing events. Among the 344 alternatively spliced genes regulated by the above 10 RBPs, 26 are transcription factors (Fig. 4A). In addition, 1475 target genes of these 26 TFs were differentially expressed (Fig. 4B). We performed a GO analysis with these differentially expressed target genes. The enriched functional pathways were found to be signal transduction, inflammatory response, immune response, leukocyte migration, innate immunity, T-cell activation, coagulation and angiogenesis (Fig. 4C). The KEGG pathway enrichment analysis revealed that crossover



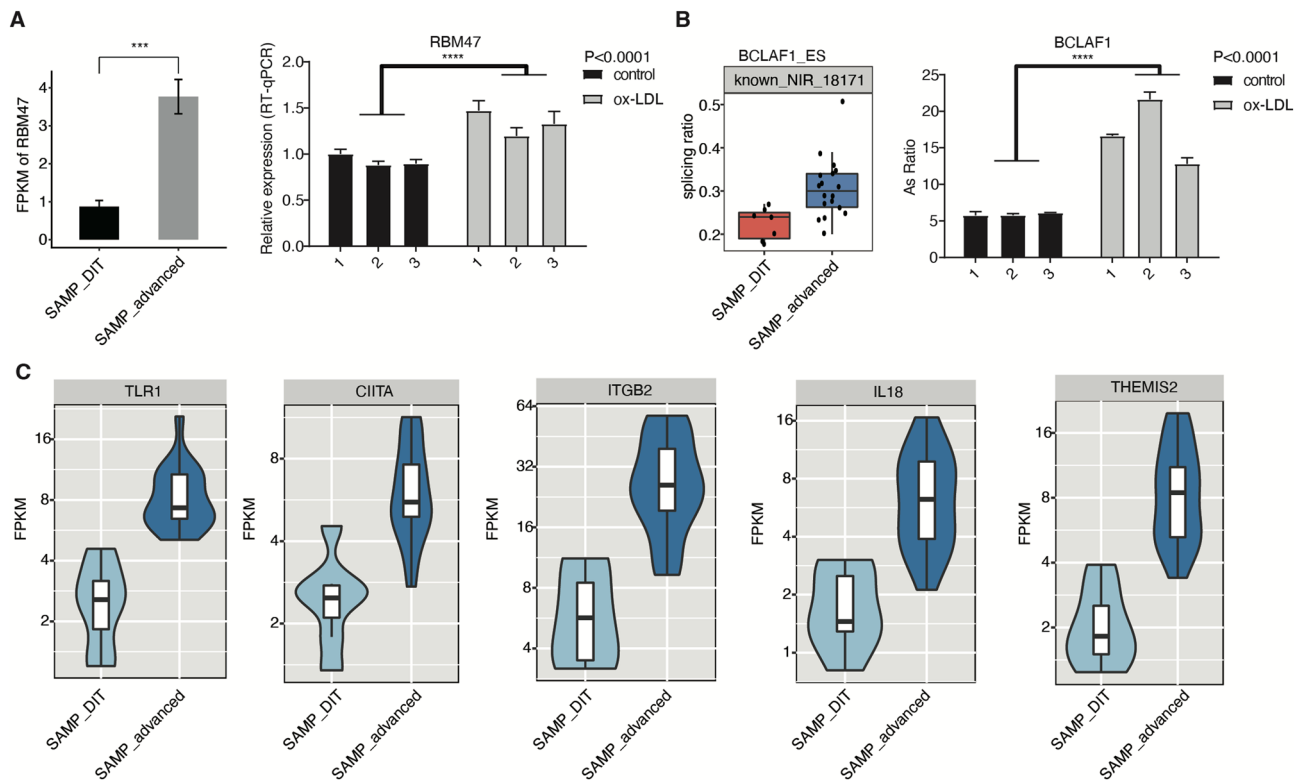
**Figure 3.** Correlation analysis of RBPs and RASEs in atherosclerotic plaques in the early disease stage (SAMP\_DIT) and advanced disease stages (SAMP\_advanced). (A) Venn diagram depicting regulated differentially expressed genes (DEGs) and RNA binding proteins (RBPs),  $p$ -value = 1. (B) heatmap showing differentially expressed RBPs with a FPKM of greater than 1 in at least 80% of samples. Cluster analysis based on the expression of differential RBP genes between groups. (C) The 20 most enriched GO terms in RASGs with expression disrupted by RBPs. (D) Deregulation of the AS network by the 10 most enriched RBPs (purple) and RASEs (blue). The 4 enriched GO terms associated with atherosclerotic plaque with RASGs (yellow) with disrupted expression shown in red.

genes were enriched in “cytokine–cytokine interaction, B-cell receptor signaling pathway, cell adhesion molecules, chemokine signaling pathway, leukocyte transendothelial migration, hematopoietic cell lineage” (Supplementary Fig. 5). The AS network deregulated by the 10 most enriched RBPs, RASEs of TFs, differentially expressed target genes of TFs and related pathways is shown in Fig. 4D.

**RBM47 is involved in the regulation of transcription factor (TF) splicing.** Among the RBPs noted above, TLR7<sup>24–26</sup> and RBM47<sup>27</sup> have been reported to be associated with atherosclerosis, while SAMHD1<sup>28</sup>, DDX60 L<sup>29</sup>, and PARP12<sup>30</sup> have been reported to be indirectly associated with atherosclerosis. Therefore, we investigated whether these RBPs (TLR7, RBM47, SAMHD1, DDX60 L, PARP12, and SMAD9) were altered in vitro (Supplementary Fig. 6). HUVECs were stimulated with ox-LDL to mimic cholesterol loading of the endothelium in atherosclerosis. The results showed that RBM47, TLR7 and SAMHD1 were significantly different between two groups, but only RBM47 was consistent with the RNA-seq data in GSE104140 (Fig. 5A). According to the results shown in Fig. 4D, ERG, ELF1, and BCLAF1 are all TFs that may be regulated by RBM47, so we verified whether alternative splicing occurs in ERG, ELF1, and BCLAF1. The results showed that the percentage of ASEs of BCLAF1 increased, which is consistent with the RNA-seq data in GSE104140 (Fig. 5B). ELF1 has a decreased probability of variable splicing events in the ox-LDL group and ERG was not significantly different in the two groups (Supplementary Figure). Fig. 5C shows the expression of five differentially expressed







**Figure 5.** RBM47 is involved in the regulation of transcription factor (TF), which mediates gene expression in inflammatory-related pathways in atherosclerotic plaques. **(A)** The left panel shows RBM47 expression levels in the early disease stage (SAMP\_DIT) samples and advanced disease stages (SAMP\_advanced) samples. \*\*\*:  $p$  value < 0.001. The right panel demonstrates the expression of RBM47 in three biological replicates of HUVEC stimulated with ox-LDL. \*\*\*\*:  $p$  value < 0.0001. **(B)** Box plots in the left panel show the splicing ratio of RAS sites located in BCLAF1. The right panel shows the validation of this RASE in three biological replicates of HUVEC stimulated with ox-LDL by RT-qPCR. \*\*\*\*:  $p$  value < 0.0001. **(C)** Illustration of five differentially expressed target genes that are regulated by ERG, BCLAF1, and ELF1 in the inflammatory pathway.

angiogenesis, and VEGF165b inhibits angiogenesis<sup>9</sup>. These studies revealed an association between AS and the development of atherosclerosis.

RBPs has been known to control and regulate the fate of all RNAs within the cell, which includes RNA splicing<sup>34</sup>. Therefore, we analyzed the differential RBPs in the dataset. RBM47, TLR7, RNASE6, SAMHD1, SMAD9, RNASE1, DDX60 L and PARP12 were identified as hub RBPs in plaque progression. RBM47 is a key player in gene posttranscriptional regulation with one or more RNA recognition motif domains<sup>35</sup> and is involved in the editing of ApoB mRNA, which affects the production of the ApoB functional proteins APOB100 and APOB48<sup>36</sup>. Apolipoprotein B, the cholesterol-carrying component of LDL, is known to have a carcinogenic effect on the retention and accumulation of subendothelial lipids in arteries<sup>37</sup>. Glykeria K et al. found that patients with higher levels of TLR7 transcripts presented with a lower risk of major cardiovascular and cerebrovascular events in the follow-up period after carotid endarterectomy, and carotid plaque immunohistochemistry showed that TLR7 was expressed in all plaques by T cells, macrophages and endothelial cells in capillaries<sup>24</sup>. Similar to TLR7, SAMHD1<sup>38</sup> initiates innate immunity mainly by binding endogenous ligands, including viral-derived ssRNA and mononucleotides<sup>39</sup>. However, recent studies have pointed to several endogenous ligands, such as RNA from apoptotic cells and extracellular miRNA<sup>40</sup>. In atherosclerosis, RNA from apoptotic and necrotic cells and from tissue damage may be reasonable sources of ligands. It has been shown that hypomethylation of the Rnase6 promoter enhances the proliferation and migration of mouse aortic vascular smooth muscle cells and aggravates atherosclerosis in mice<sup>41</sup>. However, it is not clear whether it affects the disease state in atherosclerosis by exerting RNA enzymatic effects. Smad signaling is important for mediating TGF- $\beta$  signaling from the cell surface to the nucleus<sup>42</sup>. Upstream BMP4 induces mouse foam cell formation via SMAD9<sup>43</sup>. DDX60 L and PARP12 are related to immune and inflammatory responses that contribute to the development of atherosclerosis<sup>44</sup>. The endothelial extracellular endonuclease ribonuclease 1 (RNase1) belongs to the ribonuclease A superfamily<sup>45</sup>, which is mainly expressed in various vascular endothelial cells<sup>46</sup>. There are many Weibel-Palade Bodies (WPBs) in the vascular endothelium, which store inflammatory mediators, including RNase1, IL-8, and chemokines<sup>47</sup>. When endothelial cells are activated, eRNA is released. eRNA acts as an immune inducer, enabling WPBs to release inflammatory mediators and trigger inflammation. At this time, the RNase 1 released at the same time can bind to eRNA so that the blood vessel can be protected from a severe inflammatory response<sup>48,49</sup>. The vascular endothelium is the initiating site of atherosclerosis, and good endothelial function and protection from inflammation are important for preventing atherosclerosis. Some RBPs have not been reported to be related to atherosclerosis, such as

MYEF2. MYEF2 is implicated in neurodegeneration; MYEF2 is a transcriptional repressor of the myelin basic protein gene and is involved in central nervous system development by controlling oligodendrocyte progenitor cell differentiation through the regulation of myelin protein expression<sup>50</sup>. Atherosclerosis is not only a vascular disease but is also associated with neurological disease<sup>51</sup>. We speculate that atherosclerosis may be related to sympathetic modulation. These differentially expressed RBPs provide a basis for our follow-up study of the RBP-AS-atherogenic regulatory network.

Through joint analysis, several AS genes (ABI1, FXR1, CHID1, PRKACB, and PPP3CB) were found to be simultaneously changed with changes in RBP expression. These AS genes were mainly enriched in RNA splicing, apoptotic processes, and signal transduction, which are related to atherosclerosis. AS abnormalities of these genes may have a direct relationship and play a role in atherosclerosis. ABI1 AS has been reported to contribute to macrophage differentiation from vascular SMCs during atherogenesis through activation of Rac1 expression and the Rac1-NOX1-ROS pathway, leading to an increase in transcription factor Kruppel-like factor 4 (KLF4) cell-like phenotype regulation<sup>52</sup>. Apolipoprotein M is an apolipoprotein that can bind to HDL-related sphingosine 1-phosphate (S1P), which transports S1P to receptors on the endothelium for endothelial protection<sup>53</sup>. FXR1 can upregulate the expression of apolipoprotein M, thereby mediating antiatherosclerosis<sup>54</sup>. CHID1 belongs to the family of chitinases and chitinase-like proteins. The chitinase-like protein family chitinase and YKL-40 reflect macrophage activation in atherosclerotic plaques<sup>55</sup>, so we speculate that CHID1 may also play a similar role in atherosclerosis. cAMP-dependent protein kinases or protein kinase A (PKA) are an important class of protein kinases in eukaryotic cells. The catalytic subunit of PKA is encoded by two major genes, PRKACA and PRKACB<sup>56</sup>. The PRKACB gene encodes several splice variants that are expressed in a highly cell- and tissue-specific manner. Among them, C $\beta$ 1 is ubiquitously expressed, C $\beta$ 2 is enriched in immune cells, while C $\beta$ 3, C $\beta$ 4 and their abc variants are only expressed in neuronal cells. Loss of C $\beta$ 2 can increase the inflammatory susceptibility of macrophages<sup>57</sup>, and C $\beta$ 2 transcripts were reduced in the advanced samples of this study, which may cause plaque progression. Angiotensin II-induced vascular smooth muscle exhibits increased PPP3CB-dependent extracellular matrix secretion, suggesting that PPP3CB destabilizes plaques during atherosclerosis<sup>58</sup>.

We analyzed the TFs that underwent alternative splicing to reveal a possible RBP-TF regulatory relationship in atherosclerosis. Notably, Urszula R et al. investigated molecular signatures in human plaques stratified by echogenicity as determined by duplex ultrasound<sup>59</sup>, and BCLAF1, which is similar to ZEB2, was identified as an important gene in atherosclerosis. Moreover, BCLAF1 appears to be functionally required for SMC survival and transdifferentiation to a macrophage-like phenotype<sup>59,60</sup>. The signal transducer and activator of transcription (STAT) family includes STAT1, STAT2, and STAT3, which transmit signals from the proinflammatory cytokines interferon (IFN) $\alpha$  and IFN $\gamma$  and pattern recognition receptor (PRR) Toll-like receptors, which are found in blood vessels. An inflammatory environment is formed under the vessel membrane to promote atherosclerosis<sup>61,62</sup>. The transcription factor ERG is essential for endothelial homeostasis, driving the expression of lineage genes and suppressing proinflammatory genes, all of which are associated with atherosclerosis<sup>63,64</sup>.

The analysis of differentially expressed RBPs and differential AS events in atherosclerosis in this study will help us to better understand the molecular mechanism of the occurrence and development of atherosclerosis and will provide new directions and ideas for discovering molecular therapeutic targets for atherosclerosis.

## Data availability

The datasets analysed during the current study are available in the Gene Expression Omnibus (GEO) repository (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE104140>).

Received: 2 August 2022; Accepted: 16 December 2022

Published online: 31 January 2023

## References

- Libby, P. et al. Atherosclerosis. *Nat. Rev. Dis. Primers* **5**(1), 56 (2019).
- Libby, P. The changing landscape of atherosclerosis. *Nature* **592**(7855), 524–533 (2021).
- Benjamin, E. J. et al. Heart disease and stroke statistics-2017 update: a report from the American Heart Association. *Circulation* **135**(10), e146–e603 (2017).
- World Health Organization. *Cardiovascular diseases (CVDs) Fact Sheet*. (2017).
- Pan, Q. et al. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* **40**(12), 1413–1415 (2008).
- Baralle, F. E. & Giudice, J. Alternative splicing as a regulator of development and tissue identity. *Nat. Rev. Mol. Cell Biol.* **18**(7), 437–451 (2017).
- Hasimbegovic, E. et al. Alternative Splicing in cardiovascular disease—a survey of recent findings. *Genes (Basel)* **12**(9), 1457 (2021).
- Joly, A. L. et al. Alternative splicing of FOXP3 controls regulatory T cell effector functions and is associated with human atherosclerotic plaque stability. *Circ. Res.* **122**(10), 1385–1394 (2018).
- Zhao, N. & Zhang, J. Role of alternative splicing of VEGF-A in the development of atherosclerosis. *Aging (Albany NY)* **10**(10), 2695–2708 (2018).
- Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**(4), R36 (2013).
- Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**(12), 550 (2014).
- Jin, L. et al. Transcriptome analysis reveals the complexity of alternative splicing regulation in the fungus *Verticillium dahliae*. *BMC Genom.* **18**(1), 130 (2017).
- Xia, H. et al. CELF1 preferentially binds to exon-intron boundary and regulates alternative splicing in HeLa cells. *Biochim. Biophys. Acta Gene Regul. Mech.* **1860**(9), 911–921 (2017).
- Xie, C. et al. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* **39**, W316–W322 (2011).
- Wolf, D. & Ley, K. Immunity and inflammation in atherosclerosis. *Circ Res* **124**(2), 315–327 (2019).

16. Santarpia, L., Lippman, S. M. & El-Naggar, A. K. Targeting the MAPK-RAS-RAF signaling pathway in cancer therapy. *Expert Opin. Ther. Targets* **16**(1), 103–119 (2012).
17. Reustle, A. & Torzewski, M. Role of p38 MAPK in atherosclerosis and aortic valve sclerosis. *Int. J. Mol. Sci.* **19**(12), 3761 (2018).
18. Sumara, G., Belwal, M. & Ricci, R. “Jnking” atherosclerosis. *Cell Mol. Life Sci.* **62**(21), 2487–2494 (2005).
19. Tabas, I. & Bornfeldt, K. E. Macrophage phenotype and function in different stages of atherosclerosis. *Circ. Res.* **118**(4), 653–667 (2016).
20. Lu, Q. & Rounds, S. Focal adhesion kinase and endothelial cell apoptosis. *Microvasc. Res.* **83**(1), 56–63 (2012).
21. Basatemur, G. L. *et al.* Vascular smooth muscle cells in atherosclerosis. *Nat. Rev. Cardiol.* **16**(12), 727–744 (2019).
22. Wu, T. *et al.* PRDM16 is a compact myocardium-enriched transcription factor required to maintain compact myocardial cardiomyocyte identity in left ventricle. *Circulation* **145**(8), 586–602 (2022).
23. Shimizu, S. *et al.* Multimerization of the GATA4 transcription factor regulates transcriptional activity and cardiomyocyte hypertrophic response. *Int. J. Biol. Sci.* **18**(3), 1079–1095 (2022).
24. Karadimou, G. *et al.* Low TLR7 gene expression in atherosclerotic plaques is associated with major adverse cardio- and cerebrovascular events. *Cardiovasc. Res.* **113**(1), 30–39 (2017).
25. Krogmann, A. O. *et al.* Toll-like receptor 7 stimulation promotes the development of atherosclerosis in apolipoprotein E-deficient mice. *Int. Heart J.* **61**(2), 364–372 (2020).
26. Liu, C. L. *et al.* Toll-like receptor 7 deficiency protects apolipoprotein E-deficient mice from diet-induced atherosclerosis. *Sci. Rep.* **7**(1), 847 (2017).
27. Liu, C. *et al.* Identifying RBM47, HCK, CD53, TYROBP, and HAVCR2 as hub genes in advanced atherosclerotic plaques by network-based analysis and validation. *Front. Genet.* **11**, 602908 (2020).
28. Chen, S. *et al.* SAMHD1 suppresses innate immune responses to viral infections and inflammatory stimuli by inhibiting the NF-kappaB and interferon pathways. *Proc. Natl. Acad. Sci. U S A* **115**(16), E3798–E3807 (2018).
29. Grunvogel, O. *et al.* DDX60L is an interferon-stimulated gene product restricting hepatitis C virus replication in cell culture. *J. Virol.* **89**(20), 10548–10568 (2015).
30. Welsby, I. *et al.* PARP12, an interferon-stimulated gene involved in the control of protein translation and inflammation. *J. Biol. Chem.* **289**(38), 26642–26657 (2014).
31. Ono, M. *et al.* Foxp3 controls regulatory T-cell function by interacting with AML1/Runx1. *Nature* **446**(7136), 685–689 (2007).
32. Ait-Oufella, H. *et al.* Natural regulatory T cells control the development of atherosclerosis in mice. *Nat. Med.* **12**(2), 178–180 (2006).
33. Klingenberg, R. *et al.* Depletion of FOXP3+ regulatory T cells promotes hypercholesterolemia and atherosclerosis. *J. Clin. Invest.* **123**(3), 1323–1334 (2013).
34. Sternburg, E. L. & Karginov, F. V. Global approaches in studying RNA-binding protein interaction networks. *Trends Biochem. Sci.* **45**(7), 593–603 (2020).
35. Shivalingappa, P. K. M. *et al.* RNA binding motif 47 (RBM47): emerging roles in vertebrate development, RNA editing and cancer. *Mol. Cell Biochem.* **476**(12), 4493–4505 (2021).
36. Fossat, N. *et al.* C to U RNA editing mediated by APOBEC1 requires RNA-binding protein RBM47. *EMBO Rep.* **15**(8), 903–910 (2014).
37. Skalen, K. *et al.* Subendothelial retention of atherogenic lipoproteins in early atherosclerosis. *Nature* **417**(6890), 750–754 (2002).
38. Li, M. *et al.* Roles of SAMHD1 in antiviral defense, autoimmunity and cancer. *Rev. Med. Virol.* **27**(4), e1931 (2017).
39. Heil, F. *et al.* Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. *Science* **303**(5663), 1526–1529 (2004).
40. Green, N. M. *et al.* Activation of autoreactive B cells by endogenous TLR7 and TLR3 RNA ligands. *J. Biol. Chem.* **287**(47), 39789–39799 (2012).
41. Zhao, B. *et al.* Six-gene signature associated with immune cells in the progression of atherosclerosis discovered by comprehensive bioinformatics analyses. *Cardiovasc. Ther.* **2020**, 1230513 (2020).
42. ten Dijke, P. & Hill, C. S. New insights into TGF-beta-Smad signalling. *Trends Biochem Sci* **29**(5), 265–273 (2004).
43. Feng, J. *et al.* BMP4 enhances foam cell formation by BMPR-2/Smad1/5/8 signaling. *Int. J. Mol. Sci.* **15**(4), 5536–5552 (2014).
44. Coquel, F. *et al.* SAMHD1 acts at stalled replication forks to prevent interferon induction. *Nature* **557**(7703), 57–61 (2018).
45. Sorrentino, S. The eight human “canonical” ribonucleases: molecular diversity, catalytic properties, and special biological actions of the enzyme proteins. *FEBS Lett.* **584**(11), 2194–2200 (2010).
46. Landre, J. B. *et al.* Human endothelial cells selectively express large amounts of pancreatic-type ribonuclease (RNase 1). *J. Cell Biochem.* **86**(3), 540–552 (2002).
47. Fischer, S. & Preissner, K. T. Extracellular nucleic acids as novel alarm signals in the vascular system. Mediators of defence and disease. *Hamostaseologie* **33**(1), 37–42 (2013).
48. Fischer, S. *et al.* Extracellular RNA mediates endothelial-cell permeability via vascular endothelial growth factor. *Blood* **110**(7), 2457–2465 (2007).
49. Bedenbender, K. & Schmeck, B. T. Endothelial ribonuclease 1 in cardiovascular and systemic inflammation. *Front. Cell Dev. Biol.* **8**, 576491 (2020).
50. Wang, Y. *et al.* PARP1-mediated PARylation activity is essential for oligodendroglial differentiation and CNS myelination. *Cell Rep.* **37**(1), 109695 (2021).
51. Mousa, R. F. *et al.* A pathway phenotype linking metabolic, immune, oxidative, and opioid pathways with comorbid depression, atherosclerosis, and unstable angina. *CNS Spectr.* **27**, 676–690 (2021).
52. Li, Y. *et al.* RNA Splicing of the Abi1 Gene by MBNL1 contributes to macrophage-like phenotype modulation of vascular smooth muscle cell during atherogenesis. *Cell Prolif.* **54**(5), e13023 (2021).
53. Christoffersen, C. *et al.* Endothelium-protective sphingosine-1-phosphate provided by HDL-associated apolipoprotein M. *Proc. Natl. Acad. Sci. U. S. A.* **108**(23), 9613–9618 (2011).
54. Zhu, C. *et al.* TO901317 regulating apolipoprotein M expression mediates via the farnesoid X receptor pathway in Caco-2 cells. *Lipids Health Dis.* **10**, 199 (2011).
55. Kzhyshkowska, J., Gratchev, A. & Goerdts, S. Human chitinases and chitinase-like proteins as indicators for inflammation and cancer. *Biomark Insights* **2**, 128–146 (2007).
56. Taylor, S. S. *et al.* PKA Cbeta: a forgotten catalytic subunit of cAMP-dependent protein kinase opens new windows for PKA signaling and disease pathologies. *Biochem. J.* **478**(11), 2101–2119 (2021).
57. Moen, L. V. *et al.* Ablation of the Cbeta2 subunit of PKA in immune cells leads to increased susceptibility to systemic inflammation in mice. *Eur. J. Immunol.* **47**(11), 1880–1889 (2017).
58. Nolze, A. *et al.* Calcineurin (PPP3CB) regulates angiotensin II-dependent vascular remodelling by potentiating EGFR signalling in mice. *Acta Physiol. (Oxf)* **233**(3), e13715 (2021).
59. Rykaczewska, U. *et al.* Plaque evaluation by ultrasound and transcriptomics reveals BCLAF1 as a regulator of smooth muscle cell lipid transdifferentiation in atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **42**(5), 659–676 (2022).
60. Cheng, P. *et al.* ZEB2 shapes the epigenetic landscape of atherosclerosis. *Circulation* **145**(6), 469–485 (2022).
61. Plens-Galaska, M. *et al.* Genome-wide inhibition of pro-atherogenic gene expression by multi-STAT targeting compounds as a novel treatment strategy of CVDs. *Front. Immunol.* **9**, 2141 (2018).

62. Laguna-Fernandez, A. *et al.* Endothelial transcriptomic changes induced by oxidized low density lipoprotein disclose an up-regulation of Jak-Stat pathway. *Vascul. Pharmacol.* **73**, 104–114 (2015).
63. Kalna, V. *et al.* The transcription factor ERG regulates super-enhancers associated with an endothelial-specific gene expression program. *Circ. Res.* **124**(9), 1337–1349 (2019).
64. Sperone, A. *et al.* The transcription factor Erg inhibits vascular inflammation by repressing NF-kappaB activation and proinflammatory gene expression in endothelial cells. *Arterioscler. Thromb. Vasc. Biol.* **31**(1), 142–150 (2011).

### Author contributions

R.W. and J.X. proposed this research. R.W. wrote this manuscript. R.W., J.Z., and Y.T. participated in the data analysis. Y.W. and L.D. participated in the design of the study. Y.P., and Z.Z. reviewed and edited the manuscript.

### Funding

These authors funded by National Key Research and Development Program of China, No. 2018YFC1311505, Clinical Cooperative Pilot Project of Traditional Chinese and Western Medicine for Major Diseases, No. Administration of State Administration of Traditional Chinese Medicine [2018] No.3, Gansu Provincial Key Laboratory of Cardiovascular Diseases Construction Project No. 1206RTSA025.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-26556-6>.

**Correspondence** and requests for materials should be addressed to Z.Z.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023