The Inflamm-Aging Model Identifies Key Risk Factors in Atherosclerosis

Yudan He¹, Yao Chen¹, Lilin Yao¹, Junyi Wang¹, Xianzheng Sha¹ and Yin Wang¹,²*

¹Department of Biomedical Engineering, School of Intelligent Sciences, China Medical University, Shenyang, China, ²Tumor Etiology and Screening Department of Cancer Institute and General Surgery, The First Affiliated Hospital of China Medical University, Shenyang, China

Background: Atherosclerosis, one of the main threats to human life and health, is driven by abnormal inflammation (i.e., chronic inflammation or oxidative stress) during accelerated aging. Many studies have shown that inflamm-aging exerts a significant impact on the occurrence of atherosclerosis, particularly by inducing an immune homeostasis imbalance. However, the potential mechanism by which inflamm-aging induces atherosclerosis needs to be studied more thoroughly, and there is currently a lack of powerful prediction models.

Methods: First, an improved inflamm-aging prediction model was constructed by integrating aging, inflammation, and disease markers with the help of machine learning methods; then, inflamm-aging scores were calculated. In addition, the causal relationship between aging and disease was identified using Mendelian randomization. A series of risk factors were also identified by causal analysis, sensitivity analysis, and network analysis.

Results: Our results revealed an accelerated inflamm-aging pattern in atherosclerosis and suggested a causal relationship between inflamm-aging and atherosclerosis. Mechanisms involving inflammation, nutritional balance, vascular homeostasis, and oxidative stress were found to be driving factors of atherosclerosis in the context of inflamm-aging.

Conclusion: In summary, we developed a model integrating crucial risk factors in inflamm-aging and atherosclerosis. Our computation pipeline could be used to explore potential mechanisms of related diseases.

Keywords: inflamm-aging, atherosclerosis, immune homeostasis, causal analysis, sensitive analysis

INTRODUCTION

The burden of cardiovascular disease (CVD) has increased, leading to it being considered one of the most expensive diseases (i.e., seriously costing the economic burden) (Lopez et al., 2021). CVD is becoming an increasingly serious problem, with a high incidence rate and accounting for nearly one-third of total deaths (Stewart et al., 2017). Atherosclerosis is considered to be the major cause of CVD

Abbreviations: AUC, area under the ROC curve; BP, biological process; CVD, cardiovascular disease; ECM, extracellular matrix; FDR, false discovery rate; GEO, Gene Expression Omnibus; GO, Gene Ontology; GSEA, gene set enrichment analysis; KEGG, Kyoto Encyclopedia of Genes and Genomes; KIR, killer immunoglobulin-like receptors; KNN, k-nearest neighbors; MCMC, Markov chain Monte Carlo; ML, machine learning; MMP, matrix metalloproteinase; MR, Mendelian randomization; NB, naive Bayes; OXPHOS, oxidative phosphorylation; PCA, principal components analysis; ROC, receiver operating characteristic; ROS, reactive oxygen species; SVM, support vector machine; TLR, toll-like receptors; VSMC, vascular smooth muscle cells.
A chronic disease, atherosclerosis, involves the accumulation of plaques, and its major traditional risk factors (i.e., blood lipid proteasemia, diabetes, smoking, hypertension, and genetic abnormalities) have been reported by many epidemiological studies (Bergheanu et al., 2017; Pahwa and Jialal, 2021). However, atherosclerosis is a complex disease, and there are multiple cellular and molecular interactions involved in the progression of atherosclerosis that need to be further investigated and integrated (Lusis, 2000). Moreover, treatment is often difficult, expensive, invasive, and risky, and there is still no effective means of preventing atherosclerosis (Fogoros and Ali, 2020).

An increasing number of reports show that aging is a driving factor of atherosclerosis (El Assar et al., 2012; Wang and Bennett, 2012). Aging is closely related to endothelial dysfunction and arterial stiffness, which are considered to be early events leading to CVD. The aging process involves promotion of a series of risk factors (i.e., oxidative stress, endothelial dysfunction, and pro-inflammatory cytokines), leading to endothelial dysfunction and vascular system damage. In addition, cellular aging induces the release of microbubbles, further contributing to the development and calcification of atherosclerotic plaques (Alique et al., 2018). Therefore, the incidence of atherosclerosis increases with chronological age, and accelerated aging is the main risk factor for the development of atherosclerotic plaques (Kahn, 2021). Furthermore, atherosclerotic plaques represent a key index of cellular aging, which is characterized by reduced cell proliferation, increased DNA damage, and telomere shortening. In summary, a growing body of evidence indicates that atherosclerosis is promoted by cellular aging (Wang and Bennett, 2012).

Atherosclerosis is closely related to inflamm-aging, along with oxidative stress, endothelial dysfunction, and inflammation (Alique et al., 2018). Inflamm-aging is defined as a chronic inflammatory process during aging and mainly characterized by chronic progressive strengthening of the pro-inflammatory response (Xia et al., 2016). In other words, inflamm-aging promotes the body’s pro-inflammatory status with advancing aging, which is closely related to many aging diseases (Xia et al., 2016). A series of studies have shown that aging can promote atherosclerosis by damaging the connection between mitochondrial function and the intravascular inflammatory pathway (Tyrrell and Goldstein, 2020). For example, chronic inflammation is the main cause of age-related atherosclerosis, possibly exerting its effect through the IL-6 signaling pathway (Tyrrell and Goldstein, 2020). Furthermore, inflammatory factors released by senescent cells result in a senescence-associated secretory phenotype and even atherosclerosis (Childs et al., 2017). Selective targeting and elimination of these senescent cells (called heterolysis) has been shown to slow the growth of atherosclerotic lesions by reducing the release of inflammatory and adhesion factors (Childs et al., 2017).

Inflamm-aging induces an immune homeostasis imbalance during atherosclerosis development. As a chronic inflammatory disease, atherosclerosis is characterized by endothelial dysfunction and abnormal immune responses; immune cells and lipids are also involved (Karunakaran et al., 2015). The normal function of the adaptive immune system appears to decline with chronological age, and an advanced inflammatory response is promoted, leading to the imbalance of immune homeostasis (Sadighi Akha, 2018). Both innate and adaptive immune responses have been identified in atherosclerosis, and components of cholesterol-carrying low-density lipoprotein can further trigger abnormal inflammation (Keeseey and Powley, 2008). Unsatuated fatty acids can also regulate both the differentiation and activation of T cells and then induce inflammation and oxidative stress (Hubler and Kennedy, 2016), and accumulation of dysregulated proteins can damage the normal function of the endoplasmic reticulum, which has an important role in oxidative stress and the immune response (So, 2018). In short, the nutritional balance is critical to immune homeostasis (Chandra, 1997). Chronic oxidative stress is also vital to immune/inflammatory cells during the inflamm-aging process (Fuente and Miquel, 2009). However, further and more systematic investigation is required to fully elucidate how atherosclerosis is triggered by inflamm-aging in the context of immune homeostasis imbalance. In order to study the potential mechanisms involved in inflamm-aging, immune homeostasis imbalance, and atherosclerosis, we hypothesized that the progression of atherosclerosis would be promoted by accelerated chronic inflammation during the aging process when immune homeostasis was disrupted (Figure 1A).

At present, there are many omics profiles to be further analyzed; thus, proper prediction models need to be developed (Munger et al., 2021). Machine learning (ML), which is regarded as an extension of classical statistical modeling, can digest large amounts of data to identify high-order correlations and generate predictions (Alber et al., 2019; Zhavoronkov et al., 2019). For example, the supervised ML method can be used to understand the aging process (Fabris et al., 2017). Recently, an ML-derived score has been reported to show satisfactory performance in the prediction of cardiovascular events (Sánchez-Cabo et al., 2020). Some studies have also shown that ML can predict CVD and identify key markers related to inflamm-aging and CVD (Alalawi and Alsuwat, 2021; Munger et al., 2021; Sayed et al., 2021). Network analysis can also be used to find clusters of highly related genes and to associate modules with each other and with external sample traits (Deshpande et al., 2016). Recently, causality analysis was applied to CVD and found to be informative, using Mendelian randomization (MR) to integrate genomic and phenotypic information (Yazdani et al., 2016). Although many studies have identified CVD-related risk markers (Shadrina et al., 2020), the inflamm-aging mechanism in atherosclerosis (thus the casual relationship between inflamm-aging and atherosclerosis as well as relative potential risk factors) needs to be explored more thoroughly. Therefore, there is an urgent need to model atherosclerosis progression based on the inflamm-aging process, integrating the potential mechanisms related to immune homeostasis imbalance at a system level.

In order to further explore the inflamm-aging mechanism in atherosclerosis, we present a computational pipeline using ML and systems biology methods (Figure 1B), comprising the following steps (Lopez et al., 2021): we developed an improved model of inflamm-aging-related predictors (the inflamm-aging
model) of atherosclerosis and identified relative risk markers (Stewart et al., 2017). An inflam-m-aging score was calculated by integrating both aging and inflammatory markers (Pahwa and Jialal, 2021). MR was carried out to explore the causal relationship between inflam-m-aging and atherosclerosis (Bergheanu et al., 2017). An inflam-m-aging differential network was constructed by summarizing the (partial) correlation between each pair of genes based on the inflam-m-aging score (Lusis, 2000). The Markov chain Monte Carlo (MCMC) method was used to further explore a sensitivity index for inflam-m-aging, immune homeostasis imbalance, and atherosclerosis (Fogoros and Ali, 2020). Enrichment analysis and network analysis were carried out to study potential risk factors in atherosclerosis. As a result, the casual relationship between inflam-m-aging and atherosclerosis was identified, and potential mechanisms related to nutritional balance, oxidative stress, and vessel homeostasis were integrated based on our computational pipeline.

RESULTS
Modeling Inflam-M-Aging Predictors and Identifying Relative Risk Markers
Gene expression data were downloaded from the Gene Expression Omnibus (GEO) database, including 926 samples and 11,313 genes (Supplementary Tables S1, S2). The RelifF algorithm was used to sort relative genes, and the k-nearest neighbors (kNN; k = 5 with the cosine distance) algorithm was selected to mode aging predictors and disease predictors (determined by 10-fold cross-validation; results in Supplementary Tables S3–S7). Other classification results are shown in Supplementary Tables S8–S10 based on the support vector machine (SVM), naive Bayes (NB), and ensemble learning (integrating 100 decision tree models) algorithms (Supplementary Tables S8–S10). The accuracy of the models was then validated using test data (Figure 2; Supplementary Figure S1; Table 1) including 126 atherosclerosis samples and 170 controls, where the aging/inflam-m-aging/disease as well as relative parameters (selected by the cross-validation based on the training set, in Supplementary Tables S5) were also used in the test set. Furthermore, top-ranked aging, inflammatory, and disease markers were used in the improved inflam-m-aging predictor (Supplementary Tables S5). The aging predictor accuracy was 0.7017, and 441 aging markers were identified; the prediction accuracy of the improved inflam-m-aging model was 0.7037, with 186 inflammatory markers and 316 disease markers identified (see Materials and Methods; relative parameters in Supplementary Tables S5). The area under the receiver operating characteristic (ROC) curve values were 0.85682 and 0.74206 for the aging predictor and the improved inflam-m-aging predictor, respectively. These results showed that our predictors were reliable and able to identify healthy young patients.
and old samples as well as distinguish between normal and disease samples with sufficient accuracy (Figure 2, Table 1).

Various biomarkers of aging/disease have been associated with biological functions (Supplementary Table S5). For example, MMP8, the top aging marker, is a collagenase with proteolytic activity against matrix proteins and is an important participant in the function of the vascular system and the development of atherosclerosis (Laxton et al., 2009; Toba et al., 2017; Yu et al., 2019). Its expression in the heart increases with chronological age, enabling it to regulate signal transduction during the aging process by regulating the expression and activity of cytokines, chemokines, growth factors, hormones, and angiogenic factors (Mota et al., 2018). The top disease marker is KIR3DL1. Killer immunoglobulin-like receptors (KIR) encode receptors expressed on some subsets of natural killer cells and T lymphocytes (Niepieklo-Miniewska et al., 2014), which can regulate
inflammatory response (Ross, 2014; Pugh et al., 2019). Among the 186 inflammatory markers identified, HSPA6 had the lowest false discovery rate (FDR = 4.38e-9). HSPA6 participates in cell protection (Noonan et al., 2007) and cell proliferation and affects the receptor binding affinity of IgG (Takai, 2002; Wang et al., 2020a). HSPA6 is also a putative target of endothelial nitric oxide synthase and is involved in the highly clinically relevant inhibition of vascular smooth muscle cell (VSMC) proliferation, which can repair vessels in human atherosclerosis (McCullagh et al., 2016). Thus, it could be speculated that VSMCs in the vascular system are closely related to atherosclerosis development in the context of inflamm-aging, consistent with previous research results (Bennett et al., 2016).

**Comparison of Inflamm-Aging Scores Between Normal and Disease Samples**

In order to investigate the accelerated inflamm-aging pattern in atherosclerosis, an inflamm-aging score was constructed by integrating 186 inflammatory markers and 441 aging markers based on the training data (details are given in the “Materials and Methods” section); the results are shown in Table 2. Both the median and mean of the inflamm-aging scores of atherosclerosis samples were greater than those of control samples. These results indicated an accelerated inflamm-aging pattern in atherosclerosis, consistent with previous findings (Lavin Plaza et al., 2020). To further verify this conclusion, the Kruskal–Wallis test (Figure 3A) was used to validate the significance of inflamm-aging in atherosclerosis. The results confirmed that atherosclerosis samples showed a significantly accelerated inflamm-aging pattern compared with normal samples ($p = 1.6856e-05$). In addition, when inflamm-aging scores were adjusted by chronological age (see Materials and Methods), the accelerated inflamm-aging pattern remained significant ($p = 0.0024$; Figure 3B; Table 2). A principal component analysis (PCA) plot is also shown in Supplementary Figure S2.

**Exploring the Causal Relationship Between Inflamm-Aging and Atherosclerosis by MR**

In order to study the inflamm-aging mechanisms in atherosclerosis more thoroughly, the MR method was used to distinguish causal relationships among aging, inflammation, and

| TABLE 2 | Inflamm-aging scores of the disease and control groups.
| --- | --- | --- | --- |
| Mean (control) | Mean (disease) | Median (control) | Median (disease) |
| Original score | 0.0129 | 0.0230 | 0.0131 | 0.0250 |
| Score adjusted by age | 0.0096 | 0.0161 | 0.0092 | 0.0162 |

| TABLE 3 | Top 10 pairs with the largest absolute difference in correlation and partial correlation between “inflammation+ atherosclerosis−” and “inflammation+ atherosclerosis+”.
| --- | --- | --- |
| Aging marker | Disease marker | Difference |
| BIRC2 | PLOD1 | 0.0650 |
| CRBN | IRF5 | 0.0617 |
| ICAM2 | IRF5 | 0.0581 |
| BIRC2 | CPNE1 | 0.0524 |
| RPRM | CEBPA | 0.0518 |
| ICAM2 | NGB | 0.0514 |
| ICAM2 | CEBPA | 0.0499 |
| ICAM2 | PARP16 | 0.0496 |
| ICAM2 | PECAM1 | 0.0491 |
| ATP6V0C | POM121 | 0.0483 |
Table 4: Top 10 markers with the highest degrees.

| Gene symbol | Degree | Function | Reference |
|-------------|--------|----------|-----------|
| PXN         | 117    | 1) Related to the attachment of actin membrane to ECM, 2) Participates in tissue remodeling, cell proliferation, and survival, 3) Important role in focal adhesion, endothelial dysfunction, inflammation, and oxidative stress | Chen et al. (2021a) |
| ZNF22       | 110    | Participates in bone marrow mesenchymal stem cell differentiation, including cardiac muscle cells, nerve cells, blood cells, and myogenic cells | Cao et al. (2021) |
| RBM10       | 100    | 1) Regulates cell apoptosis, cell proliferation, cell invasion and metastasis, and inflammatory response, 2) Affects the pathogenesis of atherosclerosis | Schierle et al. (2015), Nie et al. (2020) |
| NADK        | 93     | 1) Important role in regulating cell aging and aging-related diseases, 2) Responsible for the production of NADP in the cytoplasmic matrix and participates in the counteraction of oxidative damage | Zhang (2013), Covarrubias et al. (2021) |
| ANKRD12     | 90     | Participates in bone marrow mesenchymal stem cell differentiation, including cardiac muscle cells, nerve cells, blood cells, and myogenic cells | Wang et al. (2021) |
| PJA2        | 89     | Promotes M1 macrophage polarization, M2 to M1 macrophage transformation, and the inflammatory response | Zheng et al. (2021) |
| ATP5I       | 89     | Induces oxidative stress and DNA damage | Liu et al. (2021), Yang et al. (2000), Zheng et al. (2017) |
| MAP3K3      | 88     | 1) Necessary for angiogenesis, 2) Related to endothelial cell proliferation and apoptosis and interacts with heart and myocardium, 3) Participates in cell proliferation, differentiation, and apoptosis, 4) Related to cellular aging | Mezzadra et al. (2017), Guan et al. (2018), Zheng et al. (2020) |
| CMTM6       | 84     | 1) Important role in regulating immune response and inflammatory activation, 2) Regulates the expression of PD-L1, PD-L1 can inhibit the function of T-cell inhibition, 3) Promotes cell migration and invasion | Mezzadra et al. (2017), Guan et al. (2018), Zheng et al. (2020) |
| SNRPG       | 82     | 1) Arrests the cell cycle, 2) Activates p53 signaling | Lan et al. (2020) |

Atherosclerosis, and partial correlation coefficients were calculated (see Materials and Methods for details). Crucial causal relationships related to inflamm-aging were identified in this way (Table 3, Supplementary Table S1, S2). For example, RPRM was the top aging marker in MR, related to the most (Takai, 2002) inflamm-aging markers. RPRM is a highly glycosylated cytoplasmic protein encoded by a p53-dependent putative tumor suppressor gene (Ohki et al., 2000). Its overexpression decreases cell proliferation, migration, and invasion and increases apoptosis (Figueroa et al., 2017). PECAM1 was the top disease marker, related to the maximum number (Merl-Pham et al., 2019) of aging markers. The expression of PECAM1 is associated with the response of various immune cell types (i.e., CD4 + T cells, B cells, CD8 + T cells, neutrophils, macrophages, and dendritic cells) as well as endothelial cells (Harry et al., 2008; Ye et al., 2021). The activation of endothelial PECAM1 has been shown to lead to nuclear translocation of nuclear factor xB (NF-xB) and expression of inflammatory and adhesion mediators, which promoted atherosclerosis (Harry et al., 2008). Polymorphism of PECAM1 is associated with the incidence rate of coronary atherosclerosis (Harry et al., 2008), indicating that inflamm-aging could accelerate atherosclerosis progression.

BIRC2–PLOD1 was the top aging–atherosclerosis pair (with the maximum differential K-S value of 0.06498; Table 3). BIRC2 is an E3 ubiquitin-protein ligase that regulates NF-xB signaling to maintain endothelial cell survival and blood vessel homeostasis during vascular development (Santoro et al., 2007; Samanta et al., 2020). PLOD1 can promote cell growth and the aerobic glycolysis process (Zhang et al., 2021). It also plays vital roles in atherosclerosis, including enhancing collagen fibril deposition (Yuan et al., 2021), modifying collagen (Wågsäter et al., 2013), and regulating the expression of extracellular matrix (ECM) (Merl-Pham et al., 2019). Thus, we conjectured that the vascular system would have an effect on atherosclerosis. In summary, key causal relationships between inflamm-aging and atherosclerosis were identified using MR.

The Inflamm-Aging Differential Network Revealed Key Inflamm-Aging Mechanisms in Atherosclerosis

In order to explore potential inflamm-aging mechanisms in atherosclerosis at a system level, an inflamm-aging differential network was constructed by calculating the (partial) correlation coefficient for each pair of genes in the training data and test data, respectively (details shown in Materials and Methods). The Pearson correlation coefficient of degree and degree frequency indicate the reliability of the network. Furthermore, when the network was verified on the test set, the p-value was close to 0 using Fisher’s exact test. All these results indicate the reliability of the network.

The network markers indicated that endothelium, oxidative stress, and inflammation have important roles in atherosclerosis; the top-ranked genes (by degree) and their corresponding functions are shown in Table 4 (Yang et al., 2000; Zhang, 2013; Schisler et al., 2015; Mezzadra et al., 2017; Zheng et al., 2017; Guan et al., 2018; Lan et al., 2020; Nie et al., 2020; Zheng.
metabolism disorders (induced by high-fat levels) and negatively roles in cardiac development; for example, it protects the heart as the main sensor of Notch signal transduction. HEY2 has a key and Chen, 2017; Jeon et al., 2021). It also protects cells from fat kinase (AMPK) signaling pathway, and cellular metabolisms (Shi of disease in maintaining cell homeostasis (Margiotta, 2021). Vesicle component in the regulation of vesicle fusion and plays a key part in maintaining cell homeostasis (Margiotta, 2021). Vesicle transport is the main mechanism of protein and lipid assembly of the oxidative phosphorylation complex (Jia et al., 2020). In short, the ribosome has an important role in the development of atherosclerosis.

The aging marker related to the most (Keesey and Powley, 2008) disease markers (Supplementary Tables S13) was HEY2, the main sensor of Notch signal transduction. HEY2 has a key role in cardiac development; for example, it protects the heart from age-induced hypertrophic changes (Röning et al., 2022). It also regulates the proliferation of VASCs (Shirvani et al., 2007). PECAM1 was the inflam-maging marker related to the maximum number (Alber et al., 2019) of aging markers (Supplementary Tables S14) in both the sensitivity analysis and MR (Harry et al., 2008; Ye et al., 2021). These results highlighted the critical role of the immune homeostasis imbalance in atherosclerosis progression.

### Sensitivity Analysis Further Highlighted the Immune Homeostasis Imbalance in Atherosclerosis

In order to further explore the mechanisms underlying inflam-maging, immune homeostasis imbalances, and atherosclerosis, the MCMC method was used to develop a sensitivity index for the relationship between inflam-maging and atherosclerosis (based on the results obtained by MR). A bipartite graph of aging–disease pairs was constructed (see Materials and Methods, and Supplementary Figure S4).

The top 10 most sensitive pairs are shown in Table 5 (based on absolute differential frequencies). The relationship pair with the greatest difference (0.08172) was CRBN–MRPL40 (Table 5). CRBN functions as a potential regulator of cell homeostasis by mediating the ubiquitination of target substrates involved in processes including ion transport, the AMP-activated protein kinase (AMPK) signaling pathway, and cellular metabolisms (Shi and Chen, 2017; Jeon et al., 2021). It also protects cells from fat metabolism disorders (induced by high-fat levels) and negatively regulates CD4+ T cell activation (Shi and Chen, 2017), and has been implicated in a variety of diseases including CVD, obesity, and fatty liver (Shi and Chen, 2017). MRPL40 is related to mitochondrial functions (Li et al., 2019) and is crucial to ensuring ribosomal translational fidelity and subsequent assembly of the oxidative phosphorylation complex (Jia et al., 2009). In short, these results indicate that oxidative stress plays an important part in the development of atherosclerosis.

The aging marker related to the most disease markers (Supplementary Tables S13) was HEY2, the main sensor of Notch signal transduction. HEY2 has a key role in cardiac development; for example, it protects the heart from age-induced hypertrophic changes (Röning et al., 2022). It...
Meng et al., 2022; De Stefano et al., 2021; Kurek et al., 2021; Kiechl et al., 1998; Tokarev et al., 2009; Hu et al., 2015; Van-Assche et al., 2011; Shimizu et al., 1997; Moore et al., 2013; Patel et al., 2000).

“Aminoglycoside antibiotic metabolic process” (GO:0030647) was the most enriched BP term (enriched in two shortest paths; Figure 4C). Aminoglycosides provide a favorable scaffold for the synthesis of various cationic lipids (Krause et al., 2016; Sainlos et al., 2003), and are used to treat and prevent endocarditis (Gonzalez and Spencer, 1998). “Vesicle-mediated transport between endosomal compartments” (GO:0098927) was the BP term with minimum FDR ($p = 9.0282e^{-08}$, $FDR = 6.7540e^{-04}$; Figure 4D). The main communication process between a cell and its environment is vesicle transport, through bidirectional trafficking among the Golgi, endosomes, and lysosomes (Tokarev et al., 2009). The endosomal–lysosomal system and

| TABLE 6 | Top 10 KEGG with the most numerous enriched paths. |
|----------|----------------------------------------------------|
| **KEGG** | **Enriched shortest path** | **Score** | **Function** | **Reference** |
| **Ribosome** | 8 | 7.4949 | 1) Important role in cell cycle progression and proliferation, closely related to coronary restenosis and atherosclerosis 2) Regulates fibrosis, which is an important factor in the occurrence of atherosclerosis 3) With increasing age, abnormality of the immune system will affect the biogenesis of ribosomes and lead to the production of ribosomal protein autoantibodies 4) Increases oxidative stress during aging, eventually causing damage to ribosomal RNA | Chothani et al. (2019), Turie et al. (2019), Manhas and Rath (2020), Jia et al. (2021), Yin et al. (2021) |
| **Parkinson’s disease** | 7 | 6.5642 | Cardiovascular disease and Parkinson’s disease share common risk factors | Li et al. (2018) |
| **Oxidative phosphorylation** | 5 | 4.7213 | 1) Regulates adiponecin secretion in epicardial adipose tissue and has anti-atherosclerotic and anti-inflammatory effects on blood vessels 2) Triggers atherosclerosis by affecting mitochondrial functions | Nakajima et al. (2019), Hosseini et al. (2021) |
| **Huntington’s disease** | 4 | 3.7964 | Myocardial dysfunction and vasoconstrictor dysfunction are involved in Huntington’s disease progression | Critchley et al. (2018) |
| **Alzheimer’s disease** | 3 | 2.8729 | Both atherosclerosis and Alzheimer’s disease are involved in inflammation, macrophage infiltration, and vascular system obstruction | Lathe et al. (2014) |
| **SNARE interactions in vesicular transport** | 1 | 0.9914 | 1) SNARE protein is an essential component that allows membrane fusion and can regulate vesicle fusion 2) SNARE has a key role in cell homeostasis. Vesicle transport is the main mechanism of protein and lipid exchange between membrane-bound organelles in eukaryotic cells 3) Vesicle transport in mitochondria affects mitochondrial function, which is one of the characteristics of aging | Nickel et al. (2002), Chen et al. (2020), Margiotta (2021) |
| **Cytosolic DNA-sensing pathway** | 1 | 0.9843 | 1) Promotes the expression of immune genes 2) Cytoplasmic DNA and cytoplasmic DNA-sensing adapter STING play a key part in aortic degeneration by promoting smooth muscle cell damage, macrophage MMP production, and ECM degeneration 3) Endogenous cytoplasmic DNA is a contributing factor to inflammation in the absence of pathogenic infection, which is associated with many chronic age-related diseases, including cancer, cardiovascular disease, and neurodegenerative diseases | Luo et al. (2020), Miller et al. (2021) |
| **Cysteine and methionine metabolism** | 1 | 0.9785 | 1) The accumulation of oxidized methionine residues of protein is related to aging 2) Hyperhomocysteinemia and increased circulating levels of homocysteine (Hcy) are generally considered to be independent risk factors for peripheral atherosclerosis 3) Hcy can cause oxidative stress, inflammation, and endothelial dysfunction 4) Cysteine is an important source of energy and biomass and has a central role in cell metabolism 5) The metabolites of cysteine and methionine are related to fat metabolism and cardiovascular disease | Stadtman et al. (2005), Kumar et al. (2017), Blachier et al. (2020), Bonifácio et al. (2021) |
| **Acute myeloid leukemia** | 1 | 0.9752 | Inflammation can affect the progression of myeloid leukemia | Récher (2021) |
| **FC epsilon RI signaling pathway** | 1 | 0.9692 | Can induce the production of pro-inflammatory factors and mast cell degranulation | Klemm and Ruland (2006) |
endomembrane system are also closely related to both lipid and protein metabolism (Hu et al., 2015). Enrichment analysis was also performed based on the results of the sensitivity analysis. “Cytosolic DNA-sensing pathway” was the KEGG pathway with the minimum FDR in the sensitivity analysis ($p = 8.4266 \times 10^{-5}$, FDR = 0.0157; Figure 4E); this can promote the expression of immune markers (Luo et al., 2020). Cytoplasmic DNA and cytoplasmic DNA-sensing adapter STING play key parts in aortic degeneration by promoting smooth muscle cell damage, matrix metalloproteinase (MMP) production in macrophages, and ECM degeneration (Luo et al., 2020). Endogenous cytoplasmic DNA is also a contributing factor to inflammation in the absence of pathogenic infection and is associated with many chronic age-related diseases, including CVD and neurodegenerative diseases (Miller et al., 2021). “Positive regulation of reactive oxygen species biosynthetic process” was the BP term with minimum FDR (GO:1903428, $p = 4.3964 \times 10^{-7}$, FDR = 0.0016; Figure 4F). Reactive oxygen species (ROS) play a critical part in the regulation of the expression of multiple promoters and anti-inflammatory genes involved in atherosclerosis, via inducing low-density lipoprotein (LDL) oxidation, foam cell formation, and activation of many redox-sensitive transcription factors including NF-κB and activating protein 1 (Van-Assche et al., 2011).

In summary, a series of risk factors in atherosclerosis were identified, including inflammation, nutritional metabolism, energy metabolism, and the vascular system.

**Network Markers Revealed Crucial Mechanisms Between Aging and Atherosclerosis**

Network markers were also identified by calculating the betweenness of each aging–disease shortest path; the top 10 markers are shown in Table 8. The network marker with the greatest betweenness (Munger et al., 2021) was RPL35 (Figure 5A; Table 8). RPL35 is an important component of the 60S ribosomal subunit, with a key role in mRNA translation and protein synthesis (Wu et al., 2021). Gene betweenness was also calculated according to the sensitivity analysis. The top
markers and their betweenness values are shown in Table 9. According to this analysis, IRAK1 and VAMP8 were the top network markers (betweenness = 8; Figures 5B, C; Table 9). IRAK1 is a threonine/serine kinase that is related to the pathogenesis of atherosclerosis (Rana et al., 2016). It regulates lipid accumulation and prevents the formation of macrophage foam cells (Rana et al., 2016), as well as participating in multiple IL-1- and Toll-like receptors (TLR) driven signaling processes and in the regulation of levels of immune pro-inflammatory cytokines (Singer et al., 2018). VAMP8 is relevant to macrophage degranulation and secretion of TNF-α (Pushparaj et al., 2009), and regulates lysosome–autophagosome interaction and exocytosis (Jones et al., 2012; Chen et al., 2021b). Its phosphorylation regulates lipid metabolism in the liver (Huang et al., 2021). VAMP8 recruits NOX2 into phagosomes, causing lipid oxidation and

| BP | Enriched shortest path | Score | Function | Reference |
|----|------------------------|-------|----------|-----------|
| Aminoglycoside antibiotic metabolic process 0030647 | 2 | 1.9180 | 1) Aminoglycoside antibiotic works by inhibiting protein synthesis 2) Aminoglycosides provide a favorable scaffold for the synthesis of various cationic lipids 3) Aminoglycosides are used to treat and prevent endocarditis | Gonzalez and Spencer (1998), Sainlos et al. (2003), Krause et al. (2016) |
| Myeloid leukocyte activation GO:0002274 | 2 | 1.8743 | 1) Leukocytes can engender protective immunity and protect the host from damage 2) Important role in inflammation, immunity, and atherosclerosis | Swirski and Nahrendorf (2014) |
| Glycoside metabolic process GO:0016137 | 2 | 1.8682 | 1) Glycosides are associated with inflammation, oxidative stress, lipid metabolism, and atherosclerosis 2) Flavonoid glycosides can protect vascular endothelial cells by inhibiting inflammation to restrict atherosclerosis 3) Luteolin-7-O-glucoside can reduce the activity of oxidative stress and inflammatory mechanisms in different physiological systems | De Stefano et al. (2021), Kurek et al. (2022) |
| Tertiary alcohol metabolic process GO:1902644 | 2 | 1.8638 | Alcohol affects the progress of atherosclerosis | Kiehl et al. (1998) |
| Vesicle-mediated transport between endosomal compartments GO:0098627 | 1 | 0.9993 | 1) The main communication process between the cell and its environment 2) Bi-directional trafficking between the Golgi, endosomes, and lysosomes connects the two major intracellular trafficking pathways 3) Endosomes gradually mature into lysosomes, and the acidification of late endosomes is accompanied by vesicle transport 4) Endosomal–lysosomal system and endomembrane system are related to lipid homeostasis and protein homeostasis | Tokarev et al. (2009), Hu et al. (2015) |
| Positive regulation of reactive oxygen species biosynthetic process GO:1903428 | 1 | 0.9984 | 1) Induces LDL oxidation and foam cell formation and activates many redox-sensitive transcription factors including NF-κB and AP1 2) Regulates the expression of multiple promoters/anti-inflammatory genes involved in atherosclerosis | Van-Assche et al. (2011) |
| Positive regulation of nitric oxide metabolic process GO:1904407 | 1 | 0.9984 | Nitric oxide (NO) is associated with atherosclerosis as it inhibits the proliferation of VSMCs, monocyte/macrophage adhesion, platelet aggregation and adhesion, and LDL oxidation | Shimizu et al. (1997) |
| Regulation of nitric oxide metabolic process GO:0080164 | 1 | 0.9982 | NO is associated with atherosclerosis as it inhibits the proliferation VSMCs, monocyte/macrophage adhesion, platelet aggregation and adhesion, and LDL oxidation | Shimizu et al. (1997) |
| Regulation of macrophage activation GO:0043030 | 1 | 0.9970 | The number and phenotype of macrophages can affect the inflammatory state of plaques 2) Lipoprotein uptake by macrophages is considered to be one of the earliest pathogenic events in new plaques 3) The M2 polarization pathway of macrophages can prevent atherosclerosis | Moore et al. (2013) |
| Reactive nitrogen species metabolic process GO:2001057 | 1 | 0.9969 | 1) Affects endothelial function 2) Low-level proteins and lipids modified by reactive nitrogen can promote the development of atherosclerosis through mechanisms involving signal transduction | Patel et al. (2000) |
membrane damage, and also mediates the transport of NOX2 to endosomes and phagosomes, thereby promoting the induction of the cytolytic T cell immune response (Dingjan et al., 2017). In short, the identification of network markers showed that the occurrence of atherosclerosis was closely related to inflammation and immunity homeostasis imbalance.

**DISCUSSION**

CVD has become a serious health problem, the incidence of which increases with age. It is well known that both aging and inflammation have vital roles in atherosclerosis, which is one of the key risk factors for CVD. However, the exact relationships...
between inflam-aging and atherosclerosis and the role of immune homeostasis have to be studied more systematically. It is particularly important to elucidate the causal relationship between inflam-aging and atherosclerosis, in order to determine the mechanisms underlying atherosclerosis and identify potential therapeutic targets. A causal analysis was used to identify the relationship between aging and atherosclerosis using MR. An ML approach was used owing to its great advantages in processing large amounts of data and generating prediction models with appropriate accuracies. Network analysis and sensitivity analysis were also performed to integrate multiple factors and identify relative risk markers, which will be conducive to more systematic study of atherosclerosis.

Our results showed that immune homeostasis is vital for the development of atherosclerosis. Disorders of the immune system can lead to atherosclerosis, and several characteristics related to immune homeostasis imbalance were identified in our study. For example, “myeloid leukocyte activation” (GO:0002274) and “regulation macrophage activation” (GO:0043030) involve the effects of inflammation on atherosclerosis. Several risk markers related to the immune response that have key roles in atherosclerosis were identified, including the top aging marker MMP8, disease markers KIR3DL1 and MAGEA6, network markers based on degree (RBM10, PJA2, and CMTM6; Table 4), and network markers based on betweenness (IRAK1 and VAMP8). Other potential markers including HEY2 (identified by sensitivity analysis) and PECAM1 (identified by MR and sensitivity analysis) and the top causal pairs BIRC2–PLOD1 and CRBN–MRPL40 also suggested a role of immunity in atherosclerosis. In brief, the abnormal immune response is a key characteristic of atherosclerosis (Karunakaran et al., 2015); our results further confirmed that immune homeostasis imbalance is crucial in atherosclerosis in the context of inflam-aging.

Vascular homeostasis imbalance was also identified as a factor in atherosclerosis progression. It has been reported that dysfunction of endothelium and VSMCs, an early marker of atherosclerosis, can disrupt the balance between vasoconstriction and vasodilation, thereby promoting the inflammatory response and exacerbating atherosclerosis (Nkoenig et al., 2022; Alberts et al., 2002; Davignon and Ganz, 2004). The top inflammation marker (HSPA6) and the top network markers according to a degree (PXN and MAPK3) in our results indicated an effect of endothelial cells and VSMCs. Dysfunction of endothelium contributes to the recruitment of inflammatory cells, and VASC proliferation is related to upregulation of pro-inflammatory cytokines (Alique et al., 2018). Consequently, there is a key link between inflammation and vascular homeostasis. The top enriched KEGG pathway was the “cytosolic DNA-sensing pathway” (Figure 4E), indicating the potential relevance of inflammation and vascular homeostasis. In addition, PECAM1 enables endothelial cells to respond to fluid shear stress and regulates inflammatory signaling pathways (Privratsky and Newman, 2014). In the causal pair BIRC2–PLOD1, BIRC2 promotes endothelial cell survival and blood vascular homeostasis via regulation of inflammatory/immune signaling pathways (Santoro et al., 2007; Samanta et al., 2020). PLOD1 encodes lysyl hydroxylase-1, which is essential for ECM maturation. ECM proteolytic fragments affect a variety of functions and properties of inflammatory and immune cells (Adairkirk and Senior, 2008); ECM was also relevant to the integrity and stability of VSMCs (Nkoenig et al., 2022). It has been reported that inflam-aging can lead to changes in vascular structure and aging of VSMCs (Chi et al., 2019). In short, vascular homeostasis has a key role in atherosclerosis and inflam-aging.

The role of oxidative stress was further investigated. Oxidative stress increases ROS generation and/or reduces the imbalance of the body’s innate antioxidant defense system and is thus a key influence on atherosclerosis (Kattoo et al., 2017). Our results provided further evidence of the association of oxidative stress with atherosclerosis, for instance, the top enriched BPs were “positive regulation of reactive oxygen species biosynthetic process” (GO:190,342) and “oxidative phosphorylation” (Table 4). Oxidative stress caused by excessive production of ROS has become the driven risk factor in atherosclerosis (Kattoo et al., 2017). The network marker (based on degree; Table 4) NADK is related to the production of NADPH. NADPH oxidase is the main source of ROS (Ganguly et al., 2021). ROS also has important roles in inflammatory response, apoptosis, cell growth, and changes in vascular tone as well as in the oxidation of LDL cholesterol (Kattoo et al., 2017). For example, ROS can cause physical damage to plaques via stimulating smooth muscle cell migration and collagen deposition, leading to endothelial dysfunction (Ganguly et al., 2021). In summary, the interaction between oxidative stress and vascular homeostasis is vital to atherosclerosis. The relationship between oxidative stress and aging was also explored. Oxidative stress is well known to have an important influence on aging. One of the aging markers identified by our MR approach was RPRM, a p53-inducible gene (Xu et al., 2012). TP53 regulates the expression of various genes involved in the maintenance of homeostasis, including cell cycle regulation and redox homeostasis (i.e., production of antioxidant enzymes) (Gambino et al., 2013; Beyfuss and Hood, 2018). Overall, these findings indicate that oxidative stress is a key factor in atherosclerosis in the context of inflam-aging.

Nutritional imbalance (i.e., lipid and protein imbalance) has a vital role in atherosclerosis. Abnormal nutrient metabolism can affect the occurrence and development of atherosclerosis, and it has been suggested that atherosclerosis could be prevented and treated by adjusting the proportion of nutrients in the diet (Wei et al., 2021). For example, vesicles represent the main transport pathway for nutritional macromolecules (i.e., lipids and protein) and thus have a vital role in the establishment of the nutritional imbalance that is related to the development of atherosclerosis. The terms identified in our enrichment analyses, including “vesicle-mediated transport between endosomal compartments” (GO:0098927; Table 7) and “SNARE interactions in vesicular transport” (Table 6), confirmed this viewpoint. Furthermore, lipids have an important role in maintaining nutritional balance. Atherosclerosis is considered to be an inflammatory disease with chronic lipid accumulations.
and poor adaptability of arterial intima (So, 2018). Our sensitivity analysis identified CRBN, a negative regulator of AMPK in vivo, which is related to the metabolism of glucose and fat and is considered to be a regulator of body metabolism and energy homeostasis (Lee et al., 2013). We also identified IRAK1 and VAMP8 (using betweenness), which are related to the metabolism of lipids (Pushparaj et al., 2009; Singer et al., 2018). There is substantial evidence that lipids are related to atherosclerosis progression, as a high cholesterol level is a major risk factor in atherosclerosis and CVD, and fatty acid metabolism is closely related to endothelial function and protein homeostasis (Malekmohammad et al., 2021; Graham et al., 2021; Hasan and Fischer, 2021). Protein homeostasis is also relevant to macrophages in atherosclerosis (Koga et al., 2011). Our results demonstrated the role of protein as a nutrient in atherosclerosis. For instance, the top network marker (based on betweenness), RPL35, has a role in protein translation and endoplasmic reticulum docking (Wu et al., 2021). In addition, “aminoglycoside antibiotic metabolic process” (GO:0030647; Table 7) and “glycoside metabolic process” (GO:0016137; Table 7) further indicated a key role related to both lipids and proteins in atherosclerosis progression. As a result, the nutritional imbalance was identified as one of the driving factors behind atherosclerosis.

In the theory of inflamm-aging, the aging process leads to a chronic progressive inflammatory state, involving both the activation of innate immunity and accelerated pro-inflammatory response. As a result, the dynamic balance between pro- and anti-inflammatory cytokines is disrupted, and atherosclerosis was induced by the resulting immune homeostasis imbalance (Kahn, 2021). Our results further confirmed the key role of inflamm-aging in atherosclerosis and also identified dysfunction of endothelium and vessels, oxidative stress, and nutritional imbalance as risk factors. Studies have shown interactions among inflammation, nutrition, and endothelial and vascular homeostasis as well as oxidative stress. For instance, lipid accumulation is involved in the production of pro-inflammatory cytokines (Malekmohammad et al., 2021) and inflammation can affect lipid metabolism in macrophages (Osada et al., 2017). Moreover, oxidized LDL (ox-LDL) can remain in blood vessels, inducing dysfunction of the endothelium, which further responds to inflammatory signals, with a crucial role in atherosclerosis development (Malekmohammad et al., 2021). In addition, increased ROS production promotes endothelial dysfunction, leading to remodeling, platelet aggregation, loss of vasodilation, inflammation, and smooth muscle cell growth (Higashi et al., 2014). In our analysis, the “ribosome” pathway was identified as the most enriched KEGG pathway (Figure 4A), which highlighted the interactions among nutrition, vessels, and oxidative stress during the aging process (Stein et al., 2022). In summary, our work integrated potential mechanisms involving the inflammatory response, nutrition imbalance, and vascular homeostasis imbalance as well as oxidative stress, consistent with and contributing further to the inflamm-aging theory (Figure 6).
CONCLUSION

In this study, an ML method was used to model predictors of aging and disease and identify relative risk markers. The accelerated inflammatory pattern in atherosclerosis was verified by comparing inflamm-aging scores. The causal relationship between inflamm-aging and atherosclerosis was explored through MR, and then through sensitivity analysis, enrichment analysis, and network analysis; the results indicated that disorders of lipid homeostasis and dysfunction of endothelial cells and vascular homeostasis could promote the development of atherosclerosis. The role of oxidative stress was also confirmed. In summary, our work revealed that in the context of inflamm-aging, nutrition imbalance and vascular homeostasis, as well as oxidative stress, are risk factors that promote the development of atherosclerosis in a coordinated manner.

MATERIALS AND METHODS

Data and Pre-Processing

All gene expression data were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) in NCBI, including the GSE7638, GSE9820, GSE12288, GSE20681, GSE22253, GSE37171, and GSE123696 datasets. These datasets were from seven different platforms: GPL571, GPL6255, GPL96, GPL1433, GPL6244, GPL570, and GPL15207. The pre-processing steps for obtaining gene expression profiles were as follows:

1) Age and disease information of each sample (disease was encoded as 1, healthy as 0) was obtained, and samples without age were deleted.
2) The gene expression matrix for each dataset was integrated by summarizing the probe number within the gene symbol.
3) The total data matrix was integrated, and the missing gene expression values were filled with values of 0.
4) Genes with zero expression values ≥ 30% were deleted.
5) Gene expression profiles were transformed by the logarithmic transformation method in containing outliers (>3 times of standard deviation or <−3 times of standard deviation).
6) Based on the mean and the standard deviation of gene expression for healthy aged individuals, z-score normalization was performed for both normal aged samples and atherosclerosis samples.
7) The singular value decomposition method was used to eliminate inter-sample variation based on the top three principal components in healthy aged individuals.
8) The z-score was then used to normalize all individuals based on the mean and standard deviation of healthy aged individuals.

As a result, a total of 926 samples were obtained, including 313 samples from healthy aged individuals (age >50 years, 210 training datasets +103 test datasets), 207 samples from healthy young individuals (age ≤50, 140 + 67), 340 samples from aged atherosclerosis patients (age >50, 235 + 105), and 66 samples from young atherosclerosis patients (age ≤50, 45 + 21), with a total of 11,313 gene symbols.

Modeling the Improved Inflamm-Aging Predictor of Atherosclerosis

Gene expression profiles were divided into a training dataset and test dataset at a ratio close to 2:1, and the inflamm-aging prediction model was established by integrating aging, inflammation, and disease (atherosclerosis) markers. Aging and disease models were first constructed, respectively; then, an improved inflamm-aging model was developed by including inflamm-aging markers and adjusting the disease predictor (Supplementary Figure S5). The following steps were performed for the inflamm-aging model:

1) To summarize the interactions among key markers according to aging/disease, the gene expression profiles were transformed (or replaced) based on the results of our previous study (Wang et al., 2020b). The Pearson correlation coefficient was used to evaluate relevance and redundancy:

\[
\text{relevance} = \text{corr}(\text{gene}_i, \text{gene}_j, \text{phenotype}),
\]

\[
\text{redundancy} = \text{corr}(\text{gene}_i, \text{gene}_j).
\]

where the phenotype was set as 0 (control) versus 1 (disease) in the disease model, or as the transformation of the age in the aging model:

\[
1/(1 + \exp(- (\text{age} - 50)/50)).
\]

The interaction is summarized as follows:

\[
\text{interaction}_i = \log(2^{n+1}) \cdot \text{corr}(\text{gene}_i, \text{gene}_j, \text{phenotype}) \cdot \text{gene}_i \cdot \text{gene}_j,
\]

where only the first log2 \((n+1)\) interactions for each gene were included.

2) The 11,313 gene profiles (training dataset) were sorted by the ReliefF algorithm.
3) The kNN \((k = 5\) with the cosine distance), SVM, NB, and ensemble learning (established 100 decision tree models) algorithms were used to select key markers (in the aging or disease model) based on the training dataset and to evaluate the performance on the test dataset, with the help of 10-fold cross-validation. The kNN \((k = 5\) with the cosine distance) algorithm was selected, and 441 aging markers and 390 candidate disease markers were identified in this step (Supplementary Tables S3, S4).
4) The inflamm-aging model was constructed by integrating the aging, inflammation, and disease markers as follows:

6) 568 candidate inflammatory markers were obtained based on the GO BPs using the gene set enrichment analysis (GSEA) platform (http://software.broadinstitute.org/gsea/downloads).
jsp, version 7.4, keyword “INFLAMMATORY”); then, aging markers or candidate disease markers overlapping with candidate inflammatory markers were removed.

② Aging markers and candidate disease markers were further filtered using the Kruskal–Wallis test ($p < 0.05$ and FDR < 0.2).

③ Inflamm-aging markers were identified by transforming selected inflammatory markers combined with the aging markers [also based on a previous study (Wang et al., 2020b) as in step (Lopez et al., 2021)], where an interaction was retained if $p < 0.05$ and FDR < 0.2 using the Kruskal–Wallis test.

④ Disease markers were also transformed based on the inflammatory markers.

⑤ The kNN method, ReliefF algorithm, and 10-fold cross-validation were used to reselect the optimized integrative disease markers, where some of the top candidate 500 disease markers were replaced by disease markers related to inflammm-aging in step ③. Therefore, the improved inflammm-aging predictor contained 116 aging markers, 186 inflammatory markers, and 316 disease markers (Supplementary Table S5).

⑥ Other classifiers including SVM, NB, and the ensemble algorithm were also used for comparison of their prediction abilities.

### Calculating the Inflamm-Aging Score

The steps to calculate the inflammm-aging score for each sample were as follows:

1) The distance between healthy young and old neighbors was calculated using the original inflammm-aging score based on the inflammatory markers in the inflammm-aging model; the cosine distance was also used:

$$ \text{inflamm} - \text{aging} \_\text{score} = \sum_{i=1}^{q} \text{dis tan ce}_{i, \text{young}} - \sum_{i=1}^{q} \text{dis tan ce}_{i, \text{old}}. $$ (5)

2) The original score was adjusted by age:

$$ \text{adjust} \_\text{score} = \text{score} - b \times \text{transformed \_age}, $$ (6)

where $b$ is the regression coefficient for an aging score, and the transformed age is calculated as follows:

$$ \text{tanh}\left(\frac{\text{age} - 50}{50}\right). $$ (7)

3) Both the original and adjusted inflammm-aging scores were used to test the accelerated pattern in atherosclerosis based on the Kruskal–Wallis test.

4) The atherosclerosis (disease) score was calculated based on the identified disease markers (with the same details as for the inflammm-aging score) in the improved inflammm-aging model for further (sensitivity/network) analysis.

### Causality Analysis Based on MR

MR is a statistical method used to evaluate the causal relationship between risk factors and results based on observed data (Burgess et al., 2019; Burgess et al., 2020). The causal relationships among the instrumental variable, the risk factor, and the outcome variable were evaluated as follows:

1) There was a correlation between the instrumental variable and the risk factor.
2) There was no correlation between the instrumental variable and the confounding factor.
3) There was no correlation between the instrumental variable and the outcome variable after deleting the effect from the risk factor.

In this work, aging markers were used as instrumental variables, inflammatory markers were used as risk factors, and inflammm-aging markers were used as outcome variables to explore the causal relationship between inflammm-aging and atherosclerosis. The steps were as follows:

1) The training dataset was divided into three groups: low inflammation in health (“inflammation− atherosclerosis−”), high inflammation in health (“inflammation+ atherosclerosis−”), and low inflammation in disease (“inflammation− atherosclerosis+”).

2) For each group, the Pearson correlation coefficient between each aging marker and the inflammm-aging score was calculated to evaluate the correlation between the instrument variable and the risk factor. The relationship was retained if the correlation for “inflammation− atherosclerosis−” was between those of “inflammation− atherosclerosis−” and “inflammation+ atherosclerosis+” (meaning that the atherosclerosis was triggered by the inflammation).

3) To filter out the correlation between the instrumental variable and confounding factors, a permutation test was performed by generating the simulated inflammm-aging score from the same number of randomly selected markers; this process was repeated 1,000 times, then the $p$-value was calculated as the proportion of occurrence times (larger than the real difference) of the absolute difference between “inflammation+ atherosclerosis−” and “inflammation+ atherosclerosis+” in 1,000 permutations. The relationship between each aging marker and the inflammm-aging score was retained if the permutation $p < 0.05$.

4) If the gene was within both the aging marker set and the disease marker set, then it was removed.

5) For each group, the Pearson correlation coefficient between each aging marker and disease marker was calculated to evaluate the correlation between the instrument variable and the outcome variable; and the partial Pearson correlation coefficient between each aging marker and disease marker based on the inflammm-aging score was calculated to evaluate
the correlation between the instrument variable and the outcome variable without the background of the risk factor. The causal relationship between each aging marker and disease marker was retained if the difference between the correlation and the partial correlation for “inflammation–atherosclerosis” was between those for “inflammation+atherosclerosis−” and “inflammation+atherosclerosis+” (meaning that the atherosclerosis was triggered by the inflammation).

6) To filter out the potential correlation between confounding factors and the outcome variable, a permutation test was performed by generating the simulated aging marker and disease marker from the same number of randomly selected markers, repeated 1,000 times; then, the permutation P-value was calculated using the same details as in step (Pahwa and Jialal, 2021).

7) To filter out potential weak instrumental variables, the aging markers were further analyzed by evaluating the correlation between the inflamm-aging score and each aging marker, using the result of the point multiplication between the inflamm-aging score and each aging marker; then, the residual of the inflamm-aging score was tested between the atherosclerosis and control subgroup, using the Kruskal–Wallis test ($p < 0.05$ and FDR < 0.2).

\[
\text{residual} = \text{inflamm} – \text{aging_score} – b_s(\text{inflamm} – \text{aging_score}\cdot\text{aging_marker}), \tag{8}
\]

where $b_s$ is the regression coefficient.

8) To filter out the effect of horizontal pleiotropy, the aging–disease causal relationship was further examined by comparing the correlation between each aging and disease marker, through the inflamm-aging score or otherwise. Steps ①–⑤ were used to calculate the correlations between instrumental variables and outcome variables without the background of the risk factor, and step ⑥ was used to calculate the correlations between instrumental variables and outcome variables without the context of the risk factor.

① The residual of each disease marker (“residual A”) is calculated based on the inflamm-aging score:

\[
\text{residual}_A = \text{disease_marker} – b_1(\text{disease_marker}\cdot\text{inflamm} – \text{aging_Gore}). \tag{9}
\]

where $b_1$ is the regression coefficient.

② The residual of each aging marker (“residual B”) is calculated based on the inflamm-aging score:

\[
\text{residual}_B = \text{disease Marker} – b_2(\text{disease.Marker}\cdot\text{aging_Marker}), \tag{10}
\]

where $b_2$ is the regression coefficient.

③ The aforementioned two residuals were further compared, and the residual of the disease marker is calculated (as “residual C”):

\[
\text{residual}_C = \text{residual}_A – b_3\cdot\text{residue} – B, \tag{11}
\]

where $b_3$ is the regression coefficient.

④ The residual of the disease marker (“residual D”) was calculated based on the aging marker.

⑤ The difference (between “residual C” and “residual D”) was tested between the atherosclerosis and control subgroups, using the Kruskal–Wallis test ($p < 0.05$ and FDR < 0.2).

Finally, the causal relationship between the aging marker and disease marker (through the inflamm-aging score) was retained. Thus, 1,340 aging–disease pairs were identified as causal relationships, including 116 aging markers and 312 inflamm-aging markers.

**Constructing the Inflamm-Aging Differential Network**

In order to further reveal the relationship between inflamm-aging and atherosclerosis, an inflamm-aging differential network was constructed by the following steps:

1) To compare the relationship of each gene pair in the context of inflamm-aging, both the Pearson correlation coefficient for each pair of genes and the partial correlation coefficient based on the inflamm-aging score were calculated based on the atherosclerosis and control groups, respectively.

2) The Benjamini–Hochberg FDR method was used to adjust the P-values of the correlation coefficient and the partial correlation coefficient.

3) The differences in the correlations and partial correlations between the atherosclerosis and control groups were calculated.

4) The edge between each gene pair was retained if the signs of the difference in values for the correlation and partial correlation were different, and values of $p < 0.05$ and FDR < 0.1 were obtained in step (Pahwa and Jialal, 2021).

5) The scale-free characteristics of the inflamm-aging acceleration differential networks were verified by the power-law distribution.

6) The shortest path between each pair of aging and inflamm-aging markers was selected based on each aging acceleration differential network using the Dijkstra algorithm.

7) The network constructed based on the training data was used for further analysis (identifying shortest paths, exploring biological functions between aging and atherosclerosis, etc.), and the network constructed based on the test data was used to validate the training network.

**Global Sensitivity Analysis Using the MCMC Method**

In order to further explore the immune homeostasis imbalance in atherosclerosis in the context of inflamm-aging, an overall sensitivity analysis based on the MCMC method was carried out, where the aging–disease pairs identified by MR were used as candidate relationships to be further evaluated. The MCMC method is used for sampling from certain posterior
where the genes identified using the shortest pathways between known genes and candidate genes, the number of common genes between known genes and candidate genes, the number of known genes (KEGG pathways or BP terms), and the hypergeometric distribution. The steps are as follows.

1) Construct a transition kernel of an ergodic Markov chain. In this step, the prior distribution for each parameter was the normal distribution based on the identified aging inflammation and inflam-aging markers for each group ("inflammation− atherosclerosis−", "inflammation+ atherosclerosis−", and "inflammation+ atherosclerosis+").

2) Simulate the chain until it reaches equilibrium. The steps are as follows.

3) Perform global sensitivity analysis. In this step, the distribution (ratio) of each aging−disease pair is calculated as follows:

$$\sum \text{sign(aging marker}\ast \text{disease marker}). \quad (13)$$

In addition, if the value for "inflammation− atherosclerosis−" was between those of "inflammation+ atherosclerosis−" and "inflammation+ atherosclerosis+", then the aging−disease pair was identified as a sensitive relationship for further network analysis. As a result, 442 aging−disease pairs were identified as causal relationships, including 93 aging markers and 81 disease markers.

**Enrichment Analysis**

Gene functions with significant correlations were further explored through enrichment analysis of the shortest paths. GO terms and KEGG pathways were identified using the GSEA platform (http://software.broadinstitute.org/gsea/downloads.jsp, version 7.4). The hypergeometric distribution was used to test the enrichment of GO BP terms and KEGG pathways using the following formula:

$$P(X \geq x) = 1 - \sum_{k=0}^{x-1} \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}. \quad (14)$$

where \(N\) is the total number of genes in the gene set, \(M\) is the number of known genes (KEGG pathways or BP terms), \(n\) is the number of genes identified in each shortest path, and \(k\) is the number of common genes between known genes and candidate genes identified in each aging−disease shortest path. The \(p\)-value for each path was controlled using the Benjamini–Hochberg method. Finally, the paths with \(p < 0.05\) and \(FDR<0.1\) were retained. The enrichment score is calculated as follows:

$$\text{score} = \sum_{FDR<0.1} (1 - FDR). \quad (15)$$

**Identifying Network Markers**

The subnetwork with the shortest pathways among the selected aging−disease pairs was constructed (after MR or MCMC), and genes in the subnetwork were sorted by their betweenness in descending order. To test whether the top betweenness genes were hubs in the background network, we ran a permutation to count the occurrences of the top genes in the shortest paths between randomly selected genes (containing the same numbers of aging−disease pairs, 1,340 in MR or 442 in MCMC) when they had greater betweennesses than those in our study. We repeated this process 1,000 times, and the \(p\)-value was calculated as the proportion of occurrences of the top betweenness genes in 1,000 permutations.

**DATA AVAILABILITY STATEMENT**

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding author.

**AUTHOR CONTRIBUTIONS**

YH, YC, LY, JW, and YW performed the algorithm and analyzed the data; YH, YC, LY, and YW wrote the manuscript; XS and YW designed and sponsored the study. All authors read and approved the manuscript.

**FUNDING**

This work was supported by the National Natural Science Foundation of China (32000478 to YW) and the National Key R&D Program of China (2016YFC0901704, 2017YFA0505500, 2017YFC0907505, and 2017YFC0908405). The funders had no role in study design, data collection, and analysis, the decision to publish, or the preparation of the manuscript.

**ACKNOWLEDGMENTS**

We thank Charlesworth Author Services (http://www.charlesworthauthorservices.com/~Frontiers) for its linguistic assistance during the preparation of this manuscript.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.865827/full#supplementary-material
Supplementary Figure S1 | ML results for control methods. (A, B) Traditional aging predictor. (C, D) Traditional atherosclerosis predictor. (E, F) Disease predictor from our previous study. (A, C, E) Learning curve for the training dataset. (B, D, F) ROC curve for the test dataset.

Supplementary Figure S2 | PCA plots of the inflammm-aging model based on (A) the inflmm-aging score and disease risk score; and (B) the first two principal components. Blue circle, normal test set; blue ‘+’, normal training set; red circle, atherosclerosis test set; red ‘+’, atherosclerosis training set.

Supplementary Figure S3 | Degree distribution of the inflmm-aging differential network.

Supplementary Figure S4 | Heatmap of the bipartite graph between aging markers and disease markers based on MCMC. The x-axis represents disease markers; the Y-axis represents aging markers.

Supplementary Figure S5 | Workflow of the improved inflmm-aging model.

REFERENCES

Adairkirk, T., and Senior, R. (2008). Fragments of Extracellular Matrix as Mediators of Inflammation. Int. J. Biochem. Cell Biol. 40 (6–7), 1101–1110. doi:10.1016/j.biocel.2007.12.005

Alalawi, H. H., and Alsuwai, M. S. (2021). Detection of Cardiovascular Disease Using Machine Learning Classification Models. IJERT 07, 14.

Alber, M., Buganza Tepole, A., Cannon, W. R., De, S., Dura-Bernal, S., Garikipati, K., et al. (2019). Integrating Machine Learning and Multiscale Modeling-Perspectives, Challenges, and Opportunities in the Biomedical, Biomedical, and Behavioral Sciences. npj Digit. Med. 2, 115. doi:10.1038/s41746-019-0193-y

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). Molecular Biology of the Cell. 4th edition. New York: Garland Science.

Alique, M., Ramírez-Carracedo, R., Bodega, G., Carracedo, J., and Ramírez, R. (2018). Senescent Microvesicles: A Novel Advance in Molecular Mechanisms of Atherosclerotic Calcification. Int. J. Mol. Sci. 19 (7), 2033. doi:10.3390/ijms19072003

Bennett, M. R., Sinha, S., and Owens, G. K. (2016). Vascular Smooth Muscle Cells in Atherosclerosis. Circ. Res. 118 (4), 692–702. doi:10.1161/CIRCRESAHA.115.306361

Berghéna, S. C., Bodde, M. C., and Jukema, J. W. (2017). Pathophysiology and Treatment of Atherosclerosis. Neth Heart J. 25 (4), 231–242. doi:10.1007/s12471-017-0959-2

Beyfuss, K., and Hood, D. A. (2018). A Systematic Review of P53 Regulation of the inflmm-aging model. Circulation 138 (12), 16248–16266. doi:10.1161/circulation.2018.031534

Chen, Q., Hao, M., Wang, L., Li, L., Chen, Y., Shao, X., et al. (2021). Prefused lysosomes cluster on autophagosomal regulated by VAMP6. Cell Death Dis. 12, 939. doi:10.1038/s41419-021-04243-0

Chen, Y., Zhao, H., Xiao, Y., Shen, P., Tan, L., Zhang, S., et al. (2021). Pan-cancer Analysis Reveals an Immunological Role and Prognostic Potential of p53 in Human Cancer. Aging 13 (12), 16248–16266. doi:10.1080/2018.031534

Chen, Q., Hao, M., Wang, L., Li, L., Chen, Y., Shao, X., et al. (2021). Prefused lysosomes cluster on autophagosomal regulated by VAMP6. Cell Death Dis. 12, 939. doi:10.1038/s41419-021-04243-0

Chen, Y., Xiao, Y., Shen, P., Tan, L., Zhang, S., et al. (2021). Pan-cancer Analysis Reveals an Immunological Role and Prognostic Potential of p53 in Human Cancer. Aging 13 (12), 16248–16266. doi:10.1080/2018.031534

Chi, C., Li, D. J., Jiang, Y. J., Tong, J., Fu, H., Wu, Y. H., et al. (2019). Vascular Smooth Muscle Cell Senescence and Age-Related Diseases: State of the Art. Biochim. Biophys. Acta Mol. Basis Dis. 1865, 1810–1821. doi:10.1016/j.bbadis.2018.08.015

Childs, B. G., Gluscevic, M., Baker, D. J., Laberge, R. M., Marquess, D., Danenberg, J., et al. (2017). Senescent Cells: an Emerging Target for Diseases of Ageing. Nat. Rev. Drug. Discov. 16, 718–735. doi:10.1038/nrd.2017.116

Chothani, S., Schafer, S., Adami, E., Viswanathan, S., Widjaja, A. A., Langley, S. R., et al. (2019). Widespread Translational Control of Fibrosis in the Human Heart by RNA-Binding Proteins. Circulation 140 (11), 937–951. doi:10.1161/CIRCULATIONAHA.119.039596

Covarrubias, A. J., Perrone, R., Grozio, A., and Verdin, E. (2021). NAD+ Metabolism and its Roles in Cellular Processes during Ageing. Nat. Rev. Mol. Cell Biol. 22 (2), 119–141. doi:10.1038/s41580-020-00313-x

Critchley, B. J., Isalan, M., and Mielcarek, M. (2018). Neuro-Cardio-Mechanisms in Huntington’s Disease and Other Neurodegenerative Disorders. Front. Physiol. 9:338. doi:10.3389/fphys.2018.00359

Davignon, J., and Ganz, P. (2004). Role of Endothelial Dysfunction in Atherosclerosis. Circulation 109 (23 Suppl. 1), III27–32. doi:10.1161/01.CIR.000131515.03336.08

De Stefano, A., Caporalì, S., Di Daniela, N., Novella, V., Cardillo, C., Schinzari, F., et al. (2021). Anti-Inflammatory and Proliferative Properties of Luteolin 7-O-Glucoside. Int. J. Mol. Sci. 22 (3), 1321. doi:10.3390/ijms22031321

Deshpande, V., Sharma, A., Mukhopadhayay, R., Thota, L. N. R., Ghatge, M., Mangala, R. K., et al. (2016). Understanding the Progression of Atherosclerosis through Gene Profiling and Co-expression Network Analysis in ApoB tm25Gy Ldr tm1Her Double Knockout Mice. Genomics 107 (6), 239–247. doi:10.1016/j.ygeno.2016.04.007

Dingjan, I., Paardekooper, L. M., Verboogen, D. R. J., von Mollard, G. F., Ter Beest, W., Deshpande, V., Thota, L. N. R., Ghatge, M., et al. (2016). Understanding the Progression of Atherosclerosis through Gene Profiling and Co-expression Network Analysis in ApoB tm25Gy Ldr tm1Her Double Knockout Mice. Genomics 107 (6), 239–247. doi:10.1016/j.ygeno.2016.04.007

Figueroa, R. J., Carrasco-Avino, G., Ichikawa, I. A., Lange, M., Owen, G. I., Siekmann, A. F., et al. (2017). Reprimo Tissue-specific Expression Pattern Is Conserved between Zebrabfish and Human. PLoS One 12 (5), e0178274. doi:10.1371/journal.pone.0178274

Fogoros, R. N. (2020). “An Overview of Atherosclerosis,” in Verywell Health. Editor Y. S. Ali.
