Frontiers | Frontiers in Molecular Biosciences

Check for updates

Exploring the shared gene signatures of smoking-related osteoporosis and chronic obstructive pulmonary disease using machine learning algorithms

Haotian Wang[1†], Shaoshuo Li[2†], Baixing Chen[3], Mao Wu[1,2], Heng Yin[1,2], Yang Shao[2*‡] and Jianwei Wang[1,2*‡]

[1]Graduate School of Nanjing University of Chinese Medicine, Nanjing, China, [2]Department of Traumatology and Orthopedics, Wuxi Affiliated Hospital of Nanjing University of Chinese Medicine, Wuxi, China, [3]Department of Development and Regeneration, University of Leuven, Leuven, Belgium

Objectives: Cigarette smoking has been recognized as a predisposing factor for both osteoporosis (OP) and chronic obstructive pulmonary disease (COPD). This study aimed to investigate the shared gene signatures affected by cigarette smoking in OP and COPD through gene expression profiling.

Materials and methods: Microarray datasets (GSE11784, GSE13850, GSE10006, and GSE103174) were obtained from Gene Expression Omnibus (GEO) and analyzed for differentially expressed genes (DEGs) and weighted gene co-expression network analysis (WGCNA). Least absolute shrinkage and selection operator (LASSO) regression method and a random forest (RF) machine learning algorithm were used to identify candidate biomarkers. The diagnostic value of the method was assessed using logistic regression and receiver operating characteristic (ROC) curve analysis. Finally, immune cell infiltration was analyzed to identify dysregulated immune cells in cigarette smoking-induced COPD.

Results: In the smoking-related OP and COPD datasets, 2858 and 280 DEGs were identified, respectively. WGCNA revealed 982 genes strongly correlated with smoking-related OP, of which 32 overlapped with the hub genes of COPD. Gene Ontology (GO) enrichment analysis showed that the overlapping genes were enriched in the immune system category. Using LASSO regression and RF machine learning, six candidate genes were identified, and a logistic regression model was constructed, which had high diagnostic values for both the training set and external validation datasets. The area under the curves (AUCs) were 0.83 and 0.99, respectively. Immune cell infiltration analysis revealed dysregulation in several immune cells, and six immune-associated genes were identified for smoking-related OP and COPD, namely, mucosa-associated lymphoid tissue lymphoma translocation protein 1 (MALT1), tissue-type plasminogen activator (PLAT), sodium channel 1 subunit alpha (SCNN1A), sine oculis homeobox 3 (SIX3), sperm-associated antigen 9 (SPAG9), and vacuolar protein sorting 35 (VPS35).

Conclusion: The findings suggest that immune cell infiltration profiles play a significant role in the shared pathogenesis of smoking-related OP and COPD. The results could provide valuable insights for developing novel therapeutic strategies for managing these disorders, as well as shedding light on their pathogenesis.

1 Introduction

Persistent airflow llimitation is a distincitve time-developing symbol of COPD, a common preventable and treatable disease (Lippi et al., 2022). It has been found that there is a causation between enhanced chronic inflammatory and COPD, especially with a plus of the effect of cigarette smoking (Brightling and Greening, 2019). COPD is regarded as a systemic disease that is accompanied byseveral comorbid conditions, such as lung cancer, muscle wasting, diabetes, atherosclerosis, orthostatic hypotension, and anxiety/ depression. (Romme et al., 2015; Sarkar et al., 2015). The clinical management of these comorbidities is crucial owing to their high correlation to the rates of hospitalization, mortality, and reduced health-related quality of life are commonly observed in patients with COPD (Frei et al., 2014).

Osteoporosis, a systemic bone disease, is considered as one of the significant comorbidities associated with COPD (Kanis et al., 2008). Recent epidemiological evidence suggests a high prevalence of OP in individuals with COPD, despite the absence of a clearly established causal or molecular link between the two disorders (Graat-Verboom et al., 2009; Lehouck et al., 2011; Regan and Jaramillo, 2012; Watanabe et al., 2015). Retrospective chart reviews of 234 male patients with OP from a single bone clinic revealed that COPD accounted for the leading cause of secondary OP (Ryan et al., 2011). Smoking, however, has been found to be a risk factor for both OP and COPD, according to recent studies (Ward and Klesges, 2001; Kanis et al., 2005; Hikichi et al., 2019; Li et al., 2022). Smoking has been shown to cause changes in the microstructure of trabecular bones, as well as reduce the resistance of skeletal muscles to mechanical stress and loading. (Brook et al., 2012). Studies have provided evidence of a functional interplay between bone and immune cells, particularly involving activated T cells and Th17 cells (Kong et al., 1999; Sato et al., 2006; Chen et al., 2016). Additionally, interleukin-17 A expression has been demonstrated to increase in the airways of patients with COPD and correlates with decreased lung function in these patients (Di Stefano et al., 2009; Eustace et al., 2011). In light of these findings, IL-17 A could serve as a common mechanism linking smoking-related OP and COPD. (Xiong et al., 2020).

The occurrence of OP in patients with COPD is asymptomatic and is often undiagnosed until the occurrence of bone fractures. Notably, fractures resulting from OP can further impair the pulmonary function of patients with COPD. Thus, the interdependence between COPD and OP gives rise to a deleterious cycle that imposes a considerable burden on affected individuals. As the occurrence of OP in patients with COPD is immensely underrated, it is essential to elucidate the pathogenesis of OP in COPD, and the early diagnosis of patients with COPD who are at a high risk of OP should be emphasized. The rapid progress in high-throughput microarray technologies has enabled the identification of putative novel biomarkers, genetic variations, and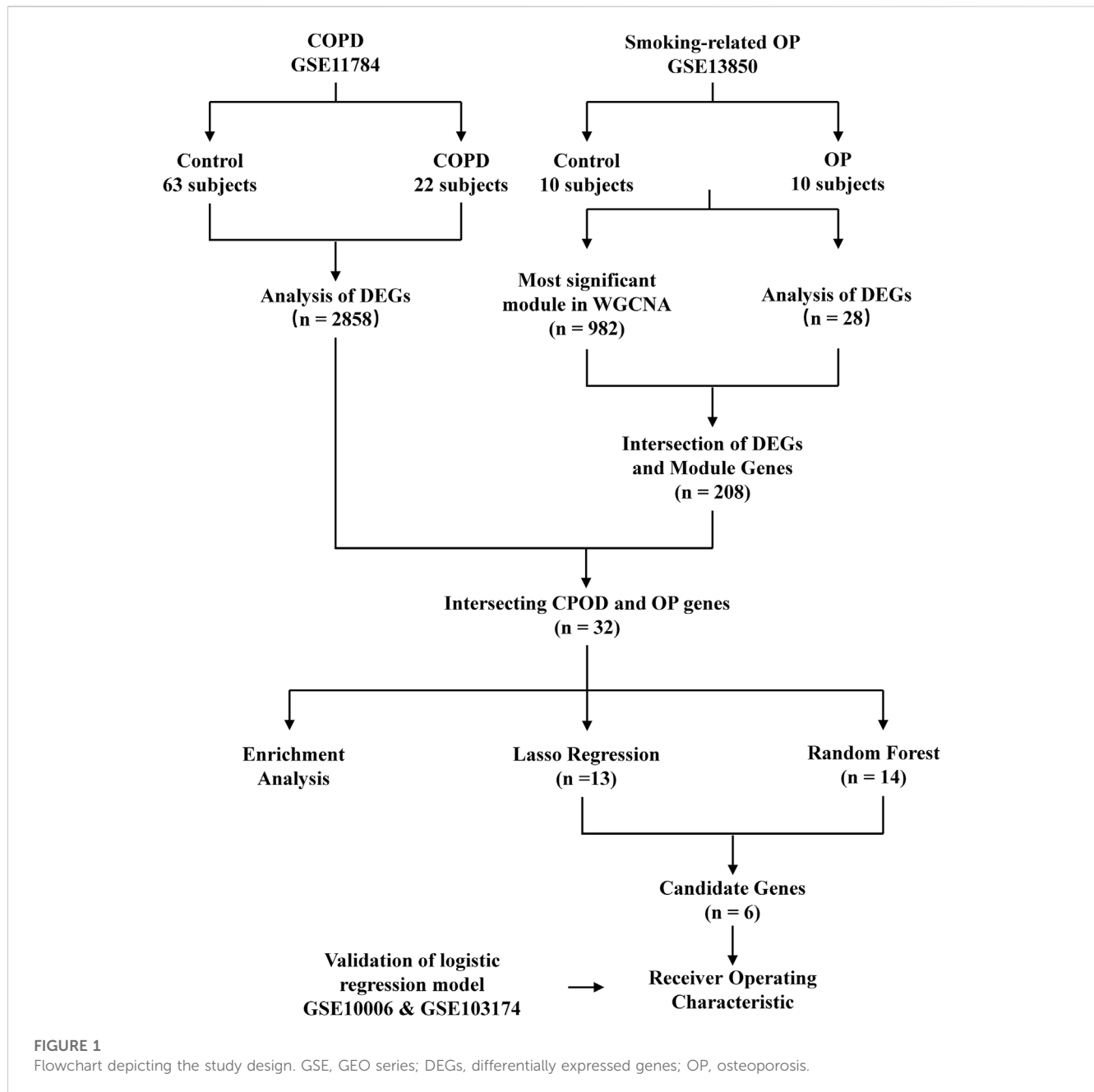 biological pathways, thereby enhancing our comprehension of the pathogenesis and therapeutic intervention of various ailments (Li et al., 2021; Zhou Y et al., 2022). WGCNA (Langfelder and Horvath, 2008; Abu-Jamous and Clust, 2018; Zhou Z et al., 2022) and machine learning techniques are gradually improving, and these bioinformatics tools can be employed for providing great prospects for identifying the potential molecular mechanisms, biomarkers, and therapeutic targets for smoking-related OP and COPD, as well as other complex diseases. These technologies enable researchers to explore large datasets and identify key molecular pathways and biomarkers, leading to the discovery of new therapeutic interventions and improved patient outcomes (Kumar et al., 2021; Chen et al., 2022).

As far as we are aware, there have been few investigations targeting the identification of immune-related biomarkers for the diagnosis of smoking-related OP and COPD and there is a scarcity of studies on the application of machine learning approaches for the diagnosis of these diseases. Smoking-related OP and COPD dataset was obtained from the GEO database and used for the identification of candidate genes in this study. In this study, we employed machine learning algorithms to identify feature genes from candidate genes and validated them using ROC curves. Our study identified potential immune-related diagnostic biomarkers for smoking-related OP in COPD patients, which could facilitate the early diagnosis and personalized treatment of COPD. Our work not only provides a approach for identifying potential biomarkers for COPD but also highlights the value of machine learning algorithms in identifying key genes for complex diseases. The findings of this study have potential implications for improving the clinical management and outcome of COPD patients, and may also provide a framework for the development of similar approaches for other complex diseases.

2 Materials and methods

2.1 Microarray data

The smoking-related OP and COPD datasets were retrieved from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) (Edgar et al., 2002) for analyzing the gene expression data in smoking-related OP and COPD. The search strategy employed in this study is described hereafter. The gene expression profiles generated by array analyses were initially retrieved from the GEO database. The OP and COPD datasets, containing samples obtained from blood and the small airway epithelium, respectively, were retrieved from the GEO database. Datasets containing samples of control groups and samples corresponding to *Homo sapiens* were also retrieved from the GEO database. The GSE11784 dataset (Tilley et al., 2011), generated using the GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array platform, contained

FIGURE 1
Flowchart depicting the study design. GSE, GEO series; DEGs, differentially expressed genes; OP, osteoporosis.

the expression data of 22 smokers with COPD and 63 healthy subjects. The GSE13850 dataset, also generated using the GPL96 [HG-U133 A] Affymetrix Human Genome U133 A Array platform, contained the expression data of 10 smokers with low bone mineral density (BMD) and 10 healthy subjects. Two validation datasets were subsequently retrieved from the GEO database; one dataset was extracted from the GSE10006 dataset (Carolan et al., 1950), also generated using the GPL570 array platform, and contained the expression data of 54 smokers with COPD and 22 healthy subjects. The other validation dataset was extracted from the GSE103174 dataset and contained the gene expression data of 10 healthy subjects and 23 smokers with COPD. The study did not require the

approval of an ethics committee or informed consent as these data are publicly available (Figure 1).

2.2 Data processing and analysis of differential gene expression

The DEGs were identified using the limma package in R (version 4.1.1; http://cran.r-project.org/) (Ritchie et al., 2015), and subsequently analyzed using different packages. The initial normalization of the data was performed by applying the normalizeBetweenArrays function from the limma package in R. Subsequently, probe IDs were transformed into gene symbols. The

lmFit and eBayes functions were used to calculate the adjusted p-values and |log2 (fold change)| of the genes, respectively. DEGs were identified based on the criteria of |log2 (fold change)| > 1.5 and adjusted p-value < 0.05.

2.3 WGCNA and selection of module genes

The WGCNA system biology strategy was used to investigate the gene-gene correlations using the WGCNA package in R (version 4.1.1; http://cran.r-project.org/), (Langfelder and Horvath, 2008). The gene expression profiles were subjected to Median Absolute Deviation (MAD) calculation for each gene. The calculation steps for MAD are as follows: first, calculate the median of the dataset, and then take the median of the absolute deviation of each data point to obtain the value of MAD. To eliminate the top 30% genes with the smallest MAD, and to remove outlier genes and samples, the goodSamplesGenes function in the WGCNA package of R was utilized. Subsequently, a scale-free co-expression network was constructed using the WGCNA package. Initially, Pearson's correlation matrices were established to examine all pairwise gene relationships, with the average linkage approach. The correlation strength between gene pairs was used to build a weighted adjacency matrix through the following power function: $A\_mn = |C\_mn|β$; where, C_mn represents the Pearson's correlation between genes m and n, A_mn represents the adjacency between genes m and n, and $β$ is a soft-thresholding parameter that emphasizes strong gene correlations while downplaying weak ones. We set $β$ to 4 and converted the adjacency matrix to a topological overlap matrix (TOM) that measured gene connectivity. A dissimilarity matrix (1-TOM) was subsequently produced. To generate gene modules, we used an average linkage hierarchical clustering method with the TOM-based dissimilarity metric to group genes with similar expression profiles. The dendrogram's gene groupings were subjected to an adjustment, setting the minimum size for each cluster to 30 while the sensitivity was established at 4. Subsequent to the clustering of modules, further analysis was conducted by computing the dissimilarity of module eigengenes. A cut-line was selected for merging the module dendrogram, and some of the modules were merged. A cut-line of 0.2 was selected for merging the modules and a total of 18 co-expression modules were finally obtained. It is worth mentioning that the grey module encompassed a cluster of genes that lacked any definitive assignment to any module (Alderden et al., 2018).

2.4 Functional enrichment analysis

The Gene Ontology (GO) resource provides structured computable data regarding the functions of genes and gene products (Consortium, 2019). The GO resource is widely used along with the Kyoto Encyclopedia of Genes and Genomes (KEGG) database for analyzing the functions of genes (Kanehisa and Goto, 2000). In this study, functional enrichment analysis was performed using the clusterProfiler package in R (Yu et al., 2012) and the results were visualized with Sangerbox (Shen et al., 2022). GO and KEGG analyses were performed for identifying the DEGs in OP that overlapped with the most significant module genes of OP,

and the intersecting DEGs were defined as the OP set. The genes associated with smoking-related OP that overlapped with the DEGs in COPD were subjected to GO and KEGG enrichment analyses. The intersecting DEGs were subsequently used for screening the candidate genes.

2.5 Application of machine learning algorithms

The least absolute shrinkage and selection operator (LASSO) regression method is widely used for selecting and regularizing variables for increasing the predictive accuracy and comprehensibility of statistical models (Yang et al., 2018). The random forest (RF) method provides an effective approach for predicting continuous variables and obtaining reliable forecasts owing to its higher accuracy, sensitivity, specificity, and no limits on variable conditions (Ellis et al., 2014). LASSO regression and RF analyses were performed in this study using the glmnet (Zhang et al., 2019) and randomForest packages, respectively, in R. The genes that were identified by both LASSO and RF were selected as predictor genes for constructing the logistic regression models.

2.6 Construction and validation of logistic regression model

A logistic regression model was constructed based on the identified predictor genes using the glmnet package in R (Friedman et al., 2010). The GSE10006 and GSE103174 validation datasets were used for evaluating the accuracy of the model based on the receiver operating characteristic (ROC) curve and the area under the curve (AUC). The results were visualized using the pROC package in R (Robin et al., 2011), and AUCs >0.8 were regarded as ideal.

3 Results

3.1 Identification of DEGs

Using the limma package in R, a total of 2,858 DEGs were detected in the COPD dataset. Of these, 2,596 genes were found to be upregulated while 262 were downregulated (Figure 2A, B). Additionally, a total of 280 DEGs were identified in the OP dataset, of which 109 and 171 genes were upregulated and downregulated, respectively (Figure 3A, B).

3.2 WGCNA and identification of key modules

The soft threshold power was set at a value of 4, and the scale-free topological index was established at 0.85. The resultant gene tree and corresponding module colors were generated and visually represented using distinct color schemes (Figure 3C, E), in 18 gene modules. The correlation between smoking-related OP and the gene modules is depicted in Figure 3D. The turquoise

FIGURE 2
Heatmap and volcano plot of the DEGs in the COPD dataset. **(A)** The rows depict the DEGs, and the columns represent the respective COPD cases or control samples. The upregulated and downregulated DEGs are depicted in red and blue, respectively. **(B)** The upregulated and downregulated DEGs are depicted in red and green, respectively.
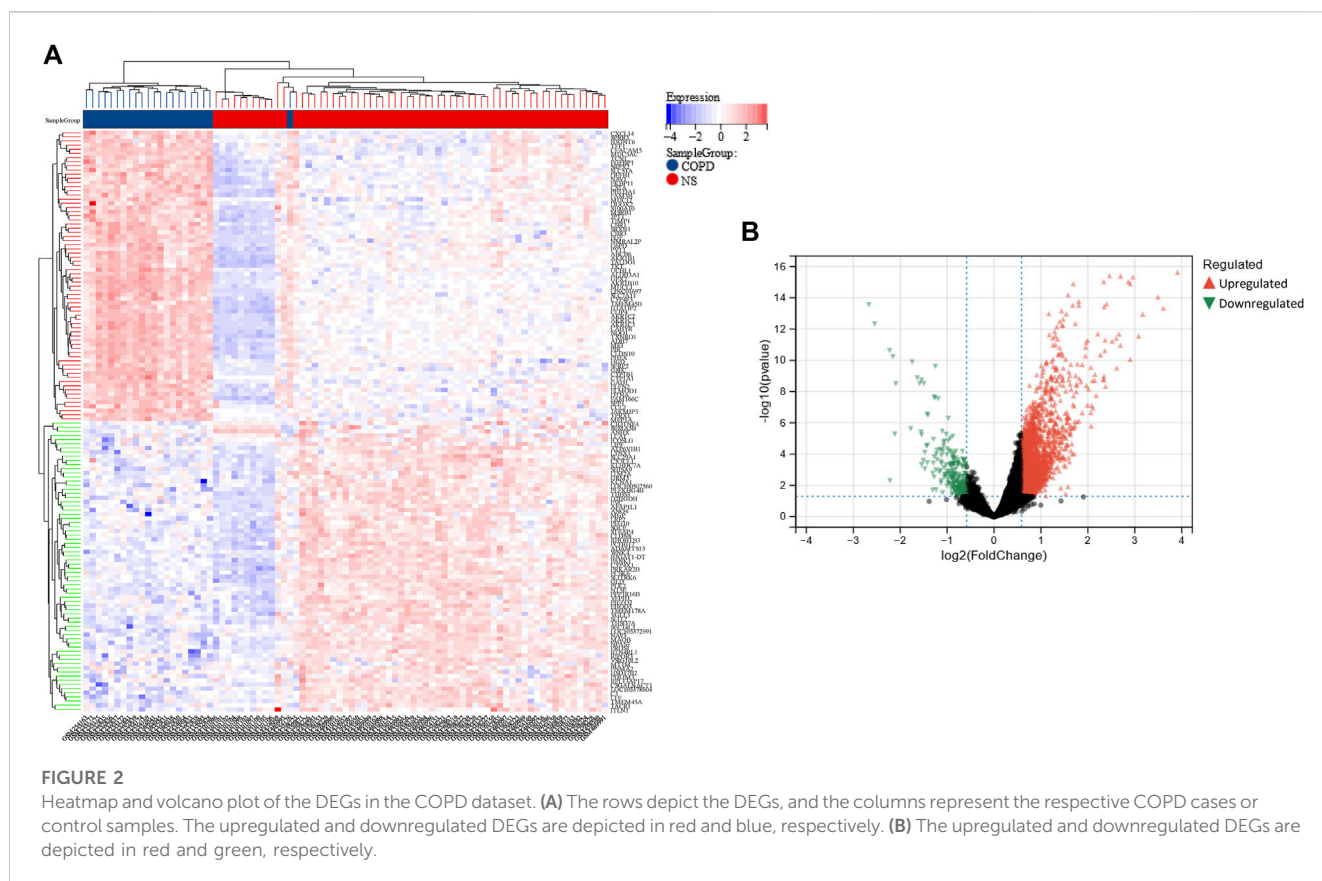
module, which comprised 982 genes, exhibited the highest correlation with OP (correlation coefficient = −0.98, p = 2.4 e-13) and this module was regarded as the pivotal module for the subsequent analyses. The correlation between module membership and gene significance in the turquoise module was determined for elucidating the association between the genes and smoking-related OP. The findings revealed a significant positive correlation between module membership and gene significance (r = 0.97, p < 0.0001) as depicted in Figure 3F. The findings showed a significant correlation between the genes within the turquoise module and smoking-related OP.

3.3 Functional enrichment analysis of DEGs associated with smoking-related OP

To evaluate the GSE13850 dataset's reliability in providing insights into the pathogenesis of smoking-related osteoporosis, functional enrichment analysis was conducted on the DEGs that intersected with the genes found within the turquoise module. By intersecting the 280 DEGs with the 982 genes in the turquoise module, a total of 208 genes were identified as shared genes (Figure 4A). The results of GO enrichment analysis revealed that the 208 intersecting genes were primarily enriched in the nucleoplasm term in the cellular component (CC) category of GO; the transcription factor binding, signaling receptor binding, and beta receptor binding terms in the molecular function (MF) category; and the regulation of cell communication and signaling

term in the biological process (BP) category (Figure 4C–E). The results of KEGG pathway analysis revealed that the 208 intersecting genes were primarily enriched in the MAPK signaling pathway and B cell receptor signaling pathway terms (Figure 4B). Altogether, the results of enrichment analyses revealed that the 208 genes of OP were primarily related to inflammatory responses, which indicated that the dataset selected herein could provide reliable insights and was therefore used for further analyses.

3.4 Enrichment analyses of intersecting DEGs in COPD and OP

In order to demonstrate whether smokers with low BMD harbor genes that may be related to the pathogenesis of smoking-related COPD, a total of 32 genes were identified as the intersection of DEGs between COPD and OP, and visualized using a Venn diagram (Figure 5A). The results of KEGG enrichment analysis (Figure 5B) revealed that the 32 intersecting genes were primarily enriched in the MAPK signaling pathway, B cell receptor signaling pathway, HIF-1 signaling pathway, toxoplasmosis, apelin signaling pathway, oxytocin signaling pathway, and cGMP-PKG signaling pathway. The results of GO enrichment analysis (Figure 5C) revealed that the 32 intersecting genes were enriched in the biosynthetic process and regulation of signaling terms in the BP category, mast cell granule and apical part of cell terms in the CC category, and signaling receptor binding and channel activity terms in the MF category.
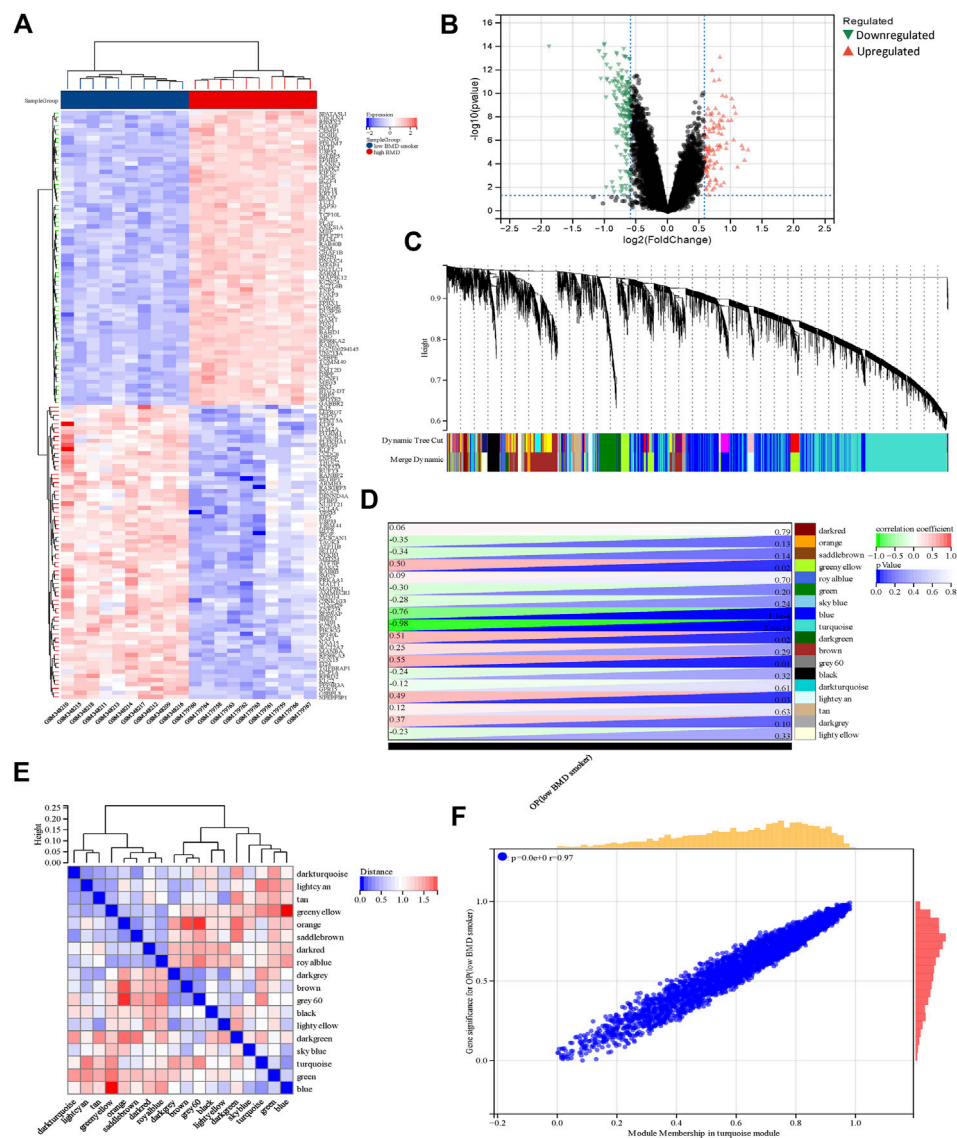
FIGURE 3
Identification of DEGs and genes associated with smoking-related OP by WGCNA. **(A)** The rows depict the DEGs, and the columns represent the respective smoking-related OP cases or control samples. The upregulated and downregulated DEGs are depicted in red and blue, respectively. **(B)** The upregulated and downregulated DEGs are depicted by red and green triangles, respectively. **(C)** The gene co-expression modules are represented in different colors in the gene tree. **(D)** Heatmap depicting the association between gene modules and OP. The turquoise module was significantly correlated with OP. The numbers on the left and right depict the correlation coefficients and *p*-values, respectively. **(E)** Heatmap depicting the adjacency of the eigengenes. **(F)** Plot depicting the correlation between module membership and gene significance of the DEGs in the turquoise module. OP, osteoporosis metabolic syndrome.

3.5 Identification of candidate genes by machine learning algorithms

The identification of candidate genes was performed by conducting logistic regression analysis with the aid of LASSO and RF machine learning algorithms. A total of 13 candidate biomarkers were identified using the LASSO regression algorithm (Figure 6A, B), and the genes were ranked using the RF algorithm by calculating the importance of each gene (Figure 6C,D). The 14 genes predicted using the RF algorithm and the 13 candidate genes identified using the LASSO algorithm were overlapped, and the intersecting genes were visualized using

a Venn diagram (Figure 6E). A total of six intersecting genes, namely, MALT1, PLAT, SCNN1A, SIX3, SPAG9, and VPS35, were identified for final validation.

3.6 Construction of logistic regression model and analysis of ROC curve

A logistic regression model was constructed based on the six candidate genes, namely, MALT1, PLAT, SCNN1A, SIX3, SPAG9, and VPS35, using a logistic regression algorithm. The results demonstrated that the predictive model constructed using these

FIGURE 4
Enrichment analyses of the intersecting genes in smoking-related OP. **(A)** Venn diagram depicting the 208 genes identified from the intersection of 280 DEGs and 982 genes in the turquoise module. **(B)** KEGG pathway analysis of the 208 intersecting genes. The different colors represent the different significantly enriched pathways and related enriched genes. **(C–E)** GO analysis of the intersecting genes that were enriched in different terms in the BP, CC, and MF categories. The x-axis represents the ratio of genes that were enriched in the different GO terms and the y-axis represents the different GO terms. The diameters of the circles correspond to the number of genes, and the colors represent the p-values.

six genes had superior diagnostic performance, with an AUC of 0.99 with the training set (Figure 7A). The results of ROC curve analysis revealed that the model achieved reliable prediction results with the GSE10006 and GSE103174 external datasets (Figure 7B, C), with AUCs of 0.93 and 0.83, respectively, and the findings were consistent with the results obtained using the training set. The findings revealed that the six candidate genes might serve as potential biomarkers of COPD and smoking-related OP; however, further experimental studies are necessary for validating these findings.
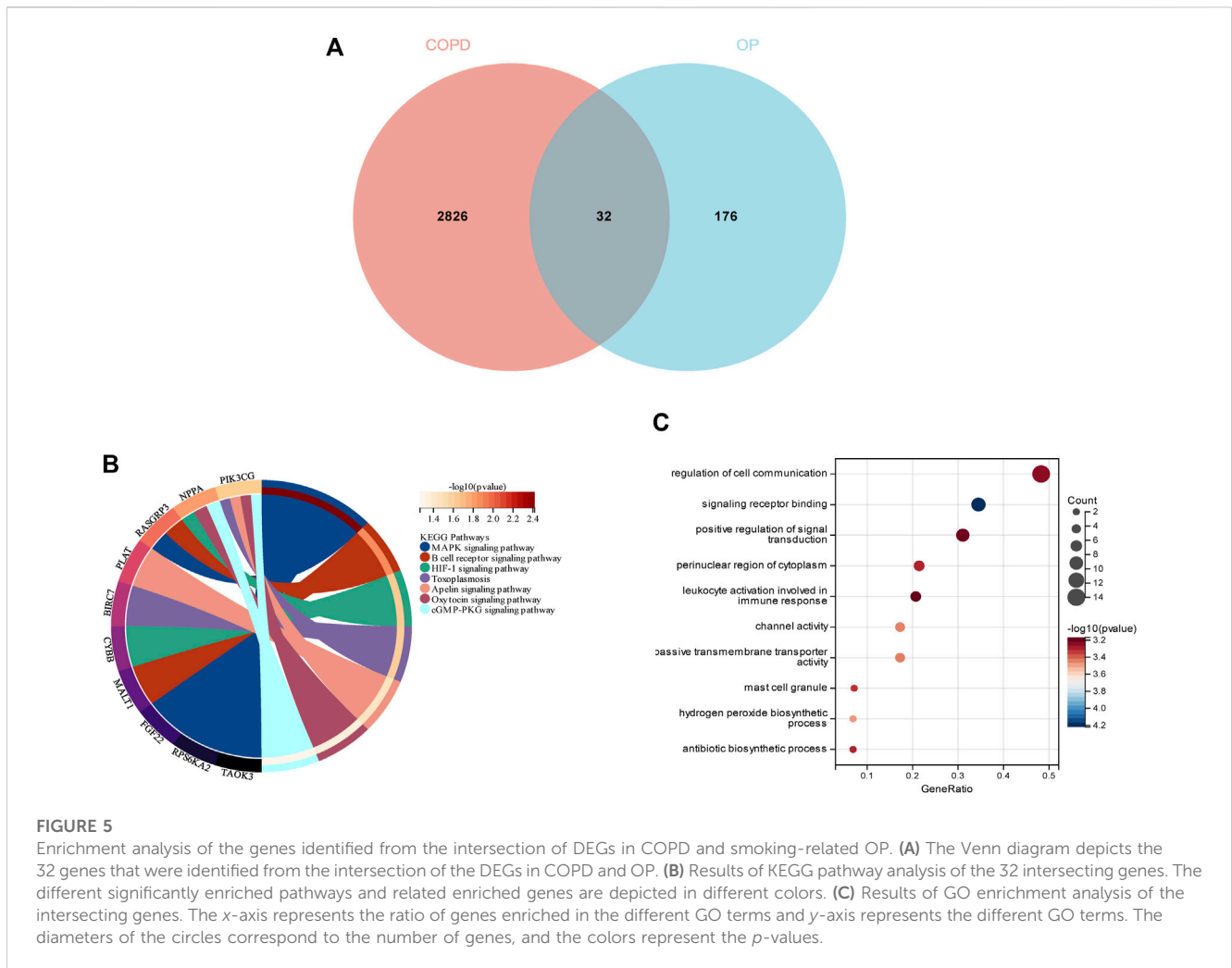
3.7 Analysis of immune cell infiltration

As the DEGs in COPD and OP were enriched in immune regulation terms, the investigation of immune cell infiltration may offer a more comprehensive understanding of the mechanisms involved in regulating immune responses in COPD. The proportions of 22 types of immune cells identified in the GSE11784 dataset were graphically represented by a bar plot (Figure 8A). The findings revealed that patients with smoking-related COPD had a higher population of regulatory T cells (Tregs) and monocytes, but a lower proportion of M2 macrophages (Figure 8B). By analyzing the correlation among the 22 different types of immune cells (Figure 8C), it

was observed that the fraction of plasma cells exhibited a positive correlation with the fraction of resting NK cells (r = 0.76), while the fraction of resting dendritic cells was positively correlated with the neutrophil fraction (r = 0.64). The fraction of M1 macrophages was negatively correlated to the fraction of resting mast cells (r = −0.52) and follicular helper T cells (r = −0.61), but positively correlated to the fraction of memory-activated CD4 T cells (r = 0.64).

4 Discussion

There has been a growing public consciousness regarding the detrimental impact of cigarette smoke exposure. Despite significant advancements in regulating tobacco usage, cigarette smoking remains a formidable global public health concern, with no clear end in sight. (Petzold et al., 2009; Alberg et al., 2014). Smoking-induced genetic alterations influence the secretion of hormones and bone metabolism in the elderly population, which regulate the pathogenesis of OP (Yoon et al., 2012; Marom-Haham and Shulman, 2016). Cigarette smoking is widely acknowledged as a major contributing factor to the development of COPD, with estimates suggesting that approximately 20%–25% of smokers eventually develop this condition (Løkke et al., 2006). Despite numerous studies on

FIGURE 5
Enrichment analysis of the genes identified from the intersection of DEGs in COPD and smoking-related OP. **(A)** The Venn diagram depicts the 32 genes that were identified from the intersection of the DEGs in COPD and OP. **(B)** Results of KEGG pathway analysis of the 32 intersecting genes. The different significantly enriched pathways and related enriched genes are depicted in different colors. **(C)** Results of GO enrichment analysis of the intersecting genes. The x-axis represents the ratio of genes enriched in the different GO terms and y-axis represents the different GO terms. The diameters of the circles correspond to the number of genes, and the colors represent the p-values.
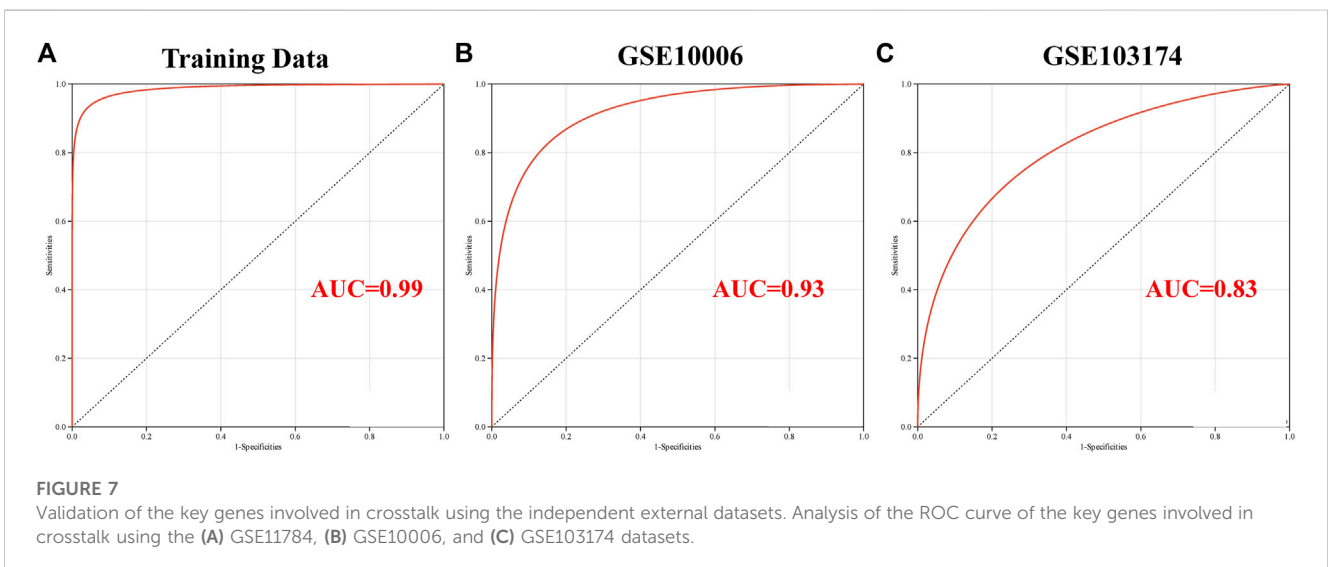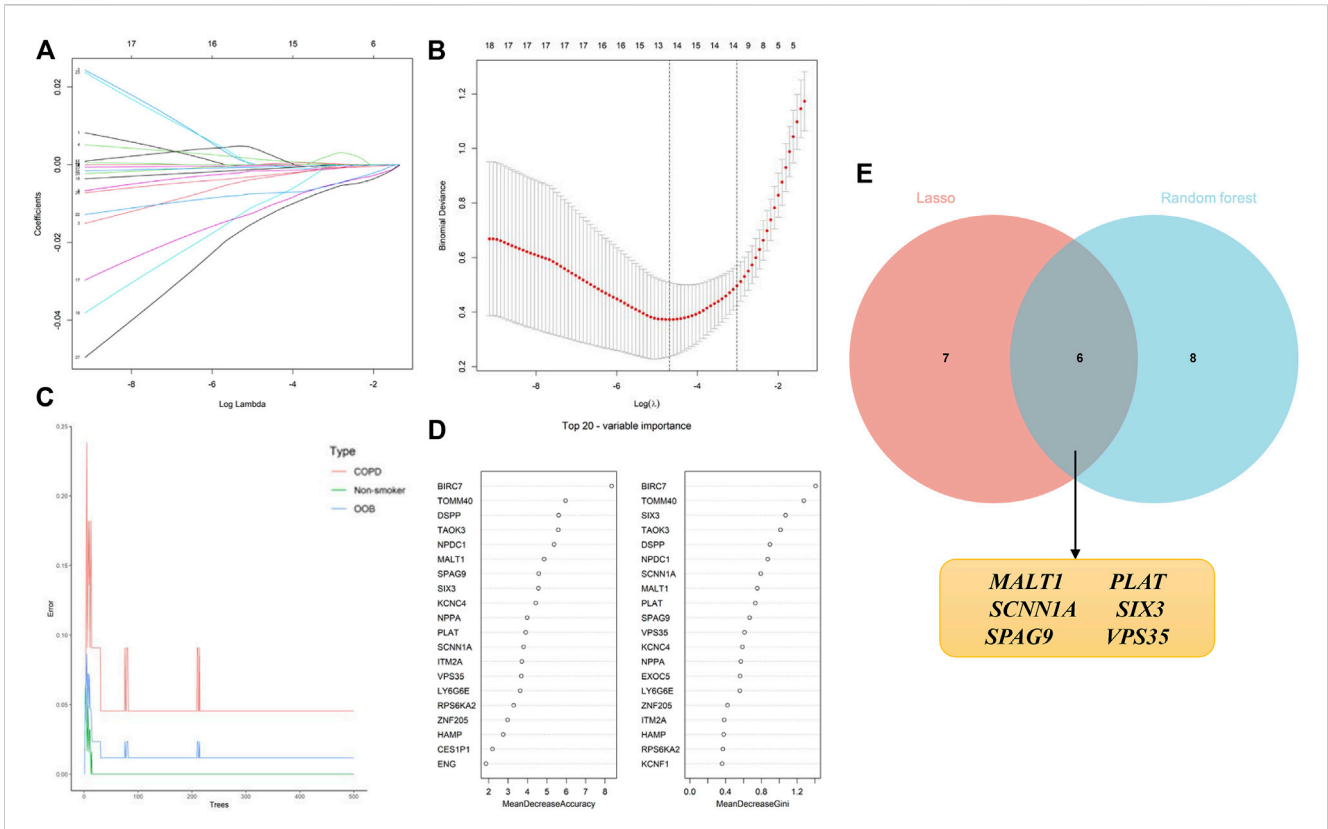
COPD in the past few decades, the mechanisms underlying the pathogenesis of COPD remains poorly understood. The precise effect of cigarette smoking on the pathogenesis of smoking-related OP and COPD, and further investigation is warranted to elucidate the potential correlation between these disorders. A growing body of evidence suggests that cigarette smoking-induced airway inflammation is linked to dysregulated immune cell infiltration in patients diagnosed with COPD (Cruz et al., 2019; Bu et al., 2020). It is worth noting that smoking has the potential to elicit detrimental alterations in the immune system, leading to diseases that stem from aberrant regulation of immune cells (Wang et al., 2021; Luan and Si, 2022). Therefore, the identification of effective and novel diagnostic biomarkers of smoking-related OP and COPD within immune cell components represents a promising domain of investigation, with potential implications for early intervention and improved clinical outcomes.

So far, however, no study has investigated the combined effects of smoking-related OP and COPD. There are no studies on the application of machine learning methods or the construction of logistic regression models for predicting the shared gene signatures of smoking-related OP and COPD. The
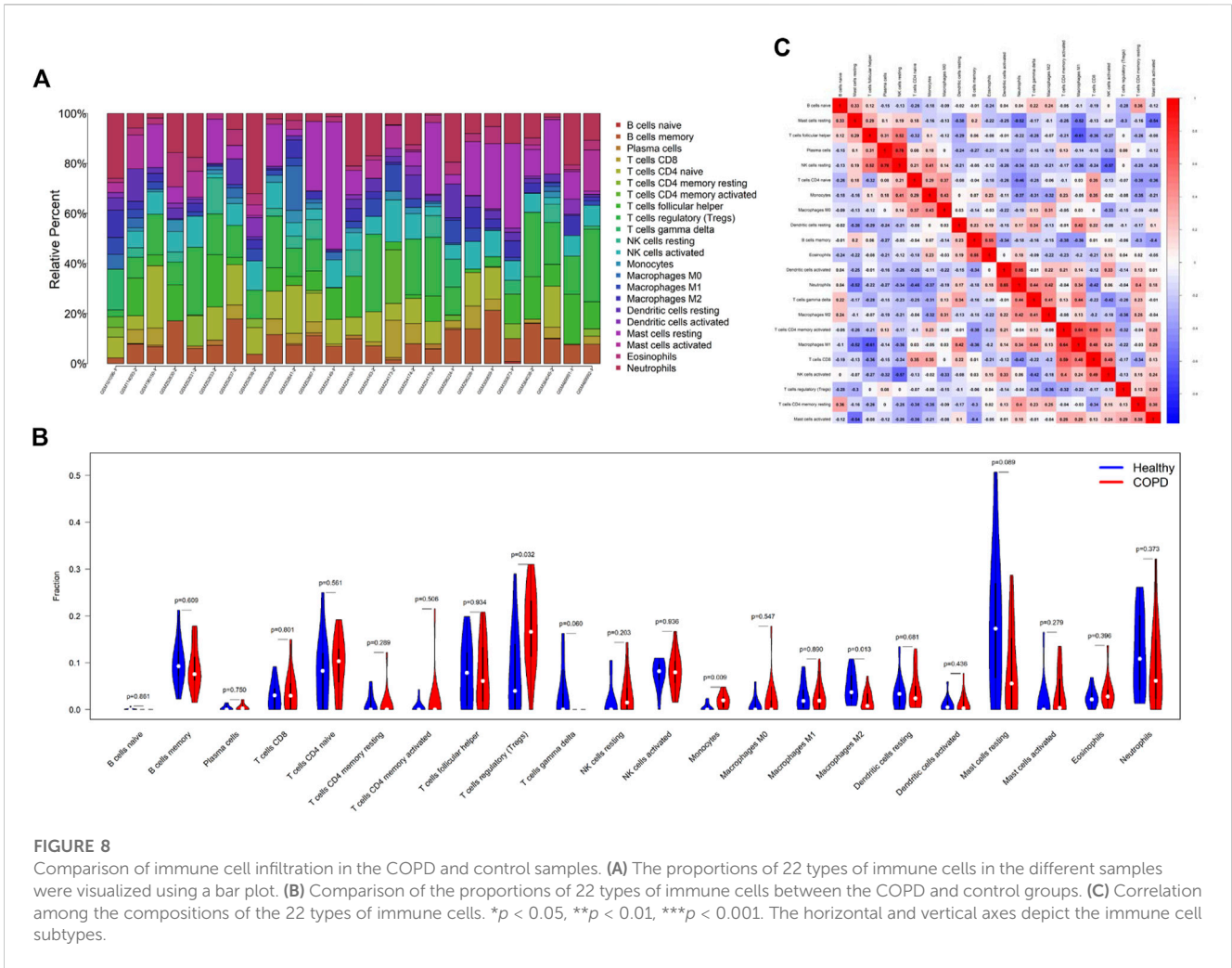
present study performed integrated bioinformatics analyses and applied machine learning algorithms for identifying potential biomarkers and evaluating the diagnostic value of smoking-related OP in patients with COPD. Notably, the study identified six immune-associated candidate diagnostic genes, namely, MALT1, PLAT, SCNN1A, SIX3, SPAG9, and VPS35, that could serve as biomarkers of OP in patients with COPD.

The MALT1 protein exhibits expression and functional activity within osteoclasts, and is stimulated by receptor activator of NF-κB ligand (RANKL) in preosteoclasts. Previous investigations have demonstrated that individuals with combined immunodeficiency attributed to MALT1 deficiency present with diminished bone mineral density, leading to the occurrence of several low-impact fractures. Monajemi et al. demonstrated that Malt1 deficient mice develop an OP phenotype that is characterized by increased osteoclastogenesis *in vivo*, and is primarily caused by the inflammation of osteoclasts (Monajemi et al., 2019). Furthermore, MALT1 is of paramount importance in regulating innate and adaptive immune signaling by adjusting the activation threshold of immune cells. In a murine model of autoimmune arthritis, it was previously reported that the deletion of MALT1 in T cells

FIGURE 6
Application of machine learning algorithms for screening candidate diagnostic biomarkers in smokers with COPD and low BMD. **(A, B)** Screening of biomarkers using the LASSO model. The number of genes (*n* = 13) corresponding to the lowest point of the curve was most suitable. **(C, D)** The error in the COPD group, as identified by the RF algorithm; the control group and genes are ranked based on the importance score. **(E)** Venn diagram depicting the six candidate diagnostic genes identified using the LASSO and RF machine learning algorithms.

FIGURE 7
Validation of the key genes involved in crosstalk using the independent external datasets. Analysis of the ROC curve of the key genes involved in crosstalk using the **(A)** GSE11784, **(B)** GSE10006, and **(C)** GSE103174 datasets.

resulted in the development of spontaneous OP, which was concomitant with changes in the frequency of dysfunctional Treg cells (Gilis et al., 2019). The CARMA1-BCL10-MALT1 (CBM) complex is known to contribute to the pathogenesis and progression of allergic inflammation and diseases, including COPD (DeVore and Khurana Hershey, 2022). PLAT has been identified as an immune checkpoint gene that exerts a critical influence on the prognosis of various cancers, such as breast cancer and hepatocellular carcinoma, by modulating the levels of immune molecules and regulating the infiltration of

FIGURE 8
Comparison of immune cell infiltration in the COPD and control samples. **(A)** The proportions of 22 types of immune cells in the different samples were visualized using a bar plot. **(B)** Comparison of the proportions of 22 types of immune cells between the COPD and control groups. **(C)** Correlation among the compositions of the 22 types of immune cells. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. The horizontal and vertical axes depict the immune cell subtypes.

immune cells in the tumor microenvironment (Wang et al., 2021; Hu et al., 2022). Using gene network analysis, a prior study identified PLAT as a SARS-CoV-2 target and elucidated its potential involvement in COVID-19's inflammatory response, metabolism of reactive oxygen species, and immune response. (Saik and Klimontov, 2022). Wan et al. conducted a large-scale genome-wide analysis of two COPD cohorts and obtained DNA methylation data for 27,578 CpG sites in 14,475 genes. The study revealed that the systemic use of steroids is associated with differential site-specific methylation of several hub genes, including SCNN1A, which partakes in various biological processes, including cellular ion homeostasis and the activation of leukocytes and lymphocytes (Wan et al., 2012). A separate investigation has provided evidence that SCNN1A is associated with unfavorable clinical outcomes in individuals afflicted with ovarian cancer, and exerts an impact on the immune cell infiltration patterns within the tumor microenvironment (Lou et al., 2022). A previous study reported that the SIX3 homeodomain transcription factor was dysregulated in a TNF-α-stimulated inflammatory model of epithelial cells, and SIX3 was found to affect cellular apoptosis and motility under inflammatory conditions (Korthagen et al.,

2015). Additionally, Zheng et al. revealed the critical role of SIX3 loss-of-function in breast cancer progression, tumorigenesis, and metastasis. (Zheng et al., 2018). SPAG9 is a newly identified member of cancer/testis antigens and has been found to induce a specific immune response in several patients with cancer. SPAG9 promotes the growth and metastasis of epithelial cells by regulating the c-Jun N-terminal kinase (JNK) and mitogen-activated protein kinase (MAPK) signaling pathways (Pan et al., 2018). The MAPK signaling pathway is a crucial mediator in the pathogenesis of OP and COPD, and has been reported to regulate the activity of osteoblasts and osteoclasts (Malakoti et al., 2022) and modulate airway inflammation and remodeling (Mei et al., 2022). Xia et al. reported that VPS35 partakes in regulating the trafficking, signaling, and functions of receptor activator of nuclear factor kappaB ligand (RANK) (Xia et al., 2013). The loss function of VPS35 alters the RANKL-induced distribution of RANK, enhances the sensitivity of RANKL, sustains RANKL signaling, and induces the formation of hyper-resorptive osteoclasts. Another study demonstrated that VPS35 can be potentially applied for the diagnosis OP and the VPS35 gene was found to play a potential role in the pathogenesis of OP (Xia et al., 2017).

Immunological and inflammatory modulation has been observed throughout all stages of COPD as reported in previous studies. The present study revealed an elevated frequency of Treg cells and monocytes and a reduced frequency of M2 macrophages in COPD patients, in agreement with previous investigations. Comprehending the underlying mechanisms that underlie inflammatory signaling can furnish valuable insights towards the development of diagnostic methodologies (Chen et al., 2022) and facilitate the identification of innovative therapeutic agents for COPD related to smoking.

The present study has certain limitations with are described hereafter. Firstly, the other related genes could not be identified in this study owing to the limited number of studies on the genetic variations induced by smoking, and the fact that only the GSE13850 dataset of GEO contained data regarding smoking-related OP. Secondly, as only a limited number of clinical samples could be included in our study, the findings obtained herein need to be confirmed using a larger cohort. Thirdly, although the six identified genes were primarily involved in regulating the inflammatory and immune pathways, further in-depth molecular biology experiments and prospective clinical trial cohorts need to be designed for validating the mechanism of action of these candidate genes.

Based on the results of this study, there are several future directions that can be explored. Firstly, further validation of the identified feature genes or exploration of other feature genes can be conducted in larger sample sizes to improve the accuracy and reliability of the results. Additionally, functional analysis can be performed on these feature genes to better understand their biological roles in COPD and smoking-related OP. Furthermore, biological experiments can be used to explore the mechanisms of action of these identified feature genes and further investigate the effectiveness and safety of targeting these genes for the treatment of these diseases. In summary, this study provides valuable insights into potential biomarkers and therapeutic targets for COPD and smoking-related OP. Further research and development in these areas may lead to progress in the diagnosis, treatment, and management of these diseases.

5 Conclusion

The present study systematically identified six candidate diagnostic genes, namely, MALT1, PLAT, SCNN1A, SIX3, SPAG9, and VPS35, which were the shared gene signatures of smoking-related OP and COPD, through a combined approach of integrated bioinformatics analyses and machine learning algorithms. The findings revealed that the association between smoking-related OP and COPD and the influence of cigarette smoking on the pathogenesis of these two diseases could be closely related to the infiltration of immune cells. The aforementioned genes and immune cells hold promise as potential targets for immunotherapeutic intervention in individuals afflicted with smoking-related OP and COPD.

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 authors.

Author contributions

HW, SL, YS, and JW developed the hypothesis and prepared the manuscript. BC and HY designed the study and analyzed the data. MW prepared the images. All the authors read and approved the final manuscript and agree to its submission. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2023.1204031/full#supplementary-material

References

Abu-Jamous, B., and Clust, K. S. (2018). Clust: Automatic extraction of optimal co-expressed gene clusters from gene expression data. *Genome Biol.* 19 (1), 172. doi:10.1186/s13059-018-1536-8

Alberg, A. J., Shopland, D. R., and Cummings, K. M. (2014). The 2014 surgeon general's report: Commemorating the 50th anniversary of the 1964 report of the advisory committee to the US surgeon general and updating the evidence on the health consequences of cigarette smoking. *Am. J. Epidemiol.* 179 (4), 403–412. doi:10.1093/aje/kwt335

Alderden, J., Pepper, G. A., Wilson, A., Whitney, J. D., Richardson, S., Butcher, R., et al. (2018). Predicting pressure injury in critical care patients: A machine-learning model. *Am. J. Crit. care official Publ. Am. Assoc. Critical-Care Nurses* 27 (6), 461–468. doi:10.4037/ajcc2018525

Brightling, C., and Greening, N. (2019). Airway inflammation in COPD: Progress to precision medicine. *Eur. Respir. J.* 54 (2). doi:10.1183/13993003.00651-2019

Brook, J. S., Balka, E. B., and Zhang, C. (2012). The smoking patterns of women in their forties: Their relationship to later osteoporosis. *Psychol. Rep.* 110 (2), 351–362. doi:10.2466/13.18.PR0.110.2.351-362

Bu, T., Wang, L. F., and Yin, Y. Q. (2020). How do innate immune cells contribute to airway remodeling in COPD progression? *Int. J. chronic Obstr. Pulm. Dis.* 15, 107–116. doi:10.2147/COPD.S235054

Carolan, B. J., Harvey, B. G., De, B. P., Vanni, H., and Crystal, R. G., Decreased expression of intelectin 1 in the human airway epithelium of smokers compared to nonsmokers. *J. Immunol. Baltim. Md* : 1950).;181(8):5760–5767. doi:10.4049/jimmunol.181.8.5760

Chen, Y., Bai, P., Liu, L., Han, J., Zeng, H., and Sun, Y. (2016). Increased RANKL expression in peripheral T cells is associated with decreased bone mineral density in patients with COPD. *Int. J. Mol. Med.* 38 (2), 585–593. doi:10.3892/ijmm.2016.2629

Chen, Y., Ma, K., Si, H., Duan, Y., and Zhai, H. (2022). Network pharmacology integrated molecular docking to reveal the autism and mechanism of baohewan heshiwei wen dan tang. *Curr. Pharm. Des.* 28 (39), 3231–3241. doi:10.2174/1381612828666220926095922

Chen, Y., Ma, K., Xu, P., Si, H., Duan, Y., and Zhai, H. (2022). Design and screening of new lead compounds for autism based on QSAR model and molecular docking studies. *Molecules* 27 (21), 7285. doi:10.3390/molecules27217285

Consortium, T. G. O. (2019). The gene Ontology resource: 20 years and still GOing strong. *Nucleic acids Res.* 47 (D1), D330–D338. doi:10.1093/nar/gky1055

Cruz, T., López-Giraldo, A., Noell, G., Casas-Recasens, S., Garcia, T., Molins, L., et al. (2019). Multi-level immune response network in mild-moderate chronic obstructive pulmonary disease (COPD). *Respir. Res.* 20 (1), 152. doi:10.1186/s12931-019-1105-z

DeVore, S. B., and Khurana Hershey, G. K. (2022). The role of the CBM complex in allergic inflammation and disease. *J. allergy Clin. Immunol.* 150 (5), 1011–1030. doi:10.1016/j.jaci.2022.06.023

Di Stefano, A., Caramori, G., Gnemmi, I., Contoli, M., Vicari, C., Capelli, A., et al. (2009). T helper type 17-related cytokine expression is increased in the bronchial mucosa of stable chronic obstructive pulmonary disease patients. *Clin. Exp. Immunol.* 157 (2), 316–324. doi:10.1111/j.1365-2249.2009.03965.x

Edgar, R., Domrachev, M., and Lash, A. E. (2002). Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic acids Res.* 30 (1), 207–210. doi:10.1093/nar/30.1.207

Ellis, K., Kerr, J., Godbole, S., Lanckriet, G., Wing, D., and Marshall, S. (2014). A random forest classifier for the prediction of energy expenditure and type of physical activity from wrist and hip accelerometers. *Physiol. Meas.* 35 (11), 2191–2203. doi:10.1088/0967-3334/35/11/2191

Eustace, A., Smyth, L. J. C., Mitchell, L., Williamson, K., Plumb, J., and Singh, D. (2011). Identification of cells expressing IL-17A and IL-17F in the lungs of patients with COPD. *Chest* 139 (5), 1089–1100. doi:10.1378/chest.10-0779

Frei, A., Muggensturm, P., Putcha, N., Siebeling, L., Zoller, M., Boyd, C. M., et al. (2014). Five comorbidities reflected the health status in patients with chronic obstructive pulmonary disease: The newly developed COMCOLD index. *J. Clin. Epidemiol.* 67 (8), 904–911. doi:10.1016/j.jclinepi.2014.03.005

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33 (1), 1–22. doi:10.18637/jss.v033.i01

Gilis, E., Gaublomme, D., Staal, J., Venken, K., Dhaenens, M., Lambrecht, S., et al. (2019). Deletion of mucosa-associated lymphoid tissue lymphoma translocation protein 1 in mouse T cells protects against development of autoimmune arthritis but leads to spontaneous osteoporosis. *Arthritis and rheumatology (Hoboken, NJ)* 71 (12), 2005–2015. doi:10.1002/art.41029

Graat-Verboom, L., Wouters, E. F., Smeenk, F. W., van den Borne, B. E., Lunde, R., and Spruit, M. A. (2009). Current status of research on osteoporosis in COPD: A systematic review. *Eur. Respir. J.* 34 (1), 209–218. doi:10.1183/09031936.50130408

Hikichi, M., Mizumura, K., Maruoka, S., and Gon, Y. (2019). Pathogenesis of chronic obstructive pulmonary disease (COPD) induced by cigarette smoke. *J. Thorac. Dis.* 11 (17), S2129-S2140–s40. doi:10.21037/jtd.2019.10.43

Hu, B., Shen, X., Qin, W., Zhang, L., Zou, T., Dong, Q., et al. (2022). A prognostic nomogram for hepatocellular carcinoma based on wound healing and immune checkpoint genes. *J. Clin. Transl. hepatology* 10 (5), 891–900. doi:10.14218/JCTH.2021.00296

Kanehisa, M., and Goto, S. (2000). Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic acids Res.* 28 (1), 27–30. doi:10.1093/nar/28.1.27

Kanis, J. A., Johnell, O., Oden, A., Johansson, H., De Laet, C., Eisman, J. A., et al. (2005). Smoking and fracture risk: A meta-analysis. *Osteoporos. Int. a J. established as result Coop. between Eur. Found. Osteoporos. Natl. Osteoporos. Found. U. S. A.* 16 (2), 155–162. doi:10.1007/s00198-004-1640-3

Kanis, J. A., McCloskey, E. V., Johansson, H., Oden, A., Melton, L. J., 3rd, and Khaltaev, N. (2008). A reference standard for the description of osteoporosis. *Bone* 42 (3), 467–475. doi:10.1016/j.bone.2007.11.001

Kong, Y. Y., Feige, U., Sarosi, I., Bolon, B., Tafuri, A., Morony, S., et al. (1999). Activated T cells regulate bone loss and joint destruction in adjuvant arthritis through osteoprotegerin ligand. *Nature* 402 (6759), 304–309. doi:10.1038/46303

Korthagen, N. M., van Bilsen, K., Swagemakers, S. M., van de Peppel, J., Bastiaans, J., van der Spek, P. J., et al. (2015). Retinal pigment epithelial cells display specific transcriptional responses upon TNF-α stimulation. *Br. J. Ophthalmol.* 99 (5), 700–704. doi:10.1136/bjophthalmol-2014-306309

Kumar, N., Narayan Das, N., Gupta, D., Gupta, K., and Bindra, J. (2021). Efficient automated disease diagnosis using machine learning models. *J. Healthc. Eng.* 2021, 9983652. doi:10.1155/2021/9983652

Langfelder, P., and Horvath, S. (2008). Wgcna: an R package for weighted correlation network analysis. *BMC Bioinforma.* 9, 559. doi:10.1186/1471-2105-9-559

Lehouck, A., Boonen, S., Decramer, M., and Janssens, W. (2011). COPD, bone metabolism, and osteoporosis. *Chest* 139 (3), 648–657. doi:10.1378/chest.10-1427

Li, S., Chen, B., Chen, H., Hua, Z., Shao, Y., Yin, H., et al. (2021). Analysis of potential genetic biomarkers and molecular mechanism of smoking-related postmenopausal osteoporosis using weighted gene co-expression network analysis and machine learning. *PloS one* 16 (9), e0257343. doi:10.1371/journal.pone.0257343

Li, Y., Gao, H., Zhao, L., and Wang, J. (2022). Osteoporosis in COPD patients: Risk factors and pulmonary rehabilitation. *Clin. Respir. J.* 16 (7), 487–496. doi:10.1111/crj.13514

Lippi, L., Folli, A., Curci, C., D'Abrosca, F., Moalli, S., Mezian, K., et al. (2022). Osteosarcopenia in patients with chronic obstructive pulmonary diseases: Which pathophysiologic implications for rehabilitation? *Int. J. Environ. Res. public health* 19 (21). doi:10.3390/ijerph192114314

Løkke, A., Lange, P., Scharling, H., Fabricius, P., and Vestbo, J. (2006). Developing COPD: A 25 year follow up study of the general population. *Thorax* 61 (11), 935–939. doi:10.1136/thx.2006.062802

Lou, J., Wei, L., and Wang, H. (2022). SCNN1A overexpression correlates with poor prognosis and immune infiltrates in ovarian cancer. *Int. J. general Med.* 15, 1743–1763. doi:10.2147/IJGM.S351976

Luan, M., and Si, H. (2022). Novel hypoxia features with appealing implications in discriminating the prognosis, immune escape and drug responses of 947 hepatocellular carcinoma patients. *Transl. Cancer Res.* 11 (7), 2097–2121. doi:10.21037/tcr-22-253

Malakoti, F., Zare, F., Zarezadeh, R., Raei Sadigh, A., Sadeghpour, A., Majidinia, M., et al. (2022). The role of melatonin in bone regeneration: A review of involved signaling pathways. *Biochimie* 202, 56–70. doi:10.1016/j.biochi.2022.08.008

Marom-Haham, L., and Shulman, A. (2016). Cigarette smoking and hormones. *Curr. Opin. obstetrics Gynecol.* 28 (4), 230–235. doi:10.1097/GCO.0000000000000283

Mei, X., Lu, R., Cui, L., Tian, Y., Zhao, P., and Li, J. (2022). Poly I:C exacerbates airway inflammation and remodeling in cigarette smoke-exposed mice. *Lung* 200 (6), 677–686. doi:10.1007/s00408-022-00574-7

Monajemi, M., Fisk, S., Pang, Y. C. F., Leung, J., Menzies, S. C., Ben-Othman, R., et al. (2019). Malt1 deficient mice develop osteoporosis independent of osteoclast-intrinsic effects of Malt1 deficiency. *J. Leukoc. Biol.* 106 (4), 863–877. doi:10.1002/JLB.5VMA0219-054R

Pan, J., Yu, H., Guo, Z., Liu, Q., Ding, M., Xu, K., et al. (2018). Emerging role of sperm-associated antigen 9 in tumorigenesis. *Biomed. Pharmacother. = Biomedecine Pharmacother.* 103, 1212–1216. doi:10.1016/j.biopha.2018.04.168

Petzold, A. M., Balciunas, D., Sivasubbu, S., Clark, K. J., Bedell, V. M., Westcot, S. E., et al. (2009). Nicotine response genetics in the zebrafish. *Proc. Natl. Acad. Sci. U. S. A.* 106 (44), 18662–18667. doi:10.1073/pnas.0908247106

Regan, E., and Jaramillo, J. (2012). It's the fracture that matters -bone disease in COPD patients. *Copd* 9 (4), 319–321. doi:10.3109/15412555.2012.708544

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., et al. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinforma.* 12, 77. doi:10.1186/1471-2105-12-77

Romme, E. A., Geusens, P., Lems, W. F., Rutten, E. P., Smeenk, F. W., van den Bergh, J. P., et al. (2015). Fracture prevention in COPD patients; a clinical 5-step approach. *Respir. Res.* 16 (1), 32. doi:10.1186/s12931-015-0192-8

Ryan, C. S., Petkov, V. I., and Adler, R. A. (2011). Osteoporosis in men: The value of laboratory testing. *Osteoporos. Int. a J. established as result Coop. between Eur. Found. Osteoporos. Natl. Osteoporos. Found. U. S. A.* 22 (6), 1845–1853. doi:10.1007/s00198-010-1421-0

Saik, O. V., and Klimontov, V. V. (2022). Gene networks of hyperglycemia, diabetic complications, and human proteins targeted by SARS-CoV-2: What is the molecular basis for comorbidity? *Int. J. Mol. Sci.* 23 (13), 7247. doi:10.3390/ijms23137247

Sarkar, M., Bhardwaj, R., Madabhavi, I., and Khatana, J. (2015). Osteoporosis in chronic obstructive pulmonary disease. *Clin. Med. insights Circulatory, Respir. Pulm. Med.* 9, 5–21. doi:10.4137/CCRPM.S22803

Sato, K., Suematsu, A., Okamoto, K., Yamaguchi, A., Morishita, Y., Kadono, Y., et al. (2006). Th17 functions as an osteoclastogenic helper T cell subset that links T cell activation and bone destruction. *J. Exp. Med.* 203 (12), 2673–2682. doi:10.1084/jem.20061775

Shen, W., Song, Z., Zhong, X., Huang, M., Shen, D., Gao, P., et al. (2022). Sangerbox: A comprehensive, interaction-friendly clinical bioinformatics analysis platform. *iMeta* 1 (3), e36. doi:10.1002/imt2.36

Tilley, A. E., O'Connor, T. P., Hackett, N. R., Strulovici-Barel, Y., Salit, J., Amoroso, N., et al. (2011). Biologic phenotyping of the human small airway epithelial response to cigarette smoking. *PloS one* 6 (7), e22798. doi:10.1371/journal.pone.0022798

Wan, E. S., Qiu, W., Baccarelli, A., Carey, V. J., Bacherman, H., Rennard, S. I., et al. (2012). Systemic steroid exposure is associated with differential methylation in chronic obstructive pulmonary disease. *Am. J. Respir. Crit. care Med.* 186 (12), 1248–1255. doi:10.1164/rccm.201207-1280OC

Wang, X., Xue, D., Zhu, X., Geng, R., Bao, X., Chen, X., et al. (2021). Low expression of PLAT in breast cancer infers poor prognosis and high immune infiltrating level. *Int. J. general Med.* 14, 10213–10224. doi:10.2147/IJGM.S341959

Ward, K. D., and Klesges, R. C. (2001). A meta-analysis of the effects of cigarette smoking on bone mineral density. *Calcif. tissue Int.* 68 (5), 259–270. doi:10.1007/BF02390832

Watanabe, R., Tanaka, T., Aita, K., Hagiya, M., Homma, T., Yokosuka, K., et al. (2015). Osteoporosis is highly prevalent in Japanese males with chronic obstructive pulmonary disease and is associated with deteriorated pulmonary function. *J. bone mineral metabolism* 33 (4), 392–400. doi:10.1007/s00774-014-0605-7

Xia, B., Li, Y., Zhou, J., Tian, B., and Feng, L. (2017). Identification of potential pathogenic genes associated with osteoporosis. *Bone and Jt. Res.* 6 (12), 640–648. doi:10.1302/2046-3758.612.BJR-2017-0102.R1

Xia, W. F., Tang, F. L., Xiong, L., Xiong, S., Jung, J. U., Lee, D. H., et al. (2013). Vps35 loss promotes hyperresorptive osteoclastogenesis and osteoporosis via sustained RANKL signaling. *J. Cell Biol.* 200 (6), 821–837. doi:10.1083/jcb.201207154

Xiong, J., Tian, J., Zhou, L., Le, Y., and Sun, Y. (2020). Interleukin-17A deficiency attenuated emphysema and bone loss in mice exposed to cigarette smoke. *Int. J. chronic Obstr. Pulm. Dis.* 15, 301–310. doi:10.2147/COPD.S235384

Yang, C., Delcher, C., Shenkman, E., and Ranka, S. (2018). Machine learning approaches for predicting high cost high need patient expenditures in health care. *Biomed. Eng. online* 17 (1), 131. doi:10.1186/s12938-018-0568-3

Yoon, V., Maalouf, N. M., and Sakhaee, K. (2012). The effects of smoking on bone metabolism. *Osteoporos. Int. a J. established as result Coop. between Eur. Found. Osteoporos. Natl. Osteoporos. Found. U. S. A.* 23 (8), 2081–2092. doi:10.1007/s00198-012-1940-y

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics a J. Integr. Biol.* 16 (5), 284–287. doi:10.1089/omi.2011.0118

Zhang, M., Zhu, K., Pu, H., Wang, Z., Zhao, H., Zhang, J., et al. (2019). An immune-related signature predicts survival in patients with lung adenocarcinoma. *Front. Oncol.* 9, 1314. doi:10.3389/fonc.2019.01314

Zheng, Y., Zeng, Y., Qiu, R., Liu, R., Huang, W., Hou, Y., et al. (2018). The homeotic protein SIX3 suppresses carcinogenesis and metastasis through recruiting the LSD1/NuRD(MTA3) complex. *Theranostics* 8 (4), 972–989. doi:10.7150/thno.22328

Zhou Y, Y., Shi, W., Zhao, D., Xiao, S., Wang, K., and Wang, J. (2022). Identification of immune-associated genes in diagnosing aortic valve calcification with metabolic syndrome by integrated bioinformatics analysis and machine learning. *Front. Immunol.* 13, 937886. doi:10.3389/fimmu.2022.937886

Zhou Z, Z., Wei, J., Kuang, L., Zhang, K., Liu, Y., He, Z., et al. (2022). Characterization of aging cancer-associated fibroblasts draws implications in prognosis and immunotherapy response in low-grade gliomas. *Front. Genet.* 13, 897083. doi:10.3389/fgene.2022.897083