BMC Public Health

RESEARCH

Open Access

An analysis of the potential association between obstructive sleep apnea and osteoporosis from the perspective of transcriptomics and NHANES

Shuzhen Li[1†], Yuxin Zan[1†], Fangzhou Li[1], Wenjing Dai[1], Liting Yang[1], Ruiping Yang[1], Xuejun He[3*] and Bei Li[1,2*]

Abstract

Background Obstructive sleep apnea (OSA) and osteoporosis (OP) are prevalent diseases in the elderly. This study aims to reveal the clinical association between OSA and OP and explore potential crosstalk gene targets.

Methods Participants diagnosed with OSA in the National Health and Nutrition Examination Survey (NHANES) database (2015–2020) were included, and OP was diagnosed based on bone mineral density (BMD). We explored the association between OSA and OP, and utilized multivariate logistic regression analysis and machine learning algorithms to explore the risk factors for OP in OSA patients. Overlapping genes of comorbidity were explored using differential expression analysis, Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis, Least Absolute Shrinkage and Selection Operator (LASSO) regression, and Random Forest (RF) methods.

Results In the OSA population, the weighted prevalence of OP was 7.0%. The OP group had more females, lower body mass index (BMI), and more low/middle-income individuals compared to the non-OP group. Female gender and lower BMI were identified as independent risk factors for OP in OSA patients. Gene expression profiling revealed 8 overlapping differentially expressed genes in OP and OSA patients. *KCNJ1*, *NPR3* and *WT1-AS* were identified as shared diagnostic biomarkers or OSA and OP, all of which are associated with immune cell infiltration.

Conclusion This study pinpointed female gender and lower BMI as OP risk factors in OSA patients, and uncovered three pivotal genes linked to OSA and OP comorbidity, offering fresh perspectives and research targets.

Keywords Obstructive sleep apnea, Osteoporosis, Body mass index, Machine learning, NHANES

[†]Shuzhen Li and Yuxin Zan contributed equally to this work.

*Correspondence:
Xuejun He
739282879@qq.com
Bei Li
libei2381@sina.com
[1]Biomedical research institute of Hubei University of Medicine, Hubei Shiyan 442000, China
[2]School of Basic Medicine of Hubei University of Medicine, Hubei Shiyan 442000, China
[3]School of Humanities and Social Science of Hubei University of Medicine, Hubei Shiyan 442000, China

Introduction

Obstructive Sleep Apnea (OSA) is a common sleep-related breathing disorder, characterized by repeated collapse of the upper airway during sleep, leading to chronic intermittent hypoxemia and sleep fragmentation [1]. It is a highly prevalent disease and has rapidly evolved into a major global public health burden [2]. The greatest number of individuals affected by this condition is found in China, followed by the United States, Brazil, and India [3]. Furthermore, an increasing body of evidence from experimental, translational, and clinical research suggests that OSA is frequently associated with the occurrence and development of various systemic diseases, such as cardiovascular and metabolic diseases [4]. However, the pathogenic mechanisms of OSA across different organs are intricate and intertwined, and their full understanding remains elusive.

Osteoporosis (OP) is an age-related systemic skeletal disease. In a normal physiological environment, a balance exists between bone formation by osteoblasts and bone resorption by osteoclasts, supporting bone mass and bone mineral density (BMD) within the normal range. However, in patients with OP, this balance is disrupted, characterized by the inhibition of osteoblasts or the overactivation of osteoclasts, leading to bone remodeling [5]. This results in decreased bone density, deterioration of bone microstructure, increased bone fragility, and risk of fractures, especially at the proximal femur and distal radius, with vertebral compression most prevalent [6]. Furthermore, 70% of vertebral fractures do not attract medical attention, mainly due to the prevalence of symptoms such as back pain and height reduction, which are often considered inevitable parts of aging [7]. The crosstalk between the immune microenvironment, inflammatory factors, and the skeletal system forms a complex interdisciplinary field of pathophysiology mechanisms.

Increasing evidence has suggested an interactive relationship between OSA and OP. It is widely accepted that OSA may be a risk factor for decreased BMD, and is associated with a higher incidence of OP [8, 9]. Apnea episodes can result in bodily hypoxia, which in turn leads to metabolic changes harmful to the skeleton, promoting the activation of osteoclasts [10]. Alternatively, sleep deprivation, heightened sympathetic nerve tension, alterations in melatonin, or comorbidities of OSA affecting the pathophysiology of skeletal biology, thereby negatively affecting the skeleton system [11]. Conversely, OP and vertebral fractures are correlated with an increased incidence of OSA, which is a secondary outcome [12, 13]. The incidence of both diseases increases with age, and they influence each other. The medical symptoms of either one can mask the non-medical symptoms of the other, which further increases the difficulty of diagnosing and treating comorbidities. However, the potential bidirectional causality between OP and OSA still needs further investigation to more accurately elucidate the precise mechanism underlying the coexistence of these two diseases.

In this study, we ingeniously combined multiple databases to analyze the relationship between OSA and OP from the perspectives of clinicians and genomics. We have established predictive models for clinical indicators and overlapping feature genes, respectively. To our knowledge, this may be the first study to explore the shared features between OSA and OP, and it is expected to help us better implement the public advocacy of early detection, early diagnosis, and early treatment.

Method

Study population

This study performed a retrospective analysis using the National Health and Nutrition Examination Survey (NHANSE) datasets from 2015–2016 and 2017–2020, of which 2020 data is available as of March of that year due to the impact of COVID-19. All data can be directly obtained from NHANES's official website (https://www.cdc.gov/nchs/nhanes/index.htm), it is important to note that the database is a study of American outpatient probability of large-scale civilian family on behalf of the survey, conducted once a year, every two years for a cycle. In this cohort study, 25,531 participants from 2015 to 2020 were recruited, and ultimately 1730 participants were included in the study analysis based on strict inclusion and exclusion criteria (Fig. 1).

Diagnosis of OSA and OP

OSA is defined as when a person answers "yes" to at least one of the following three NHANES questions: (1) Snoring is greater than or equal to three nights per week; (2) wheezing, sneezing, or stopping breathing three or more nights per week; (3) Sleep about 7 h or more per night on weekdays or work nights, but feel excessively sleepy 16–30 times per month during the day [14].

The diagnosis of OP is based on the WHO consensus that a T score ≤ -2.5 is defined as OP [15]. BMD (g/cm2) obtained by dual-energy X-ray Absorptiometry (DXA) for all participants during 2015–2020 was collected and converted into T scores using the following formula: T score = (BMD respondents - mean BMD (reference value)/ Standard deviation (reference value)).

Covariates

The following covariates were used based on previous studies and clinical practice: (1) demographic characteristics: Age, sex, race, body mass index (BMI), education level (Lower than High school, High school, College or more), poverty to income ratio (PIR), smoking status, alcohol consumption status; (2) Medical conditions:

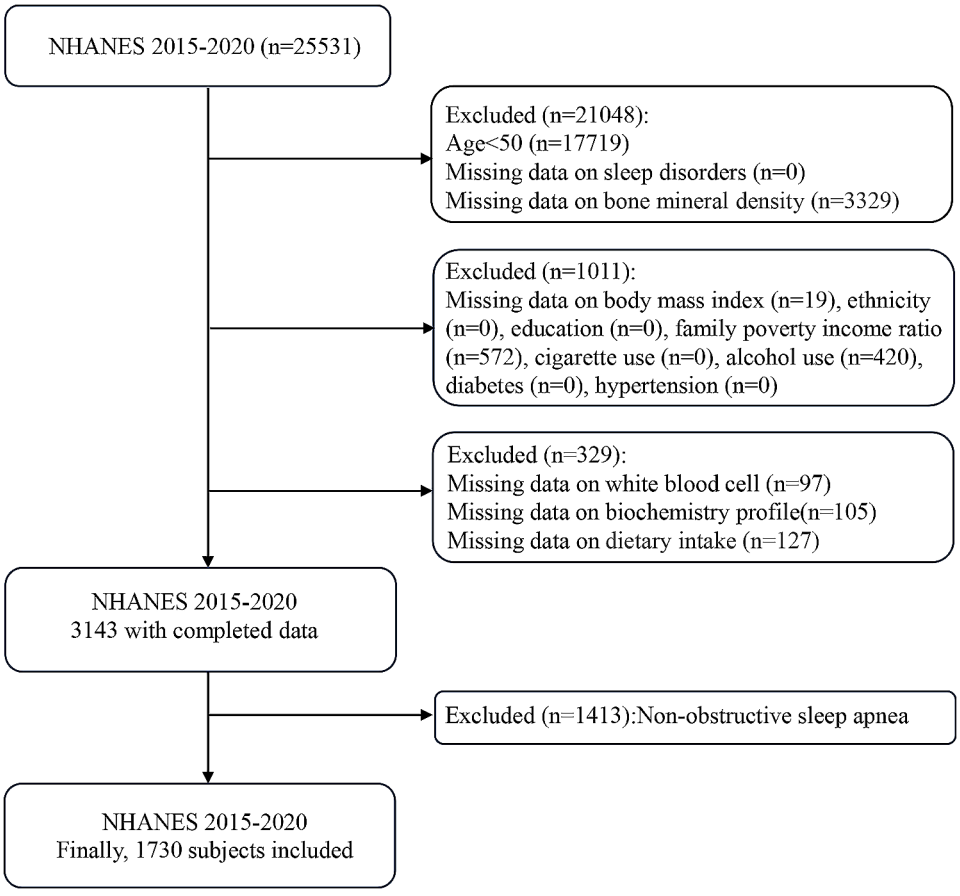

Fig. 1 Flowchart of the cross-sectional study. We selected 1730 subjects diagnosed with Obstructive Sleep Apnea (OSA) from two consecutive cycles (2015–2020) in the National Health and Nutrition Examination Survey (NHANES) database. These subjects had complete information and met the diagnostic criteria for OSA

diabetes mellitus, hypertension; (3) Laboratory Data: WBC count, ALP, AST, TBIL, GGT, Phosphorus, Total Calcium; (4) Dietary Data: Vitamin D Intake, Calcium intake. Serum total bilirubin and gamm-glutamyl transferase (GGT) were used to reflect the state of oxidative stress [16]. As an indicator of socioeconomic status, participants were divided into low income (PIR < 1.3), middle income (PIR = 1.30–3.50), and high income (PIR ≥ 3.50) groups.

Data acquisition

Transcriptome data and platform files related to OSA and OP were downloaded from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). The training dataset is represented by GSE135917, which includes 8 control samples and 34 OSA samples, and GSE7158, which comprises 14 control samples and 12 OP samples. The validation dataset is represented by GSE38792, including 8 control samples and 10 OSA samples, and GSE100609, which contains 4 control samples and 4 OP samples. It is noteworthy that the OSA dataset originates from the GPL6244 platform, while GSE7158 and GSE100609 are derived from the GPL570 and GPL16699 platforms, respectively.

Differentially expressed genes (DEGs) analysis

The original microarray data underwent normalization and analysis using R statistical software. The limma package in R facilitated the differential expression analysis of OSA and OP in comparison with normal control samples [17]. The screening criteria were set at $P < 0.05$ and $|\log2FC| > 0.5$. Heat maps and Volcano maps of DEGs in OSA and OP cohorts were generated using the "pheatmap" and "ggplot2" R packages, respectively. Venn diagrams were employed to cross-screen OSA-DEGs and OP-DEGs, yielding final DEGs (Com-DEGs) that bear a close relation to the pathophysiology of both diseases.

Functional enrichment analysis

Enrichment analysis, a widely employed bioinformatics method, is utilized to shed light on the potential functions of identified targets. In this study, the "clusterProfiler" R package facilitated the Gene Ontology (GO) analysis, encompassing Molecular Function (MF),

Biological Pathway (BP), and Cell Component (CC). Furthermore, the Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was conducted with the assistance of the online analysis tool, Sangerbox 3.0, to investigate significant signaling pathways associated with Com-DEGs.

Machine learning algorithm

The machine learning algorithm Least Absolute Shrinkage and Selection Operator (LASSO) regression and the Random Forest (RF) method, both of which are supervised learning methods, were used to screen disease characteristic genes. These were determined by the "glmnet" and "randomForest" R packages, respectively.

LASSO regression, by controlling regularization parameters, balances model complexity and performance. In this process, it employs a technique of 10-fold cross-validation, along with an optimal lambda, to determine which genes play a key role in the model. These key genes, whose coefficients in the model are non-zero, are referred to as 'marker genes'. Simply put, these 'marker genes' are genes that have a significant impact on the disease under study. The inherent sparsity of LASSO regression gives it an advantage in feature selection, dimensionality reduction, and model interpretability. It can effectively handle multicollinearity and improve generalization performance, making it a valuable choice for various data analysis applications.

For the RF model, the number of decision trees was set to 500. The point with the minimum error was identified, and a model was built to determine the importance of genes. Genes with importance scores of 1.5 or higher were selected for subsequent analysis. The random forest method is a suitable ensemble learning algorithm and machine learning method. It is not subject to variable conditions, is resistant to overfitting, and has higher accuracy, sensitivity, and specificity than decision trees. Therefore, it is a powerful and versatile choice for various predictive modeling tasks.

Finally, considering that various machine learning methods each have their advantages, and the two mentioned above are widely used in the field and perform excellently, to avoid subjective selection bias and the limitations of a single machine learning algorithm, and considering that too many methods may make the interpretation of results complex and not necessarily bring better performance, we cross-analyze the results obtained from LASSO regression and the RF method. This is to ensure the accuracy and reliability of our research results to the greatest possible extent, and to determine the final set of central genes for subsequent analysis.

Verification of characteristic genes

To quantify diagnostic performance, the receiver operating characteristic (ROC) curve was plotted and the area under the curve (AUC), along with the corresponding 95% confidence interval (CI), was calculated using the "pROC" R package. This package employs the DeLong method for AUC calculation, a method widely recognized for ROC analysis in biomedical research [18]. The predictive accuracy of the feature genes was validated in both the training (GSE135917 and GSE7158) and validation cohorts (GSE38792 and GSE100609).

Correlation of immune-infiltrating cells with signature genes

Initially, the correlation coefficient was calculated to measure the relationship between the expression of characteristic genes and immune infiltrating cells. Subsequently, Spearman rank correlation analysis was employed to further investigate this relationship. To visually demonstrate this correlation, a lollipop plot was generated using the "ggplot" R package.

Statistical analysis

This study adopted a weighted approach to account for the complex sampling techniques and research design of NHANES. Dietary weights were utilized as survey weights in accordance with NHANES Analysis Guidelines. Continuous variables, expressed as median (IQR), were analyzed using the Mann–Whitney U test. Categorical variables, expressed as absolute values (n) or percentages (%), were analyzed using the chi-square test. A multivariate logistic regression analysis was conducted to identify risk factors for OP in OSA patients. The "glmnet, corrplot, caret" packages in R software were used to generate LASSO regression results, and the nomogram was constructed using the "rms" package. All statistical analyses were performed using R software version 4.2.1, with a double-sided P-value of less than 0.05 indicating statistical significance.

Result

Baseline characteristics of participants

Among the 1730 OSA participants enrolled in the NHANES study from 2015 to 2020, Table 1 outlines the baseline characteristics of the two groups, categorized by the presence or absence of OP. The weighted prevalence of OP was 7.0%, with 124 classifieds as OP and 1606 as non-OP. When compared to the non-OP group, the OP group demonstrated a higher proportion of women, a lower BMI, and a greater number of individuals with low/middle income. There were no statistically significant differences between the two groups in terms of age, race, education level, smoking status, drinking status, medical

Table 1 Characteristics of osteoporosis and non-osteoporosis cases among OSA participants

Characteristic	Non-OP $N=1606$ (92.8%)[1]	OP $N=124$ (7.2%)[1]	*p*-value[2]
Age (years)	59.0 (55.0, 67.0)	59.0 (54.0, 69.5)	0.771
Gender			**<0.001**
Male	941(57.2%)	17 (13.3%)	
Female	665(42.8%)	107(86.7%)	
BMI (kg/m^2)	29.8 (26.5, 34.3)	25.9 (23.6, 29.6)	**<0.001**
Ethnicity, n (%)			0.300
Mexican American	205.0 (5.5%)	14.0 (3.8%)	
Other Hispanic	222.0 (7.1%)	15.0 (6.5%)	
Non-Hispanic White	622.0 (72.0%)	64.0 (70.8%)	
Non-Hispanic Black	365.0 (8.3%)	15.0 (5.3%)	
Non-Hispanic Asian	132.0 (3.4%)	10.0 (4.6%)	
Other multiracial	60.0 (3.6%)	6.0 (9.0%)	
Education level, n (%)			0.713
Lower than high school	305.0 (9.9%)	21.0 (13.5%)	
High school	396.0 (31.4%)	42.0 (34.7%)	
College or more	905.0 (58.7%)	61.0 (51.8%)	
PIR, n (%)			**0.016**
Low-income	386.0 (16.3%)	36.0 (18.1%)	
Middle-income	617.0 (30.2%)	52.0 (48.5%)	
High-income	603.0 (53.5%)	36.0 (33.4%)	
Smoking status, n (%)			0.387
Smokers	839.0 (52.3%)	64.0 (58.7%)	
Nonsmokers	767.0 (47.7%)	60.0 (41.3%)	
Drinking status, n (%)			0.536
Drinkers	861.0 (60.2%)	63.0 (65.2%)	
Nondrinkers	745.0 (39.8%)	61.0 (34.8%)	
Diabetes mellitus, n (%)	403.0 (21.5%)	21.0 (11.0%)	0.066
Hypertension, n (%)	902.0 (53.5%)	65.0 (48.3%)	0.601
WBC count (1000 cells/μL)	7.10 (5.70, 8.60)	7.10 (6.00, 8.60)	0.679
ALP (IU/L)	74 (62, 89)	79 (66, 95)	0.205
TBIL (μmol/L)	8.6 (5.1, 12.0)	6.8 (5.1, 10.3)	0.429
GGT (IU/L)	23 (16, 37)	19 (16, 31)	0.229
AST(U/L)	21 (17, 26)	21 (19, 25)	0.978
Phosphorus (mmol/L)	1.16 (1.06, 1.26)	1.23 (1.10, 1.36)	0.027
Total Calcium (mmol/L)	2.33 (2.28, 2.40)	2.35 (2.29, 2.38)	0.151
Vitamin D Intake (mg)	3.9 (2.2, 6.0)	3.3 (1.9, 6.0)	0.412
Calcium intake (mg)	889 (618, 1,205)	808 (635, 1,061)	0.286

[1]Median (IQR); n (unweighted) (%)

[2]Wilcoxon rank-sum test for complex survey samples; chi-squared test with Rao & Scott's second-order correction

Abbreviations: OSA, obstructive sleep apnea; OP, Osteoporosis; BMI, Body Mass Index; PIR, family poverty income ratio; ALP, Alkaline Phosphatase; TBIL, Total Bilirubin; GGT, Gamma Glutamyl Transferase; AST, Aspartate Aminotransferase

condition, blood test indicators, and Vitamin D or Calcium intake in the diet ($p>0.05$).

Risk factors for OP in OSA and nomogram construction

Upon employing a multivariate logistic regression model and adjusting for all covariates, it was discerned that being female, possessing a lower BMI, and having a lower PIR emerged as independent risk factors for the development of OP in OSA patients (Table 2). Specifically, within the OSA patients, the risk of females developing OP was 12.3 times higher than that of males (95% CI: 4.83–31.1, $p<0.001$). Furthermore, for every unit increase in BMI, the risk of developing OP decreased by 15%. LASSO regression was utilized to select the most predictive variables, setting a 10-fold cross-validation. Upon utilizing the optimal lambda value, two primary variables were identified: gender and BMI (Fig. 2A, B). These two independent factors were then used to construct a nomogram to predict the risk of OP in OSA patients. As depicted in

Table 2 Multivariate logistic regression analysis assessing the risk of osteoporosis in participants with OSA.

Characteristic	OR	95% CI	*p*-value
Age	1.00	0.95, 1.06	0.920
Gender			**<0.001**
Male	Reference	Reference	
Female	12.3	4.83, 31.1	
BMI	0.85	0.79, 0.91	**<0.001**
Ethnicity			0.385
Non-Hispanic White	Reference	Reference	
Non-Hispanic Black	0.65	0.26, 1.62	
Non-Hispanic Asian	1.00	0.29, 3.50	
Other Hispanic	0.75	0.27, 2.08	
Mexican American	0.61	0.12, 2.95	
Other multiracial	2.53	0.90, 7.12	
Education level			0.367
lower than high school	Reference	Reference	
High school graduate	0.59	0.09, 3.85	0.477
higher than High school	0.41	0.06, 2.89	0.306
PIR			**0.044**
low-income	Reference	Reference	
middle-income	1.77	0.84, 3.71	
high-income	0.75	0.23, 2.41	
Drinking status			0.056
Nondrinkers	Reference	Reference	
Drinkers	1.81	0.95, 3.46	0.127
Smoking status			0.680
Nonsmokers	Reference	Reference	
Smokers	1.15	0.57, 2.34	
Diabetes mellitus			0.399
No	Reference	Reference	
Yes	0.71	0.31, 1.65	
Hypertension			0.715
No	Reference	Reference	
Yes	0.88	0.44, 1.79	
Calcium intake	1.00	1.00, 1.00	0.690
Vitamin D Intake	0.95	0.84, 1.07	0.366
Phosphorus	2.69	0.44, 16.6	0.256
Total Calcium	1.28	0.05, 32.0	0.871

The multivariate logistic regression was adjusted for all covariates

Abbreviations: OSA, obstructive sleep apnea; OR, odds ratio; CI, confidence interval; BMI, body mass index; PIR, family poverty income ratio;

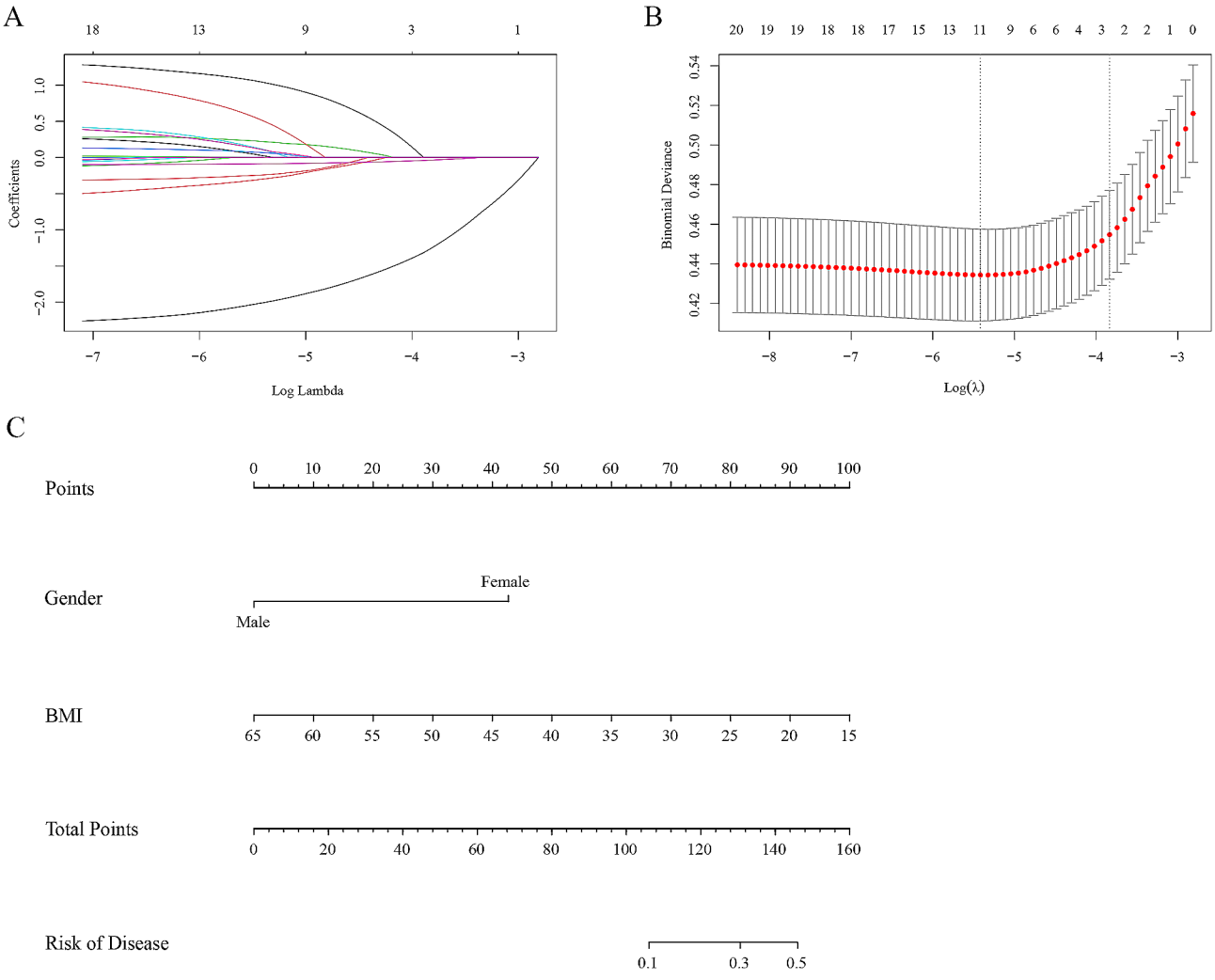

Fig. 2 Selection of optimal variables and construction of a nomogram. (**A**) According to the logarithmic (lambda) sequence, a coefficient profile was generated, and non-zero coefficients were produced by the optimal lambda, key genes corresponding to non-zero coefficients are characteristic genes that have a significant impact on the disease studied; (**B**) The optimal parameter (lambda) in the LASSO model was selected via 10-fold cross-validation using minimum criterion plus one standard error (right vertical line); (**C**)The nomogram estimates the risk of OP occurrence in patients with OSA. The nomogram includes two predictive variables: gender and BMI. Each variable is assigned a specific point value based on its position on their respective scales. These points are then summed up in the 'Total Points' scale, which correlates directly with the 'Risk of Disease' scale at the bottom of the figure

Fig. 2C, the total score of the two variables, gender and BMI, was converted into individual disease risk. The higher the total score, the higher the disease risk, indicating that females with lower BMI are often more susceptible to the disease.

Identification and functional enrichment analysis of common DEGs

In the OSA dataset GSE135917, a total of 668 DEGs were identified, including 431 upregulated genes and 237 downregulated genes. Similarly, in the OP dataset GSE7158, a total of 479 DEGs were identified, comprising 232 upregulated genes and 247 downregulated genes. The expression patterns of these DEGs in both diseases are visually represented in the volcano plots (Fig. 3A, C). The heatmaps (Fig. 3B, D) further illustrate the top 30 DEGs in both diseases. Further analysis revealed that there are 8 overlapping DEGs (*WT1-AS*, *KCNJ1*, *ZNF542P*, *CDKN1A*, *ACOT4*, *OSR2*, *NPR3* and *IL1RN*) between OSA and OP (Fig. 3E).

GO and KEGG enrichment analyses were performed on the aforementioned eight genes to explore common regulatory pathways. The BP subset revealed that these genes primarily regulate the proliferation regulation of muscle cells and osteocytes, as well as cytokine-mediated signaling pathways. The CC subset showed that they were primarily located on the cell membrane, particularly related to potassium channels and protein kinase complexes. The MF subset indicated that at the molecular level, they mainly participated in regulating the activity

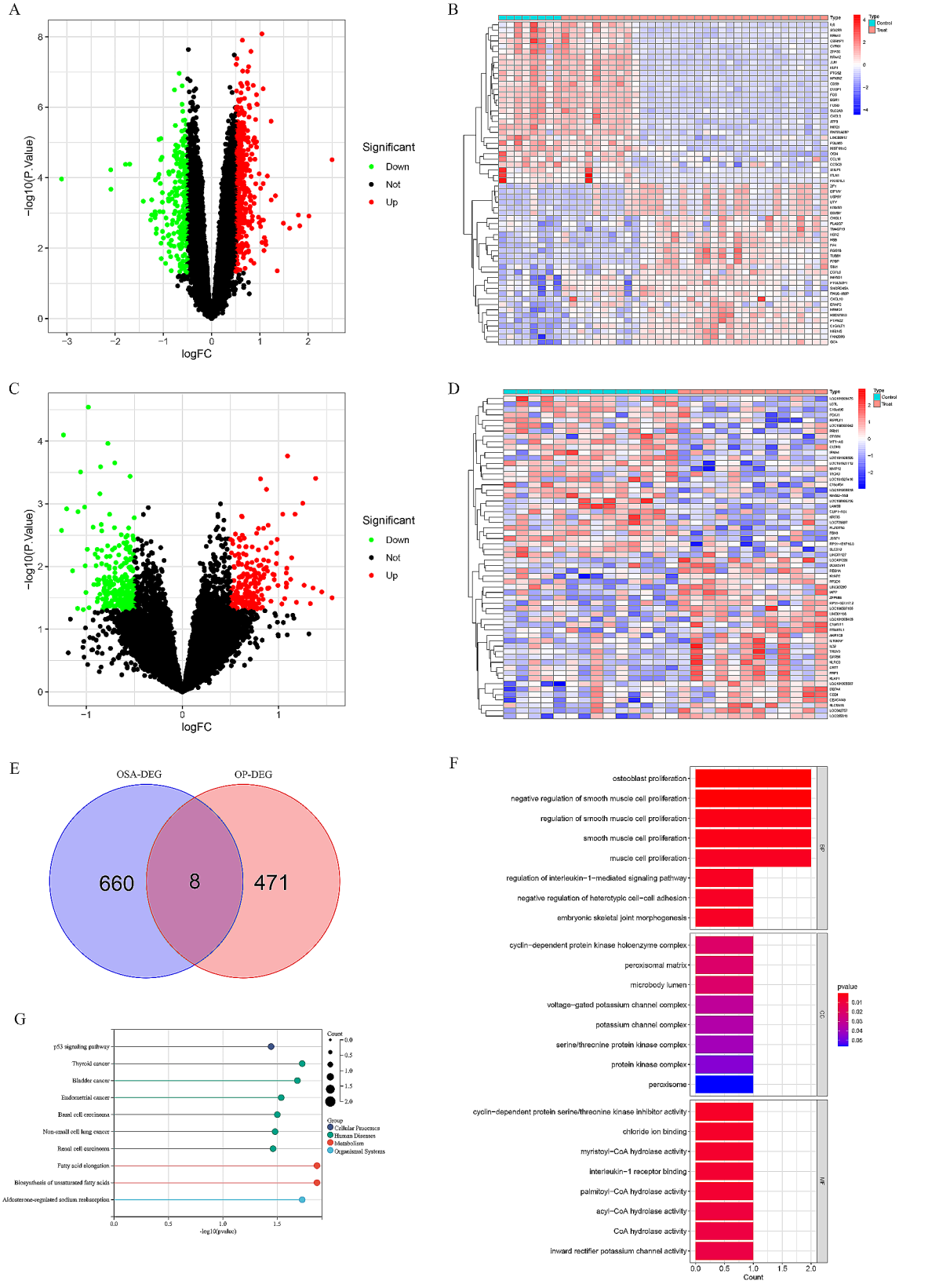

Fig. 3 Identification of key differentially expressed genes (DEGs) in patients with OSA or OP. (**A**) Volcano plot of DEGs in the OSA dataset (GSE135917); (**B**) Heatmap of the top 30 upregulated and 30 downregulated DEGs in the OSA dataset; (**C**) Volcano plot showing DEGs in the OP dataset (GSE7158); (**D**) Heatmap showing the top 30 upregulated and 30 downregulated DEGs in the OP dataset; (**E**) Intersection of OSA-DEGs and OP-DEGs through a Venn diagram; Bar graphs showing the GO (**F**) and KEGG (**G**) enrichment analysis of the eight overlapping DEGs.

of enzymes involved in biochemical reactions (Fig. 3F). KEGG analysis further suggested that these genes might be mainly associated with the p53 signaling pathway, cancer, biosynthesis of fatty acids, and Aldosterone-regulated sodium reabsorption (Fig. 3G).

Feature genes are selected by machine learning

Two machine learning algorithms were used to select feature genes from the eight overlapping genes in OSA and OP. Specifically, in the OSA dataset, five out of the eight overlapping genes were selected using LASSO regression (Fig. 4A, B), and four were selected based on an importance score of ≥1.5 in the RF analysis (Fig. 4E, F), with three common genes present in both OSA-LASSO and OSA-RF (Fig. 4I). In the OP dataset, LASSO regression selected six genes (Fig. 4C, D), and RF analysis based on the same importance score criteria selected four (Fig. 4G, H). Ultimately, three standout genes (*KCNJ1*, *NPR3*, and

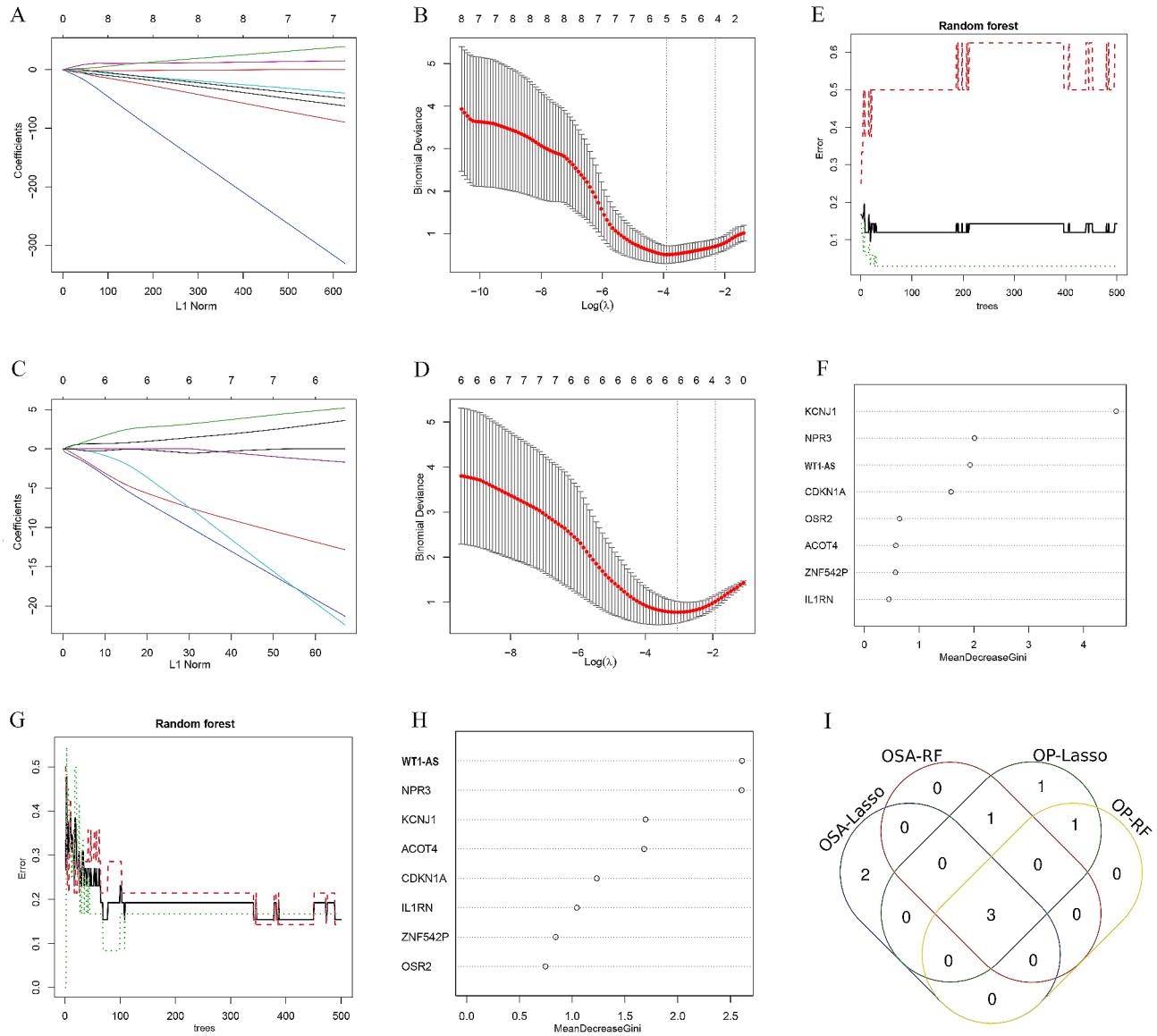

Fig. 4 Identification of potential diagnostic biomarkers for OSA and OP through machine learning methods. LASSO Coefficient Profile in OSA (**A**) and OP (**C**), these plots show the trajectory of each coefficient against the log lambda sequence. Each line, represented by a unique color, corresponds to a different gene. The lines start at zero and either increase or decrease in value as the L1 Norm increases. Cross-Validation for Tuning Parameter Selection in OSA (**B**) and OP (**D**), the red dotted line shows the binomial deviance for different values of log(lambda). The grey area represents the standard error, with vertical dotted lines indicating the minimum criteria and one standard error criteria for selecting the optimal lambda. RF coefficient profiles of candidate genes in OSA (**E**) and OP (**G**), each line represents the change in error rate as more trees are added to the model. The scatter plot illustrates the key genes identified in OSA (**F**) and OP (**H**) using the RF algorithm: each point on the plot represents a gene. The position of each point on the x-axis, indicates the importance of that gene according to the Mean Decrease Gini metric. Higher scores suggest greater importance. (**I**) Venn diagram showing the intersection of DEGs obtained from various machine learning algorithms for OSA and OP.

WT1-AS) were identified as the promising shared diagnostic biomarkers for both OSA and OP (Fig. 4I).

Diagnostic value of biomarkers

To better demonstrate the diagnostic value of KCNJ1, NPR3, and WT1-AS, we evaluated the sensitivity and specificity of these three candidate biomarkers. In the GSE135917 dataset, KCNJ1 showed a higher diagnostic value for OSA (AUC=0.939) compared to NPR3 (AUC=0.846) and WT1-AS (AUC=0.904) (Fig. 5A). In the GSE7158 dataset, WT1-AS demonstrated a higher diagnostic value for OP (AUC=0.881) (Fig. 5B). Subsequently, we performed external validation of the three genes in the validation datasets for OSA (GSE38792) and OP (GSE100609), which also showed similar good predictive performance (Fig. 5C, D). We noted that the diagnostic value of KCNJ1, NPR3, and WT1-AS varied across different datasets. This could be due to the fact that these datasets come from different research centers with varying experimental methods and techniques. Additionally, different populations in the datasets, with varying ages, genders, disease stages, etc., could also influence the size of the AUC values. However, despite these differences, KCNJ1, NPR3, and WT1-AS all demonstrated high diagnostic value across all datasets (training and testing cohorts), further emphasizing their importance as potential common diagnostic markers for OSA and OP.

Characteristic genes and immune infiltration

The gene expression of the three diagnostic genes and their correlation with immune cell infiltration were analyzed separately in the GSE135917 and GSE7158 databases. The results showed that in the OSA data, KCNJ1 was positively correlated with activated Dendritic cells and Plasma cells, but negatively correlated with memory B cells, resting NK cells, and M0 Macrophages (Fig. 6A). WT1-AS was positively correlated with activated Dendritic cells and Plasma cells, but negatively correlated with NK cells resting and M0 Macrophages (Fig. 6C). NPR3 seemed to have little correlation with various immune cell infiltrations (Fig. 6B). In the OP dataset, KCNJ1 was only positively correlated with resting Mast cells (Fig. 6D). NPR3 was positively correlated with activated Dendritic cells but negatively correlated with CD8+T cells (Fig. 6E). WT1-AS was positively correlated with M1 Macrophages but negatively correlated with regulatory T cells (Tregs) (Fig. 6F). In summary, these key diagnostic genes may play a significant role in the immune regulation of the shared disease mechanism between OSA and OP.

Discussion

This study, based on a combination of clinical variables and transcriptome data, revealed the susceptibility of female patients with lower BMI to OP in OSA patients. A nomogram model was constructed based on this finding. Furthermore, *KCNJ1*, *NPR3*, and *WT1-AS* were identified as overlapping key genes for OSA and OP.

OSA and OP are two prevalent diseases affecting a significant portion of the population over 50 years old. A meta-analysis of 15 studies, including 113,082 patients, has shown a close association between OSA and increased risk of OP, as well as a decrease in lumbar spine BMD [19]. Age, gender, and BMI are hot topics as independent risk factors for OSA. In our study, gender and BMI emerged as predictors of OP in OSA patients. Previous studies have shown that both psychological and physiological symptoms of OSA patients have gender specificity [20]. Overall, compared to males, females seem to have more severe symptoms, and vary with age and physiological states (such as menopause and pregnancy). Gender specificity also exists in the OP population, with a significantly higher prevalence in females than in males, related to hormonal metabolic imbalance and a significant decrease in estrogen levels after menopause [21]. Estrogen levels are positively correlated with BMD, can inhibit osteoclast activity, and play a protective role in preventing osteoporotic fractures, but estrogen deprivation in postmenopausal women will eliminate this inhibition and lead to bone loss. However, some studies have shown that menarche age over 17 years and menopause age less than 48 years are risk factors for female OP [22]. The incidence in postmenopausal females is three to four times that of premenopausal females, and estrogen plays an irreplaceable role in this change [23]. Therefore, the cyclical changes of estrogen in females may be a bridge linking OSA and OP.

It has been shown that the severity of OSA increases with weight gain and decreases with weight loss, an increase in BMI can affect respiration in many ways. Infiltration of fat in the upper respiratory tract or its muscles can reduce the size of the pharynx and the strength of the upper respiratory muscles, altering the structure of the upper respiratory tract [24]. However, the relationship between BMI and BMD is complex. It has been fully demonstrated that there is a significant positive correlation between BMI and BMD. Conventionally, gravitational load and muscles can affect the function and metabolism of bone cells (mainly including osteoblasts and osteoclasts), stimulate bone formation, people with low weight have less muscle mass, lighter load on the bones, and are more likely to develop OP [25]. But this does not mean that the risk of fracture is lower for those with a higher BMI, which is also called the "obesity paradox" phenomenon [26]. In children and

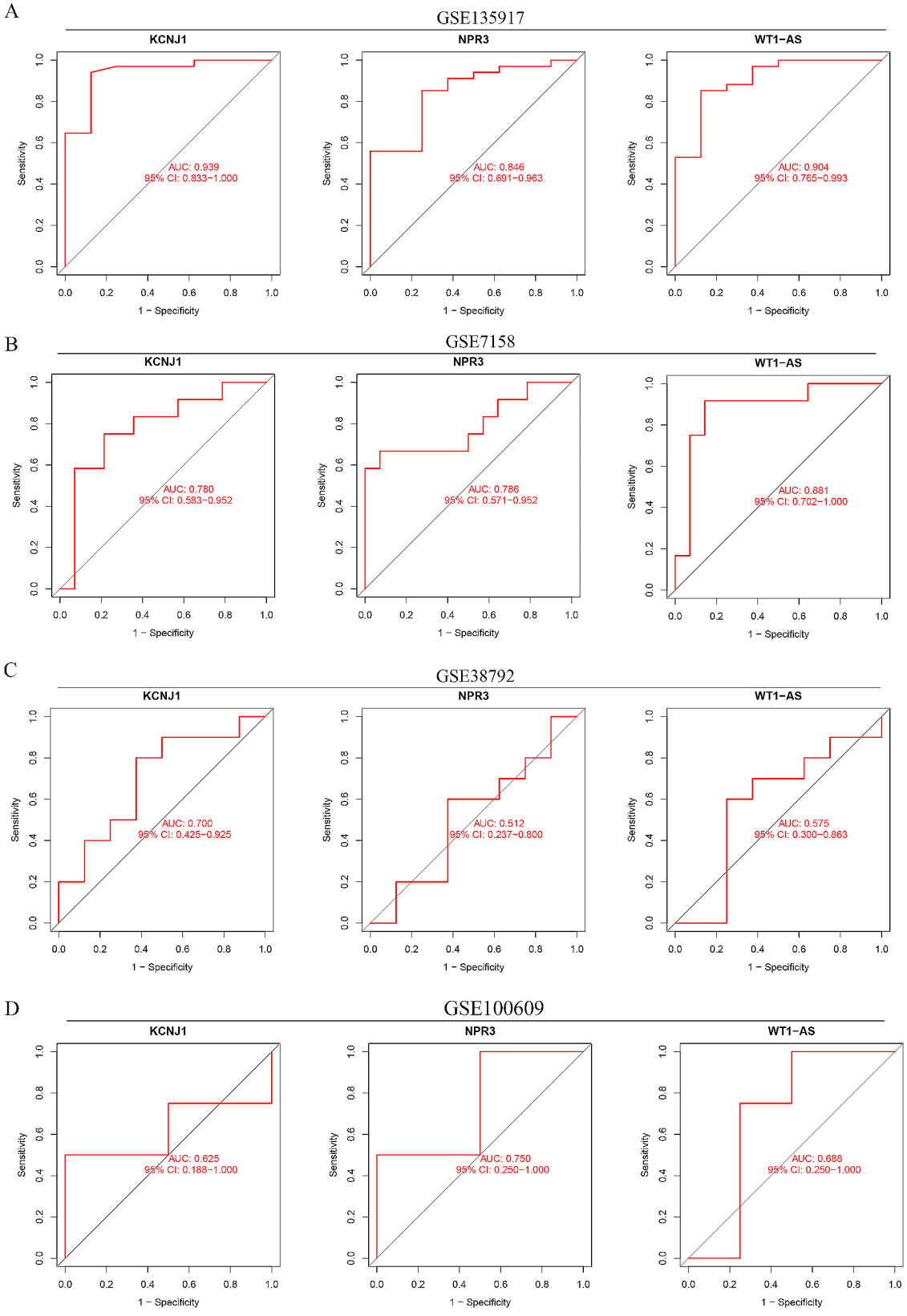

Fig. 5 Verification of the diagnostic values of characteristic genes. The ROC curves show the diagnostic performance of characteristic genes in the training dataset, including OSA (**A**) and OP (**B**); The ROC curves show the diagnostic performance of characteristic genes in the validation dataset, including OSA (**C**) and OP (**D**). Each ROC curve is depicted as a red line. The Area Under the Curve (AUC) value, accompanied by its 95% confidence interval, serves as a reflection of the predictive performance. A higher AUC value signifies superior diagnostic performance

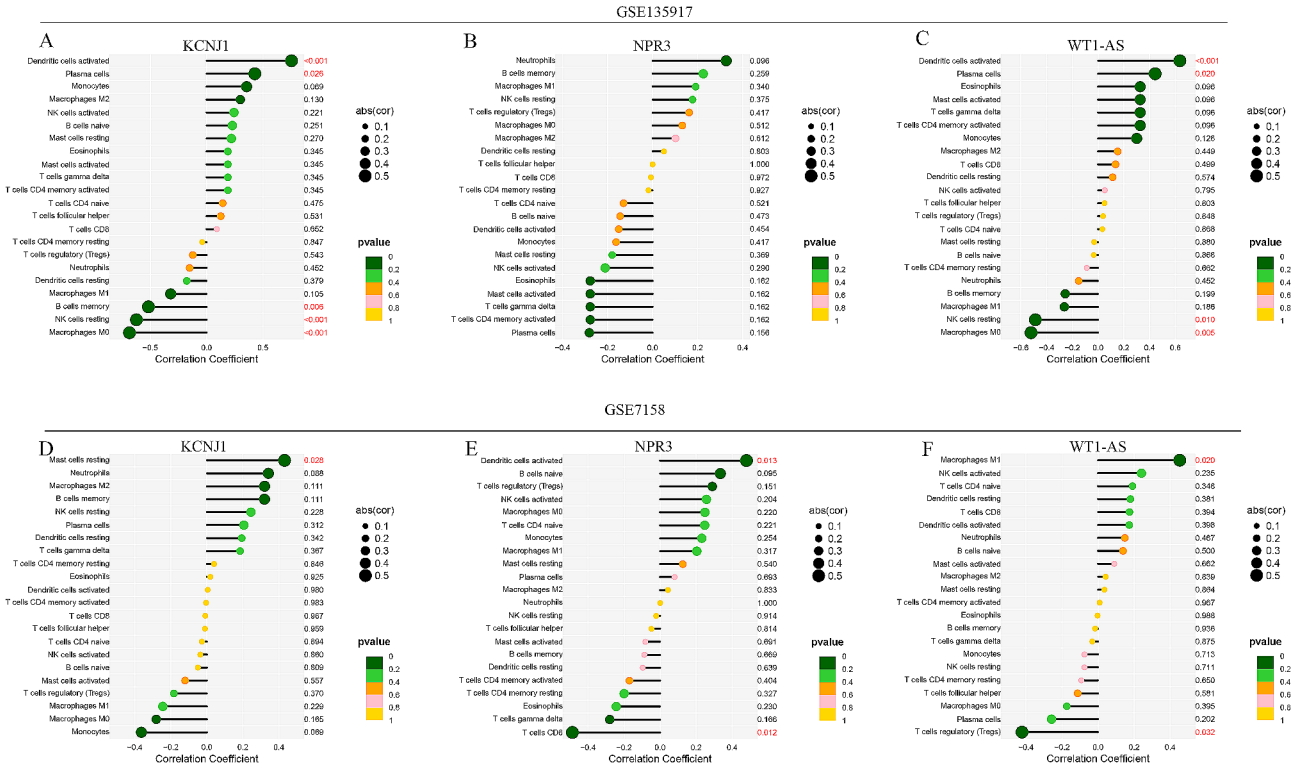

Fig. 6 Analysis of the correlation between immune infiltration and the expression of characteristic genes. (**A-C**) Lollipops represent the correlation between the expression of genes KCNJ1 (**A**), NPR3 (**B**), and WT1-AS (**C**) and immune infiltrating cells in the OSA dataset (GSE135917); (**D-F**) Lollipops represent the correlation between the expression of genes KCNJ1 (**D**), NPR3 (**E**), and WT1-AS (**F**) and immune infiltrating cells in the OP dataset (GSE7158). Each lollipop in the chart represents a distinct type of immune cell. The color of the lollipop approaches green as the statistical significance of the correlation increases. Notably, correlations with a p-value less than 0.05 are highlighted in red. The position of the lollipop on the x-axis indicates the correlation coefficient

adolescents, being overweight can have a positive effect on BMD, but the incidence of fractures is higher than in non-obese individuals. This may be due to the excessive mechanical load produced by extra adipose tissue, and it also shows that there is a saturation value for the appropriate BMD corresponding to BMI. Therefore, we should be cautious about the conclusion that BMI is a risk factor for OP in OSA patients, and future research may need to carefully consider the existence of BMI saturation value or as a critical value of risk factors.

We further intersected the DEGs of OSA and OP relative to the control group, and the eight overlapping genes are mainly enriched in the proliferation regulation of muscle cells and bone cells, cytokine-mediated signaling pathways, and metabolic-related regulation, which lays the foundation for our understanding of the commonality of OSA and OP. This also shows that OSA is likely to affect bone metabolism through multiple mechanisms, possibly partially but not strongly dependent on the related inflammation or oxidative stress we are familiar with. At the same time, the disease prediction factors (*KCNJ1*, *NPR3*, and *WT1-AS*) screened out through machine learning, whether in the training set or validation set of OSA and OP, all show super diagnostic ability. In the immune infiltration analysis, we found that these three markers are mainly positively or negatively correlated with T cells, B cells, and macrophages. As is well known, in addition to the aforementioned effects on bone metabolism, estrogen also interacts with various immune cells, leading to a chronic low-grade pro-inflammatory phenotype under estrogen deficiency, and various immune cells interact with osteoblasts and osteoclasts through direct cell contact or more likely through paracrine mechanisms [27]. Chronic intermittent hypoxia is a marker of OSA and promotes or induces a pro-inflammatory environment. OSA patients also show immune dysfunction to a certain extent, especially in specific cell types such as T cells, monocytes, T cells, dendritic cells, B cells, and neutrophils [28]. Our results seem to suggest that the balance of the immune microenvironment may mediate the occurrence and development of comorbidities.

Natriuretic peptide receptor 3 (*NPR3*), as a clearance receptor that regulates the activity of C-type natriuretic peptide (CNP), is expressed in chondrocytes and osteoblasts. Overproduction of CNP can lead to excessive bone growth and abnormal child skeleton, especially affecting stature, vertebrae, and finger length [29],

which indicates that NPR3 directly participates in bone metabolism regulation. NPR3 plays an important role in regulating blood volume and vascular tension, and may be a feasible therapeutic target to reduce the risk of cancer and cardiovascular diseases in OSA patients, and participate in OSA diagnosis prediction as a gene related to mitochondrial dysfunction [30]. Potassium inwardly rectifying channel subfamily J member 1 (*KCNJ1*), as an ATP-dependent potassium channel, plays an important role in potassium balance. KCNJ1 is related to various diseases, such as prenatal Bartter syndrome, newly diagnosed diabetes, hypotension, and can even inhibit tumor cell proliferation and metastasis and serve as a prognostic factor for renal clear cell carcinoma [31]. At present, there is no report on the direct association of KCNJ1 with OSA or OP, but potassium channels can adjust neuronal activity by affecting resting membrane potential or repolarization, and the development of OSA drugs targeting potassium channel blockers has also achieved remarkable results [32]. Wilms tumor 1 (WT1) is a transcription factor that plays an important role in the development of the urinary and reproductive system [33]. WT1 antisense RNA (*WT1-AS*) originates from the intron region of WT1, has been reported to interact with p53 to regulate the apoptosis of osteoblasts [34]. It is worth noting that current literature lacks reports on the association of these genes with OSA/OP. However, we speculate that these genes may be involved in the onset and progression of OSA and its complications, and then form a regulatory network to collectively influences the pathophysiological changes associated with OP.

Although previous studies have separately reported the risk factors and respective pivotal genes of OSA and OP, few studies have explored the clinical characteristics and overlapping genes of comorbidity. Our exploration further clarifies the comorbidity mechanism of OSA and OP, the shared biomarkers between OSA and OP could assist clinicians in detecting diseases earlier based on patient specificity, potentially enabling precision treatment or the development of gene-targeted therapies. These findings could also guide us in improving disease prevention strategies. However, there are certain limitations, such as the sample data are all ≥50 years old, which is determined by the sample information of the BMD detection in the NHANES database, not a deliberate selection bias; due to the difficulty of actual sample collection, it is impossible to carry out external validation, but the large sample size included in NHANES may compensate for this limitation; in addition, the feature genes screened out may need further experimental verification in the future, which will be the focus of our future work.

Acknowledgements
This work was funded by Hubei Province's Outstanding Medical Academic Leader Program and the Advantages Discipline Group (Public health) Project in Higher Education of Hubei Province (2021–2025).

Author contributions
B. L. and X. H. conceived the research procedure and revised the draft. S. L. and Y. Z. assisted in analyzing the study data and drafting the manuscript. F. L. and W. D. were responsible for software and data downloading. L. Y. and R. Y. were responsible for preparing tables and polishing the language. All authors have read and approved the final manuscript. All authors have contributed to the article and approved the submitted version.

Availability of data and materials
Publicly available datasets were analyzed in this study. These data can be found here: NHANES website (https://www.cdc.gov/nchs/nhanes/) and GEO database (https://www.ncbi.nlm.nih.gov/geo/).

Declarations

Ethics approval and consent to participate
The survey was approved by the National center for health statistics (NCHS) Research Ethics Review Board. All participants provided written informed consent. NCHS Research Ethics Review Board approval was waived for this analysis because of the publicly available and de-identified data. All methods were carried out in accordance with relevant guidelines and regulations (Declaration of Helsinki).

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 27 May 2024 / Accepted: 19 July 2024
Published online: 26 July 2024

References

1. Lv R, Liu X, Zhang Y, Dong N, Wang X, He Y, et al. Pathophysiological mechanisms and therapeutic approaches in obstructive sleep apnea syndrome. Signal Transduct Target Ther. 2023;8(1):218.
2. Lyons MM, Bhatt NY, Pack AI, Magalang UJ. Global burden of sleep-disordered breathing and its implications. Respirology. 2020;25(7):690–702.
3. Benjafield AV, Ayas NT, Eastwood PR, Heinzer R, Ip MSM, Morrell MJ, et al. Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis. Lancet Respir Med. 2019;7(8):687–98.
4. Giampá SQC, Lorenzi-Filho G, Drager LF. Obstructive sleep apnea and metabolic syndrome. Obes (Silver Spring). 2023;31(4):900–11.
5. Noh JY, Yang Y, Jung H. Molecular mechanisms and emerging therapeutics for osteoporosis. Int J Mol Sci. 2020;21(20).
6. Xiao PL, Cui AY, Hsu CJ, Peng R, Jiang N, Xu XH, et al. Global, regional prevalence, and risk factors of osteoporosis according to the World Health Organization diagnostic criteria: a systematic review and meta-analysis. Osteoporos Int. 2022;33(10):2137–53.
7. Alsoof D, Anderson G, McDonald CL, Basques B, Kuris E, Daniels AH. Diagnosis and management of Vertebral Compression fracture. Am J Med. 2022;135(7):815–21.
8. Chen YL, Weng SF, Shen YC, Chou CW, Yang CY, Wang JJ, et al. Obstructive sleep apnea and risk of osteoporosis: a population-based cohort study in Taiwan. J Clin Endocrinol Metab. 2014;99(7):2441–7.
9. Liguori C, Mercuri NB, Izzi F, Romigi A, Cordella A, Piccirilli E, et al. Obstructive sleep apnoea as a risk factor for osteopenia and osteoporosis in the male population. Eur Respir J. 2016;47(3):987–90.
10. Uzkeser H, Yildirim K, Aktan B, Karatay S, Kaynar H, Araz O, et al. Bone mineral density in patients with obstructive sleep apnea syndrome. Sleep Breath. 2013;17(1):339–42.
11. Swanson CM, Shea SA, Stone KL, Cauley JA, Rosen CJ, Redline S, et al. Obstructive sleep apnea and metabolic bone disease: insights into the relationship between bone and sleep. J Bone Min Res. 2015;30(2):199–211.

12. Foley D, Ancoli-Israel S, Britz P, Walsh J. Sleep disturbances and chronic disease in older adults: results of the 2003 National Sleep Foundation Sleep in America Survey. J Psychosom Res. 2004;56(5):497–502.
13. Silverman SL. The clinical consequences of vertebral compression fracture. Bone. 1992;13(Suppl 2):S27–31.
14. Cavallino V, Rankin E, Popescu A, Gopang M, Hale L, Meliker JR. Antimony and sleep health outcomes: NHANES 2009–2016. Sleep Health. 2022;8(4):373–9.
15. Kanis JA, McCloskey EV, Johansson H, Oden A, Melton LJ 3rd, Khaltaev N. A reference standard for the description of osteoporosis. Bone. 2008;42(3):467–75.
16. Huang Q, Wan J, Nan W, Li S, He B, Peng Z. Association between manganese exposure in heavy metals mixtures and the prevalence of Sarcopenia in US adults from NHANES 2011–2018. J Hazard Mater. 2024;464:133005.
17. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
18. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S + to analyze and compare ROC curves. BMC Bioinformatics. 2011;12:77.
19. Wang C, Zhang Z, Zheng Z, Chen X, Zhang Y, Li C, et al. Relationship between obstructive sleep apnea-hypopnea syndrome and osteoporosis adults: a systematic review and meta-analysis. Front Endocrinol (Lausanne). 2022;13:1013771.
20. Bonsignore MR, Saaresranta T, Riha RL. Sex differences in obstructive sleep apnoea. Eur Respir Rev. 2019;28(154).
21. Salari N, Ghasemi H, Mohammadi L, Behzadi MH, Rabieenia E, Shohaimi S, et al. The global prevalence of osteoporosis in the world: a comprehensive systematic review and meta-analysis. J Orthop Surg Res. 2021;16(1):609.
22. He Y, Huang J, Jiang G, Wang H, Zhao J, Chen Z, et al. Menarche age exceed 17 years and menopausal age smaller than 48 years may affect prevalence of osteoporosis for Chinese women. Arch Osteoporos. 2021;16(1):123.
23. Jehan S, Auguste E, Zizi F, Pandi-Perumal SR, Gupta R, Attarian H et al. Obstructive Sleep Apnea: Women's Perspective. J Sleep Med Disord. 2016;3(6).
24. Leppänen T, Kulkas A, Mervaala E, Töyräs J. Increase in body Mass Index decreases duration of Apneas and Hypopneas in Obstructive Sleep Apnea. Respir Care. 2019;64(1):77–84.
25. Lane NE. Epidemiology, etiology, and diagnosis of osteoporosis. Am J Obstet Gynecol. 2006;194(2 Suppl):S3–11.
26. Farr JN, Dimitri P. The impact of Fat and obesity on bone microarchitecture and strength in children. Calcif Tissue Int. 2017;100(5):500–13.
27. Fischer V, Haffner-Luntzer M. Interaction between bone and immune cells: implications for postmenopausal osteoporosis. Semin Cell Dev Biol. 2022;123:14–21.
28. Ludwig K, Huppertz T, Radsak M, Gouveris H. Cellular Immune Dysfunction in Obstructive Sleep Apnea. Front Surg. 2022;9:890377.
29. Gregson CL, Newell F, Leo PJ, Clark GR, Paternoster L, Marshall M, et al. Genome-wide association study of extreme high bone mass: contribution of common genetic variation to extreme BMD phenotypes and potential novel BMD-associated genes. Bone. 2018;114:62–71.
30. Liu Q, Hao T, Li L, Huang D, Lin Z, Fang Y, et al. Construction of a mitochondrial dysfunction related signature of diagnosed model to obstructive sleep apnea. Front Genet. 2022;13:1056691.
31. Guo Z, Liu J, Zhang L, Su B, Xing Y, He Q, et al. KCNJ1 inhibits tumor proliferation and metastasis and is a prognostic factor in clear cell renal cell carcinoma. Tumour Biol. 2015;36(2):1251–9.
32. Wirth KJ, Steinmeyer K, Ruetten H. Sensitization of upper airway mechanoreceptors as a new pharmacologic principle to treat obstructive sleep apnea: investigations with AVE0118 in anesthetized pigs. Sleep. 2013;36(5):699–708.
33. Zhang H, Chen L, Wang Z, Sun Z, Shan Y, Li Q, et al. Long noncoding RNA KCNQ1OT1 inhibits osteoclast differentiation by regulating the miR-128-3p/NFAT5 axis. Aging. 2022;14(10):4486–99.
34. Zhang Y, Na R, Wang X. LncRNA WT1-AS up-regulates p53 to inhibit the proliferation of cervical squamous carcinoma cells. BMC Cancer. 2019;19(1):1052.

Publisher's Note