Observational Study

Medicine®

OPEN

Screening and identification of potential biomarkers for obstructive sleep apnea via microarray analysis

Yuanyuan Cao, PhD[a], Xintian Cai, MS[b], Qing Zhu, MD[b], Nanfang Li, PhD[a,b,*]

Abstract

Obstructive sleep apnea (OSA) is a common chronic disease and increases the risk of cardiovascular disease, metabolic and neuropsychiatric disorders, resulting in a considerable socioeconomic burden. This study aimed to identify potential key genes influence the mechanisms and consequences of OSA.

Gene expression profiles related to OSA were obtained from Gene Expression Omnibus (GEO) database. Differentially expressed genes (DEGs) in subcutaneous adipose tissues from OSA compared with normal tissues were screened using R software, followed by gene ontology (GO) and pathway enrichment analyses. Subsequently, a protein-protein interaction (PPI) network for these DEGs was constructed by STRING, and key hub genes were extracted from the network with plugins in Cytoscape. The hub genes were further validated in another GEO dataset and assessed by receiver operating characteristic (ROC) analysis and Pearson correlation analysis.

There were 373 DEGs in OSA samples in relative to normal controls, which were mainly associated with olfactory receptor activity and olfactory transduction. Upon analyses of the PPI network, GDNF, SLC2A2, PRL, and SST were identified as key hub genes. Decreased expression of the hub genes was association with OSA occurrence, and exhibited good performance in distinguishing OSA from normal samples based on ROC analysis. Besides, the Pearson method revealed a strong correlation between hub genes, which indicates that they may act in synergy, contributing to OSA and related disorders.

This bioinformatics research identified 4 hub genes, including GDNF, SLC2A2, PRL, and SST which may be new potential biomarkers for OSA and related disorders.

Abbreviations: AHI = apnea-hypopnea index, DEGs = differentially expressed genes, GDNF = glial cell derived neurotrophic factor, GEO = Gene Expression Omnibus, GO = gene ontology, KEGG = Kyoto Encyclopedia of Genes and Genomes, OSA = obstructive sleep apnea, PCA = principal component analysis, PPI = protein-protein interaction, PRL = prolactin, ROC = receiver operating characteristic, SLC2A2 = solute carrier family 2 member 2, SST = somatostatin.

Keywords: biomarkers, differentially expressed genes, microarray analysis, obstructive sleep apnea

1. Introduction

Obstructive sleep apnea (OSA) is a sleep respiratory disturbance disease characterized by nocturnal sleep snoring, apnea, and daytime sleepiness, which leads to intermittent hypoxemia, transitory hypercapnia and sleep structural disorder. The main clinical risk associated with OSA is multiple organ system damage, such as cardio-cerebrovascular diseases, metabolic syndrome, and cognitive dysfunction.[1–3] The prevalence of OSA has increased among the general population, and moderate to severe OSA has been estimated to be as high as 49.7% in males and 23.4% in females.[4] Unfortunately, the vast majority patients with OSA (70%–90%) remain undiagnosed, resulting in a heavy health and socioeconomic burden.[5] Therefore, it is an urgent task to study the etiology and pathogenesis of OSA and to find indicators for early diagnosis and treatment targets.

Strong genetic influence has been reported for OSA, with more than 1.5-fold increased risk in first-degree relatives of patients.[6] Approximately 35% to 40% of variation in apnea-hypopnea index (AHI), which measures apnea severity, can be explained by genetic factors.[7] Genetic research has yielded new insights into understanding the mechanism underlying OSA to develop a more effective target therapy and optimize treatment strategies. Most efforts to date to identify the genetic cause of OSA have taken the candidate gene approach, based on 4 intermediate pathogenic pathways: obesity and body fat distribution, craniofacial morphology, ventilatory control, control of sleep and circadian

Editor: Giuseppe Insalaco.

This study was supported by the Natural Science Foundation of Xinjiang Uygur Autonomous Region (grant no. 2017D01C129), and the National Health Committee Key Laboratory of Hypertension Clinical Research of China.

The authors declare that there are no conflicts of interests.

Supplemental Digital Content is available for this article.

The datasets generated during and/or analyzed during the current study are publicly available.

[a] Xinjiang Medical University, [b] Hypertension Center of People's Hospital of Xinjiang Uygur Autonomous Region, Xinjiang Hypertension Institute, National Health Committee Key Laboratory of Hypertension Clinical Research, Urumqi, Xinjiang, China.

[*] Correspondence: Nanfang Li, Hypertension Center of People's Hospital of Xinjiang Uygur Autonomous Region, Xinjiang Hypertension Institute, National Health Committee Key Laboratory of Hypertension Clinical Research, No.91 Tianchi Road Tianshan District, Urumqi 830001, Xinjiang, China (e-mail: lnanfang2016@sina.com).

How to cite this article: Cao Y, Cai X, Zhu Q, Li N. Screening and identification of potential biomarkers for obstructive sleep apnea via microarray analysis. Medicine 2021;100:4(e24435).

Received: 18 September 2020 / Received in final form: 30 December 2020 / Accepted: 5 January 2021

http://dx.doi.org/10.1097/MD.0000000000024435

rhythm.[7,8] However, this method limits the validity of the findings because these studies rely on prior hypotheses about disease mechanism, which precludes discovery of genetic variation in previously unknown pathways.[9]

Recently, the genome-wide DNA microarray based on high-throughput platforms for gene expression analysis, has emerged as an efficient and relatively economical tool to study complex disease genetics.[10] Therefore, we compared gene expression profiles in subcutaneous adipose tissue between OSA and control from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo) for screening differentially expressed genes (DEGs). Subsequently, the DEGs were identified using a combination of functional enrichment and protein-protein interaction (PPI) analyses. The reliabilities of the identified hub genes were further validated in another independent dataset and assessed by receiver operating characteristic (ROC) analysis to determine the predictive, diagnostic and therapeutic value of these genes for OSA.

2. Materials and methods

2.1. Microarray dataset source

A systematic retrieval of gene expression microarray datasets from the National Center for Biotechnology Information GEO (http://www.ncbi.nlm.nih.gov/geo/) was performed to examine DEGs between OSA and normal. The key words "obstructive sleep apnea" was used for the screening. The gene expression profile of GSE135917 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE135917)[11] was included in the present study, which was based on GPL6244 platform (HuGene-1_0-st; Affymetrix Human Gene 1.0 ST Array). This dataset contained 18 subcutaneous adipose tissue samples obtained from the abdominal subcutaneous fat biopsy of 10 OSA patients and 8 normal controls during ventral hernia repair surgery, and the baseline characteristics of the OSA and control samples are presented in Supplemental Table S1, http://links.lww.com/MD/F598.[11] OSA severity was assessed using the ARESTM Unicorder (Watermark Medical, Boca Raton, FL) a previously validated portable sleep monitor worn 2 consecutive nights prior to surgery.[11]

A further independent dataset, GSE38792 (https://www.ncbi.nlm.nih.gov/geo/ query/acc.cgi?acc=GSE38792)[12] including microarray data from visceral fat biopsies obtained intraoperatively from the momentum of 10 OSA patients and 8 normal controls, was used to validate the results obtained from the GSE135917 dataset. Similarly, the GSE38792 dataset was based on the GPL6244 platform. All samples enrolled in our study were obtained from publicly available database which allowed researchers to download and analyze datasets for scientific purposes, and ethical review and approval were thus not required.

2.2. Data preprocessing and differential expression analysis

R software (version 3.6.2; https://www.r-project.org/) together with packages available from Bioconductor (http://www.bioconductor.org/) was used to perform statistical analyses. According to the Affymetrix platform, preprocessing, normalization and quality control of the microarray datasets with raw data (.CEL files) were performed using the Affy package[13] in R. The robust multichip average (RMA) method was used for background correction, quantile normalization and median polish summarization. According to the annotation platform, probe IDs were converted into gene symbols. Probe sets without corresponding gene symbols or genes with more than one probe set were removed or averaged.

The Linear Models for Microarray Data (LIMMA) package[14] in R Bioconductor was used to identify DEGs [$|\log_2$Fold Change$| \geq 1$; false discovery rate (FDR) <0.05] between OSA and normal controls. FDR was controlled based on the Benjamini-Hochberg method and empirical Bayes-modified t-tests were performed to select sets of DEGs. Moreover, principal component analysis (PCA) was performed and visualized with the rgl and pca3d R packages to examine separation of the OSA and normal groups based on differential gene expression. Heat map and volcano plot of DEGs were generated in R using the ggplot2 and pheatmap packages.

2.3. Biological functions and pathway enrichment analyses

To further elucidate the biological functions of DEGs between OSA and normal controls, the identified DEGs were subjected to Gene Ontology (GO) term enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses using cluster Profiler package (version 3.10.0)[15] and org.Hs.e.g,.db annotation package (version 3.10.0)[16] in R. The significant GO terms or KEGG pathways were enriched by more than 2 genes with the threshold of FDR <0.05. Furthermore, pathway network analysis may reveal the possible crosstalk interactions among pathways associated with the development of OSA. Therefore, functionally grouped GO term/pathway networks were constructed based on the identified DEGs on Cytoscape (version 3.7.2) using Clue Go (version 2.5.6)[17] with kappa score = 0.4 and P value $\leq .05$.

2.4. Protein-protein interaction network analysis and hub genes selection

Potential interactions among the DEGs encoding proteins were predicted with the aid of the online database STRING (version 11.0; http://string-db.org), and a combined interaction score >0.4 was considered statistically significant. Subsequently, the protein-protein interaction (PPI) network was visualized using Cytoscape (version 3.7.2). Molecular Complex Detection (MCODE; version 1.6) plugin of Cytoscape was used for searching the most denset and significant module in the PPI network with criteria as follows: degree cutoff = 2, node score cutoff = 0.2, K-core = 2, and max depth = 100. CytoHubba (version 0.1) plugin was used to select and identify genes ranking top ten based on eleven topological analysis algorithms.[18] The intersection of the top gene sets revealed the final hub genes displayed through a Venn diagram created with OmicShare tools (http://www.omicshare.com/tools). Further, functional annotation of hub genes was carried out using DAVID database (version 6.8; https://david.ncifcrf.gov/).

2.5. Validation and analysis of hub genes

GSE38792, another independent microarray dataset, was utilized to verify the differential expression of these hub genes between OSA and normal controls, and the gene expression visualization was performed by pheatmap R package. Meanwhile, the expression level of hub genes based on GSE135917 and GSE38792 datasets was analyzed by ggpubr and ggplot2 R packages. Subsequently, to identify diagnostic value of hub genes

and optimum cutoff points at the gene expression level for predicting OSA, we integrated the above two microarray datasets to screen out expression of hub genes in all samples and divided them into OSA and control groups; the gene expression data was imported into R software and receiver operating characteristic (ROC) curves were plotted, and the area under the curve (AUC) was calculated for each hub gene using the pROC package (AUC >0.5 indicated good diagnostic value). Furthermore, to explore co-expression among hub genes, a heat map of Pearson correlation coefficient matrix was plotted to visualize the correlations between different genes by using the ggcorrplot and ggthemes packages of R. Genes with an absolute value of correlation coefficient >0.4 and P value <.05 were identified as the related genes.

3. Results

3.1. Identification of DEGs

In this reanalysis, the raw microarray data from GSE135917 was subjected to RMA preprocessing and then normalized to the median of all samples. The box plot of each sample before and after normalization demonstrated that the chip data had been normalized and were available for DEGs selection (Fig. 1). Following data preprocessing, a total of 23281 gene expression values were obtained from the 18 samples. Based on the calculating criteria of absolute $\log_2$FC ≥1 and FDR <0.05, there were 373 DEGs in OSA samples in relative to normal controls, including 342 down-regulated genes and 31 up-regulated genes. As shown in Figure 2, the volcano plot (Fig. 2A) intuitively exhibited the distribution of DEGs and the heat map (Fig. 2B) based on the expression level of DEGs showed the result of bidirectional hierarchical clustering of DEGs and samples. Moreover, DEGs clearly separated OSA samples from controls on the PCA plots (Fig. 2C, D).

3.2. Functional and pathway enrichment analyses of DEGs

GO enrichment and KEGG pathway analyses were performed using the cluster Profiler R package, and only the GO terms and KEGG pathways enriched with an adjusted *P* value <.05 (Benjamini-Hochberg correction for multiple testing) were considered. The GO functional enrichment resulted in a total of 373 DEGs mapped to 4 GO terms including 3 biological process (BP) terms and 1 molecular functional (MF) terms (Fig. 3A). Enriched GO terms are mainly associated with olfactory receptor activity and detection of chemical stimulus involved in sensory perception of smell. Furthermore, KEGG

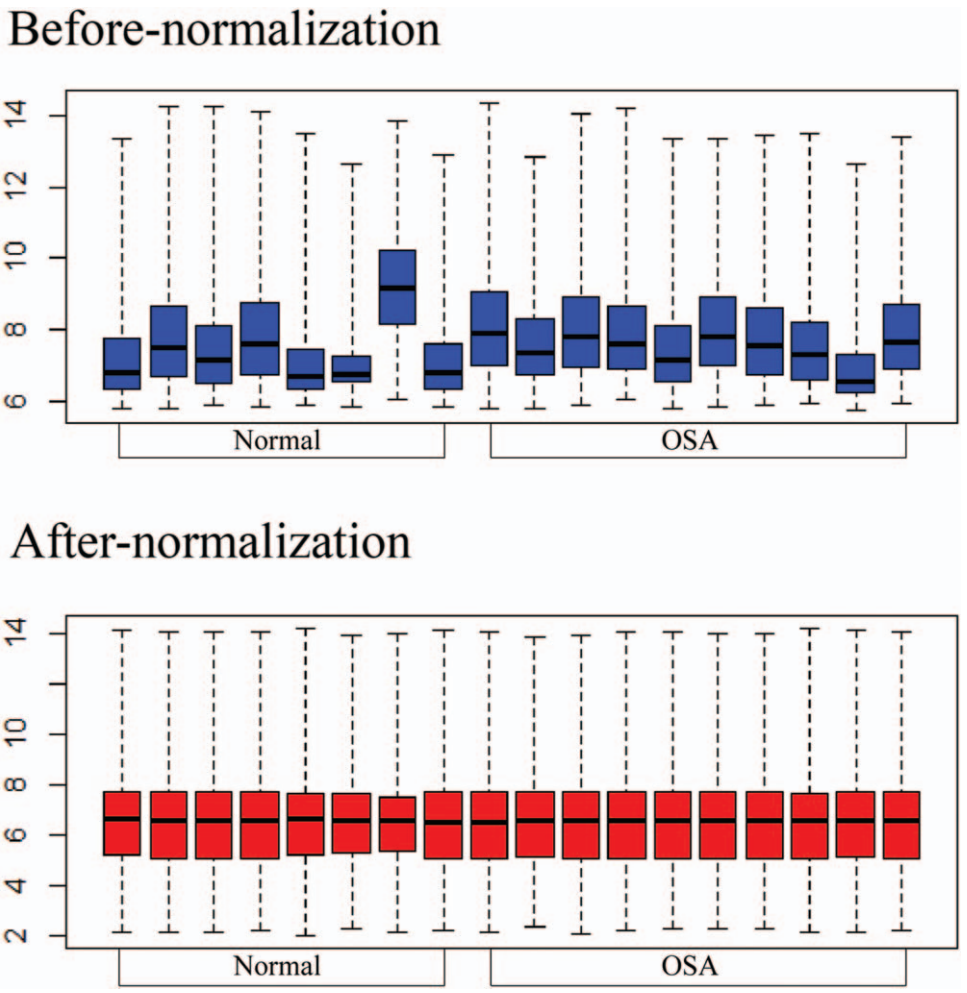

Figure 1. Data preprocessing. The distribution diagrams of gene expression values before and after normalization.

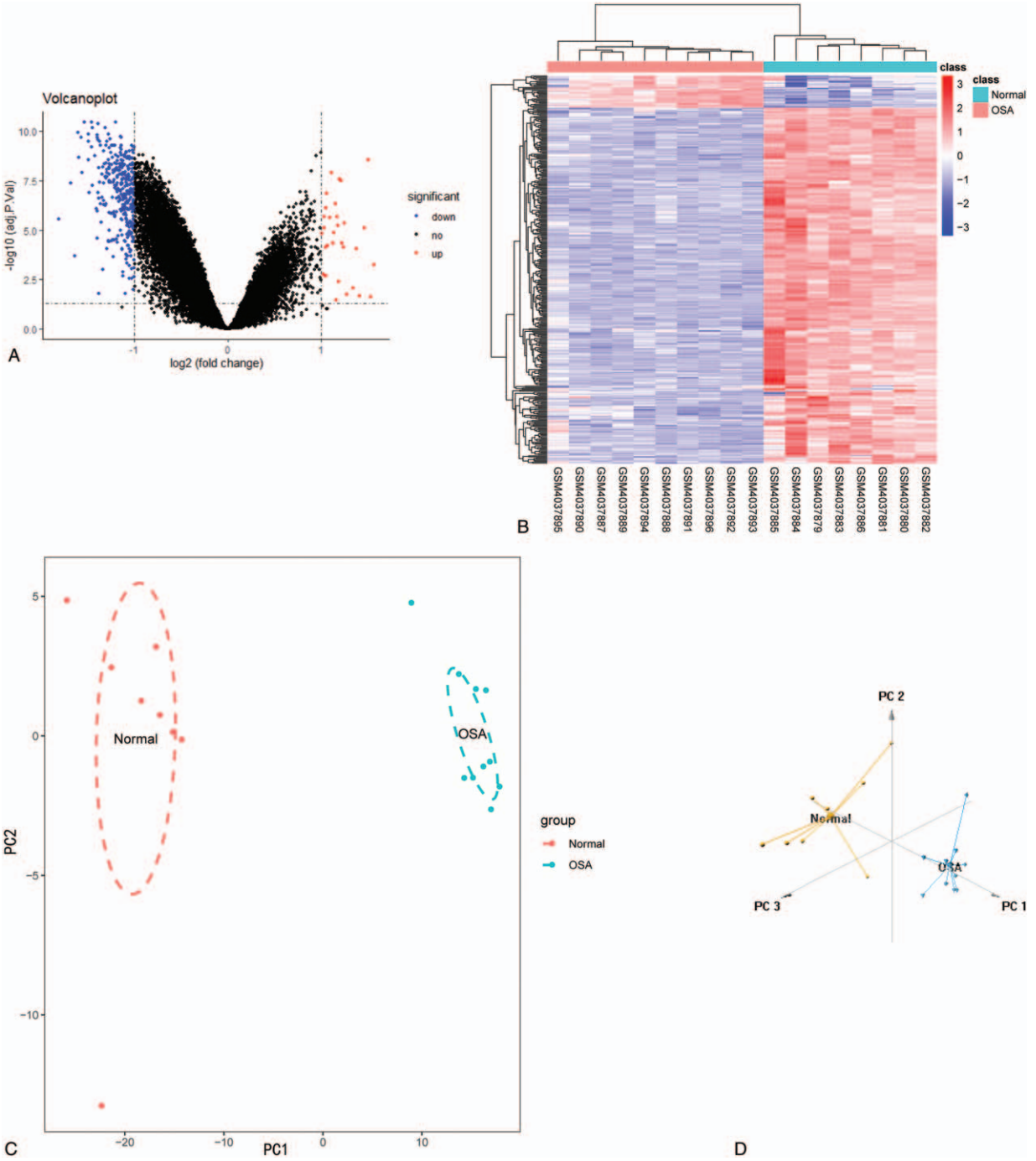

Figure 2. DEGs screening and visualization for GSE135917. (A) Volcano plot of the DEGs for OSA samples vs. Normal controls. Each dot represents a gene (red: up-regulated gene; blue: down-regulated gene; black: no significant difference). (B) Heat map and clustering pattern of the DEGs including 342 down-regulated and 31 up-regulated genes. Red or blue indicates either higher or lower expression levels of DEGs. Samples were separated into the normal and OSA cluster. (C) 2-dimensional PCA and (D) 3-dimensional PCA of DEGs. The scatter plots showed a good separation of OSA and normal samples along the first two or three PC. DEGs = differentially expressed genes, OSA = obstructive sleep apnea, PCA = principal component analysis.

pathway analysis demonstrated that DEGs were significantly enriched in olfactory transduction and neuroactive ligand-receptor interaction pathways (Fig. 3B). In addition, to examine possible networks and relationships among enriched terms, ClueGO visualization and analysis of biological role (GO, KEGG pathways) was undertaken, shown in Figure 3C,D, respectively.

3.3. PPI network analysis and hub genes selection

A PPI network comprising 90 nodes and 86 edges was constructed based on the biological interactions of 373 DEGs to further identify their associations at the protein level (Fig. 4A). As shown in the PPI network, 11 up-regulated and 79 down-regulated genes were screened out in patients with OSA. Subsequently, we combined the results of MCODE and

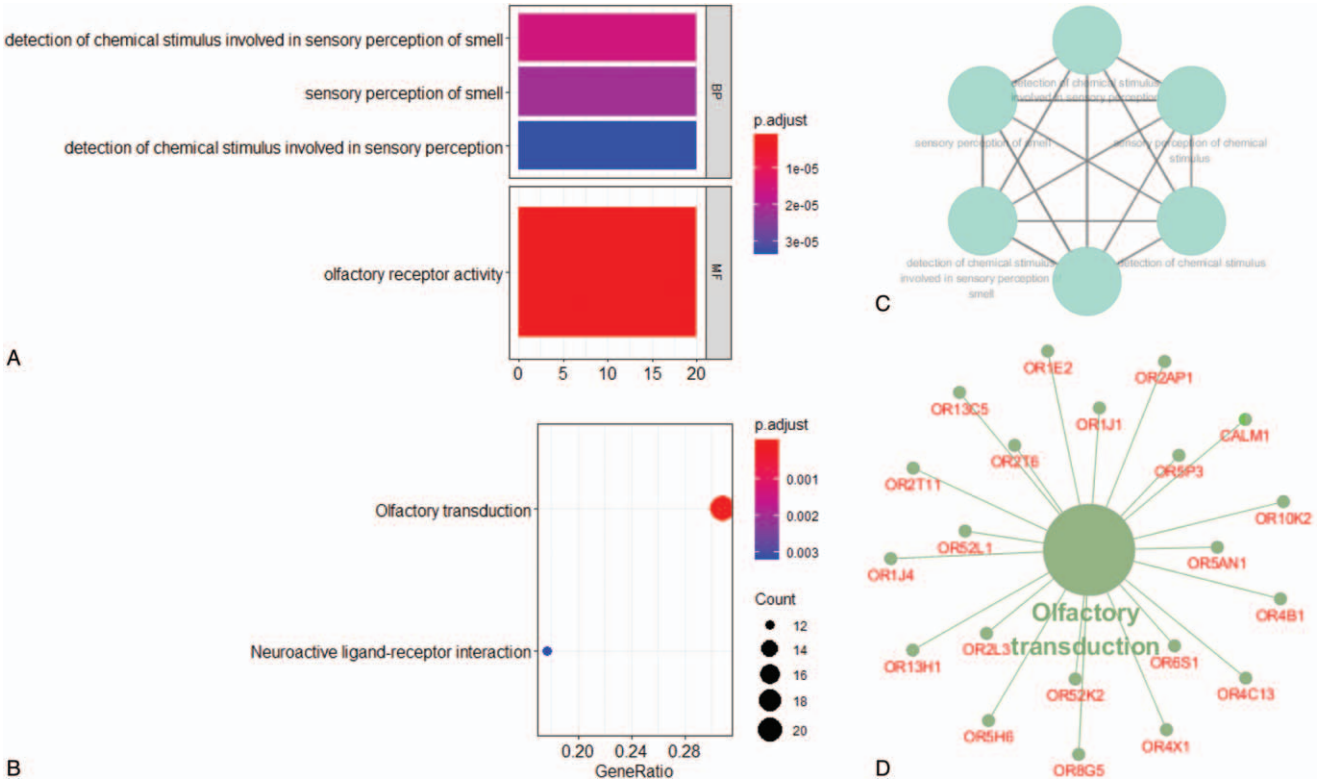

Figure 3. Enrichment analysis of the DEGs using cluster Profiler R package. (A) GO enrichment analysis. (B) KEGG pathway analysis. Four significant GO terms (3 BP and 1MF) and two KEGG pathways were identified with the cutoff criteria of adjusted *P* value <.05. DEGs = differentially expressed genes, GO = gene ontology, KEGG = Kyoto Encyclopedia of Genes and Genomes, BP = biological process, MF = molecular function. Enrichment analysis of the DEGs using ClueGO plugin. Functionally grouped GO term (C) and KEGG pathway (D) networks were constructed based on the DEGs (kappa score = 0.4 and *P* value ≤.05). DEGs = differentially expressed genes, GO = gene ontology, KEGG = Kyoto Encyclopedia of Genes and Genomes.

cytoHubba analyses, and selected genes ranking top ten based on nine topological analysis algorithms (Table 1). Venn diagram analysis was used to determine the intersection of the nine gene sets, as presented in Figure 4B. Finally, the final hub genes including GDNF, SLC2A2, PRL, and SST were identified and presented in Table 2, providing functional annotation for these genes by DAVID.

3.4. Validation of hub genes for OSA identification

To validate the reliability of the results obtained from GSE135917 dataset, the heat map and hierarchical clustering analysis of hub genes, based on data from an independent GSE38792 dataset was performed, and the results showed that the 4 hub genes with low expression can basically distinguish OSA from normal samples (Fig. 5A). Meanwhile, we compared the expression level of these hub genes between OSA and normal controls in the 2 datasets respectively, and found that all 4 genes showed a significant difference (all *P* values <.05), as presented in Figure 6. Subsequently, the ROC curves were constructed to identify diagnostic value of per hub gene and optimum cutoff points at the gene expression level for predicting OSA (Fig. 5B). The AUC as well as sensitivity, specificity and cut-off value at the best performance were summarized in Table 3, which suggested that the 4 hub genes can be used as indicators for OSA identification. Furthermore, the heat map on Pearson correlation was used to show the co-expression among hub genes, indicating that the expression between any 2 hub genes were highly positive correlated, with correlation coefficients ranging from 0.83 to 0.94 and all *P* values <.001(Fig. 5C).

4. Discussion

Currently, bioinformatics has become increasingly important for exploring the pathogenesis of multifactorial disorders.[19] The identification of gene expression changes in tissues relevant to a disease is an important step toward facilitating our understanding of pathogenesis, and eventually lead to better diagnosis and treatment.[19] In the present study, we identified 31 up-regulated and 342 down-regulated genes in subcutaneous adipose tissue of OSA patients compared with normal controls, which suggested that the occurrence and development of OSA is a complex biological process involving multiple genes and steps. Biological function and pathway enrichment analyses indicated that DEGs were mainly involved in the GO terms or KEGG pathways associated with olfactory receptor activity and olfactory transduction. However, the association of the olfactory pathway with OSA has not been previously studied. Besides, GDNF, SLC2A2, PRL, and SST were identified as the key hub genes through PPI network construction. The low expression levels of these 4 hub genes were further validated in an additional dataset, and can basically distinguish OSA from normal samples. Furthermore, ROC curves of these genes exhibited larger AUCs, suggesting that they had good diagnostic value for OSA. Therefore, the present study considered these 4 genes were significant and needed to be discussed.

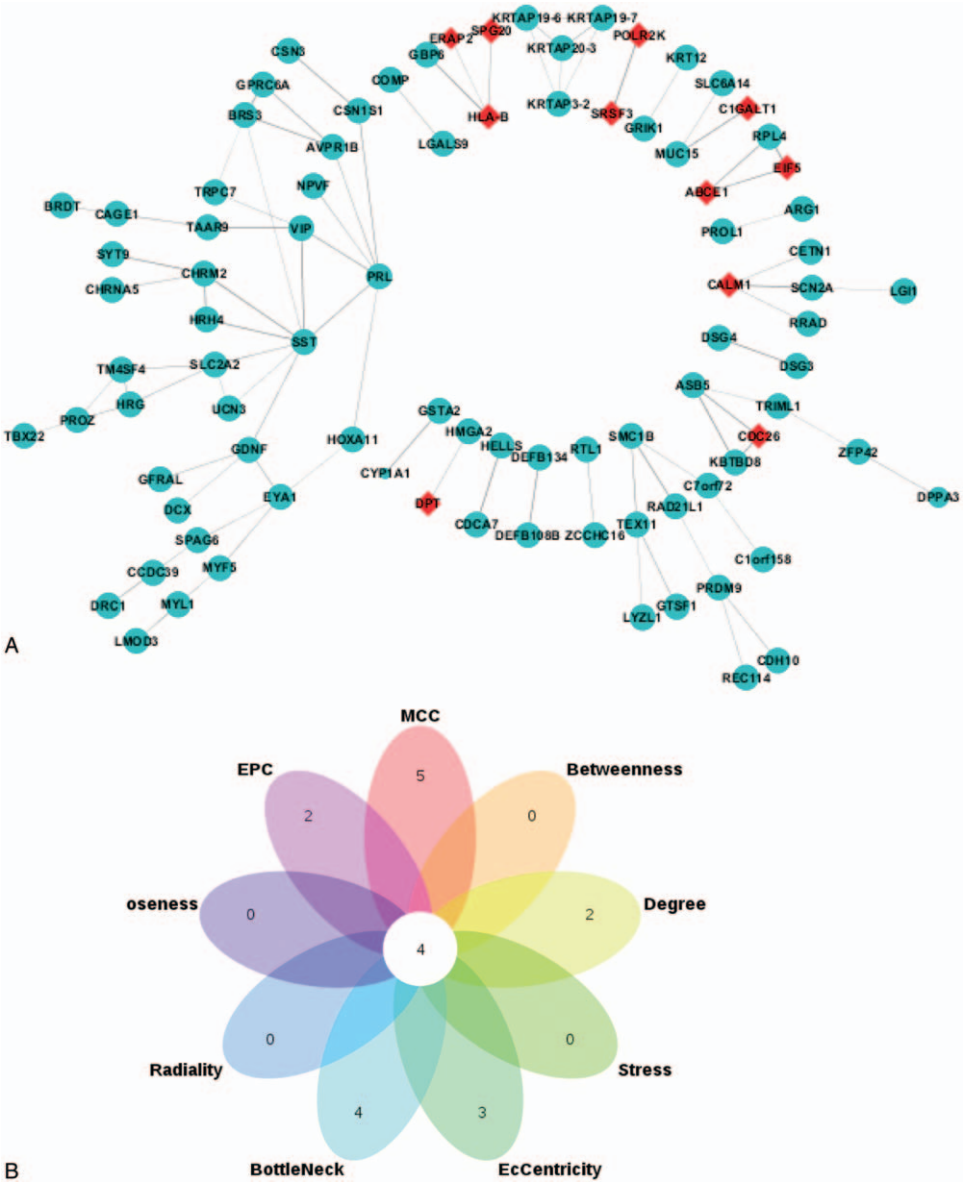

Figure 4. PPI Network Analysis and Hub Genes Selection. (A) PPI network of the DEGs. Blue circles and red diamonds represent down-regulated genes and up-regulated genes, respectively; the size of circles or diamonds indicates *P* values, with larger circles or diamonds representing smaller *P* values. (B) Venn diagram of gene sets ranking top ten based on nine topological analysis algorithms in cytoHubba plugin. The intersection of the top gene sets revealed four hub genes. PPI = protein-protein interaction, DEGs = differentially expressed genes.

Table 1

Top 10 genes by nine ranked methods respectively in cytoHubba.

	Rank methods in cytoHubba								
	MCC	**Betweens**	**Degree**	**Stress**	**Eccentricity**	**Bottleneck**	**Radiality**	**Closeness**	**EPC**
TOP 10 gene	SST	SST	SST	SST	PRL	SST	SST	SST	SST
	PRL	EYA1	PRL	EYA1	GDNF	GDNF	PRL	PRL	PRL
	KRTAP20-3	GDNF	SLC2A2	GDNF	SST	PRL	GDNF	GDNF	VIP
	KRTAP19-7	PRL	GDNF	PRL	SLC2A2	RAD21L1	VIP	VIP	BRS3
	KRTAP19-6	SLC2A2	EYA1	SLC2A2	EYA1	SMC1B	SLC2A2	SLC2A2	SLC2A2
	KRTAP3-2	VIP	CHRM2	VIP	GFRAL	EYA1	HOXA11	BRS3	AVPR1B
	HRG	CHRM2	BRS3	TAAR9	NPVF	HOXA11	EYA1	EYA1	GDNF
	SLC2A2	MYF5	VIP	MYF5	UCN3	TRIML1	BRS3	CHRM2	CHRM2
	GDNF	SPAG6	TEX11	SPAG6	HOXA11	ASB5	UCN3	HOXA11	UCN3
	EYA1	TAAR9	PROZ	HOXA11	DCX	SLC2A2	CHRM2	UCN3	HRH4

MCC, Between, Degree, Stress, Eccentricity, Bottleneck, Radiality, Closeness and EPC are different algorithm names.

Table 2

Functional roles of 4 hub genes.

Gene symbol	Full name	Function
GDNF	Glial cell derived neurotrophic factor	Neurotrophic factor that enhances survival and morphological differentiation of dopaminergic neurons and increases their high-affinity dopamine uptake
SLC2A2	Solute carrier family 2 member 2	Facilitative glucose transporter
PRL	Prolactin	Prolactin acts primarily on the mammary gland by promoting lactation
SST	Somatostatin	Somatostatin inhibits the release of somatotropin

GDNF, glial cell-derived neurotrophic factor, belongs to the transforming growth factor family, which is crucial for motor neurons, dopaminergic and peripheral neurons.[20] A large candidate gene study for OSA[8] supported a potential pathogenic role for polymorphisms in GDNF gene in European Americans. The associations of GDNF with OSA still remained after adjustment for BMI, implying that these genetic variants influence OSA susceptibility through obesity-independent pathways. Larkin et al[8] argued that GDNF variants are related to the pathogenesis of OSA via ventilatory control abnormalities for the following reasons. First, ventilatory control abnormalities may predispose to OSA by worsening ventilatory instability, impairing the arousal response to airway obstruction, or contributing to imbalanced activation of upper airway muscles compared with chest wall muscles. Additionally, GDNF plays a critical role in the development of neural pathways (e.g., noradrenergic neurons development and differentiation) vital for normal respiration. Moreover, GDNF seems to play a trophic role for sensory afferent neurons in the carotid body, which may be important in development of hypoxic responses. Finally, knockout of GDNF gene results in abnormal central respiratory output, and severe mutations in GDNF are associated with the central congenital hypoventilation syndrome. Likewise, as shown in our present study, the GDNF expression was decreased in OSA samples, and had the highest AUC for predicting OSA. Notably, Larkin et al results were not successfully replicated in a large Icelandic OSA cohort,[21] probably because they had different study population or environmental factors. So the association between GDNF and OSA remains controversial and warrants further research.

SLC2A2 (solute carrier family 2 subfamily A member 2) encodes the facilitative glucose transporter isoform GLUT2, which is expressed in liver, kidney, intestine, pancreatic islet beta cells and the central nervous system.[22] GLUT2 facilitates the passive transport of glucose across plasma membranes, which is important for the basis of an integrated inter-organ communication system to control glucose homeostasis, including the control

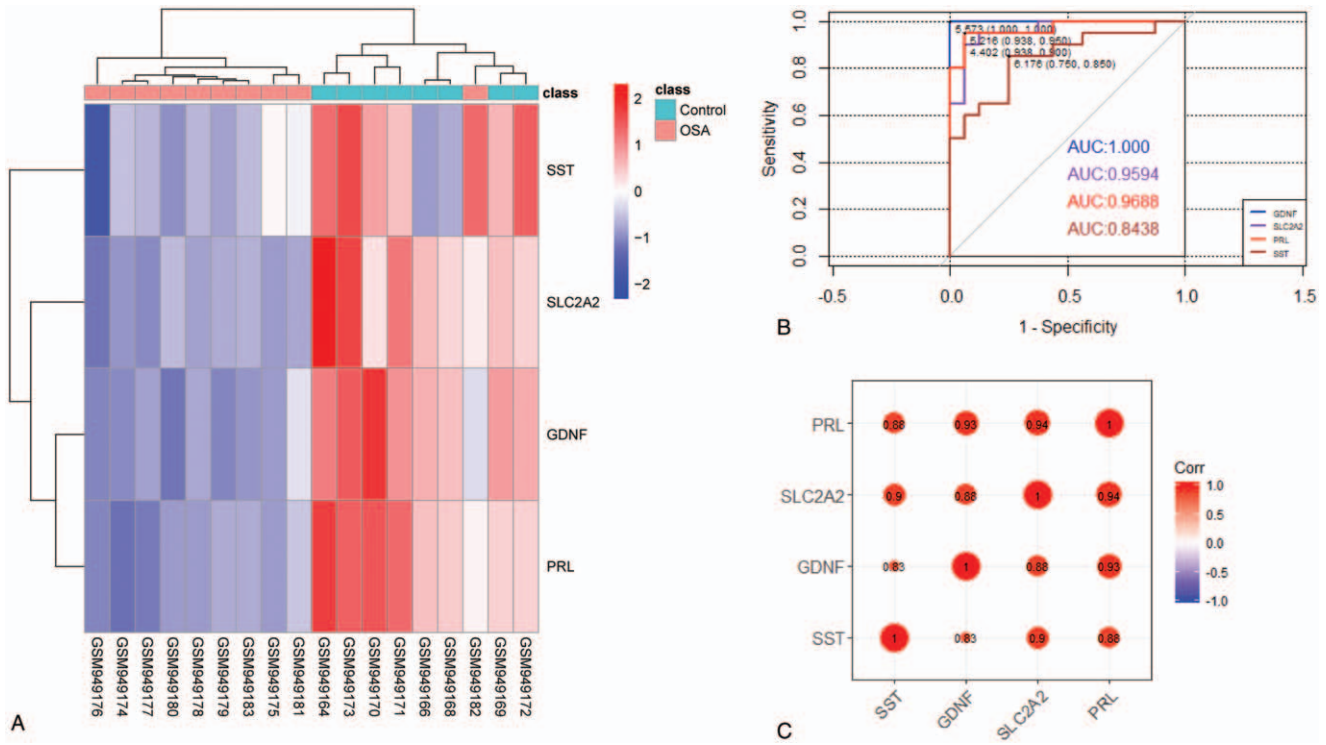

Figure 5. Validation of the hub genes. (A) Hierarchical clustering heat map of the four hub genes in the GSE38792 validation dataset. Red color, up-regulated; blue color, down-regulated. (B) ROC curves of the four hub genes to distinguish OSA from normal samples using data from GSE135917 and GSE38792 datasets. (C) Correlation analysis among the four hub genes. Heat maps showing the correlations between hub genes in GSE135917 and GSE38792 datasets. Color and size of circles indicates the correlations, with Pearson correlation coefficients superimposed on circles. ROC = receiver operating characteristic, OSA = obstructive sleep apnea, GDNF = glial cell derived neurotrophic factor, SLC2A2 = solute carrier family 2 member 2, PRL = prolactin, SST = somatostatin.

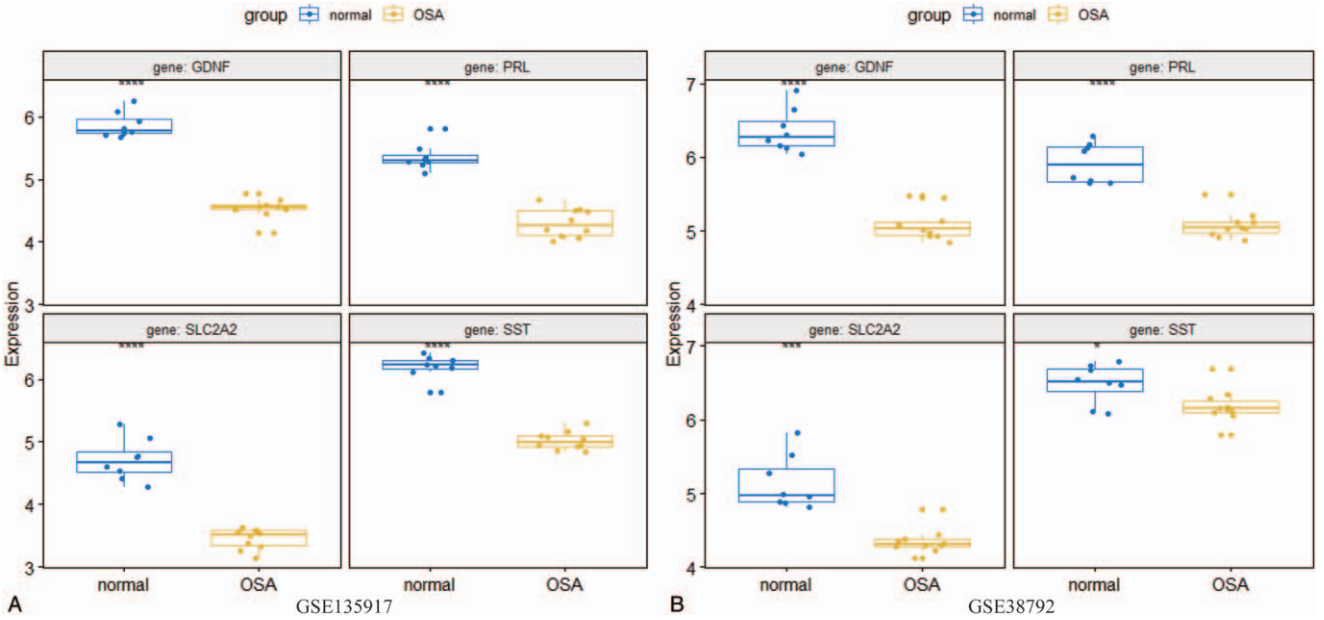

Figure 6. Continued. Validation of the hub genes. Differential expression of four hub genes in OSA samples and normal controls based on GSE135917 (A) and GSE38792 (B). These results indicate that our findings are reliable. $^{*}P < .05$, $^{***}P < .001$ and $^{****}P < .0001$.

of cellular mechanisms impinging on gene expression, regulation of intracellular metabolic pathways, and induction of hormonal and neuronal signals.[22] Single nucleotide polymorphisms (SNPs) in SLC2A2 were associated with the conversion from impaired glucose tolerance to type 2 diabetes in the Finnish Diabetes Prevention Study, and this relationship was independent of weight change.[23] Furthermore, Anders et al. found only the minor risk allele of SCL2A2 was significantly associated with increased risk of incident cardiovascular disease (CVD), when assessing the individual and cumulative effect of 46 type 2 diabetes related genetic variants, and the association was independent of diabetes status at baseline.[24] The present study combined two different microarray datasets for analysis and showed that SLC2A2 expression was low in OSA patients, and exhibited good discrimination (AUC = 0.9594). Given the increasingly recognized adverse impact of OSA on CVD, metabolic syndrome and neuropsychiatric disorders, some genetic variants may prove also to be important in determining whether pathways disrupted in the pathogenesis of OSA causally contribute to these outcomes.[7,8] It might be of great significance to investigate whether SLC2A2 variants are associate both with OSA and co-morbid disorders (e.g., diabetes, CVD), that is, SLC2A2 variants may contribute simultaneously to both phenotypes through different mechanistic effects.

Prolactin (PRL) is a luteotropic and pleiotropic hormone involved in many biological processes, such as lactation, reproduction, angiogenesis, immune response and osmoregulation.[25] Besides, PRL can influence sleep structure, and PRL-deficient mice display less rapid eye movement (REM) sleep than wild-type mice.[26] Additionally, patients with hyperprolactinemia have been observed to have metabolic alterations and are more susceptible to weight gain, and weight loss in these patients is associated with normalization of their prolactin levels.[27] PRL may modulate energy metabolism by regulating the LPL activity and lipogenesis, and reducing the release of adiponectin in human adipose tissue.[28] On the basis of a genome-wide association study (GWAS) for early onset and morbid obesity, a variant near PRL was identified relevant for common obesity and BMI variation.[29] Nilsson et al. successfully replicated the association in a large population-based study from Western Finland, finding that the variant near PRL gene is associated with increased adiposity in males.[28] Herein, our study demonstrated that PRL gene expression was significantly decreased in OSA samples compared with normal controls. As we know, obesity increases the risk of OSA by 10 to 14-fold, and is expected to explain up to 40% of AHI variation.[7] Identification of genes determining intermediate phenotypes (potentially on a causal pathway leading to OSA) may be helpful to determine susceptibility

Table 3

ROC curve analysis of hub gene expression for OSA.

Gene	Cut-off value	Specificity,%	Sensitivity,%	AUC	AUC 95%CI
GDNF	5.573	100.0	100.0	1.00	1.000–1.000
SLC2A2	4.402	93.8	90.0	0.9594	0.9019–1.000
PRL	5.216	93.8	95.0	0.9688	0.9189–1.000
SST	6.176	75.0	85.0	0.8438	0.7152–0.9723

OSA = obstructive sleep apnea, ROC = receiver operating characteristic, AUC = area under curve, CI = confidence interval, GDNF = glial cell derived neurotrophic factor, SLC2A2 = solute carrier family 2 member 2, PRL = prolactin, SST = somatostatin.

genes.[7] Thus, further research is warranted to explore whether PRL variants contribute to OSA via sleep structural alterations or intermediate phenotypes such as obesity.

Somatostatin (SST), a cyclic peptide hormone, affects growth hormone release, and gastrointestinal function. SST can influence somatic growth and body weight by regulating the gastrointestinal motility, intestinal absorption of nutrients, and energy homeostasis. It has emerged as a therapeutic approach for obesity and type 2 diabetes.[30] Moreover, as a neurotransmitter or neuromodulator abundant in the central nervous system, SST plays a role in the fine-tuning of synaptic plasticity and neuronal activity.[31] $SST^+/nNOS^+$ neuron dysfunction may contribute to pathophysiological changes observed in various diseases associated with both disturbed slow-wave activity (SWA) production and cognitive impairments including Alzheimer's disease (AD), epilepsy, schizophrenia and traumatic brain injury.[32] Evidence has suggested that the down-regulation of SST expression in the brain of early aging initiates a gradual decline in neprilysin activity, leading to amyloid-β (Aβ) peptide accumulation in patients with AD. Thus, SST variants might alter the expression or function of somatostatin and be involved in the AD process.[33] Our study demonstrated that decreased expression of SST was observed in OSA samples. As mentioned above, OSA has been strongly associated with CVD, metabolic syndrome and neuropsychiatric disorders, and some genetic variants may prove also to be important in determining whether pathways disrupted in the pathogenesis of OSA causally contribute to related consequences.[7,8] That is, SST variants may contribute simultaneously to both OSA and cognitive dysfunction through different mechanistic effects, which are required to further validate.

Finally, correlation analysis revealed a strong correlation between these 4 hub genes, which indicates that they may play a synergistic role in the occurrence and development of OSA. Previous study has shown that GDNF promotes survival and axonal regeneration in a wide variety of neuronal populations, and SST promoted neurite outgrowth in rat cerebellar granule cells.[34] Besides, Morel et al demonstrated that GDNF gene therapy is effective in ameliorating chronic hyperprolactinemia, possibly by stimulating the tuberoinfundibular dopamine neuron function.[35] However, the co-expression relationship between these hub genes has not been well explained. Although this study presented here based on rigorous bioinformatics methods, several limitations should be acknowledged. First, the findings of our study resulted from public database and bioinformatics analyses, which still need clinical data and experimental verification. Second, the data we utilized were obtained from subcutaneous adipose tissue samples, which were of poor quantity.

In conclusion, the present study has identified DEGs and key hub genes in subcutaneous adipose tissues from OSA patients compared with normal controls, which may help us understand the mechanisms behind OSA development, and provide more clues for its health consequences. Further research is needed to validate their potential role as future diagnostic, prognostic, or therapeutic biomarkers for OSA.

Author contributions

Conceptualization: Yuanyuan Cao.
Data curation: Yuanyuan Cao, Xintian Cai.
Formal analysis: Qing Zhu.
Funding acquisition: Nanfang Li.
Project administration: Qing Zhu.
Software: Yuanyuan Cao.
Supervision: Nanfang Li.
Validation: Yuanyuan Cao.
Visualization: Xintian Cai.
Writing – original draft: Yuanyuan Cao.
Writing – review & editing: Qing Zhu, Nanfang Li.

References

[1] Sánchez-de-la-Torre M, Campos-Rodriguez F, Barbé F. Obstructive sleep apnoea and cardiovascular disease. Lancet Respir Med 2013;1:61–72.
[2] Javaheri S, Barbe F, Campos-Rodriguez F, et al. Sleep apnea: types, mechanisms, and clinical cardiovascular consequences. J Am Coll Cardiol 2017;69:841–58.
[3] Li JJ, Appleton SL, Wittert GA, et al. The relationship between functional health literacy and obstructive sleep apnea and its related risk factors and comorbidities in a population cohort of men. Sleep 2014;37:571–8.
[4] Heinzer R, Vat S, Marques-Vidal P, et al. Prevalence of sleep-disordered breathing in the general population: the HypnoLaus study. Lancet Respir Med 2015;3:310–8.
[5] Simpson L, Hillman DR, Cooper MN, et al. High prevalence of undiagnosed obstructive sleep apnoea in the general population and methods for screening for representative controls. Sleep Breath 2013;17:967–73.
[6] Strohl KP, Saunders NA, Feldman NT, et al. Obstructive sleep apnea in family members. N Engl J Med 1978;299:969–73.
[7] Mukherjee S, Saxena R, Palmer LJ. The genetics of obstructive sleep apnoea. Respirology (Carlton, Vic) 2018;23:18–27.
[8] Larkin EK, Patel SR, Goodloe RJ, et al. A candidate gene study of obstructive sleep apnea in European-Americans and African-Americans. Am J Respir Crit Care Med 2010;182:947–53.
[9] Farias Tempaku P, Leite Santoro M, Bittencourt L, et al. Genome-wide association study reveals two novel risk alleles for incident obstructive sleep apnea in the EPISONO cohort. Sleep Med 2020;66:24–32.
[10] Kang S, Song J. Robust gene selection methods using weighting schemes for microarray data analysis. BMC Bioinformatics 2017;18:389.
[11] Gharib SA, Hurley AL, Rosen MJ, et al. Obstructive sleep apnea and cpap therapy alter distinct transcriptional programs in subcutaneous fat tissue. Sleep 2020;43:6.
[12] Gharib SA, Hayes AL, Rosen MJ, et al. A pathway-based analysis on the effects of obstructive sleep apnea in modulating visceral fat transcriptome. Sleep 2013;36:23–30.
[13] Gautier L, Cope L, Bolstad BM, et al. Affy–analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 2004;20:307–15.
[14] Ritchie ME, Phipson B, Wu D, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.
[15] Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS J Integr Biol 2012;16:284–7.
[16] Carlson M. org.Hs.eg.db: Genome wide annotation for Human. R package version 3.10.0,2019.
[17] Bindea G, Mlecnik B, Hackl H, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 2009;25:1091–3.
[18] Chin CH, Chen SH, Wu HH, et al. CytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol 2014;8 Suppl 4:S11.
[19] Can T. Introduction to bioinformatics. Methods Mol Biol 2014;1107: 51–71.
[20] Kao CH, Chen SH, Chio CC, et al. Exogenous administration of glial cell line-derived neurotrophic factor improves recovery after spinal cord injury. Resuscitation 2008;77:395–400.
[21] Gislason T, Pack AI, Helgadottir HT, et al. The CRP and GDNF genes do not contribute to apnea-hypopnea index or risk of obstructive sleep apnea. Am J Respir Crit Care Med 2011;184:143–5.
[22] Thorens B. GLUT2, glucose sensing and glucose homeostasis. Diabetologia 2015;58:221–32.
[23] Laukkanen O, Lindström J, Eriksson J, et al. Polymorphisms in the SLC2A2 (GLUT2) gene are associated with the conversion from impaired glucose tolerance to type 2 diabetes: the Finnish Diabetes Prevention Study. Diabetes 2005;54:2256–60.

[24] Borglykke A, Grarup N, Sparsø T, et al. Genetic variant SLC2A2 is associated with risk of cardiovascular disease-assessing the individual and cumulative effect of 46 type 2 diabetes related genetic variants. PloS One 2012;7:e50418.
[25] Bernard V, Young J, Chanson P, et al. New insights in prolactin: pathological implications. Nature reviews. Endocrinology 2015;11: 265–75.
[26] Machado RB, Rocha MR, Suchecki D. Brain prolactin is involved in stress-induced REM sleep rebound. Horm Behav 2017;89:38–47.
[27] Greenman Y, Tordjman K, Stern N. Increased body weight associated with prolactin secreting pituitary adenomas: weight loss with normalization of prolactin levels. Clin Endocrinol 1998;48:547–53.
[28] Nilsson L, Olsson AH, Isomaa B, et al. A common variant near the PRL gene is associated with increased adiposity in males. Mol Genet Metab 2011;102:78–81.
[29] Meyre D, Delplanque J, Chèvre JC, et al. Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. Nat Genet 2009;41:157–9.
[30] Tremblay M, Brisson D, Gaudet D. Association between a polymorphic poly-T repeat sequence in the promoter of the somatostatin gene and hypertension. Hypertens Res 2016;39:467–74.
[31] Iwasawa C, Kuzumaki N, Suda Y, et al. Reduced expression of somatostatin in GABAergic interneurons derived from induced pluripotent stem cells of patients with parkin mutations. Mol Brain 2019;12:5.
[32] Zielinski MR, Atochin DN, McNally JM, et al. Somatostatin+/nNOS+ neurons are involved in delta electroencephalogram activity and cortical-dependent recognition memory. Sleep 2019;42:10.
[33] Xue S, Jia L, Jia J. Association between somatostatin gene polymorphisms and sporadic Alzheimer's disease in Chinese population. Neurosci Lett 2009;465:181–3.
[34] Kishimoto Y, Yabuta C, Shearer TR, et al. FK962 promotes neurite elongation and regeneration of cultured rat trigeminal ganglion cells: possible involvement of GDNF. Invest Ophthalmol Vis Sci 2012;53:5312–9.
[35] Morel GR, Sosa YE, Bellini MJ, et al. Glial cell line-derived neurotrophic factor gene therapy ameliorates chronic hyperprolactinemia in senile rats. Neuroscience 2010;167:946–53.