Bioinformatic Analysis Identifies Biomarkers and Treatment Targets in Primary Sjögren's Syndrome Patients with Fatigue

Guangshu Chen
Guangzhou Red Cross Hospital https://orcid.org/0000-0001-6718-8667
Li Che
Jinan University First Affiliated Hospital
Xingdong Cai
Jinan University First Affiliated Hospital
Ping Zhu
Guangzhou Red Cross Hospital
Jianmin Ran (ran_jianmin@ext.jnu.edu.cn)
Guangzhou Red Cross Hospital, Jinan University https://orcid.org/0000-0003-1364-9469
Shengming Liu
Jinan University First Affiliated Hospital

Research

Keywords: Primary Sjögren's syndrome (pSS), Fatigue, Differentially expressed genes(DEGs), Protein-protein interactions (PPIs), Gene ontology(GO), Kyoto Encyclopedia of Genes and Genomes(KEGG), Treatment targets, Chlorophyllin

Posted Date: August 30th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-829258/v1

Version of Record: A version of this preprint was published at BioMed Research International on January 15th, 2022. See the published version at https://doi.org/10.1155/2022/7697558.

Abstract

We aim to identify the common genes, biological pathways and treatment targets for primary Sjögren's syndrome patients with varying degrees of fatigue features. We select datasets about transcriptomic analyses of Primary Sjögren's syndrome(pSS) patients with different degrees of fatigue features and normal controls in peripheral blood. We identify common differentially expressed genes(DEGs) to find shared pathways and treatment targets for pSS patients with fatigue and design protein-protein interactions (PPIs) network by some practical bioinformatics tools. And hub genes are detected based on the PPIs network. We perform biological pathways analysis of common genes by gene ontology(GO) terms and Kyoto Encyclopedia of Genes and Genomes(KEGG) pathway. Lastly, potential treatment targets for pSS patients with fatigue are found by the Enrichr platform. We discovered 27 DEGs are identified in pSS patients with fatigue features and the severe fatigued pSS-specific gene being RTP4. DEGs are mainly localized in the mitochondria, endosomes, endoplasmic reticulum, cytoplasm and are involved in the biological process by which interferon acts on cells and cells defend themselves against viruses. Molecular functions mainly involve the process of RNA synthesis. The DEGs of pSS are involved in the signaling pathways of viruses such as hepatitis C, influenza A, measles, and EBV. Acetohexamide PC3 UP, suloctidil HL60 UP, prenylamine HL60 UP, chlorophyllin CTD 00000324 are the four most polygenic drug molecules. PSS patients with fatigue features have specific gene regulation and chlorophyllin may alleviate fatigue symptoms in pSS patients.

Introduction

Primary Sjögren's syndrome (pSS) is an all-body autoimmune disease that mainly affects middle-aged women [1]. The main clinical features of the disease are dryness of mouth and eyes, and the pathophysiology is characterized by focal lymphocytes infiltration in exocrine glands [2, 3]. Fatigue is commonly seen in pSS patients as an extra-glandular manifestation and closely links with poor life quality [4–6]. Fatigue affects approximately 70% of pSS patients [7, 8]. Normally, fatigue and depression are considered manifestations of psychological disorders and interact with physical pain and discomfort, which create a vicious cycle. Fatigue in pSS is induced and regulated by genetic and molecular mechanisms, with the innate immune system playing an important role in the production of fatigue [9–11]. Although pSS always comes with fatigue, not all patients exhibit fatigue, which provides a good model for exploring the underlying biological mechanisms.

High-throughput methods play an increasingly essential role in biology spheres and microarray data analysis highlights its advantage in large-scale analysis of gene expression among high-throughput applications [12, 13]. Former studies [14, 15]have shown the high-throughput sequencing analysis result for pSS patients with fatigue features, but don’t offer further analysis based on varying degrees of fatigue. This study tries to present characteristic genes and biological pathways in pSS patients with manifestations of fatigue, as well as drugs of potential benefit.

The GSE66795 dataset from the GPL10558 platform on the GEO database is selected for gene expression of pSS with fatigue. The GSE66795 dataset was first identified for differentially expressed genes (DEGs) in

pSS patients with different levels of fatigue and based on the co-expressed genes, further analysis including gene ontology(GO) terms and Kyoto Encyclopedia of Genes and Genomes(KEGG) pathway are performed to understand the biological process. The top ten target genes from the protein-protein interaction (PPI) network will be obtained to identify potential drugs that may alleviate fatigue in pSS patients.

Materials And Methods

Dataset collection

We search “Primary Sjögren's syndrome” and “fatigue” in the GEO database [16] and select the dataset (GSE66795) demonstrating gene expression in pSS patients with varying degrees of fatigue characteristics and normal controls. The GSE66795 dataset is extracted from the GPL10558 platform (Illumina HumanHT-12 V4.0 expression microbead chip) for RNA sequence analysis. The data of GSE66795 is obtained from the UK registry of primary Sjögren's syndrome. It includes whole genome microarray profiles of pSS patients with varying degrees of fatigue characteristics and normal controls in peripheral blood. One hundred and thirty-one patients with pSS are involved, including 21 patients with mild fatigue, 74 patients with moderate fatigue, 36 patients with severe fatigue, and 29 normal controls.

Differential expression analysis

Differential expression analysis is performed using the online analysis tool GEO2R; gene expression profiles of pSS patients with mild, moderate, and severe fatigue were compared with normal controls separately to identify DEGs. P-values and adjusted p-values are calculated using t-tests. Genes with the following criteria were retained for each sample: (1) log2-fold change(log2FC) absolute value greater than 1; (2) adjusted P value less than 0.05. After identifying DEGs in pSS patients with varying degrees of fatigue, the online website (https://www.xiantao.love/gds) is used to plot a Venn diagram.

Gene ontology and pathway discovery in gene set enrichment analysis

Gene set enrichment analysis is used to understand the general biological function and the chromosomal location of a gene [17]. For gene product annotation, the terms of gene ontology (GO) are used, including biological process (BP), molecular function (MF) and cellular component (CC) [18]. The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways are commonly used to describe metabolic pathways [19]. GO terms and KEGG pathways were got through the platform Enrichr(https://amp.

pharm. mssm. edu/Enrichr/) based on the DEGs [20].

Protein-protein interaction (PPI) network

The information generated from the PPI network improves the understanding of protein function[21]. PPI networks are made by STRING(https://string-db.org/) after inputting the common DEGs. We analyze PPIs through Cytoscape(https://cytoscape.org/) to further present the network and identify target genes.

Transcription factor(TF)-gene interactions

We use the Network Analyst (https://www.networkanalyst.ca/) to identify interactions of TF-genes with DEGs [22]. Network Analyst plays a comprehensive network platform for gene expression across a wide range of species and enables them to be subjected to a meta-analysis [23].

Identification of potential treatment targets

Identification of drug molecules is a vital component of genomics research. We input the DEGs in the Drug Signature Database (DSigDB). Then we get the designed drug molecules, which may have promising clinical applicants. DSigDBis obtained through the Enrichr(https://amp.pharm.mssm.edu/Enrichr/) platform. Enrichr is primarily used as an enrichment analysis platform, providing extensive visual details of the common functions of inputted genes [24].

Results

DEGs identifications

We use the GSE66795 dataset to identify the DEGs of pSS with fatigue. 37, 29, and 33 DEGs are obtained for pSS with mild, moderate and severe fatigue, respectively. The collected DEGs are further compared by the online website (https://www.xiantao.

love/gds) for gathering common genes in pSS with varying degrees of fatigue. And 27(OAS1, OAS2, GBP1, IRF7, EIF2AK2, IFIT2, USP18, SAMD9L, HES4, IFI44L,

SERPING1, IFIT3, IFITM3, IFI6, XAF1, MX1, OASL, OTOF, HERC5, LY6E, EPSTI1, OAS3, ISG15, IFIT1, RSAD2, IFI44, IFI27) common DEGs are identified, which is shown in Fig. 2. The specific genes to pSS with mild fatigue are DDX60, IFIH1, GBP5, LAP3, TIMM10, moderate fatigue are HLA-DRB4, HLA-DRB6 and that severe fatigue is RTP4. The Venn diagram(Fig. 3) shows that common DEGs accounted for 67.5% out of a total of 40 DEGs.

GO terms and KEGG pathways

We analyzed 27 common DEGs for both GO and KEGG pathways. Both of the results are taken from the top 10 GO entries. GO terms in Table 1 suggest that DEGs are mainly localized in the mitochondria, endosomes, endoplasmic reticulum and cytoplasm.They are involved in the biological processes of interferon action on cells and cellular defense against viruses. And the molecular functions are mainly engaged in the process of RNA synthesis. KEGG pathways in Table 2 suggest that the DEGs of pSS with fatigue are involved in the signaling pathways of viruses such as hepatitis C, influenza A, measles and EBV. Both are seen in Fig. 3 (A and B).

Table 1
Top 10 GO pathways and their corresponding P-values and genes

GO	GO_ID	Description	P-value	Genes
MF	GO:0001730	2'-5'-oligoadenylate synthetase activity	1.305e-12	OAS2 OAS3 OASL OAS1
	GO:0003725	double-stranded RNA binding	1.225e-09	OAS1 OAS2 EIF2AK2 OAS3 OASL
	GO:0016779	nucleotidyltransferase activity	2.993e-06	OAS1 OAS3 OAS2 OASL
	GO:0003723	RNA binding	2.089e-04	EIF2AK2 OAS1 OAS3 HERC5 IFIT2 IFIT3 IFIT1 OASL OAS2
	GO:0035639	purine ribonucleoside triphosphate binding	8.781e-03	EIF2AK2 MX1 OAS1 OAS3 GBP1 OASL OAS2
	GO:0001883	purine nucleoside binding	9.156e-03	EIF2AK2 MX1 OAS1 OAS3 GBP1 OASL OAS2
	GO:0032549	ribonucleoside binding	9.156e-03	EIF2AK2 MX1 OAS1 OAS3 GBP1 OASL OAS2
	GO:0001882	nucleoside binding	9.363e-03	EIF2AK2 MX1 OAS1 OAS3 GBP1 OASL OAS2
	GO:0017076	purine nucleotide binding	1.032e-02	EIF2AK2 MX1 OAS1 OAS3 GBP1 OASL OAS2
	GO:0044822	poly(A) RNA binding	5.088e-02	EIF2AK2 HERC5 IFIT2 OASL
CC	GO:0048471	perinuclear region of cytoplasm	3.888e-03	EIF2AK2 MX1 HERC5 OAS2
	GO:0005829	cytosol	9.919e-06	EIF2AK2 OAS1 OAS3 IRF7 HERC5 ISG15 OAS2 MX1 USP18 IFIT2 OTOF IFIT3 GBP1 IFIT1 OASL XAF1
	GO:0005739	mitochondrion	4.810e-03	OAS1 RSAD2 IFI27 IFIT3 IFI6 XAF1 OAS2
	GO:0005783	endoplasmic reticulum	1.583e-02	MX1 OAS1 IFIT2 RSAD2 OTOF OAS2
	GO:0012505	endomembrane system	4.168e-03	SAMD9L OAS1 SERPING1 IRF7 RSAD2 IFITM3 OAS2 MX1 IFIT2 OTOF IFI27 GBP1
	GO:0005737	cytoplasm	1.488e-02	SERPING1 RSAD2 ISG15 USP18 OTOF IFI27 IFIT3 OASL SAMD9L OAS1 OAS3 HERC5 IFI44 IFI6 IRF7 OAS2 GBP1 IFIT1 EIF2AK2 IFITM3 MX1 IFIT2 IFI44L XAF1
	GO:0071357	cellular response to type I interferon	2.702e-30	IFIT1 OAS3 MX1 USP18 IFIT2 IRF7 OAS1 OAS2 IFITM3 OASL IFIT3 XAF1 ISG15 RSAD2 IFI27 IFI6

GO	GO_ID	Description	P-value	Genes
	GO:0060337	type I interferon signaling pathway	2.702e-30	IFIT1 OAS3 MX1 USP18 IFIT2 IRF7 OAS1 OAS2 IFITM3 OASL IFIT3 XAF1 ISG15 RSAD2 IFI27 IFI6
	GO:0045071	negative regulation of viral genome replication	2.204e-18	OAS1 EIF2AK2 IFIT1 OAS3 MX1 IFITM3 OASL ISG15 RSAD2
	GO:0034340	response to type I interferon	3.905e-30	IFIT1 OAS3 MX1 USP18 IFIT2 IRF7 OAS1 OAS2 IFITM3 OASL IFIT3 XAF1 ISG15 RSAD2 IFI27 IFI6
	GO:0045069	regulation of viral genome replication	9.301e-17	EIF2AK2 IFIT1 OAS3 MX1 OAS1 IFITM3 OASL ISG15 RSAD2
	GO:0048525	negative regulation of viral process	4.916e-16	EIF2AK2 IFIT1 OAS3 MX1 OAS1 IFITM3 OASL ISG15 RSAD2
	GO:0019079	viral genome replication	9.276e-16	EIF2AK2 IFIT1 OAS3 MX1 OAS1 IFITM3 OASL ISG15 RSAD2
	GO:0060333	interferon-gamma-mediated signaling pathway	1.854e-10	OAS3 GBP1 IRF7 OAS1 OAS2 OASL
	GO:0051607	defense response to virus	6.646e-09	HLA-DPA1 GBP2 IFNGR1 VCAM1 HLA-DRA CCL22 ICAM1

Table 2
Top 10 Top KEGG pathways and their corresponding P-values and genes

pathwayID	Description	P-value	Genes
hsa05160	Hepatitis C	2.02329E-11	IFIT1 IRF7 MX1 OAS1 OAS2 OAS3 EIF2AK2 RSAD2
hsa05164	Influenza A	2.60307E-09	IRF7 MX1 OAS1 OAS2 OAS3 EIF2AK2 RSAD2
hsa05162	Measles	3.6267E-08	IRF7 MX1 OAS1 OAS2 OAS3 EIF2AK2
hsa05169	Epstein-Barr virus infection	3.37146E-07	IRF7 OAS1 OAS2 OAS3 EIF2AK2 ISG15
hsa04621	NOD-like receptor signaling pathway	5.95267E-06	GBP1 IRF7 OAS1 OAS2 OAS3
hsa05165	Human papillomavirus infection	0.000110032	MX1 EIF2AK2 OASL ISG15 HES4
hsa05168	Herpes simplex virus 1 infection	0.000742976	IRF7 OAS1 OAS2 OAS3 EIF2AK2
hsa04622	RIG-I-like receptor signaling pathway	0.005431097	IRF7 ISG15
hsa05167	Kaposi sarcoma-associated herpesvirus infection	0.037274554	IRF7 EIF2AK2
hsa05203	Viral carcinogenesis	0.041244055	IRF7 EIF2AK2

Identification of hub genes by PPI networks

We put common DEGs into the STRING website and the files generated after analysis are further entered into Cytoscape software for visual analysis. PPIs networks are designed to detect hub genes for identifying drug molecules for pSS with fatigue. PPIs networks involve 24 nodes and 552 edges, which is showed in Fig. 4A. We present the top 20 genes in Fig. 4B and Table 3.

Table 3
Topological result exploration for top 20 genes

Hub Gene	Degree	Stress	Closeness	Betweenness	EcCentricity	ClusteringCoefficient
MX1	23	20	23	1.09235	1	0.96047
IFIT1	23	20	23	1.09235	1	0.96047
ISG15	23	20	23	1.09235	1	0.96047
RSAD2	23	20	23	1.09235	1	0.96047
IFI44L	23	20	23	1.09235	1	0.96047
IFI44	23	20	23	1.09235	1	0.96047
IFIT3	23	20	23	1.09235	1	0.96047
OAS2	23	20	23	1.09235	1	0.96047
OAS1	23	20	23	1.09235	1	0.96047
IFI6	23	20	23	1.09235	1	0.96047
XAF1	23	20	23	1.09235	1	0.96047
IFIT2	23	20	23	1.09235	1	0.96047
GBP1	23	20	23	1.09235	1	0.96047
OAS3	23	20	23	1.09235	1	0.96047
HERC5	22	14	22.5	0.75624	0.5	0.9697
IRF7	22	10	22.5	0.51637	0.5	0.97835
OASL	22	10	22.5	0.51637	0.5	0.97835
IFI27	22	14	22.5	0.77598	0.5	0.9697
EIF2AK2	22	10	22.5	0.51637	0.5	0.97835
LY6E	21	10	22	0.55833	0.5	0.97619

TF-gene interactions

The interactions of TF and genes are shown in Fig. 7. The network has 60 nodes and 108 edges. Sixteen TF-genes regulate IFIT1and IFIT3 is handled by 14 TF-genes. The network involves 60 TF-genes. Figure 5 shows the network of TF-gene interactions.

Identification of drug candidates

We identify drug molecules for the top 10 hub genes on the Enrichrplatform. We collect drug candidates judged on adjusted P-values. The analysis reveals acetohexamide PC3 UP, suloctidil HL60 UP, prenylamine

HL 60 UP and chlorophyllin CTD 00000324 are the four most drug molecules that interact with genes. Figure 5 and Table 4 present the drug candidates in the DSigDB database.

Table 4
Candidate drug for pSS with fatigue

Name of drugs	P-value	Adjusted P-value	Target Genes
acetohexamide PC3 UP	3.47E−23	8.16E−21	RSAD2;OAS1;OAS2;MX1;IFI6;IFI44;IFIT1;IFI44L
suloctidil HL60 UP	2.19E−22	2.57E−20	RSAD2;OAS1;OAS2;MX1;IFI6;IFI44;ISG15;IFIT1;IFI44L;IFIT3
prenylamine HL60 UP	3.46E−21	2.71E−19	OAS1;MX1;IFI6;IFI44;ISG15;IFIT1;IFI44L;IFIT3
chlorophyllin CTD 00000324	9.59E−19	5.64E−17	OAS1 OAS2 MX1 IFI6 ISG15 IFIT1 IFIT3
terfenadine HL60 UP	1.81E−16	8.52E−15	OAS1 MX1 IFI6 IFI44 ISG15 IFIT1 IFIT3
3'-Azido−3'-deoxythymidine CTD 00007047	6.05E−13	2.37E−11	RSAD2 OAS1 OAS2 MX1 IFI6 ISG15 IFIT1 IFI44L
clioquinol PC3 UP	1.79E−11	6.00E−10	OAS1 MX1 IFI44 IFIT1 IFIT3
etoposide HL60 UP	2.92E−11	8.56E−10	OAS1 MX1 IFI6 IFI44 ISG15 IFIT1
testosterone enanthate CTD 00000155	1.03E−09	2.68E−08	RSAD2 OAS2 MX1 IFI6 ISG15 IFI44L IFIT3
Gadodiamide hydrate CTD 00002623	3.85E−09	9.05E−08	RSAD2 IFIT1 IFI44L IFIT3

Discussion

Fatigue is an annoying experience that means physical and mental tiredness [25]. Mengshoelet al. [26] reveal that most pSS patients literally suffer fluctuating fatigue out of control regardless of their health condition. Fatigue has a significant influence on patients' daily life and patients must adapt their behavior and lives. Although the underlying mechanisms are still unclear, former studies take depression and pain as the prominent factors associated with fatigue [5, 27]. Currently, growing evidence suggests fatigue has a molecular and genetic basis on its production and regulation. Therefore, most scholars view fatigue as a biological and brain phenomenon [9–11].

IL-1β tends to increase rapidly secreted from macrophages to activate the immune system when meeting tissue injury or infection. IL-1β plays its role by binding with the IL-1 receptor coming with the downstream of IL-1 response [28]. Then immune and inflammation systems are activated, which induce the behavior of disease, fatigue being involved as an important component [29]. All these inflammatory signaling pathways go on working and turn fatigue into a chronic state. In the brain, IL-1 β signaling pathways may explain the ultimate pathway of fatigue [30, 31], and IL-1 blockers treatment may effectively release fatigue [32, 33]. Thus, fatigue and other unpressant mood in those patients with autoimmune disease should be not only understood by the unfortunate development of chronic illness, but also may be related to some signalling pathways and activation of genes that regulate the mood in cerebral system.

Genome-wide association analysis of pSS patients has been conducted and a gene (RTP4) is identified as highly relevant. Similarly, we confirm that RTP4 is highly expressed in pSS patients with severe fatigue through bioinformatic analysis, suggesting that this gene is critical in the mechanisms of fatigue. RTP4 encodes a protein associated with the expression of opioid receptors on the cell surface. These receptors are also expressed in the lymph system and pain-regulated pathways in the brain [34]. However, the former study did not stratify pSS based on the degree of fatigue and it is unclear which degree of fatigue expresses the RTP4 gene. Our study finds that pSS patients with severe fatigue specifically express the RTP4 gene, providing clues for further studies on the genomics of fatigue features in pSS patients.

OAS1, a co-expressed gene for pSS in our study, has been established in previous studies as a risk locus of pSS and impacts the flaw of virus clearance because of the altering response of IFN [35]. Our gene pathway analysis points out that DEGs for pSS with fatigue are mainly localized intracellularly and involve in signaling pathways of common viruses in the respiratory and digestive tracts, suggesting that pSS is a systemic disease with an uncertain etiology and that viral infection may be a predisposing factor.

Fatigue always accompanies pSS patients, but it is hard work to manage these bad feelings [36]. The clinical practice guidelines(CPG) committee emphasizes the many causes of fatigue in pSS, therefore the comprehensive evaluation for diagnosis is essential. So far, the treatment for fatigue in pSS with solid recommendation is mere taking exercise, which is also practical in other autoimmune diseases [37]. In America, hydroxychloroquine(HCQ) is the most widely used drug therapy for pSS with fatigue, but the recommendation strength is not strong enough [34]. It is not recommended to release fatigue in pSS using dehydroepiandrosterone (DHEA) [34]. Both the tumor necrosis factor inhibitor is discouraged for the treatment of fatigue in pSS [38, 39]. Our bioinformatic study reveals that besides chloroquine and testosterone drugs that help improve fatigue, chlorophyllin, the sulphonylurea hypoglycaemic drug acetylhexane, and the anti-allergic drug terfenadine may have improved fatigue in pSS. However, chloroquine and testosterone are not strongly recommended we mention before. Acethexamide has been discontinued in the American market due to its significant hypoglycaemic risk. Terfenadine is not suitable for long-term use since its central depression as an anti-allergy drug. And chlorophyllin appears to hold some promise for reducing fatigue in pSS.

Chlorophyll is an ingredient of derifil drug which is available in the over-the-counter [40]. And chlorophyllin, obtained by hydrolyzing chlorophyll to remove phytyl alcohol, is a water-soluble derivative. Chlorophyll has

been shown to exert its anticancer properties by playing a role as an antioxidant [41], a CYP inhibitor [42], an apoptosis inducer [43], a phase II enzymes stimulator [44] and a carcinogen transport modulator [45]. Currently, COVID-19 has swept the world and may last for a long time because of its rapid mutation. Almost 4,200,000 people have died in this epidemic [46], and the reduction of lymphocytes in COVID-19 patients is considered as an important risk factor for poor prognosis [47–49]. Recent studies suggest that the chlorophyll derivative sodium copper chlorophyllin (SCC) may improve survival in critically ill COVID-19 patients by increasing the total number of lymphocytes [50]. Increasing consumers choose dietary chlorophyll which is derived from SCC for diet supplements for the sake of keeping healthy [51, 52]. Dietary chlorophyll is safe and has been shown to have a higher absorption rate in human body, which may trigger ionic compounds chelation [53, 54]. Zeng et al. [55] cognize one functional food called barley grass powder which is rich in chlorophyll and other nutrients can effectively alleviate fatigue in chronic patients. The mechanism of chlorophyll's role in relieving fatigue in pSS patients is unclear. It may be related to the nature of the hepatic enzyme inhibitors that increase the concentrations of immuno-suppressant like hydroxychloroquine, which has better control of fatigue. And the capacity of scavenging the oxygen radical as an antioxidant that may somewhat improve the fatigue of body.

We have identified gene expression profiles in peripheral blood specific to pSS with fatigue characteristics. The analysis of identified DEGs and pathways in this study will deepen our understanding of the essence of fatigue in pSS. The discovery that chlorophyllin may improve fatigue symptoms provides a theoretical basis for better improving the quality of life in pSS patients.

Declarations

Ethics approval and consent to participate

GEO belongs to public databases. The patients we choose involved in the database have obtained ethical approval. It is available for all users to download relevant data for free. Our study is based on open-source data, so there is no need to offer ethics approval and consent to participate.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

Guangdong Science and Technology Project Fund for Key Scientific Research Base under Grant(no. 2019B020230001).

Authors' contributions

Guangshu Chen: Conceptualization, Methodology, Formal analysis,Writing-Original draft. Li Che: Data curation, Software, Visualization. Xingdong Cai and Ping Zhu: Visualization, Investigation. Jianmin Ran and Shengming Liu: Supervision, Validation, Funding acquisition, Writing-Reviewing and Editing.

Acknowledgements

We sincerely acknowledge the GEO database for providing their platforms and contributors for uploading their meaningful datasets.

Availability of data and material

The dataset supporting the conclusions of this article is available in the UK registry of primary Sjögren's syndrome repository and hyperlink to dataset in https://www.ncbi.nlm.nih.gov/geo/geo2r/?acc=GSE66795.

References

1. Fox RI, Howell FV, Bone RC, et al., Primary Sjogren syndrome: clinical and immunopathologic features. *Seminars in arthritis and rheumatism* 1984, *14* (2), 77–105.
2. Gabriel SE, Michaud K. Epidemiological studies in incidence, prevalence, mortality, and comorbidity of the rheumatic diseases. Arthritis research therapy. 2009;11(3):229.
3. Asmussen K, Andersen V, Bendixen G, et al., A new model for classification of disease manifestations in primary Sjögren's syndrome: evaluation in a retrospective long-term study. *Journal of internal medicine* 1996, *239* (6), 475–82.
4. Haldorsen K, Bjelland I, Bolstad AI, et al. A five-year prospective study of fatigue in primary Sjögren's syndrome. Arthritis research therapy. 2011;13(5):R167.
5. Karageorgas T, Fragioudaki S, Nezos A, et al. Fatigue in Primary Sjögren's Syndrome: Clinical, Laboratory, Psychometric, and Biologic Associations. Arthritis Care Res. 2016;68(1):123–31.
6. Overman CL, Kool MB, Da Silva JA, et al., The prevalence of severe fatigue in rheumatic diseases: an international study. *Clinical rheumatology* 2016, *35* (2), 409–15.
7. Lerdal A, Wahl A, Rustøen T, et al. Fatigue in the general population: a translation and test of the psychometric properties of the Norwegian version of the fatigue severity scale. Scandinavian journal of public health. 2005;33(2):123–30.

8. Howard Tripp N, Tarn J, Natasari A, et al., Fatigue in primary Sjögren's syndrome is associated with lower levels of proinflammatory cytokines. *RMD open* 2016, *2* (2), e000282.
9. Brække Norheim K, Imgenberg-Kreuz J, Jonsdottir K, et al. Epigenome-wide DNA methylation patterns associated with fatigue in primary Sjögren's syndrome. Rheumatology. 2016;55(6):1074–82.
10. Bårdsen K, Nilsen MM, Kvaløy JT, et al. Heat shock proteins and chronic fatigue in primary Sjögren's syndrome. Innate immunity. 2016;22(3):162–7.
11. Dantzer R, Heijnen CJ, Kavelaars A, et al., The neuroimmune basis of fatigue. *Trends in neurosciences* 2014, *37* (1), 39–46.
12. Sturn A, Quackenbush J, Trajanoski Z. Genesis: cluster analysis of microarray data. Bioinformatics. 2002;18(1):207–8.
13. Lee ML, Kuo FC, Whitmore GA, et al. Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations. Proc Natl Acad Sci U S A. 2000;97(18):9834–9.
14. Zhang L, Xu P, Wang X, et al., Identification of differentially expressed genes in primary Sjögren's syndrome. *Journal of cellular biochemistry* 2019, *120* (10), 17368–17377.
15. Song GG, Kim JH, Seo YH, et al., Meta-analysis of differentially expressed genes in primary Sjogren's syndrome by using microarray. *Human immunology* 2014, *75* (1), 98–104.
16. Clough E, Barrett T. The Gene Expression Omnibus Database. Methods Mol Biol. 2016;1418:93–110.
17. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545–50.
18. Doms A, Schroeder M, GoPubMed: exploring PubMed with the Gene Ontology. *Nucleic Acids Res* 2005, *33* (Web Server issue), W783-6.
19. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27–30.
20. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 2016;44(W1):W90-7.
21. Ben-Hur A, Noble WS. Kernel methods for predicting protein-protein interactions. Bioinformatics. 2005;21(Suppl 1):i38–46.
22. Ye Z, Wang F, Yan F, et al., Bioinformatic identification of candidate biomarkers and related transcription factors in nasopharyngeal carcinoma. *World journal of surgical oncology* **2019**, *17* (1), 60.
23. Zhou G, Soufan O, Ewald J, et al. NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. Nucleic acids research. 2019;47:W234–41.
24. Chen EY, Tan CM, Kou Y, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinform. 2013;14:128.
25. Krupp LB, Pollina DA. Mechanisms and management of fatigue in progressive neurological disorders. Curr Opin Neurol. 1996;9(6):456–60.

26. Mengshoel AM, Norheim KB, Omdal R. Primary Sjögren's syndrome: fatigue is an ever-present, fluctuating, and uncontrollable lack of energy. Arthritis Care Res. 2014;66(8):1227–32.
27. Pollard LC, Choy EH, Gonzalez J, et al. Fatigue in rheumatoid arthritis reflects pain, not disease activity. Rheumatology. 2006;45(7):885–9.
28. Dinarello CA, Simon A, van der Meer JW. Treating inflammation by blocking interleukin-1 in a broad spectrum of diseases. Nat Rev Drug Discov. 2012;11(8):633–52.
29. Hart BL. Biological basis of the behavior of sick animals. Neurosci Biobehav Rev. 1988;12(2):123–37.
30. Rossi S, Studer V, Motta C, et al., Neuroinflammation drives anxiety and depression in relapsing-remitting multiple sclerosis. *Neurology* **2017**, *89* (13), 1338–1347.
31. Kim JM, Kang HJ, Kim JW, et al., Associations of Tumor Necrosis Factor-α and Interleukin-1β Levels and Polymorphisms with Post-Stroke Depression. *The American journal of geriatric psychiatry: official journal of the American Association for Geriatric Psychiatry* 2017, *25* (12), 1300–1308.
32. Omdal R, Gunnarsson R. The effect of interleukin-1 blockade on fatigue in rheumatoid arthritis–a pilot study. Rheumatol Int. 2005;25(6):481–4.
33. Cavelti-Weder C, Furrer R, Keller C, et al. Inhibition of IL-1beta improves fatigue in type 2 diabetes. Diabetes Care. 2011;34(10):e158.
34. Carsons SE, Vivino FB, Parke A, et al., Treatment Guidelines for Rheumatologic Manifestations of Sjögren's Syndrome: Use of Biologic Agents, Management of Fatigue, and Inflammatory Musculoskeletal Pain. *Arthritis care & research* 2017, *69* (4), 517–527.
35. Li H, Reksten TR, Ice JA, et al., Identification of a Sjögren's syndrome susceptibility locus at OAS1 that influences isoform switching, protein expression, and responsiveness to type I interferons. *PLoS genetics* 2017, *13* (6), e1006820.
36. Segal B. Fatigue in primary Sjogren's syndrome. In: Ramos-Casals M, editor. Sjogren's syndrome: diagnosis and therapeutics. London: Springer Verlag; 2012. pp. 129–43.
37. Strömbeck BE, Theander E, Jacobsson LT. Effects of exercise on aerobic capacity and fatigue in women with primary Sjogren's syndrome. Rheumatology. 2007;46(5):868–71.
38. Sankar V, Brennan MT, Kok MR, et al., Etanercept in Sjögren's syndrome: a twelve-week randomized, double-blind, placebo-controlled pilot clinical trial. *Arthritis and rheumatism* 2004, *50* (7), 2240–5.
39. Mariette X, Ravaud P, Steinfeld S, et al. Inefficacy of infliximab in primary Sjögren's syndrome: results of the randomized, controlled Trial of Remicade in Primary Sjögren's Syndrome (TRIPSS). Arthritis rheumatism. 2004;50(4):1270–6.
40. Suryavanshi S, Sharma D, Checker R, et al. Amelioration of radiation-induced hematopoietic syndrome by an antioxidant chlorophyllin through increased stem cell activity and modulation of hematopoiesis. Free Radic Biol Med. 2015;85:56–70.
41. Boloor KK, Kamat JP, Devasagayam TP. Chlorophyllin as a protector of mitochondrial membranes against gamma-radiation and photosensitization. *Toxicology* **2000**, *155* (1–3), 63–71.
42. Yun CH, Jeong HG, Jhoun JW, et al., Non-specific inhibition of cytochrome P450 activities by chlorophyllin in human and rat liver microsomes. *Carcinogenesis* **1995**, *16* (6), 1437–40.

43. Chiu LC, Kong CK, Ooi VE. The chlorophyllin-induced cell cycle arrest and apoptosis in human breast cancer MCF-7 cells is associated with ERK deactivation and Cyclin D1 depletion. Int J Mol Med. 2005;16(4):735–40.
44. Fahey JW, Stephenson KK, Dinkova-Kostova AT, et al. Chlorophyll, chlorophyllin and related tetrapyrroles are significant inducers of mammalian phase 2 cytoprotective genes. Carcinogenesis. 2005;26(7):1247–55.
45. Mata JE, Yu Z, Gray JE, et al., Effects of chlorophyllin on transport of dibenzo(a, l)pyrene, 2-amino-1-methyl-6-phenylimidazo-[4,5-b]pyridine, and aflatoxin B(1) across Caco-2 cell monolayers. *Toxicology* **2004**, *196* (1–2), 117 – 25.
46. Coronavirus Resource Center, Johns Hopkins University. COVID-19 Dashboard. Available online at: https://coronavirus.jhu.edu/map.html (accessed July 31, 2021).
47. Terpos E, Ntanasis-Stathopoulos I, Elalamy I, et al., Hematological findings and complications of COVID-19. *American journal of hematology* 2020, *95* (7), 834–847.
48. Zhao Q, Meng M, Kumar R, et al. Lymphopenia is associated with severe coronavirus disease 2019 (COVID-19) infections: A systemic review and meta-analysis. International journal of infectious diseases: IJID : official publication of the International Society for Infectious Diseases. 2020;96:131–5.
49. Zheng M, Gao Y, Wang G, et al., Functional exhaustion of antiviral lymphocytes in COVID-19 patients. *Cellular & molecular immunology* 2020, *17* (5), 533–535.
50. Clark NF, Taylor-Robinson AW. COVID-19 Therapy: Could a Copper Derivative of Chlorophyll a Be Used to Treat Lymphopenia Associated With Severe Symptoms of SARS-CoV-2 Infection? Frontiers in medicine. 2021;8:620175.
51. Ulbricht C, Bramwell R, Catapang M, et al., An evidence-based systematic review of chlorophyll by the Natural Standard Research Collaboration. *Journal of dietary supplements* 2014, *11* (2), 198–239.
52. Mysliwa-Kurdziel B, Solymosi K. Phycobilins and Phycobiliproteins Used in Food Industry and Medicine. *Mini reviews in medicinal chemistry* 2017, *17* (13), 1173–1193.
53. Ferruzzi MG, Failla ML, Schwartz SJ. Sodium copper chlorophyllin: in vitro digestive stability and accumulation by Caco-2 human intestinal cells. *Journal of agricultural and food chemistry* 2002, *50* (7), 2173–9.
54. Bhatia S, Prabhu PN, Benefiel AC, et al., Galacto-oligosaccharides may directly enhance intestinal barrier function through the modulation of goblet cells. *Molecular nutrition & food research* 2015, *59* (3), 566–73.
55. Zeng Y, Pu X, Yang J, et al. Preventive and Therapeutic Role of Functional Ingredients of Barley Grass for Chronic Diseases in Human Beings. Oxidative Med Cell Longev. 2018;2018:3232080.

Figures

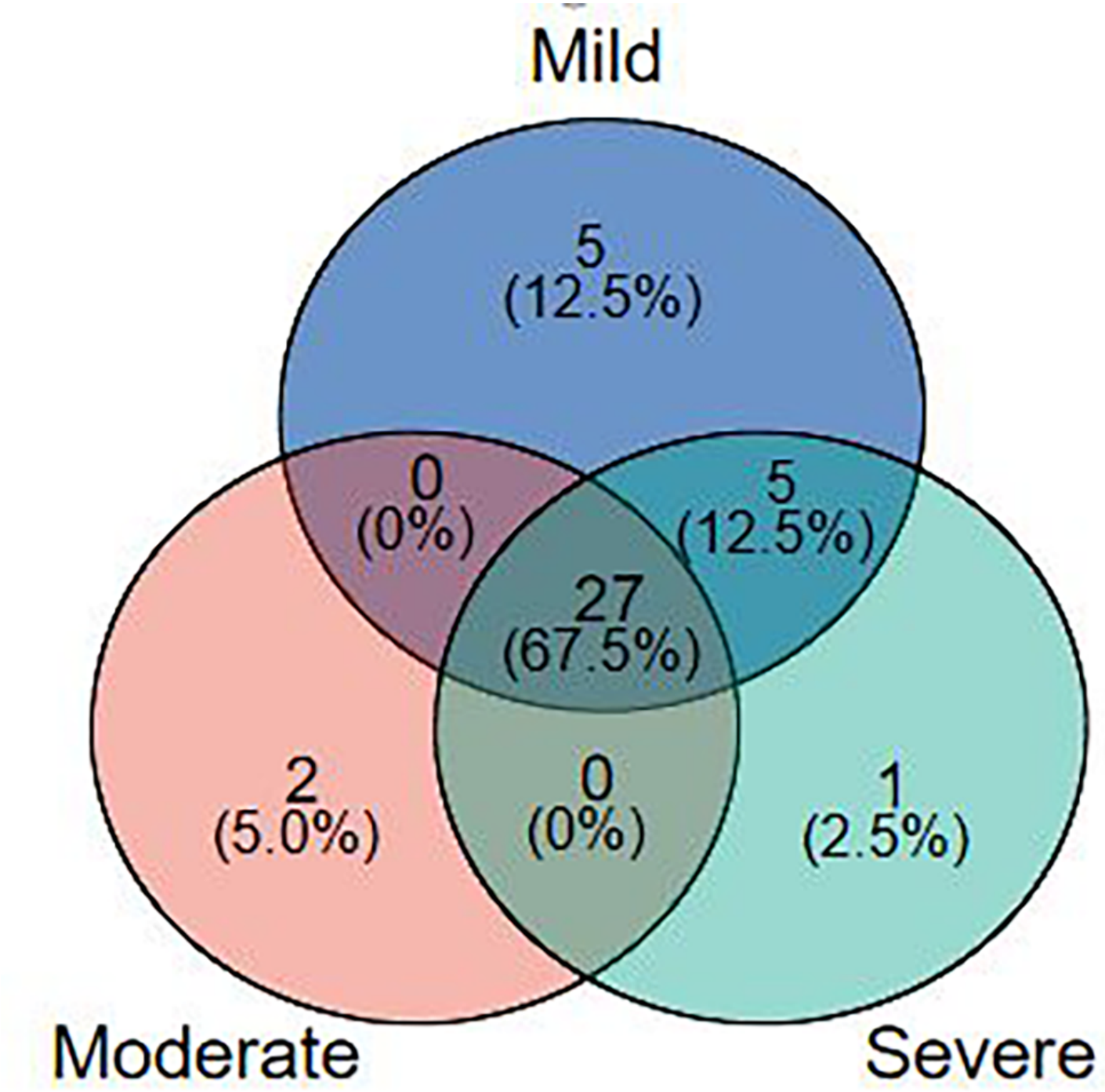

Figure 1

Venn diagram A Venn diagram represents common DEGs. Twenty-seven genes are found common among the 40 DEGs of pSS patients with varying degrees of fatigue. The common DEGs accounted for 67.5% out of a total of 40 DEGs.

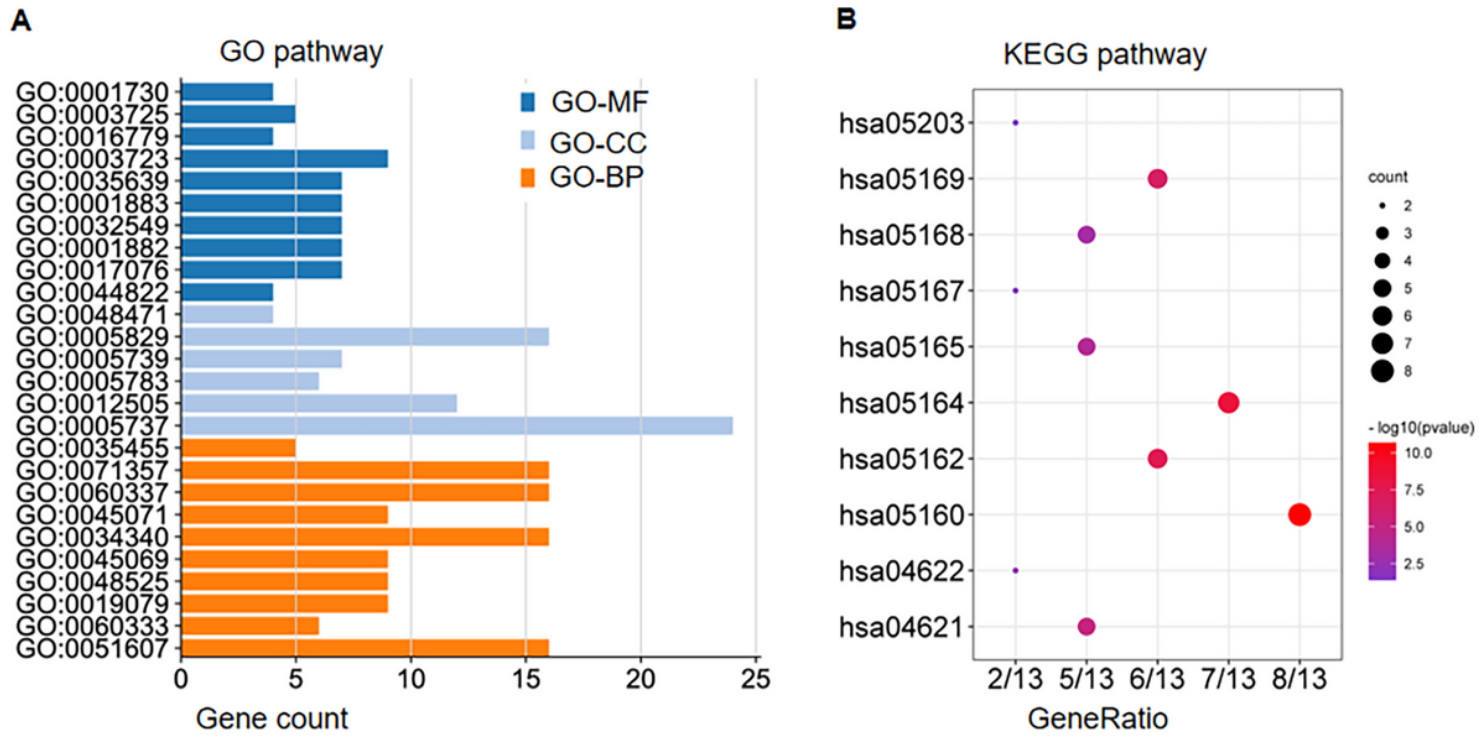

Figure 2

GO-KEGG diagram (A) The result of biological process, molecular function and cellular component associated GO terms; (B) The result of KEGG pathway analysis result.

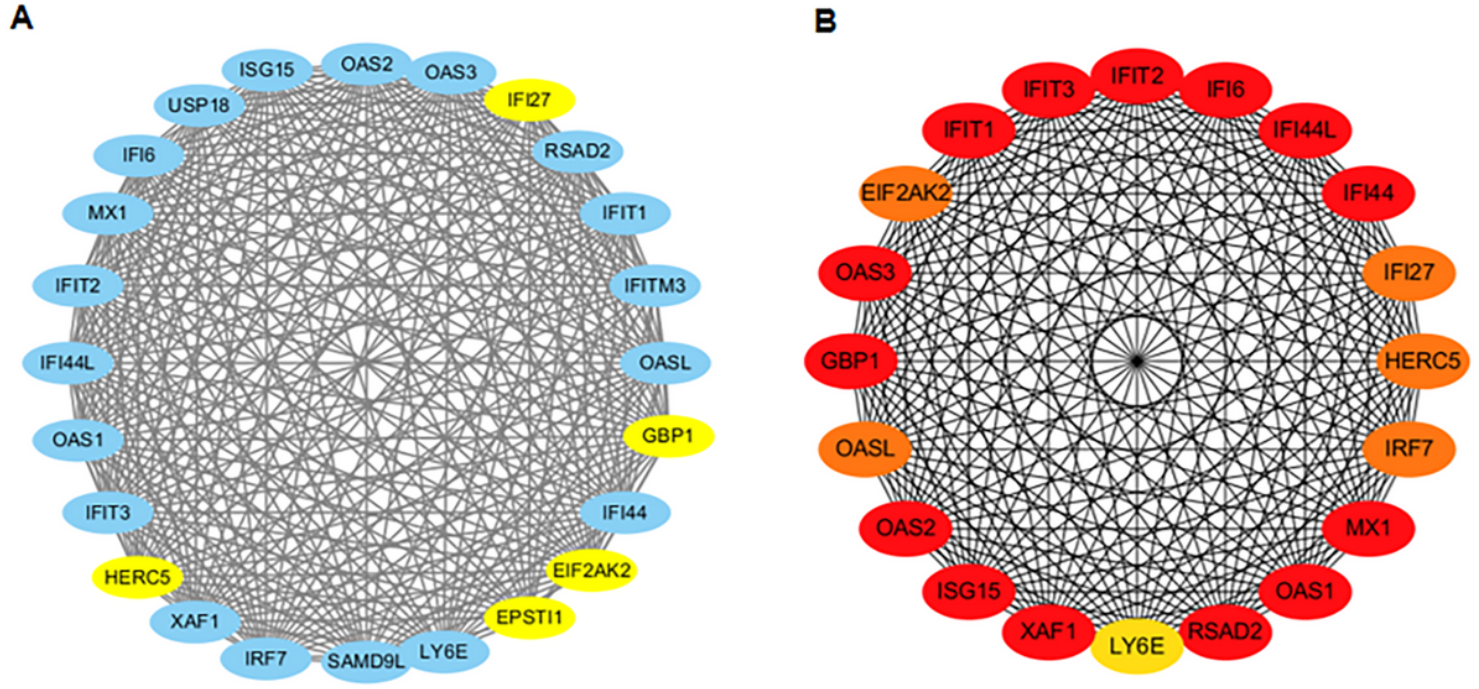

Figure 3

Protein-protein interactions (PPIs) network and the top 20 genes (A) Protein-protein interactions (PPIs) network for DEGs of patients with varying degrees of fatigue. Nodes in blue color indicate common DEGs

with the highest connection degree value and edges specify the interconnection in the middle of two genes. The network involves 24 nodes and 552 edges; (B) The top 20 genes are detected from the PPIs network of common DEGs. The nodes in red color represent 14 genes with the highest degree value which are MX1, IFIT1, ISG15, RSAD2, IFI44L, IFI44, IFIT3, OAS2, OAS1, IFI6, XAF1, IFIT2, GBP1 and OAS3. Based on the topological analysis, the degree value of these 14 genes is both 23.

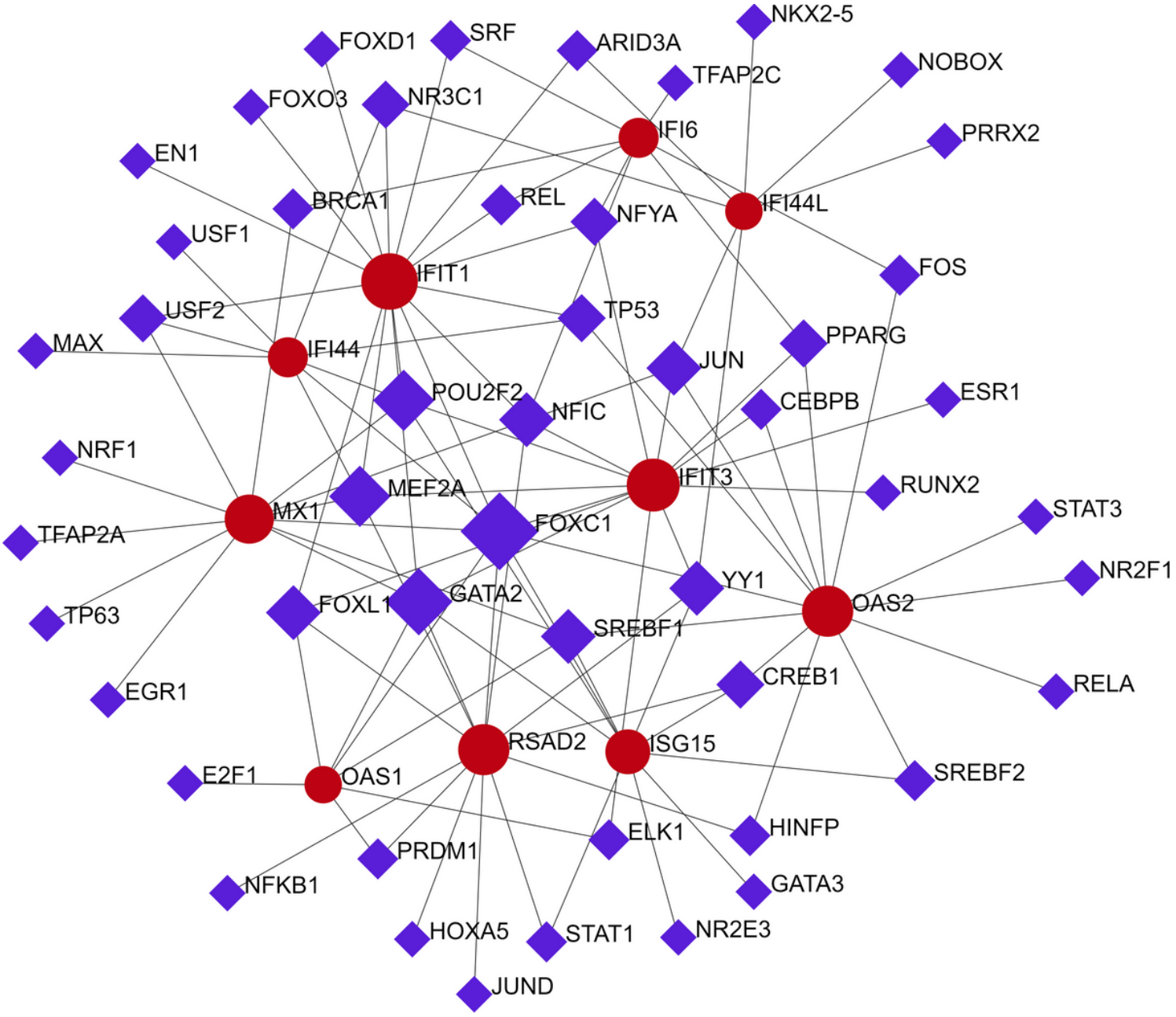

Figure 4

The network presents the TF-gene Network for TF-gene interaction with common DEGs. The highlighted 10 red color nodes represent the common genes and other nodes represent TF-genes. The network has 60 nodes and 108 edges.

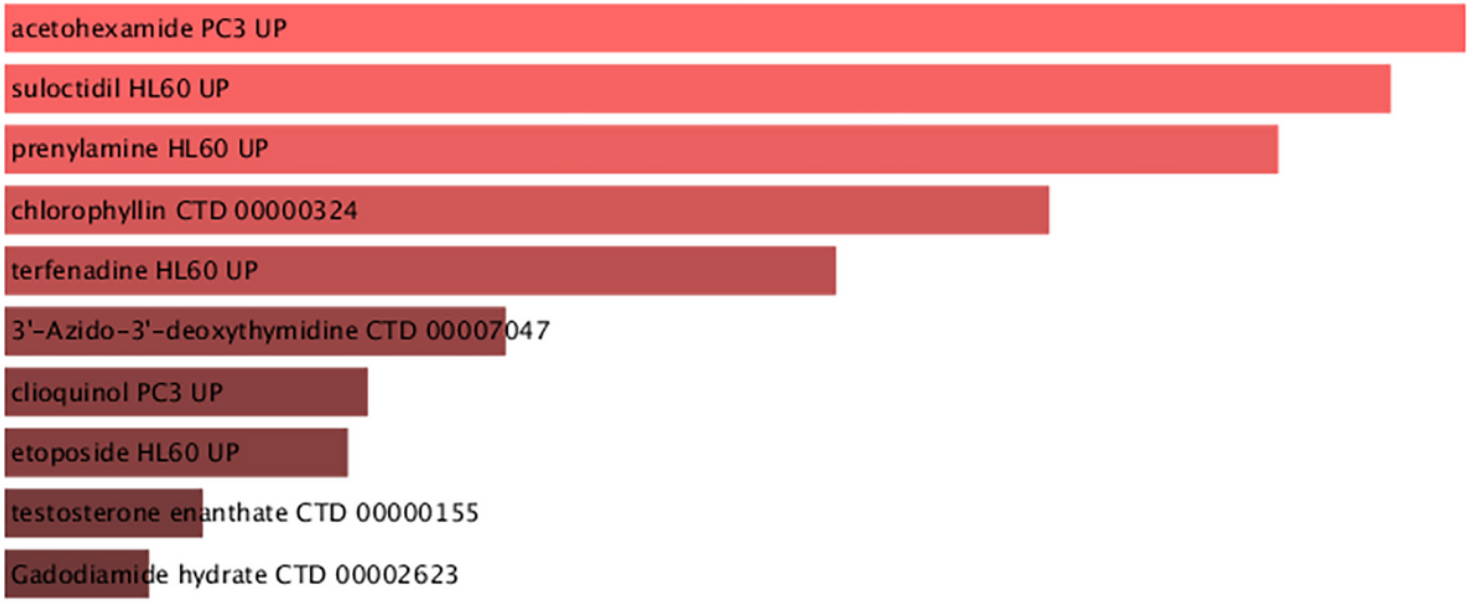

Figure 5

Suggested top drug compounds for pSS with fatigue Acetohexamide PC3 UP, suloctidil HL60 UP, prenylamine HL 60 UP, chlorophyllin CTD 00000324 are the four most drug molecules that connect with the top 10 hub genes.