 doi:10.20944/preprints202206.0176.v1

Article

Diagnostic Biomarkers Screened by Machine Learning Algorithms in Ankylosing Spondilytis

Jian Wen[1], Lijia Wan[2], Xieping Dong[3,*]

[1] Medical College of Nanchang University; Department of Orthopedics, Jiangxi Provincial People's Hospital; JXHC Key Laboratory of Digital Orthopaedics (Jiangxi Provincial People's Hospital); finerain120@163.com
[2] Department of Pediatrics, The Second Xiangya Hospital, Central South University; wanlijia0225@163.com
[3] Medical College of Nanchang University; Department of Orthopedics, Jiangxi Provincial People's Hospital; JXHC Key Laboratory of Digital Orthopaedics (Jiangxi Provincial People's Hospital); jxsrmyy-dongxieping@163.com
* Correspondence: jxsrmyydongxieping@163.com; Tel.: +8613576030901

Abstract: Ankylosing spondylitis (AS) is a chronic inflammatory disorder with unknown etiology and hard to early diagnose. It's imperative to investigate the changes in AS patients' peripheral blood, which may contribute to the diagnosis and further understanding of AS. Common differential expressed genes between normal and AS patients in GSE73754 and GSE25101 were screened by machine learning algorithms. IL2RB and ZDHHC18 were hubgenes screened and a diagnostic model was established. C-indexes and calibration analyses suggested high prediction accuracy of the model in training and validation cohorts. The AUC values of the model in GSE73754, GSE25101, GSE18781 and GSE11886 were 0.86, 0.84, 0.85 and 0.89 respectively. Decision curve analyses suggested high net benefit by the model. Functional analysis of the differential expressed genes indicated that they were mainly clustered in processes related to immune response. Immune microenvironment analysis revealed that neutrophils were expanded and activated in AS, while some T cells were decreased. IL2RB and ZDHHC18 were potential blood biomarkers for AS and might be used for early diagnosis and a supplementary diagnostic tool to the existing methods. Our study deepened the insight into the pathogenesis of AS.

Keywords: machine learning; ankylosing spondylitis; diagnostic model; immune microenvironment; informatics.

1. Introduction

Ankylosing spondylitis (AS), also known as radiographic axial spondyloarthritis, is one of the two types of axial spondyloarthritis[1-4]. It's a chronic inflammatory disorder mainly affecting the axial joints and entheses, usually characterized by typical features: like inflammatory back pain, limitation of the motion of the lumbar spine, restricted chest expansion and advanced sacroiliitis on plain radiographs. Some AS patients are also accompanied by peripheral spondyloarthritis symptoms, like dactylitis, achilles tendinitis and extra-articular manifestations, like uveitis, psoriasis, inflammatory bowel disease and etc. at the same time or in some period of the disease. The diagnosis of AS is based on the Modified New York criteria: advanced sacroiliitis on plain radiographs with any one of three typical features mentioned above[5]. Patients don't meet the criteria of advanced sacroiliitis on plain radiographs, but with sacroiliitis on MRI or HLA-B27 positivity plus the clinical criteria are classified into non-radiographic axial spondyloarthritis[6,7].

The prevalence of AS reported is varied with geography, ranged from 0.02-0.35%, while the prevalence of axial spondyloarthritis is estimated to be 0.20-1.61%, much higher than AS, which indicates a high ratio of non-radiographic axial spondyloarthritis patients[8-10]. Especially, with the development of diagnostic tools and further understanding of axial spondyloarthritis, patients without advanced sacroiliitis on plain radiographs

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2022 doi:10.20944/preprints202206.0176.v1

raise more attention and more non-radiographic axial spondyloarthritises are detected together with its definition updating[3,11]. However, even with modern diagnostic methods, the diagnostic sensitivity and specificity for axial spondyloarthritis is not higher than about 80%[2]. This means that a significant number of patients are still excluded from the current criteria and there is still a lot of room for improvement in our diagnostic methods. More importantly, it's reported that about 10-20% non-radiographic axial spondyloarthritis patients will progress to AS in one year after first diagnosis and 20.3% in two to six years[2]. Therefore, it's necessary to identify pre-AS and save time for clinical interventions.

At present, our measures to identify axial spondyloarthritis are still limited beyond clinical features. Imaging (radiograph, CT, MRI), HLA-B27 and C reactive protein features are the main indices for clinical diagnosis of axial spondyloarthritis[2,11,12]. More methods with high sensitivity and specificity are eagerly expected. Although with the rapid development of genomics technology, many biomarkers were identified for AS, but still lack of reliable indices for clinical practice besides HLA-B27. Therefore, it's not only of realistic needs and great practical value to explore gene biomarkers of AS in peripheral blood, but also can deepen our knowledge in the pathophysiology changes of AS, even do some help to understand its etiology.

Thereby, in this study we intend to screen potential gene biomarkers in peripheral blood by machine learning algorithms and build a diagnostic model. Then, preliminarily explore the immune microenvironment of AS aiming to find some differences in immune cell proportion and potential explanation for our hubgenes. To date, this work has not been done and reported, it's valuable and imperative to bridge the gap in this area.

2. Materials and Methods

2.1 Data collection.

We searched Gene Expression Omnibus (GEO) dataset (https://www.ncbi.nlm.nih.gov/geo/) for datasets containing whole-blood RNA expression data of normal and AS patients with at least 15 samples in each group. Only GSE73754, GES25101 and GSE18781 were qualified and their expression and phenotype data were downloaded for next study. GES73754 and GSE18781 contained whole-blood RNA expression data of 20 normal and 52 AS patients, 25 normal and 18 AS patients respectively, together with their corresponding basic information like gender and age. Expression data of GSE73754 were detected by Illumina HumanHT-12 V4.0 expression beadchip, University of Toronto, Canada, submitted on Oct 06, 2015. Expression data of GSE18781 were detected by Affymetrix Human Genome U133 Plus 2.0 Array, Oregon Health & Science University, USA, submitted on Oct 28, 2009. GSE25101 contained whole-blood RNA expression data of 16 normal and 16 AS patients, which were detected by Illumina HumanHT-12 V3.0 expression beadchip, University of Queensland Diamantina Institute, Australia, submitted on Nov 03, 2010. However, the basic information of the subjects from GSE25101 was unavailable, so it's only used as one of the validation sets. GSE11886 referred to RNA expression data of in vitro cultured macrophages, which were obtained from 9 normal and 8 AS patients' peripheral blood. They were detected by Affymetrix Human Genome U133 Plus 2.0 Array, Cincinnati Childrens Hospital Medical Center, USA, submitted on Jun 25, 2008. Although RNA expression data of each set were normalized data, while in the quality control process, we found samples of GSE18781 came from two batches, so we used removebatcheffect function of limma package to recalculate the expression data[13].

2.2 Identify common differentially expressed genes.

Differentially expressed genes (DEGs) in GSE73754 and GSE25101 between normal and AS patients were identified by limma package[13] (cutoff value: the absolute value of log2foldchange > 0.3 and p-value < 0.05). Then, common DEGs in GSE73754 and GSE25101 were selected as candidates for next screening.

2.3 Screen genes for diagnostic model by machine learning algorithms.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2022 doi:10.20944/preprints202206.0176.v1

Set GSE73754 as training set. Common DEGs were firstly screened by univariate logistic regression in training set. Genes with p value < 0.05 were retained. Then, three machine learning algorithms: the least absolute shrinkage and selection operator (LASSO) logistic regression[14], a support vector machine recursive feature elimination (SVM-RFE)[15] and random forests (RF)[16] were adopted to screen hubgenes. Common hubgenes were selected as final genes for a diagnostic model.

2.4 Establish a diagnostic model and evaluate it in training and a related validation set.

A diagnostic model was established by common hubgenes and visualized by nomograms. Then, the prediction accuracy and discriminatory capacities were first assessed in GSE73754 and GSE25101 by C-index, calibration analysis, receiver operating characteristic (ROC) curves and decision curve analysis (DCA).

2.5 Validate the model in validation sets.

GSE18781 was set as an in vivo external validation set and GSE11886 was set as an in vitro external validation set. The prediction accuracy and discriminatory capacities of the model were also assessed in the two cohorts above by C-index, calibration analysis, ROC analysis and DCA.

2.6 Functional analysis of the DEGs between normal and AS groups.

GO and KEGG clustering, gene set enrichment analysis (GSEA) were used to explore the potential functions of the DEGs, which might indicate the causes for the difference between normal and AS patients. With the same consideration, protein protein interact (PPI) network analysis was also adopted to investigate the interaction among proteins encoded by the DEGs.

2.7 Immune microenvironment analysis.

CIBERSORT package was employed to investigate the immune microenvironment (IME) of the samples. Meanwhile, the correlation between different types of immune cells and the hubgenes were also explored.

2.8 Statistical analysis

In this study, R software v3.63 was used to process data and generate charts. PPI network analyses were explored on STRING website (https://cn.string-db.org/) and visualized by Cytoscape software v3.7.1. Flexible statistical methods were adopted for the statistical analysis.

3. Results

3.1 Clinical characteristics of the enrolled AS patients.

The basic information of samples from GSE73754 and GSE18781 were shown in table 1. The clinical characteristics like age and gender of the two sets were similar (p value < 0.05).

Table 1. Clinical characteristics in training and validation sets

Characteristics	level	GSE18781	GSE73754	*P* value	test
sample size (n)		43	72		
gender	Female	25 (58.1)	35 (48.6)	0.342	fisher.test
	Male	18 (41.9)	37 (51.4)		
age, median (interquartile range)		45.0 [32.5, 58.5]	41.5 [28.8, 51.2]	0.324	kruskal.test
group	Normal	25 (58.1)	20 (27.8)		
	AS	18 (41.9)	52 (72.2)		

3.2 Identification of hubgenes.

64 down-regulated and 132 up-regulated DEGs were identified by limma in GSE73754 (Figure 1A). And 278 down-regulated and 345 up-regulated DEGs were identified in GSE25101 (Figure 1B). Then, common up-regulated and down-regulated genes were chosen. And there were 3 common down-regulated: IL2RB, GZMM, CXXC5 (Figure

 doi:10.20944/preprints202206.0176.v1

1C) and 4 common up-regulated: S100A12, ANXA3, PROS1, ZDHHC18 (Figure 1D) genes.

Took GSE73754 as training set, in univariate logistic regression, the p value of the 7 genes were all lower than 0.05, which meant that all the 7 genes were qualified for the next screening. Then, they were screened by three machine learning algorithms, respectively. IL2RB, GZMM, S100A12, ZDHHC18 were screened as hubgenes by LASSO (λ= lambda.min) (Figure 1E-F). IL2RB, ZDHHC18 were filtered by SVM-RFE (Figure 1G). ZDHHC18, CXXC5, PROS1, IL2RB were screened by RF with MeanDecreaseAccuracy > 3 and MeanDecreaseGini > 2 (mtry = 3, ntree = 200) (Figure 1H-I). Obviously, IL2RB and ZDHHC18 were the common hubgenes screened by the three algorithms and they were selected as the final hubgenes for a diagnostic model in AS.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2022 doi:10.20944/preprints202206.0176.v1

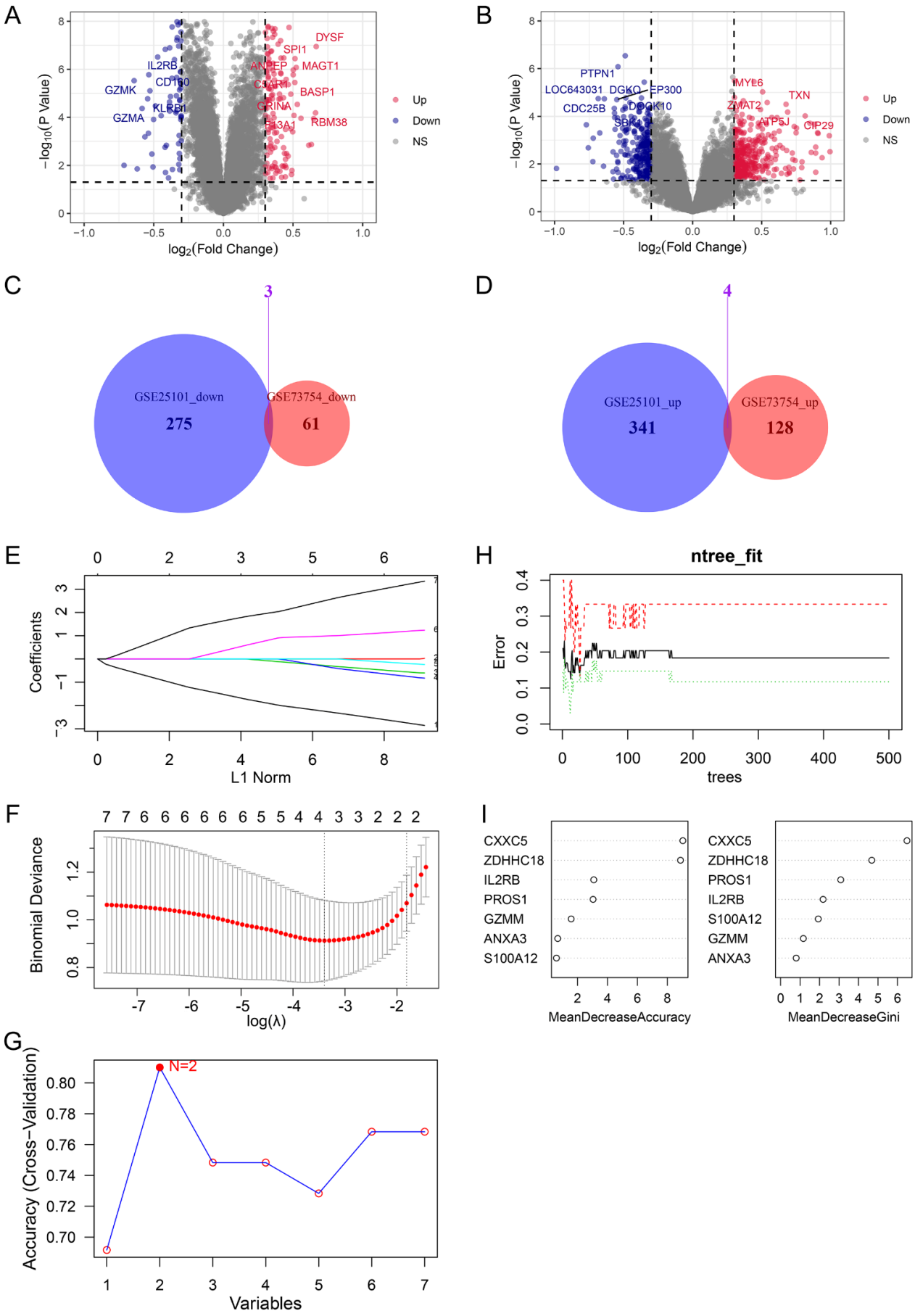

Figure 1. Screening hubgenes from DEGs between normal and AS patients. Volcano plot for DEGs in GSE73754 (A) and GSE25101 (B): x-axis was log_2(Fold change) of gene expressions in AS patients

 doi:10.20944/preprints202206.0176.v1

compared with normal controls, y-axis was -$\log_{10}$(P Value) of gene expressions between AS patients and normal controls. (C) Venn plot for down regulated DEGs in GSE73754 and GSE25101. (D) Venn plot for up regulated DEGs in GSE73754 and GSE25101. (E) The LASSO coefficient profiles for the 7 common DEGs in the 10-fold cross-validations. (F) Partial likelihood deviance with changing of log (λ) plotted by LASSO regression in 10-fold cross-validations. (G) Filtering characteristic genes by SVM-RFE algorithm: accuracy for models with different number of variables: x-axis was the number of variables involved in the models, y-axis was the corresponding accuracy of cross-validation of the models. (H) Relationship between the number of decision trees and the error rate of the model in RF. (I) Selecting hubgenes by variable importance measures for RF.

3.3 Evaluation of the diagnostic model in training set (GSE73754) and GSE25101.

A diagnostic model was established by IL2RB and ZDHHC18, then visualized by a nomogram in GSE73754 (Figure 2A) and GSE25101 (Figure 2B) respectively. Here, their nomograms suggested that the higher the expression level of ZDHHC18, the higher the risk of AS, while the reverse was true for IL2RB. C-index of the diagnostic model in GSE73754 was 0.86 (95%CI: 0.76-0.96), and 0.84 (95%CI: 0.71-0.97) in GSE25101. Calibration analysis showed that the predicted probability was in high agreement with the observed probability, which suggested high accuracy of the model both in training and an external cohort (Figure 2C-D).

ROC analysis in GSE73754 showed that the area under curves (AUCs) for nomogram, IL2RB and ZDHHC18 were 0.86, 0.83, 0.83 respectively (Figure 2E). The optimal truncation value of Y was 0.713, and the corresponding specificity and sensitivity were 0.85 and 0.827, respectively (Formula : y = 2.9111*EXPZDHHC18 — 2.3256*EXPIL2RB — 2.2376, EXPZDHHC18 referred to the expression value of ZDHHC18, EXPIL2RB referred to the expression value of IL2RB). In this model, value of Y >= 0.713 was predicted to be AS, otherwise normal. The actual prediction accuracy of the model in GSE73754 was 0.82. While in GSE25101 the AUCs for nomogram, IL2RB and ZDHHC18 were 0.84, 0.79, 0.76 respectively (Formula : y = 2.320052*EXPZDHHC18 — 1.728388*EXPIL2RB — 6.902309) (Figure 2F). There were three optimal truncation value for Y: 0.589 with corresponding specificity 0.875, sensitivity 0.688, 0.521 with corresponding specificity 0.75, sensitivity 0.812, 0.452 with corresponding specificity 0.688, sensitivity 0.875. The actual prediction accuracy of the model in GSE25101 was 0.72. DCA for nomogram and models involved only one of these genes indicated that the nomogram got higher net benefit than other models (Figure 2G-H).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2022 doi:10.20944/preprints202206.0176.v1

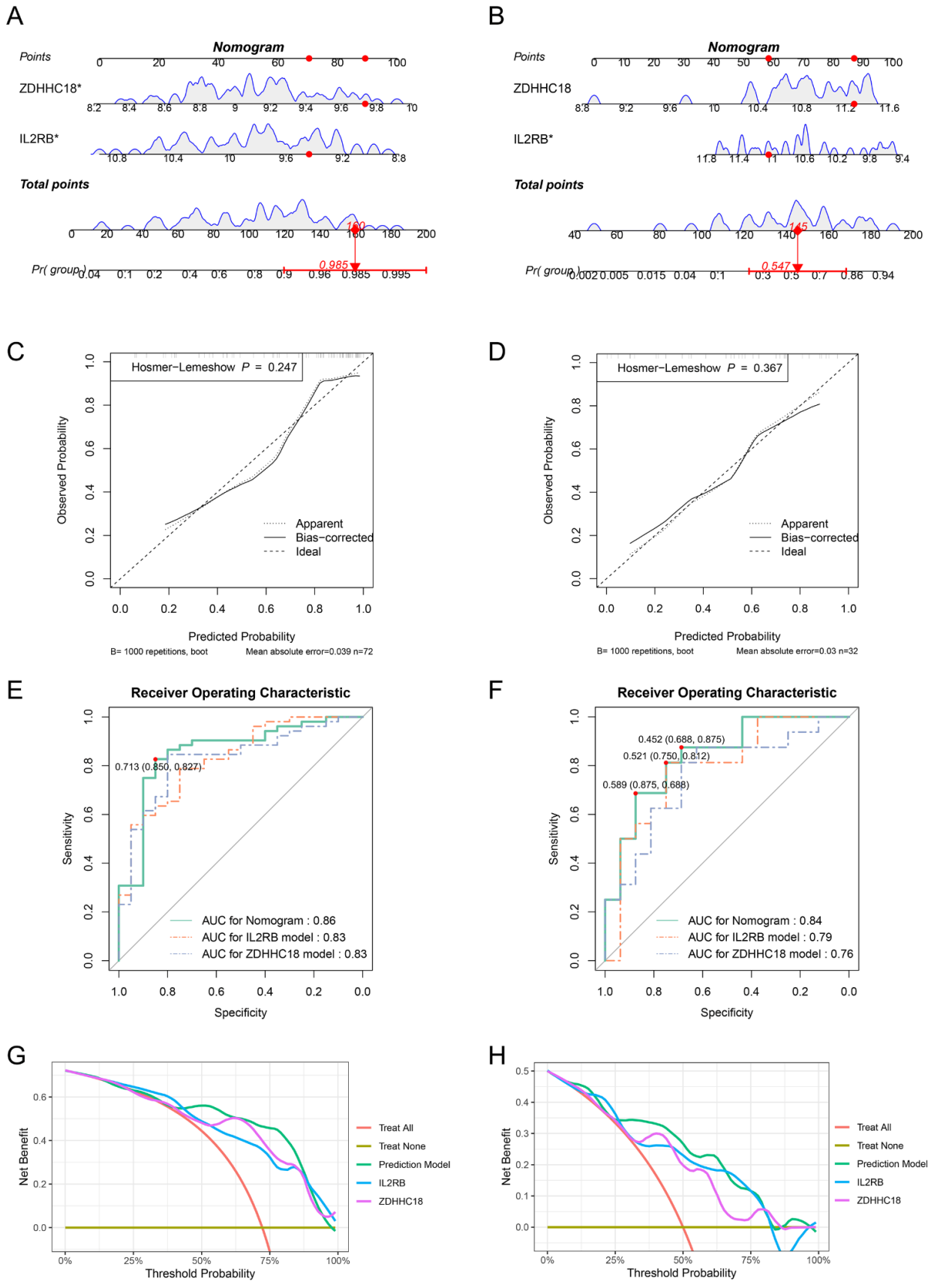

Figure 2. Evaluating the diagnostic model in training and a related validation set. Nomograms for the diagnostic model in GSE73754 (A) and GSE25101 (B). Calibration plots for the diagnostic model

 doi:10.20944/preprints202206.0176.v1

in GSE73754 (C) and GSE25101 (D): x-axis referred to the predicted probability of AS by the model, y-axis referred to the observed probability of AS, the diagonal (dashed line) referred to the ideal status that the predicted probability equaled to the observed probability, the solid and dotted lines referred to the apparent and bias-corrected statuses of the predicted and observed probabilities respectively. ROC plots for the diagnostic model in GSE73754 (E) and GSE25101 (F): x-axis was 1 - specificity of the model, y-axis was the sensitivity of the model. DCA in GSE73754 (G) and GSE25101 (H): x-axis referred to threshold probability for treatment or intervention, y-axis referred to the net benefit.

3.4 Validating the model in an independent cohort and in vitro.

The model was validated in an independent cohort: GSE18781 and an in vitro cohort: GSE11886. Nomogram for GSE18781 supported the conclusion reached in training set that AS patients got higher expression of ZDHHC18 and lower expression of IL2RB (Figure 3A). The function of IL2RB in GSE11886 was in accordance with that in the other sets, but the function of ZDHHC18 in vitro was opposite to that in vivo, while this might be due to the lack of the in vivo microenvironment (Figure 3B). According to the coverage of points in the nomogram, IL2RB showed higher weight in the validation sets and the alteration between the nomograms also indicated that it's a more robust indicator than ZDHHC18. C-index of the diagnostic model in GSE18781 was 0.85 (95%CI:0.73-0.96), and 0.89 (95%CI: 0.73-1.05) in GSE11886. Calibration analysis showed that the prediction accuracy of the model was lower than that in GSE73754 and GSE25101, but still with acceptable accuracy (Figure 3C-D).

ROC analysis in GSE18781 showed that the area under curves (AUCs) for nomogram, IL2RB and ZDHHC18 were 0.85, 0.79, 0.67 respectively (Figure 3E). The optimal truncation value of Y was 0.305, and the corresponding specificity and sensitivity were 0.72 and 0.994, respectively (Formula : y = 1.29499*EXPZDHHC18 − 2.582298*EXPIL2RB + 20.055204). The actual prediction accuracy of the model in GSE18781 was 0.72. While in GSE11886 the AUCs for nomogram, IL2RB and ZDHHC18 were 0.89, 0.89, 0.65 respectively (Formula : y =−6.49159*EXPZDHHC18−6.13506*EXPIL2RB−0.01334) (Figure 3F). The optimal truncation value of Y was 0.395, and the corresponding specificity and sensitivity were 0.778 and 1. The actual prediction accuracy of the model in GSE11886 was 0.76. DCA showed that patients could get high net benefit from nomogram (Figure 3G-H). Besides, the model established only by IL2RB also exhibited high net benefit for patients with AS in this set.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2022 doi:10.20944/preprints202206.0176.v1

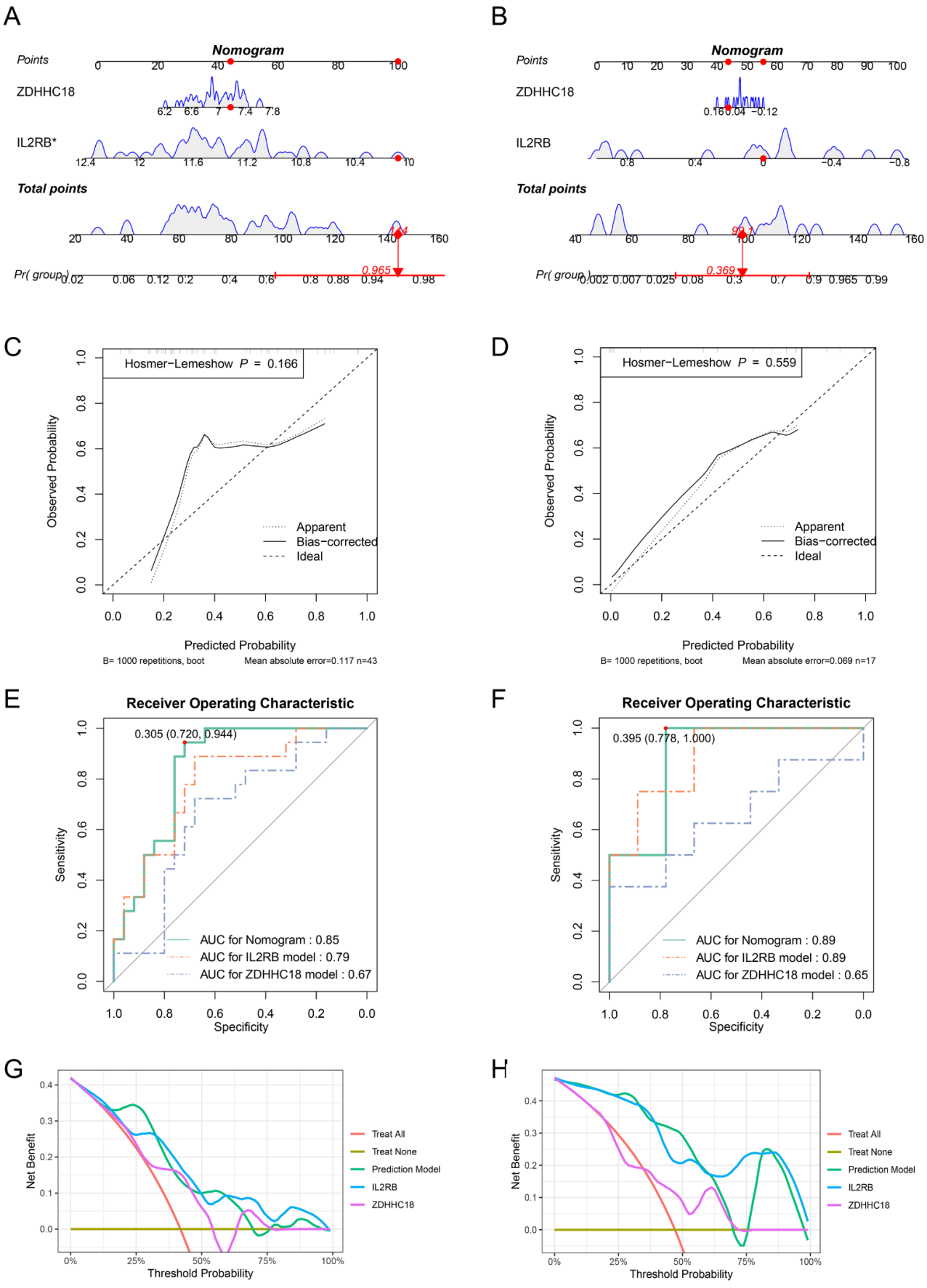

Figure 3. Validating the diagnostic model in validation sets. Nomograms for the diagnostic model in GSE18781 (A) and GSE11886 (B). Calibration plots for the diagnostic model in GSE18781 (C) and

 doi:10.20944/preprints202206.0176.v1

GSE11886 (D): x-axis referred to the predicted probability of AS by the model, y-axis referred to the observed probability of AS, the diagonal (dashed line) referred to the ideal status that the predicted probability equaled to the observed probability, the solid and dotted lines referred to the apparent and bias-corrected statuses of the predicted and observed probabilities respectively. ROC plots for the diagnostic model in GSE18781 (E) and GSE11886 (F): x-axis was 1 - specificity of the model, y-axis was the sensitivity of the model. DCA in GSE18781 (G) and GSE11886 (H): x-axis referred to threshold probability for treatment or intervention, y-axis referred to the net benefit.

3.5 Results for functional analysis of the DEGs between normal and AS groups.

There were a total of 196 DEGs between normal and AS patients in GSE73754. Biological processes (BP) clustering of the DEGs showed that they were mainly clustered in neutrophil activation, degranulation, immune response and migration (Figure 4A). Myeloid cell differentiation, leukocyte and granulocyte migration was also clustered BPs. Gene clustering of cellular components (CC) were mostly in the area of membranes, like endocytic vesicle, secretory granule membrane, membrane microdomain, cytoplasmic vesicle lumen and etc. (Figure 4A). Molecular functions (MF) of the DEGs were mostly clustered in serine type peptidase activity, serine hydrolase activity, serine type endopeptidase activity, MHC protein complex binding (Figure 4A). In KEGG clustering of the DEGs, hematopoietic cell lineage, human T-cell leukemia virus 1 infection, Th1 and Th2 cell differentiation and Th17 cell differentiation were the top clustered pathways (Figure 4B). Circle plot for BP clustering showed that neutrophil activation, degranulation, immune response and migration were up regulated in AS (Figure 4C). By GSEA, antigen processing and presentation, natural killer cell mediated cytotoxicity, graft-versus-host disease, Epstein-Barr virus infection, rheumatoid arthritis were top enriched gene sets, which were all down regulated in AS patients (Figure 4D). While the top 3 up regulated pathways enriched with core enrichment genes > 3 were neutrophil extracellular trap formation, complement and coagulation cascades and rap1 signaling pathway. GO chord plot showed that DYSF, DMTN, ITGA2B, MAGT1, SPI1, CXCL8, ID2, CD81, IKZF1 and etc., were involved in the top 7 GO terms (Figure 4E). KEGG chord plot showed that ITGA2B, SPI1, ANPEP, BCL2L1, STAT5B, IL2RB, GZMB, HLA-DQA2, CXCL8 and etc., were involved in the top 7 KEGG terms (Figure 4F).

PPI network of the proteins encoded by DEGs showed that MMP1, ID2, MBD4, GNLY, EOMES, PUF60 and APOBEC3G were seed proteins in the network by MCODE application in cytoscape (Figure 4G). The cyan nodes were also pivotal nodes in the net, such as IL2RB, GZMA, SPI1 and etc. Then, GZMA, IL2RB, CD247, KLRB1, GZMH, GZMB, GZMK, KLRD1, NKG7, GNLY were the top 10 hub proteins screened by Cytohubba.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2022 doi:10.20944/preprints202206.0176.v1

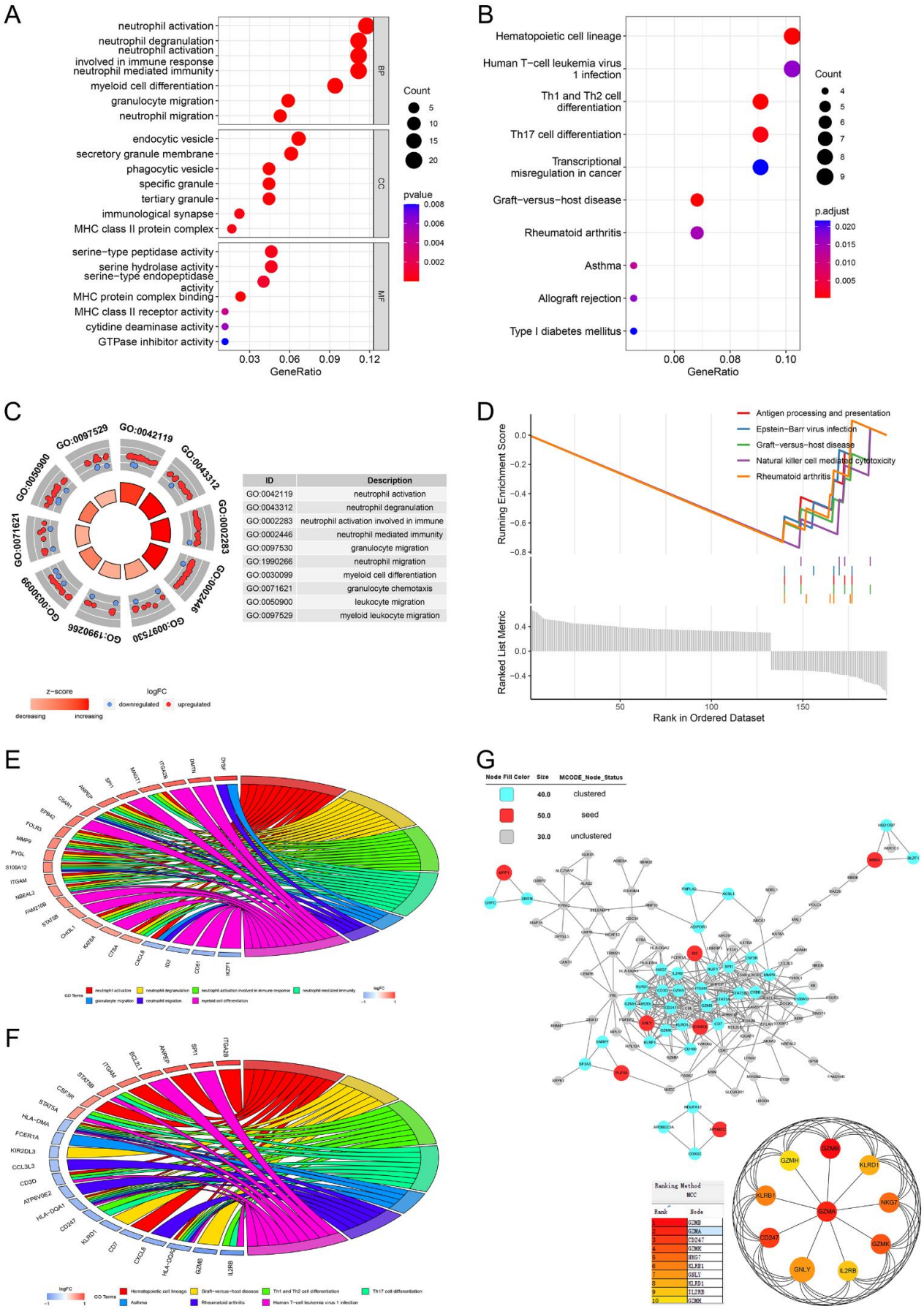

Figure 4. Functional analysis of the DEGs between normal and AS patients. Dot plots for GO (A) and KEGG (B) analysis of DEGs. (C) Circle plot for BP clustering of the DEGs. (D) GSEA analysis

 doi:10.20944/preprints202206.0176.v1

for the DEGs. (E) Chord plot for top 7 clustered GO terms. (F) Chord plot for top 7 clustered KEGG pathways. (G) PPI network analysis for DEGs.

3.6 Results for IME analysis.

IME analyses were performed in GSE73754, GSE25101 and GSE18781 by CIBERSORT. The proportion of the 22 immune cells for samples were shown in Figure 5A-C. In all the three sets, neutrophils and monocytes accounted for the top 2 highest proportion and together made up the majority of the immune cells, while other granulocytes, B cells, dendritic cells and macrophages each made up a very small proportion of the immune cells. The relative quantity of different immune cells in normal and AS patients were shown in Figure 5D-F. In GSE73754, compared to normal subjects, there were more neutrophils and T CD4 naive cells detected in AS patients' blood, while less NK resting, T CD8 and T gamma delta cells (Figure 5D). In GSE25101, monocytes were found to be higher in AS patients' blood, while T regulatory cells (Tregs) were lower. In this set, relative quantity of neutrophils was also found to be higher in AS group, but with no statistical significance (Figure 5E). In GSE18781, the result was similar to that in GSE73754, the relative quantity of neutrophils was increased, while T CD8 and T gamma delta cells were decreased in AS patients (Figure 5F).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2022 doi:10.20944/preprints202206.0176.v1

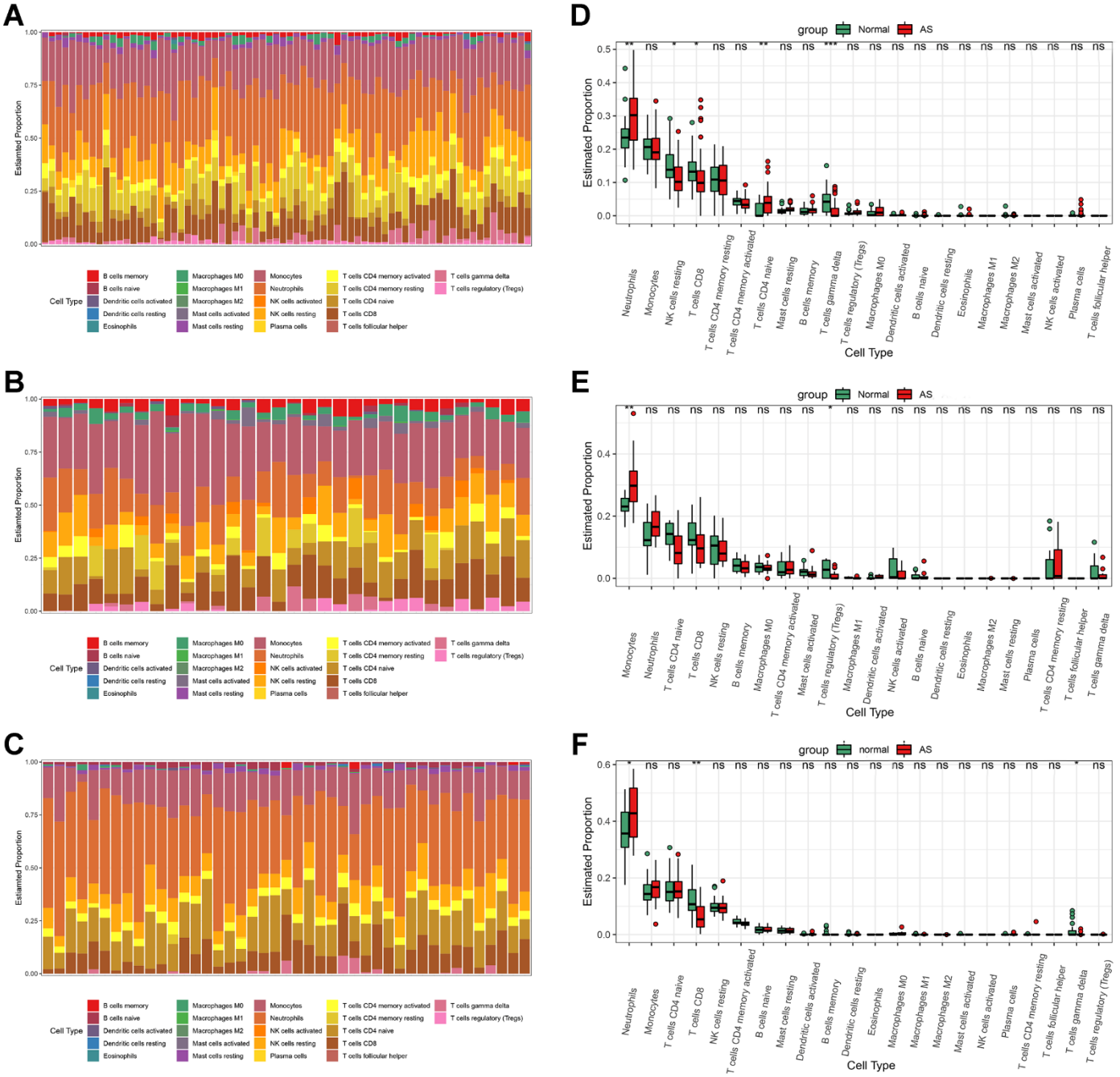

Figure 5. IME analysis of the sets by CIBERSORT. The proportion of the 22 immune cells for samples in GSE73754 (A), GSE25101 (B) and GSE18781 (C). Boxplots for the 22 immune cells between normal and AS patients in GSE73754 (D), GSE25101 (E) and GSE18781 (F). (p significance level: no significance (ns), p⩾0.05; *, p< 0.05; **, p<0.01; ***, p<0.001, ****, p<0.0001.)

The correlation between our hubgenes (IL2RB and ZDHHC18) and immune cells were also explored. In GSE73754, the expression of IL2RB was positively correlated with the relative quantities of NK resting, T CD8 and T gamma delta cells (Figure 6A-C), while negatively correlated with the relative quantities of neutrophils, T CD4 naive cells and monocytes (Figure 6D-E). Meanwhile, the expression of ZDHHC18 was positively correlated with the relative quantity of neutrophils (Figure 6F), but negatively correlated with the relative quantities of T CD8 and NK resting cells (Figure 6G-H). In GSE25101, the expression of IL2RB was positively correlated with the relative quantities of NK resting and T CD4 memory activated cells (Figure 6I-J), while negatively correlated with the relative quantity of monocytes (Figure 6K). Besides, the expression of ZDHHC18 was positively

 doi:10.20944/preprints202206.0176.v1

correlated with the relative quantity of neutrophils (Figure 6L), but negatively correlated with the relative quantity of activated NK cells (Figure 6M). There was no significant correlation between Tregs and the hubgenes. Lastly, in GSE18781, the expression of IL2RB was positively correlated with the relative quantities of NK resting and T CD8 cells (Figure 6N-O), while negatively correlated with the relative quantity of neutrophils (Figure 6P). Moreover, the expression of ZDHHC18 was positively correlated with the relative quantity of neutrophils (Figure 6Q), but negatively correlated with the relative quantities of T CD8, T gamma delta and T CD4 memory activated cells (Figure 6R-T).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2022 doi:10.20944/preprints202206.0176.v1

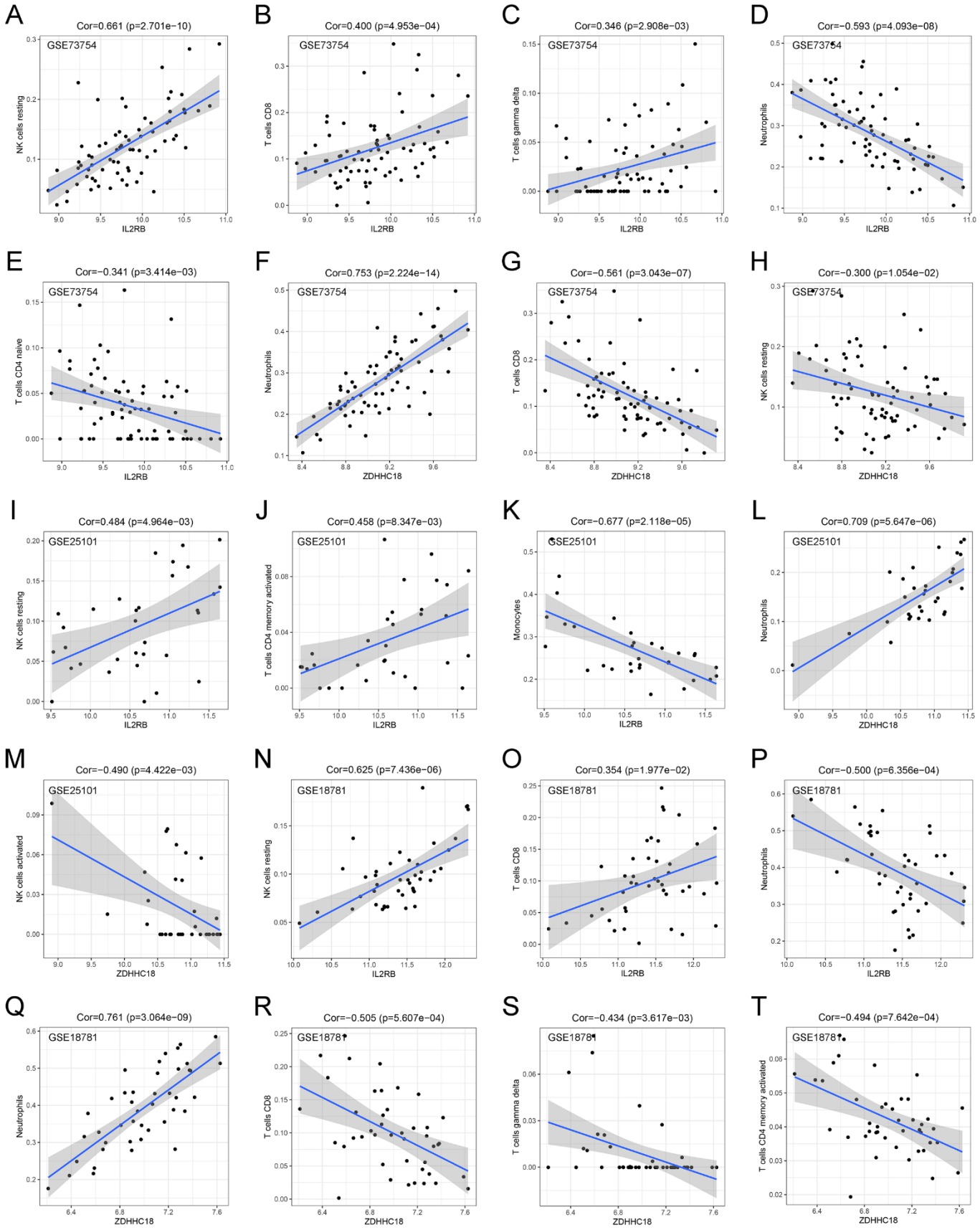

Figure 6. The correlation between hubgenes and IME cells. The correlation between the expression of IL2RB and the estimated proportion of NK resting cells (A), T CD8 cells (B), T gamma delta cells (C), neutrophils (D), T CD4 naive cells (E) by CIBERSORT in GSE73754. The correlation between the expression of ZDHHC18 and the estimated proportion of neutrophils (F), T CD8 cells (G), NK

 doi:10.20944/preprints202206.0176.v1

resting cells (H) by CIBERSORT in GSE73754. The correlation between the expression of IL2RB and the estimated proportion of NK resting cells (I), T CD4 memory activated cells (J), monocytes (K) by CIBERSORT in GSE25101. The correlation between the expression of ZDHHC18 and the estimated proportion of neutrophils (L), activated NK cells (M) by CIBERSORT in GSE25101. The correlation between the expression of IL2RB and the estimated proportion of NK resting cells (N), T CD8 cells (O), neutrophils (P) by CIBERSORT in GSE18781. The correlation between the expression of ZDHHC18 and the estimated proportion of neutrophils (Q), T CD8 cells (R), T gamma delta cells (S), T CD4 memory activated cells (T) by CIBERSORT in GSE18781.

4. Discussion

It's known that AS is an inflammatory disease mainly involve axial skeleton joints and entheses. The essential change of AS is the dysregulation of inflammation by innate and adaptive immune responses[17]. Although AS is primarily associated with axial skeleton, but recent researches indicate that it may initiate in the gut[18]. Besides, the peripheral and the extra-articular manifestations of AS also suggest that it's a systemic disorder. Therefore, DEGs in peripheral blood of AS patients can also reflect some features of AS. While as to RNAs extracted from peripheral blood, they are mostly from the nucleated cells in the blood, like leukocytes and immature red blood cells, so it's rational to explore the immune microenvironment of AS patients' blood. More importantly, compared with the focal tissue, peripheral blood is easier to obtain and a more commonly used clinical detection material, which is also conducive to the transition from experimental results to applications.

To date, HLA-B27 was still considered as the most important factor in the pathogenesis of AS29-33. Firstly, many evidences supported that alternation of amino acid sequence in antigenic peptide-binding groove of HLA-B27 might induce changes of the binding specificity for peptides and result in CD8+ T cell mediated immune cross-reactivity in AS focus34,35. Secondly, endoplasmic reticulum stress induced by the accumulation of misfolded HLA-B27 which leaded to an unfolded protein response (UPR) and autophagy36. Thirdly, HLA-B27 homodimer hypothesis, which suggested HLA-B27 homodimer could activate CD8+ T cells and NK cells by the specific receptors on their surface, then activate of IL-23/IL-17 axis37. Certainly, there were also many other hypotheses, including non-MHC hypothesis. While the common thing was that all the hypotheses are focused on the antigen presenting process and the failure or dysfunction of it mostly would result in activation of TNF signaling pathway and IL23/IL17 axis, and eventually leaded to the phenotypes of AS.

Here, in order to enhance the reliability and stability of the results, only common genes screened by the three machine learning algorithms were selected as hubgenes for a diagnostic model. Results of C-index, calibration analysis, ROC analysis and DCA in training and validation sets suggested excellent prediction accuracy and discriminatory capacities of the linear model. Meanwhile, the performance of the model in 3 different validation datasets: one related dataset, one independent dataset and one dataset of in vitro samples hinted it's a model with good applicability.

Functional analyses of DEGs and IME analyses indicated that neutrophils activation, migration and degranulation were activated in AS patients. Meanwhile, the relative quantity or proportion of neutrophils was significantly higher in AS patients. And this was also confirmed by other researchers, who even suggested neutrophil to lymphocyte ratio to be used as an indicator for AS activity[19-21]. Besides, neutrophil extracellular trap formation and complement and coagulation cascades were up regulated in AS, which might induce autoimmune response and was in agreement with the IME analysis result and our current knowledge of AS[22,23]. Moreover, in GSE25101, monocytes were also found to be higher in AS patients and myeloid cell differentiation and leukocyte migration were also clustered in GO clustering. It's known that monocytes shared some similar functions

 doi:10.20944/preprints202206.0176.v1

with neutrophil in immune response and there were also reports that monocytes to lymphocytes ratio increased in AS patients[24-26]. If the increasement of lymphocytes and monocytes were two different subtypes of AS remained unknown. Finally, how these differences in immune cells correlated with AS? A potential explanation was that the increased neutrophils might release excessive IL-17A which was the key cytokine in the pathogenesis of AS. Although mature neutrophils lack the transcriptional machinery to produce IL-17A, but IL-17A was detected in neutrophils, because they could produce and store IL-17A before mature and accumulate it from the extracellular environment[27,28].

Lastly, IL2RB was a hubgene both in GO/KEGG clustering and PPI network analysis. The expression of it was positively correlated with the relative quantities/proportion of NK resting cells and negatively correlated with the relative quantities/proportion of neutrophils and monocytes in our study, which was in coincidence with the data from the Human Protein Atlas (HPA) website[29] (Figure 7A: available from v21.1.proteinatlas.org, https://www.proteinatlas.org/ENSG00000100385-IL2RB/single+cell+type). While ZDHHC18 was observed to positively correlated with the relative quantities/proportion of neutrophils in all the three sets, though it didn't seem to be highly expressed in granulocytes based on the data from HPA website[29] (Figure 7B: available from v21.1.proteinatlas.org, https://www.proteinatlas.org/ENSG00000100385-IL2RB/single+cell+type). Above all, our results suggested that IL2RB might correlated with AS by suppressing the function of NK resting cells and ZDHHC18 might correlated with AS through the function of neutrophils, but the detailed underlying mechanism of that still needed further study.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 June 2022 doi:10.20944/preprints202206.0176.v1

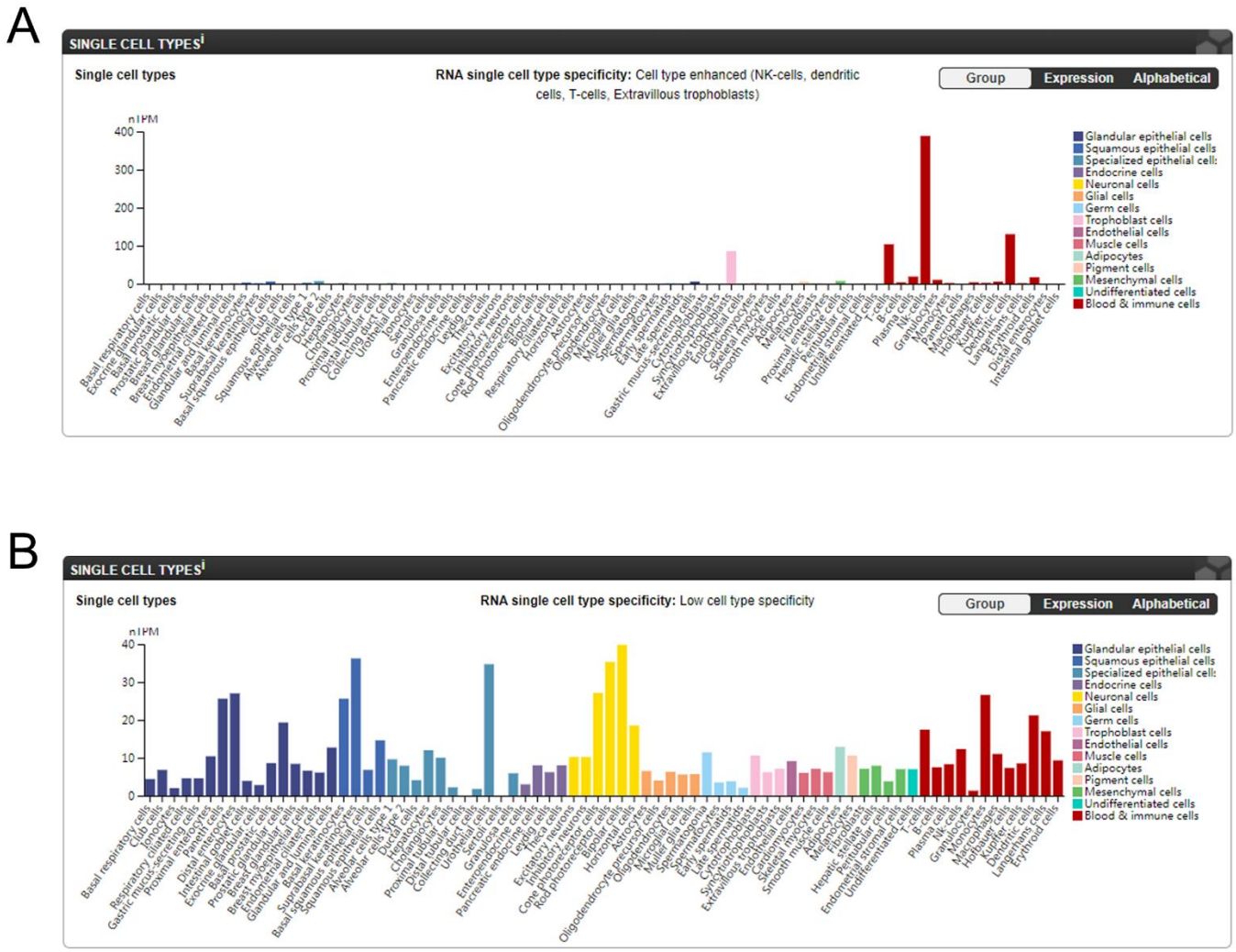

Figure 7. The expression of hubgenes in different single cell types of normal subjects from the HPA website29. (A) The expression of IL2RB in different single cell types (available from v21.1.proteinatlas.org: https://www.proteinatlas.org/ENSG00000100385-IL2RB/single+cell+type). (B) The expression of ZDHHC18 in different single cell types (available from v21.1.proteinatlas.org: https://www.proteinatlas.org/ENSG00000204160-ZDHHC18/single+cell+type).

In this study, IL2RB and ZDHHC18 were the two final screened hubgenes. The former had already been reported by other researchers to be one of the hubgenes in AS[30,31], while the latter was first reported here by us.

First of all, IL2RB, interleukin 2 receptor subunit beta, encoded beta subunit of a heterodimer or heterotrimer receptor involved in T cell-mediated immune responses and probably involved in the stimulation of neutrophil phagocytosis by IL15[32,33]. This protein was a type I membrane protein primarily expressed in NK, T and dendritic cells. According to KEGG database (https://www.kegg.jp/), IL2RB was involved in many pathways, like Endocytosis, PI3K-Akt signaling pathway, JAK-STAT signaling pathway, Th1 and Th2 cell differentiation, Th17 cell differentiation, and etc. Obviously, Th1 and Th2 cell differentiation and Th17 cell differentiation seemed to be most related to AS, that IL2 signaling could inhibit the differentiation of Th17 by the inhibition of transcription factor RORγt[34-38]. Therefore, with the down regulation of IL2RB in this study, Th17 was anticipated to be expanded. However, Th1 and Th2 cell differentiation and Th17 cell differentiation were observed to be down regulated in GSEA (Figure 6D). It's contradictory to our knowledge of AS. Then, something should be noticed. For one thing, the pathogenesis

 doi:10.20944/preprints202206.0176.v1

changes of AS mainly involved in the focus of AS, not in circulation system, and our knowledge was largely based on that, so it's common if there were some differences between the two sites. For another thing, the role of IL2 signaling in the differentiation of Th17 still hadn't been fully clarified[39]. The question was what's the minimum IL2 signal required to maintain the Tregs numbers. Isabel Z Fernandez, et al, reported a hypomorphic mutation of IL2RB in two infant siblings which resulted in an anticipated reduction in Tregs and an expansion of immature NK cells[40]. Here, in two of the 3 sets, the relative quantities of T CD8 and T gamma delta were decreased, while Tregs was not significantly reduced, which might indicate that the reduced IL2 signal was still adequate for the proliferation of Tregs and the suppression of effector T cell expansion (Figure 5D, F). Besides, by blockade of IL-2 in vitro and in vivo, Kenjiro Fujimura, et al. found that the numbers of Th17 was not significantly increased, while the proportion of Th17 cells did, which suggested that it might increase the proportion of Th17 by suppressing the total number of immune cells [41]. Here in our study, some kinds of T cells, like T CD8 cells, T gamma delta cells and Tregs were observed to be decreased in AS patients and this might overwhelm the effect of the down regulating of Th17 cell differentiation. However, which type of immune cells was decreased and if it would affect the synthesizing of IL17 by Th17 cells in AS patients, still needed further research. In the end, the potential function of IL2RB in AS remained unclear, it might contribute to AS by reducing Tregs and relatively increasing the proportion of Th17 and thereby activating IL17 signaling to form phenotypes of AS.

Secondly, ZDHHC18, Zinc Finger DHHC-Type Palmitoyltransferase 18, encoded a palmitoyltransferase, which involved in peptidyl-L-cysteine S-palmitoylation[42]. Researches on ZDHHC18 were rare and insufficient. Currently, it's reported to be associated with innate immunity[43], glioma[44], ovarian cancer[45], schizophrenia[46]. The common palmitoylation substrates of ZDHHC18 was HRAS and LCK[47-49]. Palmitoylated HRAS could be translocated and stable anchored to the plasma membrane[50], while palmitoylation-defective HRAS was trapped in Golgi and unable to traffic to plasma membrane. Meanwhile, ZDHHC18 could activate rap1 signaling pathway by the palmitoylation of Ras and promote the proliferation of cells, which was consistent with our GSEA result. Besides, Rac1, which was also involve in rap1 signaling pathway mainly regulating cell adhesion, migration, polarity, could also be palmitoylated by ZDHHC family[50]. Though we currently didn't know the exact role of ZDHHC18 played in this process, but it's essential for neutrophil motility as well as directional sensing during migration, which was clustered by GO clustering in our study. In addition, palmitoylation of LCK could promote the T cell receptor signaling to activate T cells, while this was not seen in our study, which meant that it's not important in the pathogenesis of AS. Furthermore, ZDHHC18 could negatively regulate CGAS-STING signaling mediated antiviral innate immunity by palmitoylation of cGAS, which meant that the antiviral immunity in AS patient's might be impaired by high expression of ZDHHC18[43]. In our study, KEGG and GSEA also indicated dysregulation in some antiviral immune pathways.

Generally speaking, our study indicated that IL2RB might involve in the pathogenesis of AS through IL2 signaling pathway and ZDHHC18 through rap1 signaling pathway. Both of they could be used as potential biomarkers in AS. Meanwhile, it also should be noticed that we only explored some changes of RNA expression in the peripheral blood of AS patients, it's just only a tip of the iceberg. More researches still needed to further elucidate the pathogenesis of AS.

5. Conclusions

IL2RB and ZDHHC18 were potential blood biomarkers for AS and might be used for early diagnosis and a supplementary diagnostic tool to the existing methods. Our study deepened our insight into the pathogenesis of AS.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Figure S1: title; Table S1: title; Video S1: title.

Author Contributions: Conceptualization, Jian Wen and Lijia Wan; methodology, Jian Wen; software, Jian Wen; validation, Jian Wen, Lijia Wan and Xieping Dong; formal analysis, Jian Wen; investigation, Jian Wen; resources, Jian Wen; data curation, Jian Wen; writing—original draft preparation, Jian Wen; writing—review and editing, Xieping Dong; visualization, Jian Wen; supervision, Jian Wen; project administration, Jian Wen. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: As stated in methods, all the original data were downloaded from the Gene Expression Omnibus (GEO) dataset (https://www.ncbi.nlm.nih.gov/geo/).

Acknowledgments: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Navarro-Compán, V.; Sepriano, A.; El-Zorkany, B.; van der Heijde, D. Axial spondyloarthritis. Ann Rheum Dis 2021, 80, 1511-1521, doi:10.1136/annrheumdis-2021-221035.

2. Sieper, J.; Poddubnyy, D. Axial spondyloarthritis. Lancet 2017, 390, 73-84, doi:10.1016/s0140-6736(16)31591-4.

3. Taurog, J.D.; Chhabra, A.; Colbert, R.A. Ankylosing Spondylitis and Axial Spondyloarthritis. N Engl J Med 2016, 374, 2563-2574, doi:10.1056/NEJMra1406182.

4. Sieper, J.; Braun, J.; Dougados, M.; Baeten, D. Axial spondyloarthritis. Nat Rev Dis Primers 2015, 1, 15013, doi:10.1038/nrdp.2015.13.

5. van der Linden, S.; Valkenburg, H.A.; Cats, A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984, 27, 361-368, doi:10.1002/art.1780270401.

6. Rudwaleit, M.; van der Heijde, D.; Landewé, R.; Akkoc, N.; Brandt, J.; Chou, C.T.; Dougados, M.; Huang, F.; Gu, J.; Kirazli, Y.; et al. The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. Ann Rheum Dis 2011, 70, 25-31, doi:10.1136/ard.2010.133645.

7. Rudwaleit, M.; van der Heijde, D.; Landewé, R.; Listing, J.; Akkoc, N.; Brandt, J.; Braun, J.; Chou, C.T.; Collantes-Estevez, E.; Dougados, M.; et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis 2009, 68, 777-783, doi:10.1136/ard.2009.108233.

8. Stolwijk, C.; van Onna, M.; Boonen, A.; van Tubergen, A. Global Prevalence of Spondyloarthritis: A Systematic Review and Meta-Regression Analysis. Arthritis Care Res (Hoboken) 2016, 68, 1320-1331, doi:10.1002/acr.22831.

9. Dean, L.E.; Jones, G.T.; MacDonald, A.G.; Downham, C.; Sturrock, R.D.; Macfarlane, G.J. Global prevalence of ankylosing spondylitis. Rheumatology (Oxford) 2014, 53, 650-657, doi:10.1093/rheumatology/ket387.

10. Ward, M.M.; Deodhar, A.; Gensler, L.S.; Dubreuil, M.; Yu, D.; Khan, M.A.; Haroon, N.; Borenstein, D.; Wang, R.; Biehl, A.; et al. 2019 Update of the American College of Rheumatology/Spondylitis Association of America/Spondyloarthritis Research and

Treatment Network Recommendations for the Treatment of Ankylosing Spondylitis and Nonradiographic Axial Spondyloarthritis. Arthritis Rheumatol 2019, 71, 1599-1613, doi:10.1002/art.41042.

11. Ritchlin, C.; Adamopoulos, I.E. Axial spondyloarthritis: new advances in diagnosis and management. Bmj 2021, 372, m4447, doi:10.1136/bmj.m4447.

12. Zochling, J.; Brandt, J.; Braun, J. The current concept of spondyloarthritis with special emphasis on undifferentiated spondyloarthritis. Rheumatology (Oxford) 2005, 44, 1483-1491, doi:10.1093/rheumatology/kei047.

13. Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015, 43, e47, doi:10.1093/nar/gkv007.

14. Simon, N.; Friedman, J.; Hastie, T.; Tibshirani, R. Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent. J Stat Softw 2011, 39, 1-13, doi:10.18637/jss.v039.i05.

15. Sanz, H.; Valim, C.; Vegas, E.; Oller, J.M.; Reverter, F. SVM-RFE: selection and visualization of the most relevant features through non-linear kernels. BMC Bioinformatics 2018, 19, 432, doi:10.1186/s12859-018-2451-4.

16. Strobl, C.; Boulesteix, A.L.; Zeileis, A.; Hothorn, T. Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinformatics 2007, 8, 25, doi:10.1186/1471-2105-8-25.

17. Mauro, D.; Thomas, R.; Guggino, G.; Lories, R.; Brown, M.A.; Ciccia, F. Ankylosing spondylitis: an autoimmune or autoinflammatory disease? Nat Rev Rheumatol 2021, 17, 387-404, doi:10.1038/s41584-021-00625-y.

18. Yang, L.; Wang, L.; Wang, X.; Xian, C.J.; Lu, H. A Possible Role of Intestinal Microbiota in the Pathogenesis of Ankylosing Spondylitis. Int J Mol Sci 2016, 17, doi:10.3390/ijms17122126.

19. Mercan, R.; Bitik, B.; Tufan, A.; Bozbulut, U.B.; Atas, N.; Ozturk, M.A.; Haznedaroglu, S.; Goker, B. The Association Between Neutrophil/Lymphocyte Ratio and Disease Activity in Rheumatoid Arthritis and Ankylosing Spondylitis. J Clin Lab Anal 2016, 30, 597-601, doi:10.1002/jcla.21908.

20. Xu, S.; Ma, Y.; Wu, M.; Zhang, X.; Yang, J.; Deng, J.; Guan, S.; Gao, X.; Xu, S.; Shuai, Z.; et al. Neutrophil lymphocyte ratio in patients with ankylosing spondylitis: A systematic review and meta-analysis. Mod Rheumatol 2020, 30, 141-148, doi:10.1080/14397595.2018.1564165.

21. Gökmen, F.; Akbal, A.; Reşorlu, H.; Gökmen, E.; Güven, M.; Aras, A.B.; Erbağ, G.; Kömürcü, E.; Akbal, E.; Coşar, M. Neutrophil-Lymphocyte Ratio Connected to Treatment Options and Inflammation Markers of Ankylosing Spondylitis. J Clin Lab Anal 2015, 29, 294-298, doi:10.1002/jcla.21768.

22. Yang, C.; Ding, P.; Wang, Q.; Zhang, L.; Zhang, X.; Zhao, J.; Xu, E.; Wang, N.; Chen, J.; Yang, G.; et al. Inhibition of Complement Retards Ankylosing Spondylitis Progression. Sci Rep 2016, 6, 34643, doi:10.1038/srep34643.

23. Gonnet-Gracia, C.; Barnetche, T.; Richez, C.; Blanco, P.; Dehais, J.; Schaeverbeke, T. Anti-nuclear antibodies, anti-DNA and C4 complement evolution in rheumatoid arthritis and ankylosing spondylitis treated with TNF-alpha blockers. Clin Exp Rheumatol 2008, 26, 401-407.

24. Huang, Y.; Deng, W.; Zheng, S.; Feng, F.; Huang, Z.; Huang, Q.; Guo, X.; Huang, Z.; Huang, X.; Pan, X.; et al. Relationship between monocytes to lymphocytes ratio and axial spondyloarthritis. Int Immunopharmacol 2018, 57, 43-46, doi:10.1016/j.intimp.2018.02.008.

25. Wang, J.; Su, J.; Yuan, Y.; Jin, X.; Shen, B.; Lu, G. The role of lymphocyte-monocyte ratio on axial spondyloarthritis diagnosis and sacroiliitis staging. BMC Musculoskelet Disord 2021, 22, 86, doi:10.1186/s12891-021-03973-8.

26. Liang, T.; Chen, J.; Xu, G.; Zhang, Z.; Xue, J.; Zeng, H.; Jiang, J.; Chen, T.; Qin, Z.; Li, H.; et al. Platelet-to-Lymphocyte Ratio as an Independent Factor Was Associated With the Severity of Ankylosing Spondylitis. Front Immunol 2021, 12, 760214, doi:10.3389/fimmu.2021.760214.

27. Tamassia, N.; Arruda-Silva, F.; Calzetti, F.; Lonardi, S.; Gasperini, S.; Gardiman, E.; Bianchetto-Aguilera, F.; Gatta, L.B.; Girolomoni, G.; Mantovani, A.; et al. A Reappraisal on the Potential Ability of Human Neutrophils to Express and Produce IL-17 Family Members In Vitro: Failure to Reproducibly Detect It. Front Immunol 2018, 9, 795, doi:10.3389/fimmu.2018.00795.

28. Lin, A.M.; Rubin, C.J.; Khandpur, R.; Wang, J.Y.; Riblett, M.; Yalavarthi, S.; Villanueva, E.C.; Shah, P.; Kaplan, M.J.; Bruce, A.T. Mast cells and neutrophils release IL-17 through extracellular trap formation in psoriasis. J Immunol 2011, 187, 490-500, doi:10.4049/jimmunol.1100123.

29. Karlsson, M.; Zhang, C.; Méar, L.; Zhong, W.; Digre, A.; Katona, B.; Sjöstedt, E.; Butler, L.; Odeberg, J.; Dusart, P.; et al. A single-cell type transcriptomics map of human tissues. Sci Adv 2021, 7, doi:10.1126/sciadv.abh2169.

30. Zhu, Z.Q.; Tang, J.S.; Cao, X.J. Transcriptome network analysis reveals potential candidate genes for ankylosing spondylitis. Eur Rev Med Pharmacol Sci 2013, 17, 3178-3185.

31. Zheng, Y.; Cai, B.; Ren, C.; Xu, H.; Du, W.; Wu, Y.; Lin, F.; Zhang, H.; Quan, R. Identification of immune related cells and crucial genes in the peripheral blood of ankylosing spondylitis by integrated bioinformatics analysis. PeerJ 2021, 9, e12125, doi:10.7717/peerj.12125.

32. Zhang, Z.; Gothe, F.; Pennamen, P.; James, J.R.; McDonald, D.; Mata, C.P.; Modis, Y.; Alazami, A.M.; Acres, M.; Haller, W.; et al. Human interleukin-2 receptor β mutations associated with defects in immunity and peripheral tolerance. J Exp Med 2019, 216, 1311-1327, doi:10.1084/jem.20182304.

33. Ratthé, C.; Girard, D. Interleukin-15 enhances human neutrophil phagocytosis by a Syk-dependent mechanism: importance of the IL-15Ralpha chain. J Leukoc Biol 2004, 76, 162-168, doi:10.1189/jlb.0605298.

34. Waldmann, T.A. The biology of interleukin-2 and interleukin-15: implications for cancer therapy and vaccine design. Nat Rev Immunol 2006, 6, 595-601, doi:10.1038/nri1901.

35. Soper, D.M.; Kasprowicz, D.J.; Ziegler, S.F. IL-2Rbeta links IL-2R signaling with Foxp3 expression. Eur J Immunol 2007, 37, 1817-1826, doi:10.1002/eji.200737101.

36. Liao, W.; Lin, J.X.; Wang, L.; Li, P.; Leonard, W.J. Modulation of cytokine receptors by IL-2 broadly regulates differentiation into helper T cell lineages. Nat Immunol 2011, 12, 551-559, doi:10.1038/ni.2030.

37. Pol, J.G.; Caudana, P.; Paillet, J.; Piaggio, E.; Kroemer, G. Effects of interleukin-2 in immunostimulation and immunosuppression. J Exp Med 2020, 217, doi:10.1084/jem.20191247.

38. Allard-Chamard, H.; Mishra, H.K.; Nandi, M.; Mayhue, M.; Menendez, A.; Ilangumaran, S.; Ramanathan, S. Interleukin-15 in autoimmunity. Cytokine 2020, 136, 155258, doi:10.1016/j.cyto.2020.155258.

39. Campbell, T.M.; Bryceson, Y.T. IL2RB maintains immune harmony. J Exp Med 2019, 216, 1231-1233, doi:10.1084/jem.20190546.

40. Fernandez, I.Z.; Baxter, R.M.; Garcia-Perez, J.E.; Vendrame, E.; Ranganath, T.; Kong, D.S.; Lundquist, K.; Nguyen, T.; Ogolla, S.; Black, J.; et al. A novel human IL2RB mutation results in T and NK cell-driven immune dysregulation. J Exp Med 2019, 216, 1255-1267, doi:10.1084/jem.20182015.

41. Fujimura, K.; Oyamada, A.; Iwamoto, Y.; Yoshikai, Y.; Yamada, H. CD4 T cell-intrinsic IL-2 signaling differentially affects Th1 and Th17 development. J Leukoc Biol 2013, 94, 271-279, doi:10.1189/jlb.1112581.

42. Ohno, Y.; Kashio, A.; Ogata, R.; Ishitomi, A.; Yamazaki, Y.; Kihara, A. Analysis of substrate specificity of human DHHC protein acyltransferases using a yeast expression system. Mol Biol Cell 2012, 23, 4543-4551, doi:10.1091/mbc.E12-05-0336.

43. Shi, C.; Yang, X.; Liu, Y.; Li, H.; Chu, H.; Li, G.; Yin, H. ZDHHC18 negatively regulates cGAS-mediated innate immunity through palmitoylation. Embo j 2022, e109272, doi:10.15252/embj.2021109272.

44. Chen, X.; Hu, L.; Yang, H.; Ma, H.; Ye, K.; Zhao, C.; Zhao, Z.; Dai, H.; Wang, H.; Fang, Z. DHHC protein family targets different subsets of glioma stem cells in specific niches. J Exp Clin Cancer Res 2019, 38, 25, doi:10.1186/s13046-019-1033-2.

45. Pei, X.; Li, K.Y.; Shen, Y.; Li, J.T.; Lei, M.Z.; Fang, C.Y.; Lu, H.J.; Yang, H.J.; Wen, W.; Yin, M.; et al. Palmitoylation of MDH2 by ZDHHC18 activates mitochondrial respiration and accelerates ovarian cancer growth. Sci China Life Sci 2022, doi:10.1007/s11427-021-2048-2.

46. Zhao, Y.; He, A.; Zhu, F.; Ding, M.; Hao, J.; Fan, Q.; Li, P.; Liu, L.; Du, Y.; Liang, X.; et al. Integrating genome-wide association study and expression quantitative trait locus study identifies multiple genes and gene sets associated with schizophrenia. Prog Neuropsychopharmacol Biol Psychiatry 2018, 81, 50-54, doi:10.1016/j.pnpbp.2017.10.003.

47. Adachi, N.; Hess, D.T.; McLaughlin, P.; Stamler, J.S. S-Palmitoylation of a Novel Site in the β2-Adrenergic Receptor Associated with a Novel Intracellular Itinerary. J Biol Chem 2016, 291, 20232-20246, doi:10.1074/jbc.M116.725762.

48. Akimzhanov, A.M.; Boehning, D. Rapid and transient palmitoylation of the tyrosine kinase Lck mediates Fas signaling. Proc Natl Acad Sci U S A 2015, 112, 11876-11880, doi:10.1073/pnas.1509929112.

 doi:10.20944/preprints202206.0176.v1

49. Baumgart, F.; Corral-Escariz, M.; Pérez-Gil, J.; Rodríguez-Crespo, I. Palmitoylation of R-Ras by human DHHC19, a palmitoyl transferase with a CaaX box. Biochim Biophys Acta 2010, 1798, 592-604, doi:10.1016/j.bbamem.2010.01.002.

50. Yang, X.; Chatterjee, V.; Ma, Y.; Zheng, E.; Yuan, S.Y. Protein Palmitoylation in Leukocyte Signaling and Function. Front Cell Dev Biol 2020, 8, 600368, doi:10.3389/fcell.2020.600368.