



OPEN

Identification of ENTPD1 as a novel biomarker linking allergic rhinitis and systemic lupus erythematosus

Min Chen^{1,2}, Yingdi Meng^{1,2}, Xiaoqiong Shi¹, Chengjing Zhu¹, Minhui Zhu¹✉, Haihong Tang¹✉ & Hongliang Zheng¹✉

Several studies reveal that allergic rhinitis (AR) is a significant risk factor of systemic lupus erythematosus (SLE). However, studies investigating the common pathogenesis linking AR and SLE are lacking. Our study aims to search for the shared biomarkers and mechanisms that may provide new therapeutic targets for preventing AR from developing SLE. GSE50223 for AR and GSE103760 for SLE were downloaded from the Gene Expression Omnibus (GEO) database to screen differentially expressed genes (DEGs). The Gene Ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis were performed to explore the functions of shared DEGs. Hub genes were screened by cytoHubba (a plugin of Cytoscape) and validated in another two datasets. Gene set enrichment analysis (GSEA) and single-sample Gene set enrichment analysis (ssGSEA) algorithm were applied to understand the functions of hub gene. *ENTPD1* was validated as a hub gene between AR and SLE. GSEA results revealed that *ENTPD1* was associated with KRAS_SIGNALING_UP pathway in AR and related to HYPOXIA, TGF_BETA_SIGNALING and TNFA_SIGNALING_VIA_NFKB pathways in SLE. The expression of *ENTPD1* was positively correlated with activated CD8 T cell in both diseases. Thus, *ENTPD1* may be a novel therapeutic target for preventing AR from developing SLE.

Keywords *ENTPD1*, CD4⁺T cells, Allergic rhinitis, Systemic lupus erythematosus, Bioinformatics

Systemic lupus erythematosus (SLE) is a systemic autoimmune disease that is characterized by various clinical manifestations ranging from mild skin lesion to disastrous organ damage (e.g., ischemic heart disease, lupus nephritis and pulmonary embolism)¹. The prevalence of SLE greatly varies in worldwide regions, such as 84.8, 47, 37.6 and 3299.5 cases per 100,000 persons in San Francisco, France, Chian and Kenya, respectively^{2–5}. Moreover, SLE is the leading death reason among young women⁶. Thus, early detection, early diagnosis and effective treatment of SLE are warranted. Currently, the clinical treatments for SLE include corticosteroids, immunomodulators, cytotoxic-immunosuppressants, and biologics such as rituximab⁷. Nevertheless, there are some limitations to current therapies. For example, the consumption of B cells induced by rituximab may result in the absence of persistent response to rituximab. Thus, further investigations into new therapeutic strategies are necessitated.

Allergic rhinitis (AR) is a chronic nasal disorder mainly mediated by immunoglobulin E (IgE), whose mechanism involves various kinds of immune cells⁸. The symptoms of AR including nasal itching, nasal obstruction, sneezing and rhinorrhea have bothered 10–40% of the population globally⁹. Superficially, AR and SLE seem to be no correlation, but accumulated pieces of research studies suggest that AR is a significant risk factor of SLE^{10–13}. A meta-analysis study reviewed 1 cohort and 7 case–control studies, and indicated that individuals with AR had a 1.36-fold risk of SLE compared with individuals without AR¹³. These intriguing findings inspired us to investigate the potential common mechanisms between AR and SLE in search of novel biological targets.

CD4⁺T cells are mainly composed of type 1 T helper (Th1) cells, Th 2 cells, Th9 cells, Th17 cells, T follicular helper (Tfh) cells, and regulatory T (Treg) cells¹⁴. They play a vital role in the mechanisms of both AR and SLE. The dysregulation of Th1/Th2 cells, Th17/Treg cells imbalance and cytokines produced by Th2 and Th17 cells lead to occurrence and development of SLE¹⁵. Furthermore, AR is an allergic reaction mainly driven by Th2 cells⁸. Recent studies indicate that Tfh cells and Th17 cells are also involved in the mechanism of AR^{16,17}. Considering the important function of CD4⁺T cells in both diseases, we then focused on the shared pathogenesis of AR and SLE based on CD4⁺T cells.

¹Department of Otorhinolaryngology-Head and Neck Surgery, The First Affiliated Hospital of Naval Medical University, 168 Changhai Road, Yangpu District, Shanghai 200433, China. ²These authors contributed equally: Min Chen and Yingdi Meng. ✉email: zmh197915@163.com; maria5318@163.com; zhenghongliang1018@163.com

Rapid development of microarray technology has enabled scientific researchers to obtain the expression profiles of thousands of genes in a short time, contributing to a better understanding of disease mechanisms at the gene level. In this study, we hypothesized that AR and SLE had a shared pathogenesis and tested the assumption by analyzing genes related to CD4⁺T cells. Finally, we identified common pathways, shared hub gene, and potential chemicals targeting the hub gene, potentially providing a new insight into clinical therapies for both diseases.

Materials and methods

Collection of datasets

The array datasets were obtained from the Gene Expression Omnibus (GEO) database. With the aim to comprehensively acquire the datasets for AR and SLE regarding CD4⁺T cells, we first employed the search terms “((allergic rhinitis) and T cell) and Homo sapiens” and “((systemic lupus erythematosus) and T cell) and Homo sapiens” to search AR datasets and SLE datasets, respectively. Next, the following selection criteria were applied: (1) The datasets should be CD4⁺T cells expression profiling. (2) The datasets should include cases and controls. (3) The stimulated allergen to CD4⁺T cells in selected datasets of AR should be the same. Finally, two array datasets for AR (GSE50223 and GSE44960) and two array datasets (GSE103760 and GSE10325) were collected. In particular, both GSE50223 and GSE44960 were divided into four groups: patients challenged with diluent (PD) group, patients challenged with allergen (PA) group, healthy controls challenged with diluent (HD) group, and healthy controls challenged with allergen (HA) group. Additionally, GSE103760 and GSE10325 were composed of SLE patients and healthy controls.

Identification of shared DEGs between AR and SLE

The four datasets were first normalized by the function of “normalizeBetweenArrays” during the preprocessing procedure. Then, the “hgu133a.db” R package was used to convert gene probe ID into gene symbol name, and probes that do not correspond to the genes were removed. Next, the expression value of genes with several gene probes was averaged by the function of “aggregate”. Finally, the limma R package were utilized to search for the differentially expressed genes (DEGs) in GSE50223 and GSE103760. In order to identify DEGs more comprehensively, we set *P* value < 0.05 and FC (fold change) > 1.5 or < 2/3 to identify upregulated and downregulated DEGs, respectively^{18,19}. The R package “ggplot2” was applied for drawing the volcano plots of the two datasets. Then, the shared DEGs between AR and SLE were visualized with the use of the web tool Draw Venn Diagrams ([https:// bioinformatics.psb.ugent.be/webtools/Venn/](https://bioinformatics.psb.ugent.be/webtools/Venn/)).

GO and KEGG analyses

Gene Ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis^{20–22} were undertaken to better understand the functions of shared DEGs via the “clusterProfiler” R package. GO analysis contained three terms: biological process (BP), cellular component (CC), and molecular function (MF). These outcomes were presented as bubble plots through “ggplot2” R package.

PPI network establishment and candidate hub genes selection

The protein–protein interaction (PPI) network was established based on STRING database (<https://cn.string-db.org/>) and visualized by Cytoscape software. The potential hub genes were selected by cytoHubba, a plugin of Cytoscape, which contains local-based scoring methods and global-based scoring methods for exploring key nodes of PPI. Due to the heterogeneity of topological features analyzed by different scoring methods, we utilized two local-based methods (MCC and Degree) and two global-based methods (EcCentricity and Bottle-Neck) in cytoHubba to search for candidate hub genes²³. The top 5 genes ranked by four methods were chosen and these genes were intersected based on Draw Venn Diagrams tool ([https:// bioinformatics.psb.ugent.be/webtools/Venn/](https://bioinformatics.psb.ugent.be/webtools/Venn/)).

Validation and ROC curves of hub gene in two independent datasets

The mRNA expressions of candidate hub genes above were verified in GSE44960 for AR and GSE10325 for SLE. The R package “limma” was performed to determine whether these genes were differentially expressed in datasets. Additionally, the Comparative Toxicogenomics Database (CTD, [http:// ctdbase.org/](http://ctdbase.org/)) includes information concerning gene–disease, chemical–disease, and chemical–gene/protein interactions²⁴. Data from CTD was utilized to validate the relationships between hub gene and diseases.

In order to evaluate the diagnostic efficacy of hub gene, we drew receiver operating characteristic (ROC) curves to calculate the area under curve (AUC) and the cutoff value of GSE44960 and GSE10325.

Study population

The participants were recruited from the Changhai Hospital. These participants were divided into four groups: healthy controls (n = 8), AR without SLE (n = 9), SLE without AR (n = 8), and SLE with AR (n = 5). The diagnosis of AR was based on Allergic Rhinitis and its Impact on Asthma (ARIA) guidelines⁸, as follows: (1) having at least one symptom of sneezing, rhinorrhea, nasal itching, and nasal obstruction; (2) having the signs of pale turbinate and nasal mucosa, and nasal secreta. (3) having at least one positive allergen by serum immunoglobulin E test. The diagnosis of SLE was based on the guideline²⁵, including evaluation of seven clinical domains (constitutional condition, haematological condition, neuropsychiatric condition, mucocutaneous condition, serosal condition, musculoskeletal condition and renal condition) and three immunological domains (antiphospholipid antibodies, low complements, anti-Smith and anti-dsDNA). Every domain had the corresponding score and the total

score was added up by the scores of domains. The diagnosis of SLE was made if the antinuclear antibodies were positive and total score was ≥ 10 .

For the AR without SLE group, they met the following criteria: (1) being diagnosed as AR; (2) no autoimmune diseases and other atopic diseases except AR; (3) no mental and malignant disorders. For the SLE without AR group, the inclusion criteria were as follows: (1) being diagnosed as SLE; (2) no atopic diseases and other autoimmune diseases except SLE; (3) no mental and malignant diseases. For the SLE with AR group, they matched the following criteria: (1) being diagnosed as AR and SLE; (2) no other autoimmune diseases and other atopic diseases except SLE and AR; (3) no mental and malignant disorders. The inclusion criteria of healthy controls were as follows: (1) no atopic diseases; (2) no autoimmune diseases; (3) no mental and malignant diseases. The clinical features of participants were summarized in Table S1.

Validation of *ENTPD1* expression by quantitative real-time PCR analysis and flow cytometry

Blood samples were collected from patients and controls to verify the hub gene expression. There were four groups in our study: healthy controls ($n = 8$), AR without SLE ($n = 9$), SLE without AR ($n = 8$), and SLE with AR ($n = 5$). The CD4⁺ T cells were purified from peripheral blood mononuclear cells (PBMCs) by using CytoSinct™ CD4 Nanobeads (GeneScript, L00863-1) and a magnetic-activated cell separation system. The total RNA was extracted from CD4⁺ T cells according to the manufacturer's instructions (Vazyme, R711). Next, the Evo M-MLV RT Kit (Accurate Biology, AG11705) was utilized to transcribe mRNA into cDNA. Finally, qRT-PCR analysis of cDNA by SYBR Green Premix Pro HS qPCR Kit (Accurate Biology, AG11701) was performed. *ENTPD1* gene primers: 5'-AGCAGCTGAAATATGCTGGC-3'(forward), 5'-GAGACAGTATCTGCCGAAGTCC-3' (reverse). *GAPDH* gene primers: 5'-CATGAGAAGTATGACAACAGCCT-3' (forward), 5'-AGTCCTTCCACGATACCA AAGT-3' (reverse). *GAPDH* was considered as an internal reference gene. The gene relative expression was calculated by the $2^{-\Delta\Delta CT}$ method.

Flow cytometry was used to determine the percentage of CD11b⁺CD4⁺CD39⁺T cells. Firstly, PBMCs were isolated from fresh peripheral blood. Subsequently, after the incubation of FITC anti-human CD11b (Biolegend, 982,614), CD4 (Biolegend, 980,802) and FITC anti-human CD39 antibody (Biolegend, 328,205), PBMCs was analyzed by flow cytometry using a FACSCalibur flow cytometer (BD Biosciences). CD11b⁺CD4⁺ CD39⁺T cells percentage was analyzed using FlowJo software. FITC anti-human CD11b was used to exclude the interference induced by myeloid cells.

Function analyzed by GeneMANIA

GeneMANIA is an online tool used for predicting interaction networks and functions of specific genes, which can provide information physical interactions, co-expression, colocalization, genetic interactions, pathway, and shared protein domains³⁶. Hence, we applied GeneMANIA to reveal the function of *ENTPD1*.

Gene set enrichment analysis and immune infiltration analysis

Gene set enrichment analysis (GSEA) was utilized to investigate the pathways that *ENTPD1* was involved. In our study, we first divided the PA in GSE50223 into high- and low-expression groups based on *ENTPD1* median expression. Similarly, SLE patients in GSE103760 were divided into groups with high and low expression of *ENTPD1*. Secondly, the gene sets of "H: hallmark gene sets" were downloaded from the Molecular Signatures Database (<http://www.broad.mit.edu/gsea/msigdb/index.jsp>) as reference sets. Finally, we performed GSEA between high and low expression groups using the R package "clusterProfiler". The GSEA results included enrichment score (ES), normalized enrichment score (NES), *P* value, and false discovery rate (FDR). Enriched pathways with $|\text{NES}| > 1$, *P* value < 0.05 , and FDR < 0.05 were defined significant.

Single-sample Gene set enrichment analysis (ssGSEA) algorithm, an extension method of GSEA, is used to evaluate the immune cell infiltration in each sample using gene expression profile²⁷. In our study, the ssGSEA algorithm was used to assess the immune infiltration in GSE50223 for AR and GSE103760 for SLE, respectively. Spearman's correlation analysis was performed to evaluate the associations between immune cells and *ENTPD1* and to calculate the correlation indices.

Identification of putative chemical compounds

As previously mentioned, the CTD contains data regarding chemical-gene interactions. The chemical-*ENTPD1* interactions was investigated in the CTD.

Ethical approval and consent to participate

This study was approved by Changhai Hospital Ethics Committee (CHEC2023-264). All participants signed the informed consents. All methods were carried out in accordance with relevant guidelines and regulations.

Statistical analysis

In this research study, statistical analyses were implemented in R software (version 4.2.2). The correlations between the expression of *ENTPD1* and immune cell were analyzed by using the Spearman method. Continuous variables were presented as mean \pm standard deviation. Unless otherwise stated, a *P* value < 0.05 was considered statistically significant.

Results

The shared DEGs between AR and SLE

By using limma R package, we identified 174 DEGs between PA and HA groups in GSE50223, including 123 downregulated DEGs and 51 upregulated DEGs (Fig. 1a). However, only 8 genes differentially expressed between PD and HD groups (Table S2), which might be because that DEGs were generated by CD4⁺T cells mainly after exposure to the allergen. Therefore, DEGs between the PA and HA groups were used for further study. Meanwhile, as shown in Fig. 1b, 464 DEGs containing 9 downregulated and 455 upregulated genes were screened between SLE patients and healthy controls in GSE103760. Subsequently, a comparative analysis was applied to confirm the shared DEGs between AR and SLE. Finally, a total of 11 shared genes (*CYBB*, *ENTPD1*, *FCER1G*, *FGL2*, *HMMR*, *IFI30*, *IFNG*, *LY96*, *MCOLN2*, *RNASE6*, *TYMS*) were determined for the further analyses (Fig. 1c).

Functional enrichment analyses of shared DEGs

To further understand the functional overview of the shared DEGs, GO analysis and KEGG pathway enrichment were utilized. The top 10 GO terms of BP, CC, and MF were depicted in Fig. 2a and Table S3. In terms of BP, the shared DEGs were mainly enriched in positive regulation of cytokine production, tumor necrosis factor production, antigen processing and presentation, and T cell activation and differentiation. In terms of CC, the shared DEGs were mostly distributed in secretory granule membrane, ficolin-1-rich granule, tertiary granule,

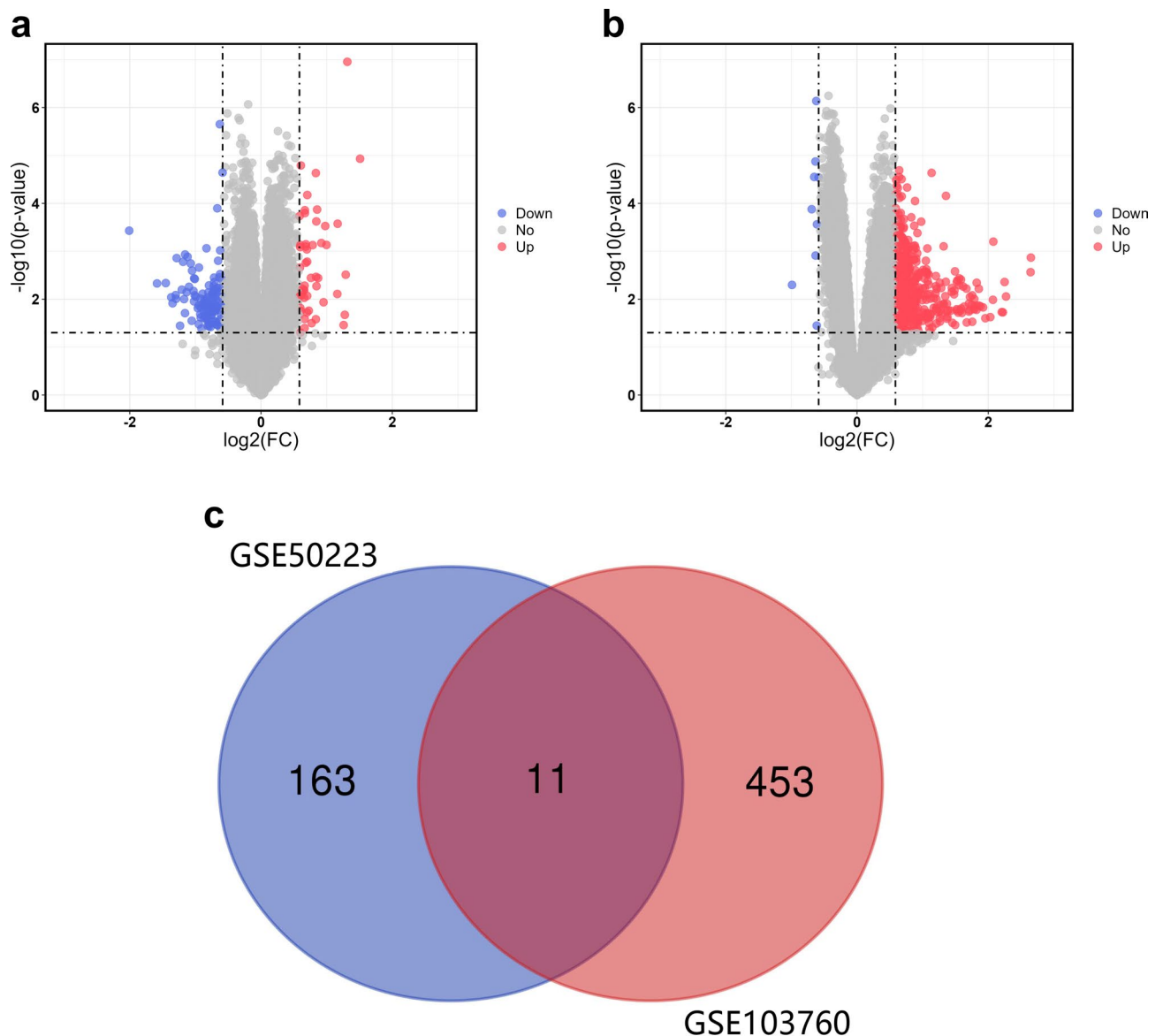


Figure 1. Identification of shared DEGs between AR and SLE. (a) The volcano plot illustrates 174 DEGs between the AR patients and healthy controls. (b) The volcano plot illustrates 464 DEGs between the SLE patients and healthy controls. (c) The Venn diagram shows a total of 11 shared DEGs between AR and SLE. *DEG* differentially expressed gene, *AR* allergic rhinitis, *SLE* systemic lupus erythematosus.

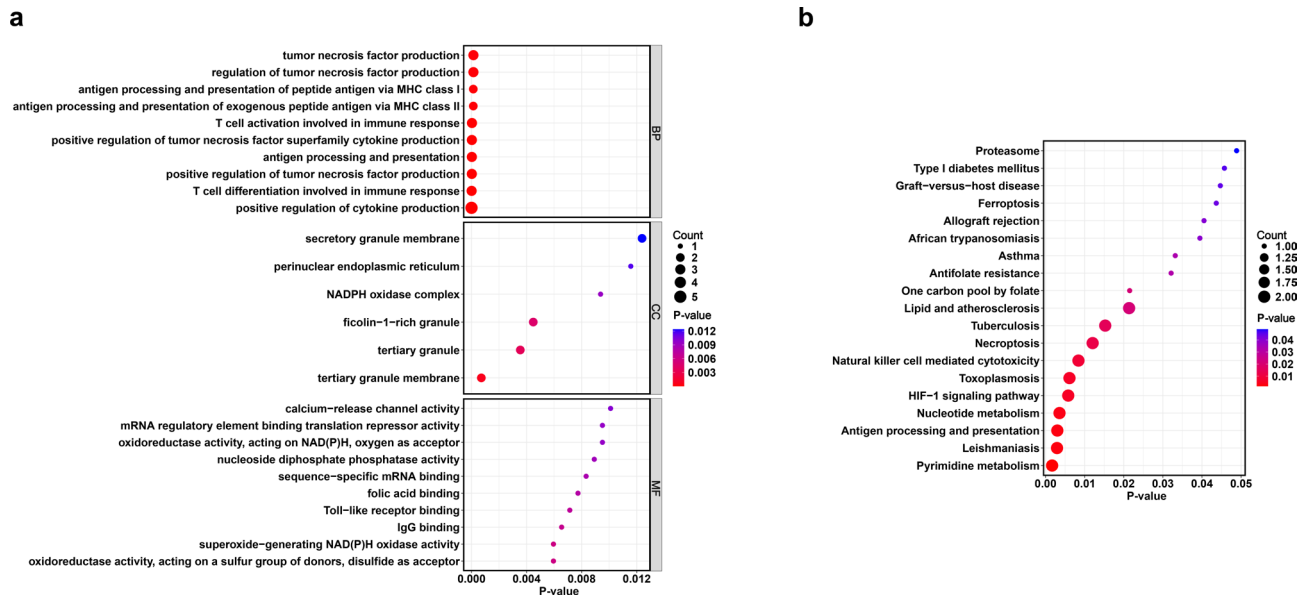


Figure 2. Functional enrichment analyses of the shared DEGs. **(a)** GO terms including BP, CC, and MF were used for functional enrichment clustering analysis on shared DEGs. **(b)** KEGG pathway enrichment analyses were performed on the shared DEGs^{20–22}. *DEG* differentially expressed gene, *GO* gene ontology, *BP* biological process, *CC* cellular component, *MF* molecular function, *KEGG* Kyoto Encyclopedia of Genes and Genomes.

and tertiary granule membrane. Besides, these DEGs involved in MF terms of superoxide-generating NAD(P)H oxidase activity, oxidoreductase activity, IgG binding, Toll-like receptor binding, folic acid binding, sequence-specific mRNA binding, nucleoside diphosphate phosphatase activity, mRNA regulatory element binding translation repressor activity, and calcium-release channel activity. Furthermore, KEGG analysis showed that shared DEGs were highly associated with antigen processing and presentation, HIF-1 signaling pathway, natural killer cell mediated cytotoxicity, and necroptosis (Fig. 2b and Table S4).

PPI network construction and selection of candidate hub genes

Using the STRING database, we established a PPI network to investigate the potential associations of proteins encoded by shared DEGs. As shown in Fig. 3a, the PPI network includes 10 nodes and 13 edges, with a PPI enrichment *P* value of 9.99E-10. Next, the PPI network was imported into Cytoscape software and analyzed in

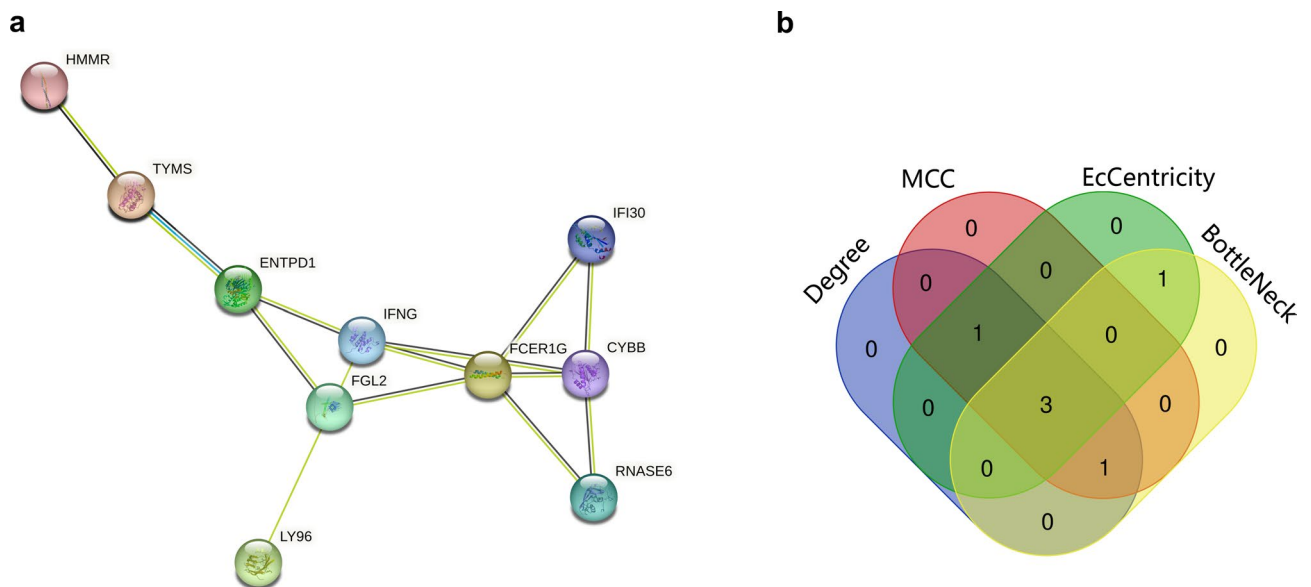


Figure 3. PPI network and identification of candidate hub genes. **(a)** PPI network of the shared DEGs constructed by STRING. **(b)** Identification of 3 candidates for hub genes by four ranked methods in Cytoscape plugin cytoHubba. *PPI* protein–protein interaction, *DEG* differentially expressed gene, *STRING* Search Tool for the Retrieval of Interacting Genes.

cytoHubba. Finally, the intersection of top 5 genes ranked by four methods (MCC, Degree, EcCentricity and BottleNeck) confirmed 3 candidate hub genes: *IFNG*, *TYMS*, and *ENTPD1* (Fig. 3b and Table 1). Additionally, the expression level of these three genes were higher in both AR and SLE patients than those in corresponding controls (Figure S1).

Validation and efficacy evaluation of hub gene in two independent datasets

The expression of candidate hub genes obtained above were analyzed in GSE44960 for AR and GSE10325 for SLE to validate our findings. Notably, the expression of *ENTPD1* gene was significantly upregulated in both AR and SLE. The expression level of *IFNG* was higher in SLE; however, such expression difference was not observed in AR. In addition, no significance of *TYMS* expression was found in both AR and SLE (Fig. 4). Furthermore, *ENTPD1* was found to be correlated with AR and SLE in CTD (Table S5). Hence, *ENTPD1* was confirmed as the hub gene of AR and SLE.

ROC curves were drawn to assess the diagnostic efficacy of *ENTPD1*. ROC curve analysis revealed that the AUC was 0.74 for *ENTPD1* in prediction of AR, with an optimal cutoff point of 4.532 (sensitivity = 70% and specificity = 90%) (Fig. S2a). Furthermore, the AUC was 0.76 for *ENTPD1* in prediction of SLE, with an optimal cutoff point of 6.531 (sensitivity = 100% and specificity = 57%) (Fig. S2b). These results suggested that *ENTPD1* might be a promising marker for diagnosing AR and SLE.

Category	Ranked methods in cytoHubba			
Gene symbol	MCC	Degree	EcCentricity	BottleNeck
	<i>FCER1G</i>	<i>FCER1G</i>	<i>IFNG</i>	<i>FGL2</i>
	<i>CYBB</i>	<i>CYBB</i>	<i>ENTPD1</i>	<i>ENTPD1</i>
	<i>IFNG</i>	<i>IFNG</i>	<i>FGL2</i>	<i>FCER1G</i>
	<i>ENTPD1</i>	<i>ENTPD1</i>	<i>TYMS</i>	<i>IFNG</i>
	<i>TYMS</i>	<i>TYMS</i>	<i>CYBB</i>	<i>TYMS</i>

Table 1. Identification of candidate hub genes in shared DEGs ranked by four methods of cytoHubba. Bold gene symbols were the top 5 shared hub genes ranked by four methods in cytoHubba. *DEG* differentially expressed gene; *MCC* Maximal clique centrality.

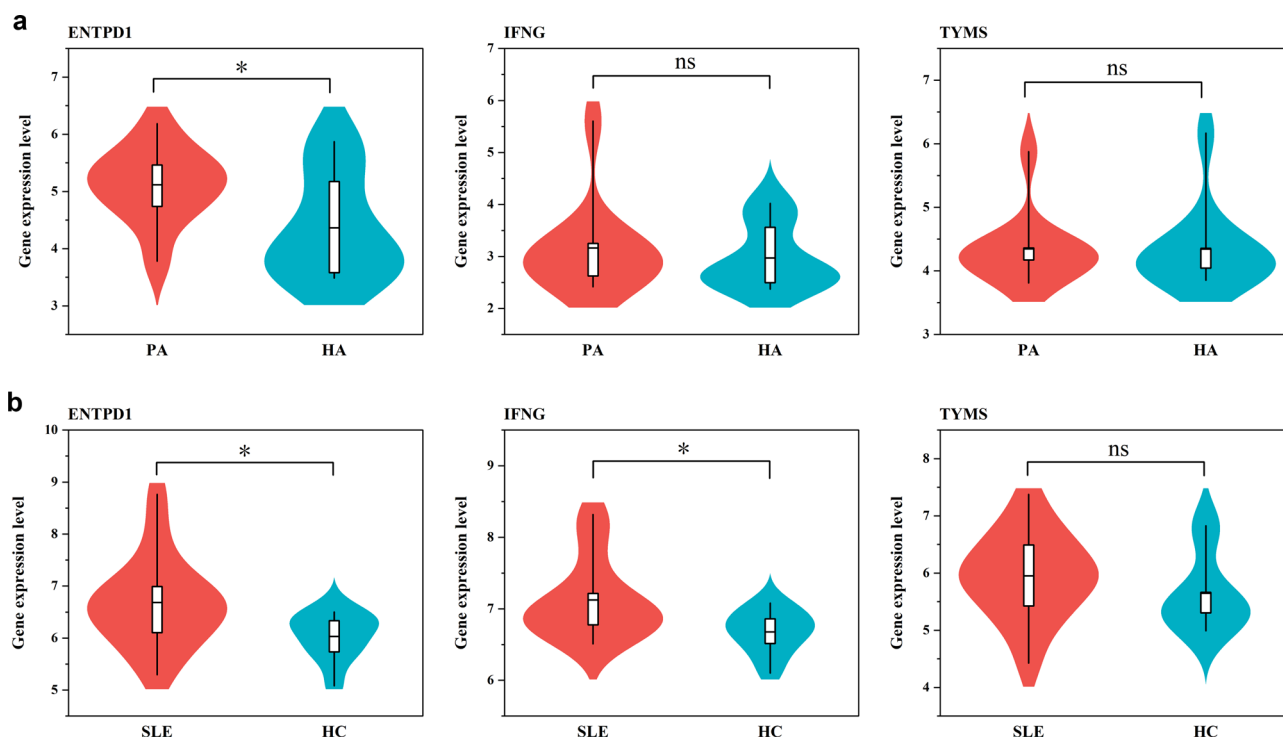


Figure 4. Validation of candidate hub genes. (a) The violin plots show the expression levels of candidate hub genes in GSE44960 for AR. (b) The violin plots show the expression levels of candidate hub genes in GSE10325 for SLE. *AR* allergic rhinitis, *SLE* systemic lupus erythematosus.

Validation of *ENTPD1* expression by qRT-PCR and flow cytometry

The clinical characteristics of individuals included in this study were summarized in Table S1. As shown in Fig. 5a, the expression level of *ENTPD1* is higher in AR, SLE and SLE with AR than that in healthy control. Additionally, the *ENTPD1* expression was the highest in SLE with AR group among 4 groups. Consistently, the percentage of CD11b⁺CD4⁺CD39⁺T cells in PBMCs was significantly lower in healthy control than that in other 3 groups (healthy control, 0.93 ± 0.37%; AR, 2.36 ± 0.97%; SLE, 2.69 ± 0.86%; SLE with AR, 8.46 ± 1.87%). The SLE with AR group had the highest percentage of CD11b⁺CD4⁺CD39⁺T cells (Fig. 5b,c).

Functional annotations of *ENTPD1* analyzed by GeneMANIA

After confirming *ENTPD1* as the hub gene, we next investigated the potential functions of *ENTPD1* via GeneMANIA (Fig. 6). The hub gene is located in the central circle, which is surrounded by predicted related genes. A sum of 20 targeted genes were predicted to construct the networks with physical interactions of 77.64%, co-expression of 8.01%, colocalization of 3.63%, genetic interactions of 2.87%, pathway of 1.88%, and shared protein domains of 0.6%. As shown in Fig. 6, the networks implicate these genes are mainly enriched in pyrimidine-containing compound catabolic, nucleoside diphosphate metabolic, organophosphate catabolic, nucleoside phosphate catabolic and nucleobase-containing small molecule biosynthetic process, and nucleoside-diphosphatase activity.

GSEA results of *ENTPD1*

To uncover the effect of *ENTPD1* on AR and SLE, GSEA was utilized to explore the significantly different pathways involved in these two diseases between the low and high *ENTPD1* expression groups. GSEA result reveals that *ENTPD1* is significantly associated with KRAS_SIGNALING_UP pathway in AR (Fig. S3a). In addition, *ENTPD1* is strongly related to HYPOXIA, TGF_BETA_SIGNALING, and TNFA_SIGNALING_VIA_NFKB pathways in SLE (Fig. S3b). The *P*-values, ES, NES and FDR for GSE50223 and GSE103760 were listed in Table S6.

Immune infiltration analyses of *ENTPD1*

We obtained the associations between the expression of *ENTPD1* and immune cell infiltration in AR and SLE via the ssGSEA algorithm and spearman's correlation analysis. The Fig. S4 shows the activated CD4 T cell and effector memory CD4 T cell are more highly expressed in AR and SLE patients compared with controls. As revealed by

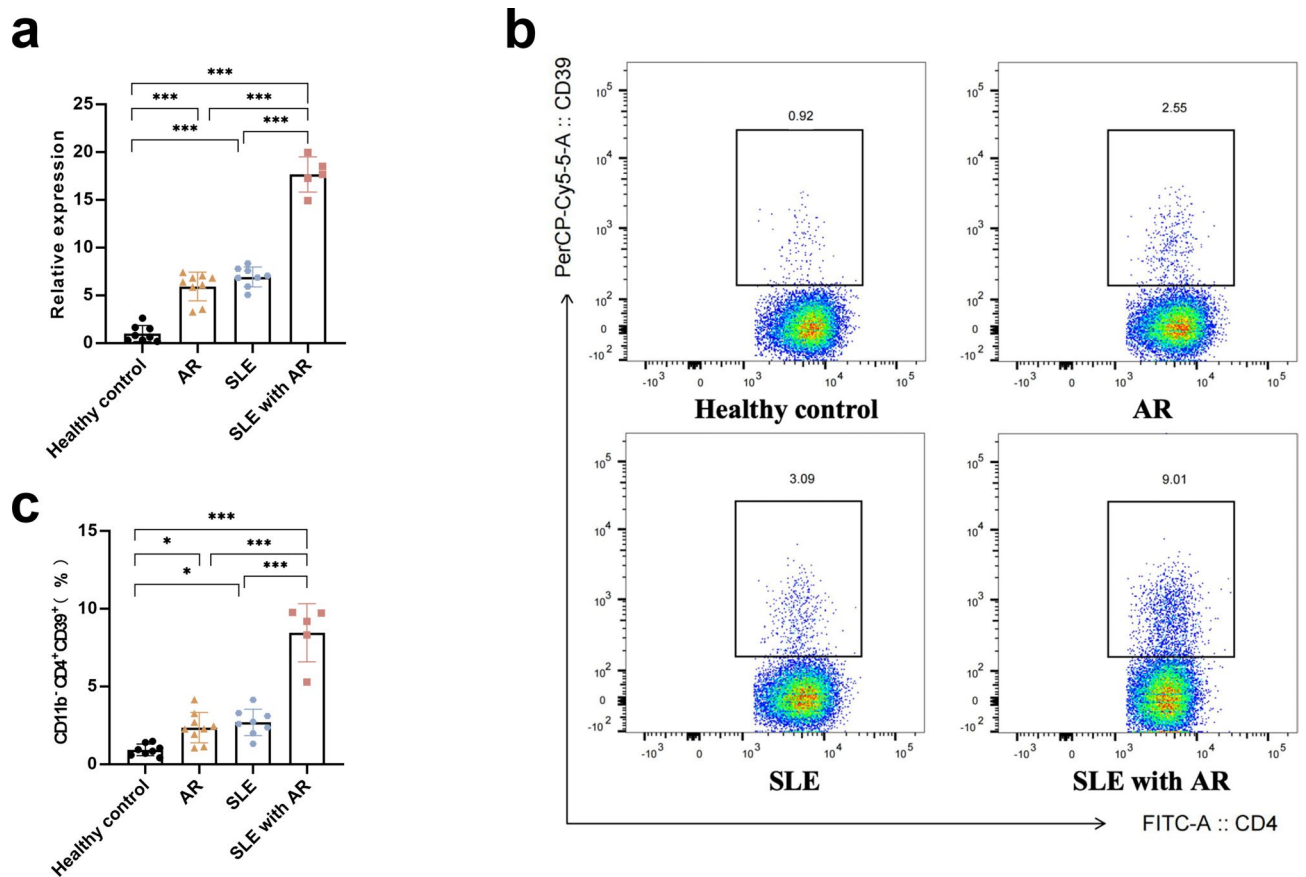


Figure 5. Validation of *ENTPD1* expression. (a) The relative mRNA expression of *ENTPD1* in four groups: healthy control, AR, SLE, and SLE with AR groups. (b) (c) The percentage of CD11b⁺CD4⁺CD39⁺T cells in PBMCs of four groups: healthy control, AR, SLE, and SLE with AR groups. **P* < 0.05, ****P* < 0.001. AR allergic rhinitis, SLE systemic lupus erythematosus, PBMC peripheral blood mononuclear cell.

11

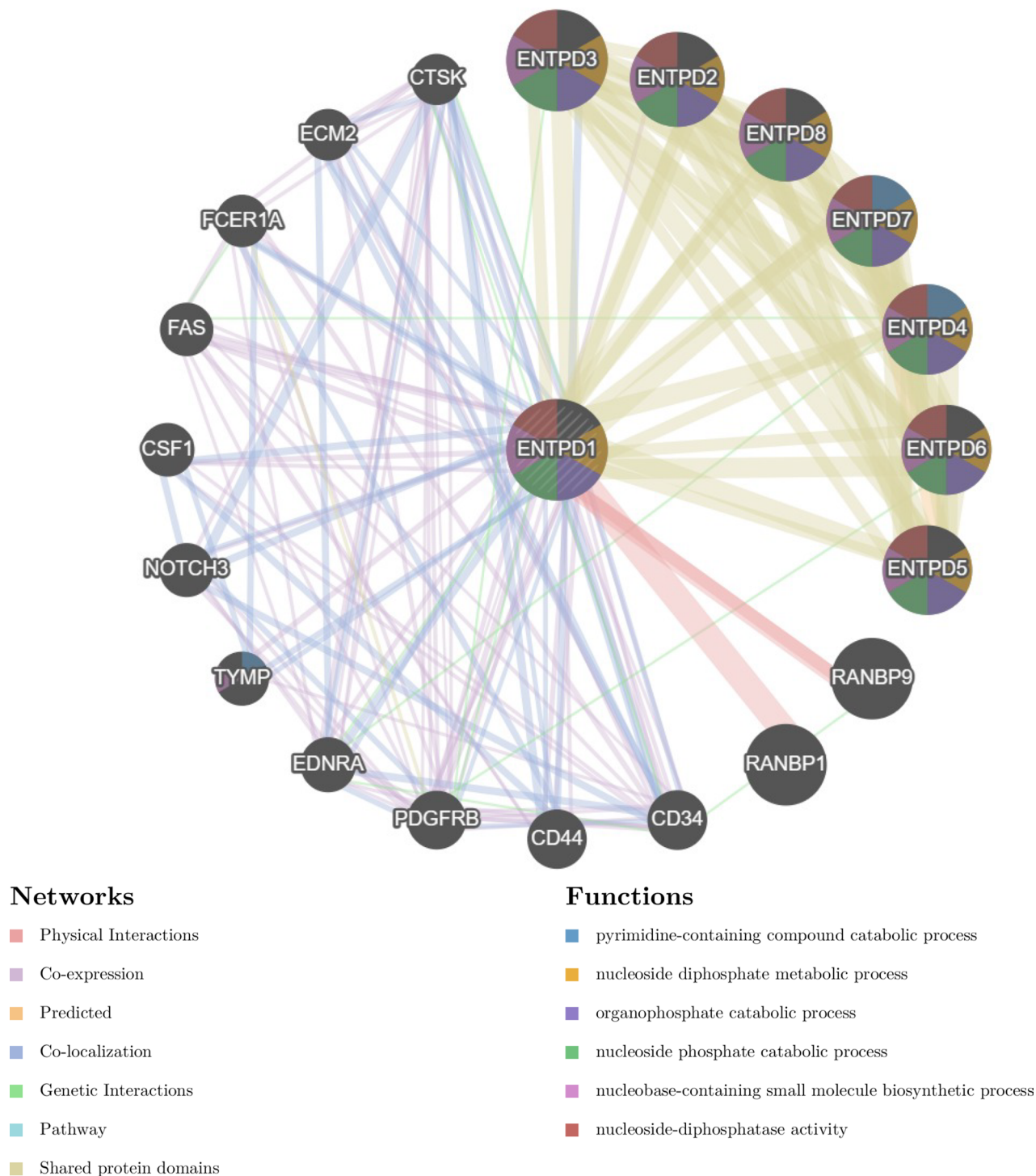


Figure 6. The gene–gene interaction network and functions for *ENTPD1* based on the GeneMANIA database.

Fig. 7a, the expression of *ENTPD1* is positively related to activated CD8 T cell and eosinophil, and is negatively linked with regulatory T cell, macrophage, gamma delta T cell and activated dendritic cell in AR. Moreover, immune infiltration analysis was also performed in SLE, which indicated that *ENTPD1* expression was positively associated with type 2 T helper cell, gamma delta T cell, memory B cell, regulatory T cell, activated CD4 T cell, effector memory CD8 T cell, plasmacytoid dendritic cell, activated CD8 T cell, and activated B cell (Fig. 7b). Clearly, the expression of *ENTPD1* was positively correlated with activated CD8 T cell in both AR and SLE.

Candidate chemical compounds targeting at *ENTPD1* in the CTD

Since *ENTPD1* might play a key role in the pathogenesis of AR and SLE, we next investigated to search for the potential compounds that regulated the expression of *ENTPD1*. Table S7 describes candidate chemical

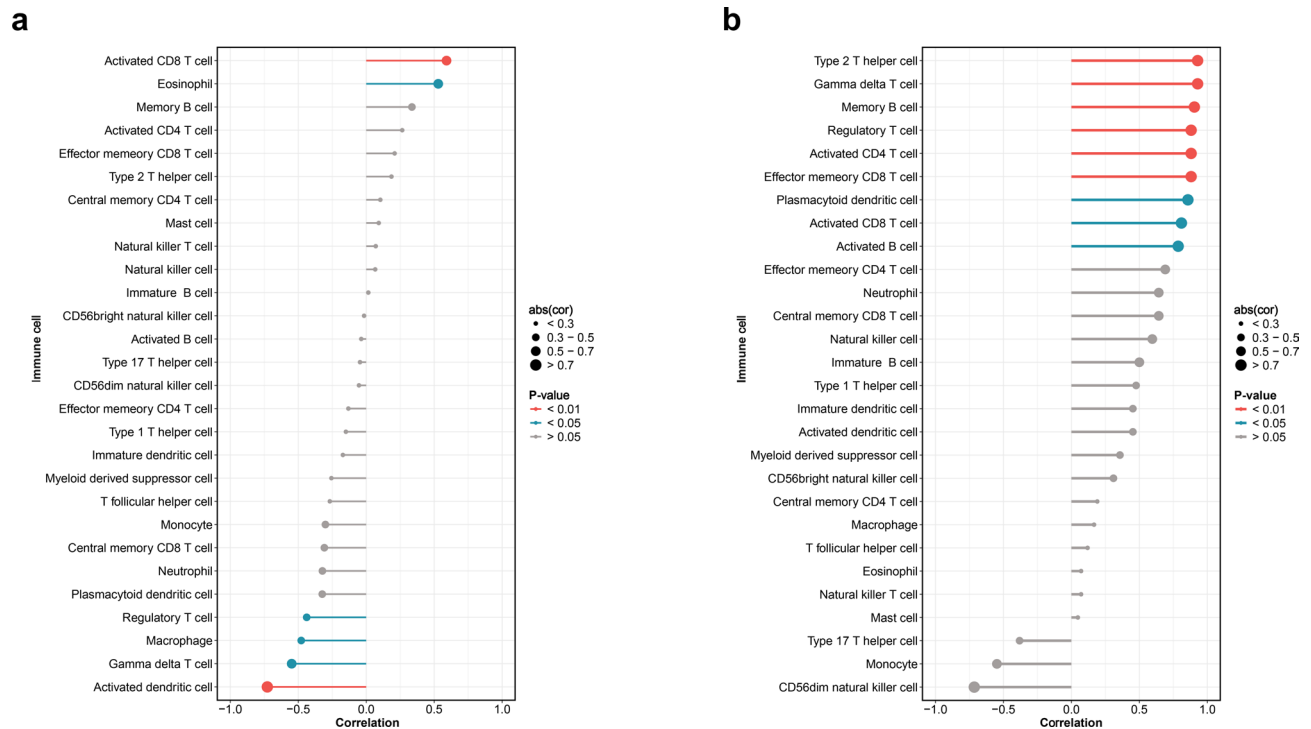


Figure 7. Distribution of immune cell infiltration in AR and SLE. The lollipop plots reveal the relationships between *ENTPD1* expression level and immune cell subtypes in AR (a) and SLE patients (b). AR allergic rhinitis, SLE systemic lupus erythematosus.

compounds which could regulate *ENTPD1* mRNA expression. There were 10 chemical compounds decreasing *ENTPD1* expression and 13 chemical compounds increasing *ENTPD1* expression.

Discussion

SLE usually causes catastrophic organ damage and is the primary reason of young female⁶. However, current therapies have no benefit in decreasing the mortality²⁸. Recent research has revealed that the prevalence of SLE was significantly higher in AR patients compared with the general population¹³. Effective AR therapy may prevent or delay the onset of SLE. Hence, identifying the common pathogenesis between AR and SLE is vital for facilitating treatment and prevention strategies. CD4⁺T cells are involved in not only the occurrence of AR but also the onset of SLE. Nevertheless, no study has focused on exploring the common pathogenesis of AR and SLE based on gene expression profiling of CD4⁺T cells. In our study, we identified shared DEGs and pathways by analyzing transcriptomic data of CD4⁺T cells in both AR and SLE. Further analyses indicated that *ENTPD1* played a crucial role in AR and SLE, and was a potential biomarker of these two diseases.

In this work, we found 11 shared DEGs between AR and SLE based on the GEO database, and performed GO and KEGG analyses on these shared DEGs. GO analysis of the shared DEGs showed that these genes were mainly enriched in positive regulation of cytokine production, tumor necrosis factor production, antigen processing and presentation, and T cell activation and differentiation in terms of BP. These results suggested that both AR and SLE were closely associated with immune response and inflammation, which were consistent with previous studies^{15,17}. Notably, in terms of MF, the shared DEGs were related to Toll-like receptor binding. Several studies have suggested that Toll-like receptor played a crucial role in the onset of AR and SLE by mediating the production of inflammatory cytokine²⁹ or the loss of B cell tolerance to autoantigens³⁰. A small number of shared DEGs involved in GO analysis might lead to unstable results, but the results might give us some idea about the functions of these genes. KEGG analysis also revealed that shared DEGs were involved in pathways regarding immune responses. Additionally, KEGG analysis indicated that shared DEGs were associated with HIF-1 signaling pathway and necroptosis, implying that these pathways might be crucial to the common pathogenesis of AR and SLE.

Using bioinformatics analysis, we identified that *ENTPD1* was the shared hub gene between AR and SLE. *ENTPD1*, ectonucleoside triphosphate diphosphohydrolase 1, encodes an ectoenzyme CD39 which hydrolyzes extracellular ATP (eATP) to AMP. CD39 is able to convert the eATP-induced pro-inflammatory responses into anti-inflammatory status in immune actions³¹. Notably, ATP level elevated in SLE patients, activating inflammasome and releasing cytokines related with disease consequently. As previous studies described, CD39 regulated the immune and inflammatory response in SLE via catalyzing ATP into adenosine, a potent immune regulator^{32,33}. And Becker et al. reported that the expression of CD39 increased in SLE compared with healthy controls and suggested that the increased CD39 expression could help control inflammation as a compensatory mechanism³³. Moreover, Treg-mediated suppression was absolutely absent in SLE patients with CD39 defect compared with controls and SLE without CD39 defect³⁴. Further experimental studies showed that deficiency of CD39 led to clinically overt pulmonary hemorrhage, potentiated splenomegaly and autoantibody production, and accelerated

the release of neutrophil extracellular traps by neutrophils in SLE model compared with corresponding controls³⁵. In lupus nephritis mice, Zheng et al. revealed that CD39 expression in iTreg cells was crucial to induce immune tolerance, and modulated the proliferation and differentiation of T cells, thus inhibiting inflammation. Besides, CD39 was effective in treating lupus nephritis and the inhibition of CD39 expression led to the failure of amelioration in lupus nephritis^{36,37}. Based on these, it could be concluded that *ENTPD1*/CD39 mediated its beneficial effects by suppressing inflammation in SLE. Previous study showed that genetic variant in *ENTPD1* was related with an increased risk of AR³⁸. The Treg cells had an effect on suppressing immune responses by increasing CD39 expression after intranasal allergen immunotherapy³⁹, which indirectly suggested the anti-inflammatory function of CD39. Although few studies have conducted in-depth exploration regarding the role of *ENTPD1*/CD39 in AR, the role of *ENTPD1*/CD39 in allergic asthma was relatively widely explored. CD39 was found to be associated with reduced airway inflammation and hyperreactivity in allergic asthma⁴⁰. Overexpression of *ENTPD1* could also mitigate airway eosinophilia, mucin deposition, and Th2 cytokines production⁴¹, while *ENTPD1* knockout in allergic asthma mice aggravated airway inflammation through increasing the cytokines level and eosinophil recruitment⁴². Likewise, allergic asthma model administered with apyrase, which has ectoenzymatic activity identical to CD39, was observed to have attenuated peribronchial eosinophilic infiltration and decreased levels of Th2 cytokines in the bronchoalveolar lavage fluid⁴³. These results indicated that *ENTPD1*/CD39 might participate in the pathogenesis of allergic asthma. Moreover, Lu et al. reported that Treg cells transfected by *ENTPD1* could secrete higher levels of IL10 and TGF- β , which are essential in modulating immune system. Furthermore, up-regulation of *ENTPD1* enhanced the immunosuppressive function and FOXP3 expression of Treg cells, contributing to the maintenance of immune homeostasis^{44,45}. It is well known that Treg cells play a leading role in the mechanism of AR¹⁷. Given that the concept of “one-airway-one-disease”⁴⁶, we speculated that *ENTPD1*/CD39 might be involved in the mechanism of AR via regulating the levels of mucin and Th2 cytokines, and eosinophil recruitment. In summary, the above studies provided compelling evidence for the crucial role played by *ENTPD1* in AR and SLE.

As described previously, GeneMANIA can predict interaction networks of *ENTPD1*²⁶. GeneMANIA results showed that 20 genes were related to *ENTPD1*, which might provide a new insight into investigating how *ENTPD1* participate in the mechanism of AR and SLE. Although CD4⁺T cells are pivotal in AR and SLE, the role of CD8⁺T cells in these two diseases could not be underestimated. CD8⁺T cell, a subset of T lymphocytes, was reported to reduce allergen-induced airway eosinophilic infiltration by secreting IFN- γ ⁴⁷. Moreover, CD8⁺T cells might mitigate inflammation in AR condition⁴⁸. Li et al. found that the frequency of CD8⁺T cells in peripheral blood was higher in SLE patients than in controls, and suggested that CD8⁺T cells might be of importance in maintaining peripheral tolerance and suppressing self-reactive⁴⁹. In addition, CD8⁺T cells were observed to expand and could suppress pathogenic CD4⁺T cells in multiple sclerosis, an autoimmune disease⁵⁰. According to our result that *ENTPD1* was positively associated with activated CD8⁺T cells in both AR and SLE, it can be deduced that *ENTPD1* might prevent the development of AR and SLE by activating CD8⁺T cells. Notably, the expression of *ENTPD1* was not correlated with the activated CD4⁺T cells. The contradictory result might due to the relatively small size of study population. Further immune infiltration analysis with a larger sample size is necessary. Additionally, the immune cell infiltration patterns of these two diseases were not similar. Previous study reported that other factors such as epigenetics and gut microbiota were involved in the mechanism of SLE⁵¹, which might contribute to the difference of immune patterns.

The importance of KRAS pathway to the development of human tumors, especially pancreatic ductal adenocarcinomas is hard to overstate⁵². However, the association between the KRAS pathway and AR has not been investigated. In our work, GSEA result showed that *ENTPD1* was involved in the KRAS_SIGNALING_UP pathway in AR, which might provide evidence for us to explore this association. In SLE, the low *ENTPD1* expression group was correlated with TGF_BETA_SIGNALING and TNFA_SIGNALING_VIA_NFKB pathways, which have been recognized as key characteristics of the occurrence of SLE^{53,54}. Interestingly, recent studies showed that hypoxia-inducible factor-1 α , a key protein of hypoxia pathway, participated in the pathogenesis of SLE by regulating microRNA, which supported our GSEA result that hypoxia pathway was involved in SLE.

There are some limitations to our studies. Firstly, the shared pathways have not been verified by basic experiments. Thus, more experimental studies are warranted as direct evidence to demonstrate the exact role of these pathways on AR and SLE. Secondly, though we confirmed as many DEGs as possible in our research, some genes correlating with diseases might be excluded because they did not match the selection criteria. Further larger sample studies are warranted to identify DEGs comprehensively. Thirdly, we were unable to analyze the association between hub gene and clinical characteristics due to a lack of such data. Lastly, we could not obtain SLE dataset including male and female individuals from the GEO database, which might because women are 9 times more frequently to develop SLE than men⁵⁵. Thus, the SLE datasets included in this study were consisted of female individuals. Notably, previous study has reported that the distribution and expression of CD4⁺T cells was not associated with sex⁵⁶. Hence, sex difference is unlikely to significantly affect the reliability of our findings.

Conclusions

In conclusion, we identified *ENTPD1* as a shared hub gene in AR and SLE by applying bioinformatic analysis based on the GEO database. The *ENTPD1* expression increased in AR and SLE patients, and the diagnostic efficacy of *ENTPD1* in two diseases was validated. Our study may provide a new therapeutic target for preventing AR from developing SLE.

Data availability

The datasets analyzed during the current study are available in the GEO repository at www.ncbi.nlm.nih.gov/geo/ (accession number: GSE103760, GSE50223, GSE44960 and GSE10325).

References

- Barber, M. R. W. *et al.* Global epidemiology of systemic lupus erythematosus. *Nat. Rev. Rheumatol.* **17**, 515–532 (2021).
- Dall'Era, M. *et al.* The incidence and prevalence of systemic lupus erythematosus in San Francisco County, California: The California lupus surveillance project. *Arthritis Rheumatol.* **69**, 1996–2005 (2017).
- Arnaud, L. *et al.* Prevalence and incidence of systemic lupus erythematosus in France: A 2010 nation-wide population-based study. *Autoimmun. Rev.* **13**, 1082–1089 (2014).
- Ekwoom, P. E. Systemic lupus erythematosus (SLE) at the Kenyatta National Hospital. *Clin. Rheumatol.* **32**, 1215–1217 (2013).
- Zou, Y. F. *et al.* Prevalence of systemic lupus erythematosus and risk factors in rural areas of Anhui Province. *Rheumatol. Int.* **34**, 347–356 (2014).
- Yen, E. Y. & Singh, R. R. Brief report: Lupus-an Unrecognized leading cause of death in young females: A population-based study using nationwide death certificates, 2000–2015. *Arthritis Rheumatol.* **70**, 1251–1255 (2018).
- Fava, A. & Petri, M. Systemic lupus erythematosus: Diagnosis and clinical management. *J. Autoimmun.* **96**, 1–13 (2019).
- Bousquet, J. *et al.* Allergic rhinitis and its impact on asthma (ARIA) 2008 update (in collaboration with the World Health Organization, GA(2)LEN and AllerGen). *Allergy* **63**(Suppl 86), 8–160 (2008).
- Hong, H., Liao, S., Chen, F., Yang, Q. & Wang, D. Y. Role of IL-25, IL-33, and TSLP in triggering united airway diseases toward type 2 inflammation. *Allergy* **75**, 2794–2804 (2020).
- Hsiao, Y. P. *et al.* Atopic diseases and systemic lupus erythematosus: An epidemiological study of the risks and correlations. *Int. J. Environ. Res. Public Health* **11**, 8112–8122 (2014).
- Hou, Y. C., Hu, H. Y., Liu, I. L., Chang, Y. T. & Wu, C. Y. The risk of autoimmune connective tissue diseases in patients with atopy: A nationwide population-based cohort study. *Allergy Asthma Proc.* **38**, 383–389 (2017).
- Krishna, M. T. *et al.* Allergic diseases and long-term risk of autoimmune disorders: Longitudinal cohort study and cluster analysis. *Eur. Respir J.* **54**(5), 1900476 (2019).
- Wongtrakul, W., Charoengam, N., Ponvilawan, B. & Ungprasert, P. Allergic rhinitis and risk of systemic lupus erythematosus: A systematic review and meta-analysis. *Int. J. Rheum. Dis.* **23**, 1460–1467 (2020).
- Ruterbusch, M., Pruner, K. B., Shehata, L. & Pepper, M. In Vivo CD4(+) T Cell differentiation and function: Revisiting the Th1/Th2 paradigm. *Annu. Rev. Immunol.* **38**, 705–725 (2020).
- Pan, L., Lu, M. P., Wang, J. H., Xu, M. & Yang, S. R. Immunological pathogenesis and treatment of systemic lupus erythematosus. *World J. Pediatr.* **16**, 19–30 (2020).
- Gupta, R. K., Gupta, K. & Dwivedi, P. D. Pathophysiology of IL-33 and IL-17 in allergic disorders. *Cytokine Growth Factor Rev.* **38**, 22–36 (2017).
- Zhang, Y., Lan, F. & Zhang, L. Update on pathomechanisms and treatments in allergic rhinitis. *Allergy* **77**, 3309–3319 (2022).
- Yan, Z. *et al.* Bioinformatics analysis and identification of underlying biomarkers potentially linking allergic rhinitis and asthma. *Med. Sci. Monit.* **26**, e924934 (2020).
- Yao, M. *et al.* Exploration of the shared gene signatures and molecular mechanisms between systemic lupus erythematosus and pulmonary arterial hypertension: Evidence from transcriptome data. *Front. Immunol.* **12**, 658341 (2021).
- Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucl. Acids Res.* **28**, 27–30 (2000).
- Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucl. Acids Res.* **51**, D587–D592 (2023).
- Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951 (2019).
- Chin, C. H. *et al.* cytoHubba: Identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* **8**(Suppl 4), S11 (2014).
- Davis, A. P. *et al.* The Comparative Toxicogenomics database: Update 2017. *Nucl. Acids Res.* **45**, D972–D978 (2017).
- Aringer, M. *et al.* 2019 European league against rheumatism/American college of rheumatology classification criteria for systemic lupus erythematosus. *Ann. Rheum. Dis.* **78**, 1151–1159 (2019).
- Warde-Farley, D. *et al.* The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucl. Acids Res.* **38**, W214–W220 (2010).
- Finotello, F. & Trajanoski, Z. Quantifying tumor-infiltrating immune cells from transcriptomics data. *Cancer Immunol. Immunother.* **67**, 1031–1040 (2018).
- Ocampo-Piraquive, V., Nieto-Aristizábal, I., Cañas, C. A. & Tobón, G. J. Mortality in systemic lupus erythematosus: Causes, predictors and interventions. *Exp. Rev. Clin. Immunol.* **14**, 1043–1053 (2018).
- Kirtland, M. E., Tsitoura, D. C., Durham, S. R. & Shamji, M. H. Toll-Like receptor agonists as adjuvants for allergen immunotherapy. *Front. Immunol.* **11**, 599083 (2020).
- Fillatreau, S., Manfroi, B. & Dörner, T. Toll-like receptor signalling in B cells during systemic lupus erythematosus. *Nat. Rev. Rheumatol.* **17**, 98–108 (2021).
- Timperi, E. & Barnaba, V. CD39 Regulation and Functions in T Cells. *Int. J. Mol. Sci.* **22**(15), 8068 (2021).
- Becker, L. V., Passos, D. F., Leal, D. B. R., Morsch, V. M. & Schetinger, M. R. C. ATP signaling and NTPDase in systemic lupus erythematosus (SLE). *Immunobiology* **224**, 419–426 (2019).
- Becker, L. V. *et al.* Activity and expression of E-NTPDase is altered in peripheral lymphocytes of systemic lupus erythematosus patients. *Clin. Chim. Acta Int. J. Clin. Chem.* **488**, 90–7 (2019).
- Loza, M. J., Anderson, A. S., O'Rourke, K. S., Wood, J. & Khan, I. U. T-cell specific defect in expression of the NTPDase CD39 as a biomarker for lupus. *Cell Immunol.* **271**, 110–117 (2011).
- Knight, J. S. *et al.* Ectonucleotidase-mediated suppression of lupus autoimmunity and vascular dysfunction. *Front. Immunol.* **9**, 1322 (2018).
- Dang, J. *et al.* Human gingiva-derived mesenchymal stem cells are therapeutic in lupus nephritis through targeting of CD39(-)CD73 signaling pathway. *J. Autoimmun.* **113**, 102491 (2020).
- Zhang, X. *et al.* CD8+CD103+ iTregs inhibit chronic graft-versus-host disease with lupus nephritis by the increased expression of CD39. *Mol. Ther J. Am. Soc. Gene Ther.* **27**, 1963–1973 (2019).
- Melchioni, R. *et al.* Genetic analysis of an allergic rhinitis cohort reveals an intercellular epistasis between FAM134B and CD39. *BMC Med. Genet.* **15**, 73 (2014).
- Moitra, S. *et al.* Modulation of regulatory T cells by intranasal allergen immunotherapy in an experimental rat model of airway allergy. *Int. Immunopharmacol.* **47**, 9–19 (2017).
- Montuori-Andrade, A. C. M. *et al.* Lactobacillus delbrueckii UFV-H2b20 increases IFN-gamma production and CD39(+)/CD73(+) Treg cell numbers in lungs, and protects mice against experimental allergic asthma. *Immunobiology* **227**, 152284 (2022).
- Huang, Y. A. *et al.* Reducing lung ATP levels and alleviating asthmatic airway inflammation through adeno-associated viral vector-mediated CD39 expression. *Biomedicines* **9**(6), 656 (2021).
- Li, P. *et al.* CD39+ regulatory T cells attenuate allergic airway inflammation. *Clin. Exp. Allergy* **45**, 1126–1137 (2015).

43. Li, P., Cao, J., Chen, Y., Wang, W. & Yang, J. Apyrase protects against allergic airway inflammation by decreasing the chemotactic migration of dendritic cells in mice. *Int. J. Mol. Med.* **34**, 269–275 (2014).
44. Lu, Y., Li, Y., Zhou, W., Ding, B. & Yu, Q. Regulatory T cells regulate the distribution of natural killer T cells through CD39 signal transduction in asthma. *Hum. Cell* **32**, 141–149 (2019).
45. Wing, J. B., Tanaka, A. & Sakaguchi, S. Human FOXP3(+) regulatory T cell heterogeneity and function in autoimmunity and cancer. *Immunity* **50**, 302–316 (2019).
46. Paiva Ferreira, L. K. D. *et al.* Combined allergic rhinitis and asthma syndrome (CARAS). *Int. Immunopharmacol.* **74**, 105718 (2019).
47. Suzuki, M. *et al.* IFN-gamma secretion by CD8T cells inhibits allergen-induced airway eosinophilia but not late airway responses. *J. Allergy Clin. Immunol.* **109**, 803–809 (2002).
48. Lin, L., Dai, F., Wei, J. & Chen, Z. Influences of CD8(+) Tregs on peripheral blood mononuclear cells from allergic rhinitis patients. *The Laryngoscope* **131**, E316–E323 (2021).
49. Li, J. *et al.* KIR(+)CD8(+) T cells suppress pathogenic T cells and are active in autoimmune diseases and COVID-19. *Science (New York, NY)*. **376**, eabi9591 (2022).
50. Saligrama, N. *et al.* Opposing T cell responses in experimental autoimmune encephalomyelitis. *Nature* **572**, 481–487 (2019).
51. Krassel, M. & Baerwald, C. Sex, symptom severity, and quality of life in rheumatology. *Clin. Rev. Allergy Immunol.* **56**, 346–361 (2019).
52. Drosten, M. & Barbacid, M. Targeting the MAPK pathway in KRAS-driven tumors. *Cancer cell*. **37**, 543–550 (2020).
53. Aringer, M. Inflammatory markers in systemic lupus erythematosus. *J. Autoimmun.* **110**, 102374 (2020).
54. Liu, Y. *et al.* The TGF- β /miR-31/CEACAM1-S axis inhibits CD4(+) CD25(+) Treg differentiation in systemic lupus erythematosus. *Immunol. Cell Biol.* **99**, 697–710 (2021).
55. Alexander, T. & Hedrich, C. M. Systemic lupus erythematosus—Are children miniature adults? *Clin. Immunol.* **234**, 108907 (2022).
56. Niu, H. Q. *et al.* Characteristics and reference ranges of CD4(+)T cell subpopulations among healthy adult Han Chinese in Shanxi Province, North China. *BMC Immunol.* **21**, 44 (2020).

Acknowledgements

The authors gratefully acknowledged the financial supports by the National Natural Science Foundation of China under Grant numbers 81200734.

Author contributions

H.Z., H.T. and M.Z. contributed to the study conception and design. Data analysis was performed by M.C. and C.Z. The first draft of the manuscript was written by M.C., X.S. and Y.M. And all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

This study was supported by a grant (No.81200734) from the National Natural Science Foundation of China.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-69228-3>.

Correspondence and requests for materials should be addressed to M.Z., H.T. or H.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024