Journal of Inflammation Research

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Identification of Key Disulfidptosis-Related Genes and Their Association with Gene Expression Subtypes in Crohn's Disease

Mingyue Fu[1,2,*], Diliaremu Aihemaiti[1,2,*], Haowen Fu[1,2,*], Na Li[1,2], Yifan Yuan[3], Mei Ye[1,2]

[1]Department of Gastroenterology, Zhongnan Hospital, Wuhan University, Wuhan, Hubei, People's Republic of China; [2]Hubei Clinical Centre and Key Laboratory of Intestinal and Colorectal Diseases, Zhongnan Hospital, Wuhan University, Wuhan, Hubei, People's Republic of China; [3]Center of Gerontology and Geriatrics, West China Hospital, Sichuan University, Chengdu, People's Republic of China

*These authors contributed equally to this work

Correspondence: Yifan Yuan; Mei Ye, Email yifanyuan95@163.com; wumeiye08@163.com

Background: Crohn's disease (CD) is a persistent inflammatory condition that impacts the gastrointestinal system and is characterized by a multifaceted pathogenesis involving genetic, immune, and environmental components. This study primarily investigates the relationship between gene expression and immune cell infiltration in CD, focusing on disulfidptosis—a novel form of cell death caused by abnormal disulfide accumulation—and its impact on various immune cell populations. By identifying key disulfidptosis-related genes (DRGs) and exploring their association with distinct gene expression subtypes, this research aims to enhance our understanding of CD and potentially other autoimmune diseases.

Methods: Gene expression data from intestinal biopsy samples were collected from both individuals with CD and healthy controls, and these data were retrieved from the GEO database. Through gene expression level comparisons, various differentially expressed genes (DEGs) were identified. Subsequently, Gene Ontology and Kyoto Encyclopedia of Genes and Genomes enrichment analyses were performed to reveal the biological processes and pathways linked to these DEGs. Later, immune cell infiltration was evaluated. Hub candidate DRGs were identified using machine learning algorithms. Validation of the expression of hub DRGs was carried out using quantitative real-time polymerase chain reaction. The hub DRGs were subjected to unsupervised hierarchical clustering to classify CD patients into subtypes. The characteristics of each subtype were then analyzed.

Results: Two hub DRGs (NDUFA11 and LRPPRC) were identified. NDUFA11 showed a significantly positive association with the abundance of Th17 cells. Conversely, higher expression levels of LRPPRC were associated with a reduced abundance of various immune cells, particularly monocytes. CD patients were classified into two disulfidptosis-related subtypes. Cluster B patients exhibited lower immune infiltration and milder clinical presentation.

Conclusion: LRPPRC and NDUFA11 are identified as hub DRGs in CD, with potential roles in disulfidptosis and immune regulation. The disulfidptosis subtypes provide new insights into disease progression.

Keywords: Crohn's disease, disulfidptosis, machine learning, expression pattern, immune cell infiltration

Background

Crohn's disease (CD) is a chronic inflammatory condition that primarily affects the gastrointestinal tract, leading to inflammation and damage in this region. The etiology of CD is multifactorial, involving immune responses, genetics, dysbiosis of gut microflora, and impaired intestinal barrier function.[1] CD has witnessed a concerning increase in incidence rates in emerging industrialized nations, leading to a significant disease burden.[2] As a chronic condition, CD profoundly impacts clinical presentation, extraintestinal manifestations, and overall quality of life.[3] Nevertheless, the precise mechanisms underlying CD remain elusive. Understanding the primary causal mechanisms may facilitate clinical diagnosis and treatment, ultimately resulting in enhanced clinical outcomes.

Received: 19 February 2024
Accepted: 16 May 2024
Published: 7 June 2024

Regulated cell death (RCD) is a common biological process, that is commonly triggered by specific molecules.[4] Recently, various emerging RCD modalities have attracted significant attention in the scientific community. These modalities encompass a diverse array of intriguing processes, including but not limited to autophagy-dependent cell death, reticulocyte death, basal prolapse, ferroptosis, necroptosis, lysosome-dependent cell death, cytoplasmic division, immunogenic cell death, apoptosis, cuproptosis, and endogenous cell death.[4–6] In the pathogenesis and progression of CD, various forms of cell death have been reported to have significant implications. These include apoptosis, pyroptosis, necrosis, necroptosis, ferroptosis, and cuproptosis, all of which collectively contribute to the intricate mechanisms driving CD.[7–9] Gaining insights into the involvement of these diverse cell death pathways is vital for comprehending the underlying pathophysiology and advancing therapeutic approaches to tackle this complex disease.

Disulfidptosis, a novel nonprogrammed cell death model entailed by disulfide, was identified by Xiaoguang Liu et al.[10] They found that glucose starvation leads to upregulated expression of solute carrier family 7 member 11 (*SLC7A11*) in kidney cancer cells, which, in turn, intensifies the utilization of nicotinamide adenine dinucleotide phosphate (*NADPH*). Under general conditions, *NADPH* participates in the reduction of disulfide. However, under glucose starvation, *NADPH* output declines substantially, and the absorption of cystine by *SLC7A11* consumes more *NADPH*, resulting in its depletion and the buildup of disulfide molecules. Consequently, abnormal aggregation of disulfide molecules occurs, instigating destabilization of the filament network through the formation of actin-cytoskeletal protein disulfide bonds. Studies have shown that disulfidptosis occurs in lung adenocarcinoma, bladder cancer, liver cancer, renal cell carcinoma, and other diseases.[11–17]

While primarily focusing on disulfidptosis in Crohn's disease, our findings also suggest potential implications for a range of autoimmune disorders like rheumatoid arthritis, systemic lupus erythematosus, and multiple sclerosis, all characterized by chronic inflammation and immune dysregulation. Understanding these connections could reveal new therapeutic targets and advance our comprehension of autoimmune mechanisms. However, disulfidptosis research is still in its infancy, and its significance in illness development is unknown.

High-throughput sequencing tools have provided invaluable insights into how genes influence disease mechanisms.[18] To examine the distinct expression patterns of disulfidptosis-related genes (DRGs) in both healthy and CD intestinal mucosal tissues, we analyzed microarray data from the Gene Expression Omnibus (GEO) database. Using Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and Gene Ontology (GO) analyses, we deciphered the molecular processes underlying CD. By employing multiple machine-learning algorithms and analyzing our own human samples, we successfully identified two hub genes, namely, Leucine Rich Pentatricopeptide Repeat Containing (*LRPPRC)* and NADH dehydrogenase 1 alpha subcomplex subunit 11 (*NDUFA11)*, which are central to the disulfidptosis-related processes in CD. A total of 410 CD patients were stratified into two distinct groups associated with disulfidptosis based on the expression patterns of hub genes. Subsequently, a comprehensive analysis of immune cell infiltration, enrichment pathways, and clinical characteristics was performed to identify the distinctions between these two groups. These discoveries imply that directing efforts toward disulfidptosis may offer a promising avenue for both diagnosis and treatment, offering novel insights into its underlying molecular mechanisms.

Materials and Methods

Dataset Acquisition and DRGs

In the pursuit of relevant gene expression data, two raw datasets, GSE112366 and GSE95095, were retrieved from the GEO database.[19] These datasets contained gene expression profiles from intestinal biopsy samples of both CD patients and control subjects. Specifically, GSE112366 consists of 362 samples from individuals with CD and 26 samples from healthy individuals, whereas GSE95095 comprises 48 CD samples and 12 control samples. To mitigate batch effects and unify the GEO dataset, we applied Bioconductor "sva" R software.[20] We derived the DRGs for this study from previous research.[10]

Detection of Differentially Expressed Genes (DEGs)

The analysis of DEGs was conducted using the R package "limma"[21] DEGs between CD patients and healthy individuals in the merged dataset were determined using a significance threshold of $P<0.05$. Volcano plots and a heatmap were

employed to visually represent the differential gene expression data. Furthermore, GO and KEGG pathway analysis, facilitated by the "clusterProfiler" package, offered comprehensive insights into the biological functions of the DEGs.[22]

Evaluation of Immune Cell Infiltration

A comprehensive understanding of the immunological microenvironment, encompassing mesenchymal tissues, immune cells, cytokines, inflammatory cells, fibroblasts, and chemokines, is essential for comprehending the disease trajectory and therapeutic efficacy in the context of CD. To evaluate the immunological characteristics across all samples, we used the "GSVA" R package with the single-sample gene set enrichment analysis (ssGSEA) technique.[23]

Machine Learning

Through the intersection of DEGs and DisRGs, we identified a subset of genes, referred to as DEDRGs, that exhibited differential expression and association with disulfidptosis in the context of CD. To further refine the list of potential hub DRGs for CD diagnosis, three distinct machine-learning techniques were employed. The initial approach involved implementing least absolute shrinkage and selection operator (LASSO) regression, which aids in enhancing predictive accuracy and model interpretability by selecting pertinent variables.[24] Subsequently, the powerful support vector machine (SVM) method was utilized to establish a threshold between two classes, enabling label prediction based on one or more feature vectors.[25] Finally, to forecast continuous variables with minimal fluctuations, the random forest (RF) technique was adopted, offering the advantages of variable condition independence and improved accuracy, sensitivity, and specificity.[26] These analyses were executed using the R packages "glmnet",[27] "kernlab",[28] and "randomForest"[29] for LASSO regression, SVM, and RF, respectively. The genes identified through the intersection of these methods were considered hub disulfidptosis genes in CD.

Specimen Collection

From December 2021 to May 2022, Zhongnan Hospital, Wuhan University, China, enrolled a total of 12 individuals diagnosed with CD. The diagnosis was confirmed through examination of tissue samples. It is crucial to note that none of the patients were administered immunosuppressants, steroids, or biologic medications throughout the duration of the research. To establish a suitable control group, 12 healthy volunteers of similar age and gender were also included in the study. All collected tissue samples were promptly preserved using liquid nitrogen.

Quantitative Real-Time PCR (qRT–PCR)

We extracted total RNA from the tissues using TRIzol reagent (Invitrogen, USA) and performed reverse transcription utilizing the TOYOBO ReverTra Ace kit (TOYOBO, Japan). Next, we quantified mRNA expression through qRT–PCR on Biorad CFX (Biorad, USA), with GAPDH being utilized as the chosen housekeeping gene. The relative gene expression level was calculated using the formula Ratio = (Etarget)ΔCPtarget (control−sample)/(Eref)ΔCPref(control −sample), as previously described in the literature. The primers used in this study were designed and synthesized by TSINGKE Biological Technology (Wuhan, China) and were as Supplementary Table 1.

Subcluster Analysis with Hub DRGs

To perform subcluster analysis based on the hub DRGs, an unsupervised hierarchical clustering analysis was performed on the mRNA expression data obtained from 410 CD samples. This analysis was conducted employing the "ConsensusClusterPlus" R package.[30] The input information for clustering comprised the expression levels of the two hub DRGs. The subclusters obtained were then visualized in a principal component analysis (PCA) plot, which revealed the geometric distance between them. To elucidate the functional differences among the disulfidptosis subclusters identified in the preceding clustering analysis, we conducted gene set variation analysis (GSVA).[23] Subsequently, a heatmap was constructed to visualize the activity of pathways associated with the two subclusters. We defined the selection criteria for DEGs between the two disulfidptosis-related subclusters as an adjusted p value < 0.05 and |log2-fold change (FC)| > 0.5, indicating statistical significance. These DEGs were displayed using a volcano plot. Furthermore, to

evaluate the biological functions associated with the identified DEGs between the two disulfidptosis-related subclusters, we performed GO and KEGG analyses.[22]

Statistical Analysis

We performed data and statistical analysis using R version 4.2.2, employing the Wilcoxon test to evaluate significant differences between the two groups. Moreover, the correlation between the expression levels of genes related to hub disulfidptosis and immune cells was investigated using the Spearman correlation. For the statistical analysis of qRT–PCR data, a *t* test was conducted, with a p value of 0.05 set as the threshold for statistical significance determination.

Results

Merging Dataset and Functional Analyses of DEGs

A dataset comprising 410 CD samples and 38 control samples was successfully integrated, and PCA confirmed the efficient removal of batch effects across the three datasets (Figures 1A and B). Through a volcano plot analysis, a total of 2030 DEGs were displayed (Figure 1C). The heatmap provides a visual representation of the 20 most upregulated and downregulated genes among the DEGs (Figure 1D). In the biological process analysis, GO enrichment analysis revealed significant enrichments in pathways related to cytokine-mediated signaling, response to oxidative stress, lymphocyte differentiation, intrinsic apoptotic defense response, positive regulation of defense response, and response to interleukin-1. In the cellular component analysis, enrichment was found in focal adhesion, cytoplasmic vesicle lumen and cell-substrate junction. In the molecular function analysis, the DEGs showed enrichment in various functions, including CXCR chemokine receptor binding, chemokine activity, chemokine receptor binding, and DNA-binding transcription activator and repressor activity (Figure 1E). KEGG pathway enrichment analysis highlighted the enrichment of pathways such as MAPK signaling, TNF signaling, chemokine signaling, PI3-Akt signaling, and HIF-1 signaling (Figure 1F).

Evaluation of Immune Cell Infiltration

To further explore the immunological regulation of CD, an immune cell infiltration investigation was performed using the Merging Dataset. In the box plot, it was evident that the CD group showed higher levels of immune cell infiltration than the control group. Figure 2A presents the differences observed in the proportions of 12 kinds of immune cell subsets, including T helper type 17 (Th17) cells, CD56dim natural killer (NK) cells, immature dendritic cells (DCs), plasmacytoid DCs, activated DCs, gamma delta cells, immature B cells, mast cells, monocytes, NK cells, neutrophils, and T helper type 2 cells, which were found to be different between the CD and control groups. Notably, only the proportion of immature B cells decreased, while others were increased. However, the other 11 immune cell subsets did not show any statistically significant differences between the groups. The association analysis revealed robust correlations among most immune cells, except for Th17 cells or CD56dim NK cells, as depicted in Figure 2B. The unique infiltration patterns of diverse immune cells in CD patients once again prove the crucial role of immunity in CD pathogenesis.

Machine Learning-Based Identification of the Disulfidptosis Signature

After intersecting the 2021 DEGs with 15 DRGs (Figure 3A), a total of 9 DEDRGs were identified. Figure 3B demonstrates the visual depiction of the overall expression levels of these DEDRGs in CD samples in comparison to the normal samples, wherein a significant portion of the identified DEDRGs displayed increased expression levels in the CD cohort. Supplementary Figures 1A-C demonstrates the utilization of sophisticated machine learning algorithms to derive a disulfidptosis signature from gene expression data, essential for understanding the molecular mechanisms underpinning this unique form of cell death. Using Support Vector Machine (SVM), LASSO (Least Absolute Shrinkage and Selection Operator) Regression, and Random Forest algorithms, we meticulously analyzed the expression profiles to discern potential biomarkers critical for disulfidptosis. This comprehensive approach led to the identification of six key genes: *DSTN*, *FLNA*, *LRPPRC*, *MYL6*, *NDUFA11*, and *SLC7A11*, which collectively form the disulfidptosis signature (Supplementary Figure 1D). Subsequently, to specifically analyze the expression levels of these genes, qRT-PCR was performed on collected samples from individuals diagnosed with CD (Figures 4A-F). Finally, NDUFA11 and LRRPRC were defined as the two key DEDRGs.

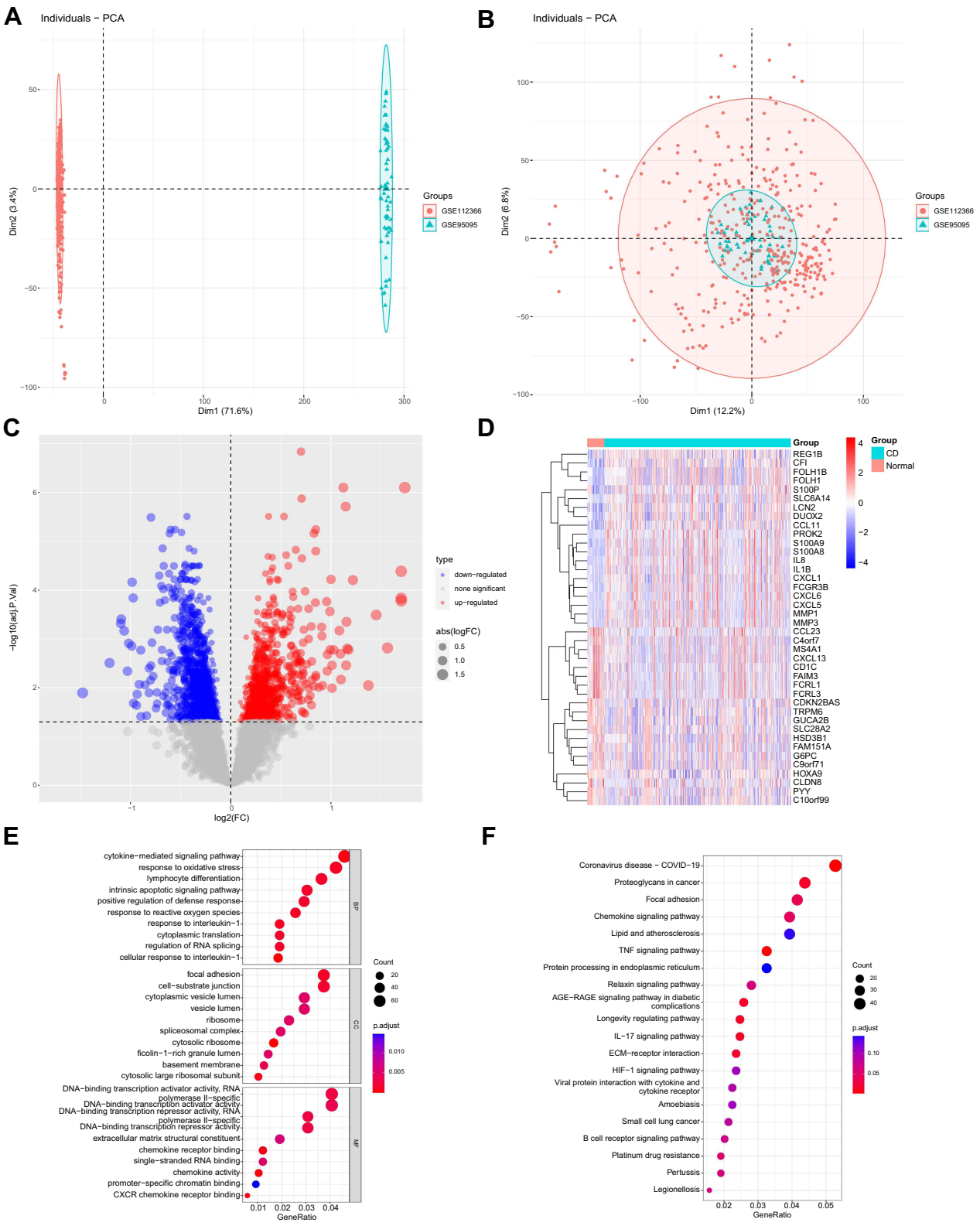

Figure 1 Integration of GEO datasets and identification of differentially expressed genes (DEGs). (**A** and **B**) illustrate principal component analysis demonstrating batch effects before and after de-batching the integrated datasets. (**C**) Depicts the volcano plot displaying DEGs related to Crohn's disease. (**D**) Presents a clustered heatmap exhibiting expression levels of DEGs. (**E**) Shows enriched Gene Ontology (GO)analysis items. (**F**) Shows enriched items from the Kyoto encyclopedia of genes and genomes (KEGG) pathway analysis items; "BP", biological processes; "MF", molecular functions; "CC", cellular component.

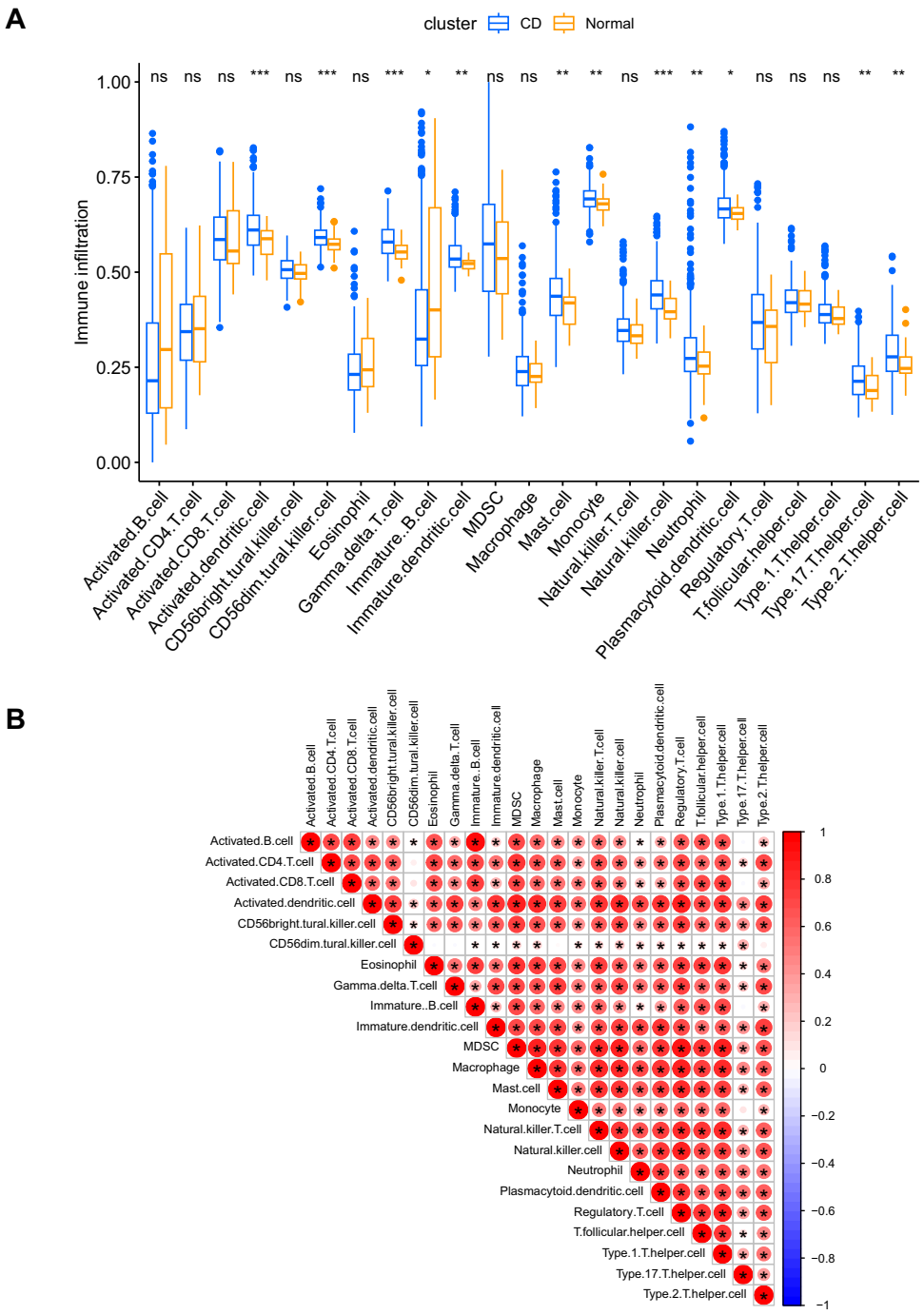

Figure 2 Immune cell infiltration analysis in merging GEO datasets. (**A**) comparing 23 immune cell subtypes between CD patients and controls, with circle size and color representing Pearson correlation coefficients. (**B**) Presents a correlation matrix illustrating the compositions of all 23 immune cell subtypes. "ns" denotes no significance, while "*", "**", and "***" correspond to p-values of <0.05, <0.01, and <0.001, respectively.

Immune Characteristics and ssGSEA of Key DEDRGs

Considering the importance of multiple immune components in the diagnosis and pathological mechanism of CD, the interrelation between immune cell infiltration and hub DRGs was examined in the Merging Dataset. We found that *NDUFA11* exhibited a positive relationship with the abundance of Th17 cells, CD56dim NK cells, neutrophils and NK cells and a negative relationship with activated CD4 T cells, T follicular helper cells, eosinophils, activated B cells, activated CD8 T cells, and immature B cells (Figure 5A). In contrast, a significant inverse association was observed

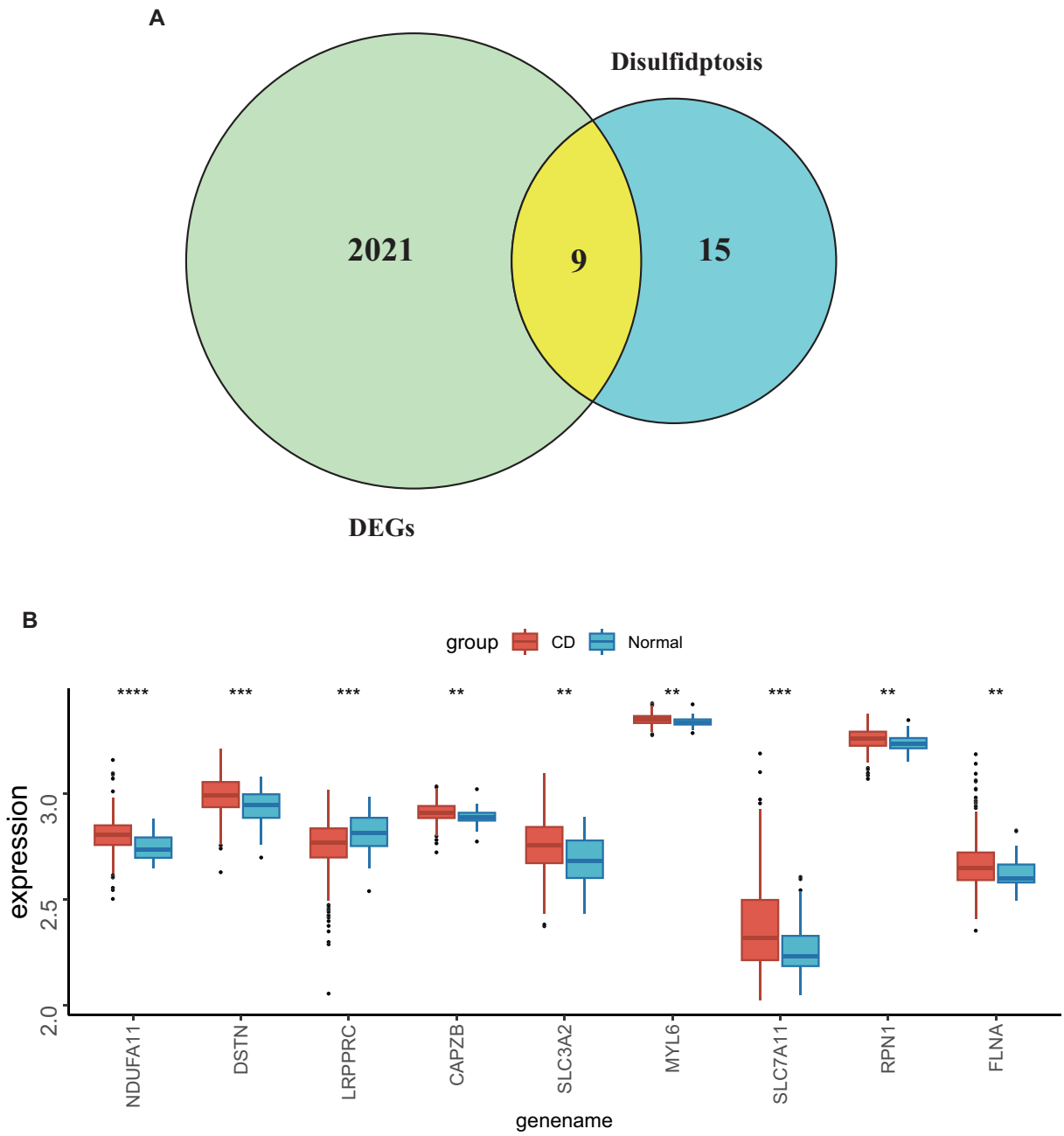

Figure 3 Determination of disulfidptosis-related genes (DRGs) in merging GEO datasets. (**A**) Demonstrates the gene overlap between DEGs and DRGs. (**B**) Depicts an overall expression histogram of differentially expressed DEGs in CD patients. "**","***", and "****" correspond to p-values of <0.01, <0.001, and <0.0001, respectively.

between *LRPPRC* and the majority of immune cells, where increased expression of this gene was consistently associated with lower abundances of these cells (Figure 5B). Figure 5C highlights *NDUFA11*'s involvement primarily in mitochondrial functions, evident from its association with mitochondrial translation and respiratory electron transport. Interestingly, there is a notable decrease in SUMOylation pathway enrichment, particularly impacting RNA binding and chromatin organization proteins. Figure 5D details LRPPRC's significant roles in RNA metabolism and translation-related pathways, underscoring its fundamental contributions to cellular processes without significant enrichment in SUMOylation.

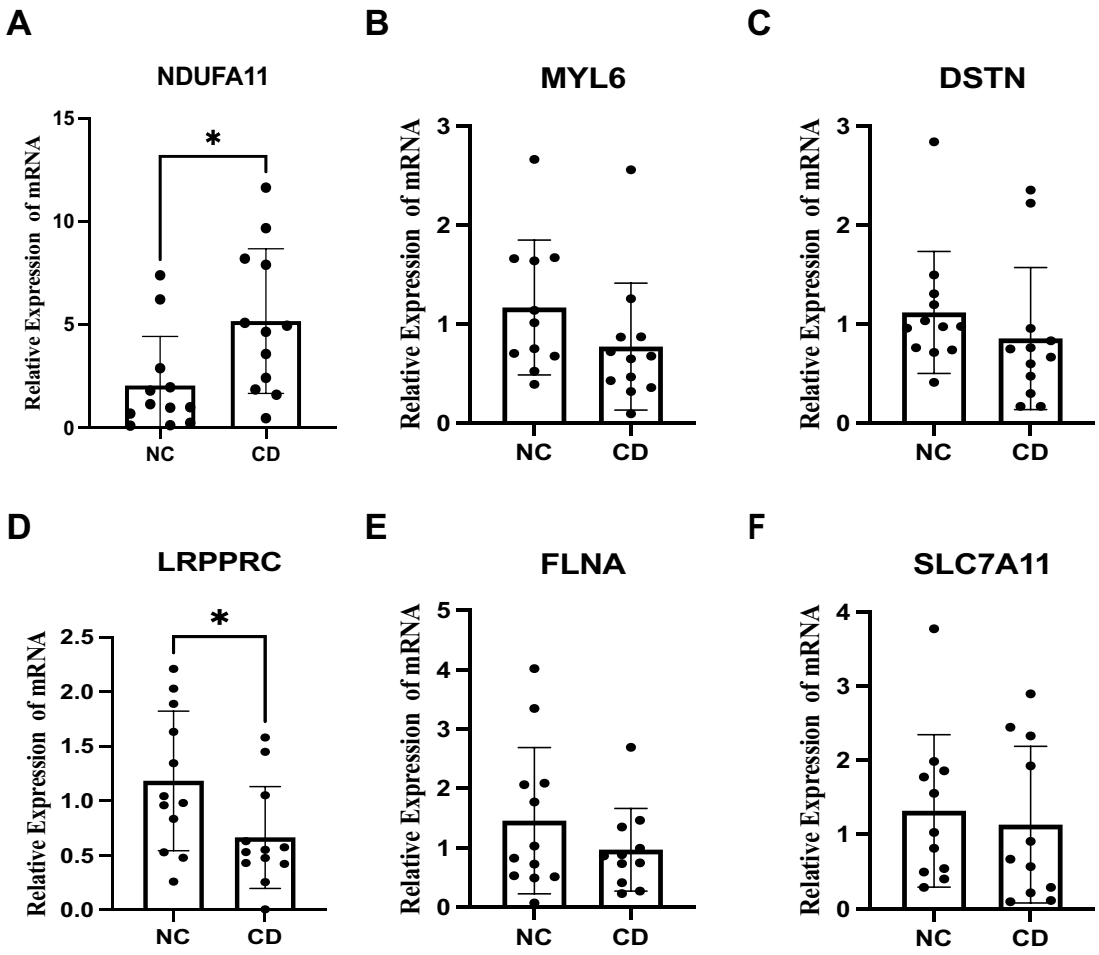

Figure 4 The mRNA levels of NDUFA11 (**A**), MYL6 (**B**), DSTN (**C**), LRPPRC (**D**), FLNA (**E**) and SLC7A11 (**F**) measured from collected samples. "*" represents p-values of <0.05.

Characteristics of Disulfidptosis-Related Clusters for CD

The analysis of consensus clustering, based on the two hub DRGs expression levels, yielded two distinct subclusters using data from CD patients in the merging dataset, showing the consensus matrix from k-means clustering analysis at k=2, where each element of the matrix represents the proportion of clustering iterations in which two samples were assigned to the same cluster. Darker blue signifies a higher consensus on sample pairing within the same cluster, demonstrating strong cluster stability and suggesting robust sub-group distinctions within the data. (Supplementary Figure 2A-C). The PCA further confirmed the distinct separation of patients within the two subclusters (Supplementary Figure 2D). Remarkably, Cluster B exhibited significantly elevated expression levels of LRPPRC in comparison to Cluster A (Supplementary Figure 2E). In total, 51 DEGs were identified, and the volcano plot effectively illustrated the disparities in gene expression (Figure 6A and Supplementary Table 2). The GO enrichment analysis revealed significant enrichment in pathways related to neutrophil migration, granulocyte migration, leukocyte migration, and signaling receptor activator activity (Figure 6B). The KEGG enrichment analysis indicated associations with graft-versus-host disease, cytokine–cytokine receptor interaction, rheumatoid arthritis, and the IL17 signaling pathway (Figure 6C). Data collected from 253 CD patients in the GSE112366 dataset, according to the Simple Endoscopic Score for CD (SES-CD), revealed that patients in cluster B exhibited lower scores, indicating a milder clinical presentation (Figure 6D). Based on the analysis of immune cell infiltration, it became evident that individuals in Cluster B displayed considerably lower levels of most immune cells, except for eosinophils, T helper type 2 cells, and activated CD4 T cells, which did not

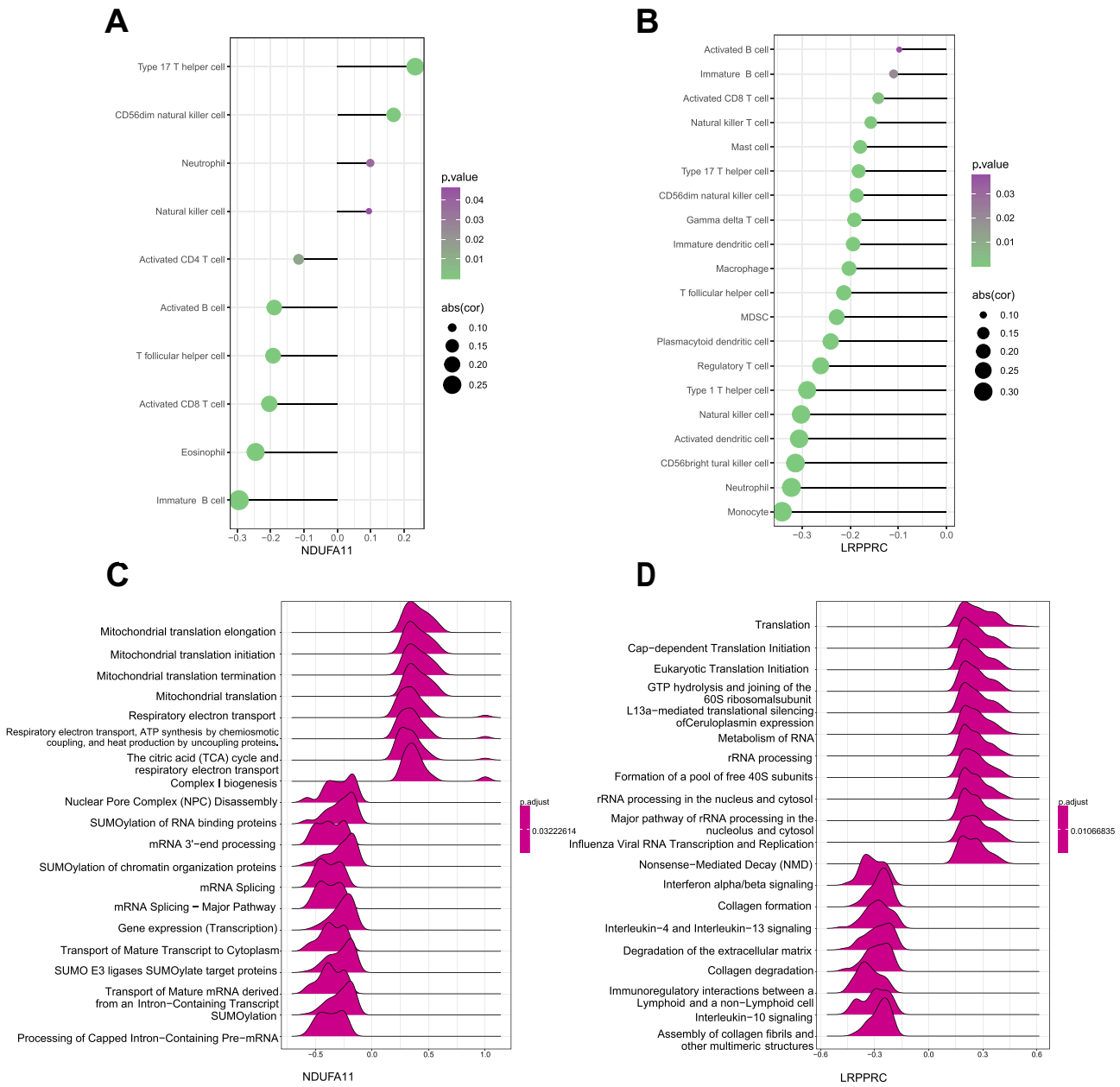

Figure 5 Correlation analysis of immune infiltrating cells and gene set enrichment analysis with NDUFA11 and LRPPRC in merging GEO datasets (**A–D**).

display significant differences (Figure 6E). This finding suggests heterogeneity between the subgroups, with Cluster B patients displaying lower immune infiltration and a milder clinical presentation.

Functional Differences Between DRGs-Clusters

A comparative investigation between Cluster A and Cluster B revealed distinctive findings regarding the KEGG activities associated with various signaling pathways in Helicobacter pylori infection. Specifically, Cluster B demonstrated noteworthy reductions in KEGG activities related to the MAPK signaling pathway, chemokine signaling pathway, epithelial cell signaling, and NK cell-mediated cytotoxicity. In contrast, Cluster B exhibited augmented activities in aminoacyl tRNA biosynthesis and RNA degradation pathways (Figure 7A). On the other hand, GSVA of the Reactome pathway revealed a significant enrichment of the anti-inflammatory response, favoring Leishmania parasite infection, other interleukin signaling, and signaling to Ras in Cluster A. In contrast, Cluster B exhibited enrichment in processes

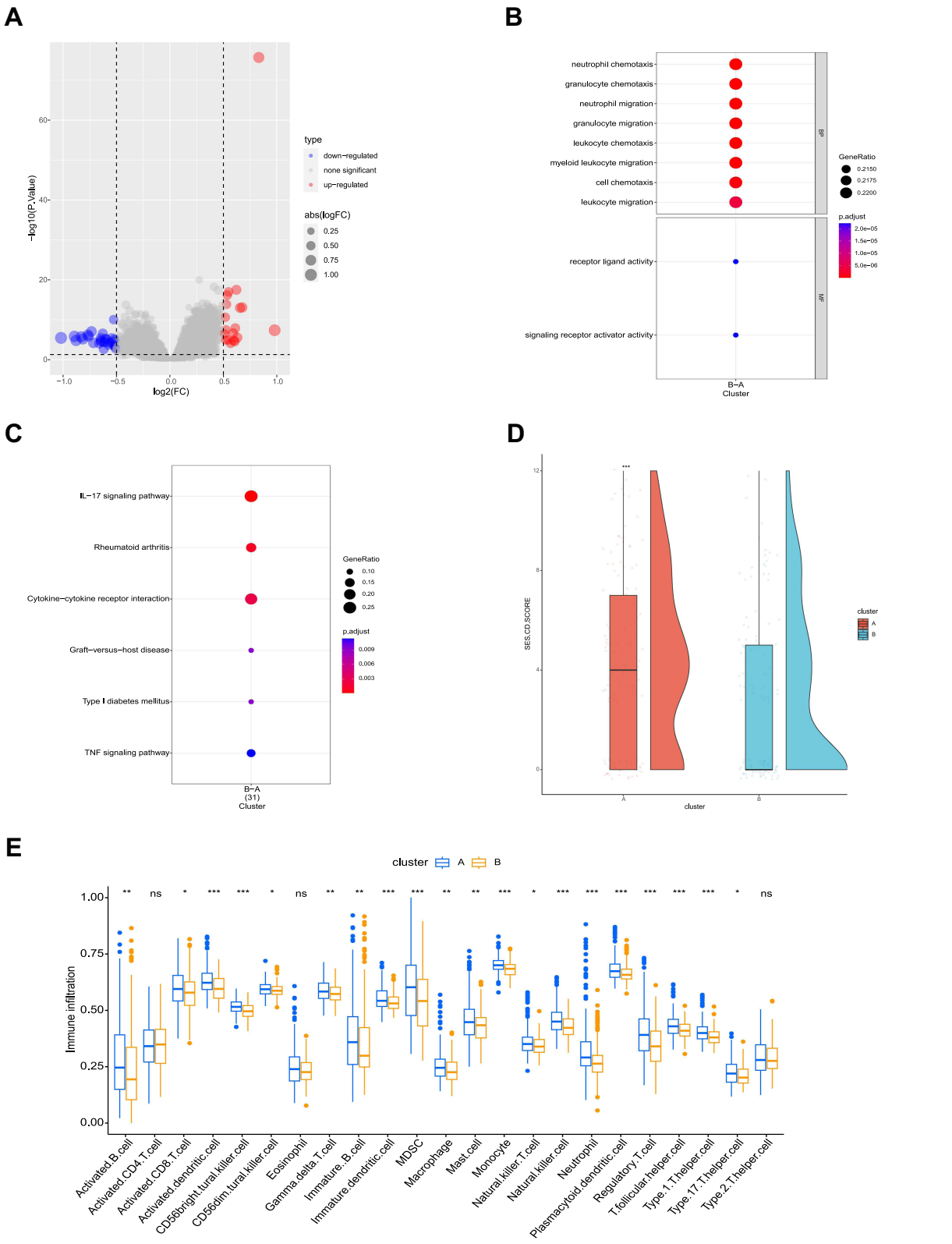

Figure 6 Characteristics of the disulfidptosis-related subtype. (**A**) Volcano plot showing DEGs. (**B**) Shows enriched GO analysis items. (**C**) Shows enriched KEGG analysis items. (**D**) Association between the disulfidptosis-related subtype and Simple Endoscopic Score for 253 CD patients (GSE112366). The shape of each violin plot indicates the density and distribution of scores within each subtype, with wider sections showing a higher density of data points. The box within each violin shows the interquartile range, and the line within the box indicates the median score. (**E**) Matrix depicting the correlation among all 23 immune cell subtypes within distinct clusters. "BP", biological processes; "MF", molecular functions; "ns" denotes no significance, while "*", "**", and "***" correspond to p-values of <0.05, <0.01, and <0.001, respectively.

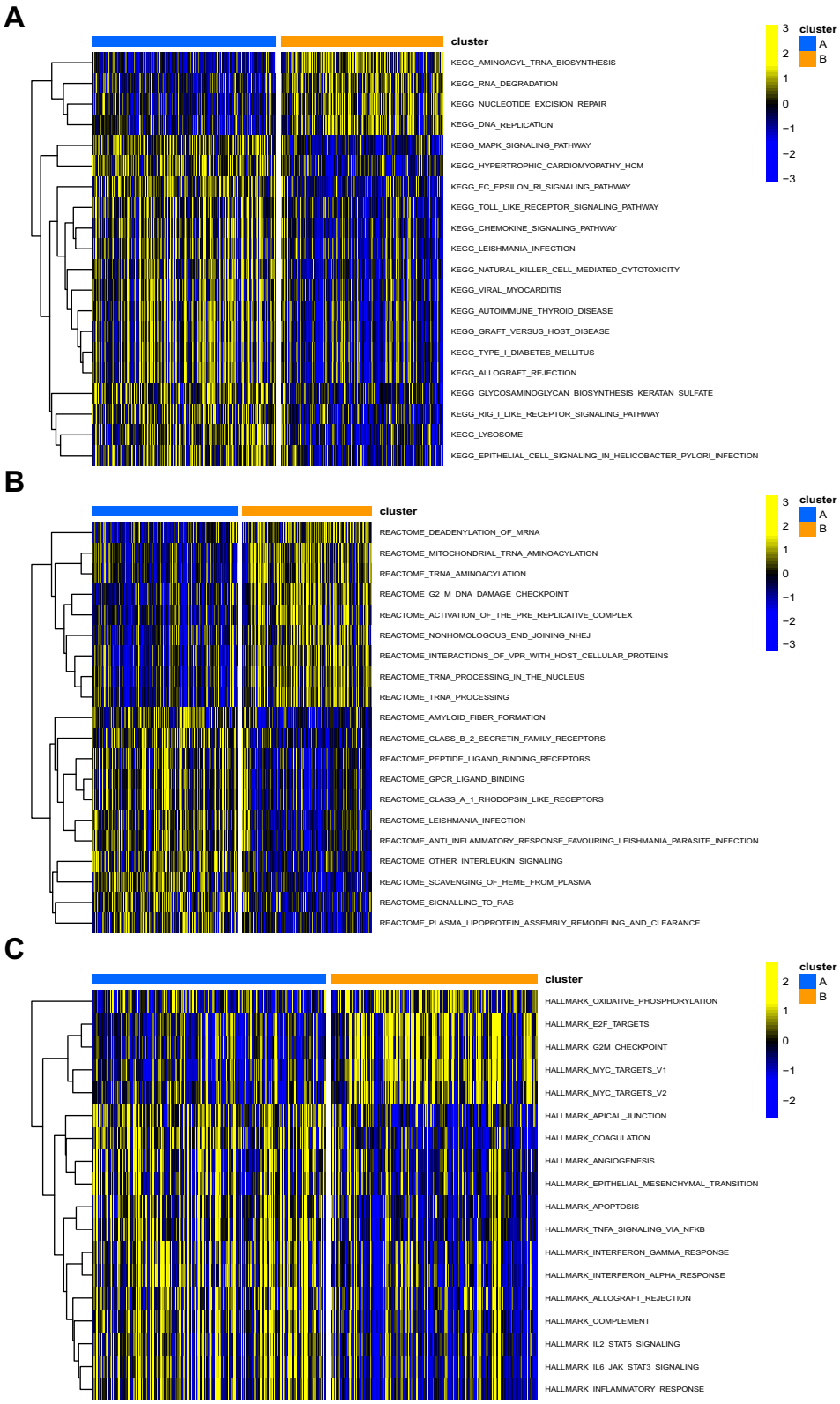

Figure 7 Gene set enrichment analysis based on the KEGG (**A**), Reactome (**B**) and HALLMARK (**C**) pathway.

related to deadenylation of mRNA, tRNA aminoacylation, and tRNA processing (Figure 7B). Additionally, Cluster A showed decreased expression of hallmark pathways associated with oxidative phosphorylation. However, it displayed higher expression levels of pathways related to IL-2-STAT5 signaling, TNF-α signaling via NF-κB, apoptosis, IL6-JAK-

STAT3 signaling, epithelial-mesenchymal transition, and inflammatory response (Figure 7C). Notably, consistent with the findings from the immune infiltration analysis showing lower immune infiltration in Cluster B, the activity of immune-related pathways was also lower in this cluster.

Discussion

In this study, we utilized bioinformatics techniques to identify the key differential expression of DRGs, *LRPPRC* and *NDUFA11*, in CD. The expression patterns of these genes are closely associated with the clinical severity and immune cell infiltration profiles in CD, suggesting their potential as biomarkers for dissecting the heterogeneous immune responses in CD. Our data not only reveal the unique expression profiles of *LRPPRC* and *NDUFA11* in CD but also demonstrate significant associations between their expression and the abundance of key immune cell subtypes. This is crucial for understanding their roles in the pathogenesis of CD. These findings provide a valuable foundation for further exploration of the biological roles of disulfidptosis in CD.

Our study corroborates established findings of significant immune cell infiltration in CD patients. Notably, the increased abundance of activated DCs, gamma delta cells, CD56 dim NK cells and natural killer cells was closely associated with CD activation. Through single-cell sequencing of intestinal mucosal tissues from CD patients, it was discovered that the mucosa of CD patients is abundant in IL-1β+ DCs.[31] Furthermore, it has been observed that patients in CD remission displayed a notable increase in CD16+CD56dim NK cell numbers.[32] Notably, CD16+CD56dim NK cells are known for their potent cytotoxicity capabilities.[33] Additionally, gamma delta T cells emerged as potential markers for poor prognosis following CD diagnosis.[34] Our study reinforces the pivotal role of immunity in CD.

Intestinal epithelial cell (IEC) death plays a crucial role in gut injury, and an imbalance between IEC death and regeneration can lead to increased intestinal permeability and compromised barrier function. This condition can make individuals more susceptible to a range of acute and chronic intestinal ailments, such as CD and necrotizing enterocolitis.[7,35] Genome-wide association studies have provided evidence suggesting that anomalies in the epithelial barrier might constitute a fundamental mechanism in the pathogenesis of CD. These studies have pinpointed susceptibility polymorphisms within genes encoding junctional proteins, including E-cadherin, Zonula occludens-1, and guanine nucleotide-binding protein alpha 12.[36,37] These discoveries indicate the vital significance of these proteins in maintaining the integrity of the gut barrier and provide insights into their potential contribution to the susceptibility of CD.

It has been reported that different forms of cell death contribute to the pathogenesis and progression of CD.[7–9] The gut epithelium is protected by a defensive mucus layer, formed by mucin polymers linked with disulfide bonds, which acts to restrict the exposure of epithelial cells to harmful toxins and bacteria. The potential manipulation of mucus integrity through the reduction of sulfide concentrations in the intestine presents a novel and promising therapeutic strategy for CD.[38] A newly identified type of cell death known as disulfidptosis can be initiated by the abnormal accumulation of intracellular disulfides in *SLC7A11*high cells during glucose deprivation. This mechanism results in the disruption of actin cytoskeleton proteins and F-actin in a manner dependent on *SLC7A11*.[10,39] Although the precise role of disulfidptosis in CD remains unclear, its emergence as a subject worthy of further investigation holds significant potential.

In the scope of our investigation, our aim was to delve into the association between DRGs and CD. Through the identification of overlapping genes between DEGs and DRGs, we employed machine learning techniques to pinpoint six candidate hub genes (*DSTN, FLNA, MYL6, NDUFA11, LRPPRC*, and *SLC7A11*). Subsequently, RT-qPCR was employed to verify these genes in clinical specimens, and *LRPPRC* and *NDUFA11* were identified as the final two hub genes. *NDUFA11*, an adjunctive subunit within the mitochondrial membrane respiratory chain, assumes a vital role in the process of mitochondrial oxidative phosphorylation. It had been reported that the expression of *NDUFA11* varies in different tumours.[40,41] On the other hand, *LRPPRC* encodes a leucine-rich protein characterized by multiple pentatricopeptide repeats, and its precise function remains enigmatic. Investigative endeavors have indicated plausible involvement in various cellular processes, such as cytoskeletal organization, vesicular transport, and the regulation of nuclear and mitochondrial gene expression.[42] Notably, mutations within this gene have been causatively linked to the French-Canadian subtype of Leigh syndrome.[43] However, the specific roles of these two genes in CD require further investigation.

Furthermore, our investigation revealed a significant association between the expression of *NDUFA11* and Th17 cells. Researchers have discovered that individuals with CD exhibited elevated levels of IL17-producing T cells.[44] In several

studies, elevated levels of IFNγ and IL17 have been observed in CD patients, which is attributed to the presence of Th1 and Th17 cells.[45] Conversely, *LRPPRC* showed a pronounced inverse association with the infiltration of numerous immune cells, with a particular emphasis on monocytes. Monocytes play a dual role in CD, as impaired monocyte function can trigger CD, while monocyte hyperactivation and adaptive immunity contribute to disease maintenance.[46] Overall, our study highlights a strong positive association between *NDUFA11* and Th17 cells in CD. Additionally, *LRPPRC* exhibited a negative correlation with monocyte infiltration, suggesting its potential role in regulating monocyte activity in CD. These findings offer valuable insights into the complex immune mechanisms that underlie the pathogenesis of CD and may open avenues for targeted therapeutic strategies.

Traditionally, the CD has been systematically classified by employing the Montreal classification, which was devised in the year 1998 as a comprehensive refinement of the antecedent Vienna classification.[47] However, recent multi-omics data suggest that CD is a heterogeneous disorder with distinct molecular subtypes.[48–52] Relying solely on traditional typing is insufficient to accurately guide CD treatment and prognosis. We propose that more precise molecular subtype identification could enrich our comprehension of the heterogeneous molecular mechanisms underlying CD and aid in the development of personalized disease management and treatments through improved clinical trial design. In our study, unsupervised clustering methodology was deployed to partition patients into two distinct disulfidptosis subtypes, enabling the identification of new feature patterns for disease progression and potential treatment targets. Examining the correlation between disulfidptosis-related expression patterns and SES-CD, we found that patients in Cluster B exhibited reduced immune infiltration and a less severe clinical presentation. Subgroup analysis suggested that the identified expression pattern could predict disease severity in patients. In summary, our findings suggest the potential for a shift in the approach to managing CD, emphasizing the value of incorporating molecular subtype identification into clinical practice. This approach may enhance patient outcomes by supporting personalized disease management strategies and could help in designing more focused clinical trials. By considering molecular distinctions beyond traditional classification systems, we aim to contribute to a deeper understanding and more effective treatment strategies for CD, although further research is necessary to fully realize these possibilities.

However, our study does have certain limitations that should be addressed. Firstly, the data utilized in our analysis were sourced from public databases without direct access to the raw sequencing data. This situation could potentially introduce certain biases in our prediction results. Secondly, our verification of the two selected hub genes was limited to their mRNA expression levels, indicating the need for further in vivo or in vitro experiments to comprehensively elucidate their underlying mechanisms of action. Additionally, in the GSE112366 dataset, there are individuals who received Ustekinumab treatment, and its influence on gene expression remains unknown. Lastly, the presence of potential samples from the same individuals in GSE112366 raises concerns regarding the validity of our study's findings and warrants careful consideration.

Conclusion

Overall, our study has unveiled the critical roles of two hub DRGs, *LRPPRC* and *NDUFA11*, suggesting their potential involvement in the development and immune cell infiltration of CD. Moreover, our study represents the pioneering effort in integrating the disulfidptosis subtype pattern with immune infiltration and SES-CD scores in CD patients. Our study's findings offer valuable understanding into the molecular biology of DRGs in CD, potentially influencing their clinical management and treatment, while also paving the way for further research and targeted therapies that may lead to enhanced outcomes for patients with CD.

Data Sharing Statement

All the information utilized in this study is available in the public domain at the GEO database (https://www.ncbi.nlm.nih.gov/geo/). Access to the R code used in this research will be provided by the corresponding author upon a reasonable request.

Ethics Approval

The ethical aspects of this article have been duly addressed and complied with. Approval for the collection of clinical specimens was obtained from the clinical research institution review committee and ethics review committee of Zhongnan

Hospital (Scientific Ethical Approval No. 2019020). All patients signed an informed consent form, and all procedures involving human participants were conducted in accordance with the principles of the Declaration of Helsinki.

Acknowledgments

Mingyue Fu, Diliaremu Aihemaiti, and Haowen Fu are listed as co-first authors for this work. The authors extend their gratitude to all the investigators of the GEO for making the transcriptomic and clinicopathological data available, and we acknowledge the generous donation of tissue specimens by the patients.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Disclosure

The authors declare that they have no competing interests in this work.

References

1. Strober W, Fuss I, Mannon P. The fundamental basis of inflammatory bowel disease. *J Clin Invest*. 2007;117(3):514–521. doi:10.1172/jci30587
2. Ray K. IBD: the changing epidemiology of IBD. *Nat Rev Gastroenterol Hepatol*. 2017;14(12):690. doi:10.1038/nrgastro.2017.159
3. Agrawal M, Spencer EA, Colombel JF, Ungaro RC. Approach to the Management of Recently Diagnosed Inflammatory Bowel Disease Patients: a User's Guide for Adult and Pediatric Gastroenterologists. *Gastroenterology*. 2021;161(1):47–65. doi:10.1053/j.gastro.2021.04.063
4. Tang D, Kang R, Berghe TV, Vandenabeele P, Kroemer G. The molecular machinery of regulated cell death. *Cell Res*. 2019;29(5):347–364. doi:10.1038/s41422-019-0164-5
5. Moujalled D, Strasser A, Liddell JR. Molecular mechanisms of cell death in neurological diseases. *Cell Death Differ*. 2021;28(7):2029–2044. doi:10.1038/s41418-021-00814-y
6. Galluzzi L, Vitale I, Aaronson SA, et al. Molecular mechanisms of cell death: recommendations of the Nomenclature Committee on Cell Death 2018. *Cell Death Differ*. 2018;25(3):486–541. doi:10.1038/s41418-017-0012-4
7. Subramanian S, Geng H, Tan XD. Cell death of intestinal epithelial cells in intestinal diseases. *Sheng Li Xue Bao*. 2020;72(3):308–324.
8. Chen Y, Li X, Sun R, et al. A broad cuproptosis landscape in inflammatory bowel disease. *Front Immunol*. 2022;13:1031539. doi:10.3389/fimmu.2022.1031539
9. Yuan Y, Fu M, Li N, Ye M. Identification of immune infiltration and cuproptosis-related subgroups in Crohn's disease. *Front Immunol*. 2022;13:1074271. doi:10.3389/fimmu.2022.1074271
10. Liu X, Nie L, Zhang Y, et al. Actin cytoskeleton vulnerability to disulfide stress mediates disulfidptosis. *Nat Cell Biol*. 2023;25(3):404–414. doi:10.1038/s41556-023-01091-2
11. Zhao S, Wang L, Ding W, et al. Crosstalk of disulfidptosis-related subtypes, establishment of a prognostic signature and immune infiltration characteristics in bladder cancer based on a machine learning survival framework. *Front Endocrinol*. 2023;14:1180404. doi:10.3389/fendo.2023.1180404
12. Yang L, Liu J, Li S, et al. Based on disulfidptosis, revealing the prognostic and immunological characteristics of renal cell carcinoma with tumor thrombus of vena cava and identifying potential therapeutic target AJAP1. *J Cancer Res Clin Oncol*. 2023. doi:10.1007/s00432-023-04877-x
13. Wang T, Guo K, Zhang D, et al. Disulfidptosis classification of hepatocellular carcinoma reveals correlation with clinical prognosis and immune profile. *Int Immunopharmacol*. 2023;120:110368. doi:10.1016/j.intimp.2023.110368
14. Liu F, Yuan D, Liu X, et al. A demonstration based on multi-omics transcriptome sequencing data revealed disulfidptosis heterogeneity within the tumor microenvironment of esophageal squamous cell carcinoma. *Discover Oncol*. 2023;14(1):96. doi:10.1007/s12672-023-00711-5
15. Qi C, Ma J, Sun J, Wu X, Ding J. The role of molecular subtypes and immune infiltration characteristics based on disulfidptosis-associated genes in lung adenocarcinoma. *Aging*. 2023;15(11):5075–5095. doi:10.18632/aging.204782
16. Chen H, Yang W, Li Y, Ma L, Ji Z. Leveraging a disulfidptosis-based signature to improve the survival and drug sensitivity of bladder cancer patients. *Front Immunol*. 2023;14:1198878. doi:10.3389/fimmu.2023.1198878
17. Feng Z, Zhao Q, Ding Y, et al. Identification a unique disulfidptosis classification regarding prognosis and immune landscapes in thyroid carcinoma and providing therapeutic strategies. *J Cancer Res Clin Oncol*. 2023. doi:10.1007/s00432-023-05006-4
18. Di Narzo AF, Brodmerkel C, Telesco SE, et al. High-Throughput Identification of the Plasma Proteomic Signature of Inflammatory Bowel Disease. *J Crohn's Colitis*. 2019;13(4):462–471. doi:10.1093/ecco-jcc/jjy190
19. Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res*. 2013;41:D991–5. doi:10.1093/nar/gks1193
20. Chakraborty S, Datta S, Datta S. Surrogate variable analysis using partial least squares (SVA-PLS) in gene expression studies. *Bioinformatics*. 2012;28(6):799–806. doi:10.1093/bioinformatics/bts022
21. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47. doi:10.1093/nar/gkv007
22. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284–287. doi:10.1089/omi.2011.0118

23. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf*. 2013;14:7. doi:10.1186/1471-2105-14-7
24. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. *J Royal Statistical Society*. 1996;58(1):267–288. doi:10.1111/j.2517-6161.1996.tb02080.x
25. Noble WS. What is a support vector machine? *Nature Biotechnol*. 2006;24(12):1565–1567. doi:10.1038/nbt1206-1565
26. Breiman L. Random Forests. *Machine Learning*. 2001;45(1):5–32. doi:10.1023/A:1010933404324
27. Friedman JH, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*. 2010;33(1):1–22. doi:10.18637/jss.v033.i01
28. Karatzoglou A, Smola A, Hornik K, Zeileis A. kernlab - An S4 Package for Kernel Methods in R. *Journal of Statistical Software*. 2004;11(9):1–20. doi:10.18637/jss.v011.i09
29. Liaw A, Matthew G. Classification and Regression by randomForest. *R News*. 2002;2(3):18–22.
30. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics*. 2010;26(12):1572–1573. doi:10.1093/bioinformatics/btq170
31. Mitsialis V, Wall S, Liu P, et al. Single-Cell Analyses of Colon and Blood Reveal Distinct Immune Cell Signatures of Ulcerative Colitis and Crohn's Disease. *Gastroenterology*. 2020;159(2):591–608.e10. doi:10.1053/j.gastro.2020.04.074
32. Lin W, Liu S, Huang Z, et al. Mass cytometry and single-cell RNA sequencing reveal immune cell characteristics of active and inactive phases of Crohn's disease. *Front Med*. 2022;9:1064106. doi:10.3389/fmed.2022.1064106
33. Berahovich RD, Lai NL, Wei Z, Lanier LL, Schall TJ. Evidence for NK cell subsets based on chemokine receptor expression. *J Immunol*. 2006;177(11):7833–7840. doi:10.4049/jimmunol.177.11.7833
34. Andreu-Ballester JC, Catalán-Serra I, Gil-Borrás R, et al. Gammadelta T cells as a predictor of surgical relapse of Crohn's disease. *Clin Res Hepatol Gastroenterol*. 2020;44(4):586–597. doi:10.1016/j.clinre.2019.11.003
35. Jozawa H, Inoue-Yamauchi A, Arimura S, Yamanashi Y. Loss of C/EBPδ enhances apoptosis of intestinal epithelial cells and exacerbates experimental colitis in mice. *Genes Cells*. 2019;24(9):619–626. doi:10.1111/gtc.12711
36. Barrett JC, Lee JC, Lees CW, et al. Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. *Nature Genet*. 2009;41(12):1330–1334. doi:10.1038/ng.483
37. Anderson CA, Boucher G, Lees CW, et al. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nature Genet*. 2011;43(3):246–252. doi:10.1038/ng.764
38. Ijssennagger N, van der Meer R, van Mil SWC. Sulfide as a Mucus Barrier-Breaker in Inflammatory Bowel Disease? *Trends in molecular medicine*. *Mar*. 2016;22(3):190–199. doi:10.1016/j.molmed.2016.01.002
39. Machesky LM. Deadly actin collapse by disulfidptosis. *Nat Cell Biol*. 2023;25(3):375–376. doi:10.1038/s41556-023-01100-4
40. Tan M, Schaffalitzky de Muckadell OB, Jørgensen MT. Gene Expression Network Analysis of Precursor Lesions in Familial Pancreatic Cancer. *J Pancreatic Cancer*. 2020;6(1):73–84. doi:10.1089/pancan.2020.0007
41. Lin Y, Fan L, Zhang R, Pan H, Li Y. ARSD is responsible for carcinoma and amyloidosis of breast epithelial cells. *Eur. J. Cell Biol*. 2022;101(2):151199. doi:10.1016/j.ejcb.2022.151199
42. Tsuchiya N, Fukuda H, Nakashima K, Nagao M, Sugimura T, Nakagama H. LRP130, a single-stranded DNA/RNA-binding protein, localizes at the outer nuclear and endoplasmic reticulum membrane, and interacts with mRNA in vivo. *Biochem Biophys Res Commun*. 2004;317(3):736–743. doi:10.1016/j.bbrc.2004.03.103
43. Bchetnia M, Tardif J, Morin C, Laprise C. Expression signature of the Leigh syndrome French-Canadian type. *Mol Gene Metabol Rep*. 2022;30:100847. doi:10.1016/j.ymgmr.2022.100847
44. Feng T, Qin H, Wang L, Benveniste EN, Elson CO, Cong Y. Th17 cells induce colitis and promote Th1 cell responses through IL-17 induction of innate IL-12 and IL-23 production. *J Immunol*. 2011;186(11):6313–6318. doi:10.4049/jimmunol.1001454
45. Harbour SN, Maynard CL, Zindl CL, Schoeb TR, Weaver CT. Th17 cells give rise to Th1 cells that are required for the pathogenesis of colitis. *Proc Natl Acad Sci U S A*. 2015;112(22):7061–7066. doi:10.1073/pnas.1415675112
46. Zhou L, Braat H, Faber KN, Dijkstra G, Peppelenbosch MP. Monocytes and their pathophysiological role in Crohn's disease. *Cell. Mol. Life Sci*. 2009;66(2):192–202. doi:10.1007/s00018-008-8308-7
47. Satsangi J, Silverberg MS, Vermeire S, Colombel JF. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. *Gut*. 2006;55(6):749–753. doi:10.1136/gut.2005.082909
48. Mohammadi A, Kelly OB, Smith MI, Kabakchiev B, Silverberg MS. Differential miRNA Expression in Ileal and Colonic Tissues Reveals an Altered Immunoregulatory Molecular Profile in Individuals With Crohn's Disease versus Healthy Subjects. *J Crohn's Colitis*. 2019;13(11):1459–1469. doi:10.1093/ecco-jcc/jjz076
49. Sudhakar P, Verstockt B, Cremer J, et al. Understanding the Molecular Drivers of Disease Heterogeneity in Crohn's Disease Using Multi-omic Data Integration and Network Analysis. *Inflamm Bowel Dis*. 2021;27(6):870–886. doi:10.1093/ibd/izaa281
50. Kanke M, Kennedy Ng MM, Connelly S, et al. Single-Cell Analysis Reveals Unexpected Cellular Changes and Transposon Expression Signatures in the Colonic Epithelium of Treatment-Naïve Adult Crohn's Disease Patients. *CMGH*. 2022;13(6):1717–1740. doi:10.1016/j.jcmgh.2022.02.005
51. Weiser M, Simon JM, Kochar B, et al. Molecular classification of Crohn's disease reveals two clinically relevant subtypes. *Gut*. 2018;67(1):36–42. doi:10.1136/gutjnl-2016-312518
52. Kong L, Pokatayev V, Lefkovith A, et al. The landscape of immune dysregulation in Crohn's disease revealed through single-cell transcriptomic profiling in the ileum and colon. *Immunity*. 2023;56(2):444–458.e5. doi:10.1016/j.immuni.2023.01.002