Computational and Structural Biotechnology Journal 23 (2024) 1348–1363

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj

Review article

The molecular subtypes of autoimmune diseases

Xiangshu Cheng [1], Xin Meng [1], Rui Chen [1], Zerun Song, Shuai Li, Siyu Wei, Hongchao Lv, Shuhao Zhang, Hao Tang, Yongshuai Jiang [*], Ruijie Zhang [*]

College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, China

ARTICLE INFO

Keywords:
Autoimmune diseases
Molecular subtype
Disease heterogeneity
Precision medicine

ABSTRACT

Autoimmune diseases (ADs) are characterized by their complexity and a wide range of clinical differences. Despite patients presenting with similar symptoms and disease patterns, their reactions to treatments may vary. The current approach of personalized medicine, which relies on molecular data, is seen as an effective method to address the variability in these diseases. This review examined the pathologic classification of ADs, such as multiple sclerosis and lupus nephritis, over time. Acknowledging the limitations inherent in pathologic classification, the focus shifted to molecular classification to achieve a deeper insight into disease heterogeneity. The study outlined the established methods and findings from the molecular classification of ADs, categorizing systemic lupus erythematosus (SLE) into four subtypes, inflammatory bowel disease (IBD) into two, rheumatoid arthritis (RA) into three, and multiple sclerosis (MS) into a single subtype. It was observed that the high inflammation subtype of IBD, the RA inflammation subtype, and the MS "inflammation & EGF" subtype share similarities. These subtypes all display a consistent pattern of inflammation that is primarily driven by the activation of the JAK-STAT pathway, with the effective drugs being those that target this signaling pathway. Additionally, by identifying markers that are uniquely associated with the various subtypes within the same disease, the study was able to describe the differences between subtypes in detail. The findings are expected to contribute to the development of personalized treatment plans for patients and establish a strong basis for tailored approaches to treating autoimmune diseases.

1. Introduction

Autoimmune diseases (ADs) arise when the immune system erroneously targets and destroys healthy body cells. Currently, over 50 different ADs are recognized in the Medical Subject Headings (MeSH), with common types including rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), inflammatory bowel disease (IBD), and multiple sclerosis (MS). Although the mortality rate associated with ADs is significantly lower than that of cancer, the impact of ADs on individuals and society is no less than that of cancer. ADs often lead to various physical symptoms such as pain, fatigue, weight changes, and joint swelling [1], which can severely limit daily activities and diminish quality of life, affecting work, educational pursuits, household responsibilities, and leisure activities. In extreme cases, some ADs may cause serious health issues, including organ damage and dysfunction, and can become life-threatening. Furthermore, the long-term treatment of ADs increases public health costs, including medical services, disability benefits, and lost productivity; thus, imposing a considerable economic burden on healthcare systems [2].

Despite significant advancements in understanding human ADs in recent years, the precise causes of many such diseases remain elusive. ADs are characterized by a wide variability in pathophysiology, clinical presentations, and responses to treatment. For instance, SLE, known for its systemic involvement, exhibits a broad spectrum of clinical manifestations [3], from mild skin rashes to severe kidney damage [4]. Variability in SLE can be seen in the disease phenotype, as evidenced by the presence or absence of specific complications such as glomerulonephritis or neurological involvement [5], and in the disease course, with some patients experiencing relapses while others suffer from ongoing activity [6]. Studies have consistently shown SLE's heterogeneity at multiple levels, including serology, epigenetics, immunophenotyping, and biomarkers [7–11]. IBD, consisting mainly of Crohn's disease (CD)

* Correspondence to: College of BioinformaticsScience and Technology, Harbin Medical University, 157Baojian Road. Nangang District, Harbin, Heilongiiang Province, China.
E-mail addresses: jiangyongshuai@gmail.com (Y. Jiang), zhangrj@ems.hrbmu.edu.cn (R. Zhang).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.csbj.2024.03.026
Received 12 November 2023; Received in revised form 27 March 2024; Accepted 27 March 2024
Available online 28 March 2024

and ulcerative colitis (UC), represents a group of inflammatory conditions affecting the gastrointestinal tract [12]. Although CD and UC share certain clinical features, they are distinct in their pathogenesis, lesion locations, histology, and endoscopy findings, suggesting different underlying mechanisms [13]. Moreover, each subtype has its specific biomarkers, highlighting the heterogeneity within CD and UC [14]. For example, mucosal healing in CD patients treated with corticosteroids is achieved in less than one-third of cases, and large retrospective cohort studies support the effectiveness of methotrexate in maintaining clinical remission for approximately one-third of children with CD [15]. A ten-year follow-up study showed that about half of the CD patients developed severe stenosis or penetrating disease[16]. Similar heterogeneity has also been reported in RA, where differences in factors like age and environment contribute to variations in disease states and clinical phenotypes [17]. MS also shows heterogeneity in clinical manifestations and disease course predictability [18]. Furthermore, there exists heterogeneity in the response to drug treatment across autoimmune diseases. The limitations of traditional immunomodulatory drugs used to treat autoimmune diseases lie in their broad and nonspecific nature. Patients with similar clinical symptoms might have different responses to the same medication, necessitating individualized drug selection and dosage adjustment [19]. Therefore, revealing the heterogeneity of autoimmune diseases is important for exploring disease diversity.

With the emergence of precision medicine, there has been an increased emphasis on addressing the complexity and diversity of diseases. Transcriptomic information has provided an extensive range of biomedical insights for investigating disease mechanisms, identifying clinical diagnostic markers, and discovering drug targets. The advancements in molecular technologies have significantly influenced the understanding of autoimmune diseases. Along with technological advances, the study of heterogeneity in autoimmune diseases has gone through the following stages: I. Researchers employed a low-throughput gene expression assay, RT-qPCR, to assess gene expression levels. Despite its lower throughput, RT-qPCR's high accuracy has been instrumental in examining the heterogeneity of autoimmune diseases. For instance, a study utilizing RT-qPCR demonstrated that the origin of different stromal cells (synoviocytes or skin fibroblasts) contributes to variability in cytokine production (such as IL-23, IL-17, etc.) and receptor expression. This variation could account for the differing responses to IL-23 or IL-17 inhibitors across various autoimmune diseases [20]. II. As technology advanced, high-throughput microarray technologies, capable of measuring extensive gene expression data, have been extensively applied to explore the heterogeneity in autoimmune diseases. For example, Daniel Toro-Domínguez et al. used microarray data alongside disease activity scores to thoroughly investigate SLE heterogeneity, eventually classifying SLE patients into three distinct subtypes [21]. III. RNA-seq technology, offering higher resolution than microarry, enables the identification of new transcripts and splice variants, thereby delivering more detailed and comprehensive insights for the molecular phenotyping of autoimmune diseases. The surge in data volume has also expanded the community of researchers dedicated to studying the heterogeneity of autoimmune diseases. A Bulk RNA-seq-based study including seven systemic autoimmune diseases (including systemic lupus erythematosus, RA, SSC, pSjS, MCTD, PAPS, and UCTD) identified four unique subtypes [22]. Furthermore, single-cell RNA sequencing technologies have unveiled the transcriptome's heterogeneity and complexity at the individual cell level, significantly advancing autoimmune disease research [23]. However, gene expression analysis involves more than just quantification; it often necessitates considering the spatial aspects of gene expression. Recent contributions from spatial transcriptomics technology to autoimmune disease research include a study on psoriasis, which revealed distinct molecular profiles between patients with mild and severe forms of the disease [24]. It is anticipated that novel transcriptomic molecular techniques will introduce fresh perspectives and methods, along with more precise results, for the classification studies of autoimmune diseases.

Therefore, in this study, we have reviewed the process and results of pathological and molecular typing studies. To define the scope of our research, a comprehensive literature search was performed on PubMed, employing 26 common autoimmune diseases as keywords, in addition to terms like "molecular subtype", "classification", "subtype", and "heterogeneity". Our selection criteria included diseases with at least two subtypes within the same disease category, enabling us to summarize and analyze the commonalities and heterogeneities of these diseases. Among them, 15 studies met our criteria, covering four types of autoimmune diseases: systemic lupus erythematosus (SLE), rheumatoid arthritis (RA), multiple sclerosis (MS), and inflammatory bowel disease (IBD). These four diseases are common and representative autoimmune diseases. Over the past 20 years, the incidence of these diseases has increased. In addition, patients with these diseases are more prone to complications from other autoimmune diseases, which significantly affects their quality of life [25]. These diseases involve various tissues and organs of the human body and have diverse manifestations. In the context of autoimmune diseases (ADs), significant heterogeneity among patients, even with the same disease, has garnered considerable attention from researchers. Recently, there has been a surge in studies focusing on the classification of these diseases, leading to an accumulation of substantial transcriptomic data and typing results. Therefore, summarizing and examining the research findings on the molecular subtypes of these four diseases, and analyzing the heterogeneity and similarities among different subtypes, hold significant importance for a deeper understanding of the pathogenesis of ADs and for advancing precise treatment strategies. In summary, this study selected four representative ADs with the goal of synthesizing existing research outcomes on the molecular subtypes of autoimmune diseases. The investigation began by reviewing the timeline of the pathological classification process for diseases such as MS and LN. Compared to pathological classification, molecular subtype methods offer advantages like objectivity, stability, and assistance in exploring mechanisms of disease onset and development. Consequently, this study outlines the general procedures for molecular typing studies of ADs, the commonly utilized data, and the biological characterization of ten molecular subtypes across four ADs to examine the heterogeneity within subtypes of the same disease and some common pathogenic mechanisms across many ADs. This research further characterizes different subtypes of the same disease, as well as similar subtypes across various autoimmune diseases, based on markers for each subtype. Lastly, it summarizes the drugs effective against each subtype. This study may provide valuable insights into disease treatment and offer a relatively generalized framework for professionals in related fields.

2. Pathological classification methods for autoimmune diseases

The clinical manifestations of ADs can vary widely among individuals, and the absence of precise assessment methods for disease course complicates the diagnosis, treatment, and research of these conditions. The emergence of pathological classification has allowed clinicians to employ unified terminology to describe disease progression. Thus, summarizing progress in pathological classification research of ADs is beneficial for understanding and diagnosing ADs. The focus here is primarily on summarizing pertinent findings of pathological subtyping in MS and LN.

In 1996, the National Multiple Sclerosis Society (NMSS) Advisory Committee established the first clinical subtyping criteria for MS, suggesting six concepts to unify the clinical manifestations of MS. Relapsing-remitting (RR) MS is characterized by abrupt impairments in brain function followed by varied degrees of recovery and stability between episodes. Primary-progressive (PP) MS progresses from onset, sometimes with stable periods and temporary minor improvements. Secondary-progressive (SP) MS progresses after the initial RR disease course, with or without occasional recurrences, mild remissions, and

periods of stability. Progressive-relapsing (PR) MS advances from the onset, with several acute relapses, with or without complete recovery. Benign MS sees patients retaining full nervous system functions 15 years after disease onset. Malignant MS progresses rapidly, causing severe impairment of multiple nervous system functions or death within a short period [26]. With advances in imaging and molecular marker technologies, Lublin et al. updated the 1996 criteria in 2013 to include two new disease processes — clinically isolated syndrome and radiologically isolated syndrome (CIS and RIS) — and incorporated disease activity (imaging and relapses) and progression in the classification criteria [27]. Although the Lublin criteria have been widely used in clinical settings, they do not accurately describe disease subtypes in early relapsing or late progressive MS patients [28–33]. To address this, David Pitt et al. proposed a pathology-based approach in 2022, centered around the concept of 'pathological axes'. This framework aims to define MS based on the extent and nature of pathological processes rather than solely on clinical presentations [34]. This approach quantifies the activity and type of each pathological process to classify MS patients. However, the limited availability of clinical biomarkers restricts the clinical application of this method.

LN, a major cause of death and disability among patients with SLE, was initially classified by the World Health Organization (WHO) in 1975 and 1978 into five classes based on clinical renal biopsy findings and glomeruli involvement [35]. Class I showed no detectable glomerular abnormalities under the microscope, while Class II involved purely mesangial immune deposition, subdivided further depending on mesangial hypercellularity presence. Class III involved less than 50% of proliferative glomerulonephritis, Class IV more than 50%, and Class V was defined as Membranous lupus nephritis. The International Study of Kidney Diseases in Children (ISKDC) revised the WHO criteria in 1982, removing the glomerular involvement criteria and further subclassifying the disease into six categories [36]. In 2003, the International Society of Nephrology/Renal Pathology Society (ISN/RPS) introduced new criteria that clarified the ambiguous aspects of the WHO classification, refined the definitions of activity and chronicity, and underscored the corresponding clinical lesions[37]. In 2018, the RPS updated the prior version of the criteria by reviewing and refining each category based on the earlier standards and incorporated the activity index (AI) and chronicity index (CI) scoring[38]. Despite these advancements and the broad acceptance of these criteria, their reliance mainly on light microscopy observation rather than on the underlying pathological mechanisms remains a limitation[39].

We have presented a chronological overview of the research progress in the pathological classification of MS and LN (Fig. 1), highlighting the significant value of these pathological classification methods in understanding and diagnosing ADs. However, conventional methods of categorizing autoimmune diseases primarily depend on pathological classification, which entails the observation of clinical manifestations, anatomical pathology changes, and other features for disease categorization. This approach has several drawbacks. First, as this classification method depends on the clinical manifestations of the disease, different clinicians may introduce subjectivity in its classification [27]. This may lead to unclear allocation of specific pathological subtypes for patients with less obvious symptoms. Second, pathological classification methods often overlook the molecular mechanisms underlying diseases, which are crucial for understanding, preventing, and treating diseases. In contrast, molecular classification methods address some of these limitations and possess objective and stable characteristics. Molecular classification categorizes diseases based on gene expression, protein composition, and other molecular features. This approach provides a novel perspective for understanding and classifying autoimmune diseases. This approach enables the exploration of the fundamental biological processes of diseases and the identification of potential mechanisms that pathological classification alone cannot achieve. Furthermore, molecular classification aids in a deeper understanding of disease progression, prediction of disease outcomes, and formulation of more accurate treatment strategies. Hence, molecular classification is gaining increasing significance in autoimmune disease research.

3. Molecular level classification methods of autoimmune diseases

The diversity of autoimmune diseases (ADs) is apparent not only in the clinical symptoms observed across different individuals but also at the molecular level. The aim of subtyping diseases at the molecular level is to uncover the biological mechanisms underlying these variations and to facilitate the classification of patient samples. Considering the variety of methods used by researchers and the different perspectives from which molecular subtypes of ADs have been studied, current research presents a wide array of study designs and results, utilizing various bioinformatics tools and techniques. This review offers a detailed overview of the typical process and the data commonly used in existing studies, focusing on the crucial steps, methods, and tools employed in research design.

3.1. Common data types for autoimmune classification

Transcriptome data, which have the advantage of high density, have become the main type of data used for many disease subtype classification studies. Nevertheless, differences exist in the transcriptome data utilized for investigating subtypes of autoimmune diseases due to variations in the affected sites of different diseases, the study period, and the direction of subtype classification.

The largest number of studies on SLE has been published, and most of these studies involve data. From 2016 to 2022, a total of six SLE classification studies covering 6172 patients were published [21,40–44]. Most of the data can be downloaded from the GEO database, with a

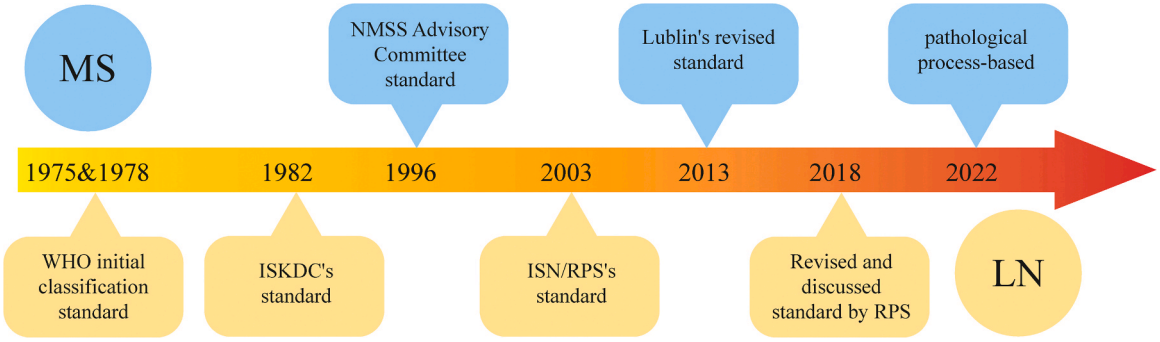

Fig. 1. The pathological classification of MS and LN. The blue part represents MS, and the yellow part represents LN. MS: multiple sclerosis. LN: lupus nephritis. WHO: World Health Organization. NMSS: National Multiple Sclerosis Society Advisory Committee. ISKDC: the International Study of Kidney Diseases in Children. ISN/RPS: International Society of Nephrology/Renal Pathology Society.

small amount coming from the PRECISESADS IMI project [43]. In recent years, with the development of single-cell transcriptomes, the classification of ADs has been combined with single-cell data to explore the molecular subtypes and subtype-specific cells of patients with SLE [44]. Given that SLE is a complex disease affecting multiple systems and organs and considering the diverse clinical manifestations and organ damage observed in different patients, it is challenging to establish a single organ or tissue as a universal standard for sequencing. Nonetheless, as nearly all SLE patients undergo hematological changes, current studies on SLE typing predominantly rely on data derived from blood samples (Table 1).

RA, an autoimmune disease, is mainly characterized by joint damage. The data for classification studies often come from synovial biopsies or blood tests of the patients. Initial studies lacked healthy control samples for RA, leading researchers to use sequenced samples from osteoarthritis(OA) patients as controls for the analysis [45]. Subsequent studies, benefiting from a larger collection of sequencing samples, some studies still retain OA data and use normal samples and OA as controls to identify RA-specific DEGs [46] (Table 1).

IBD, is a chronic gastrointestinal inflammatory disorder influenced by multiple factors. Intestinal mucosal biopsy, particularly from the organ most affected, has become the primary data source for IBD typing studies. Depending on the IBD pathological classification, different sections of the intestine are selected for mucosal collection. For UC, the sigmoid colon, located 15 to 20 cm from the anal verge, is often chosen for biopsy. For CD, biopsies are usually taken from the ileum site (Table 1).

MS predominantly impacts the optic nerve, brain, and cervical spinal cord's white matter. The unique nature of the pathogenesis sites makes sampling challenging. Currently, the primary data source for MS typing studies is microarray data from patients' blood samples [47] (Table 1).

3.2. The general process for subtyping of autoimmune diseases

The traditional classification of ADs includes five steps: data processing; feature matrix generation; clustering; subtype functional definition, and classifier construction, as depicted in Fig. 2.

3.2.1. Data processing

During the initial analysis step, data processing commonly encompasses quality control, batch effect removal, probe annotation, data normalization, and data scaling. After eliminating abnormal data and ineffective samples, the final gene expression profile was obtained, which served as the raw material for subsequent classification analysis.

Table 1
The SLE, RA, IBD, and MS classification of research data statistics.

Disease	PMID	Patients	Normal	Data Type	Source
SLE	27040498	924	72	Microarray	BLOOD
	29938934	1216	92		
	34305454	1123	102		
	35858908	404	43		
	35947992	6232	791		
		99	18	RNA-seq	
IBD	36923403	362	26	Microarray	ILEUM
		268	42		COLON
	36845139	630	68		
	37443022	208	0	Microarray	
	27742763	21	11	RNA-seq	ILEAL
RA	30831253	300	40(OA 51)	Microarray	SYNOVIAL
	33230538	62	0		
		190	28	RNA-seq	
	29468833	123	0(OA 6)		
	31461658	87	0		
		67	0		BLOOD
MS	16837931	363	0	Microarray	BLOOD
	23019656	29	28		

SLE: Systemic Lupus Erythematosus; RA: Rheumatoid Arthritis; IBD: Inflammatory Bowel Disease; MS: Multiple Sclerosis.

3.2.2. Feature matrix generation

Upon acquiring the gene expression profile, researchers often select one or more features pertinent to the disease or their research interests, such as differentially expressed genes (DEGs), disease activity indices, and specific biological markers like neutrophil content. Methods like differential analysis, correlation calculations, and Weighted Gene Co-expression Network Analysis(WGCNA)[48] are employed to create feature-related matrices from these features. WGCNA builds an undirected weighted co-expression network from gene expression data to identify gene modules with high cooperation. It is frequently used to pinpoint candidate genes associated with phenotypes through the module-phenotype relationships. Notably, some studies bypass the second step and directly employ the gene expression profile post-data processing as a feature matrix for the clustering analysis.

3.2.3. Clustering

Clustering aims to categorize a dataset into distinct groups, grouping similar items together to simplify complex data structures. This method can identify the disease subtypes. Researchers apply various clustering algorithms to divide complex data with unknown class labels into similar groups, which often display molecular-level differences such as variations in immune cell infiltration, dysregulated genes, and enriched pathways. The heterogeneity within these groups may indicate different pathogenic factors, disease mechanisms, and varying drug response and treatment effectiveness within the same disease. Thus, precise clustering of disease data is vital for developing more effective, personalized treatment strategies.

Three essential steps condense cluster analysis. Initially, evaluating the dataset's clustering tendency is imperative, determining if the data intrinsically possesses a meaningful clustering structure. This evaluation is critical as clustering methods can produce results even without clear inherent clusters. The Hopkins statistic is used to assess clustering propensity. Secondly, identifying the optimal cluster number is essential. Various statistical metrics, including gap statistics [49], silhouette coefficients [50], sum of squared error (SSE), Bayesian information criteria (BIC), cumulative distribution function (CDF), and Calinski-Harabasz index, assist in determining the most appropriate cluster number. The R package NbClust [51], provides 30 indicators to help select the optimal number. Finally, choosing the suitable clustering algorithm for implementation follows the establishment of the optimal number of clusters. These algorithms include K-means clustering, hierarchical clustering, consensus clustering, non-negative matrix factorization (NMF), and affinity propagation(AP) [52], with several R packages such as ConsensusClusterPlus [53], mclust [54], NMF [55], CrossICC [56], and apcluster [57], facilitating these steps' practical application.

The significance of clustering trends in disease data and the chosen clustering algorithms profoundly affects classification outcomes. In each included study, researchers utilized specific methods to test the classification's accuracy and reliability. I. Employing multiple clustering methods to achieve consistent typing results enhances result reliability. For instance, Cui et al. determined the optimal number of subtypes using three unsupervised clustering methods (NbClust, PAM, and Vegan) after identifying labeled genes for subtype classification and assigning category labels to each sample [41]. II. Statistical metrics determine the optimal subtyping results. After subtyping, metrics such as *gap statistics* [49]*, silhouette coefficients, cumulative distribution function (CDF)*, etc., were used to determine the optimal subtyping results. For example, a study classifying systemic lupus erythematosus (SLE) subtypes based on neutrophils utilized CDF for optimal subtype determination[42]. III. Subtype results are validated using validation sets. In a study on ulcerative colitis (UC), iterative consensus clustering (crossICC) categorized patients into three subtypes using bulk transcriptomic data from three sources. The genes characterizing these subtypes were validated in three other independent bulk transcriptome datasets, showing stable

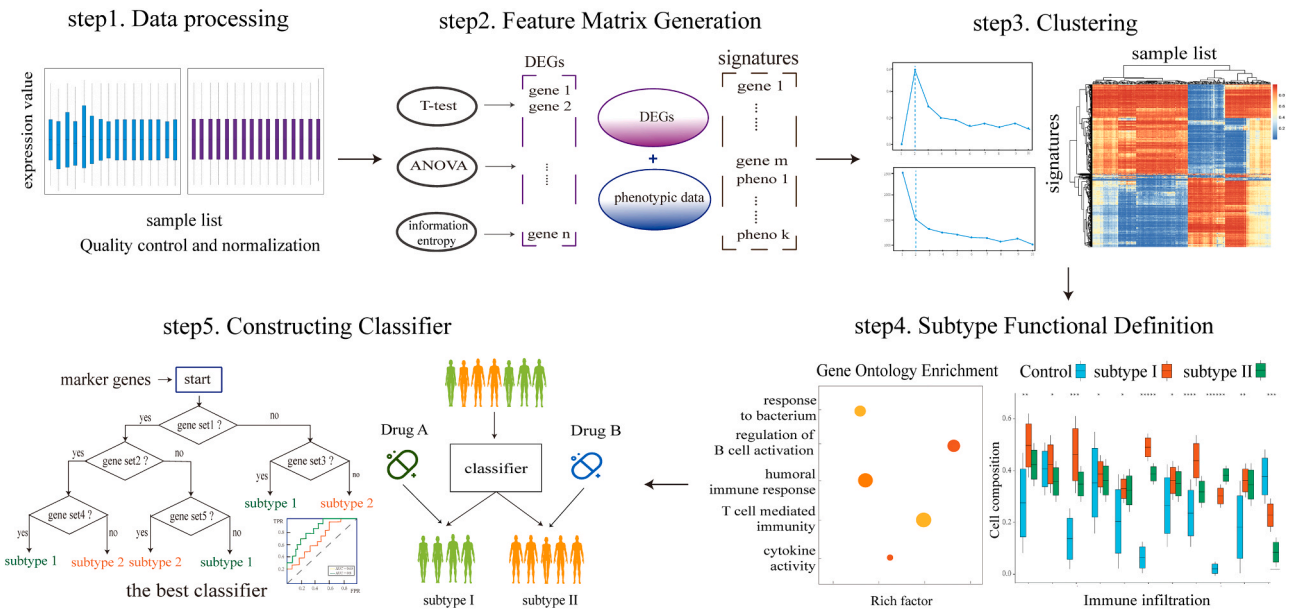

Fig. 2. Common flow chart for subtype classification. The traditional subtype division procedure is divided into the following five steps. Step 1. Data Preprocessing: Perform data quality control and standardization; Step 2. Feature Matrix Generation: Choose crucial biomarkers for subtype classification; Step 3. Clustering: Employ clustering algorithms to determine distinct disease subtypes; Step 4. Subtype functional definition: Analyze the functions and heterogeneity of each subtype; Step 5. Classifier construction: Construct classifier models for precision medicine and drug response prediction.

performance across the validation sets. This result ensures the stability and accuracy of these three isoforms in the disease population [58]. Therefore, we recommend the use of multiple clustering methods for subtyping and incorporating validation mechanisms to determine reliable subtyping results. This approach not only would help assess data quality but also ensure the reliability of the subtypes produced by the chosen clustering algorithm.

3.2.4. *Subtype functional definition*

After identifying the disease subtype, understanding its function and heterogeneity becomes crucial. This insight aids in exploring disease pathogenesis and informs personalized medicine in drug development. Specifically, the researchers will first identify gene sets specific to each subtype based on differential expression analysis, correlation analysis, and WGCNA. Here, the researchers entered the subtype as phenotypic information into the WGCNA [48] to identify the subtype-specific functional modules that make up the specific gene set of the subtype. Based on these gene sets, researchers have conducted functional enrichment analyses, including overrepresentation analysis (ORA), gene set enrichment analysis (GSEA)[59], and gene set variation analysis (GSVA)[60]. Immune infiltration analyses were also performed using tools such as CIBERSORTx[61] and IOBR[62] to delve deeper into the functional characteristics and heterogeneity of the subtypes.

3.2.5. *Constructing the classifier*

Finally, researchers have applied methods such as decision trees, logistic regression, naive Bayes, support vector machines (SVMs), and neural networks to construct subtype classifier models or drug response prediction models. These models assist in identifying disease subtypes and predicting aspects such as drug responses, recurrence, and remission for newly diagnosed patients. Integrating subtype research findings with clinical applications enhances our understanding of disease characteristics, paving the way for tailored medical services for patients.

3.3. *Experimental design for different disease classifications*

3.3.1. *Feature selection*

Selecting features for disease subtype classification is essential prior to clustering. Selecting all genes for classification is impractical from both biological and algorithmic perspectives. Moreover, choosing genes that do not reflect disease heterogeneity can obscure the true disease subtypes, yielding inaccurate results. Hence, employing feature selection to extract and transform expression profile information is crucial. In feature selection, researchers pick features that are either strongly associated with disease onset and progression or capable of distinguishing subtypes. Bioinformatics techniques then convert the expression profile into a feature matrix linked to selected features. Clustering this feature matrix identifies the disease subtypes. The feature matrix plays a pivotal role in determining the final disease subtype outcome, forming the foundation of subtype studies.

In order to retain the expression profile information, certain studies opted not to perform additional operations on the expression profile, instead utilizing the preprocessed expression profile directly as a matrix for clustering. For example, peripheral blood mononuclear cell data was used by Ottoboni et al. to classify all genes in the expression profile as features of disease subtypes. In 141 untreated patients, two disease subtypes (MSA and MSB) of multiple sclerosis (MS) exhibiting varying disease activity were identified[47]. Another study also used peripheral blood profiles from relapsing-remitting multiple sclerosis (RRMS) patients to classify them into two distinct MS subtypes [63].

Currently, the generation of feature matrices predominantly relies on DEGs between the disease and normal control. In one investigation, the DESeq2 package [64] was utilized to identify 1241 upregulated DEGs in the synovial tissue expression profiles of patients with RA compared to healthy controls. These DEGs served as features for subtyping, leading to the delineation of three distinct disease subtypes [65]. In cancer tissue typing research, the most variably expressed features are often employed for unsupervised subdivision into distinct subtypes [66]. However, due to the inherent variability of the immune system [67], this approach is not suitable for blood samples from patients with immune disorders. In blood samples, characteristics with the highest variability may relate to processes not associated with systemic autoimmune diseases (SADs), such as infections. Barturen et al. opted to concentrate on DEGs and Differentially Methylated CpGs (DMCs), disregarding highly variable features unrelated to the pathology under examination. A total of 1821 DEGs and 4144 DMCs were selected to categorize seven SADs

into four subtypes [22].

To screen important DEGs more precisely, researchers commonly integrate data like disease activity indicators or protein interaction networks to create a feature matrix for distinguishing disease subtypes. For example, Cui et al. identified 92 DEGs associated with disease activity by assessing the correlation between disease activity score and longitudinal gene expression data. These DEGs were then utilized as identifying markers for SLE [41]. Another study screened and clustered genes linked with the SLEDAI and revealed three subtypes [21]. Based on the functional enrichment analysis of 876 genes that were differentially expressed between the RA group and the control group and between the RA group and the OA group, Kim et al. constructed protein—protein interaction networks using open databases such as BIOGRID[68], HPRD[69], IntAct[70], Reactome [71], and STRING[72] to analyze PPIN. They further identified 56 genes through the network centrality analysis [46].

Furthermore, two studies have classified disease subtypes based on neutrophil abundance, given their crucial role in initiating and sustaining intestinal inflammation. Neutrophils, by generating chemokines and reactive oxygen species (ROS), disrupting the epithelial barrier, recruiting and activating other immune cells, and triggering redox-sensitive inflammatory pathways [73], may exacerbate intestinal inflammation and contribute to the formation of malignant lesions through the release of neutrophil extracellular traps (NETs) [74]. NETs may also intensify autoimmune diseases like rheumatoid arthritis and systemic lupus erythematosus [75]. Thus, in a study on ulcerative colitis (UC) subtype study, researchers determined subtypes based on neutrophil-associated differential expression genes [76]. Initially, the limma [77] tool was used to identify DEGs between disease and normal controls. Subsequently, CIBERSORT was applied to estimate the proportion of patients with immune cell infiltration, and these findings were used as phenotypic input for WGCNA. Following WGCNA analysis, a red module associated with neutrophils was identified, and genes within this module were intersected with DEGs to define subtype features. In another study, SLE patients were classified into two subtypes: high neutrophilic (NEUT_H) and low neutrophilic (NEUT_L), based on neutrophil levels. Initially, ssGSEA analysis with the immune cell gene set [78] was conducted to assess enrichment scores for various immune cell subpopulations. Data from the Immunology Database and Analysis Portal (ImmPort; https://immport.niaid.nih.gov)[79] was then used to acquire immune-related genes. These gene expression profiles were the input for WGCNA, combined with ssGSEA results as phenotypic data for each immune cell. Through this process, functional modules with strong correlations to neutrophils were identified. This methodology was applied across four independent datasets, identifying eight features from the intersection of neutrophil-related functional modules to classify disease subtypes [42]. In addition, when we have identified the functional modules of a specific cell/pathway, we can also use Cytoscape's CytoHubba [80] for further characterization screening.

Beyond common features, some studies construct specific typing matrices for analysis by creating scores and reducing dimensionality. For example, MyPROSLE, a molecular patient portrait model by Toro-Dominguez et al., involved collecting data from 6134 whole-blood SLE patients and 757 healthy controls across 10 cohorts. Utilizing the Tmod R package [81], and the premise that genes may co-express in specific biological functions, gene expression data at the individual level was summarized into 606 co-expressed gene modules related to regulatory biology and immune mechanisms. To quantify each gene module's deviation from the normal state in patients, the module score (M-score) was calculated. Disease subtypes were classified using the M-score matrix, resulting in the identification of two clinically and molecularly distinct SLE subtypes [43]. Additionally, a study employing K-means clustering and module partitioning to map the blood fingerprint selected the module most closely associated with the disease activity index. Subsequently, the correlation coefficient matrix of this module and blood module was obtained, and hierarchical clustering of the correlation coefficient matrix was performed to identify disease subtypes [40].

3.3.2. Functional definition

For the functional definition, researchers typically employ functional enrichment, immune infiltration, and pathway activity assessment to unravel the pathogenesis and define the functions of each subtype. If different subtypes do not provide distinct functional definitions, classification efforts are reassessed. This reevaluation aids in the selection of subtype features and refines the classification approach. In this section, we outline the methodologies employed to define functions and characterize heterogeneity in the typing studies of ADs (Table 2).

4. The review of the molecular subtypes biological results for the four ADs

Currently, due to variations in the research timeline, data selection, and methodology, different studies examining the subtypes of the same disease may produce different results. Nevertheless, we observed consistency among the various subtyping results for the same disease. Therefore, in this review, we compiled and summarized the molecular subtyping results for four ADs (SLE, IBD, RA, and MS), focusing on the same disease, to explore the similarities and connections between different studies.

4.1. Systemic lupus erythematosus

SLE is a complex autoimmune disease characterized by immune system dysregulation, including aberrant immune activation, impaired regulatory cell function, antinuclear antibody production, and cytokine balance disturbances. These factors significantly contribute to SLE's onset and progression. Molecular classification studies have refined SLE's different functional subtypes, enhancing the understanding of disease mechanisms. This summary outlines common characteristics among SLE functional subtypes, categorizing SLE into four primary functional subtypes: neutrophil, interferon, lymphocyte, and plasma cell subtypes.

4.1.1. Neutrophil subtype

Neutrophils, as the most abundant immune cells in humans, play a crucial role in inflammation and various systemic autoimmune diseases. Recent studies have shown that neutrophils, especially low-density granulocytes (LDGs), are key in triggering autoimmune responses and organ damage in SLE patients[47]. Neutrophil subtypes were described

Table 2
The tools used for functional definition and heterogeneity of subtypes in different studies.

Disease	PMID	Defining Function / Characterizing Heterogeneity
SLE	27040498	modular analysis[82]; QuSAGE[83]
	29938934	GO[84]; EnrichR[85]; CIBERSORT[86]
	34305454	GO[84];gene expression deviation (GED)
	35858908	ssGSEA; GO[84]; KEGG[87]
	35947992	modular analysis
IBD	36923403	GSVA[60]; CIBERSORT[86]; IOBR[62]
	37443022	GO[84]; KEGG[87]; CIBERSORT[86]; GSVA[60]
	27742763	Reactome[71]; GSAA[88]
	36845139	GSVA[60]; ssGSEA; ESTIMATION
RA	30831253	DAVID[89]; GSEA[59]
	33230538	GSEA[59]; GO[84]; KEGG[87]; xCell[90]; ssGSEA
	29468833	DAVID[89]; CIBERSORT[86]
	31461658	Single cell annotation; Ingenuity Pathway Analysis (IPA); Reactome[71]; KEGG[87]
MS	23019656	Ingenuity Pathway Analysis (IPA http://www.ingenuity.com)
	16837931	PANTHER[91]; GSEA[59]

SLE: Systemic Lupus Erythematosus; RA: Rheumatoid Arthritis; IBD: Inflammatory Bowel Disease; MS: Multiple Sclerosis.

in five SLE-related categorization studies[21,40–43], and numerous investigations found a strong connection between the neutrophil subtype and immunological, inflammatory, bacterial infections, and renal manifestations, such as hematuria or proteinuria, particularly linked to severe nephritis risk. This subtype is associated with platelets and disease progression through various cytokines, including BCL, IL1RA, MMP8, IL6, TGF-β, and BAFF, with a positive correlation between SLEDAI scores and neutrophil percentage. Elevated neutrophil counts in remission patients may increase relapse risk. Additionally, in a subtyping study[42], neutrophil subtypes exhibited relatively less $CD8^+$ T cells infiltration than other subtypes. Researchers believe that a reduction in $CD8^+$ T cells may weaken the T-cell response to SLE infection, thereby exacerbating the severity of SLE. (Fig. 3-A).

4.1.2. Interferon subtype

Type I interferons, a class of cytokines, play a crucial role in the human immune response. Type I interferon, a major pathogenic factor in SLE, is associated with more active disease levels and severe manifestations, like nephritis [92]. Two studies [40,41]. defined interferon-related subtypes, one identifying an IFN-centric viral infection-related pattern due to strong dysregulation in virus-related modules. Further research into gender and age revealed a higher prevalence of the interferon subtype in children and women, potentially due to children's PBMC increased sensitivity to viral RNA transfection by the influenza virus and the effect of female estrogen secretion on SLE activity. Additionally, ethnic differences were noted, with black patients showing a stronger tendency towards the interferon subtype [41]. Despite limited studies directly defining interferon subtypes, interferon signatures are commonly identified in other SLE typing studies. In the Systemic Lupus Erythematosus (SLE) typing study published in 2022 [43], findings revealed dysregulation of interferon along with several other factor modules. It was observed that significant dysregulation of neutrophilic granulocyte/inflammation, plasma cell/cell cycle, and platelet signals coincided with high dysregulation of interferon signals, but not in the absence of such conditions. In other words, interferon signatures can be dysregulated alone, but when other signatures are dysregulated, interferon is co-dysregulated. This result may explain why interferon dysregulation is a common feature in SLE. Another study on SLE[93] focused on the Type I interferon-stimulating gene (ISG), observing that neutrophils and low-density granulocytes (LDGs) demonstrated increased ISG activity, potentially offering evidence for a pattern of co-dysregulation between interferon and neutrophils (Fig. 3-B).

4.1.3. Lymphocyte subtype

Lymphocytes, including T, B, and NK cells, are central to immune response regulation and execution. Abnormalities in lymphocyte function in SLE patients have provided targets for treatment and intervention. For instance, the T cell-related cytokine IL-2 is a therapeutic target for SLE and other ADs[94]. Studies investigating B cells have become a prominent area of SLE research. This focus has led to substantial progress in developing B cell-targeted therapies for SLE, exemplified by Belimumab, a BAFF inhibitor [95], and Telitacicept, a B lymphocyte stimulator inhibitor [96]. Dysregulated genes on NK cells are also potential new therapeutic targets for SLE[97]. In conclusion, lymphocytes play a critical role in both the development and treatment of SLE.

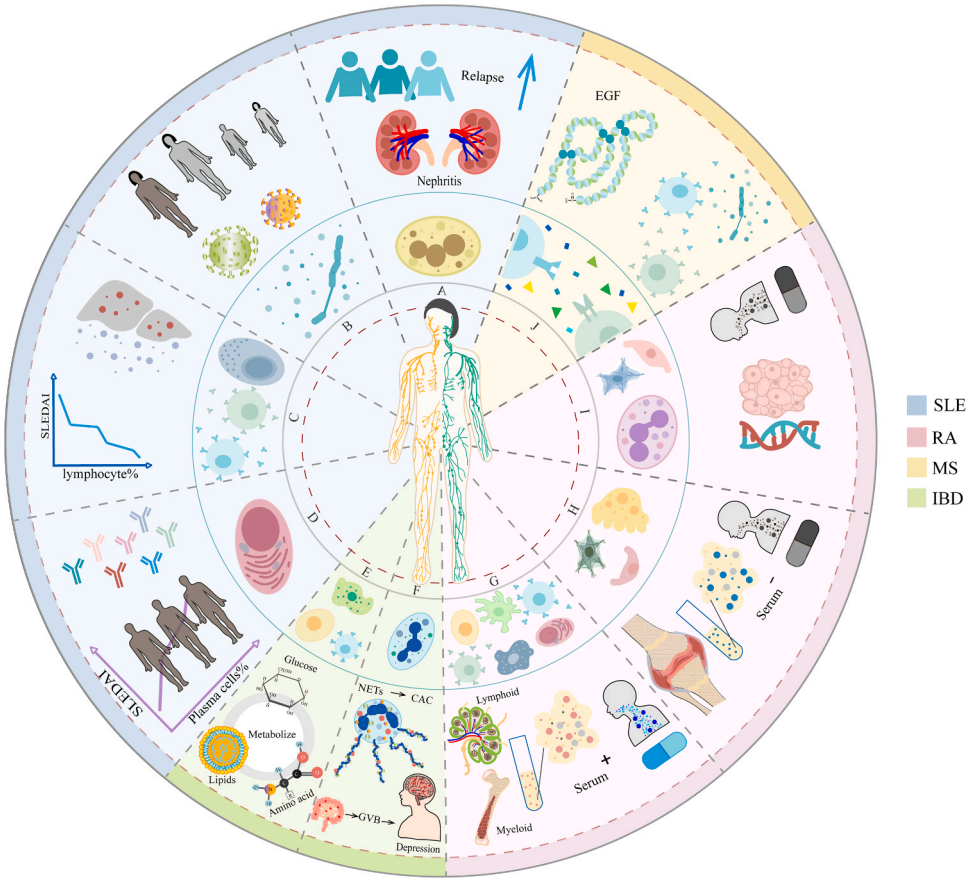

Fig. 3. Summary of the biological conclusions of ADs subtypes research. The figure describes the biological characteristics of each disease subtype at the molecular, cellular, organ and other levels. The blue part represents SLE (A: Neutrophil subtype, B: Interferon subtype, C: Lymphocyte subtype, D: Plasma cell subtype). The green part represents IBD (E: High metabolism subtype, F: High inflammation subtype); The pink part represents RA (G: Inflammation subtype, H: Joint damage subtype, I: Neutrophil subtype); The yellow part represents MS (J: Inflammation & EGF).

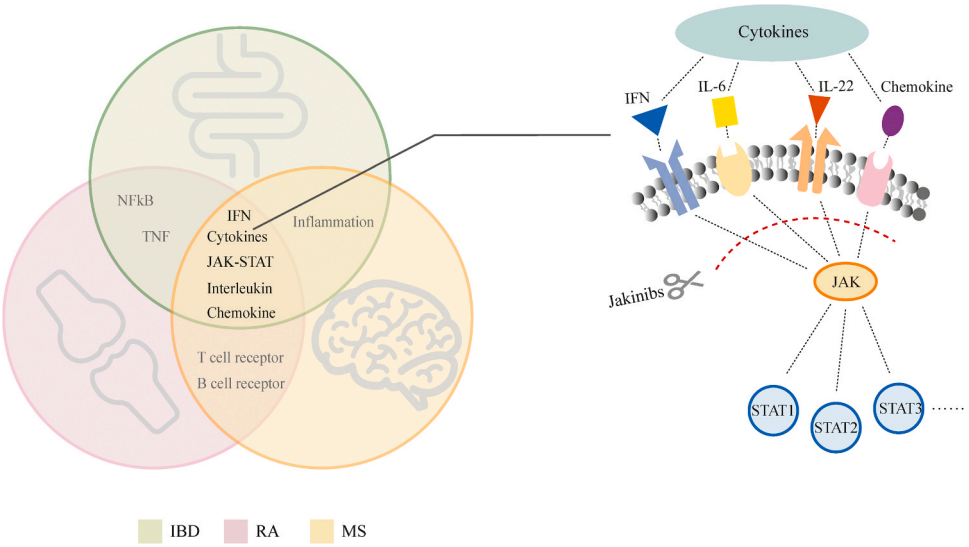

Fig. 4. The common features of the three inflammation subtypes. The summary on the left depicts commonalities among similar subtypes based on markers. In the Venn diagram, green represents the High Inflammation subtype of Inflammatory Bowel Disease (IBD), pink represents the Inflammation subtype of Rheumatoid Arthritis (RA), and yellow represents the "Inflammation & EGF" subtype of Multiple Sclerosis (MS). On the right, the figure illustrates how Jakinibs block the release of inflammatory cytokines, such as interleukins and interferons, by acting on Janus kinase (JAK).

Two SLE subtyping studies [21,40] have defined lymphocyte subtypes in SLE, linking them with clinical manifestations like secondary Sjögren's syndrome, photosensitive rash, antiphospholipid syndrome, and elevated aspartate aminotransferase activity, indicating hepatic dysfunction. A study observed a decrease in lymphocyte proportion with increasing disease activity index [21] and over-expression of the type I interferon-related pathway in lymphocyte-driven subtypes with low SLEDAI values. A Mendelian R4andomization study further established a causal relationship between lymphocyte abundance-related variations and SLE[44](Fig. 3-C).

4.1.4. *Plasma cell subtype*

Plasma cells, also known as effector B cells, are a type of human immune cell and can be classified as long- or short-lived. In patients with SLE, long-lived plasma cells are thought to be responsible for producing anti-RNA and anti-cardiolipin antibodies [98], which are markers of SLE [99]. Subtypes associated with plasma cells were identified in two studies [40,43]. A study based on co-expressed gene modules found that the plasma cell/cell cycle and neutrophil/inflammation modules showed opposite patterns of dysregulation in patients, identifying two distinct subtypes. The plasma cell subtype is predominantly associated with autoantibody-mediated diseases and is linked to clinical manifestations in the dermis, musculoskeletal components, and arthritis. Another study revealed that type I interferon promotes T-cell-independent B-cell proliferation and differentiation into early plasma cells, possibly through a previously described interferon-plasma cell co-dysregulation mechanism [100]. Another study revealed that plasmacyte signaling was stronger in African Americans, who also exhibited the highest levels of disease activity. In addition, anti-dsDNA antibody titers were most strongly correlated with the SLEDAI in African Americans. Therefore, the researchers concluded that African American patients might respond better to B-cell depletion therapy than white patients (Fig. 3-D).

4.2. *Inflammatory bowel disease*

The clinical manifestations of IBD primarily include diarrhea, abdominal pain, decreased appetite, fatigue, intestinal bleeding, intestinal obstruction, and other extraintestinal manifestations (EIMs) involving the musculoskeletal system, skin, and hepatobiliary tract [101]. These manifestations significantly impact patients' quality of life. Based on molecular subtyping studies of IBD, we categorized IBD into two molecular subtypes: the high inflammation subtype and the high metabolism subtype.

4.2.1. *High metabolism subtype*

Three studies [58,102,103] defined the high metabolic subtype as having an abundance of anti-inflammatory immune cells, such as M2 macrophages and regulatory T cells (Tregs). These studies also suggested an enrichment of multiple metabolic pathways, including retinol metabolism, steroid hormone metabolism, niacin and nicotinamide metabolism, bile acid metabolism, and fatty acid metabolism, as well as lower levels of proinflammatory cytokines. In one study, the gene expression pattern of the metabolism-related subtypes was associated with the ileum and thus named the ileum-like subtype [103]. This subtype is primarily distinguished by the activation of lipid and xenobiotic metabolism pathways [94]. Importantly, previous research has linked abnormal lipid metabolism and specific lipid levels to IBD[94]. For example, fatty acids not only influence the composition of intestinal microbiota in the human body [104], but they may also influence the occurrence and progression of IBD via inflammatory status and immune signaling transduction[105].

Furthermore, another study revealed that the metabolism-related subtype exhibited activation of immune cell pathways, especially adaptive immunity pathways, such as B cells and T follicular helper (Tfh) cells. B cells and Tfh cells were found to be associated with elevated IgG levels and disease deterioration in patients with UC [106,107] (Fig. 3-F).

4.2.2. *High inflammation subtype*

The high inflammation subtype, identified in three studies [58,76,102], is characterized by marked inflammatory features, including significant enrichment in inflammation and immune-related pathways such as interferon-γ/α response, TNF-α signaling pathway, IL-6 JAK--STAT3 signaling pathway [102], and others including complement and coagulation cascades, cytokine-cytokine receptor interaction, and Toll-like receptor signaling pathway [76]. It also involves the upregulation of proinflammatory markers at the mRNA level, such as S100A8, S100A9, TNF-α, IL-1β, IL-6, IL-17, INF-γ. Furthermore, this subtype is characterized by the infiltration of proinflammatory immune cells,

notably neutrophils, and mast cells [58], with neutrophils forming extracellular traps that may contribute to the progression of UC towards cancer [75] (Fig. 3-E).

Another study highlighted a general downregulation of metabolic pathways in the high inflammation subtype, alongside decreased levels of intestinal barrier-related proteins, suggesting a potential link between this subtype and poorer prognoses due to increased inflammatory activity and compromised intestinal barriers. Additionally, genes upregulated in the high inflammation subtype showed consistency with those upregulated in depression, hinting at a possible association between this subtype and the occurrence of depression in patients with IBD. The damage to the intestinal barrier might allow inflammation to impact the brain[108], providing a rationale for the observed link between this subtype and depression, as well as the noted improvement in IBD patients with high inflammation upon treatment with the antidepressant drug paroxetine[102](Fig. 3-E).

4.3. Rheumatoid arthritis

RA is characterized as a common autoimmune arthritis marked by chronic inflammation of the synovium, leading to complex disease heterogeneity and nonlinear dynamic interactions. As RA progresses, it often results in irreversible bone tissue damage, causing persistent pain and significantly impaired joint functionality. Molecular subtyping studies have categorized RA patients into multiple subtypes based on the enrichment of different cells and pathways, promoting the development of subtype-specific treatments. This study compiles current molecular subtyping research in RA into three functional subtypes: inflammation, joint damage, and neutrophil subtypes.

4.3.1. Inflammation subtype

The characteristics of inflammation subtypes have been described in three studies [45,46,65], which are associated with both synovial and systemic inflammation levels, as well as the presence of autoantibodies. Although specific characteristics of the inflammation subtypes vary across studies, a common observation is the significant increase in various immune cell populations, such as B cells, T cells, $CD8^+$ T cells, $CD4^+$ T cells, Th2 cells, regulatory T cells, dendritic cells, macrophages, and plasma cells. Additionally, there is an enrichment of proinflammatory pathways, including immune cell-related pathways, IL-17, IFN-a/B, IFN-c, JAK-STAT, NF-κB, TNF, and toll-like receptor (TLR) signaling. Certain studies have taken further steps to provide a more detailed classification of inflammation subtypes. This classification is based on functional variances in proinflammatory pathways, with pathways like P53-PI3K-AKT and BCR, JAK-STAT, NF-κB, and TCR being utilized as markers to identify distinct inflammation subtypes [46].

Observations have been made that the traits of both purely inflammatory and higher inflammation subtypes are remarkably alike, showing elevated immune cell counts, enhancement of proinflammatory pathways, and significant expression of lymphoid and myeloid eigengenes. A shared characteristic between purely inflammatory and higher inflammation subtypes is their association with seropositivity, for instance, ACPA, which in some studies has been linked to improved treatment outcomes compared to other subtypes[45,46,65]. This finding is consistent with the results of a longitudinal cohort study spanning 25 years of RA[109], where patients with autoantibody-positive RA experienced more significant improvements than those with autoantibody-negative RA when treated with the same strategy, especially regarding long-term outcomes (mortality and functional disability). It suggests that autoantibody-positive and autoantibody-negative RA might have distinct pathogeneses and could, therefore, represent different subtypes (Fig. 3-G).

4.3.2. Joint damage subtype

The joint damage subtypes have been defined in two studies [46,65] and is characterized by moderate scores in most pro-inflammatory signaling pathways or only high scores in some pro-inflammatory pathways, hence being considered low-inflammatory. However, this subtype often leads to more severe joint damage and less favorable treatment results compared to the inflammation subtypes.

Subtypes related to joint damage are notably involved in pathways primarily concerning joint damage, tissue remodeling, and synovial proliferation. There is also a marked increase in the expression of regulatory factors critical for fibroblast invasion and molecules essential for osteoclastogenesis[65], particularly in synovial fibroblast linings. These cell-related pathways and markers, prevalent in joint damage subtypes, specifically contribute to bone and cartilage degradation in the body [46,65,110]. Another distinct trait of this subtype is the experience of pain not linked to systemic inflammation markers. Despite minimal tissue inflammation and low inflammation-related gene expression, patients with this subtype still report tenderness and swelling in multiple joints [45], leading to a significantly high pain severity score, adversely affecting their quality of life[45] (Fig. 3-H).

4.3.3. Neutrophil subtype

The neutrophil subtype, as defined in one study [65], shows considerable neutrophil infiltration and raised fibroid eigengene scores. Predominant enrichment pathways include MAPK, PI3K-Akt, p53, TGF-β, VEGF, and Wnt signaling pathways. Due to a significant overlap in related pathways, neutrophil and joint damage subtypes exhibit some similar features, such as poorer outcomes, and more active endothelial and fibroblasts than inflammation subtypes [65] (Fig. 3-I).

4.4. Multiple sclerosis

We summarized the results of two molecular subtyping studies of MS and found that both identified subtypes with strong enrichment of genes related to inflammation and the epithelial growth factor EGF. The lymphocyte-related pathways include T cell receptor, B cell receptor, TLR signaling pathway, interleukins, NFAT, and IFN signaling pathways. EGF-related pathways include ILK, JAK-STAT, PI3K, Ras, EGF, and EGFR, along with pathways linked to angiogenesis and apoptosis. Complex interactions exist between lymphocyte-related pathways and EGF-related pathways. For instance, the JAK-STAT pathway, acting downstream of EGFR, is involved in the IL-6 (gp130) receptor family, influencing B cell differentiation, plasma cell production, and acute-phase response [111]. PI3K interacts with EGFR to regulate cell proliferation, differentiation, apoptosis, and migration and plays a critical role in B cell development, function, and B cell receptor signaling [112].

In the early inflammatory stage of MS, B cells and T cells play essential roles. Patients with these MS subtypes may present increased lymphocytes in their bloodstream, suggesting a potential for improved outcomes with early, aggressive treatment [47]. Further research indicates that these patients, compared to those with less pronounced phenotypes, are more likely to relapse and show more active disease course when treated with GA and IFN-β. One study highlighted a significant enrichment of IFN-induced and viral response gene sets in the high-performing MS subset through GSEA analysis [63]. Type I IFN activates transcription of interferon-stimulated genes (ISGs) such as IFP35 through the JAK-STAT pathway, contributing to MS inflammation via innate immune mechanisms [113].

In summary, the main subtypes identified in the MS subtyping studies are characterized by enrichment of lymphocyte and EGF-related pathways. These pathways are primarily involved in immunity, inflammation, and cellular proliferation and growth in the body. Yet, the specific connections between these pathways in MS remain unclear, suggesting heterogeneity within MS that indicates different subtypes requiring further study (Fig. 3-J).

5. Characteristics and biomarkers of heterogeneity and commonalities among different molecular subtypes

To more comprehensively characterize subtypes and explore the heterogeneity among them within the same disease, ten subtype-specific markers for four autoimmune diseases were summarized. Markers for all subtypes of the same disease were systematically collected, and shared characteristics between subtypes were eliminated, retaining only those markers unique to a single subtype (Supplementary Table 1). Additionally, subtypes with analogous names across various diseases have been identified, such as neutrophil subtypes in SLE and RA, as well as inflammation-associated subtypes in IBD, RA, and MS. To further examine the concordance of these similar subtypes across different diseases, an intersection of biomarkers for similar subtypes was analyzed, delving into the potential mechanisms at the molecular level of shared markers between similar subtypes. Hence, this review presents a detailed analysis of the characteristics of each disease subtype using subtype-specific markers and delineates similar subtypes across various autoimmune diseases.

5.1. The heterogeneity among SLE subtypes

In this research, SLE has been categorized into four subtypes: neutrophil, interferon, lymphocyte, and plasma cell. Given that the plasma cell subtype displays only three specific traits, and the specific markers of the lymphocyte subtype are mainly related to pathway dysregulation common to various autoimmune diseases, the focus is predominantly on elucidating the heterogeneity within the neutrophil and interferon subtypes.

Specific markers for neutrophil subtypes include BAFF, BCL2, MMP8, IL-6, TNF-α, and antinuclear antibody production, alongside cytokine-mediated signaling pathways. BCL-2, a protein that inhibits apoptosis, when overexpressed, extends the lifespan of neutrophils and blocks apoptosis [114]. This prolongation allows the release of inflammatory cytokines like IL-6 and TNF-α [115], which activate pathways such as NF-κB[116] to regulate immune response, thereby promoting neutrophil overactivation and inflammatory responses [117]. Additionally, TNF-α induced abnormal neutrophil apoptosis [118], and excessive MMP8 may lead to extracellular matrix degradation, cell membrane rupture, and nuclear component release, intensifying the inflammatory response [119]. Moreover, BAFF overexpression can increase B cell survival and differentiation into plasma cells, raising antibody production such as antinuclear antibodies (ANA) [120], and the binding of ANAs to nuclear components [121] further stimulates abnormal neutrophil activation, triggering inflammation. Therefore, in neutrophil subtypes, excessive neutrophil survival with aberrant apoptosis promotes abnormal neutrophil activation and inflammation.

Interferon, a significant pathogenic factor in SLE, is closely associated with its pathogenesis [92,122]. In the interferon subtype, specific characteristics are observed, primarily as abnormal interferon pathway activation and abnormal chemokine gene expression, such as IP10 (CXCL10) and MCP2 (CCL8). The production of these chemokines is induced by interferon. Their overexpression under interferon stimulation prolongs autoimmune processes, leading to leukocyte recruitment at inflammation sites, thereby exacerbating autoimmune disease progression like SLE [123,124]. Furthermore, this subtype significantly expresses components of the complement pathway, such as C4, and autoantibodies like anti-Sm/ribonucleoprotein (RNP), anti-DNA, and anti-SSA are progressively increased in response to interferon stimulation [125], These chemokines and autoantibodies may be affected by interferon regulation, leading to specific manifestations in patients with the interferon subtype, which distinguishes their clinical presentation from that of patients with other subtypes of SLE.

5.2. The heterogeneity between IBD subtypes

IBD can be clearly categorized into two disease subtypes, namely, the high metabolism subtype and the high inflammation subtype. As expected, the biomarkers for the high-metabolism subgroup predominantly included genes related to lipid metabolism and transport, such as APOA1, APOB, MTTP, SLC2A2, and TM6SF2, as well as various pathways associated with metabolism, including amino acid and oligopeptide SLC transporters, bile acid metabolism, biological oxidation, fatty acid metabolism, glucose metabolism, lipid and lipoprotein metabolism, nicotinate and nicotinamide metabolism, retinol metabolism, and steroid hormone metabolism.

In contrast, the biomarkers for the high inflammation subtype primarily include proinflammatory factors such as IL-17, IL-1β, IL-22, IL-6, INF-γ, and TNF-α and proinflammatory markers such as S100A8 and S100A9. The specific pathways involved are mainly related to immune responses and inflammation, including pathways such as the IL-6 JAK-STAT3 signaling pathway, inflammatory response, TNF-alpha signaling via NF-κB, and Toll-Like Receptor signaling pathway.

5.3. The heterogeneity among RA subtypes

We categorized rheumatoid arthritis (RA) into three subtypes, namely, the inflammation subtype, joint damage subtype and neutrophil subtype.

By analyzing specific markers within inflammation and joint injury subtypes (Supplementary Table 1), it was found that most markers of the inflammation subtype, such as IL-17, IFN, JAK-STAT, NF-κB, TNF, TCR, BCR, and TLR, are associated with inflammation and immunity. Conversely, the joint injury subtype's specific traits are mainly characterized by the TGF-β family, its receptor, and their coding genes (e.g., BMP6). The inflammation subtype's clinical and molecular features involve increased inflammation and immune cells, with its main markers typically associated with inflammation. The joint injury subtype, characterized by milder inflammatory symptoms and more pronounced pain, includes cellular markers primarily of fibroblasts and osteoclasts. The interaction between these two cells affects the Treg and Th17 cell balance and exacerbates RA bone destruction by promoting RANKL expression on synovial fibroblasts [126]. Osteoclasts are crucial in bone loss and joint destruction in RA, suggesting the joint damage subtype may cause more joint damage than the inflammation subtype. TGF-β, a key feature of the joint damage subtype, promotes synoviocyte growth and differentiation and is essential for articular cartilage metabolic homeostasis and structural integrity [127]. TGF-β signaling can trigger bone remodeling and destruction, potentially leading to joint damage and dysfunction. Inhibiting TGF-β signaling might also alleviate osteoarthritis [128]. In summary, TGF-β as an important feature is highly associated with osteoarthritis and arthritis and is an important feature and marker of heterogeneity of joint injury subtypes. Additionally, the "neuronal neuron" factor observed in the joint injury subtype may be linked to more severe pain symptoms in patients with this subtype.

In RA, there is both commonality and heterogeneity between neutrophil and joint injury subtypes. The neutrophil subtype shares some pathways with the joint injury subtype, and the two subtypes share many similar features and markers, including TGF-β. As a crucial marker of the joint damage subtype, TGF-β might play a similar role in the neutrophil subtype by influencing the growth and differentiation of synovial cells, leading to increased bone resorption and decreased bone formation. This can further exacerbate osteoporosis and bone destruction in RA patients [129], resulting in clinical features in patients with the neutrophil subtype similar to those of the joint damage subtype, such as suboptimal treatment outcomes and the activation of endothelial cells and fibroblasts. However, there is heterogeneity between the neutrophil and joint damage subtypes. The hypothesis is that the distinctive feature of the neutrophil subtype is the extensive infiltration of neutrophils and

a decrease in TGF-β expression, whereas the joint damage subtype is characterized mainly by an increase in TGF-β expression. Existing research indicates a complex regulatory relationship between neutrophils and TGF-β, where neutrophils can inhibit the differentiation of mesenchymal progenitor cells (MPCs) into osteoblasts through TGF-β and promote osteoclastogenesis [129]. Additionally, neutrophil extracellular traps (NETs) released by neutrophils can inhibit the expression of TGF-β and TGFBR2 [130]. Conversely, TGF-β can inhibit neutrophil activity [131], and in some disease states, blocking TGF-β activity leads to a significant increase in neutrophil levels[132,133], indicating that an increase in neutrophil levels might reflect a decrease in TGF-β expression, a mechanism that might play a significant role in the neutrophil subtype and contribute to the heterogeneity between the neutrophil and joint damage subtypes.

5.4. The commonalities among the molecular subtypes of the four ADs

In four autoimmune diseases, three inflammation-related subtypes were defined: the high inflammation subtype in IBD, the inflammation subtype in RA, and the "inflammation & EGF" subtype in MS. In autoimmune diseases, inflammation is closely associated with disease onset and progression. The immune system's erroneous attack on normal cells triggers an inflammatory response, leading to pain and damage, while chronic inflammation results in abnormal immune system activation and dysregulation, initiating autoimmune diseases. These three subtypes exhibit a marked enrichment of inflammatory pathways and increased immune cell presence in the diseases. By summarizing subtype-specific markers and factors (including genes, pathways, clinical indicators, etc.), it was observed that the markers of these three subtypes are closely related, with many markers shared among them.

We found that IFNs, cytokines, multiple proinflammatory interleukins (IL-6, IL-22, IL-17, and IL-1β), the JAK-STAT signaling pathway, and chemokines were shared among the inflammation subtypes of the three disease groups. These features may either promote inflammatory responses or play a role in the regulation of inflammation. They are not only strongly associated with inflammation but also interlinked; for instance, cytokines such as interleukins, interferons, and chemokines are soluble messengers that allow immune cells to communicate. Abnormalities in their signaling can lead to an imbalance in the immune response, contributing to autoimmune diseases [134]. IL-6, as an inflammatory cytokine, can significantly influence the differentiation and activation of T-lymphocytes through the induction of the JAK-STAT3 pathway [135], and interferon can promote the transcription of IFN-stimulated genes (ISGs) by activating the JAK-STAT signaling pathway, affecting the body's autoimmune and inflammatory responses [136]. It has been shown that the JAK-STAT signaling pathway plays a role in the development of both inflammatory and autoimmune diseases, with many cytokines involved in these diseases using JAK and STAT to interrupt intracellular signals [137], The shared features that we observed in the three inflammation subtypes can just be categorized as Jak-STAT and its associated cytokines. We suggest that the features shared by these inflammation subtypes may exhibit a unified inflammatory pattern centered on the activation of the Jak-STAT pathway in IBD, RA, and MS. The effectiveness of JAK inhibitors (Jakinibs) in treating autoimmune diseases also indicates that the JAK-STAT-related signaling pathway could be a potential therapeutic target for these conditions. Jakinibs work by reducing inflammation and the activation of immune cells through the inhibition of Jak activity and blocking various cytokines [138]. The therapeutic action of Jakinibs is aimed at alleviating inflammation. In the section "Drugs or Treatment Methods Beneficial for Four AD Molecular Subtypes," it was found that most drugs specific to the three types of inflammation are linked to Jakinibs, which have been demonstrated to limit Th1 and Th17 differentiation by targeting the JAK2/TYK2-STAT3/STAT4 axis [139]; IFN-β activates the JAK-STAT (JAK1/TYK2-STAT2) signaling pathway, thereby modulating the immune response [140]; and paroxetine acts through the immune 5-HT system and JAK2-STAT3 pathway in cells to provide anti-inflammatory benefits [141]. In summary, infliximab, glatiramer acetate, and paroxetine are identified as more effective drugs for the inflammation subtypes of RA, MS, and IBD, respectively. The mechanisms of these drugs are linked to the JAK-STAT signaling pathway. Hence, it is hypothesized that Jakinibs might offer greater therapeutic benefits to patients with these three types of inflammation subtypes.

In molecular subtype studies of four autoimmune diseases, both SLE and RA patients exhibited neutrophil subtypes.Through comparison, We found that TGF-β is a common marker for both neutrophil subtypes.The roles of neutrophils and TGF-β, which are important factors in immune regulation and inflammation modulation, have been described in many studies of autoimmune diseases, and their mechanisms of involvement in SLE and RA have been revealed by an increasing number of studies. Studies have shown that TGF-β is involved in the production and regulation of γδ Tregs in patients with SLE and may be associated with the pathogenesis of SLE[142]. In RA, TGF-β's role in regulating the activity of synovial fibroblasts by enhancing MMP-11 expression, leading to matrix degradation and remodeling, indicates its contribution to disease occurrence[143]. The relationship between neutrophils and TGF-β in disease has been discussed previously, neutrophil extracellular traps generated by neutrophils have the effect of inhibiting the expression of TGF-β, and blocking TGF-β can increase the level of neutrophils. Therefore, we speculate that the mutual influence of neutrophils and TGF-β may be one of the important reasons for the heterogeneity of the neutrophil subtype. However, further research is needed to understand the common mechanism of action in autoimmune diseases.

The study not only delved into the heterogeneity of the same disease subtype but also focused on the similarity among different subtypes of autoimmune diseases. Given that various autoimmune diseases share common causative genes and risk factors, investigating their similarities can enhance understanding of the pathogenesis of autoimmune diseases (AD). The research conducted by Barturen et al.[22] provided valuable insights by examining seven systemic autoimmune diseases—including SLE, RA, SSC, pSjS, MCTD, PAPS, and UCTD—as a collective unit, and identified four principal clusters: the inflammatory, lymphoid, interferon, and undefined clusters, each presenting distinct immune cell compositions and molecular characteristics. For instance, increased levels of MMP8 and CRP were observed in the inflammatory cluster, while the interferon cluster showed an expression of various interferon-related molecules (e.g., anti-dsDNA, anti-SM, anti-SSB, anti-U1RNP, PFLC, anti-SSA, TNFα, MCP2, BAFF, IP10). Some shared features like CCP2, CENTB, and PC.IGM were found to be slightly enriched in the lymphoid and undefined clusters, with a notable depletion in the interferon cluster. Similar patterns, such as the interferon subtype in SLE, inflammation subtype in RA, high inflammation subtype in IBD, and "inflammation & EGF" subtype in MS, were observed in this study. These consistencies underscore our findings, suggesting the importance of these shared features across different autoimmune diseases. Future research will likely deepen the understanding of molecular typing in ADs and aid in a more precise exploration of disease heterogeneity and shared pathogenesis.

6. Drugs or treatment methods beneficial for four AD molecular subtypes

In practical medicine, due to the complexity of ADs, increasingly intensive treatment regimens are often necessary, and it may be required to attempt various treatments multiple times. If a patient does not respond well to conventional treatments, new drug regimens may be considered. However, these drugs typically act through a broad range of mechanisms and can have significant side effects. Unlike the personalized therapies in oncology, treatments for autoimmune diseases generally lack individualization [144]. Recognizing that different drugs may elicit varied responses across populations, and even the same drug may

produce different effects and side effects within a single patient population, this study has categorized different subtypes of autoimmune diseases. This categorization aims to assist clinicians in making more precise therapeutic decisions for individual patients and in the rational use of multiple drugs and treatment methods.

6.1. Beneficial drugs and treatment methods in SLE

To illustrate the specificity of drug and treatment responses across different subtypes, the results of typing studies related to drug responses were summarized, along with the drugs and treatments beneficial to various disease subtypes (Table 3). Corticosteroids were found to have the most significant effect in the neutrophilic subtype of SLE. As glucocorticoids, corticosteroids are the main treatment for various inflammatory and autoimmune diseases, effectively controlling symptoms such as joint swelling and pain. However, they can also cause adverse effects like gastrointestinal bleeding, osteoporosis, and psychiatric issues [145,146], with some patients experiencing glucocorticoid resistance. Since corticosteroids inhibit a wide range of inflammatory cells and the neutrophil subtype is characterized by an enrichment of neutrophils, these drugs may reduce the inflammatory response in this subtype by inhibiting neutrophils and decreasing the inflammatory mediators they release.

In the interferon subtype, hydrocortisone was particularly effective in reducing disease activity, likely because it enhances IFN-γ mediated indoleamine 2,3-dioxygenase (IDO) activity; thus, increasing IDO mRNA and protein levels in the body [147]. This mechanism, which reduces immune response intensity and suppresses inflammation, may explain the beneficial effect of hydrocortisone on patients with the interferon subtype.

Although the plasma cell subtype of SLE did not have a subtype-specific drug mentioned, this subtype tends to occur more frequently in individuals of Black ethnicity, who have been found to respond better to B-cell depletion therapy than Caucasian patients [148]. Thus, it is hypothesized that B-cell depletion therapy might be more effective in patients of this subtype. Additionally, because B-cell depletion therapy targets the entire B-cell population and affects plasma cells, it may be an important treatment for individuals with the plasma cell subtype, regardless of ethnicity.

6.2. Beneficial drugs in IBD

In studies on IBD, the antidepressant paroxetine proved effective in reducing bowel severity, alleviating weight loss, and significantly lowering Disease Activity Index (DAI) scores for colitis symptom severity in patients with the high inflammation subtype of IBD. Paroxetine, a medication commonly used to treat depression, is of particular interest as patients with IBD are more likely to experience depressive symptoms than the general population [149], offering a novel perspective on treating IBD.

Table 3
Drugs/treatments applicable to the SLE, RA, IBD, and MS subtypes.

Disease	Subtype	Drugs/Treatment
SLE	Neutrophil subtype	Corticosteroids
	Interferon subtype	Hydrocortisone
	Lymphocyte subtype	
	Plasma cell subtype	B-cell depletion therapy
IBD	High metabolism subtype	
	High inflammation subtype	Paroxetine
RA	Inflammation subtype	Infliximab
	Joint damage subtype	
	Neutrophil subtype	
MS	"inflammation & EGF" subtype	GA (Glatiramer acetate); IFN-β (Interferon beta)

SLE: Systemic Lupus Erythematosus; RA: Rheumatoid Arthritis; IBD: Inflammatory Bowel Disease; MS: Multiple Sclerosis.

6.3. Beneficial drugs in RA

In research on RA, it was found that infliximab had a greater positive response in subtypes primarily characterized by inflammatory symptoms, significantly more than in subtypes with non-inflammatory symptoms. Infliximab, a biologic preparation produced through biotechnology from biological macromolecules, is frequently used in treating autoimmune, and inflammatory diseases, and tumors. Despite its effectiveness, biologic preparations, including infliximab, can cause side effects such as infections and allergic reactions. Since infliximab can inhibit the TNF-dependent cytokine cascade in RA patients [150] and TNF is one of the characteristics of the inflammation subtype of RA, indicates it may suitability for this specific RA patient group(Table 3).

6.4. Beneficial drugs in MS

In the present study, we summarized strongly performing MS patients as the "inflammation & EGF" subtype, which has a greater likelihood of relapse than weakly performing MS patients and shows a more active course of disease related to glatiramer acetate (GA) and interferon-β (IFN-β) treatment. Researchers speculate that this may be related to the greater number of active B cells and T cells in this group of patients. cells in such patients. GA and IFN-β belong to the group of immunomodulatory drugs (IVDs), both of which are commonly used as first-line therapeutic agents for MS and have the advantages of reducing disease recurrence, delaying disability progression, and minimizing side effects [151]; however, in previous treatments, approximately 20%–50% of patients did not respond well to treatment with GA and IFN-beta. Our study may help clinicians screen MS patients who are better treated with GA and IFN-β therapies.

7. Discussion

Our research group has committed many years to the study of molecular subtyping in ADs, with a focus on SLE[41], LN, and IgA nephropathy (IgAN) [152], while also exploring overarching similarities among various ADs. This paper reviews 15 molecular typing studies across four different ADs. We have systematically arranged pathological classification studies of Multiple Sclerosis (MS) and LN chronologically, highlighting some advantages of molecular classification over pathological classification. The analysis of molecular subtypes of ADs is summarized in five steps: I) Quality control and normalization of original gene expression data. II) Establishment of feature matrices based on differentially expressed genes (DEGs) or features identified through other methods. III) Classification of subtypes and validation of their reliability. IV) Exploration of heterogeneity between subtypes and characterization of subtype-specific functions and features through differential expression analysis, functional enrichment analysis, immune infiltration analysis, among other methods. V) Development of subtype classification models using techniques like decision trees and support vector machines.

Recent progress in molecular subtyping research of autoimmune diseases (ADs) has facilitated linking distinct subtyping outcomes across various studies focusing on the same condition. This integration involves identifying common dysregulated genes, immune cells, and pathways among subtypes, enabling a deeper characterization of subtypes within each disease. This review concentrates on four ADs—SLE, RA, IBD, and MS—and categorizes them into several subtypes based on their molecular profiles. For SLE, we identify four subtypes: neutrophil subtype, interferon subtype, lymphocyte subtype, and plasma cell subtype. The neutrophil subtype is associated with immune responses, inflammation, bacterial infections, and renal manifestations; the interferon subtype is characterized by widespread dysregulation and a higher prevalence among females, children, and black individuals; the lymphocyte subtype is associated with various clinical manifestations and liver dysfunction; and the plasma cell subtype is mediated by multiple autoantibodies and

correlated with clinical manifestations such as cutaneous, muscular, skeletal, and arthritic involvement. IBD is classified into two subtypes: high metabolism subtype and high inflammation subtype. The high inflammation subtype is characterized by neutrophils displaying significant inflammatory features, whereas the high metabolism subtype exhibits low inflammatory features and upregulation of multiple metabolic and metabolic homeostasis-related pathways. RA is divided into three subtypes: inflammation subtype, joint damage subtype, and neutrophil subtype. The inflammation subtype exhibits high inflammatory features and serum positivity; the joint damage subtype is associated with various factors causing joint damage, displaying high pain scores and a tendency toward seronegativity; and the neutrophil subtype is characterized by prominent neutrophil infiltration and a poor prognosis. MS is summarized as an "inflammation & EGF" subtype with lymphocyte and other inflammatory characteristics, with epidermal growth factor related to cell proliferation and apoptosis. Previous studies have integrated transcriptome data and methylation data of seven ADs for comprehensive molecular stratification of ADs, stratified multiple ADs as a whole, and finally uniformly divided multiple ADs into four clusters: inflammatory cluster, lymphatic cluster, interferon cluster, and undefined cluster[22]. Among these, the inflammatory, lymphatic, and interferon clusters represent distinct molecular patterns and can be considered stable subtypes of the disease. Our study elucidates multiple subtypes characterized by inflammation, lymph, and interferon, offering a nuanced perspective beyond joint subtype classification. Different ADs often exhibit significant variations in organ involvement, genetic or environmental risk factors, and pathophysiological mechanisms [153]. Our study adopts a more nuanced approach, focusing on individual ADs and categorizing each into specific subtypes distinguished by characteristic cells and different disease manifestations. These findings hold significant implications for future clinical trials and personalized drug utilization in ADs.

Despite notable progress in molecular subtyping of autoimmune diseases (ADs), the field remains in its early stages. This is primarily due to the heterogeneity stemming from diverse autoantibodies in individual patients, which complicates the molecular subtyping process of ADs. Furthermore, molecular subtyping research has predominantly focused on multi-organ ADs such as SLE and RA, with limited attention given to organ-specific subtypes. Multi-organ ADs exhibit greater diversity, offering a more robust biological basis for subtyping. SLE, in particular, has been extensively studied in subtyping efforts, facilitated by the ease of blood sample collection, which enables extensive data accumulation. Single-cell transcriptome sequencing has emerged as a powerful tool in ADs research [23], allowing for the exploration of cellular diversity in immune-inflammatory tissues, the identification of pathogenic cell populations, and the elucidation of disease development mechanisms. However, the limited availability of single-cell RNA sequencing samples presents a significant challenge to advancing subtyping studies. Although several tools for integrating bulk and single-cell transcriptomes have been developed [154–156], there is a shortage of matched bulk and single-cell sequencing data in ADs research, making it challenging to integrate these datasets. This scarcity of data poses challenges in transferring subtype information during the analysis process, potentially leading to information loss.

The integration of multiomics data, including epigenomic, transcriptomic, and proteomic data, could unveil intricate details of the molecular subtypes of ADs. This thorough analysis will enhance our understanding of disease mechanisms and heterogeneity, ultimately establishing a more precise foundation for personalized treatment approaches. By using these enriched multi-omics datasets, researchers can delve deeper into ADs, advancing treatment protocols and fostering the development of new pharmaceuticals. In this review, we delineate the specificity of different subtypes within the same disease based on subtype-specific markers, thereby investigating disease-specific pathogenesis and summarizing the drugs beneficial for each subtype. For instance, we observed that in the neutrophil subtype of SLE, the combination of excessive neutrophil survival and abnormal apoptosis leads to abnormal neutrophil activation and inflammatory responses, potentially contributing to disease development. The inflammation subtype of RA is primarily linked to inflammation and immunity. The distinctive features of the joint damage subtype of RA are mainly associated with the TGF-β family, TGF-β receptor, and their related coding genes. Building upon this, we identified similar subtypes across different diseases and explored common autoimmune disease pathogenesis. We discovered that both SLE and RA have neutrophilic subtypes, while the high inflammation subtype of IBD, the inflammation subtype of RA, and the "inflammation & EGF" subtype of MS exhibit similarities. Additionally, we found that the inflammation subtype of these three diseases and their beneficial drugs are associated with the JAK-STAT signaling pathway. However, most studies did not provide detailed drug dosage information to further investigate differences in drug dosage among subtypes. Variations in drug targets, individual differences in drug metabolism, and drug tolerance may result in differing therapeutic effects and adverse reactions in patients during drug administration. Exploring the optimal dosage of beneficial drugs for different patient subtypes, considering patient-specific conditions and disease characteristics, can enhance treatment efficacy, minimize adverse reactions, and improve patient quality of life. In the future, through the utilization of emerging bioinformatics technologies, we anticipate expanding our study sample size and conducting more comprehensive molecular typing of additional autoimmune diseases. This will furnish us with more extensive data to accurately define the functional subtypes of these diseases and explore shared molecular mechanisms and potential connections among a broader range of autoimmune diseases.

CRediT authorship contribution statement

Hao Tang: Writing – review & editing. **Yongshuai Jiang:** Project administration, Supervision, Writing – review & editing. **Ruijie Zhang:** Project administration, Supervision, Writing – review & editing. **Xiangshu Cheng:** Visualization, Writing – original draft, Writing – review & editing. **Xin Meng:** Visualization, Writing – original draft, Writing – review & editing. **Rui Chen:** Visualization, Writing – original draft, Writing – review & editing. **Zerun Song:** Writing – review & editing. **Shuai Li:** Writing – review & editing. **Siyu Wei:** Writing – review & editing. **Hongchao Lv:** Writing – review & editing. **Shuhao Zhang:** Writing – review & editing.

Declaration of Competing Interest

The authors declare no competing interests.

Acknowledgments

This work was supported by the Natural Science Foundation of Heilongjiang Province, China (Grant No. LH2019C043) and the National Natural Science Foundation of China (31970651 and 12026414).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.csbj.2024.03.026.

References

[1] Goldblatt F, O'Neill SG. Clinical aspects of autoimmune rheumatic diseases. Lancet 2013;382(9894):797–808.
[2] Bieber K, et al. Autoimmune pre-disease. Autoimmun Rev 2023;22(2):103236.
[3] Tan EM, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1982;25(11):1271–7.
[4] Kiriakidou M, Ching CL. Systemic lupus erythematosus. Ann Intern Med 2020; 172(11). p. Itc81-itc96.
[5] Agmon-Levin N, et al. Systemic lupus erythematosus one disease or many? Autoimmun Rev 2012;11(8):593–5.

[6] Kaul A, et al. Systemic lupus erythematosus. Nat Rev Dis Prim 2016;2:16039.
[7] Gordon C, et al. Post Hoc Analysis of the Phase II/III APRIL-SLE Study: Association Between Response to Atacicept and Serum Biomarkers Including BLyS and APRIL. Arthritis Rheuma 2017;69(1):122–30.
[8] de Oliveira NFP, de Souza BF, de Castro Coêlho M. UV Radiation and Its Relation to DNA Methylation in Epidermal Cells: A Review. Epigenomes 2020;4(4).
[9] Adams DE, Shao WH. Epigenetic alterations in immune cells of systemic lupus erythematosus and therapeutic implications. Cells 2022;11(3).
[10] Kubo S, et al. Peripheral immunophenotyping identifies three subgroups based on T cell heterogeneity in Lupus patients. Arthritis Rheuma 2017;69(10):2029–37.
[11] Fasano S, et al. Precision medicine in systemic lupus erythematosus. Nat Rev Rheuma 2023;19(6):331–42.
[12] Zhang YZ, Li YY. Inflammatory bowel disease: pathogenesis. World J Gastroenterol 2014;20(1):91–9.
[13] de Lange KM, Barrett JC. Understanding inflammatory bowel disease via immunogenetics. J Autoimmun 2015;64:91–100.
[14] Modigliani R, et al. Clinical, biological, and endoscopic picture of attacks of Crohn's disease. Evolution on prednisolone. Groupe d'Etude Thérapeutique des Affections Inflammatoires Digestives. Gastroenterology 1990;98(4):811–8.
[15] Turner D, et al. Methotrexate following unsuccessful thiopurine therapy in pediatric Crohn's disease. Am J Gastroenterol 2007;102(12):2804–12. quiz 2803, 2813.
[16] Solberg IC, et al. Clinical course in Crohn's disease: results of a Norwegian population-based ten-year follow-up study. Clin Gastroenterol Hepatol 2007;5 (12):1430–8.
[17] Klareskog L, Catrina AI, Paget S. Rheumatoid arthritis. Lancet 2009;373(9664): 659–72.
[18] Hansen MR, Okuda DT. Precision medicine for multiple sclerosis promotes preventative medicine. Ann N Y Acad Sci 2018;1420(1):62–71.
[19] Jiang J, et al. Type I Interferons in the Pathogenesis and Treatment of Autoimmune Diseases. Clin Rev Allergy Immunol 2020;59(2):248–72.
[20] Noack M, Miossec P. Synoviocytes and skin fibroblasts show opposite effects on IL-23 production and IL-23 receptor expression during cell interactions with immune cells. Arthritis Res Ther 2022;24(1):220.
[21] Toro-Domínguez D, et al. Stratification of Systemic Lupus Erythematosus Patients Into Three Groups of Disease Activity Progression According to Longitudinal Gene Expression. Arthritis Rheuma 2018;70(12):2025–35.
[22] Barturen G, et al. Integrative Analysis Reveals a Molecular Stratification of Systemic Autoimmune Diseases. Arthritis Rheuma 2021;73(6):1073–85.
[23] Zeng L, et al. Research progress of single-cell transcriptome sequencing in autoimmune diseases and autoinflammatory disease: a review. J Autoimmun 2022;133:102919.
[24] Castillo RL, et al. Spatial transcriptomics stratifies psoriatic disease severity by emergent cellular ecosystems. Sci Immunol 2023;8(84):eabq7991.
[25] Conrad N, et al. Incidence, prevalence, and co-occurrence of autoimmune disorders over time and by age, sex, and socioeconomic status: a population-based cohort study of 22 million individuals in the UK. Lancet 2023;401(10391): 1878–90.
[26] Lublin FD, Reingold SC. Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. Neurology 1996;46(4):907–11.
[27] Lublin FD, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. Neurology 2014;83(3):278–86.
[28] Kuhlmann T, et al. Acute axonal damage in multiple sclerosis is most extensive in early disease stages and decreases over time. Brain 2002;125(Pt 10):2202–12.
[29] Kuhlmann T. Relapsing-remitting and primary progressive MS have the same cause(s)–the neuropathologist's view: 2. Mult Scler 2013;19(3):268–9.
[30] Lassmann H. Relapsing-remitting and primary progressive MS have the same cause(s)–the neuropathologist's view: 1. Mult Scler 2013;19(3):266–7.
[31] Luchetti S, et al. Progressive multiple sclerosis patients show substantial lesion activity that correlates with clinical disease severity and sex: a retrospective autopsy cohort analysis. Acta Neuropathol 2018;135(4):511–28.
[32] Tallantyre EC, et al. Greater loss of axons in primary progressive multiple sclerosis plaques compared to secondary progressive disease. Brain 2009;132(Pt 5):1190–9.
[33] Stevenson VL, et al. Primary and transitional progressive MS: a clinical and MRI cross-sectional study. Neurology 1999;52(4):839–45.
[34] Pitt D, et al. Toward precision phenotyping of multiple sclerosis. Neurol Neuroimmunol Neuroinflamm 2022;9(6).
[35] Appel GB, et al. Renal involvement in systemic lupud erythematosus (SLE): a study of 56 patients emphasizing histologic classification. Med (Baltim) 1978;57 (5):371–410.
[36] Painter DM. Renal disease: classification and atlas of glomerular diseases. Pathology 1996;28:215.
[37] Weening JJ, et al. The classification of glomerulonephritis in systemic lupus erythematosus revisited. Kidney Int 2004;65(2):521–30.
[38] Bajema IM, et al. Revision of the International Society of Nephrology/Renal Pathology Society classification for lupus nephritis: clarification of definitions, and modified National Institutes of Health activity and chronicity indices. Kidney Int 2018;93(4):789–96.
[39] Yu F, et al. Redefining lupus nephritis: clinical implications of pathophysiologic subtypes. Nat Rev Nephrol 2017;13(8):483–95.
[40] Banchereau R, et al. Personalized immunomonitoring uncovers molecular networks that stratify lupus patients. Cell 2016;165(3):551–65.
[41] Cui M, et al. Blood Genomics identifies three subtypes of systemic lupus erythematosus: "IFN-High," "NE-High," and "Mixed". Mediat Inflamm 2021;2021: 6660164.
[42] Li H, Yang P. Identification of biomarkers related to neutrophils and two molecular subtypes of systemic lupus erythematosus. BMC Med Genom 2022;15 (1):162.
[43] Toro-Domínguez D, et al. Scoring personalized molecular portraits identify Systemic Lupus Erythematosus subtypes and predict individualized drug responses, symptomatology and disease progression. Brief Bioinform 2022;23(5).
[44] Perez RK, et al. Single-cell RNA-seq reveals cell type-specific molecular and genetic associations to lupus. Science 2022;376(6589)). p. eabf1970.
[45] Orange DE, et al. Identification of Three Rheumatoid Arthritis Disease Subtypes by Machine Learning Integration of Synovial Histologic Features and RNA Sequencing Data. Arthritis Rheuma 2018;70(5):690–701.
[46] Kim KJ, et al. Compendium of synovial signatures identifies pathologic characteristics for predicting treatment response in rheumatoid arthritis patients. Clin Immunol 2019;202:1–10.
[47] Ottoboni L, et al. An RNA profile identifies two subsets of multiple sclerosis patients differing in disease activity. Sci Transl Med 2012;4(153). 153ra131.
[48] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinforma 2008;9:559.
[49] Tibshirani R, Walther G, Hastie T. Estimating the Number of Clusters in a Data Set Via the Gap Statistic. J R Stat Soc Ser B 2001;63:411–23.
[50] Kaufman, L. and P. Rousseeuw, Finding Groups in Data: An Introduction To Cluster Analysis. 1990.
[51] Charrad M, et al. NbClust: an R Package for Determining the Relevant Number of Clusters in a Data Set. J Stat Softw 2014;61(6). p. 1 - 36.
[52] Brusco MJ, et al. Affinity propagation: an exemplar-based tool for clustering in psychological research. Br J Math Stat Psychol 2019;72(1):155–82.
[53] Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics 2010;26(12):1572–3.
[54] Scrucca L, et al. mclust 5: clustering, classification and density estimation using gaussian finite mixture models. R J 2016;8(1):289–317.
[55] Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. BMC Bioinforma 2010;11:367.
[56] Zhao Q, et al. CrossICC: iterative consensus clustering of cross-platform gene expression data without adjusting batch effect. Brief Bioinform 2020;21(5): 1818–24.
[57] Bodenhofer U, Kothmeier A, Hochreiter S. APCluster: an R package for affinity propagation clustering. Bioinformatics 2011;27(17):2463–4.
[58] Mo S, et al. A precise molecular subtyping of ulcerative colitis reveals the immune heterogeneity and predicts clinical drug responses. J Transl Med 2023;21(1):466.
[59] Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005; 102(43):15545–50.
[60] Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinforma 2013;14:7.
[61] Steen CB, et al. Profiling cell type abundance and expression in bulk tissues with CIBERSORTx. Methods Mol Biol 2020;2117:135–57.
[62] Zeng D, et al. IOBR: multi-omics immuno-oncology biological research to decode tumor microenvironment and signatures. Front Immunol 2021;12:687975.
[63] van Baarsen LG, et al. A subtype of multiple sclerosis defined by an activated immune defense program. Genes Immun 2006;7(6):522–31.
[64] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15(12):550.
[65] Jung SM, Park KS, Kim KJ. Deep phenotyping of synovial molecular signatures by integrative systems analysis in rheumatoid arthritis. Rheumatol (Oxf) 2021;60(7): 3420–31.
[66] Hoadley KA, et al. Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. Cell 2014;158(4):929–44.
[67] Liston A, Carr EJ, Linterman MA. Shaping Variation in the Human Immune System. Trends Immunol 2016;37(10):637–46.
[68] Oughtred R, et al. The BioGRID interaction database: 2019 update. Nucleic Acids Res 2019;47(D1). p. D529-d541.
[69] Keshava Prasad TS, et al. Human protein reference database–2009 update. Nucleic Acids Res 2009;37(Database issue):D767–72.
[70] Kerrien S, et al. The IntAct molecular interaction database in 2012. Nucleic Acids Res 2012;40(Database issue):D841–6.
[71] Fabregat A, et al. The Reactome Pathway Knowledgebase. Nucleic Acids Res 2018;46(D1). p. D649-d655.
[72] Szklarczyk D, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res 2021;49(D1). p. D605-d612.
[73] Biasi F, et al. Inflammatory bowel disease: mechanisms, redox considerations, and therapeutic targets. Antioxid Redox Signal 2013;19(14):1711–47.
[74] Dinallo V, et al. Neutrophil extracellular traps sustain inflammatory signals in ulcerative colitis. J Crohns Colitis 2019;13(6):772–84.
[75] Mutua V, Gershwin LJ. A review of neutrophil extracellular traps (NETs) in disease: potential Anti-NETs therapeutics. Clin Rev Allergy Immunol 2021;61(2): 194–211.
[76] Zhang C, et al. Identifying neutrophil-associated subtypes in ulcerative colitis and confirming neutrophils promote colitis-associated colorectal cancer. Front Immunol 2023;14:1095098.
[77] Ritchie ME, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43(7):e47.

[78] Bindea G, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity 2013;39(4):782–95.
[79] Bhattacharya S, et al. ImmPort: disseminating data to the public for the future of immunology. Immunol Res 2014;58(2-3):234–9.
[80] Chin CH, et al. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol 2014;8(Suppl 4(Suppl 4):S11.
[81] Zyla J, et al. Gene set enrichment for reproducible science: comparison of CERNO and eight other algorithms. Bioinformatics 2019;35(24):5146–54.
[82] Chaussabel D, et al. A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus. Immunity 2008;29(1):150–64.
[83] Yaari G, et al. Quantitative set analysis for gene expression: a method to quantify gene set differential expression including gene-gene correlations. Nucleic Acids Res 2013;41(18):e170.
[84] Gene Ontology Consortium: going forward. Nucleic Acids Res, 2015. 43(Database issue): p. D1049-D1056.
[85] Kuleshov MV, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 2016;44(W1):W90–7.
[86] Newman AM, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 2015;12(5):453–7.
[87] Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000;28(1):27–30.
[88] Xiong Q, et al. Integrating genetic and gene expression evidence into genome-wide association analysis of gene sets. Genome Res 2012;22(2):386–97.
[89] Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4(1):44–57.
[90] Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol 2017;18(1):220.
[91] Mi H, et al. The PANTHER database of protein families, subfamilies, functions and pathways. Nucleic Acids Res 2005;33(Database issue):D284–8.
[92] Postal M, et al. Type I interferon in the pathogenesis of systemic lupus erythematosus. Curr Opin Immunol 2020;67:87–94.
[93] Deng Y, et al. Expression characteristics of interferon-stimulated genes and possible regulatory mechanisms in lupus patients using transcriptomics analyses. EBioMedicine 2021;70:103477.
[94] Agouridis AP, Elisaf M, Milionis HJ. An overview of lipid abnormalities in patients with inflammatory bowel disease. Ann Gastroenterol 2011;24(3):181–7.
[95] Levy RA, et al. 10 Years of belimumab experience: What have we learnt? Lupus 2021;30(11):1705–21.
[96] Dhillon S. Telitacicept: First Approval. Drugs 2021;81(14):1671–5.
[97] Humbel M, et al. Restoration of NK Cell Cytotoxic Function With Elotuzumab and Daratumumab Promotes Elimination of Circulating Plasma Cells in Patients With SLE. Front Immunol 2021;12:645478.
[98] Liu Z, Zou Y, Davidson A. Plasma cells in systemic lupus erythematosus: the long and short of it all. Eur J Immunol 2011;41(3):588–91.
[99] Clarke J. IL-17 sustains plasma cells in SLE. Nat Rev Rheuma 2020;16(12):666.
[100] Banchereau J, Pascual V. Type I interferon in systemic lupus erythematosus and other autoimmune diseases. Immunity 2006;25(3):383–92.
[101] Rogler G, et al. Extraintestinal Manifestations of Inflammatory Bowel Disease: Current Concepts, Treatment, and Implications for Disease Management. Gastroenterology 2021;161(4):1118–32.
[102] Ning L, et al. Identification and investigation of depression-related molecular subtypes in inflammatory bowel disease and the anti-inflammatory mechanisms of paroxetine. Front Immunol 2023;14:1145070.
[103] Weiser M, et al. Molecular classification of Crohn's disease reveals two clinically relevant subtypes. Gut 2018;67(1):36–42.
[104] den Besten G, et al. The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. J Lipid Res 2013;54(9):2325–40.
[105] Shores DR, et al. New insights into the role of fatty acids in the pathogenesis and resolution of inflammatory bowel disease. Inflamm Bowel Dis 2011;17(10):2192–204.
[106] Long Y, et al. The Imbalance of Circulating Follicular Helper T Cells and Follicular Regulatory T Cells Is Associated With Disease Activity in Patients With Ulcerative Colitis. Front Immunol 2020;11:104.
[107] Long Y, et al. Activated inducible co-stimulator-positive programmed cell death 1-positive follicular helper T cells indicate disease activity and severity in ulcerative colitis patients. Clin Exp Immunol 2020;202(1):106–18.
[108] Carloni S, et al. Identification of a choroid plexus vascular barrier closing during intestinal inflammation. Science 2021;374(6566):439–48.
[109] Matthijssen XME, et al. Enhanced treatment strategies and distinct disease outcomes among autoantibody-positive and -negative rheumatoid arthritis patients over 25 years: a longitudinal cohort study in the Netherlands. PLoS Med 2020;17(9):e1003296.
[110] Croft AP, et al. Distinct fibroblast subsets drive inflammation and damage in arthritis. Nature 2019;570(7760):246–51.
[111] Aliyu M, et al. Interleukin-6 cytokine: An overview of the immune regulation, immune dysregulation, and therapeutic approach. Int Immunopharmacol 2022;111:109130.
[112] Jellusova J, Rickert RC. The PI3K pathway in B cell metabolism. Crit Rev Biochem Mol Biol 2016;51(5):359–78.
[113] Cohen JR, et al. Altered aortic protease and antiprotease activity in patients with ruptured abdominal aortic aneurysms. Surg Gynecol Obstet 1987;164(4):355–8.
[114] Graninger WB, et al. Cytokine regulation of apoptosis and Bcl-2 expression in lymphocytes of patients with systemic lupus erythematosus. Cell Death Differ 2000;7(10):966–72.
[115] Kolaczkowska E, Kubes P. Neutrophil recruitment and function in health and inflammation. Nat Rev Immunol 2013;13(3):159–75.
[116] Hayden MS, Ghosh S. Regulation of NF-κB by TNF family cytokines. Semin Immunol 2014;26(3):253–66.
[117] Hayden MS, Ghosh S. NF-κB in immunobiology. Cell Res 2011;21(2):223–44.
[118] Maianski NA, Roos D, Kuijpers TW. Tumor necrosis factor alpha induces a caspase-independent death pathway in human neutrophils. Blood 2003;101(5):1987–95.
[119] Pittayapruek P, et al. Role of Matrix Metalloproteinases in Photoaging and Photocarcinogenesis. Int J Mol Sci 2016;17(6).
[120] Möckel T, et al. B cell activating factor (BAFF): Structure, functions, autoimmunity and clinical implications in Systemic Lupus Erythematosus (SLE). Autoimmun Rev 2021;20(2):102736.
[121] Pisetsky DS, Bossuyt X, Meroni PL. ANA as an entry criterion for the classification of SLE. Autoimmun Rev 2019;18(12):102400.
[122] Liu W, Zhang S, Wang J. IFN-γ, should not be ignored in SLE. Front Immunol 2022;13:954706.
[123] Antonelli A, et al. Chemokine (C-X-C motif) ligand (CXCL)10 in autoimmune diseases. Autoimmun Rev 2014;13(3):272–80.
[124] Van Coillie E, et al. Transcriptional control of the human MCP-2 gene promoter by IFN-gamma and IL-1beta in connective tissue cells. J Leukoc Biol 1999;66(3):502–11.
[125] Tumurkhuu G, et al. Oxidative DNA Damage Accelerates Skin Inflammation in Pristane-Induced Lupus Model. Front Immunol 2020;11:554725.
[126] Komatsu N, Takayanagi H. Mechanisms of joint destruction in rheumatoid arthritis - immune cell-fibroblast-bone interactions. Nat Rev Rheuma 2022;18(7):415–29.
[127] Blaney Davidson EN, van der Kraan PM, van den Berg WB. TGF-beta and osteoarthritis. Osteoarthr Cartil 2007;15(6):597–604.
[128] Zhen G, et al. Inhibition of TGF-β signaling in mesenchymal stem cells of subchondral bone attenuates osteoarthritis. Nat Med 2013;19(6):704–12.
[129] Li J, et al. TGFβ1(+)CCR5(+) neutrophil subset increases in bone marrow and causes age-related osteoporosis in male mice. Nat Commun 2023;14(1):159.
[130] Liu B, et al. Neutrophil extracellular traps promote intestinal barrier dysfunction by regulating macrophage polarization during trauma/hemorrhagic shock via the TGF-β signaling pathway. Cell Signal 2024;113:110941.
[131] Shen L, et al. Inhibition of human neutrophil degranulation by transforming growth factor-beta1. Clin Exp Immunol 2007;149(1):155–61.
[132] Allen SS, et al. Altered inflammatory responses following transforming growth factor-beta neutralization in experimental guinea pig tuberculous pleurisy. Tuberc (Edinb) 2008;88(5):430–6.
[133] Fridlender ZG, et al. Polarization of tumor-associated neutrophil phenotype by TGF-beta: "N1" versus "N2" TAN. Cancer Cell 2009;16(3):183–94.
[134] Lampé L, et al. International cooperation to collect obstetric data. Zent Gynakol 1979;101(17):1097–103.
[135] Neurath MF, Finotto S. IL-6 signaling in autoimmunity, chronic inflammation and inflammation-associated cancer. Cytokine Growth Factor Rev 2011;22(2):83–9.
[136] Ivashkiv LB, Donlin LT. Regulation of type I interferon responses. Nat Rev Immunol 2014;14(1):36–49.
[137] Banerjee S, et al. JAK-STAT Signaling as a Target for Inflammatory and Autoimmune Diseases: Current and Future Prospects. Drugs 2017;77(5):521–46.
[138] Schwartz DM, et al. JAK inhibition as a therapeutic strategy for immune and inflammatory diseases. Nat Rev Drug Discov 2017;17(1):78.
[139] Ghoreschi K, et al. Fumarates improve psoriasis and multiple sclerosis by inducing type II dendritic cells. J Exp Med 2011;208(11):2291–303.
[140] O'Shea JJ, et al. The JAK-STAT pathway: impact on human disease and therapeutic intervention. Annu Rev Med 2015;66:311–28.
[141] Kabiri M, et al. Paroxetine modulates immune responses by activating a JAK2/STAT3 signaling pathway. J Biochem Mol Toxicol 2020;34(5):e22464.
[142] Li X, et al. Generation of human regulatory gammadelta T cells by TCRgammadelta stimulation in the presence of TGF-beta and their involvement in the pathogenesis of systemic lupus erythematosus. J Immunol 2011;186(12):6693–700.
[143] Pohlers D, et al. Constitutive upregulation of the transforming growth factor-beta pathway in rheumatoid arthritis synovial fibroblasts. Arthritis Res Ther 2007;9(3):R59.
[144] Fugger L, Jensen LT, Rossjohn J. Challenges, progress, and prospects of developing therapies to treat autoimmune diseases. Cell 2020;181(1):63–80.
[145] Galati A, et al. Glucocorticoids for therapeutic immunosuppression: clinical pearls for the practicing neurologist. J Neurol Sci 2021;430:120004.
[146] Caplan A, et al. Prevention and management of glucocorticoid-induced side effects: a comprehensive review: A review of glucocorticoid pharmacology and bone health. J Am Acad Dermatol 2017;76(1):1–9.
[147] Türck J, et al. Enhancement of antimicrobial effects by glucocorticoids. Med Microbiol Immunol 2005;194(1-2):47–53.
[148] Navarra SV, et al. Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomised, placebo-controlled, phase 3 trial. Lancet 2011;377(9767):721–31.
[149] Bisgaard TH, et al. Depression and anxiety in inflammatory bowel disease: epidemiology, mechanisms and treatment. Nat Rev Gastroenterol Hepatol 2022;19(11):717–26.
[150] Monaco C, et al. Anti-TNF therapy: past, present and future. Int Immunol 2015;27(1):55–62.
[151] O'Gorman C, Lucas R, Taylor B. Environmental risk factors for multiple sclerosis: a review with a focus on molecular mechanisms. Int J Mol Sci 2012;13(9):11718–52.

[152] Ren H, et al. Identifying functional subtypes of IgA nephropathy based on three machine learning algorithms and WGCNA. BMC Med Genom 2024;17(1):61.
[153] Frazzei G, et al. Preclinical autoimmune disease: a comparison of rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis and type 1 diabetes. Front Immunol 2022;13:899372.
[154] Zhang Q, Jin S, Zou X. scAB detects multiresolution cell states with clinical significance by integrating single-cell genomics and bulk sequencing data. Nucleic Acids Res 2022;50(21):12112–30.
[155] Sun D, et al. Identifying phenotype-associated subpopulations by integrating bulk and single-cell sequencing data. Nat Biotechnol 2022;40(4):527–38.
[156] Liao J, et al. De novo analysis of bulk RNA-seq data at spatially resolved single-cell resolution. Nat Commun 2022;13(1):6498.