



Integration of *Immunome* With Disease-Gene Network Reveals Common Cellular Mechanisms Between IMIDs and Drug Repurposing Strategies

Abhinandan Devaprasad^{1,2}, Timothy R. D. J. Radstake^{1,2} and Aridaman Pandit^{1,2*}

¹ Division Internal Medicine and Dermatology, University Medical Center Utrecht, Utrecht, Netherlands, ² Center for Translational Immunology, University Medical Center Utrecht, Utrecht, Netherlands

OPEN ACCESS

Edited by:

Moncef Zouali, Institut National de la Santé et de la Recherche Médicale (INSERM), France

Reviewed by:

Mohan Maddur, Pfizer, United States Giuseppe Murdaca, University of Genoa, Italy

*Correspondence: Aridaman Pandit a.pandit@umcutrecht.nl

Specialty section:

This article was submitted to Molecular Innate Immunity, a section of the journal Frontiers in Immunology

Received: 18 February 2021 Accepted: 04 May 2021 Published: 24 May 2021

Citation:

Devaprasad A, Radstake TRDJ and Pandit A (2021) Integration of Immunome With Disease-Gene Network Reveals Common Cellular Mechanisms Between IMIDs and Drug Repurposing Strategies. Front. Immunol. 12:669400. doi: 10.3389/fimmu.2021.669400 **Objective:** Development and progression of immune-mediated inflammatory diseases (IMIDs) involve intricate dysregulation of the disease-associated genes (DAGs) and their expressing immune cells. Identifying the crucial disease-associated cells (DACs) in IMIDs has been challenging due to the underlying complex molecular mechanism.

Methods: Using transcriptome profiles of 40 different immune cells, unsupervised machine learning, and disease-gene networks, we constructed the Disease-gene IMmune cell Expression (DIME) network and identified top DACs and DAGs of 12 phenotypically different IMIDs. We compared the DIME networks of IMIDs to identify common pathways between them. We used the common pathways and publicly available drug-gene network to identify promising drug repurposing targets.

Results: We found CD4⁺Treg, CD4⁺Th1, and NK cells as top DACs in inflammatory arthritis such as ankylosing spondylitis (AS), psoriatic arthritis, and rheumatoid arthritis (RA); neutrophils, granulocytes, and BDCA1⁺CD14⁺ cells in systemic lupus erythematosus and systemic scleroderma; ILC2, CD4⁺Th1, CD4⁺Treg, and NK cells in the inflammatory bowel diseases (IBDs). We identified lymphoid cells (CD4⁺Th1, CD4⁺Treg, and NK) and their associated pathways to be important in HLA-B27 type diseases (psoriasis, AS, and IBDs) and in primary-joint-inflammation-based inflammatory arthritis (AS and RA). Based on the common cellular mechanisms, we identified liftegrast as a potential drug repurposing candidate for Crohn's disease and other IMIDs.

Conclusions: Existing methods are inadequate in capturing the intricate involvement of the crucial genes and cell types essential to IMIDs. Our approach identified the key DACs, DAGs, common mechanisms between IMIDs, and proposed potential drug repurposing targets using the DIME network. To extend our method to other diseases, we built the DIME tool (https://bitbucket.org/systemsimmunology/dime/) to help scientists uncover

1

the etiology of complex and rare diseases to further drug development by betterdetermining drug targets, thereby mitigating the risk of failure in late clinical development.

Keywords: IMID (immune-mediated inflammatory diseases), disease-associated cells, disease-associated genes, drug repurposing, machine learning, *immunome*

INTRODUCTION

Genetic and epigenetic heterogeneity plays a significant role in the development and progression of complex diseases. The past two decades have seen a major surge in studies that characterize genes and loci associated with diseases (1). The use of highthroughput omics technology and functional screenings have boosted our knowledge about genetic, epigenetic, and metabolic factors underlying complex diseases (1). As a result of these genetic and epigenetic screenings, we now know that most complex diseases and genes/loci have a many-to-many relationship. Meaning that complex diseases are linked to many different genes, and a gene/loci might be associated with several diseases (2). Thus, it is essential to identify and characterize these disease-associated genes (DAGs) to understand diseases better and develop therapy accordingly.

Extensive high-throughput screening studies and multi-omics have helped in the identification of DAGs. However, in most studies, DAGs were identified using bulk tissue or whole blood, a caveat since each gene's expression is known to vary between tissues and cell types (3, 4). Thus, bulk tissue- or blood-based studies on DAGs do not consider the role of different cells and tissues in disease biology. To improve the understanding and molecular basis of complex diseases, a large number of research groups and consortiums have started to functionally identify disease-associated cells (DACs) or tissue types (3-7). The Genotype-Tissue Expression (GTEx) is one such valuable project, which maps gene expression profiles of 54 different human tissue types and the corresponding expression quantitative trait loci (eQTLs) (5-7). Furthermore, the growth of single-cell technologies propelled our understanding of diseases and helped in identifying DACs for complex conditions, including cancer (8), Alzheimer's (9), rheumatoid arthritis (10), among others. Among these studies, the role of immune cells has been central to disease etiology and progression.

The immune system plays a vital role in developing and progressing immune-mediated and non-immune mediated chronic diseases. Many association and functional studies have shown that immune cells express multiple DAGs, and perturbing these DAGs can modulate immune cell functions (11). However, very few studies have explored the impact of DAGs on specific cell types and even fewer on immune cells, many of which focus on a limited number of cell subsets (12–16). Recently, Schmiedel et al. studied the effect of genetic variants on gene expression in 13 different immune cell types (17). However, this study primarily focused on analyzing genetic variants and their impact on a total of 13 immune cell types: monocytes (classical and non-classical), NK cells, naïve B-cells, and nine subpopulations of T-cells. The study identified several genetic variants to have a role in specific immune cell subsets in autoimmune disorders. For example, the modulatory effects of the variant rs12936231 in asthma and other autoimmune diseases are seen in lymphoid rather than myeloid subsets as previously described (17). Such new insights into specific immune cells' role led us to believe that specific immune cells and their DAGs remain poorly understood even in immune disorders.

The immune-mediated inflammatory diseases (IMIDs) are complex among the immune disorders, involving several immune cells. For example, in rheumatoid arthritis, the immune cells such as B-cells, T-cells, macrophages, mast cells, dendritic cells, and NK cells play a significant role (18). However, the exact mechanism of these cell types remains unknown. Insights on the precise mechanism of action are crucial for developing successful therapies, which becomes particularly challenging for IMIDs due to several cell types involved. The massive undertaking of GWAS has enabled the mapping of some of the molecular mechanisms of the IMIDs (19-22). However, further research is required to understand the etiology of IMIDs taking into account the several different immune cells at play and the contributing DAGs for each immune cell type. By identifying the critical immune cells and their mechanism, we would set a robust rationale for identifying any mechanistic overlap between diseases and exploiting them to develop therapeutic strategies.

This study mapped the largest available and expert-curated disease-gene network (from the DisGeNet curated from 16 different databases) (23) on the most extensive immunome data comprising gene expression profiles of 40 different immune cell types, curated by us. We then used an unsupervised machine learning algorithm, the disease-gene network, and the immunome to create the Disease-gene IMmune cell Expression (DIME) network. Using this approach, we built a tool called DIME. Using DIME, we quantified the effects of 3957 DAGs on the immunome to identify DACs for 12 phenotypically different IMIDs. We used the DIME to (1) study the underlying cell-specific mechanisms (2); identify common DACs and their topweighted DAGs between disease pairs (referred to as the common cell-gene network); and (3) identify drug repurposing targets using the common cell-gene network. The DIME is available as a user-friendly R tool (https:// bitbucket.org/systemsimmunology/dime), to determine the top genes and cells associated with the disease of interest for (1): diseases from the DisGeNet (2), diseases from the EBI genome-wide association study (GWAS) catalog, or (3) custom set of genes defined by the user.

Disease-Gene Networks for IMIDs

METHODS

Transcriptome Data - Immunome

The transcriptome data consists of RNA-sequencing datasets of 40 different immune cell types curated using 316 samples from a total of 27 publicly available datasets (see **Supplementary Table S1** for list of GEO datasets and samples used). The 40 different immune cells cover the entire hematopoietic stem cell differentiation tree comprising nine progenitors, 19 lymphoid, and 12 myeloid cell types. The samples used here were manually curated considering only the unstimulated (except for monocyte-derived macrophages) immune cells that were sorted using Fluorescence-activated cell sorting (FACS) and were isolated from either blood, bone marrow, or cord blood from healthy donors.

All the selected datasets (**Supplementary Table S1**) were downloaded as FASTQ files using the fastq-dump tool from sratoolkit. The "split-files" option was given if the library type was paired-end sequencing. FASTQ files were then aligned to the reference genome (GRCH.Hg38.79) using the STAR aligner (24). The result is a SAM file, which was then converted into a sorted BAM file using the samtools program (25). These were then used to calculate the count of aligned reads using the HTSeq program (26) with the "intersection non-empty" option. HTSeq was run for all possible stranded mode options, the count file with the maximum counts was chosen as the respective count file for the sample.

The data was then filtered by removing all genes that had less than 20 read counts in 95 percent of the samples using R programming. The filtered data was then lane normalized using the "betweenLaneNormalization" function from the RUVSeq package (27). The RUVr method from RUVSeq was used to identify residual factors contributing to the batch effect. The resulting filtered, batch corrected, and normalized data had expression for 34,906 genes that were void of any observable batch effect. We calculated counts per million (CPM) for the filtered genes and used log2(CPM +1) as the gene expression measure. We then used the median gene expression for each cell type for the rest of the analysis. This processed, batch corrected, and normalized data of the 40 immune cells is referred to here as the *immunome*.

Disease-Gene Network From DisGeNet

The disease-gene network from DisGeNet (23) was downloaded from the DisGeNet database (www.disgenet.org/downloads). All HLA associated genes were removed from the network; this was done to ensure that bias towards myeloid cells and B cells are removed since the HLA genes are primarily expressed by these cells. The resulting network was further filtered to include only those genes that were present in the *immunome*.

IMID Disease-Gene Network

To study and identify the DACs of the IMIDs, we extracted the DAGs of 12 IMIDs extracted from the above DisGeNet. The IMID gene network for the 12 diseases comprised of 3579 DAGs. The 12 diseases that broadly represent the IMIDs in this study include: ankylosing spondylitis (CUI: C0038013), arthritis (CUI: C0003864), Crohn's disease (CUI: C0010346), diabetes mellitus -

non-insulin-dependent (CUI: C0011860), systemic lupus erythematosus (CUI: C0024141), multiple sclerosis (CUI: C0026769), psoriasis (CUI: C0033860), psoriatic arthritis (CUI: C0003872), rheumatoid arthritis (CUI: C0003873), Sjogren's syndrome (CUI: C1527336), systemic scleroderma (CUI: C0036421), and ulcerative colitis (CUI: C0009324). CUI, used in DisGeNet, is the concept-unique-identifier for the disease term defined by the unified medical language system (28). The disease term arthritis (CUI: C0003864) comprises DAGs that pan over several arthropathies such as spondyloarthropathy, osteoarthritis, gout, allergic arthritis, etc., that fall under the broad arthritis MeSH term.

Identification of Top DAC and DAG Using Machine Learning

Briefly, we used an unsupervised machine learning algorithm called non-negative matrix factorization (NMF) to map the disease-gene network to the immunome and identify the top DACs and DAGs of the 12 IMIDs. The NMF algorithm clusters the input gene expression data into 'k' clusters, such that the DAGs of a cluster are expressed by the DACs of the same cluster, thus forming DAC-DAG pairs in each cluster (29). We used the coefficients and weights identified by the NMF algorithm as the DAC and DAG scores, respectively. The scores were scaled between 0 and 1, with 1 being the highest score. Those in the top 25 percentile of the scores were regarded as the top DACs and DAGs, respectively. We calculated the Frobenius norm for each cluster to weigh and rank the clusters; the rank 1 cluster is the top cluster having the highest Frobenius norm value. The top cluster comprises the DAC-DAG pair that maximally captures/ represents the input gene expression matrix. Using the top DAC-DAG pairs of all clusters, we constructed the Disease-gene IMmune cell Expression (DIME) network for the 12 IMIDs. Detailed description of the DIME method is as follows.

Mapping Disease-Gene Network to Immunome Data

For a given disease D and its DAG, we first extracted the corresponding *immunome* expression matrix (X_D) . X_D comprised the gene expression of the DAG across the 40 cell types. X_D was used as an input matrix for the NMF algorithm.

Using NMF to Cluster X_D Into k Classes

We used the NMF package (30) in R and applied Brunet's NMF algorithm (29) on X_D to factor it into two matrices, namely W_D and H_D such that.

$$X_D \approx W_D H_D \tag{1}$$

$$W_{D}H_{D} = \begin{bmatrix} | & | & | & | \\ w_{D}^{1} w_{D}^{2} \dots w_{D}^{k} \\ | & | & | & | \end{bmatrix} \begin{bmatrix} -h_{D}^{1} - \\ -h_{D}^{2} - \\ -h_{D}^{2} - \\ -\vdots - \\ -h_{D}^{k} - \end{bmatrix}$$
(2)

$$W_D H_D = \sum_{i=1}^k w_D^i h_D^i; \ i \in \{1, ..., k\}$$
(3)

Where W_D and H_D are the basis and coefficient matrices computed by NMF. Here, k is the number of classes/clusters that splits the data such that it satisfies the above NMF equations. The W_D matrix comprises the weights of the DAGs across the k clusters (in each column), and the H_D matrix comprises the weights of the cells in the corresponding k clusters (in each row). We used Brunet's method to identify the ideal k value using the cophenetic correlation coefficient (29).

Identifying the Top DAG and DAC From W_D and H_D

The NMF algorithm clusters the data into k clusters (as shown in Equations 2 and 3) such that, in each cluster 'i', where $i \in \{1, ..., k\}$, the genes that have high values in w_D^i are constitutively expressed by the cells that have high values in h_D^i . Where, w_D^i is the *i*th column of W_D and h_D^i is the *i*th row of H_D . We used the scaled (between 0 and 1) values of h_D^i and w_D^i as the DAC and DAG scores, respectively. For each cluster *i*, we chose the DACs and DAGs that were in the top 25th percentile range of their DAC and DAG scores, respectively. These filtered DACs and DAGs are regarded as the top DACs and DAGs, respectively. The top DACs and DAGs were extracted for all clusters of $i \in \{1, ..., k\}$. The DIME network was constructed using the top DAC-DAG pairs from all clusters.

Identifying the Top Cluster

We then identified the largest weighted cluster (referred to as the top cluster) among the *k* clusters identified by the NMF. That is, the subset of DACs and DAGs of X_D that can capture most of its expression pattern. We did this by calculating the Frobenius norm of each $w_D^i h_D^i$ for all values of $i \in (1, ..., k)$ from Equation 3. We then identified the top cluster for which $||w_D^i h_D^i||_F$ is the maximum. This can be represented as:

top cluster = argmax(
$$||w_D^i h_D^i||_F$$
); $i \in \{1, \dots k\}$ (4)

Where the top cluster represents that which maximally captures/represents the expression matrix X_D . Thus, the top cluster is the rank 1 cluster of DIME. Subsequent ranks are the next highest weighted clusters.

Evaluating the Consistency of Top DACs and DAGs Identified by DIME

To check the consistency of the results from DIME, we performed 1000 jackknife simulations for each of the 12 IMIDs. For each simulation of each disease, we ran DIME with 70% random subsampling of either the DACs or the DAGs. And in each simulation, we identified the top cluster and the top DACs when DAGs were subsampled and vice versa. We compared the consistency of the top DACs (**Supplementary Figure S2**) and the top 10 DAGs (**Supplementary Figure S3**) identified by the original DIME run (100% of the sample) against the 1000 simulations. We computed the Pearson correlation coefficient between the DAC/DAG score of the top cluster of the original run to the number of times the DAC/DAG was found as the top DAC/ DAG in the top cluster of the 1000 simulations. We used the pvalue from the Pearson correlation test to state the significance of correlation and thus the statistical significance of the top DAC/DAG of the top cluster.

The Common Cell-Gene Network Between Diseases

To identify the common cell-gene network between two diseases, we looked at their overlapping DAC-DAG pairs in their corresponding DIME networks. We refer to the overlapping DAC-DAG pairs as the common cell-gene network between the two diseases. We then used the Jaccard index (JI) to measure the overlap between the two diseases and Fisher's exact test (FET) to obtain a confidence p-value for the given overlap.

Integrating Drug-Gene Network

We extracted the drug-gene target network from (1) DGIdb with the filter set to contain CHEMBL interactions of the drugs approved by the food and drug administration (FDA) of USA (31) (2); all drug-gene of CLUE database (32) and (3); all druggene of hPDI (33). The genes with drugs associated with them are labeled in the common cell-gene networks to highlight drugability (**Figures 6C–E**).

Statistical Analysis

We performed 1000 jackknife simulations to assess the consistency of the results from the DIME (Supplementary Methods and **Supplementary Figures S2–S4**). We used the Pearson correlation coefficient to measure the significance of the jackknife simulations compared to the original run (**Supplementary Figure S4**).

RESULTS

The Disease-Gene Network of the 12 IMIDs Reveal Several Common DAGs

In this study, we analyzed the DAGs of 12 different types of IMIDs that broadly include inflammatory arthropathies, spondyloarthropathies, rheumatic diseases, systemic IMIDs, and inflammatory bowel diseases (IBDs). And specifically, the 12 different IMIDs studied here are ankylosing spondylitis (298 DAGs), arthritis (567 DAGs), Crohn's disease (786 DAGs), diabetes mellitus - non-insulin-dependent (1415 DAGs), systemic lupus erythematosus (963 DAGs), multiple sclerosis (961 DAGs), psoriasis (689 DAGs), psoriatic arthritis (177 DAGs), rheumatoid arthritis (1612 DAGs), Sjogren's syndrome (229 DAGs), systemic scleroderma (494 DAGs), and ulcerative colitis (796 DAGs) (Figures 1A, B). The 12 IMIDs had a total of 3957 DAGs. Among these, several genes were linked to several IMIDs; for example, 74 DAGs were linked to only Crohn's disease (CD) and ulcerative colitis (UC), both IBDs. Calculating the Jaccard index and Fisher's exact test (FET) on all the overlapping DAGs between all IMIDs revealed that CD and UC had the highest significant overlap (Figure 1C). Interestingly, genes associated with CD had significant overlap (FET p-value \leq 0.05) with all diseases except psoriatic arthritis and non-insulindependent diabetes mellitus. Rheumatoid arthritis (RA) had a significant overlap of DAGs with all IMIDs except non-insulindependent diabetes mellitus. However, non-insulin-dependent





diabetes mellitus did not have a significant overlap of DAGs with any of the IMIDs. Arthritis, psoriasis, CD, and RA had a significant overlap of DAGs between each other. We found 12 DAGs that were associated with all the 12 IMIDs (Figures 1A, E). These DAGs were related to processes typically associated with inflammation, such as cytokine signaling (GO:0001817; GO:0019221), regulation of inflammatory response (GO:0050727), and regulation of interleukin-6 (GO:0032675; GO:0032635). We further explored these DAGs in the immunome and found the expression of TNFAIP3 to be the highest in CD8⁺ T-cells, ILC3 and CD4⁺ Tcells (Figures 1D, E). Likewise, IL1B was expressed by myeloid and progenitor cells; TNF was expressed by lymphoid and myeloid cells. Overall, specific myeloid and lymphoid cells specifically expressed some of the 12 genes linked to all the 12 IMIDs. Such cell-specific expression of the DAGs led us to question the immune cell types and genes essential for the 12 IMIDs. Hence, we used the DIME on the 12 IMIDs to identify their top DACs and DAGs. Briefly, DIME uses the immunome, input disease-gene network, and an unsupervised machine learning algorithm (NMF) to determine the clusters of top DACs and DAGs.

Inflammatory Arthritis Is Driven by CD4⁺ Treg, CD4⁺ Th1, and NK Cells

Inflammatory arthritis is an autoimmune disorder characterized by joint inflammation. And joint inflammation is the primary clinical feature observed in inflammatory arthritis types such as ankylosing spondylitis (AS) and RA. However, in other inflammatory arthritis types, such as psoriatic arthritis, inflammation is present in both the skin and joints. Interestingly, AS and psoriatic arthritis are both seronegative spondyloarthropathies (negative for rheumatoid factor and auto nuclear antibodies) characterized by enthesitis and predominant HLA-B27 genotype (34, 35). Such shared clinical features led us to question if the inflammatory arthritis types shared molecular mechanisms. So, we performed DIME on the different types of inflammatory arthritis to identify the critical DACs and DAGs and compare the molecular mechanism shared between them. As a reference, we used the broader arthritis disease term that encompassed several different kinds of arthropathies (including inflammatory arthritis).

In the DIME analysis, the clusters of the DIME network (Figures 2-5) are ordered based on the Frobenius norm. Cluster with the highest Frobenius norm represents the most crucial cluster and is designated as the top cluster comprising the most crucial DACs and DAGs. In each cluster, the DACs and DAGs are ordered based on the DIME score, high scoring nodes signify higher importance. The DIME analysis of AS revealed lymphoid cells such as NK cells, ILC3, CD4⁺ T-cells (Th1, Treg, TEMRA) as the top DACs in the top cluster (Figure 2A). The top cluster's top DAGs comprised ETS1, HSPA5, TNFAIP3, IL2RG, WNK1, etc. that were associated with pathways such as interleukin signaling, antigen presentation, regulation of RUNX3, and BCR signaling (Figure 2C). We find RUNX3 expression highest in the NK cells, followed by CD8⁺ Tcells and Th1 cells (Supplementary Figure S1A). The exact role of RUNX3 in AS is unclear and possibly involves regulation, differentiation, and activation of Th1 and T-bet cells (36). Further research is required to establish the exact role of RUNX3 in AS and

the above-identified lymphoid cell subsets. In the second cluster, the top DACs included myeloid cells, and the top DAGs were associated with pathways such as interleukin (IL-4, IL-10, IL-13) signaling, MAPK3 activation, and MyD88 (**Figures 2A, C**). Thus, the key DACs of AS are diverse, as reported in the literature (37). However, according to DIME, the top DACs were NK cells, ILC3, $CD4^+$ T-cells (Th1, Treg, TEMRA).

The DIME analysis of psoriatic arthritis revealed lymphoid cells such as NK cells, ILC3, T-cells (CD8⁺, TEMRA, Th1), ILC2 and myeloid cells like the macrophages and BDCA1⁺ DC as the top DACs in the top cluster (Figure 2B). Likewise, T-cells, NK cells, and antigen-presenting cells have played a crucial role in psoriatic arthritis pathology (38). The top cluster's top DAGs were associated with interleukin (IL-4, IL-10, IL-13) signaling, PI3K, and NF-KB activation. (Figure 2D). Furthermore, the downstream genes of TNF-alpha such as TNFAIP3, TRAF5, NFKB1, and ICAM1 were top DAGs in the top cluster. Interestingly, TNF-alpha is a therapeutic target for psoriatic arthritis (39, 40), perhaps the downstream genes identified by DIME could also be explored as a therapeutic target for psoriatic arthritis. S100 calcium-binding proteins like S100A8 and S100A9 play a role in regulating inflammation in psoriatic arthritis (41). In the second cluster, we found the top DAGs included the S100 calcium-binding proteins, such as S100A9 and S100A8, which were highly expressed by neutrophils, granulocytes, monocytes, and dendritic cells (Figure 2B and Supplementary Figure S1B).

The crucial immune cells involved in RA are T-cells, B-cells, and APCs (42). While activation of CD4⁺ Th1 and impairment of CD4⁺ Tregs are essential for the pathology of rheumatoid arthritis (43), the DIME analysis of RA revealed several lymphoid cells such as CD4⁺ Tregs, CD4⁺ Th1, NK cells, etc., as the top DACs in the top cluster (**Figure 3A**). The top cluster's top DAGs were associated with pathways such as interleukin, TCR, FCERI, and BCR signaling (**Figure 3C**). In the second cluster, the top DACs included myeloid cells, and the top DAGs were associated with pathways such as interleukin (IL-10, IL-13) signaling, neutrophil degranulation, and ECM organization (**Figures 3A, C**). Activation, recruitment, and apoptosis of neutrophils are altered in RA, and under chronic inflammatory conditions, they release protease-rich granules (44).

The DIME analysis of the broader arthritis disease term revealed macrophages as the top DAC in the top cluster (**Figure 3B**). Macrophages play a central role in arthropathies, where they release cytokines and activate several immune cells such as T-cells, monocytes, neutrophils, and synovial fibroblasts. Besides, they are also the most abundant cells at the site of inflammation (45). The top DAGs of the top cluster were associated with pathways such as interleukin (IL-4, IL-13) signaling, extracellular matrix (ECM) related pathways, neutrophil degranulation, and toll-like receptor (TLR) cascades (**Figure 3D**). In the second cluster, the top DACs comprise neutrophils, granulocytes, and the top DAGs associated with pathways similar to the top cluster and included inflammasomes-related pathways (**Figures 3B, D**).

In conclusion, using DIME, we found that in addition to the shared clinical features, the three inflammatory arthritis types



also had a similar DAC profile comprising CD4⁺ Treg, CD4⁺ Th1, and NK cells as the top DACs, distinguishing them from the broader arthritis disease that showed macrophages as its top DAC. Perhaps these lymphoid cells contribute to inflammation in these arthropathies, and targeting them to reduce inflammation could be explored as a therapeutic strategy (46).

Myeloid Cells Are Essential to Systemic IMIDs

We performed the DIME analysis on the systemic IMIDs such as systemic lupus erythematosus (SLE) and systemic scleroderma

(SSc) (**Figure 4**). SLE and SSc are type I interferon-mediated systemic autoimmune diseases, that unlike RA, primarily affect not just the joints but also the skin, kidney, heart, and other organs (47). In SLE, the continuous IFN production by pDC and neutrophils leads to activation of monocytes, T-cells, and B-cells (48). The DIME analysis of SLE revealed the myeloid cells (granulocytes, macrophages, BDCA1⁺ CD14⁺, monocytes) as the top DACs in the top cluster (**Figure 4A**). The top DAGs in the top cluster were associated with interleukin signaling pathways (IL-4, IL-13), neutrophil degranulation, cell-surface interactions at the vascular wall, and the TLR cascades



(Figure 4C). Incidentally, the neutrophils in SLE undergo spontaneous NETosis (a form of suicidal cell death), and this process is dependent on TLR signaling (48). Additionally, T-cells in SLE have altered cytokine production with higher IL6, IL7, and IL10 secretions (48). In the second cluster, we found the top DACs included CD4⁺ T-cells (TEMRA, TEM, TCM), and the top DACs were associated with pathways such as immunoregulatory interactions, Nef-associated factors (TNIP1, TNFAIP3), ZAP-70, VAV1 pathway (Figures 4A, C). Nef-associated factors (TNIP1,

TNFAIP3) have played a role in T-cell activation *via* TCR signaling in SLE (49).

The DIME analysis of SSc revealed myeloid cells (neutrophils, granulocytes, BDCA1⁺ CD14⁺ cells) and lymphoid cells (NK cells, CD4⁺ Treg, ILC2, and ILC3) as the top DACs in the top cluster (**Figure 4B**). The top DAGs in the top cluster were associated with interleukin signaling pathways (IL-4, IL-13), TGF beta signaling, NLR signaling, etc. (**Figure 4D**). In the second cluster, the top DACs included macrophages and the top



DAGs were associated with pathways that included IL-10 signaling and degradation of ECM (**Figures 4B, D**). As described in the review by Caam et al., several studies have shown neutrophils, macrophages, NK cells, and Tregs to play a role in the profibrotic events in SSc by the production of profibrotic cytokines such as TGF beta, IL-4, IL-10, IL-13, etc., thus corroborating our findings (50).

Thus, the top DACs of the systemic IMIDs comprised myeloid cells such as neutrophils, granulocytes, BDCA1⁺ CD14⁺ cells, CD11c⁺ myeloid cells, and BDCA1⁺ DC.

Exploring the role of these cells, their corresponding DAGs and pathways in the systemic IMIDs may be useful for gaining mechanistic insights into disease and for successful therapeutic strategy. Exploring the role of neutrophils and their degranulation genes may serve as stronger targets in SLE, as neutrophil degranulation precedes type 1 interferon signature (often observed in SLE patients) (51–54). While the DACs of the SSc are diverse populations of myeloid and lymphoid cells, neutrophils are still an interesting candidate and exploring the role of TGF beta may shed light on the profibrotic events in SSc. Thus, based on the DIME analysis, attenuation of neutrophilic inflammation may be the key therapeutic strategy for the systemic IMIDs.

IBDs Are Primarily Lymphoid Driven

We then looked at IMIDs that involve chronic inflammation of the digestive system, categorized as IBDs. The two primary forms of IBDs are CD and UC. CD is known to be driven by CD4⁺ Th1 cells, with a dominant Th1 cytokine profile leading to a pro-inflammatory effect (55). The DIME analysis of CD revealed lymphoid cells (CD4⁺ Treg, ILC2, CD4⁺ TEMRA, CD4⁺ Th1) as the top DACs in the top cluster (**Figure 5A**). The top DAGs of the top cluster were associated with pathways such as interleukin (IL-4, IL-10, IL-13) signaling, TLR (TLR-5, TLR-10) signaling, MyD88, and neutrophil degranulation (**Figure 5C**). In the second cluster, the top DACs included granulocytes, neutrophils, monocytes, macrophages, etc., and the top DAGs were associated with pathways such as



interleukin signaling, neutrophil degranulation, and TLR cascades (Figures 5A, C).

The T-cell profile of UC has been difficult to categorize due to discrepancies in its response among patients. However, there is evidence of Th2 cells, NK cells, macrophages, and neutrophils involved in the pathogenesis of UC (55). The DIME analysis of UC revealed lymphoid cells (ILC2, NK, ILC3, CD4⁺ Th1, etc.) as the DACs in the top cluster (Figure 5B). The top DAGs of the top cluster were associated with pathways such as interleukin (IL-4, IL-13) signaling, TLR cascades, NLR signaling, neutrophil degranulation, etc. (Figure 5D). In the second cluster, the top DACs included granulocytes, BDCA1⁺ CD14⁺ cells, etc. The top DAGs were associated with interleukin signaling pathways (IL-4, IL-10, IL-13), neutrophil degranulation, and TLR cascades (Figures 5B, D). The DIME analysis of UC comprised 5 clusters in total, with myeloid cells in the third cluster, thrombocytes and TEMRA in the fourth cluster, and bone marrow plasma cells in the fifth cluster. This shows that the disease-gene network of UC involves genes that participate in many different cell types and are more complex to uncover their etiology. However, the highest weighted cluster signifies that the lymphoid cells may be the primary candidates for further analysis.

Thus, the top DACs of the IBDs include the ILC2, CD4⁺ Th1, CD4⁺ Treg, and the NK cells. The T-cells, innate lymphoid cells, and NK cells play an important role in the pathogenesis of IBDs (56–58), thus corroborating our findings. We hypothesize that Crohn's disease is driven primarily by lymphoid inflammatory response and downregulating TLR signaling pathways could be a potential therapeutic strategy. The DAC profile of ulcerative colitis was found to be diverse, and would require a deeper analysis into the participation of the different cell types involved in their etiology.

Statistically Significance of DIME Results

To evaluate DIME's consistency, we performed 1000 Jackknife simulations with random subsampling of DAC/DAG. We reidentified the top DAC/DAG for all IMIDs (see Supplementary Methods). The jackknife simulations revealed that the top DACs identified across all clusters in the simulations (Supplementary Figure S2A) showed a similar pattern compared to top DACs identified in the original run (Supplementary Figure S2C). For the top DACs of the top cluster, the pattern from the simulations (Supplementary Figure S2B) was comparable to the original run's DAC score (Supplementary Figure S2D). We used Pearson correlation to compare the pattern observed between the simulations and the original run. The Pearson correlation between the pattern observed in the simulated run (Supplementary Figure S2B) and the DAC scores of the original run for the top cluster revealed that the top DACs in the top cluster were significantly correlated (p-value ≤ 0.05) for all the IMIDs except ulcerative colitis (Supplementary Figure S4). This correlation shows that the top DACs of the top cluster identified by DIME is statistically significant for all IMIDs, except UC.

Likewise, we evaluated the consistency of the top DAGs. In all simulations, the top 10 DAGs of the original run's top cluster

were present as the top DAG in any of the simulated run clusters (**Supplementary Figure S3**). The Pearson correlation between the pattern observed in the simulated run and the DAG scores of the original run for the top cluster was significantly correlated for all the IMIDs (**Supplementary Figure S4B**). Thus, the top DAGs of the top cluster identified by DIME is statistically significant for all IMIDs.

Why Are the top DACs of UC Insignificant?

In the case of UC, the top DACs were statistically insignificant from our 1000 jackknife simulations; the top DAGs, however, were significant (Supplementary Figures S2-S4). We found from 1000 simulations that the lymphoid cells identified by the original run (Figure 5B) were indeed present in the simulations. The myeloid cells were also part of the top DACs of the top cluster in the simulations (Supplementary Figure S2B). Furthermore, we found that the top cluster's top DAGs included genes associated with neutrophil degranulation pathways and other myeloid cell-related pathways (Figures 5B, D). Thus, owing to the non-convergence of NMF in accurately predicting the top DACs of the top cluster in UC. The top DACs of the top cluster of UC were ambiguous, as reported in the literature (55). From our simulations, we propose the inclusion of the myeloid cells in the top DACs of the top cluster in addition to the lymphoid cells previously identified (Figure 5B).

Common Cell-Gene Networks Reveal Common Mechanisms Between IMIDs and Potential Drug Targets

The DIME analysis revealed that several top DAGs and their corresponding DACs were present in many IMIDs. For example, in many IMIDs, the gene FOS was present as top DAG in the cluster typically containing myeloid cells (granulocytes, neutrophils, and dendritic cells) as the top DACs. We found several genes, like FOS, that were present as the top DAG in the same top DAC cluster between different pairs of diseases. We refer to these top DACs and DAGs present between the two diseases as the common cell-gene network (represented schematically in Figure 6A). Using the common cell-gene network, we suggest that these diseases may have a similar mechanism of action. We could exploit such common mechanisms to gain mechanistic insights between diseases and identify drug repurposing targets. Hence, we integrated the publicly available drug-gene networks to identify and reinforce drug repurposing targets based on the common mechanisms (cell-gene networks) determined from the DIME analysis (Figure 6A).

To identify the common mechanisms across the 12 IMIDs, we identified the common cell-gene networks between all disease comparisons (**Figure 6B**). We then used the Jaccard index and FET to measure the extent and significance of the overlap in the common cell-gene networks between the disease pairs. Compared to the analysis that looked at all DAGs, which showed several diseases to be statistically significant in the overlap, the common cell-gene network overlap was restricted to fewer diseases (**Figure 1C** and **Figure 6B**).

The comparative analysis revealed that CD had statistically significant common cell-gene networks with several diseases such as, psoriasis, RA, and UC (**Figure 6B**). CD and UC's common cell-gene network had the highest Jaccard index among all the IMIDs, both being IBDs with an aggressive T-cell response (55). CD and UC's common cell-gene network revealed that the top DACs included the lymphoid cells such as CD4⁺ T-cell, CD4⁺ Th1, CD4⁺ Treg, ILC1, ILC2, ILC3, and NK cells in one cluster (**Figure 6C**). CD4⁺ Th1 and NK cells are known to be implicated in both CD and UC (55). The top DAGs (represented by green border) such as CD44, CXCR4, SELL, HSP90AA1 etc., were highly expressed by cells of the

lymphoid cluster and were also drug targets (genes that are druggable or have drug targeting them in our drug-gene network). Thus, making them potential drug repurposing candidates for CD and UC. Other potentially interesting DAGs (that were not drug targets) included PTPRC, ETS1, IL7R, TNFAIP3, etc. Collectively, the DAGs in the lymphoid cluster were enriched in interleukin signaling pathways (IL-4 and IL-13), NLR signaling, etc. (**Supplementary Figure S5A**). The other clusters consisted of myeloid cells such as the granulocytes, dendritic cells, monocytes, and macrophages, among which dendritic cells have been crucial for regulating the T-cell responses in IBDs. The top DAGs, such as SH2B3,





12

JAK2, CD86, MMP9, CXCL8, PTGS2, TREM1, LRRK2, TYK2, etc., were highly expressed by the myeloid cluster cells and were potential drug repurposing candidates. These DAGs were enriched in interleukin signaling pathways (IL-10), TLR signaling, ECM degradation, etc. (Supplementary Figure S5A).

We next explored the common cell-gene network of the two distinct IMIDs that belonged to different pathophysiology, namely CD and RA. The common cell-gene network of CD and RA revealed that the top DACs comprised of the lymphoid cells that included all CD4⁺ T-cells and NK cells in one cluster (Figure 6D). The top DAGs such as CXCR4, CD44, SELL, RAC2, ANXA1, etc., were highly expressed by this cluster's cells and were potential drug repurposing candidates. These DAGs were enriched for pathways associated with interleukin, TLR, MyD88 signaling, etc. (Supplementary Figure S5B). The other cluster comprised myeloid cells such as granulocytes, dendritic cells, monocytes, and macrophages. The top DAGs such as CTSS, ITGB2, MCL1, LYZ, SH2B3, etc., were highly expressed by this cluster's cells and were potential drug repurposing candidates. These DAGs were enriched for pathways associated with interleukin (IL-4, IL-13) signaling, neutrophil degranulation, etc. (Supplementary Figure S5B).

In addition to CD's common cell-gene networks, we found a statistically significant common cell-gene network between the two inflammatory arthropathies that have joint pain as the primary feature, namely AS and RA. The common cell-gene network of AS and RA revealed that the top DACs comprised of the lymphoid cells that included all the T-cells and NK cells in one cluster (Figure 6E). The top DAGs such as IL2RG, ITGAL, IL2Rb, PTGER4, PRKCQ, etc., were highly expressed by this cluster's cells and were potential drug repurposing candidates. These DAGs were enriched for interleukin (IL-1) signaling pathways, FCERI mediated NF-kB activation, TCR signaling, etc. The other clusters comprised myeloid cells such as granulocytes, dendritic cells, monocytes, and macrophages. The top DAGs, such as SH2B3, TYK2, TLR2, TLR4, IL6R, were highly expressed by this cluster's cells and were potential drug repurposing candidates. These DAGs were enriched for interleukin signaling pathways (IL-4, IL-10, IL-13) (Supplementary Figure S5C).

Thus, using the common cell-gene networks, we could uncover the common mechanisms between the IMID pairs (**Figure 6** and **Supplementary Figure S5**) and use them to identify potential drug targets. This novel method of computational drug repurposing is a combination of targetbased and mechanism-based drug repurposing strategies (59). We found several DAGs such as IL1B, IL6R, ITGAL, PTGS2, TYK2, NFKB1, NLRP3, PRKCQ, PTGER4, PTPN2, RELA, SH2B3, SMAD3, TLR2, TLR4, and TREM1, that were drug targets and present in all the common cell-gene networks (**Figures 6C, E**). Among these DAGs, ITGAL was the only DAG that was a drug target and present as the top DAG of the top cluster (lymphoid cell cluster) in the DIME networks of CD, UC, AS, and RA. This is interesting as the lymphoid cells were identified as the top DAC of the top cluster for all of the above diseases. Using the drugs associated with these drug targets specifically for these diseases (CD, UC, AS, and RA) in therapy would require extensive experimental validation and clinical trials. Therefore, we explored (in the next section) the possibility of using some of these drug targets for repurposing based on existing studies and drugs that are already approved by the FDA. Thus, reinforcing and strengthening these targets and also the validity of our approach in identifying them.

Common Cell-Gene Networks Reveal Drug Repurposing Targets

To explore and validate the drug targets for repurposing, we focused on the top DAGs of the statistically significant (FET p-value ≤ 0.05) common cell-gene networks of all IMIDs (**Figure 6B**). To identify drug targets that were targets of FDA-approved drugs, we used the drug-gene network of CHEMBL. We found several drug targets such as IL1B, IL6R, ITGAL, and TYK2 to be present in all the statistically significant common cell-gene networks (**Table 1**). Here, we explore the possibility of using these drug targets for repurposing based on their current use as a therapeutic target in certain IMIDs and suggest where they can be repurposed based on their presence in the common cell-gene networks of those IMIDs and others.

Drug targets identified by us that are already in use for the different IMIDs include drugs that target IL1B, IL6, TYK2, and JAK2.IL1B was present in the common cell-gene network of AS, CD, psoriasis, RA, and UC, indicating a similar mechanism between these diseases (Table 1). Anti-IL1 therapy is used for psoriasis and RA (60-62), suggesting that it could be potentially used in IMIDs like AS, CD, and UC. Incidentally, preliminary studies indicate that anti-IL1 treatment has shown promising clinical response for treating AS, CD, and UC (63, 64). Anti-IL6 therapy (tocilizumab) showed a positive clinical response in a small group of patients in AS, CD, and RA, suggesting its application in other IMIDs like psoriasis and UC (65-67). However, anti-IL6 therapy had side effects in smaller studies on psoriasis and UC and must be explored carefully (68, 69). Tofacitinib, a TYK2 and JAK2 inhibitor developed for RA, is now making way to treatment options in other diseases such as, CD, UC, and psoriasis (70-73). Plerixafor (drug target: CXCR4) is a drug now used in cancer (lymphoma and multiple myeloma) after stem cell transplantation to initiate migration of stem cells in the bloodstream (74). This drug is now in clinical trials (NCT01413100) to be evaluated for use after autologous transplant in patients with SSc. We suggest extending such trials based on exploiting the CXCR4 mediated dysregulation of the immune system to other IMIDs like psoriasis, CD, RA, and UC.

Integrin based therapies (such as natalizumab and vedolizumab that targets ITGB2) are already used for CD (75). Exploring other integrin based therapies (such as Lifitegrast that targets ITGAL and ITGB2) for CD may be beneficial since ITGAL and ITGB2 are top DAGs in the DIME network and are also implicated in CD (76, 77). Lifitegrast could be a promising drug repurposing candidate for CD and perhaps for

TABLE 1 | The top DAGs and their FDA approved drug candidates identified from the DIME-based common cell-gene networks of the different IMIDs.

DAG/ Drug target	Diseases	Drugs
IL1B	AS, CD, Psoriasis,	canakinumab, rilonacept, anakinra
IL6R	RA, UC AS, CD, Psoriasis,	tocilizumab
ITGAL	RA, UC AS, CD, Psoriasis,	lifitegrast
TYK2	RA, UC AS, CD, Psoriasis,	tofacitinib citrate
PSMB9	RA, UC AS, CD, Psoriasis,	bortezomib, carfilzomib, ixazomib citrate
DNMT3A	AS, CD, Psoriasis,	azacitidine, decitabine
HDAC7	AS, CD, Psoriasis,	belinostat, panobinostat lactate, romidepsin
JAK2	AS, CD, Psoriasis,	baricitinib, ruxolitinib phosphate, tofacitinib citrate
PTGS2	AS, CD, RA, UC	acetaminophen, aminosalicylate potassium, aminosalicylate sodium, aspirin, balsalazide disodium, bismuth subsalicylate, bromfenac sodium, carprofen, diclofenac, diclofenac epolamine, diclofenac potassium, diclofenac sodium, diflunisal, etodolac, etoricoxib, fenoprofen calcium, flurbiprofen, flurbiprofen sodium, ibuprofen, ibuprofen lysine, ibuprofen sodium, indomethacin, indomethacin sodium, ketoprofen, ketorolac tromethamine, meclofenamate sodium, meloxicam, mesalamine, nabumetone, naproxen, naproxen etemesil, naproxen sodium, nepafenac, electronethamine, organzatio, organzatio potassium, piravicam, eufgaglazio, euficada talmatin sodium.
BCL2	CD, Psoriasis,	venetoclax
CXCR4	RA, UC CD, Psoriasis,	plerixafor
IL4R	RA, UC CD, Psoriasis,	dupilumab
IL17RA	RA, UC CD, Psoriasis,	brodalumab
ITGB2	RA CD, Psoriasis,	lifitegrast
PSMD7	RA CD, Psoriasis,	bortezomib, carfilzomib, ixazomib citrate
CD86	RA CD, RA,	abatacept, belatacept
CSF2RA	UC CD, RA, UC	sargramostim
NR3C1	CD, RA, UC	alclometasone dipropionate, amcinonide, beclomethasone dipropionate, betamethasone, betamethasone acetate, betamethasone benzoate, betamethasone dipropionate, betamethasone sodium phosphate, betamethasone valerate, budesonide, ciclesonide, clobetasol propionate, clocortolone pivalate, cortisone acetate, deflazacort, desonide, desoximetasone, dexamethasone, dexamethasone acetate, dexamethasone sodium phosphate, flumethasone pivalate, flunisolide, fluocinonide, fluorometholone, fluorometholone acetate, diflupredniace, fluticasone furoate, fluticasone propionate, hydrocortisone acetate, hydrocortisone butyrate, hydrocortisone cypionate, hydrocortisone probutate, hydrocortisone sodium phosphate, hydrocortisone sodium succinate, hydrocortisone valerate, loteprednol etabonate, medrysone, meprednisolone, methylprednisolone, methylprednisolone sodium succinate, prednisolone sodium succinate, prednisolone sodium phosphate, prednisolone acetate, prednisolone diacetate, triamcinolone hexacetonide

(Continued)

TABLE 1 | Continued

DAG/ Drug target	Diseases	Drugs	
P4HB	CD, RA, UC	lomitapide mesylate	
IL2RB, IL2RG	AS, RA	basiliximab, daclizumab	
PSMB8 ALOX5	AS, RA CD, RA	bortezomib, carfilzomib, ixazomib citrate balsalazide disodium, meclofenamate sodium, mesalamine, olsalazine sodium, sulfasalazine, zileuton	

UC, AS, and RA. Since its target gene ITGAL, was the only top DAG of the top cluster (lymphoid cell cluster) that was also a drug target in the DIME networks of these diseases (**Figures 2**, **3**, **6C–E**). Thus, we propose liftegrast as a novel drug repurposing candidate to be tested for CD, UC, AS, and RA.

DISCUSSION

Despite decades of experimental data, the knowledge of important cell types involved in the disease's pathogenesis remains limited. To address this gap, we used the *immunome* comprising 40 immune cells, the disease-gene network, and computational methods to identify the important DACs and DAGs of the disease. The integration of these parts resulted in the novel mechanisms being captured by our method, using which we built a tool called the DIME. Here, we highlight the important DACs, DAGs, and common mechanisms captured using DIME for 12 phenotypically different IMIDs. Using DIME, the top DACs were found to be CD4⁺ Treg, CD4⁺ Th1, and NK cells in inflammatory arthritis (AS, PsA, and RA); neutrophils, granulocytes, and BDCA1⁺ CD14⁺ cells in SLE and SSc; ILC2, NK, CD4⁺ Th1, and CD4⁺ Treg in the IBDs.

Lymphoid cells such as CD4⁺ Th1, CD4⁺ Treg, and NK cells were the key players in inflammatory arthritis (AS, PsA, and RA) and IBD (CD and UC). These diseases have been reported to have an intricate cross-play of the above lymphoid cells, where the NK cells influence the differentiation of CD4⁺ Th cells into CD4⁺ Th1 and CD4⁺ Tregs; CD4⁺ Th1 plays a crucial role in the initiation of inflammation by cytokine production; the CD4⁺ Tregs are crucial for immune response modulation (78). Interestingly, the top DAGs of these diseases show pathways associated with the signaling of IL-4 and IL-13 that are crucial in this cross-play, thus corroborating DIME results.

Although we excluded HLA genes to prevent myeloid and B cell bias, the IMIDs associated with the HLA-B27, such as psoriasis, AS, and IBDs were found to have statistically significant common-cell gene networks. However, PsA (also associated with HLA-B27) is omitted here since it did not have a statistically significant common-cell gene network with any IMIDs (**Figure 6B**). Additionally, AS and RA, the two inflammatory arthritis with joint inflammation as the primary feature, also had a statistically significant common-cell gene network. Thus, the diseases with these shared clinical features also had common mechanisms as identified by DIME. The

common mechanisms from these networks revealed several lymphoid and myeloid cells and their expressing DAGs. The lymphoid cells such as CD4⁺ Th1, CD4⁺ Treg, and NK were predominant in all the statistically significant common-cell gene networks, showing that these diseases were indeed mainly driven by the aggressive T-cell response (36–38, 55). Pathways such as interleukin (IL-4 and IL-13), TLR, TCR signaling, etc., was found to be enriched in the top DAGs of the common cell-gene networks of these IMIDs. Thus, the common cell-gene network revealed several common mechanisms between the diseases in accordance with the top DACs, DAGs, and their associated pathways.

We used the common mechanism from the common cellgene network and the drug-gene networks to propose potential drug targets for repurposing. This novel computational drug repurposing strategy, a combination of target-based (literature drug-gene network) and mechanismbased (inferred from DIME), revealed several potential drug targets such as IL1B, IL6R, ITGAL, PTGS2, TYK2, NFKB1, NLRP3, PRKCQ, PTGER4, PTPN2, RELA, SH2B3, SMAD3, TLR2, TLR4, and TREM1. Further, we used these mechanismbased drug targets from DIME, and the FDA approved druggene network to propose several drug targets and their drugs that could expedite the drug repurposing process (Table 1). Thus, we were able to capture drug targets and their drugs currently being targeted or being explored for use in therapy for the IMIDs. We also found a few novel targets, such as the drug lifitegrast (used for dry eyes) for CD, UC, AS, and RA as an alternative to other integrin-based therapies already in use for CD. Lifitegrast is particularly interesting because it targets ITGAL, which was found to be important in the lymphoid cell cluster of CD, UC, AS, and RA. Thus, effectively targeting the exact mechanism. Perhaps the effect of lifitegrast could be used for down-regulating lymphoid cellmediated inflammation in these diseases (79). Although Lifitegrast is currently available as an eye drop application and used to treat eye complications only, different formulations of this drug can be explored to treat CD, UC, AS, and RA. To the best of our knowledge, lifitegrast in the axis of ITGAL has not been investigated to treat CD, UC, AS, and RA. Thus, using DIME, we were able to propose a novel drug repurposing strategy from the analysis of the 12 IMIDs.

In conclusion, DIME helped identify (1) top DACs, DAGs of the IMIDs (2), Common mechanisms between the IMIDs, and (3) drug targets for repurposing. To enable DIME analysis

for other diseases from the DisGeNet, the GWAS network, or a user-defined set of genes, we built the DIME tool as a userfriendly shinyapp. We believe that this tool will help scientists uncover the etiology of complex and rare diseases and facilitate drug development by better-determining drug targets, thereby mitigating the risk of failure in late clinical development.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. The DIME tool is available on https://bitbucket.org/systemsimmunology/dime. Further inquiries can bedirected to the corresponding author.

AUTHOR CONTRIBUTIONS

AD and AP were involved in the conception of the study. AD was involved in the data curation, visualization, and R shiny package development. AD and AP were involved in the data analysis and interpretation. AD and AP drafted the manuscript. TR helped in writing and revising the manuscript and discussions about

REFERENCES

- Sun YV, Hu Y-J. Integrative Analysis of Multi-omics Data for Discovery and Functional Studies of Complex Human Diseases. *Adv Genet* (2016) 93:147– 90. doi: 10.1016/bs.adgen.2015.11.004
- Li Y, Agarwal P. A Pathway-Based View of Human Diseases and Disease Relationships. *PloS One* (2009) 4:e4346. doi: 10.1371/journal.pone.0004346
- Camp JG, Platt R, Treutlein B. Mapping Human Cell Phenotypes to Genotypes With Single-Cell Genomics. Sci (80-) (2019) 365:1401–5. doi: 10.1126/science.aax6648
- Rozenblatt-Rosen O, Stubbington MJT, Regev A, Teichmann SA. The Human Cell Atlas: From Vision to Reality. *Nature* (2017) 550:451–3. doi: 10.1038/ 550451a
- Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The Genotype-Tissue Expression (Gtex) Project. Nat Genet (2013) 45:580–5. doi: 10.1038/ng.2653
- Ardlie KG, Deluca DS, Segre AV, Sullivan TJ, Young TR, Gelfand ET, et al. The Genotype-Tissue Expression (Gtex) Pilot Analysis: Multitissue Gene Regulation in Humans. *Sci (80-)* (2015) 348:648–60. doi: 10.1126/ science.1262110
- GTEx Consortium, Laboratory, Data Analysis &Coordinating Center (LDACC)—Analysis Working Group, Statistical Methods groups—Analysis Working Group, Enhancing GTEx (eGTEx) groups, NIH Common Fund and NIH/NCI, et al. Genetic Effects on Gene Expression Across Human Tissues. *Nature* (2017) 550:204–13. doi: 10.1038/nature24277
- Tirosh I, Suvà ML. Deciphering Human Tumor Biology by Single-Cell Expression Profiling. Annu Rev Cancer Biol (2019) 3:151–66. doi: 10.1146/ annurev-cancerbio-030518-055609
- Keren-Shaul H, Spinrad A, Weiner A, Matcovitch-Natan O, Dvir-Szternfeld R, Ulland TK, et al. A Unique Microglia Type Associated With Restricting Development of Alzheimer's Disease. *Cell* (2017) 169:1276–90.e17. doi: 10.1016/j.cell.2017.05.018
- Kuo D, Ding J, Cohn IS, Zhang F, Wei K, Rao DA, et al. HBEGF+ Macrophages in Rheumatoid Arthritis Induce Fibroblast Invasiveness. *Sci Transl Med* (2019) 11:eaau8587. doi: 10.1126/scitranslmed.aau8587
- Knight JC. Genomic Modulators of the Immune Response. Trends Genet (2013) 29:74–83. doi: 10.1016/j.tig.2012.10.006

clinical perspective. All authors contributed to the article and approved the submitted version.

FUNDING

This project was supported by the Netherlands Organization for Scientific Research (NWO; Grant number 016.Veni.178.027).

ACKNOWLEDGMENTS

The authors thank Ajinkya Kadu for sharing his expertise in linear algebra and multivariate analysis. AP would like to acknowledge the Netherlands Organization for Scientific Research (NWO; Grant number 016.Veni.178.027) for financial support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2021.669400/ full#supplementary-material

- 12. Raj T, Rothamel K, Mostafavi S, Ye C, Lee MN, Replogle JM, et al. Polarization of the Effects of Autoimmune and Neurodegenerative Risk Alleles in Leukocytes. *Sci (80-)* (2014) 344:519–23. doi: 10.1126/science.1249547
- Quach H, Rotival M, Pothlichet J, Loh Y-HE, Dannemann M, Zidane N, et al. Genetic Adaptation and Neandertal Admixture Shaped the Immune System of Human Populations. *Cell* (2016) 167:643–56.e17. doi: 10.1016/j.cell.2016.09.024
- Lee MN, Ye C, Villani A-C, Raj T, Li W, Eisenhaure TM, et al. Common Genetic Variants Modulate Pathogen-Sensing Responses in Human Dendritic Cells. Sci (80-) (2014) 343:1246980–1246980. doi: 10.1126/science.1246980
- Fairfax BP, Humburg P, Makino S, Naranbhai V, Wong D, Lau E, et al. Innate Immune Activity Conditions the Effect of Regulatory Variants Upon Monocyte Gene Expression. *Sci (80-)* (2014) 343:1246949–1246949. doi: 10.1126/science.1246949
- Chen L, Ge B, Casale FP, Vasquez L, Kwan T, Garrido-Martín D, et al. Genetic Drivers of Epigenetic and Transcriptional Variation in Human Immune Cells. *Cell* (2016) 167:1398–414.e24. doi: 10.1016/j.cell.2016.10.026
- Schmiedel BJ, Singh D, Madrigal A, Valdovino-Gonzalez AG, White BM, Zapardiel-Gonzalo J, et al. Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. *Cell* (2018) 175:1701–15.e16. doi: 10.1016/ J.CELL.2018.10.022
- Yap H-Y, Tee S, Wong M, Chow S-K, Peh S-C, Teow S-Y. Pathogenic Role of Immune Cells in Rheumatoid Arthritis: Implications in Clinical Treatment and Biomarker Development. *Cells* (2018) 7:161. doi: 10.3390/cells7100161
- Lim J, Kim K. Genetic Variants Differentially Associated With Rheumatoid Arthritis and Systemic Lupus Erythematosus Reveal the Disease-Specific Biology. Sci Rep (2019) 9:1–7. doi: 10.1038/s41598-019-39132-2
- Verstockt B, Smith KG, Lee JC. Genome-Wide Association Studies in Crohn's Disease: Past, Present and Future: Past. *Clin Transl Immunol* (2018) 7:e1001. doi: 10.1002/cti2.1001
- Bentham J, Morris DL, Cunninghame Graham DS, Pinder CL, Tombleson P, Behrens TW, et al. Genetic Association Analyses Implicate Aberrant Regulation of Innate and Adaptive Immunity Genes in the Pathogenesis of Systemic Lupus Erythematosus. *Nat Genet* (2015) 47:1457–64. doi: 10.1038/ ng.3434
- Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, et al. Genetics of Rheumatoid Arthritis Contributes to Biology and Drug Discovery. *Nature* (2014) 506:376–81. doi: 10.1038/nature12873

- Piñero J, Bravo À, Queralt-Rosinach N, Gutiérrez-Sacristán A, Deu-Pons J, Centeno E, et al. DisGeNET: A Comprehensive Platform Integrating Information on Human Disease-Associated Genes and Variants. *Nucleic Acids Res* (2017) 45:D833–9. doi: 10.1093/nar/gkw943
- 24. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast Universal RNA-seq Aligner. *Bioinformatics* (2013) 29:15–21. doi: 10.1093/bioinformatics/bts635
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map Format and Samtools. *Bioinformatics* (2009) 25:2078–9. doi: 10.1093/bioinformatics/btp352
- Anders S, Pyl PT, Huber W. Htseq–a Python Framework to Work With High-Throughput Sequencing Data. *Bioinformatics* (2015) 31:166–9. doi: 10.1093/ bioinformatics/btu638
- Risso D, Ngai J, Speed TP, Dudoit S. Normalization of RNA-seq Data Using Factor Analysis of Control Genes or Samples. *Nat Biotechnol* (2014) 32:896– 902. doi: 10.1038/nbt.2931
- Bodenreider O. The Unified Medical Language System (Umls): Integrating Biomedical Terminology. *Nucleic Acids Res* (2004) 32:D267. doi: 10.1093/nar/ gkh061
- Brunet J-P, Tamayo P, Golub TR, Mesirov JP. Metagenes and Molecular Pattern Discovery Using Matrix Factorization. *Proc Natl Acad Sci U S A* (2004) 101:4164–9. doi: 10.1073/pnas.0308531101
- Gaujoux R, Seoighe C. A Flexible R Package for Nonnegative Matrix Factorization. BMC Bioinf (2010) 11:367. doi: 10.1186/1471-2105-11-367
- Cotto KC, Wagner AH, Feng Y-Y, Kiwala S, Coffman AC, Spies G, et al. Dgidb 3.0: A Redesign and Expansion of the Drug–Gene Interaction Database. *Nucleic Acids Res* (2018) 46:D1068–73. doi: 10.1093/nar/gkx1143
- Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, et al. A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* (2017) 171:1437–52.e17. doi: 10.1016/j.cell.2017.10.049
- Xie Z, Hu S, Blackshaw S, Zhu H, Qian J. Hpdi: A Database of Experimental Human protein-DNA Interactions. *Bioinformatics* (2010) 26:287–9. doi: 10.1093/bioinformatics/btp631
- 34. Armstrong RD, Panayi GS, Welsh KI. Histocompatibility Antigens in Psoriasis, Psoriatic Arthropathy, and Ankylosing Spondylitis. Ann Rheum Dis (1983) 42:142–6. doi: 10.1136/ard.42.2.142
- McGonagle D, Gibbon W, Emery P. Classification of Inflammatory Arthritis by Enthesitis. *Lancet* (1998) 352:1137–40. doi: 10.1016/S0140-6736(97)12004-9
- Vecellio M, Cohen CJ, Roberts AR, Wordsworth PB, Kenna TJ. RUNX3 and T-bet in Immunopathogenesis of Ankylosing Spondylitis - Novel Targets for Therapy? *Front Immunol* (2019) 10:3132. doi: 10.3389/fimmu.2018.03132
- Rezaiemanesh A, Abdolmaleki M, Abdolmohammadi K, Aghaei H, Pakdel FD, Fatahi Y, et al. Immune Cells Involved in the Pathogenesis of Ankylosing Spondylitis. *BioMed Pharmacother* (2018) 100:198–204. doi: 10.1016/ j.biopha.2018.01.108
- Veale DJ, Ritchlin C, FitzGerald O. Immunopathology of Psoriasis and Psoriatic Arthritis. In: Annals of the Rheumatic Diseases. Annals of the Rheumatic Diseases: BMJ Publishing Group Ltd (2005). p. ii26–9. doi: 10.1136/ard.2004.031740
- Murdaca G, Gulli R, Spanò F, Lantieri F, Burlando M, Parodi A, et al. Tnf-α Gene Polymorphisms: Association With Disease Susceptibility and Response to Anti-TNF-α Treatment in Psoriatic Arthritis. *J Invest Dermatol* (2014) 134: P2503–9. doi: 10.1038/jid.2014.123
- Murdaca G, Colombo BM, Cagnati P, Gulli R, Span F, Puppo F. Update Upon Efficacy and Safety of TNF-α Inhibitors. *Expert Opin Drug Saf* (2012) 11:1–5. doi: 10.1517/14740338.2012.630388
- Fröhling M, Vogl T, Loser K, Paruzel P, Blackshear P, Stumpo D, et al. A1.30 A Key Role of S100A9 in the Pathogenesis of Psoriatic Arthritis in TTP/S100 Deficient Mice. Ann Rheum Dis (2016) 75:A13.1–A13. doi: 10.1136/ annrheumdis-2016-209124.30
- Hu XX, Wu YJ, Zhang J, Wei W. T-Cells Interact With B Cells, Dendritic Cells, and Fibroblast-Like Synoviocytes as Hub-Like Key Cells in Rheumatoid Arthritis. *Int Immunopharmacol* (2019) 70:428–34. doi: 10.1016/j.intimp.2019.03.008
- Cope AP, Schulze-Koops H, Aringer M. The Central Role of T Cells in Rheumatoid Arthritis. *Clin Exp Rheumatol* (2007) 25:S4–11.
- Cascão R, Rosário HS, Souto-Carneiro MM, Fonseca JE. Neutrophils in Rheumatoid Arthritis: More Than Simple Final Effectors. *Autoimmun Rev* (2010) 9:531–5. doi: 10.1016/j.autrev.2009.12.013

- Udalova IA, Mantovani A, Feldmann M. Macrophage Heterogeneity in the Context of Rheumatoid Arthritis. Nat Rev Rheumatol (2016) 12:472–85. doi: 10.1038/nrrheum.2016.91
- Louis C, Souza-Fonseca-Guimaraes F, Yang Y, D'Silva D, Kratina T, Dagley L, et al. NK Cell–Derived GM-CSF Potentiates Inflammatory Arthritis and is Negatively Regulated by CIS. J Exp Med (2020) 217:e20191421. doi: 10.1084/ jem.20191421
- Assassi S, Mayes MD, Arnett FC, Gourh P, Agarwal SK, McNearney TA, et al. Systemic Sclerosis and Lupus: Points in an Interferon-Mediated Continuum. *Arthritis Rheum* (2010) 62:589–98. doi: 10.1002/art.27224
- Matta B, Barnes BJ. Coordination Between Innate Immune Cells, Type I Ifns and IRF5 Drives SLE Pathogenesis. *Cytokine* (2020) 132:154731. doi: 10.1016/ j.cyto.2019.05.018
- Vaughn SE, Kottyan LC, Munroe ME, Harley JB. Genetic Susceptibility to Lupus: The Biological Basis of Genetic Risk Found in B Cell Signaling Pathways. J Leukoc Biol (2012) 92:577–91. doi: 10.1189/jlb.0212095
- Van Caam A, Vonk M, Van Den Hoogen F, Van Lent P, Van Der Kraan P. Unraveling SSc Pathophysiology; The Myofibroblast. *Front Immunol* (2018) 9:2452. doi: 10.3389/fimmu.2018.02452
- Denny MF, Yalavarthi S, Zhao W, Thacker SG, Anderson M, Sandy AR, et al. A Distinct Subset of Proinflammatory Neutrophils Isolated From Patients With Systemic Lupus Erythematosus Induces Vascular Damage and Synthesizes Type I Ifns. J Immunol (2010) 184:3284–97. doi: 10.4049/jimmunol.0902199
- Garcia-Romo GS, Caielli S, Vega B, Connolly J, Allantaz F, Xu Z, et al. Netting Neutrophils are Major Inducers of Type I Ifn Production in Pediatric Systemic Lupus Erythematosus. *Sci Transl Med* (2011) 3:73ra20–0. doi: 10.1126/ SCITRANSLMED.3001201
- Darrah E, Andrade F. Nets: The Missing Link Between Cell Death and Systemic Autoimmune Diseases? Front Immunol (2012) 3:428. doi: 10.3389/ fimmu.2012.00428
- 54. Lindau D, Mussard J, Rabsteyn A, Ribon M, Kötter I, Igney A, et al. TLR9 Independent Interferon α Production by Neutrophils on NETosis in Response to Circulating Chromatin, a Key Lupus Autoantigen. *Ann Rheum Dis* (2014) 73:2199–207. doi: 10.1136/annrheumdis-2012-203041
- Sartor RB. Mechanisms of Disease: Pathogenesis of Crohn's Disease and Ulcerative Colitis. Nat Clin Pract Gastroenterol Hepatol (2006) 3:390–407. doi: 10.1038/ncpgasthep0528
- Giuffrida P, Di Sabatino A. Targeting T Cells in Inflammatory Bowel Disease. *Pharmacol Res* (2020) 159:105040. doi: 10.1016/j.phrs.2020.105040
- Geremia A, Arancibia-Cárcamo CV. Innate Lymphoid Cells in Intestinal Inflammation. Front Immunol (2017) 8:1296. doi: 10.3389/fimmu.2017.01296
- Poggi A, Benelli R, Venè R, Costa D, Ferrari N, Tosetti F, et al. Human Gut-Associated Natural Killer Cells in Health and Disease. *Front Immunol* (2019) 10:961. doi: 10.3389/fimmu.2019.00961
- Park K. A Review of Computational Drug Repurposing. Transl Clin Pharmacol (2019) 27:59. doi: 10.12793/tcp.2019.27.2.59
- Ito H, Takazoe M, Fukuda Y, Hibi T, Kusugami K, Andoh A, et al. A Pilot Randomized Trial of a Human Anti-Interleukin-6 Receptor Monoclonal Antibody in Active Crohn's Disease. *Gastroenterology* (2004) 126:989–96. doi: 10.1053/j.gastro.2004.01.012
- Tsai Y-C, Tsai T-F. Anti-Interleukin and Interleukin Therapies for Psoriasis: Current Evidence and Clinical Usefulness. *Ther Adv Musculoskelet Dis* (2017) 9:277–94. doi: 10.1177/1759720X17735756
- 62. Alten R, Gomez-Reino J, Durez P, Beaulieu A, Sebba A, Krammer G, et al. Efficacy and Safety of the Human anti-IL-1beta Monoclonal Antibody Canakinumab in Rheumatoid Arthritis: Results of a 12-Week, Phase II, Dose-Finding Study. *BMC Musculoskelet Disord* (2011) 12:153. doi: 10.1186/1471-2474-12-153
- Haibel H, Rudwaleit M, Listing J, Sieper J. Open Label Trial of Anakinra in Active Ankylosing Spondylitis Over 24 Weeks. Ann Rheum Dis (2005) 64:296–8. doi: 10.1136/ard.2004.023176
- Hügle B, Speth F, Haas J-P. Inflammatory Bowel Disease Following antiinterleukin-1-treatment in Systemic Juvenile Idiopathic Arthritis. *Pediatr Rheumatol* (2017) 15:16. doi: 10.1186/s12969-017-0147-3
- 65. Merashli M, De Marco G, Podgorski M, McGonagle D, Marzo-Ortega H. Evidence of Response to IL-6 Inhibition in Some Cases of Refractory Spondyloarthritis Associated Peripheral Synovitis. Ann Rheum Dis (2016) 75:1418–20. doi: 10.1136/annrheumdis-2016-209275

- 66. Danese S, Vermeire S, Hellstern P, Panaccione R, Rogler G, Fraser G, et al. Randomised Trial and Open-Label Extension Study of an Anti-Interleukin-6 Antibody in Crohn's Disease (ANDANTE I and II). *Gut* (2019) 68:40–8. doi: 10.1136/gutjnl-2017-314562
- 67. Nishimoto N, Hashimoto J, Miyasaka N, Yamamoto K, Kawai S, Takeuchi T, et al. Study of Active Controlled Monotherapy Used for Rheumatoid Arthritis, an IL-6 Inhibitor (SAMURAI): Evidence of Clinical and Radiographic Benefit From an X Ray Reader-Blinded Randomised Controlled Trial of Tocilizumab. *Ann Rheum Dis* (2007) 66:1162–7. doi: 10.1136/ard.2006.068064
- Blauvelt A. Il-6 Differs From TNF-α: Unpredicted Clinical Effects Caused by IL-6 Blockade in Psoriasis. J Invest Dermatol (2017) 137:541–2. doi: 10.1016/ j.jid.2016.11.022
- Atreya R, Billmeier U, Rath T, Mudter J, Vieth M, Neumann H, et al. First Case Report of Exacerbated Ulcerative Colitis After anti-interleukin-6R Salvage Therapy. World J Gastroenterol (2015) 21:12963–9. doi: 10.3748/ wjg.v21.i45.12963
- Van Der Heijde D, Deodhar A, Wei JC, Drescher E, Fleishaker D, Hendrikx T, et al. Tofacitinib in Patients With Ankylosing Spondylitis: A Phase II, 16-Week, Randomised, Placebo-Controlled, Dose-Ranging Study. *Ann Rheum Dis* (2017) 76:1340–7. doi: 10.1136/annrheumdis-2016-210322
- Rogler G. JAK Efficacy in Crohn's Disease. J Crohn's Colitis (2019) 14:S746– 54. doi: 10.1093/ecco-jcc/jjz186
- Dhillon S. Tofacitinib: A Review in Rheumatoid Arthritis. Drugs (2017) 77:1987–2001. doi: 10.1007/s40265-017-0835-9
- Tian F, Chen Z, Xu T. Efficacy and Safety of Tofacitinib for the Treatment of Chronic Plaque Psoriasis: A Systematic Review and Meta-Analysis. J Int Med Res (2019) 47:2342–50. doi: 10.1177/0300060519847414
- 74. Hübel K, Fresen MM, Salwender H, Basara N, Beier R, Theurich S, et al. Plerixafor With and Without Chemotherapy in Poor Mobilizers: Results From

the German Compassionate Use Program. Bone Marrow Transplant (2011) 46:1045-52. doi: 10.1038/bmt.2010.249

- Ley K, Rivera-Nieves J, Sandborn WJ, Shattil S. Integrin-Based Therapeutics: Biological Basis, Clinical Use and New Drugs. *Nat Rev Drug Discovery* (2016) 15:173–83. doi: 10.1038/nrd.2015.10
- Ellinghaus D, Jostins L, Spain SL, Cortes A, Bethune J, Han B, et al. Analysis of Five Chronic Inflammatory Diseases Identifies 27 New Associations and Highlights Disease-Specific Patterns At Shared Loci. *Nat Genet* (2016) 48:510–8. doi: 10.1038/ng.3528
- Bernstein CN, Sargent M, Rector E. Alteration in Expression of Beta 2 Integrins on Lamina Propria Lymphocytes in Ulcerative Colitis and Crohn's Disease. *Clin Immunol* (2002) 104:67–72. doi: 10.1006/clim.2002.5223
- Gianchecchi E, Delfino DV, Fierabracci A. NK Cells in Autoimmune Diseases: Linking Innate and Adaptive Immune Responses. *Autoimmun Rev* (2018) 17:142–54. doi: 10.1016/j.autrev.2017.11.018
- Perez VL, Pflugfelder SC, Zhang S, Shojaei A, Haque R. Lifitegrast, a Novel Integrin Antagonist for Treatment of Dry Eye Disease. *Ocul Surf* (2016) 14:207–15. doi: 10.1016/j.jtos.2016.01.001

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Devaprasad, Radstake and Pandit. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.