Functional genomics implicates natural killer cells in the pathogenesis of ankylosing spondylitis

Marcos Chiñas[1,3], Daniela Fernandez-Salinas[1,3,4], Vitor R. C. Aguiar[1,2,3], Victor E. Nieto-Caballero[1,3,4], Micah Lefton[5], Peter A. Nigrovic[1,2,5], Joerg Ermann*[2,5], Maria Gutierrez-Arcelus*[1,2,3]

1. Division of Immunology, Boston Children's Hospital, Boston, MA, USA.
2. Harvard Medical School, Boston, MA, USA.
3. Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA.
4. Licenciatura en Ciencias Genomicas, Centro de Ciencias Genomicas, Universidad Nacional Autónoma de México (UNAM), Morelos 62210, Mexico.
5. Division of Rheumatology, Inflammation and Immunity, Brigham and Women's Hospital, Boston, MA 02115, USA.

*Co-corresponding authors:

Maria Gutierrez-Arcelus
Boston Children's Hospital
Division of Immunology
Karp Family Research Laboratories, Room 102016.1
1 Blackfan Circle
Boston, MA, 02115
mgutierr@broadinstitute.org
857-559-3229

Joerg Ermann
Brigham and Women's Hospital
Division of Rheumatology, Inflammation and Immunity
Hale Building for Transformative Medicine, Room 6002P
60 Fenwood Road
Boston, MA, 02115
jermann@bwh.harvard.edu
617-525-1227

Abstract

Objective: Multiple lines of evidence indicate that ankylosing spondylitis (AS) is a lymphocyte-driven disease. However, which lymphocyte populations are critical in AS pathogenesis is not known. In this study, we aimed to identify the key cell types mediating the genetic risk in AS using an unbiased functional genomics approach.

Methods: We integrated genome-wide association study (GWAS) data with epigenomic and transcriptomic datasets of human immune cells. To quantify enrichment of cell type-specific open chromatin or gene expression in AS risk loci, we used three published methods that have successfully identified relevant cell types in other diseases. We performed co-localization analyses between GWAS risk loci and genetic variants associated with gene expression (eQTL) to find putative target genes.

Results: Natural killer (NK) cell-specific open chromatin regions are significantly enriched in heritability for AS, compared to other immune cell types such as T cells, B cells, and monocytes. This finding was consistent between two AS GWAS. Using RNA-seq data, we validated that genes in AS risk loci are enriched in NK cell-specific gene expression. Using the human Space-Time Gut Cell Atlas, we also found significant upregulation of AS-associated genes predominantly in NK cells. Co-localization analysis revealed four AS risk loci affecting regulation of candidate target genes in NK cells: two known loci, *ERAP1 and TNFRSF1A*, and two under-studied loci, *ENTR1* (aka *SDCCAG3*) and *B3GNT2*.

Conclusion: Our findings suggest that NK cells may play a crucial role in AS development and highlight four putative target genes for functional follow-up in NK cells.

Introduction

Axial spondyloarthritis (axSpA) is a chronic inflammatory rheumatic disease characterized by inflammation of the spine and sacroiliac joints, with a proportion of patients also presenting with arthritis in peripheral joints, uveitis, psoriasis or inflammatory bowel disease (1). Historically, most genetic and pathogenetic studies in axSpA have been carried out in ankylosing spondylitis (AS), a severe and well-characterized subtype of axSpA. The heritability of AS is high, with estimates ranging between 40-90% (2). HLA-B27 is the major risk allele for AS (OR = 21.4) (3). Additionally, genome wide-association studies (GWAS) have revealed >100 non-MHC risk loci for AS, most of them implicating non-coding variants (4–8).

Many immune cell-types have been associated with axSpA (9,10). However, which ones are "driver" cell types actively contributing to the pathogenesis of the disease, as opposed to "bystanders" that become involved as a consequence of the disease, remains unclear. Studies leveraging genetic risk variants and their overlap with epigenomic and transcriptomic features variably suggested CD8+ T cells, CD4+ T cells, NK (natural killer) cells, monocytes, and gastrointestinal cells as potential mediators of AS genetic risk (10–14). However, these studies did not apply the new functional genomics datasets generated from human cells or the latest methodologies designed to integrate functional genomics with GWAS data. This new generation of methods takes advantage of the full range of SNPs examined in a GWAS (not just those surpassing the genome-wide significance threshold) and robustly control for genomic and linkage disequilibrium biases (15–17).

For several immune-mediated diseases, these integrative functional genomics methods have successfully identified specific cell types as drivers of disease development. For example, for rheumatoid arthritis (RA), multiple studies have found a significant enrichment of genetic risk in open or active chromatin regions (marking regulatory elements) specific for T cells (11,18–20). Both mouse and human studies corroborate the role of T cells as central players in RA pathogenesis (21–23). Similarly, for systemic lupus erythematosus (SLE), studies have identified an enrichment of B cell-specific putative regulatory elements and gene expression in SLE risk loci (19,20,24,25) consistent with the well-established role of B cells in SLE pathogenesis (26,27). Hence, there is precedence that the integration of GWAS with functional genomics datasets can identify cellular drivers in inflammatory diseases with complex pathogenesis.

Here we sought to investigate which immune cell populations could be drivers of AS development. We integrated GWAS summary statistics from two different AS cohorts with epigenomic and transcriptomic datasets of human leukocytes from peripheral blood and tissue using established methods that control for biases in genomic enrichment analyses. Our results bring forward NK cells as potential key drivers in the pathogenesis of AS.

Results

In order to assess which immune cell types might be mediating the genetic susceptibility to AS, we first utilized a dataset of open chromatin profiles of immune cell subsets from peripheral blood of four healthy subjects (19) (**Fig. 1A**). Sorted cell subsets were analyzed using ATAC-seq with or without prior *in vitro* stimulation. For our study, we grouped the cells analyzed by Calderon et al. into 7 main immune cell types: T cells, B cells, NK cells, plasma cells, dendritic cells (DCs), plasmacytoid DCs, and monocytes. We identified cell type-specific open chromatin regions and assessed whether these were significantly enriched in AS genetic risk. We used the LDSC-SEG method (15) to quantify enrichment of partitioned heritability in each of these cell type-specific annotations (conceptual scheme in **Fig. 1B,** data in **Fig. 1C**) compared to baseline and control annotations, while taking into account the effects of linkage disequilibrium. We excluded the MHC region from our analyses given the unusually high linkage disequilibrium in this region and the fact that genetic associations with this locus are mostly driven by coding variants of the HLA-B gene. Using the Immunochip association study summary statistics from the International Genetics of Ankylosing Spondylitis Consortium (IGAS) (7), we found that NK cell-specific open chromatin regions were significantly enriched in genetic risk for AS ($P = 0.026$), while this was not the case for the other six immune cell types (**Fig. 1D**).

We validated this finding in a GWAS with genome-wide genotyping using the summary statistics for AS from the UK Biobank. With this GWAS, we confirmed that open chromatin regions specific for NK cells were significantly enriched in AS heritability ($P = 0.034$, **Fig. 1E**). To evaluate the reliability of our results, we included four control traits that have been extensively examined in similar studies integrating GWAS with functional genomics (15,18–20). As expected, RA presented the highest enrichment for T cell-specific open chromatin regions ($P = 0.0018$), Alzheimer's for myeloid DC ($P = 0.00018$), and SLE for B cells ($P = 0.0015$). We selected body height as a negative control trait anticipating no significant enrichment for immune cells, a

prediction that was confirmed by our data (all P > 0.1, **Supplementary Fig. 1**). Collectively, these epigenomic analyses suggest that AS risk alleles are preferentially located in regions that may influence gene regulation in NK cells.

To corroborate these findings using an alternative experimental approach, we used our previously published RNA-seq dataset of sorted peripheral CD4+ T, CD8+ T, MAIT, invariant NKT (iNKT), γδ T cells expressing Vδ1 TCR chain (Vd1), γδ T cells expressing Vδ2 TCR chain (Vd2), and NK cells (each in duplicate from 6 healthy donors, **Fig. 2A**) (28). We applied the SNPsea method, which quantifies enrichment of cell type-specific gene expression in risk loci for a given trait (conceptual scheme in **Fig. 2B**) by employing a non-parametric statistical method to calculate empirical P-values through comparison with sets of null SNPs (29). We used the AS risk SNPs reported by Brown and Wordsworth in 2017 which were curated from multiple AS genetic studies (30). SNPsea analysis revealed a significant enrichment of NK cell-specific gene expression in AS risk loci (P = 0.01), which was not observed in the other lymphocyte subsets included in the dataset (**Fig. 2C**).

We then performed a differential expression analysis comparing NK cells with the six T cell subsets (Supplementary Table 1). Genes in AS risk loci with significant upregulation in NK cells are presented in **Fig 2D**. Two of these genes, *RUNX3* and *TBX21* (Tbet), encode transcription factors with important roles in lymphocytes. *TNFRSF1A* encoding TNF Receptor I has a well-established association with AS that has been validated by multiple studies (31–33). *FCGR2A* codes for the low-affinity Fcγ receptor IIA, an activating receptor involved in orchestrating immune response. Less studied genes included *NPEPPS,* which encodes a puromycin-sensitive aminopeptidase, and *LNPEP*, which encodes a zinc-dependent aminopeptidase (34). Both genes are paralogs of *ERAP1* and belong to the MHC Class I antigen processing and presentation pathway, along with other known AS risk genes (35–37). Collectively, the results of our second integrative analysis indicate that several genes within AS risk loci are highly expressed in NK cells relative to T cells, providing additional support for the emerging hypothesis that AS risk alleles exert their effects, at least in part, via NK cells.

The transcriptomic phenotype of immune cells commonly differs between blood and tissue (38–41). Hence, in addition to analyzing peripheral blood as in the previous analyses, we sought to evaluate disease-relevant cell subsets from a tissue relevant for AS. We used the human Space-

Time Gut Cell Atlas (42) which includes scRNA-seq data for samples from various locations of fetal (N = 16), pediatric (N = 8) and adult (N = 13, including 6 healthy and 7 Crohn's disease patients) intestine (**Fig. 3A)**. We applied the scDRS method (16) which identifies cells that over-express a significant proportion of genes implicated by GWAS, weighted on their strength of association with disease, compared to null sets of control genes in the dataset (conceptual scheme in **Fig. 3B**). The Space-Time Gut Cell Atlas investigators identified the following broad cell types: mesenchymal, epithelial, endothelial, neuronal, myeloid, red blood cells, B cells, plasma cells, T cells, NK cells and other innate lymphoid cells (ILCs)(**Fig. 3C)**. scDRS identified 1,852 cells with significantly enriched expression of AS GWAS genes (20% FDR, **Fig. 3D**). Of these, 765 were T cells, 264 myeloid cells, 320 NK cells and 319 other ILCs. Normalized for cell type abundance in the dataset, NK cells showed the highest enrichment (39-fold), followed by other ILCs (34-fold), T cells (5-fold), and myeloid cells (5-fold, **Fig. 3E**). In contrast, non-immune cell types exhibited a depletion of disease relevant cells relative to their abundance in the entire dataset (**Supplementary Fig. 2**). We then used the fine-grained annotations of the Space-Time Gut Cell Atlas to identify the particular cell subsets that had significant expression enrichment of AS-associated genes. This revealed NK cells as the most abundant (N = 320), followed by Lti-like NCR+ ILC3 cells (N = 147), activated CD8+ T cells (N = 132), macrophages (N = 130), Lti-like NCR- ILC3 cells (N = 112), γδ T cells (N = 94), and other T cells, ILCs and myeloid subsets (**Fig. 3F**). Genes in AS risk loci with high expression in gut NK cells include *GNLY*, *CCL4* and *CCL3* (**Fig. 3G**). Using the control traits specified earlier, we confirmed T cells as the main disease-relevant cell type for RA and monocytes for Alzheimer's disease (**Supplementary Fig. 2**). No significant disease-relevant cells were identified for height (as expected) and for SLE, which could mean that B cells in the gut are in a state not pertinent to SLE or that the dataset lacked sufficient power to detect an association for this disease (**Supplementary Fig. 2**). In sum, our analyses indicate that tissue-resident NK cells exhibit significant expression of AS-associated genes.

Lastly, we sought to find putative target genes for AS risk variants in NK cells. To this end we performed co-localization analyses between AS GWAS risk loci and genetic variants associated with gene expression (expression quantitative trait loci, eQTLs) using coloc (43). We leveraged eQTL summary statistics from the eQTL Catalogue (44) drawing upon data from a study on the transcriptomic profiling of peripheral NK cells from 91 genotyped individuals (45) as well as a microarray QTL study that profiled NK cells from 245 genotyped individuals (46). We found four AS risk loci with genome-wide significance ($P < 5 \times 10^{-8}$) and a high posterior probability (>0.8) of

sharing a causal variant with an NK cell eQTL (PP4, **Table 1, Fig. 4**). An additional 10 loci with suggestive AS association P-values ($3.56 \times 10^{-5} < P < 5.40 \times 10^{-8}$) showed evidence of co-localization with NK cell eQTLs for 18 genes (PP4 > 0.75, **Table 1**). Within the genome-wide significant loci we identified the established target genes *ERAP1* and *TNFRSF1A*, as well as the putative target genes *ENTR1* (a.k.a. *SDCCAG3*) and *B3GNT2*, which have been studied less.

Discussion

In this study we integrated epigenomic and transcriptomic datasets with AS genetic risk data to find candidate cellular drivers of AS pathogenesis. Our unbiased approach, applying three different methods to datasets from both peripheral blood and tissue, consistently identified NK cells as the dominant disease-relevant cell type. Specifically, we found that NK-specific open chromatin regions and NK-specific gene expression were significantly enriched for non-MHC AS genetic risk. This suggests that a significant portion of AS risk variants affect gene regulation in NK cells, pointing to NK cells as potential key mediators of AS pathogenesis.

NK cells have the ability to directly destroy target cells through cell lysis, and in addition play a significant role in shaping immune responses by releasing cytokines (47). Previous studies support a role for NK cells in AS. AS patients with chronic subclinical intestinal inflammation were found to have an increased abundance of NKp44+ NK cells in their gut, and these cells were the major producers of IL-22 in the lamina propria, suggesting a possible role in tissue protection (48). One could speculate that dysfunctional NK cells “drive” AS development by contributing to intestinal inflammation, in line with the gut-joint axis hypothesis (49). Alternatively, NK cells may play a critical role through activities in spinal tissues. Cuthbert *et al.* studied entheseal immunology using discarded surgical specimens from patients with back pain (not axSpA) undergoing laminectomy and reported that NK cells are present in both entheseal soft tissue and peri-entheseal bone (50). We are not aware of any data assessing the presence of NK cells at spinal enthesis in AS patients or in the subchondral bone marrow in patients with sacroiliitis.

HLA-B27 can bind to the killer cell immunoglobulin-like receptor (KIR) KIR3DL1 and affect the function of NK cells, including their ability to lyse cells (51–53). HLA-B27 homodimers can also bind KIR3DL2 (54). Chan and colleagues showed an expansion of KIR3DL2+ NK and CD4+ T cells in AS patients (Chan Arthritis Rheum 2005, PMID 16255049). Subsequent studies by the same group focused on CD4+ T cells demonstrating that KIR3DL2+ CD4+ T cells were major IL-

17A producers (55). However, an expansion of KIR3DL2+ NK and T cells has not been observed in other axSpA cohorts (56,57). Multiple risk loci for AS include genes relevant for NK cell function, including *KIR2DL1*, *KIR3DL1*, *KIR2DS5*, *KIR3DS1* and *KIR2DL5* (58–62). In another study, investigators co-cultured ERAP1-inhibited M1 macrophages with NK cells from AS patients, and found that patients with ERAP1 protective alleles led to decreased CD69 and CD107a on NK cells and a lower number of IFN-γ+ NK cells compared to patients carrying non-protective alleles (63).

Our findings do not rule out the involvement of other cell types in AS pathogenesis. Indeed, in the human Space-Time Gut Cell Atlas, we identified significant expression of AS-associated genes in T cell subsets and ILC subsets (Fig. 3D-F), which share transcriptional programs with NK cells (28,64–66). Indeed, it is likely that genetic risk to AS is mediated through multiple cell types, as is the case for other complex diseases such as multiple sclerosis, for which studies have found risk enrichment in open/active chromatin regions specific to both T cells and B cells (20). We and others have shown that eQTLs often exhibit impact in multiple cell types (67,68). Hence, determining the specific cell type through which a disease risk variant is exerting its pathogenic effects can be challenging.

Our co-localization analyses using two eQTL NK cell datasets identified four putative target genes for AS risk variants: *ERAP1*, *TNFRSF1A*, *ENTR1* (a.k.a. *SDCCAG3*) and *B3GNT2*. The importance of *ERAP1* in AS risk is well established, and polymorphisms affecting its expression have been reported for multiple cell types, including macrophages, monocytes, T cells, and induced pluripotent stem cells, fibroblasts, and immortalized B cells (69–71). Similarly, multiple studies have found significant associations between non-coding polymorphisms at or near *TNFRSF1A* and AS, including in European and East Asian populations (31,32,72,73). While there are multiple genes in this genomic locus, including *PLEKHG6*, *SCNN1A*, and *LTBR*, our co-localization results suggest that *TNFRSF1A*, which encodes TNF receptor I, is the target gene of the causal variant in this locus, and its dysregulation can happen in NK cells. This is consistent with the therapeutic efficacy of TNF inhibitors in AS and the known function of TNF as a booster of the cytolytic capacity of NK cells (74). Interestingly, TNFRSF1A has been functionally linked to ENTR1, a less extensively studied putative target gene identified in this study. *ENTR1*, which encodes an endosome-associated trafficking regulator, is needed for TNF receptor expression on the cell surface (75). Lastly, *B3GNT2* encodes an acetylglucosaminyltransferase enzyme that is a type II transmembrane protein. A recent study in a Taiwanese cohort demonstrated that a non-coding genetic variant near *B3GNT2* is associated with AS susceptibility, and that *B3GNT2* blood

mRNA levels were negatively correlated with C-reactive protein (CRP), erythrocyte sedimentation rate, syndesmophyte formation, and Bath ankylosing spondylitis functional index (BASFI) (76).

While the 4 putative target genes identified here make sense in the context of AS and potential impact on NK cell function, they are not as numerous as we would have expected given the 32 genome-wide significant risk loci included in the co-localization analyses. However, similar challenges have been encountered with non-coding risk variants in other complex diseases, where only 20-47% of risk variants co-localized with eQTLs (77–79). Our research, along with that of others, suggests that many regulatory effects might remain undetected due to their presence in cell states of activation or differentiation that have not been thoroughly explored (80–83). Moreover, the sample size of typical eQTL studies is likely insufficient to find the regulatory effects of most risk variants identified by GWAS (84). Hence, we believe that better-powered eQTL studies ascertaining multiple activation states in NK cells are needed to find additional target genes for AS risk variants.

One limitation of our study, due to a lack of published data, is the incomplete assessment of the spectrum of potentially relevant immune cell subsets and states, particularly those present in inflamed sacroiliac joints and spine. Consequently, if the real driver for AS pathogenesis is a cell subset or state that was not present in the analyzed datasets, but has transcriptomic and epigenomic similarities to NK cells, then our results may suffer from “guilt-by-association” bias. To our knowledge, current transcriptomic datasets profiling multiple immune cell types from AS patients are limited to peripheral blood (14,85–90). When we applied scDRS to a recently published single-cell RNA-seq dataset of 98,884 PBMCs cells from 10 AS patients and 29 healthy controls (88), we found no significant cells for the disease-relevant gene expression score (data not shown), possibly due to lack of power in the study for this type of analysis.

Our study encompassed a broad spectrum of immune cell states within the gastrointestinal tract and peripheral blood of healthy subjects and consistently pointed to NK cells. Since GWAS pinpoint genetic regions implicated in the onset of disease, including early stages when future patients are still asymptomatic, the study of samples from healthy subjects is relevant, despite the possibility that not all cell states are represented. Future investigations, particularly larger-scale studies of samples from blood and inflamed tissue from AS patients including untreated patients in the early phases of the disease, will be key to establish whether NK cells are indeed drivers of AS pathogenesis.

Materials and Methods

Genome-wide association studies

We used the GWAS ImmunoChip summary statistics from the International Genetics of Ankylosing Spondylitis Consortium (IGAS). The IGAS study, led by (7), performed high-density genotyping of 9,069 AS cases and 13,578 healthy controls. In addition, we used the GWAS summary statistics from the UK Biobank, which involved a case-control design with 1,185 AS cases and 419,276 controls, providing genome-wide coverage for AS susceptibility loci (91).

We lifted the genomic positions of the genetic variants to genome build hg19 or hg38 according to the version compatible with subsequent analyses. Given the complexity and strong genetic association signals within the Major Histocompatibility Complex (MHC) region, we excluded variants located on chr6:25,000,000–34,000,000.

We additionally used GWAS summary statistics for rheumatoid arthritis, Alzheimer's disease, and systemic lupus erythematosus as positive control traits for which we know the disease relevant immune cell types, and height as a negative control trait for which we do not expect immune cells to be relevant. The summary statistics for control traits were preprocessed by the Alkes Price laboratory, they included HapMap 3 SNPs and SNPs that are in the 1000 Genomes Project, and they excluded the MHC region (chr6:25Mb-34Mb). These summary statistics are available at https://alkesgroup.broadinstitute.org/.

Epigenomic and transcriptomic datasets

To identify cell type-specific open chromatin regions in different immune cell types, we used the Calderon *et al.* study *(19)*, in which the authors collected blood from 4 healthy subjects, sorted immune cell types, and generated chromatin accessibility profiles using Assay for Transposase-Accessible Chromatin sequencing (ATAC-seq, GSE118189).

To find AS risk enrichment for cell type-specific expression, we incorporated data from the study conducted by Gutierrez-Arcelus *et al.* (28), which involved low-input mRNA-seq data from sorted

NK cells and six T cell subsets isolated from 6 healthy subjects (each with two replicates per cell-type, GSE124731).

We used the Space-Time Gut Cell Atlas to identify cells exhibiting significant upregulation of disease-associated genes. This dataset includes single-cell RNA-seq profiling of 428,000 intestinal cells obtained from fetal (N = 16), pediatric (N = 8), and adult donors (N = 13). The dataset covers 11 different intestinal regions (42), https://www.gutcellatlas.org/.

Differential accessibility analysis

We used the counts of open chromatin consensus peaks called by Calderon et al. First, we transformed counts into Reads Per Kilobase per Million mapped reads (RPKM), then normalized by quantiles using the preprocess Core R package and finally scaled to their log2 (normalized RPKM+1), thus we account for differences in library size across samples and peak length variability. We pooled sorted samples into 7 main immune cell types, aiming for a similar number of samples per cell type to avoid biases in the differential accessibility analyses: T cells (stimulated and unstimulated CD8+ T, unstimulated naïve CD4 T and memory CD4 T), B cells (stimulated and unstimulated bulk B cells, unstimulated memory and naïve B cells), natural killer cells (stimulated and unstimulated mature NKs, unstimulated memory NK and immature NK cells), monocytes (stimulated and unstimulated monocytes), plasma cells (unstimulated plasma cells), dendritic cells (unstimulated myeloid DCs), plasma dendritic cells (unstimulated plasmacytoid DCs). The latter three cell-types had less samples available, however, this did not impede our control trait Alzheimer disease to show significant heritability enrichment for myeloid DC-specific open chromatin regions, as expected (see Methods below and main text).

Next, we employed linear mixed model regression to identify regions that exhibited differential accessibility between each cell type and the rest of the cell types. To account for potential donor-specific effects, we incorporated the donor ID variable as a random effect in our analysis.
For each cell type comparison, we tested peaks that had counts greater than the mean for that cell type in at least half of the samples, this yielded between 400 to 600 thousand tested peaks depending on the cell type. To select the “cell type-specific open chromatin peaks” for each cell type, we sorted open chromatin peaks by their t-statistic and chose the positive top 10%.

Partitioned heritability enrichment analysis with LDSC-SEG

Linkage Disequilibrium score regression applied to specifically expressed genes (LDSC-SEG) v1.0.1 method was applied to determine disease-relevant cell types (15,92) for AS.

Cell type-specific open chromatin peaks were extended by 225bp to each side, to match the genomic coverage recommended by the LDSC-SEG authors. These annotations were then utilized as input for the partitioned heritability enrichment analysis by LDSC-SEG. We used the baseline annotation v1.2 provided by Price Lab for LDSC-SEG, comprising 75 background annotations. Additionally, we used all consensus peaks (N = 829,942) of Calderon et al. as the control annotation. Using other baselines or controls did not affect our results. We utilized SNP weight files derived from the HapMap 3 project (HM3) European population.

Analysis of cell type-specific gene expression enrichment in risk loci using SNPsea

SNPsea analysis aimed to assess the association between risk SNPs and genes expressed specifically for a given cell type (29). We incorporated a curated list of risk SNPs for ankylosing spondylitis (AS), compiled by Brown and Wordsowrth 2017 (30), which includes genetic variants that have been associated with AS susceptibility. This list was derived from multiple AS studies conducted until 2017.

We utilized the expression data obtained from Gutierrez-Arcelus et al. (2019). The gene expression counts in this dataset were normalized to transcripts per million (TPM) and transformed to log2(TPM+1) values. To identify the genes with meaningful expression levels, we included those with log2(TPM+1) > 2 in at least 10 samples. SNPsea was then run for the normalized expression matrix and AS risk SNPs, using recombination intervals from Myers et al. (93), null SNPs from Lango et al. (94), and the following parameters: --score single --slop 10000 --threads 2 --null-snpsets 0 --min-observations 100 --max-iterations 10000000.

Integration of GWAS with single-cell RNA-seq with scDRS

We used the single-cell Disease Relevance Score (scDRS) by combining scRNA-seq and GWAS to identify cells with significant up-regulation of disease-associated genes, which are scored based on their strength of association with disease, and are compared with null sets of genes present in the dataset.

As recommended by scDRS authors, we first created disease-relevant genesets using Multi-marker Analysis of GenoMic Annotation (MAGMA) version 1.10 (95). First, we generated gene annotations with MAGMA setting a window of 10kb using the following parameters: "--annotate window=10,10 --snp-loc ./g1000_eur/g1000_eur.bim --gene-loc ./NCBI37.3/NCBI37.3.gene.loc". Then we ran MAGMA using GWAS summary statistics for traits of interest with the following parameters: --bfile ./magma_v1.10/g1000_eur/g1000_eur --pval GWAS.pval use='SNP,P' ncol='N' --gene-annot ./magma_v1.10/out/step1.genes.annot.

We ran scDRS using the disease-relevant gene sets from MAGMA, the expression data obtained from the Space-Time Gut Cell Atlas (42) and corrected for biases by adding as covariates the number of genes expressed per cell and sample batch. Next, for visualization purposes and downstream analysis, we processed the single-cell dataset using Seurat (96), we performed integration across batches with Harmony (97), and we visualized cells in two dimensions with Uniform Manifold Approximation and Projection (UMAP). We labeled cells plotted in UMAP by the annotations defined by the Space-Time Gut Cell Atlas. Additionally, we colored cells by their scDRS score when cells passed the 0.20 FDR threshold.

eQTL co-localization analysis

To select genomic loci for colocalization analysis, GWAS summary statistics were sorted by P-values, then starting from the variant with the smallest P-value, variants within a 50 Kb window were removed. The process was repeated with the next most significant variant among the remaining variants until no variant with a P-value below 5×10^{-5} was left. We performed colocalization analysis for GWAS studies against the eQTL Catalogue (44). We imported eQTL summary statistics from RNA-seq and microarray from Schmiedel et al. (45) and Gilchrist et al. (2022). We fetched the summary statistics data using the tabix method with the seqminer R package (v8.5). For each region tested, we included all biallelic SNPs that were ascertained in both the GWAS and eQTL study and performed the analysis only for genes within a window of ±500,000 base pairs from the GWAS top variant, and for which there was at least one eQTL passing the 5×10^{-5} P-value threshold. Before merging GWAS and QTL data, the variant coordinates of the GWAS were lifted to the GRCh38 version of the reference genome using liftOver with the UCSC chain file. We used the coloc v5.1.0.1 package (98) in R v4.1.0 to test for colocalization at each gene and dataset. The code is available at <https://github.com/gutierrez-arcelus-lab/>.

Each locus was plotted using plotgardener (99), and we recovered the LD of the top SNP in a given region in the GWAS dataset using the locuscomparer package (100). Then we used plotgardener functions to display the regions near the lead variant and colored the genes tested using the posterior probability that the two traits share a causal variant (PP4).

Data availability

All data and methods are publicly available as specified above.

Acknowledgments

This study was supported by a seed grant from the Spondyloarthritis Research and Treatment Network (SPARTAN) and a microgrant from the Joint Biology Consortium (1P30AR070253-01). PAN was supported by P30AR070253 and R01AR073201. JE was supported by NIH Grant R21 AR076040-01 and an ASPIRE grant from Pfizer. MGA was supported by P30AR070253, the Arthritis National Research Foundation, the Lupus Research Alliance, and the Gilead Sciences Rheumatology Research Scholars Award. We thank Soumya Raychaudhuri and Kamil Slowikowski for guidance on implementing SNPsea, and Steven Gazal for guidance on implementing LDSC-SEG. We thank the Gutierrez-Arcelus and Nigrovic laboratories for feedback on this study.

Disclosures

The authors declare they have no relevant conflicts of interest.

References

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

2. Díaz-Peña R, Castro-Santos P, Durán J, Santiago C, Lucia A. The Genetics of Spondyloarthritis. *J Pers Med* 2020;10. Available at: http://dx.doi.org/10.3390/jpm10040151.

3. Reveille JD, Zhou X, Lee M, Weisman MH, Yi L, Gensler LS, et al. HLA class I and II alleles in susceptibility to ankylosing spondylitis. *Ann Rheum Dis* 2019;78:66–73. Available at: http://dx.doi.org/10.1136/annrheumdis-2018-213779.

4. Wellcome Trust Case Control Consortium, Australo-Anglo-American Spondylitis Consortium (TASC), Burton PR, Clayton DG, Cardon LR, Craddock N, et al. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet* 2007;39:1329–1337. Available at: http://dx.doi.org/10.1038/ng.2007.17.

5. Australo-Anglo-American Spondyloarthritis Consortium (TASC), Reveille JD, Sims A-M, Danoy P, Evans DM, Leo P, et al. Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. *Nat Genet* 2010;42:123–127. Available at: http://dx.doi.org/10.1038/ng.513.

6. Evans DM, Spencer CCA, Pointon JJ, Su Z, Harvey D, Kochan G, et al. Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat Genet* 2011;43:761–767. Available at: http://dx.doi.org/10.1038/ng.873.

7. International Genetics of Ankylosing Spondylitis Consortium (IGAS), Cortes A, Hadler J, Pointon JP, Robinson PC, Karaderi T, et al. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nat Genet* 2013;45:730–738. Available at: http://dx.doi.org/10.1038/ng.2667.

8. 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;48:510–518. Available at: http://dx.doi.org/10.1038/ng.3528.

9. Mauro D, Simone D, Bucci L, Ciccia F. Novel immune cell phenotypes in spondyloarthritis pathogenesis. *Semin Immunopathol* 2021;43:265–277. Available at: http://dx.doi.org/10.1007/s00281-021-00837-0.

10. Li Z, Haynes K, Pennisi DJ, Anderson LK, Song X, Thomas GP, et al. Epigenetic and gene expression analysis of ankylosing spondylitis-associated loci implicate immune cells and the gut in the disease pathogenesis. *Genes Immun* 2017;18:135–143. Available at: http://dx.doi.org/10.1038/gene.2017.11.

11. Farh KK-H, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* 2015;518:337–343. Available at: http://dx.doi.org/10.1038/nature13835.

12. Costantino F, Breban M, Garchon H-J. Genetics and Functional Genomics of Spondyloarthritis. *Front Immunol* 2018;9:2933. Available at: http://dx.doi.org/10.3389/fimmu.2018.02933.

13. Yazar S, Alquicira-Hernandez J, Wing K, Senabouth A, Gordon MG, Andersen S, et al. Single-cell eQTL mapping identifies cell type-specific genetic control of autoimmune disease. *Science* 2022;376:eabf3041. Available at: http://dx.doi.org/10.1126/science.abf3041.

14. Brown AC, Cohen CJ, Mielczarek O, Migliorini G, Costantino F, Allcock A, et al. Comprehensive epigenomic profiling reveals the extent of disease-specific chromatin states and informs target discovery in ankylosing spondylitis. *Cell Genom* 2023;3:100306. Available at: http://dx.doi.org/10.1016/j.xgen.2023.100306.

15. Finucane HK, Reshef YA, Anttila V, Slowikowski K, Gusev A, Byrnes A, et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat*

Genet 2018;50:621–629. Available at: http://dx.doi.org/10.1038/s41588-018-0081-4.

16. Zhang MJ, Hou K, Dey KK, Sakaue S, Jagadeesh KA, Weinand K, et al. Polygenic enrichment distinguishes disease associations of individual cells in single-cell RNA-seq data. *Nat Genet* 2022;54:1572–1580. Available at: http://dx.doi.org/10.1038/s41588-022-01167-z.

17. Jagadeesh KA, Dey KK, Montoro DT, Mohan R, Gazal S, Engreitz JM, et al. Identifying disease-critical cell types and cellular processes by integrating single-cell RNA-sequencing and human genetics. *Nat Genet* 2022;54:1479–1492. Available at: http://dx.doi.org/10.1038/s41588-022-01187-9.

18. Trynka G, Sandor C, Han B, Xu H, Stranger BE, Liu XS, et al. Chromatin marks identify critical cell types for fine mapping complex trait variants. *Nat Genet* 2013;45:124–130. Available at: http://dx.doi.org/10.1038/ng.2504.

19. Calderon D, Nguyen MLT, Mezger A, Kathiria A, Müller F, Nguyen V, et al. Landscape of stimulation-responsive chromatin across diverse human immune cells. *Nat Genet* 2019;51:1494–1505. Available at: http://dx.doi.org/10.1038/s41588-019-0505-9.

20. Soskic B, Cano-Gamez E, Smyth DJ, Rowan WC, Nakic N, Esparza-Gordillo J, et al. Chromatin activity at GWAS loci identifies T cell states driving complex immune diseases. *Nat Genet* 2019;51:1486–1493. Available at: http://dx.doi.org/10.1038/s41588-019-0493-9.

21. Banerjee S, Webber C, Poole AR. The induction of arthritis in mice by the cartilage proteoglycan aggrecan: roles of CD4+ and CD8+ T cells. *Cell Immunol* 1992;144:347–357. Available at: http://dx.doi.org/10.1016/0008-8749(92)90250-s.

22. Kobezda T, Ghassemi-Nejad S, Mikecz K, Glant TT, Szekanecz Z. Of mice and men: how animal models advance our understanding of T-cell function in RA. *Nat Rev Rheumatol* 2014;10:160–170. Available at: http://dx.doi.org/10.1038/nrrheum.2013.205.

23. Rao DA, Gurish MF, Marshall JL, Slowikowski K, Fonseka CY, Liu Y, et al. Pathologically expanded peripheral T helper cell subset drives B cells in rheumatoid arthritis. *Nature* 2017;542:110–114. Available at: http://dx.doi.org/10.1038/nature20810.

24. Hu X, Kim H, Stahl E, Plenge R, Daly M, Raychaudhuri S. Integrating autoimmune risk loci with gene-expression data identifies specific pathogenic immune cell subsets. *Am J Hum Genet* 2011;89:496–506. Available at: http://dx.doi.org/10.1016/j.ajhg.2011.09.002.

25. Khunsriraksakul C, Li Q, Markus H, Patrick MT, Sauteraud R, McGuire D, et al. Multi-ancestry and multi-trait genome-wide association meta-analyses inform clinical risk prediction for systemic lupus erythematosus. *Nat Commun* 2023;14:668. Available at: http://dx.doi.org/10.1038/s41467-023-36306-5.

26. Caielli S, Wan Z, Pascual V. Systemic Lupus Erythematosus Pathogenesis: Interferon and Beyond. *Annu Rev Immunol* 2023;41:533–560. Available at: http://dx.doi.org/10.1146/annurev-immunol-101921-042422.

27. Vinuesa CG, Shen N, Ware T. Genetics of SLE: mechanistic insights from monogenic disease and disease-associated variants. *Nat Rev Nephrol* 2023. Available at: http://dx.doi.org/10.1038/s41581-023-00732-x.

544 28. Gutierrez-Arcelus M, Teslovich N, Mola AR, Polidoro RB, Nathan A, Kim H, et al. Lymphocyte
545 innateness defined by transcriptional states reflects a balance between proliferation and effector
546 functions. *Nat Commun* 2019;10:687. Available at: http://dx.doi.org/10.1038/s41467-019-08604-
547 4.

548 29. Slowikowski K, Hu X, Raychaudhuri S. SNPsea: an algorithm to identify cell types, tissues
549 and pathways affected by risk loci. *Bioinformatics* 2014;30:2496–2497. Available at:
550 http://dx.doi.org/10.1093/bioinformatics/btu326.

551 30. Brown MA, Wordsworth BP. Genetics in ankylosing spondylitis - Current state of the art and
552 translation into clinical outcomes. *Best Pract Res Clin Rheumatol* 2017;31:763–776. Available at:
553 https://www.sciencedirect.com/science/article/pii/S1521694218300548.

554 31. Davidson SI, Liu Y, Danoy PA, Wu X, Thomas GP, Jiang L, et al. Association of STAT3 and
555 TNFRSF1A with ankylosing spondylitis in Han Chinese. *Ann Rheum Dis* 2011;70:289–292.
556 Available at: http://dx.doi.org/10.1136/ard.2010.133322.

557 32. Karaderi T, Pointon JJ, Wordsworth TWH, Harvey D, Appleton LH, Cohen CJ, et al. Evidence
558 of genetic association between TNFRSF1A encoding the p55 tumour necrosis factor receptor,
559 and ankylosing spondylitis in UK Caucasians. *Clin Exp Rheumatol* 2012;30:110–113. Available
560 at: https://www.ncbi.nlm.nih.gov/pubmed/22272576.

561 33. Sode J, Bank S, Vogel U, Andersen PS, Sørensen SB, Bojesen AB, et al. Genetically
562 determined high activities of the TNF-alpha, IL23/IL17, and NFkB pathways were associated with
563 increased risk of ankylosing spondylitis. *BMC Med Genet* 2018;19:165. Available at:
564 http://dx.doi.org/10.1186/s12881-018-0680-z.

565 34. Robinson PC, Brown MA. Genetics of ankylosing spondylitis. *Mol Immunol* 2014;57:2–11.
566 Available at: http://dx.doi.org/10.1016/j.molimm.2013.06.013.

567 35. Vitulano C, Tedeschi V, Paladini F, Sorrentino R, Fiorillo MT. The interplay between HLA-B27
568 and ERAP1/ERAP2 aminopeptidases: from anti-viral protection to spondyloarthritis. *Clin Exp*
569 *Immunol* 2017;190:281–290. Available at: http://dx.doi.org/10.1111/cei.13020.

570 36. Tsui FW, Tsui HW, Akram A, Haroon N, Inman RD. The genetic basis of ankylosing
571 spondylitis: new insights into disease pathogenesis. *Appl Clin Genet* 2014;7:105–115. Available
572 at: http://dx.doi.org/10.2147/TACG.S37325.

573 37. Agrawal N, Brown MA. Genetic associations and functional characterization of M1
574 aminopeptidases and immune-mediated diseases. *Genes Immun* 2014;15:521–527. Available at:
575 http://dx.doi.org/10.1038/gene.2014.46.

576 38. Szabo PA, Levitin HM, Miron M, Snyder ME, Senda T, Yuan J, et al. Single-cell
577 transcriptomics of human T cells reveals tissue and activation signatures in health and disease.
578 *Nat Commun* 2019;10:1–16. Available at: https://www.nature.com/articles/s41467-019-12464-3.
579 Accessed August 4, 2023.

580 39. Meininger I, Carrasco A, Rao A, Soini T, Kokkinou E, Mjösberg J. Tissue-Specific Features of
581 Innate Lymphoid Cells. *Trends Immunol* 2020;41:902–917. Available at:
582 http://dx.doi.org/10.1016/j.it.2020.08.009.

583 40. Domínguez Conde C, Xu C, Jarvis LB, Rainbow DB, Wells SB, Gomes T, et al. Cross-tissue

584 immune cell analysis reveals tissue-specific features in humans. *Science* 2022;376:eabl5197.
585 Available at: http://dx.doi.org/10.1126/science.abl5197.

586 41. Mass E, Nimmerjahn F, Kierdorf K, Schlitzer A. Tissue-specific macrophages: how they
587 develop and choreograph tissue biology. *Nat Rev Immunol* 2023:1–17. Available at:
588 http://dx.doi.org/10.1038/s41577-023-00848-y.

589 42. Elmentaite R, Kumasaka N, Roberts K, Fleming A, Dann E, King HW, et al. Cells of the human
590 intestinal tract mapped across space and time. *Nature* 2021;597:250–255. Available at:
591 http://dx.doi.org/10.1038/s41586-021-03852-1.

592 43. Wang G, Sarkar A, Carbonetto P, Stephens M. A simple new approach to variable selection
593 in regression, with application to genetic fine mapping. *J R Stat Soc Series B Stat Methodol*
594 2020;82:1273–1300. Available at: http://dx.doi.org/10.1111/rssb.12388.

595 44. Kerimov N, Hayhurst JD, Peikova K, Manning JR, Walter P, Kolberg L, et al. A compendium
596 of uniformly processed human gene expression and splicing quantitative trait loci. *Nat Genet*
597 2021;53:1290–1299. Available at: http://dx.doi.org/10.1038/s41588-021-00924-w.

598 45. Schmiedel BJ, Singh D, Madrigal A, Valdovino-Gonzalez AG, White BM, Zapardiel-Gonzalo
599 J, et al. Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. *Cell*
600 2018;175:1701–1715.e16. Available at: http://dx.doi.org/10.1016/j.cell.2018.10.022.

601 46. Gilchrist JJ, Makino S, Naranbhai V, Sharma PK, Koturan S, Tong O, et al. Natural Killer cells
602 demonstrate distinct eQTL and transcriptome-wide disease associations, highlighting their role in
603 autoimmunity. *Nat Commun* 2022;13:4073. Available at: http://dx.doi.org/10.1038/s41467-022-
604 31626-4.

605 47. Vivier E, Tomasello E, Baratin M, Walzer T, Ugolini S. Functions of natural killer cells. *Nat*
606 *Immunol* 2008;9:503–510. Available at: http://dx.doi.org/10.1038/ni1582.

607 48. Ciccia F, Accardo-Palumbo A, Alessandro R, Rizzo A, Principe S, Peralta S, et al. Interleukin-
608 22 and interleukin-22-producing NKp44+ natural killer cells in subclinical gut inflammation in
609 ankylosing spondylitis. *Arthritis Rheum* 2012;64:1869–1878. Available at:
610 http://dx.doi.org/10.1002/art.34355.

611 49. Gracey E, Vereecke L, McGovern D, Fröhling M, Schett G, Danese S, et al. Revisiting the
612 gut–joint axis: links between gut inflammation and spondyloarthritis. *Nat Rev Rheumatol*
613 2020;16:415–433. Available at: https://www.nature.com/articles/s41584-020-0454-9. Accessed
614 January 12, 2024.

615 50. Cuthbert RJ, Fragkakis EM, Dunsmuir R, Li Z, Coles M, Marzo-Ortega H, et al. Brief Report:
616 Group 3 Innate Lymphoid Cells in Human Enthesis. *Arthritis Rheumatol* 2017;69:1816–1822.
617 Available at: http://dx.doi.org/10.1002/art.40150.

618 51. Peruzzi M, Wagtmann N, Long EO. A p70 killer cell inhibitory receptor specific for several
619 HLA-B allotypes discriminates among peptides bound to HLA-B*2705. *J Exp Med*
620 1996;184:1585–1590. Available at: http://dx.doi.org/10.1084/jem.184.4.1585.

621 52. Stewart-Jones GBE, Gleria K di, Kollnberger S, McMichael AJ, Jones EY, Bowness P. Crystal
622 structures and KIR3DL1 recognition of three immunodominant viral peptides complexed to HLA-
623 B*2705. *Eur J Immunol* 2005;35:341–351. Available at: http://dx.doi.org/10.1002/eji.200425724.

53. Malnati MS, Peruzzi M, Parker KC, Biddison WE, Ciccone E, Moretta A, et al. Peptide specificity in the recognition of MHC class I by natural killer cell clones. *Science* 1995;267:1016–1018. Available at: http://dx.doi.org/10.1126/science.7863326.

54. Kollnberger S, Chan A, Sun M-Y, Chen LY, Wright C, Gleria K di, et al. Interaction of HLA-B27 homodimers with KIR3DL1 and KIR3DL2, unlike HLA-B27 heterotrimers, is independent of the sequence of bound peptide. *Eur J Immunol* 2007;37:1313–1322. Available at: http://dx.doi.org/10.1002/eji.200635997.

55. Ridley A, Hatano H, Wong-Baeza I, Shaw J, Matthews KK, Al-Mossawi H, et al. Activation-Induced Killer Cell Immunoglobulin-like Receptor 3DL2 Binding to HLA-B27 Licenses Pathogenic T Cell Differentiation in Spondyloarthritis. *Arthritis Rheumatol* 2016;68:901–914. Available at: http://dx.doi.org/10.1002/art.39515.

56. Jansen DTSL, Hameetman M, Bergen J van, Huizinga TWJ, Heijde D van der, Toes REM, et al. IL-17-producing CD4+ T cells are increased in early, active axial spondyloarthritis including patients without imaging abnormalities. *Rheumatology* 2015;54:728–735. Available at: http://dx.doi.org/10.1093/rheumatology/keu382.

57. Larid G, Trijau S, Barral C, Lafforgue P, Pham T. Absence of overexpression of KIR3DL2 on CD4+ T cells and NK cells in patients with axial spondyloarthritis. *Rheumatology* 2023;62:e114–e116. Available at: http://dx.doi.org/10.1093/rheumatology/keac546.

58. Wang S, Li G, Ge R, Duan Z, Zeng Z, Zhang T, et al. Association of KIR genotype with susceptibility to HLA-B27-positive ankylosing spondylitis. *Mod Rheumatol* 2013;23:538–541. Available at: http://dx.doi.org/10.1007/s10165-012-0692-z.

59. Jiao Y-L, Zhang B-C, You L, Li J-F, Zhang J, Ma C-Y, et al. Polymorphisms of KIR gene and HLA-C alleles: possible association with susceptibility to HLA-B27-positive patients with ankylosing spondylitis. *J Clin Immunol* 2010;30:840–844. Available at: http://dx.doi.org/10.1007/s10875-010-9444-z.

60. Díaz-Peña R, Blanco-Gelaz MA, Suárez-Alvarez B, Martínez-Borra J, López-Vázquez A, Alonso-Arias R, et al. Activating KIR genes are associated with ankylosing spondylitis in Asian populations. *Hum Immunol* 2008;69:437–442. Available at: http://dx.doi.org/10.1016/j.humimm.2008.04.012.

61. Lopez-Larrea C, Blanco-Gelaz MA, Torre-Alonso JC, Bruges Armas J, Suarez-Alvarez B, Pruneda L, et al. Contribution of KIR3DL1/3DS1 to ankylosing spondylitis in human leukocyte antigen-B27 Caucasian populations. *Arthritis Res Ther* 2006;8:R101. Available at: http://dx.doi.org/10.1186/ar1988.

62. Harvey D, Pointon JJ, Sleator C, Meenagh A, Farrar C, Sun JY, et al. Analysis of killer immunoglobulin-like receptor genes in ankylosing spondylitis. *Ann Rheum Dis* 2009;68:595–598. Available at: http://dx.doi.org/10.1136/ard.2008.095927.

63. Babaie F, Mohammadi H, Salimi S, Ghanavatinegad A, Abbasifard M, Yousefi M, et al. Inhibition of ERAP1 represses HLA-B27 free heavy chains expression on polarized macrophages and interrupts NK cells activation and function from ankylosing spondylitis. *Clin Immunol* 2023;248:109268. Available at: http://dx.doi.org/10.1016/j.clim.2023.109268.

64. Sun JC, Lanier LL. NK cell development, homeostasis and function: parallels with $CD8^+$ T

665 cells. *Nat Rev Immunol* 2011;11:645–657. Available at: http://dx.doi.org/10.1038/nri3044.

666 65. Cohen NR, Brennan PJ, Shay T, Watts GF, Brigl M, Kang J, et al. Shared and distinct
667 transcriptional programs underlie the hybrid nature of iNKT cells. *Nat Immunol* 2013;14:90–99.
668 Available at: http://dx.doi.org/10.1038/ni.2490.

669 66. Robinette ML, Fuchs A, Cortez VS, Lee JS, Wang Y, Durum SK, et al. Transcriptional
670 programs define molecular characteristics of innate lymphoid cell classes and subsets. *Nat*
671 *Immunol* 2015;16:306–317. Available at: http://dx.doi.org/10.1038/ni.3094.

672 67. Gutierrez-Arcelus M, Ongen H, Lappalainen T, Montgomery SB, Buil A, Yurovsky A, et al.
673 Tissue-specific effects of genetic and epigenetic variation on gene regulation and splicing. *PLoS*
674 *Genet* 2015;11:e1004958. Available at: http://dx.doi.org/10.1371/journal.pgen.1004958.

675 68. Chen L, Ge B, Casale FP, Vasquez L, Kwan T, Garrido-Martín D, et al. Genetic Drivers of
676 Epigenetic and Transcriptional Variation in Human Immune Cells. *Cell* 2016;167:1398–1414.e24.
677 Available at: http://dx.doi.org/10.1016/j.cell.2016.10.026.

678 69. Costantino F, Talpin A, Evnouchidou I, Kadi A, Leboime A, Said-Nahal R, et al. ERAP1 Gene
679 Expression Is Influenced by Nonsynonymous Polymorphisms Associated With Predisposition to
680 Spondyloarthritis. *Arthritis Rheumatol* 2015;67:1525–1534. Available at:
681 http://dx.doi.org/10.1002/art.39072.

682 70. Chen L, Ridley A, Hammitzsch A, Al-Mossawi MH, Bunting H, Georgiadis D, et al. Silencing
683 or inhibition of endoplasmic reticulum aminopeptidase 1 (ERAP1) suppresses free heavy chain
684 expression and Th17 responses in ankylosing spondylitis. *Ann Rheum Dis* 2016;75:916–923.
685 Available at: http://dx.doi.org/10.1136/annrheumdis-2014-206996.

686 71. Hanson AL, Cuddihy T, Haynes K, Loo D, Morton CJ, Oppermann U, et al. Genetic Variants
687 in ERAP1 and ERAP2 Associated With Immune-Mediated Diseases Influence Protein Expression
688 and the Isoform Profile. *Arthritis Rheumatol* 2018;70:255–265. Available at:
689 http://dx.doi.org/10.1002/art.40369.

690 72. Zhao S, Chen H, Wu G, Zhao C. The association of NLRP3 and TNFRSF1A polymorphisms
691 with risk of ankylosing spondylitis and treatment efficacy of etanercept. *J Clin Lab Anal* 2017;31.
692 Available at: http://dx.doi.org/10.1002/jcla.22138.

693 73. Xing-Rong W, Sheng-Qian X, Wen L, Shan Q, Fa-Ming P, Jian-Hua X. Role of TNFRSF1A
694 and TNFRSF1B polymorphisms in susceptibility, severity, and therapeutic efficacy of etanercept
695 in human leukocyte antigen-B27-positive Chinese Han patients with ankylosing spondylitis.
696 *Medicine* 2018;97:e11677. Available at: http://dx.doi.org/10.1097/MD.0000000000011677.

697 74. Wang R, Jaw JJ, Stutzman NC, Zou Z, Sun PD. Natural killer cell-produced IFN-γ and TNF-
698 α induce target cell cytolysis through up-regulation of ICAM-1. *J Leukoc Biol* 2012;91:299–309.
699 Available at: http://dx.doi.org/10.1189/jlb.0611308.

700 75. Neznanov N, Neznanova L, Angres B, Gudkov AV. Serologically defined colon cancer antigen
701 3 is necessary for the presentation of TNF receptor 1 on cell surface. *DNA Cell Biol* 2005;24:777–
702 785. Available at: http://dx.doi.org/10.1089/dna.2005.24.777.

703 76. Wang C-M, Jan Wu Y-J, Lin J-C, Huang L-Y, Wu J, Chen J-Y. Genetic effects of B3GNT2 on
704 ankylosing spondylitis susceptibility and clinical manifestations in Taiwanese. *J Formos Med*

Assoc 2022;121:1283–1294. Available at: http://dx.doi.org/10.1016/j.jfma.2021.09.010.

77. Chun S, Casparino A, Patsopoulos NA, Croteau-Chonka DC, Raby BA, De Jager PL, et al. Limited statistical evidence for shared genetic effects of eQTLs and autoimmune-disease-associated loci in three major immune-cell types. *Nat Genet* 2017;49:600–605. Available at: http://dx.doi.org/10.1038/ng.3795.

78. Mu Z, Wei W, Fair B, Miao J, Zhu P, Li YI. The impact of cell type and context-dependent regulatory variants on human immune traits. *Genome Biol* 2021;22:1–28. Available at: https://genomebiology.biomedcentral.com/articles/10.1186/s13059-021-02334-x. Accessed September 20, 2023.

79. Barbeira AN, Bonazzola R, Gamazon ER, Liang Y, Park Y, Kim-Hellmuth S, et al. Exploiting the GTEx resources to decipher the mechanisms at GWAS loci. *Genome Biol* 2021;22:49. Available at: http://dx.doi.org/10.1186/s13059-020-02252-4.

80. Gutierrez-Arcelus M, Rich SS, Raychaudhuri S. Autoimmune diseases - connecting risk alleles with molecular traits of the immune system. *Nat Rev Genet* 2016;17:160–174. Available at: http://dx.doi.org/10.1038/nrg.2015.33.

81. Gutierrez-Arcelus M, Baglaenko Y, Arora J, Hannes S, Luo Y, Amariuta T, et al. Allele-specific expression changes dynamically during T cell activation in HLA and other autoimmune loci. *Nat Genet* 2020;52:247–253. Available at: http://dx.doi.org/10.1038/s41588-020-0579-4.

82. Umans BD, Battle A, Gilad Y. Where Are the Disease-Associated eQTLs? *Trends Genet* 2021;37:109–124. Available at: http://dx.doi.org/10.1016/j.tig.2020.08.009.

83. Connally NJ, Nazeen S, Lee D, Shi H, Stamatoyannopoulos J, Chun S, et al. The missing link between genetic association and regulatory function. *Elife* 2022;11. Available at: http://dx.doi.org/10.7554/eLife.74970.

84. Mostafavi H, Spence PJ, Naqvi S, Jonathan PK. Limited overlap of eQTLs and GWAS hits due to systematic differences in discovery. *bioRxiv* 2022. Available at: ;

85. Venken K, Jacques P, Mortier C, Labadia ME, Decruy T, Coudenys J, et al. RORγt inhibition selectively targets IL-17 producing iNKT and γδ-T cells enriched in Spondyloarthritis patients. *Nat Commun* 2019;10:9. Available at: http://dx.doi.org/10.1038/s41467-018-07911-6.

86. Yu H, Wu H, Zheng F, Zhu C, Yin L, Dai W, et al. Gene-regulatory network analysis of ankylosing spondylitis with a single-cell chromatin accessible assay. *Sci Rep* 2020;10:19411. Available at: http://dx.doi.org/10.1038/s41598-020-76574-5.

87. Simone D, Penkava F, Ridley A, Sansom S, Al-Mossawi MH, Bowness P. Single cell analysis of spondyloarthritis regulatory T cells identifies distinct synovial gene expression patterns and clonal fates. *Commun Biol* 2021;4:1395. Available at: http://dx.doi.org/10.1038/s42003-021-02931-3.

88. Alber S, Kumar S, Liu J, Huang Z-M, Paez D, Hong J, et al. Single Cell Transcriptome and Surface Epitope Analysis of Ankylosing Spondylitis Facilitates Disease Classification by Machine Learning. *Front Immunol* 2022;13:838636. Available at: http://dx.doi.org/10.3389/fimmu.2022.838636.

744 89. Ren C, Li M, Zheng Y, Cai B, Du W, Zhang H, et al. Single-cell RNA-seq reveals altered NK
745 cell subsets and reduced levels of cytotoxic molecules in patients with ankylosing spondylitis. *J*
746 *Cell Mol Med* 2022;26:1071–1082. Available at: http://dx.doi.org/10.1111/jcmm.17159.

747 90. Yi K, Jo S, Song W, Lee H-I, Kim H-J, Kang J-H, et al. Analysis of Single-Cell Transcriptome
748 and Surface Protein Expression in Ankylosing Spondylitis Identifies OX40-Positive and
749 Glucocorticoid-Induced Tumor Necrosis Factor Receptor-Positive Pathogenic Th17 Cells.
750 *Arthritis Rheumatol* 2023;75:1176–1186. Available at: http://dx.doi.org/10.1002/art.42476.

751 91. Pan-UKB team. Pan-UK Biobank. 2020. Available at: https://pan.ukbb.broadinstitute.org.
752 Accessed 2023.

753 92. Finucane HK, Bulik-Sullivan B, Gusev A, Trynka G, Reshef Y, Loh P-R, et al. Partitioning
754 heritability by functional annotation using genome-wide association summary statistics. *Nat Genet*
755 2015;47:1228–1235. Available at: http://dx.doi.org/10.1038/ng.3404.

756 93. Myers S, Bottolo L, Freeman C, McVean G, Donnelly P. A fine-scale map of recombination
757 rates and hotspots across the human genome. *Science* 2005;310:321–324. Available at:
758 http://dx.doi.org/10.1126/science.1117196.

759 94. Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, et al. Hundreds
760 of variants clustered in genomic loci and biological pathways affect human height. *Nature*
761 2010;467:832–838. Available at: http://dx.doi.org/10.1038/nature09410.

762 95. Leeuw CA de, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of
763 GWAS data. *PLoS Comput Biol* 2015;11:e1004219. Available at:
764 http://dx.doi.org/10.1371/journal.pcbi.1004219.

765 96. Yuhan H, Stephanie H, Erica A-N, Mauck WM, Shiwei Z, Andrew B, et al. Integrated analysis
766 of multimodal single-cell data. *Cell* 2021;184:3573–3587.e29. Available at:
767 http://dx.doi.org/10.1016/j.cell.2021.04.048. Accessed May 31, 2023.

768 97. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and
769 accurate integration of single-cell data with Harmony. *Nat Methods* 2019;16:1289–1296.
770 Available at: http://dx.doi.org/10.1038/s41592-019-0619-0.

771 98. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al.
772 Bayesian test for colocalisation between pairs of genetic association studies using summary
773 statistics. *PLoS Genet* 2014;10:e1004383. Available at:
774 http://dx.doi.org/10.1371/journal.pgen.1004383.

775 99. Kramer NE, Davis ES, Wenger CD, Deoudes EM, Parker SM, Love MI, et al. Plotgardener:
776 cultivating precise multi-panel figures in R. *Bioinformatics* 2022;38:2042–2045. Available at:
777 http://dx.doi.org/10.1093/bioinformatics/btac057.

778 100. Liu B, Gloudemans MJ, Rao AS, Ingelsson E, Montgomery SB. Abundant associations with
779 gene expression complicate GWAS follow-up. *Nat Genet* 2019;51:768–769. Available at:
780 http://dx.doi.org/10.1038/s41588-019-0404-0.

781

782 **Table 1. Putative target genes identified by co-localization analysis**
783 **between AS-associated loci and eQTLs in NK cells.**

Lead GWAS Variant	P-value	GWAS	Putative Target Gene	Posterior probability of shared causal variant	Quantification method	eQTL study
rs27529	1.24E-40	IGAS	ERAP1	0.99186	microarray	Gilchrist et al. 2021
rs6759298	2.07E-38	IGAS	B3GNT2	0.97245	RNA-seq	Schmiedel et al. 2018
rs1128905	3.17E-10	IGAS	ENTR1	0.8207	microarray	Gilchrist et al. 2021
rs1860545	8.66E-10	IGAS	TNFRSF1A	0.99673	RNA-seq	Schmiedel et al. 2018
rs11065898	5.41E-08	IGAS	TMEM116	0.87563	microarray	Gilchrist et al. 2021
rs9619386	4.42E-07	IGAS	UBE2L3	0.96397	microarray	Gilchrist et al. 2021
rs1250542	2.07E-06	IGAS	ZMIZ1	0.87272	microarray	Gilchrist et al. 2021
rs1250542	2.07E-06	IGAS	ZMIZ1	0.83268	RNA-seq	Schmiedel et al. 2018
rs952594	2.08E-06	IGAS	APEH	0.75689	microarray	Gilchrist et al. 2021
rs952594	2.08E-06	IGAS	RBM6	0.91981	RNA-seq	Schmiedel et al. 2018
rs952594	2.08E-06	IGAS	UBA7	0.94237	microarray	Gilchrist et al. 2021
rs952594	2.08E-06	IGAS	UBA7	0.8851	RNA-seq	Schmiedel et al. 2018
rs6565217	2.82E-06	IGAS	AC135050.3	0.91556	RNA-seq	Schmiedel et al. 2018
rs6565217	2.82E-06	IGAS	STX4	0.94653	RNA-seq	Schmiedel et al. 2018
rs7191548	3.13E-06	IGAS	EIF3CL	0.7822	microarray	Gilchrist et al. 2021
rs7191548	3.13E-06	IGAS	NPIPB8	0.97545	microarray	Gilchrist et al. 2021
rs7191548	3.13E-06	IGAS	SGF29	0.85225	microarray	Gilchrist et al. 2021
rs7191548	3.13E-06	IGAS	TUFM	0.82246	microarray	Gilchrist et al. 2021
rs7191548	3.13E-06	IGAS	TUFM	0.95694	RNA-seq	Schmiedel et al. 2018
rs6583441	3.84E-06	IGAS	IKZF1	0.96244	RNA-seq	Schmiedel et al. 2018
rs4690326	6.49E-06	IGAS	DGKQ	0.91883	microarray	Gilchrist et al. 2021
rs4690326	6.49E-06	IGAS	DGKQ	0.89989	RNA-seq	Schmiedel et al. 2018
rs4690326	6.49E-06	IGAS	IDUA	0.98988	microarray	Gilchrist et al. 2021
rs4690326	6.49E-06	IGAS	SLC49A3	0.801	RNA-seq	Schmiedel et al. 2018
rs26481	9.17E-06	UK Biobank	CAST	0.95638	microarray	Gilchrist et al. 2021
rs26481	9.17E-06	UK Biobank	ERAP1	0.95565	microarray	Gilchrist et al. 2021
rs2236167	3.57E-05	IGAS	PPP2R3C	0.9491	microarray	Gilchrist et al. 2021

784

Figure legends

Figure 1. Human NK cell-specific open chromatin regions are enriched in AS genetic risk. (**A**) Cartoon depicting the Calderon *et al.* study design. Peripheral blood cells from 4 healthy subjects were sorted into immune cell populations that we grouped *in silico* into seven cell types (see Methods). Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) was performed with and without prior *in vitro* activation. (**B**) Graphical representation of LDSC-SEG analysis: identification of cell type-specific annotations (in our case open chromatin regions), followed by the integration with GWAS summary statistics to obtain a risk enrichment coefficient β and P-value. (**C**) Volcano plots showing results of differential accessibility analyses for each cell type compared to the other cell types. Colored dots indicate open chromatin peaks in the top decile of the t-statistic for each cell type, which were used for LDSC-SEG analysis. (**D-E**) Bar graphs displaying the AS genetic risk enrichment coefficient β and block jackknife standard error for cell type-specific open chromatin accounting for control peaks and baseline annotations. Summary statistics from the International Genetics of Ankylosing Spondylitis Consortium (IGAS) (**D**) and UK Biobank (**E**) GWAS were used. * indicates $P < 0.05$.

Figure 2. NK cells show enrichment of cell type-specific expression of AS-associated genes. (**A**) Cartoon depicting the Gutierrez-Arcelus *et al.* study. Peripheral blood cells from 6 healthy subjects were sorted into NK cells (orange) and six T cell populations (purple): CD4+ T, CD8+ T, MAIT, iNKT, and two γδ T cell populations. Bulk RNA sequencing was performed on two replicates per sample. (**B**) Graphical representation of the SNPsea method illustrating the integration of gene expression profiles with risk loci obtained from GWAS. (**C**) Bar graphs showing -log10(P-value) for enrichment of cell type-specific expression of genes in AS risk loci using SNPsea. (**D**) Heatmap showing expression levels for genes in AS risk loci that were significantly upregulated in NK cells compared to six T cells subsets. Expression levels are scaled by row. Tpm: transcripts per million. ** indicates $P < 0.01$.

Figure 3. Human gut single-cell atlas reveals significant upregulation of AS-associated genes in NK cells. (**A**) Cartoon depicting the generation of the Space-Time Gut Cell Atlas with samples from fetal, pediatric and adult subjects. (**B**) Graphical representation of the scDRS method, which integrates GWAS risk genes with single cell data to identify disease-relevant cells. (**C**) Visualization of the Space-Time Gut Cell Atlas data using UMAP on the top 20 principal components from 1,997 variable genes from the scRNA-seq expression matrix. (**D**) Same UMAP visualization as in C. Cells with significant scDRS score (20% FDR) are colored in red. (**E**) Bar graph showing enrichment of scDRS significant cells per cell type (cell-type percent in whole dataset over cell-type percent within scDRS significant cells). (**F**) Bar graph showing the number of significant scDRS cells for each cell type using the fine-grained annotations from the Space-Time Gut Cell Atlas. Cell populations with at least 15 significant scDRS cells are shown. (**G**) Scaled average expression levels and percent of cells expressing a given gene for 50 genes associated with AS that had significant upregulation (5% FDR) in NK cells compared to the other cell types. Genes are sorted by multiplying their MAGMA score (strength of association with AS) by their average level of expression in NK cells.

Figure. 4. Co-localization of AS risk loci and NK cell eQTLs points to putative target genes for AS risk variants. (**A-D**) Manhattan plots showing AS GWAS and NK cell eQTL -log10(P-values) for SNPs within 500 kb of a lead GWAS SNP. The color of each SNP indicates its level of linkage disequilibrium (LD) between with the lead GWAS SNP (purple diamond). Genes in the region are colored according to their posterior probability of hypothesis four (PP4), i.e. that the same causal variant is shared between AS and the eQTL for that gene. (**A**) Manhattan plots identifying putative target gene *ERAP1* using AS IGAS GWAS (top) and NK microarray gene

835 expression QTL (eQTL) data obtained from Gilchrist *et al*. (bottom) (**B**) Manhattan plots identifying
836 putative target gene *TNFRSF1A* using AS IGAS GWAS (top) and NK gene expression QTL
837 (eQTL) data obtained from Schmiedel *et al*. (bottom) (**C**) Manhattan plots identifying putative
838 target gene *ENTR1* using AS IGAS GWAS (top) and NK microarray gene expression QTL (eQTL)
839 data obtained from Gilchrist *et al*. (bottom) (**D**) Manhattan plots identifying putative target gene
840 *B3GNT2* using AS IGAS GWAS (top) and NK gene expression QTL (eQTL) data obtained from
841 Schmiedel *et al*. (bottom). All QTL summary statistics taken from eQTL Catalogue.
842
843 **Supplementary Figure 1. Heritability enrichment results for control traits.** (**A**) Bar graphs
844 display the genetic risk enrichment coefficient (y-axis) and standard error for cell-type specific
845 open chromatin accounting for control peaks and baseline annotations. Open chromatin data
846 were taken from the Calderon *et al*. study. Risk enrichment was assessed using GWAS summary
847 statistics for the positive control traits rheumatoid arthritis, Alzheimer's disease, systemic lupus
848 erythematosus, and the negative control trait height. Bars marked with "*" indicate $P < 0.05$, "**"
849 indicates $P < 0.01$, "***" indicates $P < 0.001$.
850
851 **Supplementary Figure 2. Single-cell disease relevant score results for control traits.** (**A**)
852 Visualization of the Space-Time Gut Cell Atlas using Uniform Manifold Approximation and
853 Projection (UMAP) on the top 20 principal components from 1,997 variable genes from the single-
854 cell RNA-seq expression matrix. Cells are colored based on the coarse cell type annotations from
855 the Space-Time Gut Cell Atlas. (**B**) Barplots shows the cell type proportions within the whole
856 Space-Time Gut Cell Atlas and within cells with significant disease relevant score (20% FDR) for
857 AS (using IGAS GWAS), Alzheimer's disease (AD) and rheumatoid arthritis (RA). (**C**) Same
858 UMAP visualization as in A, where cells with significant scDRS score (20% FDR) are colored in
859 red and non-significant cells are colored in gray, for each control trait.

Figure 1

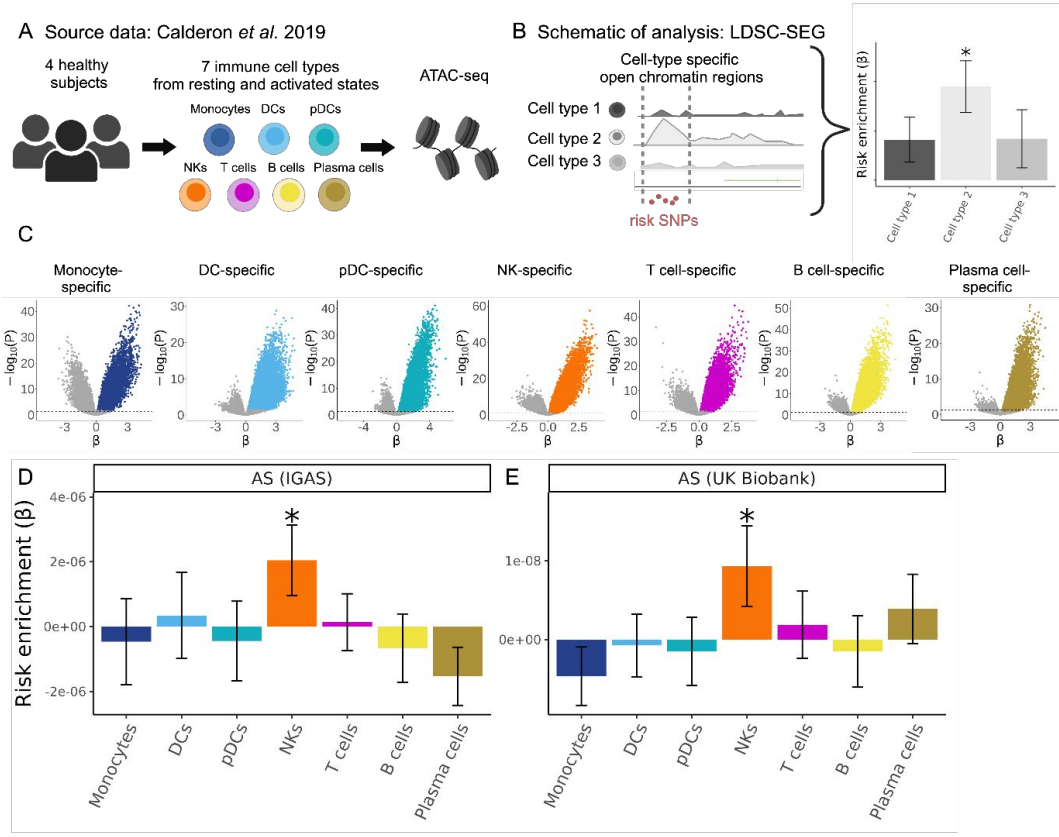

A Source data: Calderon et al. 2019
4 healthy subjects
7 immune cell types from resting and activated states
Monocytes
DCs
pDCs
NKs
T cells
B cells
Plasma cells
ATAC-seq
B Schematic of analysis: LDSC-SEG
Cell-type specific open chromatin regions
Cell type 1
Cell type 2
Cell type 3
risk SNPs
Risk enrichment (β)
Cell type 1
Cell type 2
Cell type 3
C
Monocyte-specific
DC-specific
pDC-specific
NK-specific
T cell-specific
B cell-specific
Plasma cell-specific
− log10(P)
β
D
AS (IGAS)
Risk enrichment (β)
4e-06
2e-06
0e+00
-2e-06
Monocytes
DCs
pDCs
NKs
T cells
B cells
Plasma cells
E
AS (UK Biobank)
1e-08
0e+00
Monocytes
DCs
pDCs
NKs
T cells
B cells
Plasma cells

Figure 2

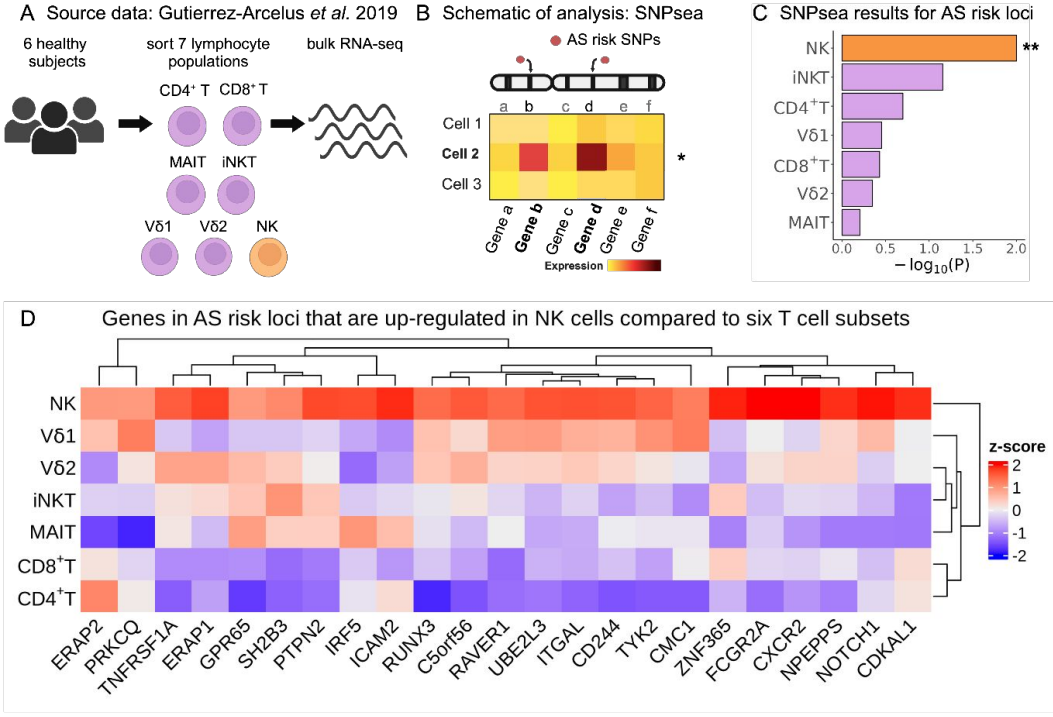
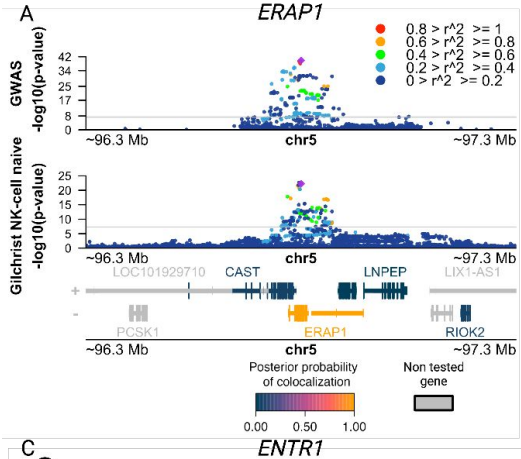
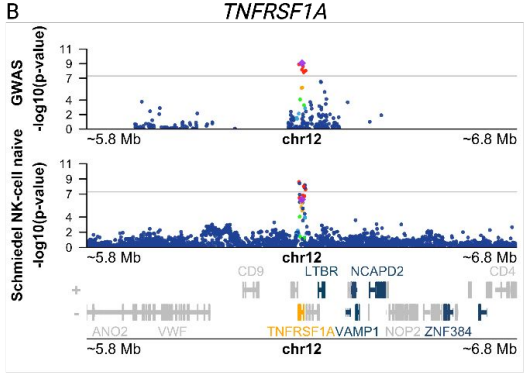

A Source data: Gutierrez-Arcelus et al. 2019
6 healthy subjects
sort 7 lymphocyte populations
bulk RNA-seq
CD4+ T
CD8+ T
MAIT
iNKT
Vδ1
Vδ2
NK
B Schematic of analysis: SNPsea
AS risk SNPs
a b c d e f
Cell 1
Cell 2
Cell 3
*
Gene a
Gene b
Gene c
Gene d
Gene e
Gene f
Expression
C SNPsea results for AS risk loci
NK
iNKT
CD4+T
Vδ1
CD8+T
Vδ2
MAIT
**
0.0 0.5 1.0 1.5 2.0
– log10(P)
D Genes in AS risk loci that are up-regulated in NK cells compared to six T cell subsets
NK
Vδ1
Vδ2
iNKT
MAIT
CD8+T
CD4+T
z-score
2
1
0
-1
-2
ERAP2
PRKCQ
TNFRSF1A
ERAP1
GPR65
SH2B3
PTPN2
IRF5
ICAM2
RUNX3
C5orf56
RAVER1
UBE2L3
ITGAL
CD244
TYK2
CMC1
ZNF365
FCGR2A
CXCR2
NPEPPS
NOTCH1
CDKAL1

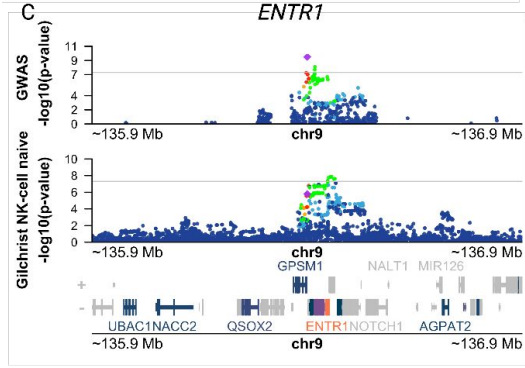
Figure 3


A Source data: Human gut cell atlas
Fetal n = 16
Pediatric n = 8,
CD n = 7, adult n = 6
Single cell RNA-seq
B Schematic of analysis: scDRS
Cell type 2
Cell type 3
Cell type 1
UMAP 2
UMAP 1
scDRS identifies cells with upregulated GWAS genes
C
Mesenchymal
Epithelial
Red blood cells
Plasma cells
Other ILCs
NK cells
B cells
Neuronal
Myeloid
T cells
Endothelial
D
cells with significant AS relevant score
E
NK cells
Other ILCs
Myeloid
T cells
Endothelial
Epithelial
Mesenchymal
B cells
Neuronal
Plasma cells
Enrichment of disease relevant cells
F
NK cells
Activated CD8 T
Macrophages
LTi-like NCR+ ILC3
LTi-like NCR- ILC3
gdT
Activated T
TRGV4 gdT
Treg
LYVE1+ Macrophage
TRGV2 gdT
CD8 Tmem
ILCP
CX3CR1+ CD8 Tmem
Monocytes
ILC3
NK T cell
Activated CD4 T
cDC2
SELL+ CD4 T
TA
Cell types
NK cells
Other ILCs
T cells
Epithelial
Myeloid
Number of disease relevant cells
G
GNLY
CCL4
CCL3
CTSW
EVL
ZAP70
HSPA6
SH2D2A
ETS1
FCGR3A
TBX21
SKAP1
NFKBIA
DENND2D
GPR65
IL18RAP
TGFB1
RASA2
IRF1
IKZF3
LNPEP
PDE4A
KIF21B
RAB27A
IL12RB2
CLN3
BCO2
PTPN2
CARHSP1
DGKD
NK cells
T cells
Other ILCs
Myeloid
Neuronal
Red blood cells
Epithelial
Mesenchymal
Endothelial
B cells
Plasma cells
Average Expression
Percent Expressed

Figure 4


A
ERAP1
0.8 > r^2 >= 1
0.6 > r^2 >= 0.8
0.4 > r^2 >= 0.6
0.2 > r^2 >= 0.4
0 > r^2 >= 0.2
GWAS -log10(p-value)
Gilchrist NK-cell naive -log10(p-value)
~96.3 Mb
chr5
~97.3 Mb
LOC101929710
CAST
LNPEP
LIX1-AS1
PCSK1
ERAP1
RIOK2
Posterior probability of colocalization
0.00 0.50 1.00
Non tested gene


B
TNFRSF1A
GWAS -log10(p-value)
Schmiedel NK-cell naive -log10(p-value)
~5.8 Mb
chr12
~6.8 Mb
CD9
LTBR
NCAPD2
CD4
ANO2
VWF
TNFRSF1A
VAMP1
NOP2
ZNF384


C
ENTR1
GWAS -log10(p-value)
Gilchrist NK-cell naive -log10(p-value)
~135.9 Mb
chr9
~136.9 Mb
GPSM1
NALT1
MIR126
UBAC1
NACC2
QSOX2
ENTR1
NOTCH1
AGPAT2

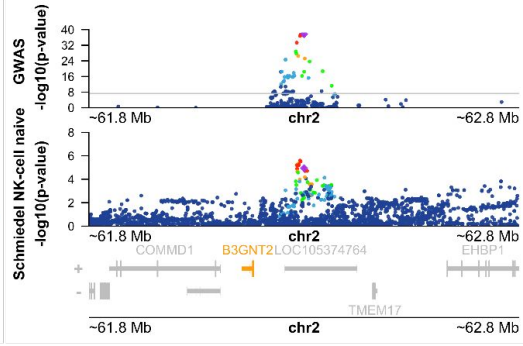

D
B3GNT2
GWAS -log10(p-value)
Schmiedel NK-cell naive -log10(p-value)
~61.8 Mb
chr2
~62.8 Mb
COMMD1
B3GNT2
LOC105374764
EHBP1
TMEM17

Figure S1

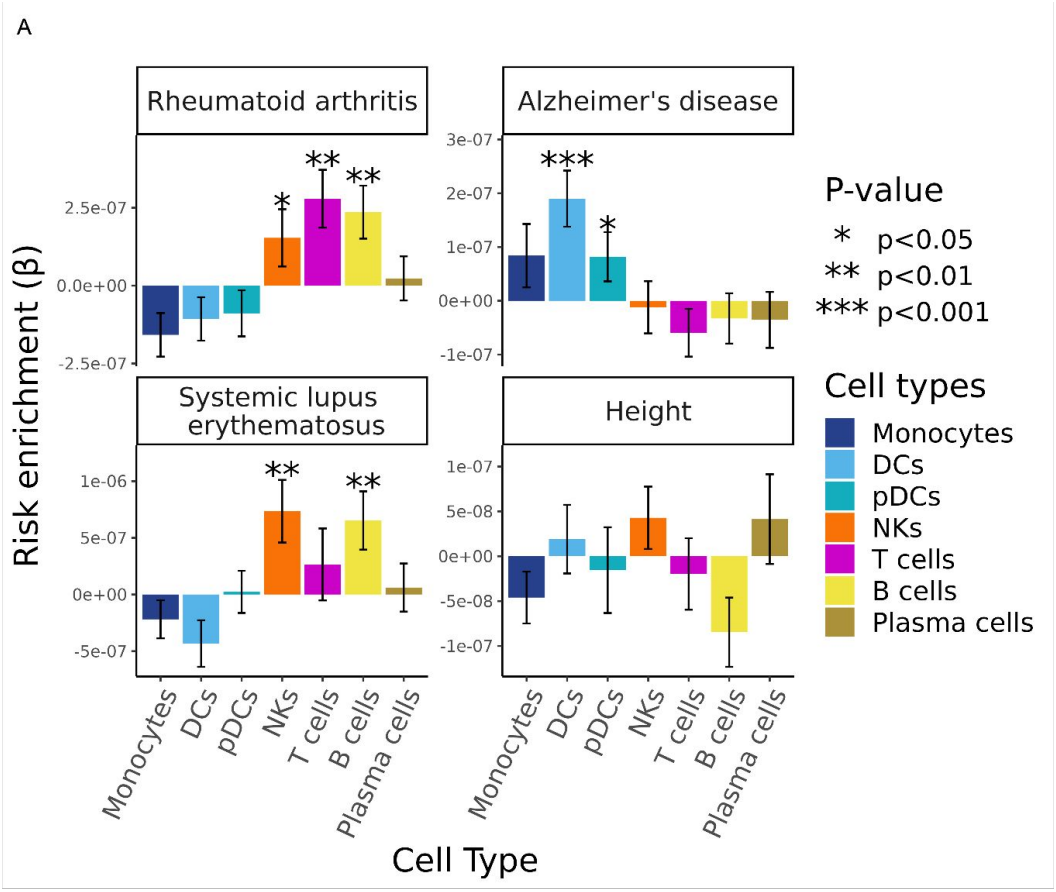
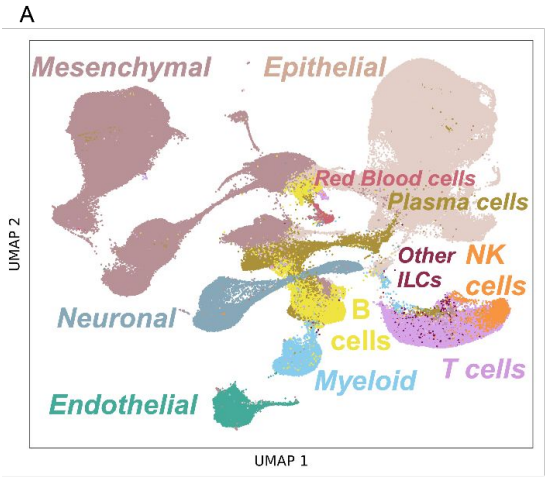
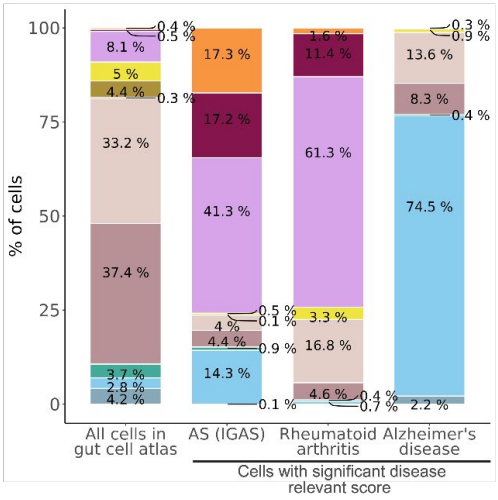
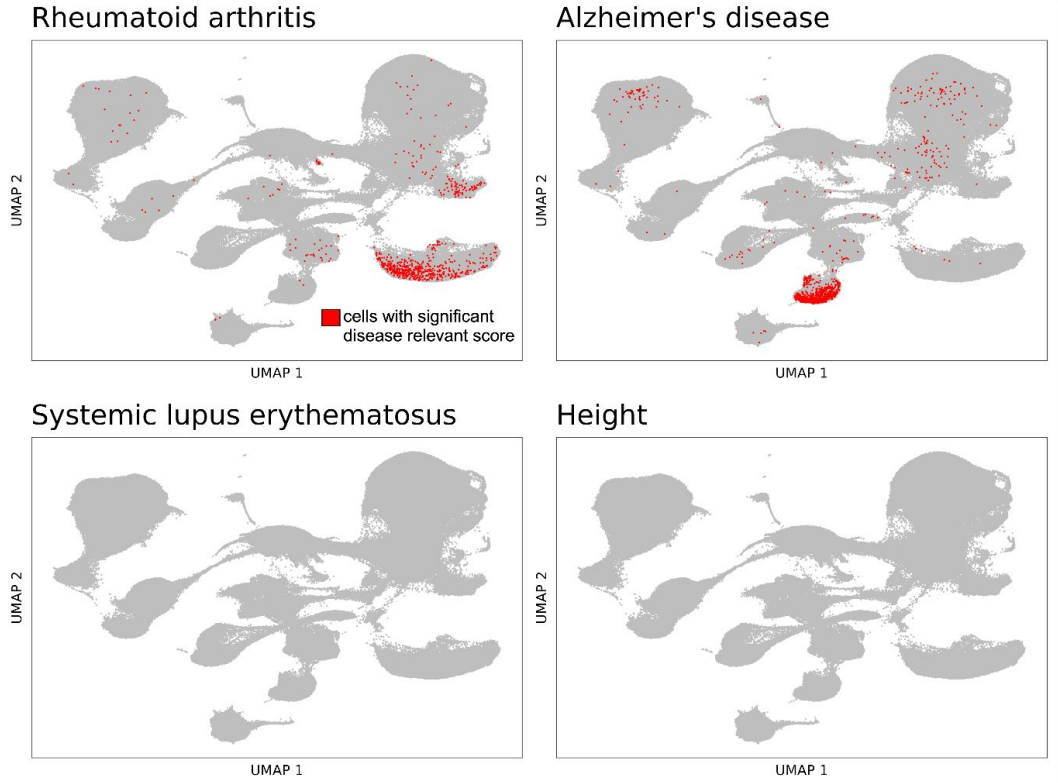

A
Rheumatoid arthritis
Alzheimer's disease
Systemic lupus erythematosus
Height
Risk enrichment (β)
Cell Type
Monocytes
DCs
pDCs
NKs
T cells
B cells
Plasma cells
P-value
* p<0.05
** p<0.01
*** p<0.001
Cell types

Figure S2


A
Mesenchymal
Epithelial
Red Blood cells
Plasma cells
Other ILCs
NK cells
Neuronal
B cells
Myeloid
T cells
Endothelial
UMAP 2
UMAP 1


B
% of cells
100
75
50
25
0
All cells in gut cell atlas
AS (IGAS)
Rheumatoid arthritis
Alzheimer's disease
Cells with significant disease relevant score


C
Rheumatoid arthritis
Alzheimer's disease
cells with significant disease relevant score
UMAP 2
UMAP 1
Systemic lupus erythematosus
Height