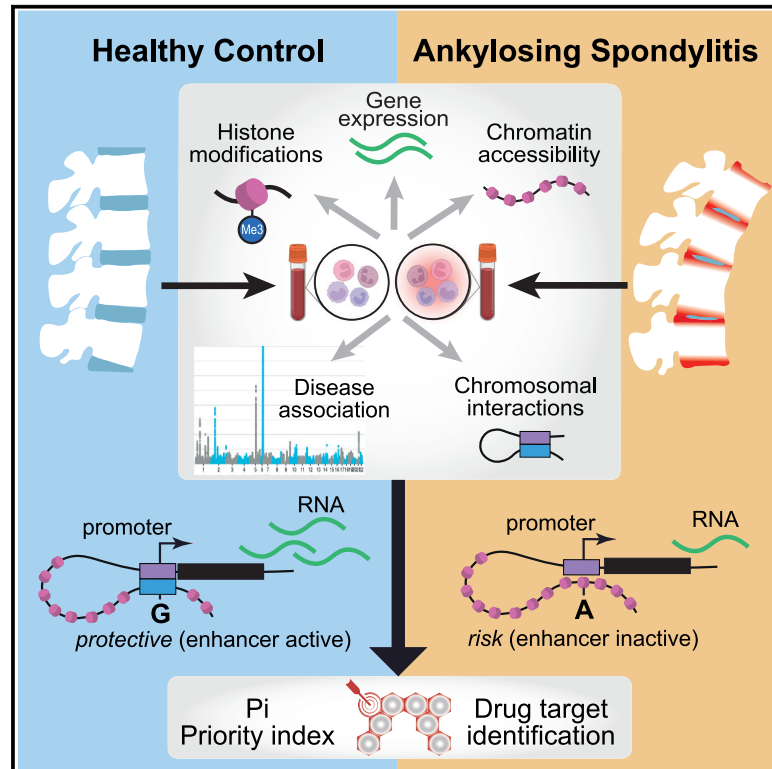


# Comprehensive epigenomic profiling reveals the extent of disease-specific chromatin states and informs target discovery in ankylosing spondylitis

## Graphical abstract



## Authors

Andrew C. Brown, Carla J. Cohen, Olga Mielczarek, ..., B. Paul Wordsworth, Paul Bowness, Julian C. Knight

## Correspondence

julian.knight@well.ox.ac.uk

## In brief

Brown et al. performed comprehensive transcriptomic and epigenomic mapping in ankylosing spondylitis. Comparison of blood immune cells from patients and healthy controls revealed disease-specific chromatin patterns with potential pathogenic effects via mechanisms such as chromosome looping. The authors used these results to prioritize genetic variants, genes, and pathways for drug discovery.

## Highlights

- Comprehensive transcriptomic and epigenomic mapping in ankylosing spondylitis
- Disease-specific chromatin patterns seen upon integration of multiple modalities
- Evidence of functional SNPs via chromosome looping events at GWAS regions
- Gene and pathway prioritization facilitates future drug discovery



## Article

# Comprehensive epigenomic profiling reveals the extent of disease-specific chromatin states and informs target discovery in ankylosing spondylitis

Andrew C. Brown,<sup>1,14</sup> Carla J. Cohen,<sup>1,2,3,14</sup> Olga Mielczarek,<sup>1,4,14</sup> Gabriele Migliorini,<sup>1,5,14</sup> F elicie Costantino,<sup>1,6,7</sup> Alice Allcock,<sup>1</sup> Connor Davidson,<sup>1,3</sup> Katherine S. Elliott,<sup>1</sup> Hai Fang,<sup>1,8</sup> Alicia Lled o Lara,<sup>1</sup> Alice C. Martin,<sup>1</sup> Julie A. Osgood,<sup>1</sup> Anna Sanniti,<sup>1</sup> Giuseppe Scozzafava,<sup>1</sup> Matteo Vecellio,<sup>3,9</sup> Ping Zhang,<sup>1,10</sup> Mary Helen Black,<sup>11</sup> Shuwei Li,<sup>11</sup> Dongnhu Truong,<sup>11</sup> Julio Molineros,<sup>11</sup> Trevor Howe,<sup>12</sup> B. Paul Wordsworth,<sup>3,13</sup> Paul Bowness,<sup>3,13</sup> and Julian C. Knight<sup>1,10,13,15,\*</sup>

<sup>1</sup>Wellcome Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK

<sup>2</sup>MRC WIMM Centre for Computational Biology, MRC Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DS, UK

<sup>3</sup>Nuffield Department of Orthopaedics, Rheumatology, and Musculoskeletal Sciences, University of Oxford, Oxford OX3 7LD, UK

<sup>4</sup>Horizon Discovery (PerkinElmer) Cambridge Research Park, 8100 Beach Dr., Waterbeach, Cambridge CB25 9TL, UK

<sup>5</sup>Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

<sup>6</sup>UVSQ, INSERM UMR1173, Infection et Inflammation, Laboratory of Excellence INFLAMEX, Universit e Paris-Saclay, Paris, France

<sup>7</sup>Rheumatology Department, AP-HP, Ambroise Par e Hospital, Paris, France

<sup>8</sup>Shanghai Institute of Hematology, State Key Laboratory of Medical Genomics, National Research Centre for Translational Medicine at Shanghai, Ruijin Hospital affiliated with Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China

<sup>9</sup>Centro Ricerche Fondazione Italiana Ricerca sull'Artrite (FIRA), Fondazione Pisana per la Scienza ONLUS, Via Ferruccio Giovannini 13, 56017 San Giuliano Terme (Pisa), Italy

<sup>10</sup>Chinese Academy of Medical Sciences Oxford Institute, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7BN, UK

<sup>11</sup>Data Science, Population Analytics, Janssen R&D, Spring House, PA 19002, USA

<sup>12</sup>Data Science, External Innovation, Janssen R&D, London W1G 0BG, UK

<sup>13</sup>National Institute for Health Research, Comprehensive Biomedical Research Centre, Oxford OX4 2PG, UK

<sup>14</sup>These authors contributed equally

<sup>15</sup>Lead contact

\*Correspondence: [julian.knight@well.ox.ac.uk](mailto:julian.knight@well.ox.ac.uk)

<https://doi.org/10.1016/j.xgen.2023.100306>

## SUMMARY

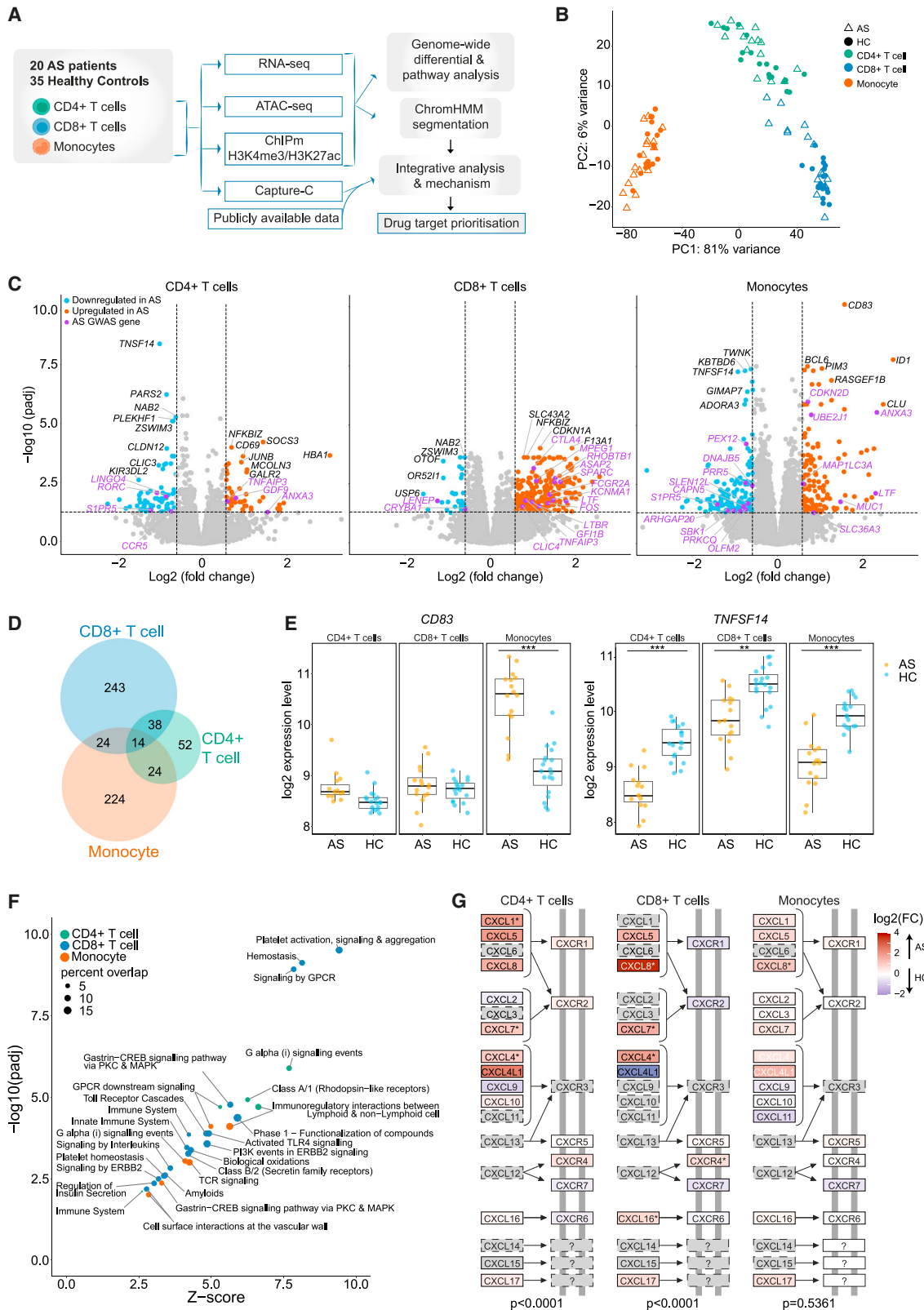
Ankylosing spondylitis (AS) is a common, highly heritable inflammatory arthritis characterized by enthesitis of the spine and sacroiliac joints. Genome-wide association studies (GWASs) have revealed more than 100 genetic associations whose functional effects remain largely unresolved. Here, we present a comprehensive transcriptomic and epigenomic map of disease-relevant blood immune cell subsets from AS patients and healthy controls. We find that, while CD14<sup>+</sup> monocytes and CD4<sup>+</sup> and CD8<sup>+</sup> T cells show disease-specific differences at the RNA level, epigenomic differences are only apparent upon multi-omics integration. The latter reveals enrichment at disease-associated loci in monocytes. We link putative functional SNPs to genes using high-resolution Capture-C at 10 loci, including *PTGER4* and *ETS1*, and show how disease-specific functional genomic data can be integrated with GWASs to enhance therapeutic target discovery. This study combines epigenetic and transcriptional analysis with GWASs to identify disease-relevant cell types and gene regulation of likely pathogenic relevance and prioritize drug targets.

## INTRODUCTION

Ankylosing spondylitis (AS) is a common inflammatory arthritis characterized by inflammation of the sacroiliac joints and spinal entheses, which causes extensive new bone formation and vertebral fusion, resulting in pain, loss of movement, and disability.<sup>1,2</sup> Combined with other systemic manifestations of the disease, such as inflammation of the gut, skin, and eyes, this leads to significant morbidity and disease burden.<sup>3,4</sup> Twin and other family studies indicate that AS is highly heritable ( $\lambda$

~50) with broad-sense heritability greater than 90%.<sup>5,6</sup> This involves strong association with the major histocompatibility complex (MHC) allele HLA-B27<sup>7,8</sup> and more than 100 other loci identified through genome-wide association studies (GWASs).<sup>9–13</sup> Several of these associations implicate genes involved in interleukin-23 (IL-23)-driven inflammation and Th17 responses; these include *IL23R* (encoding the IL-23 receptor), *IL6R* (IL-6 receptor), *TYK2* (tyrosine kinase 2 receptor), and *IL27R*.<sup>9</sup> In a few cases, such as *ERAP1* (endoplasmic reticulum aminopeptidase 1) and *IL23R*, functional non-synonymous single-nucleotide





(legend on next page)

polymorphisms (SNPs) have been described.<sup>9,13,14</sup> However, most AS associations involve non-coding SNPs, which may be regulatory in nature and act in a cell-specific manner to modulate a variety of epigenetic, transcriptional, and post-transcriptional mechanisms.<sup>15–18</sup> Recently we demonstrated how AS-associated SNPs at *RUNX3* modulate the binding of transcription factors (TFs) and regulatory complexes in T cells and monocytes,<sup>15–18</sup> but for other associated loci, the causal genes and pathways remain largely unresolved.<sup>19</sup> The expression and coordination of regulatory mechanisms for genes involved in the disease pathophysiology are likely to be cell type specific. Here we focus on CD4<sup>+</sup> and CD8<sup>+</sup> T cells and monocytes from patients with active AS and healthy controls (HC) because these cell types have been implicated previously in the pathogenesis of AS,<sup>20–26</sup> and previous studies have largely sampled whole peripheral blood mononuclear cells (PBMCs).<sup>27–30</sup>

The outlook for patients with more severe forms of AS has been greatly improved in recent years by the introduction of new biologic treatments inhibiting the inflammatory cytokines tumor necrosis factor alpha (TNF- $\alpha$ ) and IL-17A. Nevertheless, fewer than half are likely to achieve sustained remission even with these targeted therapies,<sup>31</sup> highlighting the need for patient stratification of potential responders and new therapeutic targets in AS. Human genetic evidence supporting the identification of therapeutic targets strongly increases the likelihood of success in late-stage clinical trials.<sup>32</sup> We and others have shown that GWASs, in combination with functional genomic evidence and knowledge of network connectivity, can be used to prioritize target genes and pathways through, for example, the priority index (Pi) algorithm.<sup>33,34</sup> Logistical and technical challenges have limited the number of studies generating omics data from patient samples to date, and it remains unresolved how best to maximize the value of such data through integration across assay modalities, including genetics.<sup>35</sup>

Here, we present a comprehensive map of the epigenomic landscape of AS defining the global transcriptome, chromatin accessibility, and enhancer- and promoter-associated histone modifications in disease-relevant subsets of immune cells from patients and HCs. We identify global changes in chromatin architecture in the AS disease state in monocytes and characterize specific GWAS loci to identify interactions between lead SNPs in enhancers and cognate genes, including prostaglandin E receptor 4 (*PTGER4*) and ETS proto-oncogene 1 (*ETS1*). Furthermore, we show how functional genomic evidence can be

integrated with GWAS data through Pi to identify candidate therapeutic targets for future study.

## Design

### Patient and control cohorts and experimental overview

To generate a comprehensive functional genomic and epigenomic atlas of the immune response in peripheral blood of AS patients, we recruited 20 adult patients with active disease who were naive to biologic therapy and fulfilled diagnostic criteria for AS and 35 HCs recruited locally or from the Oxford Biobank (Table S1; Table S2; STAR Methods). Cell populations of interest (CD4<sup>+</sup> and CD8<sup>+</sup> T cells and CD14<sup>+</sup> monocytes) were freshly isolated from peripheral blood by positive selection using immunomagnetic cell separation with more than 98% purity (Figure S1; STAR Methods). Each cell type was then processed immediately for total RNA sequencing (RNA-seq), chromatin accessibility (assay for transposase-accessible chromatin with next-generation sequencing [ATAC-seq]), informative histone modifications for promoter (histone H3 lysine 4 trimethylation [H3K4me3]) and enhancer (histone H3 lysine 4 trimethylation [H3K27ac]) activity (ChIPmentation), and high-resolution chromosomal conformation capture (Capture-C) (Figure 1A; Table S2; STAR Methods).

## RESULTS

### Differential gene expression in active AS is cell type specific

We first investigated the nature of differential gene expression between AS patients and HCs for specific immune cell types. We focused on three major immune cell types previously implicated in AS (CD4<sup>+</sup> and CD8<sup>+</sup> T cells and CD14<sup>+</sup> monocytes).<sup>20–26</sup> Analysis of gene expression by RNA-seq comparing the three cell types for each individual showed that gene expression segregated by cell type more strongly than by disease state (Figures 1B and S1B). However, for each cell type, we found hundreds of differentially expressed genes between AS patients and HC (CD4<sup>+</sup> T cells, 122 genes; CD8<sup>+</sup> T cells, 299 genes; monocytes, 300 genes; padj (adjusted p value) < 0.05, fold change [FC] > 1.5) (STAR Methods; Figures 1C and S1C; Table S3). The majority of differentially expressed genes were cell type specific (Figure 1D), and where differentially expressed genes involved more than one cell type, the direction of effect was the same in the majority of cases. For example, *CD83* (encoding

**Figure 1. Gene expression levels consistently delineate cell type and show differences between AS patients and HCs**

(A) Workflow of the study.

(B) PCA of RNA-seq data in CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, and monocytes from AS patients and HCs.

(C) Volcano plot showing differentially expressed genes calculated using DESeq2 (padj < 0.05, FC > 1.5) between AS patients and HCs in CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, and monocytes. Genes in AS-associated GWAS regions are purple. Red genes are upregulated and blue downregulated in AS patients. CD4<sup>+</sup> T cells had 122 differentially expressed genes, CD8<sup>+</sup> T cells 299 genes, and monocytes 300 genes.

(D) Cell type specificity of differentially expressed genes; numbers of differentially expressed genes are given.

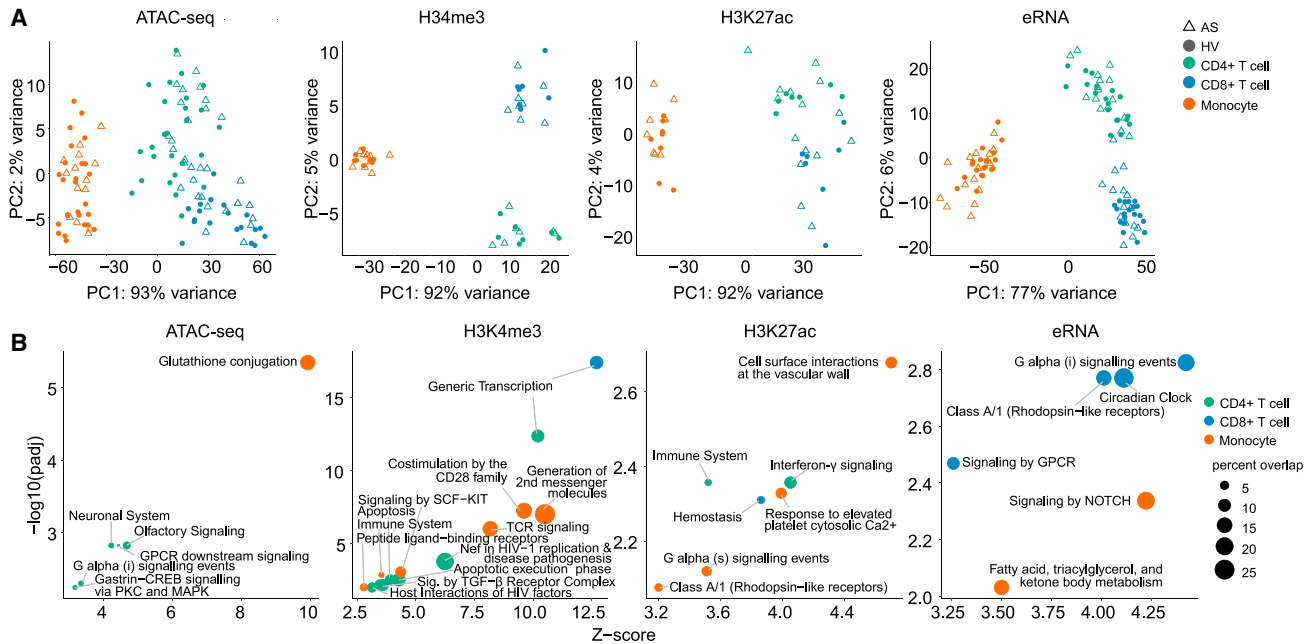
(E) Examples of differential gene expression at *CD83* and *TNFSF14*; \*\*padj < 0.01, \*\*\*padj < 10<sup>-7</sup> (from DESeq2).

(F) Enriched pathways in the Reactome database (FDR < 0.01 from XGR output) from significant differentially expressed genes in each cell type. Dot size represents percentages of genes represented in that pathway, and colors represent cell types.

(G) CXC subfamily of the Kyoto Encyclopedia of Genes and Genomes (KEGG) “cytokine-cytokine reception interaction” pathway colored by gene expression log<sub>2</sub> FC. Significantly differentially expressed genes are marked by asterisks. The p value of CXC family subset over-representation is shown below for each cell type, calculated by chi-squared test with Yates’ correction.

See also Figure S1 and Table S3.





**Figure 2. Individual epigenomic mapping methods in immune cell subsets have limited capacity to differentiate AS patients and HCs**  
(A) PCA of genome-wide distribution of ATAC-seq, ChIPm H3K4me3, ChIPm H3K27ac, and eRNA peaks in CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, and monocytes from AS patients and HCs.  
(B) Enriched Reactome pathways (FDR < 0.01 from XGR output) within genes associated with the top 200 differentially expressed peaks in each cell type. Note that each modality is plotted with different x and y scales to maximize clarity. Dot size represents percentages of genes represented in that pathway, and colors represent cell types. Numbers of significant differential peaks ( $\text{padj} < 0.05$ ,  $\text{FC} > 1.5$ ) are shown at the top left for each modality, colored by cell type. See also [Figure S2](#) and [Tables S4–S7](#).

CD83, a cell-surface glycoprotein involved in regulation of antigen presentation) has significantly higher expression only in AS patient monocytes, *TNFSF14* (encoding TNF superfamily member 14) has significantly lower expression in AS patients in all three cell types ([Figure 1E](#)), and *SCAMP5* is an example of a gene upregulated in CD8<sup>+</sup> T cells and downregulated in CD4<sup>+</sup> T cells. We defined disease-enriched pathways from differentially expressed genes using exploring Genomic Relations (XGR)<sup>36</sup> ([STAR Methods](#)). All cell types showed enrichment of immune-related pathways and G protein-coupled receptor (GPCR) signaling pathways in AS ([Figure 1F](#)). We identified significant upregulation of the CXC subfamily of chemokine receptors in CD4<sup>+</sup> and CD8<sup>+</sup> T cells ([Figures 1G and S1D](#)), which links to the important role of IL-17-producing cells in AS pathogenesis. We investigated cell subset composition by deconvolution of RNA-seq data using CIBERSORTx<sup>37</sup> and found no difference in abundance of the three cell types or major cell subsets within these when comparing AS patients and HCs ([Figure S1E](#)).

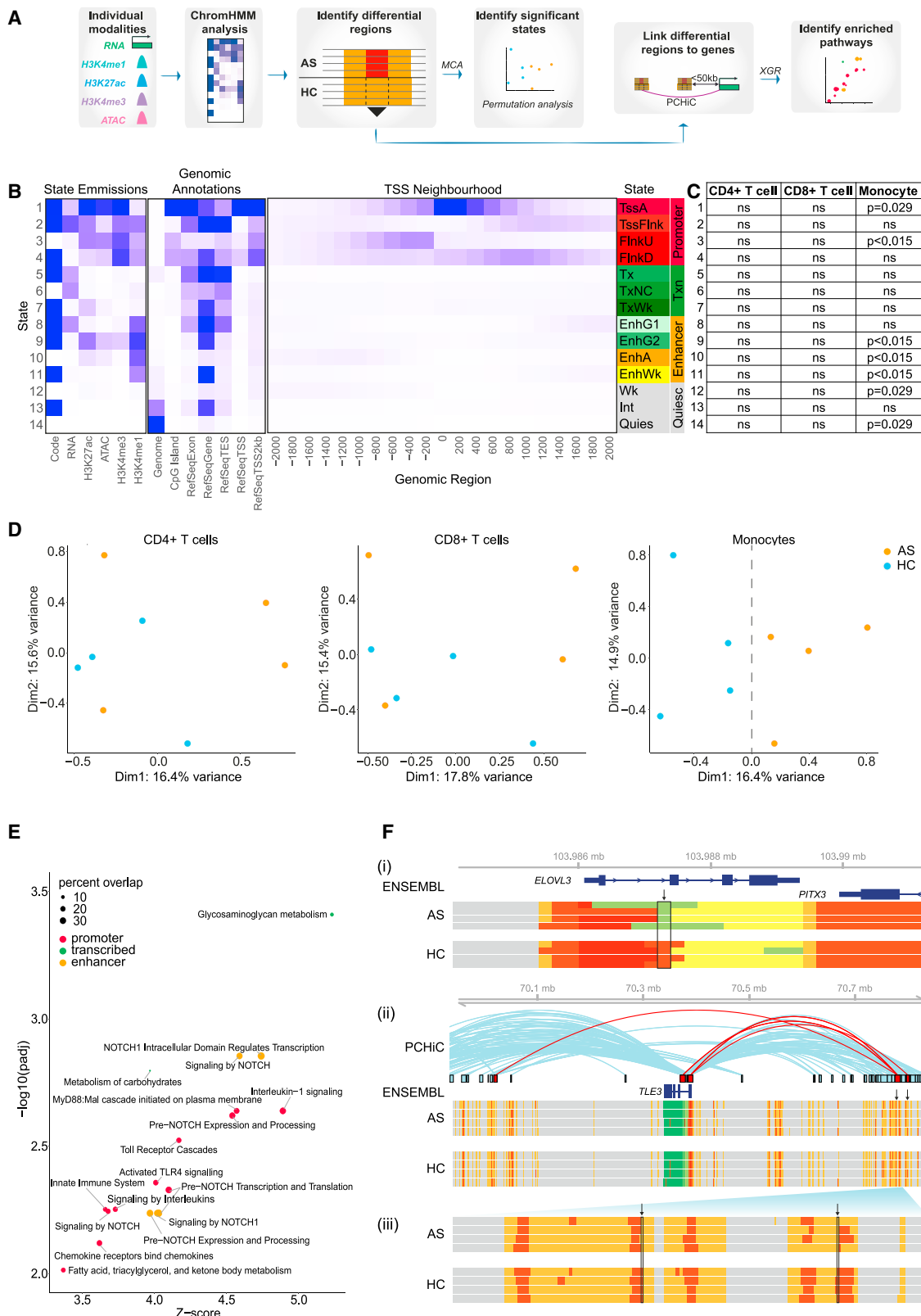
### Cell-type-specific epigenomic marks show limited differences between AS patients and controls

We then resolved genomic regulatory features in CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, and monocytes from AS patients. To do this, we assayed open chromatin with ATAC-seq and enrichment of H3K4me3 and H3K27ac histone modifications using ChIPmentation (ChIPm) and identified non-coding enhancer RNAs (eRNAs) within ATAC peaks outside of coding genes

([STAR Methods](#); [Figures S2A and S2B](#)). Each omics modality showed cell type specificity on principal-component analysis (PCA) that outweighed the effect of disease state ([Figure 2A](#)). Only a very small number of significant differential signals were observed between AS patients and HCs ([Figure 2B](#); [Tables S4–S6](#)), and disease state could not be clearly distinguished on PCA ([Figure S2C](#)). The transcription start site (TSS) score for ATAC and chromatin immunoprecipitation (ChIP) correlated with expression levels of their corresponding gene ([Figure S2D](#)). We assigned differential ATAC, ChIPm, and eRNA signals to genes by proximity or overlap with promoter capture Hi-C (PChIP) looping interactions identified in relevant cell types<sup>38</sup> ([STAR Methods](#)). Pathway enrichment analysis of genes linked to the top 200 differential regions between AS patients and HCs for each modality implicated immunological pathways and GPCR signaling along with transcriptional pathways and NOTCH signaling across cell types and modalities, consistent with our analysis of differentially expressed genes ([Figure 2B](#)).

### Disease-specific regulatory chromatin states are found in monocytes from patients with AS

We further investigated whether there were disease-specific differences in the chromatin landscape by maximizing the informativeness of different sources of omics information using ChromHMM<sup>39</sup> ([Figure 3A](#)), a machine learning algorithm that enables definition of chromatin conformational states based on combinations of regulatory features. The derived emission model consisted of 14



(legend on next page)

chromatin states (Figure 3B) in four functional categories: promoter (states 1–4), transcribed (states 5–7), enhancer (states 8–11), and quiescent (states 12–14). Multiple correspondence analysis (MCA) on global ChromHMM data revealed major differences between cell types, albeit with considerable variation between individuals (Figure S3A). Within each cell type, we identified thousands of 200-bp segments that were assigned to different chromatin states in AS patients and HCs (Table S7). We found that differential segments in two promoter states (1 and 3), three enhancer states (9, 10, and 11), and two quiescent states (12 and 14) were significantly over-represented in AS monocytes (Figure 3C), as determined by permutation analysis (Figure S3B). Consistent with this, MCA comparison of AS patients and HCs showed separation on dimension (dim) 1 in monocytes but not CD4<sup>+</sup> and CD8<sup>+</sup> T cells (Figure 3D). Within monocytes, all states except state 14 (quiescent) showed separation of AS patients and HCs on dim 1 or dim 2 (Figure S3C). In contrast, only three states showed separation in CD8<sup>+</sup> T cells (states 7, 8, and 11), and none in CD4<sup>+</sup> T cells. Comparative state transitions between patients and HCs revealed a complex pattern of differential states, including enrichment of enhancer state 10 (EnhA) in AS patients corresponding to promoter state 1 (TssA) in HCs (Figure S3D). Overall, this analysis demonstrates significant changes in the epigenomic landscape of active AS disease, notably in monocytes.

We next sought to determine which molecular pathways might be altered by these global changes in monocyte chromatin architecture. We assigned the differential ChromHMM fragments to genes based on proximity or PChI-C looping events<sup>38</sup> (STAR Methods; Table S7) and performed pathway enrichment analysis (Figure 3E). The highest number of enriched pathways contained genes linked to promoter state, followed by those linked to enhancer state. Only two pathways were enriched in genes linked to transcribed states and none with quiescent states. Six of 15 enriched pathways from promoter and enhancer states related to NOTCH signaling (Figure 3E), further implicating this pathway in monocytes in AS. To aid interpretation of these findings, we further defined by flow cytometry which monocyte subpopulations were represented in our sorted monocyte population, showing that these are 80% CD14<sup>+</sup> CD16<sup>+</sup> classical monocytes with the remainder intermediate and non-classical monocytes (Figure S3E).

To illustrate our results, we show two examples where differential ChromHMM segments correlate with alterations in gene expression (Figure 3F). *ELOVL3* (encoding elongation of very-long-chain fatty acids) is involved in fatty acid metabolism and downregulated in psoriasis,<sup>40</sup> a common extra-articular manifestation of AS.<sup>41</sup> The *ELOVL3* promoter is marked by state 1 (TssA, promoter-like) in AS patients and HCs, but this mark spans a shorter genomic interval in AS patients (Figure 3F). Consistent with this, expression of *ELOVL3* is lower in AS patients (padj = 0.0003). *TLE3* (encoding transducin-like enhancer family member 3) is a transcriptional co-repressor involved in the NOTCH signaling pathway. We identified up- and downstream enhancers that form looping interactions with the *TLE3* promoter (Figure 3F). The enhancers are defined by ChromHMM state 10 (enhancer) containing punctate regions of state 1 (promoter) that correspond with non-coding eRNA transcription. In two segments these promoter elements are narrower in AS patients compared with HC, which may contribute to the observed reduction in *TLE3* expression in AS patients (padj = 0.045).

### Regulatory chromatin signatures are enriched at AS GWAS regions

Having generated comprehensive epigenomic maps in three cell types from AS patients and HCs, we next addressed whether this could inform the functional basis of observed genetic associations in AS from GWASs, specifically seeking evidence to implicate/delineate the gene(s) responsible for the genetic association. We identified 35 differentially expressed genes located within GWAS regions (<500 kb from the lead SNP), of which 31 were cell type specific (Figure 1C; STAR Methods). We found that regulatory ChromHMM states are over-represented in or near AS GWAS regions, and this is independent of the HLA-B27 association (Figures 4A and S4). Enhancer state 10 is enriched near GWAS regions in all three cell types, while in monocytes, states 2, 6, 10, 11, 12, 13, and 14 are all enriched in GWAS regions. We also found significant enrichment of ATAC, H3K4me3, and H3K27ac signals and eRNAs at GWAS loci in all cell types (Figure 4B).

The chr3p21.31 locus illustrates the intersection between an AS genetic association (*rs1001007*<sup>10</sup>) and disease-specific chromatin state in monocytes (Figure 4C). This intergenic SNP is

### Figure 3. Disease-specific regulatory regions are found in monocytes and implicate NOTCH signaling in AS pathogenesis

(A) Workflow of the ChromHMM analysis.

(B) ChromHMM emission model showing 14 states annotated according to state emissions, genomic annotations, and TSS neighborhood. State names were assigned according to predicted function, aligned to The Encyclopedia of DNA Elements (ENCODE) labels where possible. Promoter states: TssA, active TSS; TssFlnk, flanking promoter; FlnkU, upstream of TSS; FlnkD, downstream of TSS. Transcribed (Txn) states: Tx, transcription; TxNC, non-coding transcription; TxWk, weak transcription. Enhancer states: EnhG1, strongly transcribed, weak enhancer; EnhG2, weakly transcribed, strong enhancer; EnhA, enhancer; EnhWk, weak intronic enhancer. Quiescent (Quiesc) states: Wk, low/weak enhancer; Int, introns; Quies, quiescent.

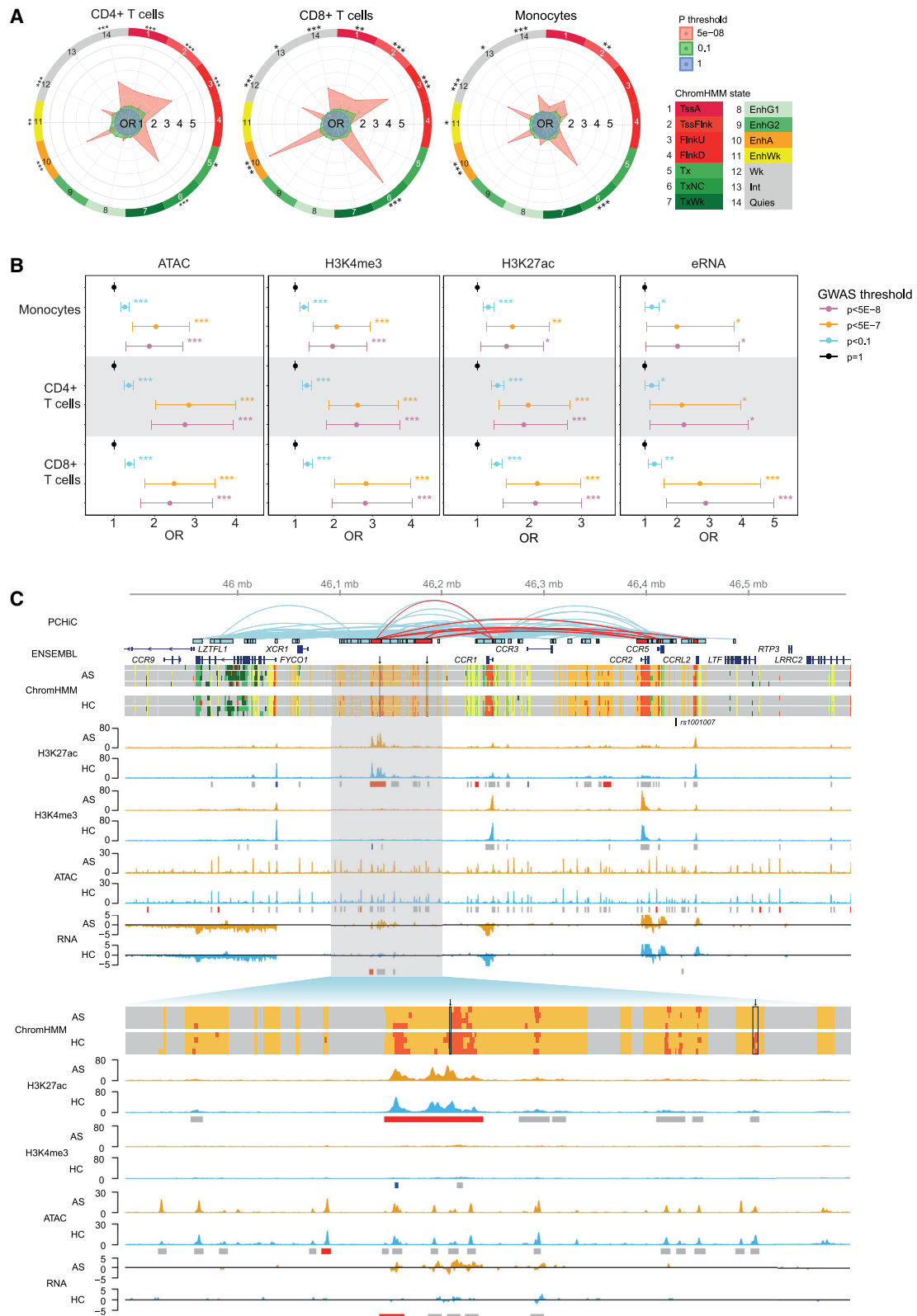
(C) Significance of disease-specific ChromHMM segment frequency for each state in each cell type, calculated using permutation analysis (ns, not significant). Only monocytes contained significantly over-represented states.

(D) MCA showing overall distribution of ChromHMM states within each cell type comparing AS patients and HCs. The monocyte plot dotted line at x = 0 delineates AS and HC samples.

(E) Enriched pathways in the Reactome database (FDR < 0.01, XGR output) performed on genes linked to promoter-like states, transcribed regions, or enhancer-like states in monocytes. Dot size represents percentage of genes represented in that pathway.

(F) Visualization of differential ChromHMM regions (vertical arrows) at (i) the promoter of *ELOVL3* and (ii) enhancer of *TLE3*, with (iii) magnified region of the *TLE3* differential enhancer. ChromHMM states in AS patients and HCs are colored as in (B). PChI-C looping interactions<sup>38</sup> are shown, with loops intersecting differential ChromHMM segments in red.

See also Figure S3 and Table S8.



(legend on next page)

located between *CCR5* and *CCRL2*, but this locus has not been fine mapped to establish a functional variant and modulated gene in AS. Consequently, we interrogated the whole monocyte-specific topologically associating domain (TAD) and discovered several regions with disease-specific chromatin signatures. We found evidence of an intergenic region between *XCR1* and *CCR1* that has an enhancer-like profile (ChromHMM state 10, EnhA) in AS patients with a promoter-like profile (ChromHMM state 3, FlnkU) in HCs, consistent with observed differences in eRNA and H3K27ac between cases and controls. Analysis of chromatin interaction data shows that this region is involved in DNA looping to *CCR1*, *CCR2*, *CCR5*, and *CCRL2* but not with *CCR3*. Looping events correlate with expression of the cognate genes in monocytes. These findings indicate complex regulation of the CCR gene cluster in monocytes with epigenetic changes specific to AS.

### Chromosome looping interactions from AS-associated SNPs in enhancers implicate disease-relevant genes

To further substantiate the relationship between disease-associated SNPs and likely modulated genes, we mapped chromosome looping events at GWAS loci in patient samples at high resolution (compared with PCHI-C). We performed Capture-C on CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, and monocytes from AS patients and controls (STAR Methods). Baits were designed at promoters of genes with known immune roles within AS GWAS regions. We found that 18 of 44 promoter viewpoints assayed demonstrated chromatin interactions in at least one cell type, although no differences were found between AS patients and HCs. Follow-up Capture-C experiments were performed with baits at AS-associated SNPs to demonstrate reciprocal interactions between SNPs and promoters. Overall, nine reciprocal interactions were found between promoters and regions containing GWAS SNPs, nine of which were marked as enhancers (state 10 or 11) in our ChromHMM analysis in the same cell type (Table 1; Figures 5 and S5–S14). These loci also contained differential ATAC, H3K4me3, H3K27ac, or eRNA peaks, and five overlapped expression quantitative trait loci (eQTLs) in the same cell types<sup>43</sup> (Figure 5A).

Two loci illustrate SNP-promoter interactions. *ETS1* encodes ETS proto-oncogene 1, a TF with numerous roles in immune cells, including regulation of cytokine and chemokine gene expression.<sup>46</sup> Our ChromHMM analysis showed enhancer regions flanking the *ETS1* gene, and Capture-C analysis showed looping events between the enhancer overlapping the AS-associated lead SNP *rs7933433*<sup>10</sup> and *ETS1* promoter in CD4<sup>+</sup> and CD8<sup>+</sup> T cells but not monocytes (Figures 5B and S11). *PTGER4* encodes prostaglandin receptor E4, a GPCR whose expression is associ-

ated with disease severity in AS.<sup>47</sup> We observed an interaction between the promoter of *PTGER4* and the enhancer overlapping *rs1992661* (AS GWAS lead SNP<sup>10</sup>) in all cell types (Figures 5C and S9). We detected an interaction between an enhancer encompassing the known functional SNP *rs9283753*<sup>48</sup> and the *PTGER4* promoter specifically in monocytes and only with the promoter bait (Figures 5C and S9). There was no detectable interaction between the lead SNP *rs12186979* and *PTGER4* gene in any of the cell types. These results support a model where associated SNPs lying in an enhancer region interact with a distal gene promoter via a chromatin looping event. This, combined with additional local disease-context-specific epigenomic modifications, may lead to alterations in cognate gene expression.

We were interested to explore whether the genotypes of individual SNPs were associated with alterations in chromatin structure. The study was underpowered to perform such an analysis genome wide, but we were able to analyze the effect of individual lead AS-associated SNPs. In doing so, we identified an association between *rs4672505* and ATAC-seq peak chr2:62559366–62561099, where the A risk allele correlates with a reduced ATAC-seq signal in CD8<sup>+</sup> T cells (Figure 5D). Analysis of next-generation sequencing (NGS) reads from the 16 heterozygous individuals with at least 5 mapping reads showed that 99.4% of ATAC peak reads encoded the G allele, strongly suggesting that the risk A allele prevents chromatin opening. This finding was not driven by mapping bias because A was the reference allele in the hg19 build used in this analysis. This SNP lies at chr2p15 and overlaps a putative enhancer that forms a looping interaction with *B3GNT2* identified by publicly available PCHI-C (Figure 5D). *rs4672505* is associated with AS, Crohn disease, and psoriasis<sup>10</sup> and is an eQTL for *B3GNT2*.<sup>45</sup>

### Use of disease-specific functional genomic datasets enhances drug target discovery in AS

The final aim of our study was to prioritize new therapeutic targets in AS. We previously developed Pi, a genetics-led approach that annotates GWASs with functional genomic data to prioritize therapeutic targets across a range of immune-mediated diseases.<sup>33,34</sup> We modified the underlying algorithm of Pi to include our new AS-specific functional genomic datasets (Figure 6A; STAR Methods), the algorithm previously having been limited to non-disease-context functional genomics data, and assessed whether this inclusion increased the power to identify potential therapeutic targets. In the original algorithm, we used disease-specific genetic associations to define seed (core) genes, including (1) nearby genes (nGene) using genomic proximity and organization, (2) expression-associated genes (eGene)

#### Figure 4. Differential chromatin regions are enriched at GWAS loci

(A) Enrichment of differential ChromHMM differential regions at AS-associated GWAS loci<sup>10</sup> with association p value thresholds as indicated in CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, and monocytes. \*\*\*p < 0.001, \*\*p < 0.01, \*p < 0.05. OR, odds ratio.

(B) Enrichment of ATAC, H3K4me3, H3K27ac, and eRNA peaks at AS-associated GWAS loci<sup>9</sup> in CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, and monocytes; GWAS association p value thresholds are indicated. Error bars represent 95% confidence interval of the OR (from GARFIELD<sup>42</sup>). \*\*\*p < 0.001, \*\*p < 0.01, \*p < 0.05.

(C) Visualization of multiple epigenomic datasets at chr3p21 (top, chr3:45890000–46600000; bottom, chr3:46090000–46200000). PCHI looping events<sup>38</sup> that intersect regions of differential chromatin are indicated in red. Selected gene transcripts from the Ensembl database are shown. ChromHMM is shown for 4 AS patients and 4 HCs with colors as in Figure 3B. The lead AS-associated SNP *rs1001007*<sup>10</sup> is shown. Representative ATAC, H3K4me3, and H3K27ac tracks (RPKM) and total RNA tracks (log2 count) are shown for AS patients (orange) and HCs (blue), with called peaks marked by gray boxes and differential peaks shown in red (upregulated in AS) or blue (downregulated in AS).

See also Figure S4.



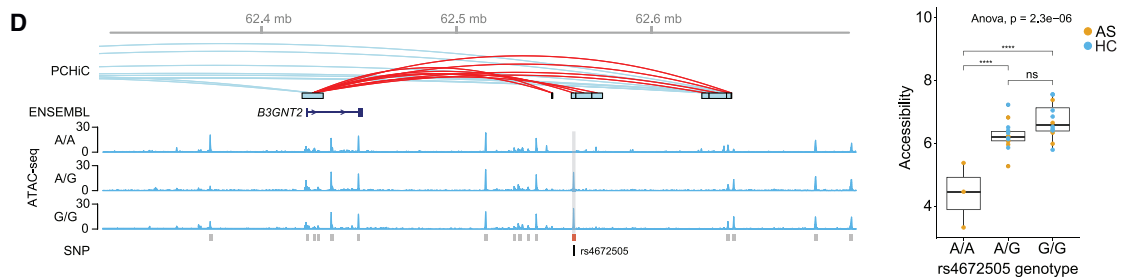
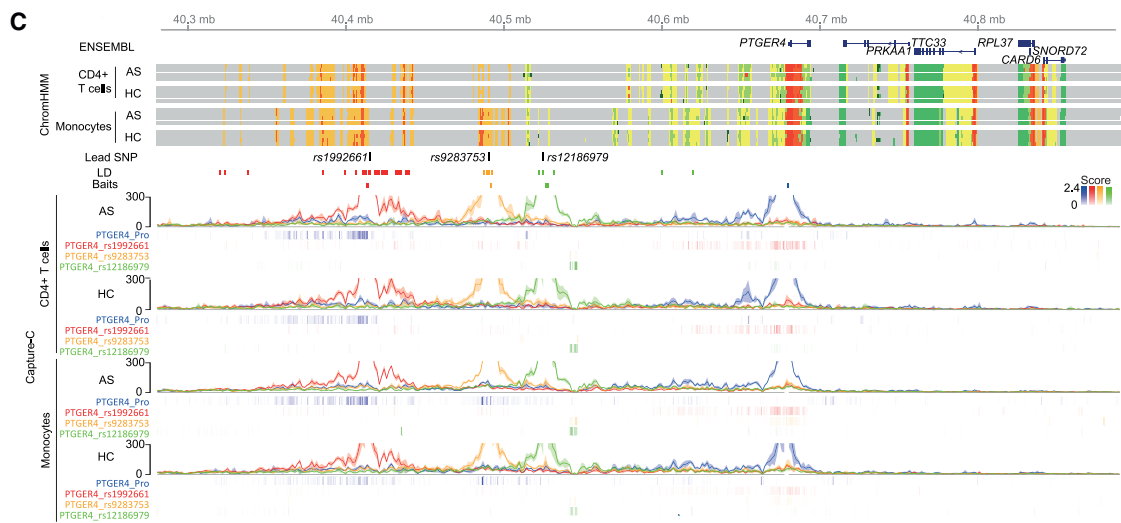
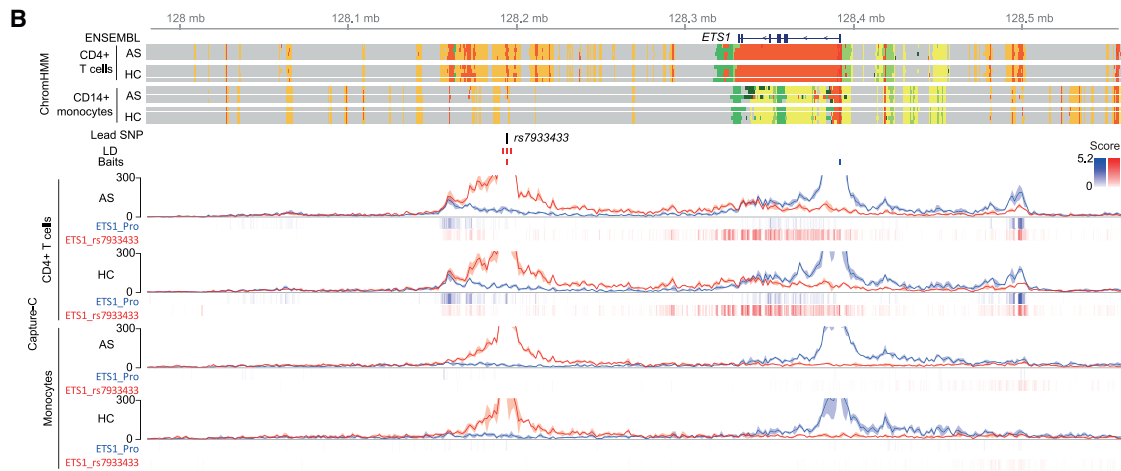
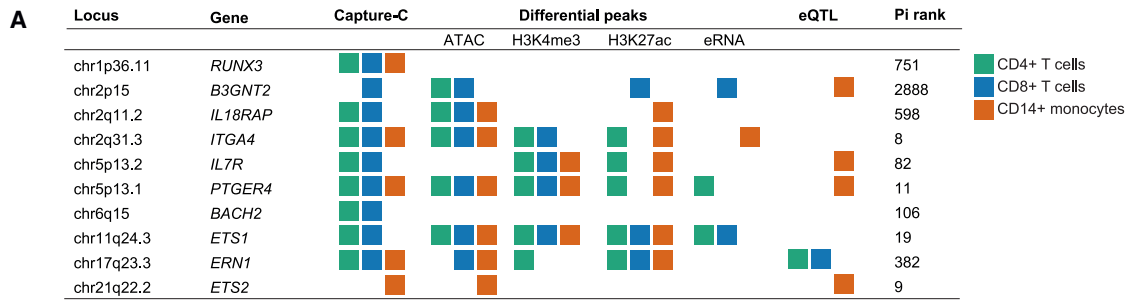
**Table 1. Evidence of enhancer-gene interactions at GWAS loci**

Locus	Lead SNPs	Interacting gene	Gene function	Gene-SNP distance (bp)	Other evidence	Putative mechanism	Figure
chr1p36.11	<i>rs6600247</i> <sup>a,b</sup>	<i>RUNX3</i> (RUNX family transcription factor 3)	TF in T cell differentiation	13,612	functional evidence of disrupted TF binding <sup>15–18</sup>	SNPs alter TF binding; local looping interactions	<a href="#">Figure S5</a>
chr2q11.2	<i>rs4851529</i> <sup>a,b</sup>	<i>IL18RAP</i> (IL-18 receptor accessory protein)	component of IL-18 receptor, binds pro-inflammatory cytokine	387,952	–	long-range enhancer	<a href="#">Figure S6</a>
chr2q31.3	<i>rs12615545</i> <sup>a</sup>	<i>ITGA4</i> (integrin subunit alpha 4)	integrin component, role in cell motility and migration	273,165	long non-coding RNA in CD4/CD8	long-range enhancer	<a href="#">Figure S7</a>
chr5p13.2	<i>rs11742270</i> <sup>a</sup>	<i>IL7R</i> (IL-7 receptor)	component of IL-7 receptor, binds pro-inflammatory cytokine	1,738	regulatory and splice variant SNPs <sup>20</sup> eRNA present	two signals: regulatory SNP; splicing SNP controls soluble IL-7R production	<a href="#">Figure S8</a>
chr5p13.1	<i>rs12186979</i> <sup>a</sup> , <i>rs1992661</i> <sup>b</sup>	<i>PTGER4</i> (prostaglandin E receptor 4)	prostaglandin receptor, role in IL-23 and TNF pathways	155,171	functional SNP <sup>44</sup> ; eRNAs present; differential ChromHMM at promoter	long-range enhancer	<a href="#">Figures 5 and S9</a>
chr6q15	<i>rs17765610</i> <sup>a</sup>	<i>BACH2</i> (BTB domain And CNC homolog 2)	TF, T cell regulation	0	–	long-range enhancer within gene	<a href="#">Figure S10</a>
chr11q24.3	<i>rs7933433</i> <sup>b</sup>	<i>ETS1</i> (ETS proto-oncogene 1)	TF, regulates cytokines and chemokines	134,204	eRNA present	long-range enhancer	<a href="#">Figures 5 and S11</a>
chr17q23.3	<i>rs196941</i> <sup>b</sup>	<i>ERN1</i> (endoplasmic reticulum to nucleus signaling 1)	unfolded protein response	0	–	intronic enhancer	<a href="#">Figure S12</a>
chr21q22.2	<i>rs2836883</i> <sup>a</sup> , <i>rs9977672</i> <sup>b</sup>	<i>ETS2</i> (ETS proto-oncogene 2)	TF, T and B cell regulation	269,866	eRNA present	long-range enhancer	<a href="#">Figure S13</a>

All SNPs except *rs1992661* were used as baits in the Capture-C experiment. See also Figures 5 and S5–S14.

<sup>a</sup>Lead SNP for each locus from Cortes et al.<sup>9</sup>

<sup>b</sup>Lead SNP for each locus from Ellinghaus et al.<sup>10</sup>



(legend on next page)

integrating eQTL datasets, and (3) conformation genes (cGene) using PCHI-C datasets. Here, we added five types of AS-specific functional genomic predictors using data from this study (denoted RNA, eRNA, H3K27ac, H3K4me3, and ATAC). The AS-specific expression predictor (RNA) was generated based on differential gene expression, and AS-specific epigenomic predictors (ATAC, H3K4me3, H3K27ac, and eRNA) were prepared on the basis of differential peaks linked to genes as above.

We benchmarked the performance of the Pi algorithm with and without the AS-specific functional genomic predictors to prioritize currently approved drug targets for AS versus simulated negative targets (STAR Methods). This showed that inclusion of disease-specific data improved the predictive power (area under the curve [AUC] = 0.869) compared with the original prediction (AUC = 0.822) and the state-of-the-art approach (Open Targets,<sup>49</sup> including text mining [AUC = 0.761] and genetic associations [AUC = 0.582] from the Open Targets Genetics Portal<sup>50</sup>) (Figure 6B). As expected, combined use of predictors performed much better than each predictor alone (Figures 6B and S15A). Among the top 1% of prioritized genes (of >17,000 ranked genes) were the known AS drug targets *IL23R*, *JAK2*, and *TYK2* (Figure S15B). Pathway enrichment analysis of the top 1% of genes identified pathogenic AS pathways, such as Janus kinase-signal transducer and activator of transcription (JAK-STAT) signaling (false discovery rate [FDR] =  $3.3 \times 10^{-32}$ ), TNF signaling (FDR =  $1 \times 10^{-16}$ ) and Th17 cell differentiation (FDR =  $3.1 \times 10^{-24}$ ) (Figure 6C). The significantly enriched pathways also included cytokine-cytokine receptor interaction, Toll-like receptor signaling, ErbB signaling, chemokine signaling, and T cell receptor (TCR) signaling, which were consistent with our findings from the RNA-seq data (Figure S2C). Pathway crosstalk analysis for potential therapeutic intervention identified a network of 53 interconnecting genes (Figure 6D), including genes that are already therapeutic targets in AS and other autoimmune conditions, such as *JAK1* and *TYK2*, alongside yet unexplored genes. The list of genes within this network (Figure 6E) are highly prioritized as candidates for therapeutic intervention, providing an input for the drug development pipelines and further study.<sup>51</sup>

## DISCUSSION

### Summary of findings

We performed transcriptomic and epigenomic profiling in specific primary immune cell populations isolated from carefully

phenotyped patients with active AS. This revealed disease-specific changes in the chromatin landscape and differential regulatory signatures significantly enriched near AS-associated loci. Pathway analysis indicated the importance of NOTCH and chemokine receptor signaling in AS in addition to known disease-associated pathways. Capture-C identified physical interactions between associated SNPs lying in enhancer regions and nearby gene promoters. Taken together, these results provide a map of the epigenetic landscape in AS and evidence of the mechanisms by which genetic associations can alter immune cell function in AS. Earlier studies have compared gene expression between AS patients and HCs in PBMCs.<sup>27–30</sup> By looking at individual cell types, in this study we demonstrated a key role of monocytes in the pathogenesis of AS. This work highlights the importance of looking for functional effects of SNPs in the appropriate cellular and disease contexts in AS and other immune-mediated conditions.

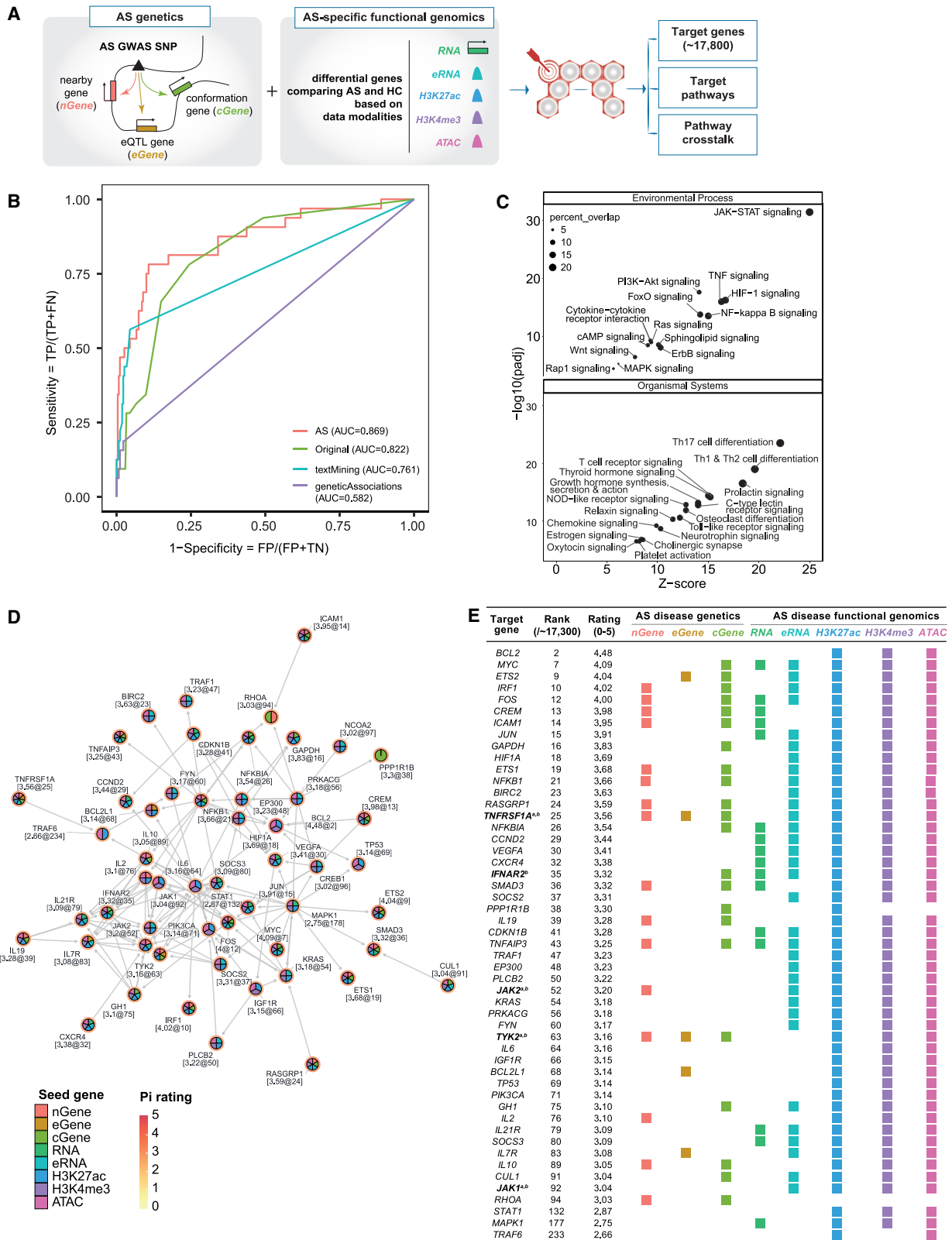
### Identification of regulatory disease-associated SNPs and cognate genes

When analyzing the effects of a genetic variant with putative enhancer-modifying activity, a key component is to understand which gene(s) are regulated by that enhancer. Here, we described 10 enhancers that overlap AS-associated SNPs and that interact with a gene promoter within the same TAD (Table 1; Figures 5 and S5–S14). These enhancers share common features such as eRNA expression, activating histone marks, and open chromatin marks.

We discovered enhancer-gene interactions with two members of the ETS proto-oncogene (ETS) family of TFs, *ETS1* and *ETS2* (Figures 5, S11, and S13). At the *ETS1* locus, we observed chromatin looping interactions with dual enhancers, up- and downstream of the gene, specifically in T cells, of which the downstream enhancer overlaps an AS-associated lead SNP (*rs7933433*<sup>10</sup>). A monocyte-specific interaction was observed between an enhancer and the *ETS2* promoter, which overlaps the 99% credible set comprising 5 AS-associated SNPs<sup>52</sup> that are also eQTLs in monocytes.<sup>43</sup> *ETS1* and *ETS2* enhancers exhibit non-coding eRNA transcription. *ETS1* and *ETS2* are TFs expressed across various immune cell types and have roles including regulation of T cell subset differentiation.<sup>53</sup> *ETS1* regulates the expression of *IL-7R*<sup>54</sup> (encoding the *IL-7R $\alpha$*  subunit), which is also associated with AS,<sup>20</sup> and *RUNX3*, which encodes Runt-related TF 3, a TF involved in T cell function with known

### Figure 5. Chromosome looping interactions link genes with genetically associated SNPs at enhancers

(A) Summary of epigenomic evidence at 10 GWAS loci where enhancer-gene interactions are observed. Colored squares indicate the presence of Capture-C interactions, differential ATAC, H3K4me3, H3K27ac, eRNA peaks, or eQTLs<sup>45</sup> in each cell type. Pi rank shows priority ranking among 17,000 genes (Figure 6). (B) Visualization of the *ETS1* locus (chr11:127980000–128560000). (C) Visualization of the *PTGER4* locus (chr5:40280000–40890000). The following data are shown for CD4<sup>+</sup> T cells and monocytes. Ensembl: selected transcripts of Ensembl genes. ChromHMM: ChromHMM data from four AS patients and four HC are shown with colors as in Figure 3B. Lead SNP: Position of lead GWAS SNPs from International Genetics of Ankylosing Spondylitis Consortium (IGAS) et al.<sup>9</sup> and/or Ellinghaus et al.<sup>10</sup> as per Table 1. LD: SNPs in LD ( $r^2 > 0.8$ ) with lead SNP. Baits: position of Capture-C baits (see also Table S9). Capture-C: mean interaction count ( $n = 3$ ) and 1 SD shading, with PeakY scores colored according to bait for AS patients and HCs in CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, and monocytes. (D) Genetic association with chromatin openness at *B3GNT2*. Left: visualization of the *B3GNT2* locus (chr2:62320000–62700000) in CD8<sup>+</sup> T cells. PCHI-C: PCHI-C looping events.<sup>38</sup> ENSEMBL: *B3GNT2* gene. ATAC-seq: representative ATAC tracks for AS patients with each *rs4672505* genotype. Called peaks are shown in gray with the differential peak chr2:62559366–62561099 in red. SNP: position of *rs4672505*. Right: association of ATAC-seq peak chr2:62559366–62561099 limma-corrected counts with the *rs4672505* genotype calculated by ANOVA with Tukey post-test. \*\*\*\*p < 0.0001. See also Figures S5–S14 and Table S9.



(legend on next page)

AS-associated functional variants.<sup>15–18</sup> *ETS1*, *ETS2*, and *IL7R* are in the Pi network output (Figure 6D), indicating that they form part of an important functional pathway with strong possibility for therapeutic intervention.

Gene-SNP interactions were also observed at the *PTGER4* locus (Figures 5 and S9). *PTGER4* is widely expressed throughout the immune system and is involved in the IL-23 and TNF- $\alpha$  pathways.<sup>55</sup> It is also expressed in osteoclasts and could potentially have a role in new bone formation in AS.<sup>56</sup> GWASs have found two independent associations at the *PTGER4* locus.<sup>9</sup> Tewhey et al.<sup>44</sup> described a functional SNP (*rs9283753*) that lies in an enhancer and alters *PTGER4* expression in lymphoblastoid cell lines. Our data show a chromatin looping interaction between AS-associated SNP *rs1992661* and the *PTGER4* promoter in all three cell types and a T cell-specific interaction between *rs9283753* and the *PTGER4* promoter. Our ChromHMM analysis shows that these SNPs overlap enhancer marks in the same cell types. Taken together, these data strongly support a functional role of *rs1992661* and *rs9283753* in regulation of *PTGER4*.

We found allele-specific differences in the ATAC-seq signal for an AS GWAS-associated SNP, *rs4672505*. This variant has been associated previously with differential abundance of *B3GNT2*, a poly-N-acetyllactosamine synthase in whole blood.<sup>45</sup> *B3GNT2* is upregulated in T cells on activation, and a recent CRISPR screen showed evidence that this enzyme is important in modulating T cell activation in the setting of cancer.<sup>57</sup> Our findings suggest that reduced expression of *B3GNT2* in individuals with the AS risk allele is likely caused by reduced chromatin openness at this locus. This may impact higher-order chromatin structures, such as the looping event identified between a distal enhancer and *B3GNT2* (Figure 5). The JASPAR 2022<sup>58</sup> database of TF binding profiles predicts that STAT1 binds to this site, a TF that plays an important role in transcriptional activation in the immune system.<sup>59</sup> Further work, such as genomic editing and functional assays, will be needed to identify the function of these or other unknown TFs at this locus.

### Cytokine and NOTCH signaling pathways

GPCR-related and cytokine signaling pathways were consistently enriched across modalities and cell types. In particular, expression of CXC cytokine subfamily genes was upregulated in CD4<sup>+</sup> and CD8<sup>+</sup> T cells, consistent with the inflammatory environment of AS. Chemokine levels have been shown previously to be increased at the gene and protein level in AS<sup>60</sup> and psoriatic arthritis, a related spondyloarthropathy.<sup>61,62</sup> This may contribute to the differentiation of pro-inflammatory Th1 and Th17 cells,

which are expanded in AS, or trafficking of leukocytes and osteoclast precursors to the sites of inflammation at the joint.<sup>63</sup>

NOTCH signaling has a wide range of functions in the innate and adaptive immune systems.<sup>64</sup> We found that NOTCH signaling was linked to changes in chromatin signatures in monocytes. NOTCH signaling has been shown to be important for monocyte-macrophage differentiation, with increased NOTCH signaling favoring inflammatory M1 macrophage development in atherosclerosis, systemic lupus erythematosus, and cancer.<sup>65</sup> NOTCH signaling has already been implicated in inflammatory states<sup>66</sup> and rheumatoid arthritis<sup>67</sup> and may be important for regulating monocyte differentiation to osteoclasts,<sup>68</sup> thus influencing the pathogenic ossification that is a key feature of late-stage AS.<sup>69</sup> Wang et al.<sup>70</sup> showed that *NOTCH1* expression is reduced in AS patients who have had anti-TNF biologic therapy. Further study of this pathway in AS is warranted and could lead to repurposing of existing NOTCH pathway inhibitors in AS.<sup>71</sup>

### Identification of novel therapeutic targets

We showed how disease-context-specific functional genomic data could be used to identify novel therapeutic targets in AS. We modified Pi, a previously published algorithm designed to identify the network of therapeutic targets in autoimmune disease from GWAS. The multi-omics approach used here supported the importance of known candidate pathways, such as Th17/IL-23 and TNF, and identifies new pathways and potential drug targets, including *PTGER4*, ErbB, phosphatidylinositol 3-kinase (PI3K)/AKT, NOTCH, and GPCR (Figure 6). Existing inhibitors of these pathways could be repurposed in AS, such as PI3K inhibitors used in lymphoma treatment.<sup>72</sup> Our results show exciting promise for development of new therapeutics in AS. The exact roles these TFs and other network genes play in AS remain to be elucidated and will be investigated in the future using, for example, genomic editing and functional assays.

### Conclusions

We demonstrated that the epigenomic landscape of immune cells is altered in AS. We used these results, together with evidence of chromosomal interactions, to inform the interpretation of GWASs for AS in terms of likely functional variants and modulated genes and prioritize potential drug targets and networks. This is important because existing therapies are only effective in a subset of AS patients and ultimately do not cure the disease.

### Limitations of the study

This study used disease-relevant immune cell subsets isolated from PBMCs of AS patients and HCs. One limitation of our study

### Figure 6. Prioritization of new drug targets in AS

(A) Overview of the modified Pi prioritization algorithm.

(B) Inclusion of disease-specific datasets (AS) improves the Pi algorithm output relative to the original Pi algorithm (Original) and outperformance relative to Open Targets textMining and geneticAssociations. AUC, area under the curve.

(C) Prioritized target pathways (FDR < 0.05) based on KEGG pathway enrichment analysis using the top 1% prioritized genes. Dot size represents percentages of genes represented in that pathway.

(D) Identification of pathway crosstalk; that is, a network of highly prioritized and interconnecting genes defined by KEGG interaction data, with nodes segmented by predictor/evidence types.

(E) Table summarizing the prioritization and evidence information associated with pathway crosstalk genes. Colored boxes underneath each predictor type represent the datasets in support. Existing therapeutic targets are indicated for <sup>a</sup>AS and <sup>b</sup>other autoimmune conditions.

See also Figure S15.



was the use of bulk CD4<sup>+</sup> T cell, CD8<sup>+</sup> T cell, and CD14<sup>+</sup> monocyte populations rather than investigating smaller subsets of these cell types. This was necessary because of the large cell numbers required to perform a panel of multi-omics experiments on the same sample, in particular mapping of chromosomal interactions. We showed that major cell subsets were present at expected frequencies and did not differ between AS patients and HCs (Figure S1), although disease-specific cell type frequencies have been reported for minor cell subtypes.<sup>20,25,26,73</sup> Future work sampling cells from sites of inflammation, such as sacroiliac joints, and utilizing single-cell-based methods will further unravel the cell types and pathogenic mechanisms of AS.<sup>74</sup> For individual modalities, signals from differential analysis were modest, especially in CD4<sup>+</sup> T cells, which could be due to the heterogeneity of this cell type. Context specificity of regulatory regions is key, so in addition to direct *ex vivo* analysis of cells from patients with the active disease state, analysis of such cells subjected to immune challenges *in vitro* (such as lipopolysaccharide stimulation of monocytes, anti-CD3/28 stimulation of T cells) may identify additional functional SNPs specifically in those activation states. The study was underpowered to perform expression and chromatin quantitative trait mapping on a genome-wide scale, and this is an important area for future work in a disease context. We were unable to study the effect of other covariates, including sex and drug regimens (although all patients were biologic therapy naive), because of the small sample size. We have shown previously that eRNAs have a role in innate immune activation.<sup>75</sup> The observed widespread bidirectional eRNA expression at genomic enhancers can be further investigated using more sensitive methods (such as global run-on sequencing, small capped RNA sequencing, and precision run-on sequencing<sup>76</sup>) to detect more subtle alterations in eRNA expression and specifically identify the role of SNPs therein. Findings from ChromHMM were limited by small sample size. We focused on the presence of activating chromatin modifications (chromatin accessibility, promoter- and enhancer-associated histone modifications), so future studies could investigate the role of repressive histone marks such as H3K9me3 and H3K27me3. Future studies will be required to further characterize the genes and pathways highlighted by this study to assess their effect at the protein level and on cellular phenotype and function; for example, through genome editing and small-molecule inhibitors.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human subjects
- METHOD DETAILS
  - Isolation of immune cell populations

- Flow cytometry
- RNA-seq
- ATAC-seq
- ChIPmentation
- Capture-C
- Genotyping

- QUANTIFICATION AND STATISTICAL ANALYSIS

- Analysis of flow cytometry data
- Downstream analysis of NGS data
- Computational deconvolution of bulk RNA-seq data
- ChromHMM
- Pathway enrichment analysis
- GWAS enrichment
- Capture-C NGS data analysis
- Genetic analysis and imputation
- Therapeutic target prioritisation
- Data visualisation

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.xgen.2023.100306>.

## ACKNOWLEDGMENTS

This research was supported by Versus Arthritis (program grant 20773 to J.C.K.), Janssen Oxford Translational fellowships (to J.C.K.), Wellcome grant 106130/Z/14/ZR (to Prof. Doug Higgs), and a Wellcome Trust grant (090532/Z/09/Z) to core facilities at the Wellcome Center for Human Genetics. Computation was supported by Wellcome Trust Core Award grant 203141/Z/16/Z with funding from the NIHR Oxford BRC. The study received support from the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC) (to J.C.K., B.P.W., and P.B.). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health. J.C.K. is supported by a Wellcome Trust Investigator Award (204969/Z/16/Z). K.S.E. and A.C.B. are supported by Janssen Oxford Translational Fellowships. C.J.C. was supported by funds from NDORMS, University of Oxford, and Oxford University Bridging Salary Support. C.D. is supported by Versus Arthritis PhD studentship 22198. F.C. was supported by research grants from FOREUM Foundation for Research in Rheumatology, Société Française de Radiologie, Fondation pour la Recherche Médicale (FRM SPE201803005058), and Fondation Bettencourt Schueller. A.L.L. was supported by Psoriasis UK PhD studentship 200314. A.S. was supported by Arthritis Research UK PhD studentship MP/20331. H.F. is supported by the National Natural Science Foundation of China (32170663) and Shanghai Pujiang Program (21PJ1409600). P.Z. is supported by the Chinese Academy of Medical Sciences (CAMS) Innovation 537 Fund for Medical Science (2018-I2M-2-002). P.B. has received research funding from AstraZeneca, Regeneron, and Novartis. The authors thank the patients and healthy donors and clinical staff involved in recruitment and sample collection. This research was funded in part by the Wellcome Trust (grant numbers 106130/Z/14/ZR, 090532/Z/09/Z, 203141/Z/16/Z, 204969/Z/16/Z). For the purpose of Open Access, the author has applied a CC BY public copyright licence to any author accepted manuscript version arising from this submission.

## AUTHOR CONTRIBUTIONS

Conceptualization, P.B., J.C.K., and B.P.W.; methodology, A.C.B., H.F., A.L.L., G.M., J.A.O., and A.S.; software, A.C.B., C.J.C., F.C., C.D., K.S.E., H.F., A.L.L., A.C.M., O.M., G.M., and P.Z.; formal analysis, A.C.B., C.J.C., F.C., O.M., and G.M.; investigation, A.A., A.C.B., C.J.C., C.D., A.L.L., A.C.M., O.M., J.A.O., A.S., and G.S.; resources, P.B., A.C.B., J.C.K., G.M., and B.P.W.; data curation, A.C.B. and G.M.; writing – original draft, C.J.C.;

writing – review & editing, P.B., A.C.B., M.H.B., C.J.C., F.C., H.F., T.H., J.C.K., S.L., J.M., O.M., D.T., M.V., and B.P.W. (all authors reviewed and approved the final version); visualization, A.C.B., C.J.C., F.C., H.F., O.M., G.M., and P.Z.; supervision, P.B., C.J.C., J.C.K., M.V., and B.P.W.; project administration, J.C.K. and G.S.; funding acquisition, P.B., C.J.C., J.C.K., and B.P.W.

#### DECLARATION OF INTERESTS

J.M., T.H., D.T., S.L., and M.H.B. are full-time employees of Janssen R&D and own shares of stock in Johnson & Johnson. K.S.E. and A.C.B. are supported by Janssen Oxford Translational Fellowships.

#### INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: April 29, 2022

Revised: January 30, 2023

Accepted: March 27, 2023

Published: April 24, 2023

#### REFERENCES

- Taurog, J.D., Chhabra, A., and Colbert, R.A. (2016). Ankylosing spondylitis and axial spondyloarthritis. *N. Engl. J. Med.* *374*, 2563–2574. <https://doi.org/10.1056/NEJMra1406182>.
- Bridgewood, C., Sharif, K., Sherlock, J., Watad, A., and McGonagle, D. (2020). Interleukin-23 pathway at the enthesis: the emerging story of enthesitis in spondyloarthropathy. *Immunol. Rev.* *294*, 27–47. <https://doi.org/10.1111/imr.12840>.
- Stolwijk, C., van Tubergen, A., Castillo-Ortiz, J.D., and Boonen, A. (2015). Prevalence of extra-articular manifestations in patients with ankylosing spondylitis: a systematic review and meta-analysis. *Ann. Rheum. Dis.* *74*, 65–73. <https://doi.org/10.1136/annrheumdis-2013-203582>.
- Rizzo, A., Ferrante, A., Guggino, G., and Ciccia, F. (2017). Gut inflammation in spondyloarthritis. *Best Pract. Res. Clin. Rheumatol.* *31*, 863–876. <https://doi.org/10.1016/j.berh.2018.08.012>.
- Carter, N., Williamson, L., Kennedy, L.G., Brown, M.A., and Wordsworth, B.P. (2000). Susceptibility to ankylosing spondylitis. *Rheumatology* *39*, 445. <https://doi.org/10.1093/rheumatology/39.4.445>.
- Brown, M.A., Kennedy, L.G., MacGregor, A.J., Darke, C., Duncan, E., Shatford, J.L., Taylor, A., Calin, A., and Wordsworth, P. (1997). Susceptibility to ankylosing spondylitis in twins: the role of genes, HLA, and the environment. *Arthritis Rheum.* *40*, 1823–1828. <https://doi.org/10.1002/art.1780401015>.
- Brewerton, D.A., Hart, F.D., Nicholls, A., Caffrey, M., James, D.C., and Sturrock, R.D. (1973). Ankylosing SPONDYLITIS and HL-a 27. *Lancet* *1*, 904–907. [https://doi.org/10.1016/S0140-6736\(73\)91360-3](https://doi.org/10.1016/S0140-6736(73)91360-3).
- Schlosstein, L., Terasaki, P.I., Bluestone, R., and Pearson, C.M. (1973). High association of an HL-A antigen, W27, with ankylosing spondylitis. *N. Engl. J. Med.* *288*, 704–706. <https://doi.org/10.1056/nejm197304052881403>.
- International Genetics of Ankylosing Spondylitis Consortium IGAS; Cortes, A., Hadler, J., Karaderi, T., Harris, J., Pointon, J.P., Cremin, K., Leo, P., Pryce, K., Robinson, P.C., et al. (2013). Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nat. Genet.* *45*, 730–738.
- Ellinghaus, D., Jostins, L., Spain, S.L., Cortes, A., Bethune, J., Han, B., Park, Y.R., Raychaudhuri, S., Pouget, J.G., Hübenthal, M., et al. (2016). Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. *Nat. Genet.* *48*, 510–518. <https://doi.org/10.1038/ng.3528>.
- Evans, D.M., Spencer, C.C.A., Pointon, J.J., Su, Z., Harvey, D., Kochan, G., Oppermann, U., Dilthey, A., Pirinen, M., Stone, M.A., et al. (2011). Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat. Genet.* *43*, 761–767. <https://doi.org/10.1038/ng.873>.
- Australo-Anglo-American Spondyloarthritis Consortium TASC; Reveille, J.D., Sims, A.M., Danoy, P., Evans, D.M., Leo, P., Pointon, J.J., Jin, R., Bradbury, L.A., Zhou, X., et al. (2010). Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. *Nat. Genet.* *42*, 123–127. <https://doi.org/10.1038/ng.513>.
- Wellcome Trust Case Control Consortium (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* *447*, 661–678. <https://doi.org/10.1038/nature05911>.
- Karaderi, T., Harvey, D., Farrar, C., Appleton, L.H., Stone, M.A., Sturrock, R.D., Brown, M.A., Wordsworth, P., and Pointon, J.J. (2009). Association between the interleukin 23 receptor and ankylosing spondylitis is confirmed by a new UK case-control study and meta-analysis of published series. *Rheumatology* *48*, 386–389. <https://doi.org/10.1093/rheumatology/ken501>.
- Cohen, C.J., Davidson, C., Selmi, C., Bowness, P., Knight, J.C., Wordsworth, B.P., and Vecellio, M. (2021). Disruption of c-MYC binding and chromosomal looping involving genetic variants associated with ankylosing spondylitis upstream of the RUNX3 promoter. *Front. Genet.* *12*, 741867. <https://doi.org/10.3389/fgene.2021.741867>.
- Vecellio, M., Chen, L., Cohen, C.J., Cortes, A., Li, Y., Bonham, S., Selmi, C., Brown, M.A., Fischer, R., Knight, J.C., et al. (2021). Functional genomic analysis of a RUNX3 polymorphism associated with ankylosing spondylitis. *Arthritis Rheumatol.* *73*, 980–990. <https://doi.org/10.1002/art.41628>.
- Vecellio, M., Cortes, A., Roberts, A.R., Ellis, J., Cohen, C.J., Knight, J.C., Brown, M.A., Bowness, P., and Wordsworth, B.P. (2018). Evidence for a second ankylosing spondylitis-associated RUNX3 regulatory polymorphism. *RMD Open* *4*, e000628. <https://doi.org/10.1136/rmdopen-2017-000628>.
- Vecellio, M., Roberts, A.R., Cohen, C.J., Cortes, A., Knight, J.C., Bowness, P., and Wordsworth, B.P. (2016). The genetic association of RUNX3 with ankylosing spondylitis can be explained by allele-specific effects on IRF4 recruitment that alter gene expression. *Ann. Rheum. Dis.* *75*, 1534–1540. <https://doi.org/10.1136/annrheumdis-2015-207490>.
- Nancy, Z., Yan, L., Hui, S., Paul, B., and Liye, C. (2021). From the genetics of ankylosing spondylitis to new biology and drug target discovery. *Front. Immunol.* *12*, 624632. <https://doi.org/10.3389/fimmu.2021.624632>.
- Al-Mossawi, H., Yager, N., Taylor, C.A., Lau, E., Danielli, S., de Wit, J., Gilchrist, J., Nassiri, I., Mahe, E.A., Lee, W., et al. (2019). Context-specific regulation of surface and soluble IL7R expression by an autoimmune risk allele. *Nat. Commun.* *10*, 4575. <https://doi.org/10.1038/s41467-019-12393-1>.
- Raine, T., Brown, D., Bowness, P., Hill Gaston, J.S., Moffett, A., Trowsdale, J., and Allen, R.L. (2006). Consistent patterns of expression of HLA class I free heavy chains in healthy individuals and raised expression in spondyloarthropathy patients point to physiological and pathological roles. *Rheumatology* *45*, 1338–1344. <https://doi.org/10.1093/rheumatology/kel305>.
- Conrad, K., Wu, P., Sieper, J., and Syrbe, U. (2015). In vivo pre-activation of monocytes in patients with axial spondyloarthritis. *Arthritis Res. Ther.* *17*, 179. <https://doi.org/10.1186/s13075-015-0694-2>.
- Taurog, J.D., Dorris, M.L., Satumtira, N., Tran, T.M., Sharma, R., Dressel, R., van den Brandt, J., and Reichardt, H.M. (2009). Spondylarthritis in HLA-B27/human beta2-microglobulin-transgenic rats is not prevented by lack of CD8. *Arthritis Rheum.* *60*, 1977–1984. <https://doi.org/10.1002/art.24599>.
- Shi, H., Chen, L., Ridley, A., Zaarour, N., Brough, I., Caucci, C., Smith, J.E., and Bowness, P. (2020). GM-CSF primes proinflammatory monocyte responses in ankylosing spondylitis. *Front. Immunol.* *11*, 1520. <https://doi.org/10.3389/fimmu.2020.01520>.
- Ciccia, F., Guggino, G., Zeng, M., Thomas, R., Ranganathan, V., Rahman, A., Alessandro, R., Rizzo, A., Saieva, L., Macaluso, F., et al.

- (2018). Proinflammatory CX3CR1+CD59+Tumor Necrosis factor-like molecule 1A+Interleukin-23+ monocytes are expanded in patients with ankylosing spondylitis and modulate innate lymphoid cell 3 immune functions. *Arthritis Rheumatol.* 70, 2003–2013. <https://doi.org/10.1002/art.40582>.
26. Mauro, D., Simone, D., Bucci, L., and Ciccia, F. (2021). Novel immune cell phenotypes in spondyloarthritis pathogenesis. *Semin. Immunopathol.* 43, 265–277. <https://doi.org/10.1007/s00281-021-00837-0>.
27. Duan, R., Leo, P., Bradbury, L., Brown, M.A., and Thomas, G. (2010). Gene expression profiling reveals a downregulation in immune-associated genes in patients with AS. *Ann. Rheum. Dis.* 69, 1724–1729. <https://doi.org/10.1136/ard.2009.111690>.
28. Assassi, S., Reveille, J.D., Arnett, F.C., Weisman, M.H., Ward, M.M., Agarwal, S.K., Gourh, P., Bhula, J., Sharif, R., Sampat, K., et al. (2011). Whole-blood gene expression profiling in ankylosing spondylitis shows upregulation of toll-like receptor 4 and 5. *J. Rheumatol.* 38, 87–98. <https://doi.org/10.3899/jrheum.100469>.
29. Pimentel-Santos, F.M., Ligeiro, D., Matos, M., Mourão, A.F., Costa, J., Santos, H., Barcelos, A., Godinho, F., Pinto, P., Cruz, M., et al. (2011). Whole blood transcriptional profiling in ankylosing spondylitis identifies novel candidate genes that might contribute to the inflammatory and tissue-destructive disease aspects. *Arthritis Res. Ther.* 13, R57. <https://doi.org/10.1186/ar3309>.
30. Li, Z., Haynes, K., Pennisi, D.J., Anderson, L.K., Song, X., Thomas, G.P., Kenna, T., Leo, P., and Brown, M.A. (2017). Epigenetic and gene expression analysis of ankylosing spondylitis-associated loci implicate immune cells and the gut in the disease pathogenesis. *Gene Immun.* 18, 135–143. <https://doi.org/10.1038/gene.2017.11>.
31. Macfarlane, G.J., Pathan, E., Jones, G.T., and Dean, L.E. (2020). Predicting response to anti-TNF $\alpha$  therapy among patients with axial spondyloarthritis (axSpA): results from BSRBR-AS. *Rheumatology* 59, 2481–2490. <https://doi.org/10.1093/rheumatology/kez657>.
32. Nelson, M.R., Tipney, H., Painter, J.L., Shen, J., Nicoletti, P., Shen, Y., Floratos, A., Sham, P.C., Li, M.J., Wang, J., et al. (2015). The support of human genetic evidence for approved drug indications. *Nat. Genet.* 47, 856–860. <https://doi.org/10.1038/ng.3314>.
33. Fang, H., ULTRA-DD Consortium; De Wolf, H., Knezevic, B., Burnham, K.L., Osgood, J., Sanniti, A., Lledó Lara, A., Kasela, S., De Cesco, S., et al. (2019). A genetics-led approach defines the drug target landscape of 30 immune-related traits. *Nat. Genet.* 51, 1082–1091. <https://doi.org/10.1038/s41588-019-0456-1>.
34. Fang, H., and Knight, J.C. (2022). Priority index: database of genetic targets in immune-mediated disease. *Nucleic Acids Res.* 50, D1358–D1367. <https://doi.org/10.1093/nar/gkab994>.
35. Davenport, E.E., Amariuta, T., Gutierrez-Arcelus, M., Slowikowski, K., Westra, H.-J., Luo, Y., Shen, C., Rao, D.A., Zhang, Y., Pearson, S., et al. (2018). Discovering in vivo cytokine-eQTL interactions from a lupus clinical trial. *Genome Biol.* 19, 168. <https://doi.org/10.1186/s13059-018-1560-8>.
36. Fang, H., Knezevic, B., Burnham, K.L., and Knight, J.C. (2016). XGR software for enhanced interpretation of genomic summary data, illustrated by application to immunological traits. *Genome Med.* 8, 129. <https://doi.org/10.1186/s13073-016-0384-y>.
37. Newman, A.M., Steen, C.B., Liu, C.L., Gentles, A.J., Chaudhuri, A.A., Scherer, F., Khodadoust, M.S., Esfahani, M.S., Luca, B.A., Steiner, D., et al. (2019). Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* 37, 773–782. <https://doi.org/10.1038/s41587-019-0114-2>.
38. Javierre, B.M., Burren, O.S., Wilder, S.P., Kreuzhuber, R., Hill, S.M., Sewitz, S., Cairns, J., Wingett, S.W., Várnai, C., Thiecke, M.J., et al. (2016). Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. *Cell* 167, 1369–1384.e19. <https://doi.org/10.1016/j.cell.2016.09.037>.
39. Ernst, J., and Kellis, M. (2017). Chromatin-state discovery and genome annotation with ChromHMM. *Nat. Protoc.* 12, 2478–2492. <https://doi.org/10.1038/nprot.2017.124>.
40. Yu, Z., Gong, Y., Cui, L., Hu, Y., Zhou, Q., Chen, Z., Yu, Y., Chen, Y., Xu, P., Zhang, X., et al. (2020). High-throughput transcriptome and pathogenesis analysis of clinical psoriasis. *J. Dermatol. Sci.* 98, 109–118. <https://doi.org/10.1016/j.jdermsci.2020.03.006>.
41. Bengtsson, K., Forsblad-d’Elia, H., Deminger, A., Klingberg, E., Dehlin, M., Exarchou, S., Lindström, U., Askling, J., and Jacobsson, L.T.H. (2021). Incidence of extra-articular manifestations in ankylosing spondylitis, psoriatic arthritis and undifferentiated spondyloarthritis: results from a national register-based cohort study. *Rheumatology* 60, 2725–2734. <https://doi.org/10.1093/rheumatology/keaa692>.
42. Iotchkova, V., Ritchie, G.R.S., Geihs, M., Morganello, S., Min, J.L., Walter, K., Timpson, N.J., UK10K Consortium; Birney, E., Dunham, I., and Soranzo, N. (2019). GARFIELD classifies disease-relevant genomic features through integration of functional annotations with association signals. *Nat. Genet.* 51, 343–353. <https://doi.org/10.1038/s41588-018-0322-6>.
43. Kerimov, N., Hayhurst, J.D., Peikova, K., Manning, J.R., Walter, P., Kolberg, L., Samovica, M., Sakthivel, M.P., Kuzmin, I., Trevanion, S.J., et al. (2021). A compendium of uniformly processed human gene expression and splicing quantitative trait loci. *Nat. Genet.* 53, 1290–1299. <https://doi.org/10.1038/s41588-021-00924-w>.
44. Tewhey, R., Kotliar, D., Park, D.S., Liu, B., Winnicki, S., Reilly, S.K., Andersen, K.G., Mikkelsen, T.S., Lander, E.S., Schaffner, S.F., and Sabeti, P.C. (2016). Direct identification of hundreds of expression-modulating variants using a multiplexed reporter assay. *Cell* 165, 1519–1529. <https://doi.org/10.1016/j.cell.2016.04.027>.
45. GTEx Consortium (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* 369, 1318–1330. <https://doi.org/10.1126/science.aaz1776>.
46. Russell, L., and Garrett-Sinha, L.A. (2010). Transcription factor Ets-1 in cytokine and chemokine gene regulation. *Cytokine* 51, 217–226. <https://doi.org/10.1016/j.cyto.2010.03.006>.
47. Klasen, C., Meyer, A., Wittekind, P.S., Waqué, I., Nabhani, S., and Kofler, D.M. (2019). Prostaglandin receptor EP4 expression by Th17 cells is associated with high disease activity in ankylosing spondylitis. *Arthritis Res. Ther.* 21, 159. <https://doi.org/10.1186/s13075-019-1948-1>.
48. Libioulle, C., Louis, E., Hansoul, S., Sandor, C., Farnir, F., Franchimont, D., Vermeire, S., Dewit, O., de Vos, M., Dixon, A., et al. (2007). Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4. *PLoS Genet.* 3, e58. <https://doi.org/10.1371/journal.pgen.0030058>.
49. Ochoa, D., Hercules, A., Carmona, M., Suveges, D., Gonzalez-Uriarte, A., Malangone, C., Miranda, A., Fumis, L., Carvalho-Silva, D., Spitzer, M., et al. (2021). Open Targets Platform: supporting systematic drug-target identification and prioritisation. *Nucleic Acids Res.* 49, D1302–d1310. <https://doi.org/10.1093/nar/gkaa1027>.
50. Ghossaini, M., Mountjoy, E., Carmona, M., Peat, G., Schmidt, E.M., Hercules, A., Fumis, L., Miranda, A., Carvalho-Silva, D., Buniello, A., et al. (2021). Open Targets Genetics: systematic identification of trait-associated genes using large-scale genetics and functional genomics. *Nucleic Acids Res.* 49, D1311–d1320. <https://doi.org/10.1093/nar/gkaa840>.
51. Fang, H., Chen, L., and Knight, J.C. (2020). From genome-wide association studies to rational drug target prioritisation in inflammatory arthritis. *The Lancet Rheumatology* 2, e50–e62. [https://doi.org/10.1016/S2665-9913\(19\)30134-1](https://doi.org/10.1016/S2665-9913(19)30134-1).
52. van de Bunt, M., Cortes, A., IGAS Consortium; Brown, M.A., Morris, A.P., and McCarthy, M.I. (2015). Evaluating the performance of fine-mapping strategies at common variant GWAS loci. *PLoS Genet.* 11, e1005535. <https://doi.org/10.1371/journal.pgen.1005535>.

53. Garrett-Sinha, L.A. (2013). Review of Ets1 structure, function, and roles in immunity. *Cell. Mol. Life Sci.* *70*, 3375–3390. <https://doi.org/10.1007/s00018-012-1243-7>.
54. Grenningloh, R., Tai, T.S., Frahm, N., Hongo, T.C., Chicoine, A.T., Brander, C., Kaufmann, D.E., and Ho, I.C. (2011). Ets-1 maintains IL-7 receptor expression in peripheral T cells. *J. Immunol.* *186*, 969–976. <https://doi.org/10.4049/jimmunol.1002099>.
55. Yokoyama, U., Iwatsubo, K., Umemura, M., Fujita, T., and Ishikawa, Y. (2013). The prostanoid EP4 receptor and its signaling pathway. *Pharmacol. Rev.* *65*, 1010–1052. <https://doi.org/10.1124/pr.112.007195>.
56. Haroon, N. (2015). Ankylosis in ankylosing spondylitis: current concepts. *Clin. Rheumatol.* *34*, 1003–1007. <https://doi.org/10.1007/s10067-015-2956-4>.
57. Joung, J., Kirchgatterer, P.C., Singh, A., Cho, J.H., Nety, S.P., Larson, R.C., Macrae, R.K., Deasy, R., Tseng, Y.Y., Maus, M.V., and Zhang, F. (2022). CRISPR activation screen identifies BCL-2 proteins and B3GNT2 as drivers of cancer resistance to T cell-mediated cytotoxicity. *Nat. Commun.* *13*, 1606. <https://doi.org/10.1038/s41467-022-29205-8>.
58. Castro-Mondragon, J.A., Riudavets-Puig, R., Rauluseviciute, I., Lemma, R.B., Turchi, L., Blanc-Mathieu, R., Lucas, J., Boddie, P., Khan, A., Manosalva Pérez, N., et al. (2022). Jaspas 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* *50*, D165–d173. <https://doi.org/10.1093/nar/gkab1113>.
59. Villarino, A.V., Kanno, Y., and O’Shea, J.J. (2017). Mechanisms and consequences of Jak–STAT signaling in the immune system. *Nat. Immunol.* *18*, 374–384. <https://doi.org/10.1038/ni.3691>.
60. Wang, J., Zhao, Q., Wang, G., Yang, C., Xu, Y., Li, Y., and Yang, P. (2016). Circulating levels of Th1 and Th2 chemokines in patients with ankylosing spondylitis. *Cytokine* *81*, 10–14. <https://doi.org/10.1016/j.cyto.2016.01.012>.
61. Muntyanu, A., Abji, F., Liang, K., Pollock, R.A., Chandran, V., and Gladman, D.D. (2016). Differential gene and protein expression of chemokines and cytokines in synovial fluid of patients with arthritis. *Arthritis Res. Ther.* *18*, 296. <https://doi.org/10.1186/s13075-016-1196-6>.
62. Affandi, A.J., Silva-Cardoso, S.C., Garcia, S., Leijten, E.F.A., van Kempen, T.S., Marut, W., van Roon, J.A.G., and Radstake, T.R.D.J. (2018). CXCL4 is a novel inducer of human Th17 cells and correlates with IL-17 and IL-22 in psoriatic arthritis. *Eur. J. Immunol.* *48*, 522–531. <https://doi.org/10.1002/eji.201747195>.
63. Simone, D., Al Mossawi, M.H., and Bowness, P. (2018). Progress in our understanding of the pathogenesis of ankylosing spondylitis. *Rheumatology* *57*, vi4–vi9. <https://doi.org/10.1093/rheumatology/key001>.
64. Radtke, F., MacDonald, H.R., and Tacchini-Cottier, F. (2013). Regulation of innate and adaptive immunity by Notch. *Nat. Rev. Immunol.* *13*, 427–437. <https://doi.org/10.1038/nri3445>.
65. Zhou, D., Huang, C., Lin, Z., Zhan, S., Kong, L., Fang, C., and Li, J. (2014). Macrophage polarization and function with emphasis on the evolving roles of coordinated regulation of cellular signaling pathways. *Cell. Signal.* *26*, 192–197. <https://doi.org/10.1016/j.cellsig.2013.11.004>.
66. Šučur, A., Filipović, M., Flegar, D., Kelava, T., Šisl, D., Lukač, N., Kovacic, N., and Grčević, D. (2020). Notch receptors and ligands in inflammatory arthritis – a systematic review. *Immunol. Lett.* *223*, 106–114. <https://doi.org/10.1016/j.imlet.2020.04.010>.
67. Wei, K., Korsunsky, I., Marshall, J.L., Gao, A., Watts, G.F.M., Major, T., Croft, A.P., Watts, J., Blazar, P.E., Lange, J.K., et al. (2020). Notch signaling drives synovial fibroblast identity and arthritis pathology. *Nature* *582*, 259–264. <https://doi.org/10.1038/s41586-020-2222-z>.
68. Sekine, C., Koyanagi, A., Koyama, N., Hozumi, K., Chiba, S., and Yagita, H. (2012). Differential regulation of osteoclastogenesis by Notch2/Delta-like 1 and Notch1/Jagged1 axes. *Arthritis Res. Ther.* *14*, R45. <https://doi.org/10.1186/ar3758>.
69. Kusuda, M., Haroon, N., and Nakamura, A. (2022). Complexity of enthesitis and new bone formation in ankylosing spondylitis: current understanding of the immunopathology and therapeutic approaches. *Mod. Rheumatol.* *32*, 484–492. <https://doi.org/10.1093/mr/roab057>.
70. Wang, X.B., Ellis, J.J., Pennisi, D.J., Song, X., Batra, J., Hollis, K., Bradbury, L.A., Li, Z., Kenna, T.J., and Brown, M.A. (2017). Transcriptome analysis of ankylosing spondylitis patients before and after TNF- $\alpha$  inhibitor therapy reveals the pathways affected. *Gene Immun.* *18*, 184–190. <https://doi.org/10.1038/gene.2017.19>.
71. Christopoulos, P.F., Gjolberg, T.T., Krüger, S., Haraldsen, G., Andersen, J.T., and Sundlisæter, E. (2021). Targeting the notch signaling pathway in chronic inflammatory diseases. *Front. Immunol.* *12*, 668207. <https://doi.org/10.3389/fimmu.2021.668207>.
72. Sapon-Cousineau, V., Sapon-Cousineau, S., and Assouline, S. (2020). PI3K inhibitors and their role as novel agents for targeted therapy in lymphoma. *Curr. Treat. Options Oncol.* *21*, 51. <https://doi.org/10.1007/s11864-020-00746-8>.
73. Shen, H., Goodall, J.C., and Hill Gaston, J.S. (2009). Frequency and phenotype of peripheral blood Th17 cells in ankylosing spondylitis and rheumatoid arthritis. *Arthritis Rheum.* *60*, 1647–1656. <https://doi.org/10.1002/art.24568>.
74. Simone, D., Penkava, F., Ridley, A., Sansom, S., Al-Mossawi, M.H., and Bowness, P. (2021). Single cell analysis of spondyloarthritis regulatory T cells identifies distinct synovial gene expression patterns and clonal fates. *Commun. Biol.* *4*, 1395. <https://doi.org/10.1038/s42003-021-02931-3>.
75. Zhang, P., Amarasinghe, H.E., Whalley, J.P., Tay, C., Fang, H., Migliorini, G., Brown, A.C., Allcock, A., Scozzafava, G., Rath, P., et al. (2022). Epigenomic analysis reveals a dynamic and context-specific macrophage enhancer landscape associated with innate immune activation and tolerance. *Genome Biol.* *23*, 136. <https://doi.org/10.1186/s13059-022-02702-1>.
76. Field, A., and Adelman, K. (2020). Evaluating enhancer function and transcription. *Annu. Rev. Biochem.* *89*, 213–234. <https://doi.org/10.1146/annurev-biochem-011420-095916>.
77. Ahern, D.J., Ai, Z., Ainsworth, M., Allan, C., Allcock, A., Angus, B., Ansari, M.A., Arancibia-Cárcamo, C.V., Aschenbrenner, D., Attar, M., et al. (2022). A blood atlas of COVID-19 defines hallmarks of disease severity and specificity. *Cell* *185*, 916–938.e958. <https://doi.org/10.1016/j.cell.2022.01.012>.
78. Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P., Chang, H.Y., and Greenleaf, W.J. (2015). Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* *523*, 486–490. <https://doi.org/10.1038/nature14590>.
79. Telenius, J.M., Downes, D.J., Sergeant, M., Oudelaar, A.M., McGowan, S., Kerry, J., Hanssen, L.L., Schwessinger, R., Eijsbouts, C.Q., Davies, J.O., et al. (2020). CaptureCompendium: a comprehensive toolkit for 3C analysis. Preprint at bioRxiv. <https://doi.org/10.1101/2020.02.17.952572>.
80. Delaneau, O., Ongen, H., Brown, A.A., Fort, A., Panousis, N.I., and Dermitzakis, E.T. (2017). A complete tool set for molecular QTL discovery and analysis. *Nat. Commun.* *8*, 15452. <https://doi.org/10.1038/ncomms15452>.
81. Loh, P.R., Danecek, P., Palamara, P.F., Fuchsberger, C., A Reshef, Y., K Finucane, H., Schoenherr, S., Forer, L., McCarthy, S., Abecasis, G.R., et al. (2016). Reference-based phasing using the haplotype reference consortium panel. *Nat. Genet.* *48*, 1443–1448. <https://doi.org/10.1038/ng.3679>.
82. Durbin, R. (2014). Efficient haplotype matching and storage using the positional Burrows-Wheeler transform (PBWT). *Bioinformatics* *30*, 1266–1272. <https://doi.org/10.1093/bioinformatics/btu014>.
83. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* *29*, 15–21. <https://doi.org/10.1093/bioinformatics/bts635>.



84. Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. <https://doi.org/10.1093/bioinformatics/btu638>.
85. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. <https://doi.org/10.1093/bioinformatics/btq033>.
86. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. <https://doi.org/10.1186/s13059-014-0550-8>.
87. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. <https://doi.org/10.1038/nmeth.1923>.
88. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience* 10, giab008. <https://doi.org/10.1093/gigascience/giab008>.
89. Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S., and Karolchik, D. (2010). BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* 26, 2204–2207. <https://doi.org/10.1093/bioinformatics/btq351>.
90. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-seq (MACS). *Genome Biol.* 9, R137. <https://doi.org/10.1186/gb-2008-9-9-r137>.
91. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. <https://doi.org/10.1093/bioinformatics/btt656>.
92. Yu, G., Wang, L.-G., and He, Q.-Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* 31, 2382–2383. <https://doi.org/10.1093/bioinformatics/btv145>.
93. Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *J. Stat Softw.* 25, 18. <https://doi.org/10.18637/jss.v025.i01>.
94. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBO J.* 17, 3. <https://doi.org/10.14806/ej.17.1.200>.
95. Magoč, T., and Salzberg, S.L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27, 2957–2963. <https://doi.org/10.1093/bioinformatics/btr507>.
96. Eijbouts, C.Q., Burren, O.S., Newcombe, P.J., and Wallace, C. (2019). Fine mapping chromatin contacts in capture Hi-C data. *BMC Genom.* 20, 77. <https://doi.org/10.1186/s12864-018-5314-5>.
97. Fang, H., and Ultra-DD consortium; and Knight, J.C. (2017). “Pi: leveraging genetic evidence to prioritise drug targets at the gene and pathway level. R package version 1.3.3. 2017.
98. Fang, H., and Gough, J. (2014). The ‘dnet’ approach promotes emerging research on cancer patient survival. *Genome Med.* 6, 64. <https://doi.org/10.1186/s13073-014-0064-8>.
99. Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* (New York: Springer-Verlag).
100. Kanehisa, M., Sato, Y., and Kawashima, M. (2022). KEGG mapping tools for uncovering hidden features in biological data. *Protein Sci.* 31, 47–53. <https://doi.org/10.1002/pro.4172>.
101. Harmston, N., Ing-Simmons, E., Perry, M., Barešić, A., and Lenhard, B. (2015). GenomicInteractions: an R/Bioconductor package for manipulating and investigating chromatin interaction data. *BMC Genom.* 16, 963. <https://doi.org/10.1186/s12864-015-2140-x>.
102. Hahne, F., and Ivanek, R. (2016). Visualizing genomic data using Gviz and bioconductor. *Methods Mol. Biol.* 1478, 335–351. [https://doi.org/10.1007/978-1-4939-3578-9\\_16](https://doi.org/10.1007/978-1-4939-3578-9_16).
103. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47. <https://doi.org/10.1093/nar/gkv007>.
104. Rudwaleit, M., van der Heijde, D., Landewé, R., Listing, J., Akkoc, N., Brandt, J., Braun, J., Chou, C.T., Collantes-Estevez, E., Dougados, M., et al. (2009). The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Ann. Rheum. Dis.* 68, 777–783. <https://doi.org/10.1136/ard.2009.108233>.
105. Hamilton, L., Barkham, N., Bhalla, A., Brittain, R., Cook, D., Jones, G., Mackay, K., Marshall, D., Marzo-Ortega, H., Murphy, D., et al. (2017). BSR and BHPR guideline for the treatment of axial spondyloarthritis (including ankylosing spondylitis) with biologics. *Rheumatology* 56, 313–316. <https://doi.org/10.1093/rheumatology/kew223>.
106. Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* 10, 1213–1218. <https://doi.org/10.1038/nmeth.2688>.
107. Corces, M.R., Buenrostro, J.D., Wu, B., Greenside, P.G., Chan, S.M., Koenig, J.L., Snyder, M.P., Pritchard, J.K., Kundaje, A., Greenleaf, W.J., et al. (2016). Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat. Genet.* 48, 1193–1203. <https://doi.org/10.1038/ng.3646>.
108. Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* 14, 959–962. <https://doi.org/10.1038/nmeth.4396>.
109. Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* 109, 21–29. <https://doi.org/10.1002/0471142727.mb2129s109>.
110. Schmidt, C., Rendeiro, A.F., Sheffield, N.C., and Bock, C. (2015). ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors. *Nat. Methods* 12, 963–965. <https://doi.org/10.1038/nmeth.3542>.
111. Davies, J.O.J., Telenius, J.M., McGowan, S.J., Roberts, N.A., Taylor, S., Higgs, D.R., and Hughes, J.R. (2016). Multiplexed analysis of chromosome conformation at vastly improved sensitivity. *Nat. Methods* 13, 74–80. <https://doi.org/10.1038/nmeth.3664>.
112. UK10K Consortium; Walter, K., Min, J.L., Huang, J., Crooks, L., Memari, Y., McCarthy, S., Perry, J.R.B., Futema, M., Xu, C., et al. (2015). The UK10K project identifies rare variants in health and disease. *Nature* 526, 82–90. <https://doi.org/10.1038/nature14962>.
113. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
114. Jia, X., Han, B., Onengut-Gumuscu, S., Chen, W.-M., Concannon, P.J., Rich, S.S., Raychaudhuri, S., and de Bakker, P.I.W. (2013). Imputing amino acid polymorphisms in human leukocyte antigens. *PLoS One* 8, e64683. <https://doi.org/10.1371/journal.pone.0064683>.
115. McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A.R., Teumer, A., Kang, H.M., Fuchsberger, C., Danecek, P., Sharp, K., et al. (2016). A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* 48, 1279–1283. <https://doi.org/10.1038/ng.3643>.
116. Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47, D607–d613. <https://doi.org/10.1093/nar/gky1131>.



117. Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., and Aiden, E.L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680. <https://doi.org/10.1016/j.cell.2014.11.021>.
118. Fairfax, B.P., Humburg, P., Makino, S., Naranbhai, V., Wong, D., Lau, E., Jostins, L., Plant, K., Andrews, R., McGee, C., and Knight, J.C. (2014). Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science* 343, 1246949. <https://doi.org/10.1126/science.1246949>.
119. Fairfax, B.P., Makino, S., Radhakrishnan, J., Plant, K., Leslie, S., Dilthey, A., Ellis, P., Langford, C., Vannberg, F.O., and Knight, J.C. (2012). Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat. Genet.* 44, 502–510. <https://doi.org/10.1038/ng.2205>.
120. Kasela, S., Kisand, K., Tserel, L., Kaleviste, E., Remm, A., Fischer, K., Esko, T., Westra, H.-J., Fairfax, B.P., Makino, S., et al. (2017). Pathogenic implications for autoimmune mechanisms derived by comparative eQTL analysis of CD4+ versus CD8+ T cells. *PLoS Genet.* 13, e1006643. <https://doi.org/10.1371/journal.pgen.1006643>.
121. Naranbhai, V., Fairfax, B.P., Makino, S., Humburg, P., Wong, D., Ng, E., Hill, A.V.S., and Knight, J.C. (2015). Genomic modulators of gene expression in human neutrophils. *Nat. Commun.* 6, 7545. <https://doi.org/10.1038/ncomms8545>.
122. Westra, H.J., Peters, M.J., Esko, T., Yaghootkar, H., Schurmann, C., Kettenen, J., Christiansen, M.W., Fairfax, B.P., Schramm, K., Powell, J.E., et al. (2013). Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet.* 45, 1238–1243. <https://doi.org/10.1038/ng.2756>.
123. Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M., and Tanabe, M. (2021). KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res.* 49, D545–d551. <https://doi.org/10.1093/nar/gkaa970>.
124. Cairns, J., Freire-Pritchett, P., Wingett, S.W., Várnai, C., Dimond, A., Plagnol, V., Zerbino, D., Schoenfelder, S., Javierre, B.-M., Osborne, C., et al. (2016). CHICAGO: robust detection of DNA looping interactions in Capture Hi-C data. *Genome Biol.* 17, 127. <https://doi.org/10.1186/s13059-016-0992-2>.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Anti-CD14-PE	BioLegend	Cat#325602; RRID:AB_830675
Anti-CD4-BUV295	BD Biosciences	Cat# 563550, RRID:AB_2738273
Anti-CD8-BUV737	BD Biosciences	Cat#612754; RRID:AB_2870085
Anti-CD3-BV786	BD Biosciences	Cat #563918; RRID:AB_2738487
Anti-CD14-FITC	BD Biosciences	Cat #555397; RRID:AB_395798
Anti-CD16-APC	BioLegend	Cat #360706; RRID:AB_2562751
Ani-HLA-DR-BV605	BD Biosciences	Cat #562845
anti-H3K27ac	Diagenode	Cat#C15410196; RRID:AB_2637079
anti-H3K4me3	Diagenode	Cat# pAb-003-050, RRID:AB_2616052
anti-H3K4me1	Diagenode	Cat# C15410194, RRID:AB_2637078
<b>Biological samples</b>		
AS patients and HC blood samples	This study	N/A
<b>Chemicals, peptides, and recombinant proteins</b>		
Ficoll-Paque Plus	Sigma-Aldrich	GE17-1440-02
CD4 MicroBeads	Miltenyi Biotec	Cat# 130-045-101, RRID:AB_2889919
CD8 MicroBeads	Miltenyi Biotec	Cat#130-045-201; RRID:AB_2889920
CD14 MicroBeads	Miltenyi Biotec	Cat#130-050-201; RRID:AB_2665482
Trypan Blue	Gibco	15250061
FACS Lysing Solution 10x concentrate	BD Biosciences	Cat# 349202, RRID:AB_2868862
eBioscience™ Flow Cytometry Staining Buffer	ThermoFisher Scientific	Cat# 00-4222-26
7-AAD Viability Staining Solution	BioLegend	Cat# 420403
eBioscience™ 1-Step Fix/Lyse Solution	ThermoFisher Scientific	Cat# 00-5333-54
PBS, pH 7.4	Gibco	Cat#10010031
BSA	New England Biolabs	Cat#B9000S
Trizma® hydrochloride solution (Tris-HCl) pH 7.4	Sigma-Aldrich	Cat#T2194
Sodium chloride solution, 5M	Sigma-Aldrich	Cat#S5150
MgCl <sub>2</sub> (1M)	Invitrogen	Cat#AM9530G
Digitonin	Promega	Cat#G9441
Tween 20	Sigma-Aldrich	Cat#P1379
TD Tagment DNA Buffer	Illumina	Cat#15027866
TDE1 Tagment DNA Enzyme	Illumina	Cat#15027865
AMPure XP Beads	Beckman Coulter	Cat#A63881
Formaldehyde solution	Sigma-Aldrich	Cat#F8775
Glycine	Sigma-Aldrich	Cat#G7403
Sodium dodecyl sulphate solution	Sigma-Aldrich	Cat#71736
EDTA (0.5M) pH8	Sigma-Aldrich	Cat#102161034
Tris (1 M), pH 8.0, RNase-free	Invitrogen	Cat#AM9855G
Complete Protease Inhibitor Cocktail	Sigma-Aldrich	Cat#COEDTAF-RO
Triton X-100	Sigma-Aldrich	Cat#T8787
Dynabeads Protein G for Immunoprecipitation	Invitrogen	Cat#10003D
10% Igepal CA-630	Sigma-Aldrich	Cat#I8896
DpnII 50,000 U/ml	New England Biolabs	Cat#R0543M
T4 DNA HC ligase (30 Weiss U/μL)	Thermo Fisher Scientific	Cat#EL0013
RNase, DNase free	Roche	Cat#1119915
Proteinase K	Thermo Fisher Scientific	Cat#EO0491

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Dynabeads M-270 Streptavidin	Invitrogen	Cat#65305
Ribonucleic acid, transfer from baker's yeast	Sigma-Aldrich	Cat#R5636
Lithium chloride solution, 8M	Sigma-Aldrich	Cat#L7026
Sodium deoxycholate monohydrate	Alfa Aesar	Cat#B20759
Tris-EDTA buffer solution	Sigma-Aldrich	Cat#T9285

**Critical commercial assays**

Live/Dead Fixable Violet Dead cell stain kit	Invitrogen	Cat#L34955
AllPrep DNA/mRNA/miRNA Universal kit	Qiagen	Cat#80224
Ribo-Zero rRNA Removal kit	Illumina	Cat#20040526
TruSeq Stranded Total RNA	Illumina	Cat#20020596
MinElute PCR purification kit	Qiagen	Cat#28004
NEBNext High-Fidelity 2x PCR master mix	New England Biolabs	Cat#M0541S/L
NEBNext DNA Library Prep Master Mix Set	New England Biolabs	Cat#E6040S/L
Herculase II Fusion Enzyme with dNTPs Combo	Agilent	Cat#600677
Nimblegen SeqCap EZ Hybridisation and wash kit	Roche	Cat#05634261001
Nimblegen SeqCap EZ Accessory kit v2	Roche	Cat#07145594001
KAPA Library Quantification Complete Kit (Universal)	KAPA	Cat#KK4824
NextSeq 500/550 High Output kit v2.5 (150 Cycles)	Illumina	Cat#20024907
TapeStation D1000 Screen Tape	Agilent	Cat#5067-5582
TapeStation High Sensitivity D1000 Screen Tape	Agilent	Cat#5067-5584
TapeStation D1000 Reagents	Agilent	Cat#5067-5583
Infinium Global Array V2.0	Illumina	Cat#20024444

**Deposited data**

PCHI-C	Javierre et al. <sup>38</sup>	<a href="https://osf.io/u8tzp/">https://osf.io/u8tzp/</a>
Ankylosing spondylitis ImmunoChip summary statistics	Cortes et al. <sup>9</sup>	N/A
Cross-disease GWAS summary statistics	Ellinghaus et al. <sup>10</sup>	N/A
eQTL catalogue	Kerimov et al. <sup>43</sup>	<a href="https://www.ebi.ac.uk/eqtl/">https://www.ebi.ac.uk/eqtl/</a>
Human PBMC scRNA-seq data	COvid-19 Multi-omics Blood Atlas (COMBAT) consortium <sup>77</sup>	<a href="https://www.combat.ox.ac.uk/">https://www.combat.ox.ac.uk/</a>
RNA-seq, ATAC-seq, ChIPm fastq data	This study	European Genome-Phenome Archive: EGAS00001006233
Genotype data	This study	European Genome-Phenome Archive: EGAS00001006945
RNA-seq, ATAC-seq, ChIPm raw and normalised count data; Capture-C count data and PeakY scores; ChromHMM data	This study	<a href="https://doi.org/10.5281/zenodo.6373353">https://doi.org/10.5281/zenodo.6373353</a>

**Oligonucleotides**

Capture-C baits	This study	Table S8
NEBNext Multiplex Oligos for Illumina (Index Primers set 1)	New England Biolabs	Cat#E7335S/L
NEBNext Multiplex Oligos for Illumina (Index Primers set 2)	New England Biolabs	Cat#E7500S/L
Nimblegen HyperCap Universal Blocking Oligos	Roche	Cat#08286396001
Modified Nextera Index primers	Buenrostro <sup>78</sup>	N/A

**Software and algorithms**

CapSequm2	Telenius et al. <sup>79</sup>	<a href="https://capsequm.molbiol.ox.ac.uk/cgi-bin/CapSequm.cgi">https://capsequm.molbiol.ox.ac.uk/cgi-bin/CapSequm.cgi</a>
CaptureCompare	Telenius et al. <sup>79</sup>	<a href="https://github.com/djdownes/CaptureCompare">https://github.com/djdownes/CaptureCompare</a>

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
QTLtools	Delaneau et al. <sup>80</sup>	<a href="https://qtltools.github.io/qtltools/">https://qtltools.github.io/qtltools/</a>
CIBERSORTx	Newman et al. <sup>37</sup>	<a href="https://cibersortx.stanford.edu/">https://cibersortx.stanford.edu/</a>
Eagle2	Loh et al. <sup>81</sup>	<a href="https://alkesgroup.broadinstitute.org/Eagle/">https://alkesgroup.broadinstitute.org/Eagle/</a>
PBWT	Durbin et al. <sup>82</sup>	<a href="https://github.com/richarddurbin/pbwt">https://github.com/richarddurbin/pbwt</a>
STAR	Dobin et al. <sup>83</sup>	<a href="https://github.com/alexdobin/STAR/releases">https://github.com/alexdobin/STAR/releases</a> ; RRID:SCR_004463
Picard tools	Picard Toolkit. 2019. Broad Institute, GitHub Repository.	<a href="https://broadinstitute.github.io/picard/">https://broadinstitute.github.io/picard/</a> ; RRID:SCR_006525
HTSeq	Anders et al. <sup>84</sup>	<a href="https://htseq.readthedocs.io/en/release_0.11.1/count.html">https://htseq.readthedocs.io/en/release_0.11.1/count.html</a> ; RRID:SCR_011867
BEDTools	Quinlan et al. <sup>85</sup>	<a href="https://bedtools.readthedocs.io/en/latest/">https://bedtools.readthedocs.io/en/latest/</a> ; RRID:SCR_006646
DESeq2	Love et al. <sup>86</sup>	<a href="https://bioconductor.org/packages/release/bioc/html/DESeq2.html">https://bioconductor.org/packages/release/bioc/html/DESeq2.html</a> ; RRID:SCR_015687
XGR	Fang et al. <sup>33</sup>	<a href="http://galahad.well.ox.ac.uk:3030/">http://galahad.well.ox.ac.uk:3030/</a>
Bowtie2 v2.3.5.1	Langmead et al. <sup>87</sup>	<a href="http://bowtie-bio.sourceforge.net/bowtie2/index.shtml">http://bowtie-bio.sourceforge.net/bowtie2/index.shtml</a> ; RRID:SCR_016368
Samtools v.1.12	Danecek et al. <sup>88</sup>	<a href="http://www.htslib.org/">http://www.htslib.org/</a> ; RRID:SCR_002105
bedGraphToBigWig	Kent et al. <sup>89</sup>	<a href="https://www.encodeproject.org/software/bedgraphbigwig/">https://www.encodeproject.org/software/bedgraphbigwig/</a>
MACS2	Zhang et al. <sup>90</sup>	<a href="https://macs3-project.github.io/MACS/">https://macs3-project.github.io/MACS/</a> ; RRID:SCR_013291
featureCounts	Liao et al. <sup>91</sup>	<a href="http://subread.sourceforge.net">http://subread.sourceforge.net</a> ; RRID:SCR_012919
ChIPseeker	Yu et al. <sup>92</sup>	<a href="http://www.bioconductor.org/packages/release/bioc/html/ChIPseeker.html">http://www.bioconductor.org/packages/release/bioc/html/ChIPseeker.html</a> ; RRID:SCR_021322
ChromHMM	Ernst et al. <sup>39</sup>	<a href="http://compbio.mit.edu/ChromHMM/">http://compbio.mit.edu/ChromHMM/</a> ; RRID:SCR_018141
FactoMineR	Lê et al. <sup>93</sup>	<a href="http://factominer.free.fr/">http://factominer.free.fr/</a> ; RRID:SCR_014602
GARFIELD	lotchkova et al. <sup>42</sup>	<a href="https://www.ebi.ac.uk/birney-srv/GARFIELD/">https://www.ebi.ac.uk/birney-srv/GARFIELD/</a>
FastQC v0.11.4	Babraham Bioinformatics	<a href="https://www.bioinformatics.babraham.ac.uk/projects/fastqc">https://www.bioinformatics.babraham.ac.uk/projects/fastqc</a> ; RRID:SCR_014583
cutadapt v 1.10	Martin et al. <sup>94</sup>	<a href="https://cutadapt.readthedocs.io/en/stable/">https://cutadapt.readthedocs.io/en/stable/</a> ; RRID:SCR_011841
TrimGalore v 0.4.4	Babraham Bioinformatics	<a href="https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/">https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/</a>
FLASH	Magoč et al. <sup>95</sup>	<a href="https://sourceforge.net/projects/flashpage/files/FLASH-1.2.11.tar.gz/">https://sourceforge.net/projects/flashpage/files/FLASH-1.2.11.tar.gz/</a> ; RRID:SCR_005531
PeakY	Eijsbouts et al. <sup>96</sup>	<a href="https://github.com/cqgd/pky">https://github.com/cqgd/pky</a>
Pi	Fang et al. <sup>33</sup>	<a href="http://pi.well.ox.ac.uk:3010/">http://pi.well.ox.ac.uk:3010/</a> ; <sup>97</sup>
Dnet	Fang et al. <sup>98</sup>	<a href="https://cran.r-project.org/package=dnet">https://cran.r-project.org/package=dnet</a>
ggplot2	Wickham et al. <sup>99</sup>	<a href="https://ggplot2.tidyverse.org">https://ggplot2.tidyverse.org</a> ; RRID:SCR_014601

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
EnhancedVolcano v1.12.0	EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labelling	<a href="https://github.com/kevinblighe/EnhancedVolcano">https://github.com/kevinblighe/EnhancedVolcano</a> ; RRID:SCR_018931
Venny 2.1	VENNY. An interactive tool for comparing lists with Venn Diagrams.	<a href="https://bioinfogp.cnb.csic.es/tools/venny/">https://bioinfogp.cnb.csic.es/tools/venny/</a> ; RRID:SCR_016561
KEGG mapper – Colour	Kanehisa et al. <sup>100</sup>	<a href="https://www.genome.jp/kegg/mapper/color.html">https://www.genome.jp/kegg/mapper/color.html</a>
GenomicInteractions	Harmston et al. <sup>101</sup>	<a href="https://www.bioconductor.org/packages/release/bioc/html/GenomicInteractions.html">https://www.bioconductor.org/packages/release/bioc/html/GenomicInteractions.html</a>
Gviz	Hahne et al 2016 <sup>102</sup>	<a href="https://bioconductor.org/packages/release/bioc/html/Gviz.html">https://bioconductor.org/packages/release/bioc/html/Gviz.html</a>
Limma	Ritchie et al. <sup>103</sup>	<a href="https://bioconductor.org/packages/release/bioc/html/limma.html">https://bioconductor.org/packages/release/bioc/html/limma.html</a> ; RRID:SCR_010943
R	R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria	<a href="https://www.R-project.org/">https://www.R-project.org/</a> ; RRID:SCR_001905
ICeCAP pipeline	This paper	<a href="https://doi.org/10.5281/zenodo.7760066">https://doi.org/10.5281/zenodo.7760066</a>

**Other**

DynaMag-2 Magnet	Invitrogen	12321D
AutoMACS Pro Separator	Miltenyi Biotech	<a href="https://www.miltenyibiotec.com/GB-en/products/automacs-pro-separator-starter-kit.html#gref">https://www.miltenyibiotec.com/GB-en/products/automacs-pro-separator-starter-kit.html#gref</a> ; RRID:SCR_018596
LSRFortessa X-20-flow cytometer	BD Biosciences	<a href="https://www.bdbiosciences.com/en-us/instruments/research-instruments/research-cell-analyzers/lrsfortessa">https://www.bdbiosciences.com/en-us/instruments/research-instruments/research-cell-analyzers/lrsfortessa</a> ; RRID:SCR_018655
HiSeq4000 platform	Illumina	<a href="https://emea.illumina.com/systems/sequencing-platforms/hiseq-3000-4000.html">https://emea.illumina.com/systems/sequencing-platforms/hiseq-3000-4000.html</a>
NextSeq 500 platform	Illumina	<a href="https://emea.illumina.com/systems/sequencing-platforms/nextseq.html">https://emea.illumina.com/systems/sequencing-platforms/nextseq.html</a>
M220 focussed ultrasonicator	Covaris	<a href="https://www.covaris.com/m220-focused-ultrasonicator-500295">https://www.covaris.com/m220-focused-ultrasonicator-500295</a>
TapeStation 4200	Agilent	<a href="https://www.agilent.com/en/product/automated-electrophoresis/tapestation-systems/tapestation-instruments/4200-tapestation-system-228263">https://www.agilent.com/en/product/automated-electrophoresis/tapestation-systems/tapestation-instruments/4200-tapestation-system-228263</a>

**RESOURCE AVAILABILITY**

**Lead contact**

Further information and request for resources and reagents should be directed to and will be fulfilled by the lead contact, Julian Knight ([julian@well.ox.ac.uk](mailto:julian@well.ox.ac.uk)).

**Materials availability**

This study did not generate new unique reagents.



### Data and code availability

- Sequence level RNA-seq, ATAC-seq, ChIPm fastq and genotype data have been deposited at the European Genome-Phenome Archive (EGA); access is managed by a Data Access Committee. Count data for RNA-seq, ATAC-seq, ChIPm and eRNA (raw and normalised), Capture-C mean pile-up tracks and PeakY scores, and ChromHMM data are deposited at Zenodo. Accession numbers are listed in the [key resources table](#). This paper analyses existing, publicly available data (GWAS data, PCHiC data and eQTL). The accession numbers for the datasets are listed in the [key resources table](#).
- All original code for the ICeCAP pipeline has been deposited at GitHub and Zenodo and is publicly available as of the date of publication. DOIs are listed in the [key resources table](#).
- Any additional information required to reanalyse the data reported in this paper is available from the lead contact upon request.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Human subjects

20 patients with AS and 35 HC were enrolled in the study; see demographic information in [Table S1](#). All samples were collected following informed consent and under ethical approval: National Research Ethics Service Committee South Central – Oxford Research Ethics Committee B (Ref 12/SC/0063 for patient samples, Ref 06/Q1605/55 for HC). AS patient samples were collected at the Nuffield Orthopaedic Centre in Oxford. Patients fulfilled the ASAS imaging criteria for axial Spondyloarthritis.<sup>104</sup> Cases were reviewed by two consultant rheumatologists to confirm the patients had active disease at the time of recruitment as judged by the British Society for Rheumatology criteria for biologic therapy eligibility.<sup>105</sup> 85% of patients and 8.6% of HC were positive for HLA-B27. All patients had axial manifestation of spondyloarthritis, and peripheral and extra-articular manifestations were present in a subset of patients. 53% of patients were taking NSAIDs, 11% were taking csDMARDs and none were taking glucocorticoids or anti-TNF biologic therapy (see clinical summary in [Table S1](#) and extended clinical information in [Table S2](#)). All participants were aged over 18 and had no other current infections. 75% of AS patients were male, which reflects the higher frequency of disease in men; 49% of HC were male, and controls were matched for age. Human samples were collected over time and next-generation sequencing (NGS) was performed in seven sets ([Table S2](#)). NGS set and sex were included as covariates in all downstream analyses except ATAC analysis, where ATAC method and sex were included as covariates. Data quality control was performed by PCA and outliers were removed as appropriate.

## METHOD DETAILS

### Isolation of immune cell populations

Up to 100 mL blood was taken from AS patients or HC into heparinised vacutainers. PBMCs were isolated from blood samples by density gradient separation using Ficoll-Paque Plus with centrifugation at 500 x g for 30 minutes at room temperature with minimum acceleration and no braking. Primary cell subpopulations were separated using magnetic-activated cell sorting following the manufacturer's instructions. Consecutive positive selection was performed using beads for CD8<sup>+</sup>, and CD4<sup>+</sup> cells using an AutoMACS Pro (Miltenyi Biotec) followed by a manual cell count with Trypan blue.

### Flow cytometry

Cell purity following MACS separation was determined by flow cytometry. Briefly, cells were incubated with Live/Dead Fixable Violet Dead cell stain kit (Invitrogen) for 30 minutes followed by incubation with either anti-CD14-PE, anti-CD4-BUV395 or anti-CD8-BUV737 antibodies for 30 minutes at room temperature. Cells were fixed for 10 minutes using FACS Lysing Solution then re-suspended in PBS + 0.2% BSA for acquisition using LSR Fortessa X-20-flow cytometer (BD Biosciences). Cells were immediately processed for downstream assays without further culture ([Figure 1A](#)). The composition of monocyte bulk populations was determined as follows. Cells were washed with PBS and pelleted, then incubated with Human BD Fc Block™ Reagent for 10 minutes at room temperature. Cells were incubated for 30 minutes with anti-CD3-BV786, anti-CD14-FITC, anti-CD16-APC and anti-HLA-DR-BV605 antibodies diluted in eBioScience™ Flow Cytometry Staining Buffer (2 μL antibody/1x10<sup>6</sup> cells). Cells were also stained for viability using 7-AAD Viability Staining Solution then fixed for 10 minutes with 1X eBioScience™ 1-Step Fix/Lyse Solution, and washed with 0.2% PBS-BSA prior to flow cytometry analysis using a BD LSRFortessa™ X-20 Cell Analyzer. 1x10<sup>5</sup> events were recorded for each sample.

### RNA-seq

Total RNA was isolated from purified CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells and monocytes (2–3 x 10<sup>6</sup> cells resuspended in RLT Plus) using the AllPrep DNA/mRNA/microRNA Universal kit (Qiagen) from 16 AS patients and 19 HC. Samples were depleted from ribosomal RNA using Ribo-Zero rRNA Removal kit (Illumina) prior to cDNA synthesis and library preparation using TruSeq Stranded Total RNA (Illumina). Libraries were sequenced using HiSeq4000 to a depth of 25 million paired-end reads per sample.

### ATAC-seq

ATAC-seq was used to assess chromatin accessibility in CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells and monocytes from 16 AS patients and 30 HC. Three different variations of the ATAC-seq protocol were used as improvements to the method were published: ATAC-seq<sup>106</sup> (5 x 10<sup>4</sup> cells); Fast-ATAC<sup>107</sup> (FATAC, 2 x 10<sup>4</sup> cells); Omni-ATAC<sup>108</sup> (OATAC, 5 x 10<sup>4</sup> cells), with minor modification to determine amplification by using 10% of the sample in qPCR prior to indexing and amplification. Following MACS isolation, ATAC was carried out. For ATAC-seq cells were re-suspended in 50  $\mu$ L lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.1% IGEPAL CA-630) for 10 minutes at 4°C. Nuclei were pelleted (500 x g 10 min at 4°C) and re-suspended in 50  $\mu$ L transposition mixture (1 x TD Tagment DNA buffer (Illumina), 2.5  $\mu$ L TDE1 (Illumina) and incubated for 40 min at 37 °C with agitation at 400 rpm. For Fast-ATAC-seq cells were incubated in 50  $\mu$ L transposition mixture (1 x TD Tagment DNA buffer (Illumina), 2.5  $\mu$ L of TDE1 (Illumina), 0.01% Digitonin) for 30 minutes at 37°C with agitation at 400 rpm. For Omni-ATAC-seq cells were re-suspended in 50  $\mu$ L lysis buffer (10 mM Tris-HCl pH7.4, 10 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.01% Digitonin, 0.1% Tween 20, 0.1% Igepal CA-630) for 3 minutes at 4°C, then 1 mL cold wash buffer (10 mM Tris-HCl pH7.4, 10 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.1% Tween 20) was added. Nuclei were pelleted (500 x g 10 min at 4°C), re-suspended in 50  $\mu$ L transposition mixture (1 x TD Tagment DNA buffer (Illumina), 2.5  $\mu$ L TDE1 (Illumina), 0.01 % Digitonin and 0.1 % Tween-20), and incubated for 30 min at 37°C with agitation at 1000 rpm. DNA was extracted with MinElute PCR Purification Kit (Qiagen). Samples were amplified and indexed as in Buenrostro et al<sup>109</sup> with NEBNext High-Fidelity PCR MasterMix (NEB) with modified Nextera indexing primers.<sup>78</sup> DNA libraries were purified using the MinElute PCR purification kit (Qiagen) and AMPure XP Magnetic Beads (Beckman Coulter). Library distribution was determined using TapeStation D1000 reagents and tape. Libraries were sequenced on a HiSeq4000 or NextSeq500 Illumina platform to a depth of 30 million paired-end reads after filtering.

### ChIPmentation

ChIPmentation (ChIPm) for H3K4me3 and H3K27ac was performed as described<sup>110</sup> on samples of 1x10<sup>5</sup> CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells and monocytes from 10 AS patients and 11 HC.

ChIPm was also performed for H3K4me1 on samples from one healthy individual not included in the other experiments. Following MACS isolation, cells were fixed with 1% formaldehyde (Sigma) in PBS for 10 min, then quenched with glycine (0.25 M for 5 min). Cells were pelleted (100 x g 2.5 minutes), washed with PBS, then pelleted again and re-suspended in 130  $\mu$ L SDS lysis buffer (0.25% SDS, 1mM EDTA, 10mM Tris-HCl pH8, 1x Protease Inhibitor) and sonicated for 8 min using M220 Focused-ultrasonicator (Covaris), duty factor 10%, peak power of 75, cycles/burst of 200 set to 7°C. 50  $\mu$ L chromatin aliquots were mixed with 75  $\mu$ L ChIP Equilibration Buffer (1.66% Triton-X100, 1 mM EDTA, 10 mM Tris-HCl pH8, 233 mM NaCl, 1x protease inhibitor) and 25  $\mu$ L ChIP buffer (0.1 % SDS, 1% Triton-X100, 1 mM EDTA, 10 mM Tris-HCl pH8, 140 mM NaCl, 1x protease inhibitor). Antibodies (1  $\mu$ g anti-H3K4me3, 2  $\mu$ g anti-H3K27ac, 1  $\mu$ g anti-H3K4me1) were added and incubated overnight at 4 °C with rotation. Protein G Dynabeads were prepared in bead wash buffer (0.1% SDS, 1 mM EDTA, 50 mM Tris-HCl pH8, 150 mM NaCl, 1% Igepal CA-630, 1 x protease inhibitor) and blocked with yeast tRNA and BSA as per the manufacturer's instructions. Samples and beads were incubated for 2 hr 4°C with rotation, then washed twice with 200  $\mu$ L each wash buffer (wash buffer 1: 0.1% SDS, 1% Triton-X100, 2 mM EDTA, 20 mM Tris-HCl pH8, 150 mM NaCl, 1 x protease inhibitor; wash buffer 2: 0.1% SDS, 1% Triton-X100, 2 mM EDTA, 20 mM Tris-HCl pH8, 500 mM NaCl, 1 x protease inhibitor; wash buffer 3: 0.25 M LiCl, 1% Triton-X100, 0.7% sodium deoxycholate monohydrate, 1 mM EDTA, 10 mM Tris-HCl pH8, 1 x protease inhibitor) and once with 10 mM Tris-HCl pH8. Samples were finally resuspended in 20  $\mu$ L Tagmentation reaction buffer (1 x TD buffer, 1  $\mu$ L TDE 1 Tagment DNA Enzyme (Illumina)) and incubated for 10 min at 37 °C. Samples were washed twice with 200  $\mu$ L wash buffer (0.1% SDS, 1% Triton-X100, 2 mM EDTA, 20 mM Tris-HCl pH8, 150 mM NaCl, 1 x protease inhibitor) and once with Tris-EDTA buffer. Samples were de-crosslinked in ChIP Elution buffer (10 mM Tris-HCl pH8, 5 mM EDTA, 300 mM NaCl, 0.4% SDS) containing 2.4 U proteinase K for 1 hour at 55 °C then overnight at 65 °C with 1400 rpm shaking. Samples were purified using the MinElute PCR purification kit (QIAGEN). Indexing was performed with modified Nextera indexing primers<sup>78</sup> and NEBNext HiFi PCR master mix using cycle number determined by qPCR. PCR product clean-up was performed using AMPureXP beads and QC was performed using TapeStation D1000 tape and reagents. DNA quantitation was performed using KAPA assay. Input libraries were prepared using 1 ng of purified chromatin without antibody incubation. Libraries were sequenced using the HiSeq4000 Illumina platform at a depth of 25 million paired-end reads after filtering.

### Capture-C

Capture-C experiments were performed as described by Davies et al<sup>111</sup> with minor changes.

Oligonucleotide baits for sequence-capture were designed for regions containing (i) 52 GWAS SNPs associated with AS,<sup>9,10</sup> (ii) 47 promoters of nearby genes, and (iii) five control regions with known genomic interactions (Table S8, Figure S14). Two biotinylated 120nt ssDNA bait sequences were designed for each viewpoint (DpnII fragment to be captured) using CapSequm2<sup>79</sup> and synthesized by Sigma-Aldrich. Some bait oligos were 80nt and some baited viewpoints had only one oligo bait due to regional sequence repetitiveness. All oligo baits were combined in equimolar amounts to make a pool containing each oligo at 2.9nM. For two regions, interactions were seen not with the baited promoter, but with other genes so results are presented with those gene names (*ITGA4* relates to bait for *UBE2E3*, and *IL12RAP* relates to bait for *IL1R2*). We were unable to capture the DpnII fragment containing rs4672505 using Capture-C due to DNA repeats.

Following MACS isolation,  $10\text{--}30 \times 10^6$  of monocytes, CD8<sup>+</sup> T cells and CD4<sup>+</sup> T cells from three AS patients and three HC were fixed with 2% formaldehyde (10 min), quenched with 0.1M glycine (10 min on ice), washed with PBS and snap frozen in 1ml of lysis buffer (10 mM Tris-HCl pH8, 10 mM NaCl, 0.3% Igepal CA-630, 1 x protease inhibitor). Cells were thawed and centrifuged at 500 x g for 5 minutes at RT. The supernatant was removed, cells were washed in 1 mL of 1 x DpnII buffer (NEB) to remove the residual lysis buffer and spun down at 500 x g for 5min at room temperature. Cells were re-suspended in 200  $\mu$ L of 1 x DpnII buffer per every  $6 \times 10^6$  cells and homogenized to free the nuclei. Chromatin was digested using 1500U of DpnII per  $6 \times 10^6$  cells overnight at 37 °C with shaking 900 rpm; the reaction was stopped by incubation for 20 minutes at 65 °C. Ligation was performed with 240 U T4 DNA ligase overnight at 16 °C. Controls were prepared without DpnII digestion (Undigested) and without ligation steps (Digested). De-crosslinking was performed with addition of 3 U Proteinase K and overnight incubation at 65 °C. Samples were treated with RNase A for 30 min at 37 °C then DNA was purified by phenol-chloroform extraction and ethanol precipitation. Covaris M220 focussed ultrasonicator was used to fragment the 3C material to 200 bp (duty cycle, 20%; intensity, 50; cycles per burst, 200; time, 280s), then AMPure XP SPRI bead cleanup was performed. Illumina TruSeq sequencing adapters (NEBNext Multiplex Oligos Index Primers sets 1 and 2, NEB) were added to 5 $\mu$ g of sonicated 3C material using NEBNext DNA Library Prep Master Mix Set (NEB) reagents for end repair, dA tailing and adaptor ligation, Herculase II Fusion Enzyme (Agilent) for indexing PCR, and AMPure® XP Beads (Beckman Coulter) for clean-up steps.

Selective enrichment of 3C libraries was performed as described<sup>111</sup> with minor changes using Nimblegen SeqCap EZ Hybridisation and wash kit, Accessory kit v2 (Roche) and their HyperCap Workflow v2.0.

Briefly, the oligo bait pool was hybridised to a pool of indexed 3C libraries (up to six 3C libraries equating to 6  $\mu$ g of material per hybridisation reaction) in the presence of Universal Blocking Oligos (Roche) and SeqCap EZ reagents (Roche) at 47°C for 72h. The captured fragments were pulled down with M-270 Streptavidin Dynabeads (Invitrogen), washed with SeqCap EZ wash buffers, cleaned up using AMPure XP Beads (Beckman Coulter) and amplified using KAPA Library Quantification Complete Kit (Roche) (9–12 cycles). The resulting enriched library was used as an input for a secondary capture following the same protocol as above, but with hybridization time of 24h and fewer final PCR cycles determined by a test qPCR reaction. Quality control was performed using TapeStation high sensitivity D1000 screen tape and reagents. Capture-C libraries were sequenced on the Illumina NextSeq 500 platform using 150bp paired-end reads at a depth of 1 million paired end reads per viewpoint.

### Genotyping

DNA was isolated from purified CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells and monocytes ( $2\text{--}3 \times 10^6$  cells resuspended in RLT Plus) using the AllPrep DNA/mRNA/microRNA Universal kit (Qiagen) from 20 AS patients and 20 HC. Samples were submitted for genome wide array genotyping at the Oxford Genomics Center and processed using the Infinium Global Array V2.0 (Illumina).

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Analysis of flow cytometry data

Flow cytometry data were analysed using FlowJo Version 10.8.1. Cells were gated by size (FSC) and granularity (SSC) and then for singlets by FSC-H vs FSC-W followed by SSC-H vs. SSC-W. CD3<sup>neg</sup> cells were selected using CD3 vs. SSC-A. A CD14 vs. CD16 plot was used to define CD14hi/CD16neg classical monocytes (cMono), CD14hi/CD16pos intermediate monocytes (iMono), and CD14lo/CD16pos non-classical monocytes (ncMono). CD14lo/CD16pos cells were defined as non-monocytic (not Mono).

### Downstream analysis of NGS data

#### RNA-seq

NGS data was mapped to human genome assembly GRCh37 (hg19) using STAR,<sup>83</sup> reads were counted using featureCount<sup>91</sup> with those mapping to X and Y chromosomes removed. Duplicates were marked and removed using Picard Tools. Genes lying within 500kb of lead AS-associated SNPs<sup>9,10</sup> were identified using BEDtools<sup>85</sup> window. PCA was performed using DESeq2 accounting for sex and sequencing set using limma<sup>103</sup> (Table S2). Differential analysis between AS patients and HC was performed using DESeq2<sup>86</sup> with sex and batch effect included in the design. Thresholds padj <0.05 and fold-change >1.5 were used to call significantly differential genes. Enriched pathways were identified in REACTOME pathways using XGR.<sup>33</sup>

#### ATAC-seq, ChIPm and eRNA analysis

Reads were aligned to the human genome assembly GRCh37 (hg19) using bowtie2.<sup>87</sup> Picard Tools was used to remove PCR duplicates, read with MAPQ score <30, non-uniquely mapping reads, non-properly paired reads and mitochondrial reads. Pileup tracks were generated using BEDTools genomCoverageBed and bedGraphToBigWig.<sup>85</sup> Normalised bigWigs were generated from normalised bedgraph files with BEDTools genomecov. Peak calling was performed using MACS2 callpeak<sup>90</sup> and peak master lists were built by union of peaks present in at least 20% of samples. Reads for ATAC and ChIPm were counted using HTSeq,<sup>84</sup> with those from X and Y chromosomes removed subsequently. Enhancer RNAs (eRNAs) were defined as uniquely mapped RNA reads within ATAC-seq peaks >3kb from a gene coding sequence and were counted using featureCounts.<sup>91</sup> Genomic distributions of ATAC peaks, ChIP peaks and eRNAs were generated using ChIPseeker.<sup>92</sup> In all cases features were filtered that did not have at least 10 reads in the smallest batch, or in at more than one sample for ChIP. PCA was performed on ATAC-seq, ChIPm and eRNA data both across and within each cell type using DESeq2<sup>86</sup> accounting for sex and batch using limma.<sup>103</sup> Differential analysis within each cell type

to compare samples from AS patients and HC was performed using DESeq2<sup>86</sup> including sex and batch in the design. QTLtools<sup>80</sup> was used for QTL discovery and trait significance was calculated using ANOVA with Tukey post-test.

### Computational deconvolution of bulk RNA-seq data

The composition of CD4<sup>+</sup> T cell, CD8<sup>+</sup> T cell and monocyte populations was analysed by deconvolution of bulk RNA-seq data using CIBERSORTx<sup>37</sup>. Reference single-cell RNA-seq data from PBMC from 10 healthy individuals was obtained from the COMBAT consortium.<sup>77</sup> Raw counts for cells annotated as CD4, CD8, or monocyte (comprising cMono or ncMono) were extracted and used as inputs for the CIBERSORTx Create Signature Matrix function, performed with parameters Min.Expression = 0.25, Replicates = 20 and Sampling = 0.5. The CIBERSORTx Impute Cell Fractions module was run using these Signature Matrices and raw bulk RNA-seq counts from the CD4<sup>+</sup> T cell, CD8<sup>+</sup> T cell and CD14<sup>+</sup> monocyte populations with S-mode batch correction and 100 permutations. Annotations were used according to the COMBAT consortium minor subsets. For visualisation, cycling classical monocytes (cMono\_cyc) were merged with classical monocytes (cMono), proliferating T Effector cells (TEFF.prolif) were merged with T effector cells (TEFF), and CCL5-expressing T central memory cells (TCM\_CCL5) were merged with T central memory cells (TCM).

### ChromHMM

ChromHMM<sup>39</sup> was run on a subset of four AS patients and four HC samples for which we had all data types, RNA, ATAC, H3K4me3 and H3K27ac and were processed together as a single batch. A generic H3K4me1 track was used that was independent of all samples. For the coding track a binarized file of the genes was generated based on the presence or absence of the gene within each 200bp bin. The default settings were used for ChromHMM and a 14-state model based on analysis of the output for varying model sizes. Manual curation was used to assign putative functions to each state according to the combination of epigenomic marks and genomic distribution. MCA was performed on global or individual chromatin states using FactoMineR.<sup>93</sup> In order to reduce the number of regions analysed (to reduce computational time), the most polymorphic regions were selected. For the global analysis (Figure 3D), 200-bp regions were kept if they contained at least two different states, and if the quiescent state 14 (Quies) was present in less than  $\frac{2}{3}$  of the samples. For individual state analysis (Figure S3B), regions were selected where that state was present in at least  $\frac{1}{3}$  of the samples.

Differences in ChromHMM states between AS patients and HC were assessed by Fisher exact tests for each state and each cell type separately. Because of the limited sample size (4 AS, 4 HC) and of the binary parameter analysed, the number of p-values obtained was limited with a minimal p value of 0.05 corresponding to a state present in 4 AS patients and no HC or 4 HC and no AS patients. To identify states with a higher number of differentially distributed regions between the two groups than expected by chance, we permuted the disease status as many times as possible (70 combinations) and compared the experimentally observed frequency of significant regions (Fisher  $p < 0.05$ ) with the permuted distribution (Figure S3B). From this we calculated a p value whose significance is reported in Figure 3C. The most significant p value possible is  $\frac{1}{70}$  i.e. 0.015.

### Pathway enrichment analysis

ATAC and ChIP peaks, eRNAs and ChromHMM regions were assigned to genes on the basis of (i) genomic proximity and (ii) genomic looping events. Proximal genes (<50 kb) were annotated using XGR.<sup>36</sup> Genomic looping events were obtained from PChi-C data<sup>38</sup> from the cell types most similar to those in our study (total CD4<sup>+</sup> and CD8<sup>+</sup> T cells, and monocytes). Regions with >50% overlap with the ends of PChi-C loops were identified using BEDtools.<sup>85</sup> Pathway enrichment analysis in REACTOME and KEGG databases was performed using XGR on gene sets from individual cell types. Gene sets were generated from the top 200 differential ATAC-seq peaks, ChIPm peaks or eRNAs, and genes associated with differential ChromHMM regions (Fisher  $p < 0.05$ ).

### GWAS enrichment

Enrichment of AS-associated variants was determined using GARFIELD<sup>42</sup> for each of the epigenomic marks identified in this study (ATAC, H3K27ac and H3K4me3) and the chromatin states defined by ChromHMM analysis. Enrichment of AS variants from European subanalysis of Cortes et al.<sup>9</sup> was assessed at four GWAS significance thresholds:  $p < 1$ , 0.1,  $5 \times 10^{-7}$  and  $5 \times 10^{-8}$ . The UK10K variant set,<sup>112</sup> pre-processed in GARFIELD, was used as the reference population for these analyses, with correction for multiple testing performed using the Bonferroni method.

### Capture-C NGS data analysis

The quality of Illumina reads was validated with FastQC v0.11.4 and the presence of Illumina adapter sequences addressed using cutadapt v1.10.<sup>94</sup> TrimGalore v.0.4.4 was used to automate quality and adaptor trimming. We extended the length of the pair-end libraries following the overlap and merge criteria of FLASH96 a fast length adjustment algorithm 100, so that two sets of merged and non-merged reads were generated for downstream analysis.

A single core version of FLASH was modified for Capture-C analysis, to allow for in-silico digestion of flashed and non-flashed reads at the DpnII recognition sequence, subsequent to the merging step. Pairs of in-silico digested, reported read sequences were enumerated, counted and aligned to hg19 using bowtie2.<sup>87</sup> All in-silico digested sub-sequence pairs were assessed and the corresponding mapping quality reported, together with the genomic location of the pair ends, within both flashed and non-flashed read set independently. Any pair containing sub-reads shorter than a minimum 20 base pairs were discarded. All pairs that passed



this stage were then considered for downstream analysis. Duplicated pairs were enumerated and excluded using samtools.<sup>113</sup> Pairs containing low quality mapping reads (Phred Quality Score <30) on either end were also excluded. De-duplicated valid paired-end reads were then considered for bona-fide Capture-C filtering, using an in-house algorithm, ICeCAP (see [key resources table](#)). In brief, after flashing and aligning reads and after in-silico digestion, artefacts, e.g. self-ligation and re-ligation events were identified, enumerated and excluded based on the principles of the HiCUP pipeline. Remaining fragments were allocated to each bait and grouped in promoter specific and enhancer specific bona-fide ditags for statistical analysis. ICeCAP combines the principles of fast length adjustment (FLASH) and bona-fide Capture-HiC filtering principles of the HiCUP pipeline. The significance of interaction frequencies was assessed using PeakY<sup>96</sup>. CaptureCompare<sup>79</sup> was used to compare AS and HC scores and to generate mean tracks for visualisation. Median PeakY interaction scores were calculated from the top 10% of scores across the region of interest, which were defined based on ChromHMM enhancer/promoter delineation. *ETS1* promoter region chr11:128310000-128410000, *ETS1* SNP region chr11:128150000-128185000. *PTGER4* promoter region chr5:40650000-40690000, rs1992661 SNP region, chr5:40370000-40420000, rs9283753 SNP region chr5:40480000-40505000.

### Genetic analysis and imputation

Data QC was performed by the Oxford Genomics Centre using Genome Studio v2.0 (Illumina, Human genome reference assembly GRCh37/hg19). To impute B27 from the genotyping data, we used snp2hla on chromosome 6 of genotyped data using the T1DGC reference panel, as described in Jia et al.<sup>114</sup> For genome wide imputation SNPs were removed if they had a minor allele frequency (MAF) < 5%, missingness >2% or a heterozygosity rate greater than 3 standard deviations from the mean. We also removed SNPs if they deviated from Hardy-Weinberg Equilibrium ( $p < 1 \times 10^{-7}$ ). Imputation was performed on the Sanger Imputation Server with the HRC reference panel<sup>115</sup> using Eagle2<sup>81</sup> and PBWT.<sup>82</sup> After imputation, we kept SNPs with an info score >0.8, MAF >0.05 and in Hardy Weinberg equilibrium.

### Therapeutic target prioritisation

Using our previously established Pi pipeline,<sup>33</sup> we first prepared three types of genomic predictors taking as inputs GWAS summary data in AS<sup>9</sup> and the knowledge of protein interactions<sup>116</sup> including: (i) nearby genes (the *nGene* predictor) using genomic proximity and organisation<sup>117</sup> (ii) expression-associated genes (the *eGene* predictor) integrating eQTL datasets,<sup>118–122</sup> and (iii) conformation genes (the *cGene predictor*) using PCHI-C datasets.<sup>38</sup> Next, we modified the Pi algorithm to prepare additional five types of genomic predictors using disease-specific functional genomic datasets arising from this study, including (iv) the AS-specific expression predictor (the *RNA* predictor) generated based on differential gene expression between AS patients and HC, and (v) AS-specific epigenomic predictors (the *ATAC*, *eRNA*, *H3K4me3* and *H3K27ac* predictors) on the basis of differential peaks (linked to genes via genomic proximity or PCHI-C looping events) between AS patients and HC. The knowledge of protein interactions was obtained from the STRING database,<sup>116</sup> and only the genes located in chromosomes 1-22 were considered; this corresponded to a total of 17,249 genes prioritised (and ranked by priority rating). The performance benchmarking (including how to define clinical proof-of-concept targets and simulate negative targets) and comparisons with naïve prioritisation and Open Targets<sup>49,50</sup> were the same as previously described<sup>33,34</sup> and uses Area Under the Curve (AUC) as a global measure of performance. We carried out pathway enrichment analysis of the top 1% prioritised genes using KEGG pathways<sup>123</sup> with the enrichments measured by Z-score and FDR (one-sided Fisher's exact test). Using the algorithm originally proposed in dnet,<sup>98</sup> we performed pathway crosstalk analysis to identify a subset of gene interactions (merged from KEGG pathways) that contained highly prioritised and interconnecting genes.

### Data visualisation

Results from PCA, gene expression, pathway analysis and GWAS enrichment were plotted with ggplot2. Volcano plots were generated with EnhancedVolcano. The Venn diagram showing cell type specificity of differentially expressed genes was generated using Venny. Colour mapping of gene expression (by fold-change) in the CXC Cytokine Receptor pathway was performed using KEGG mapper – color pathway.<sup>100</sup>

Genomic data including location of ENSEMBL genes, ChromHMM data, RNA-seq, ATAC-seq data, ChIPm data, PCHIc data, and Capture-C interactions were plotted in R (<https://www.R-project.org/>) using Gviz<sup>102</sup> and GenomicInteractions.<sup>101</sup> For visualisation RNA-seq was plotted in bins of 1000bp using the log<sub>2</sub> of the count. ATAC-seq and ChIPm-seq were visualised with sliding windows of 200bp and 400bp respectively. Capture-C tracks are plotted as mean counts from three AS patients and three HC samples in bins of 400bp, with the PeakY interaction score (-log<sub>10</sub>(FDR)) plotted below each pile-up track with a sliding window of 400bp. PCHIc interactions were plotted with a CHiCAGO<sup>124</sup> score of ≥5. For ATAC-seq and ChIPm tracks, data from one representative AS patient and one HC are shown. For ChromHMM data from all four AS patients and controls is shown.

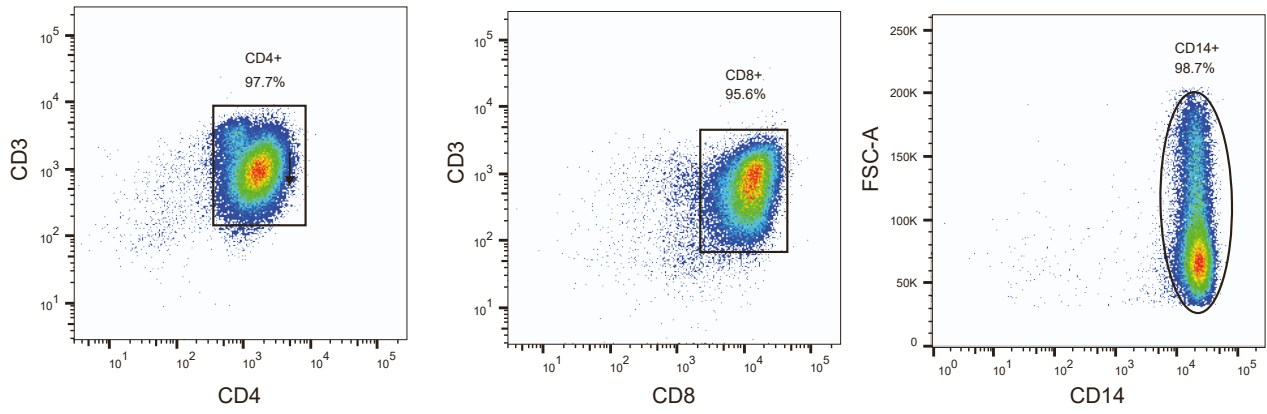
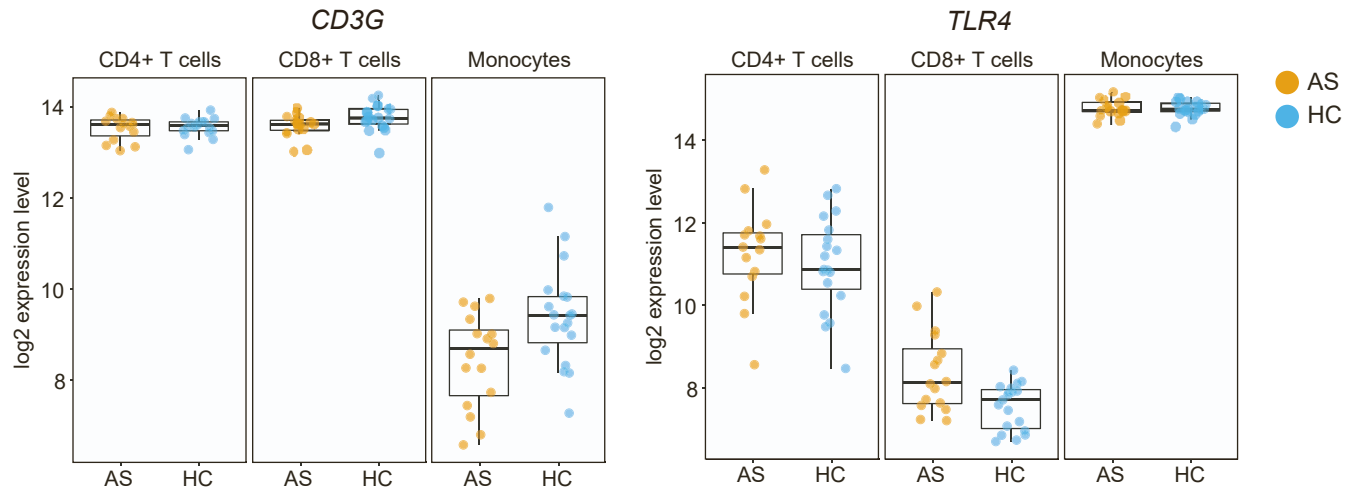
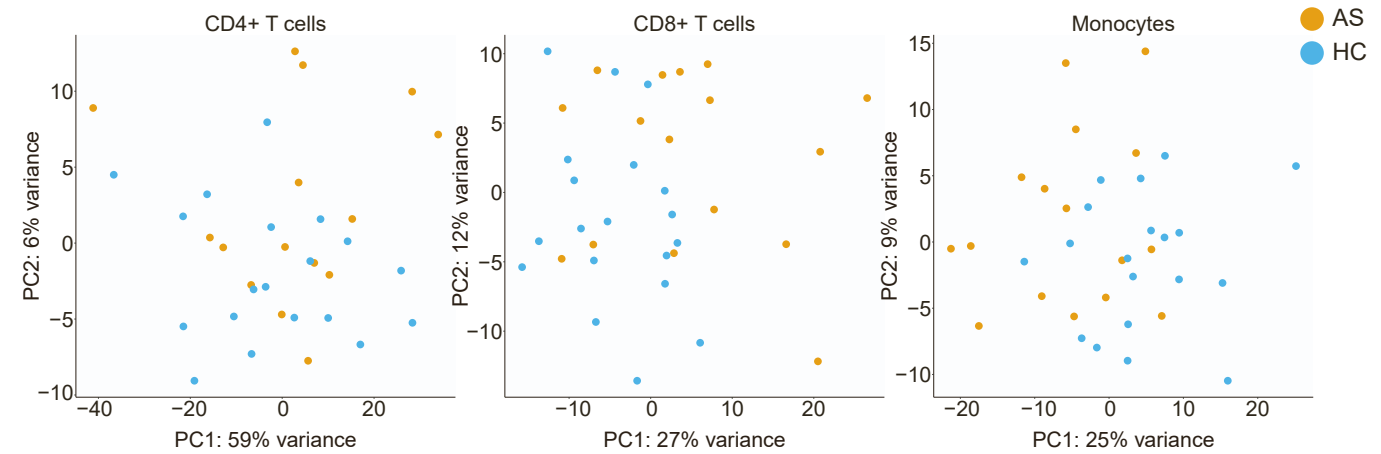
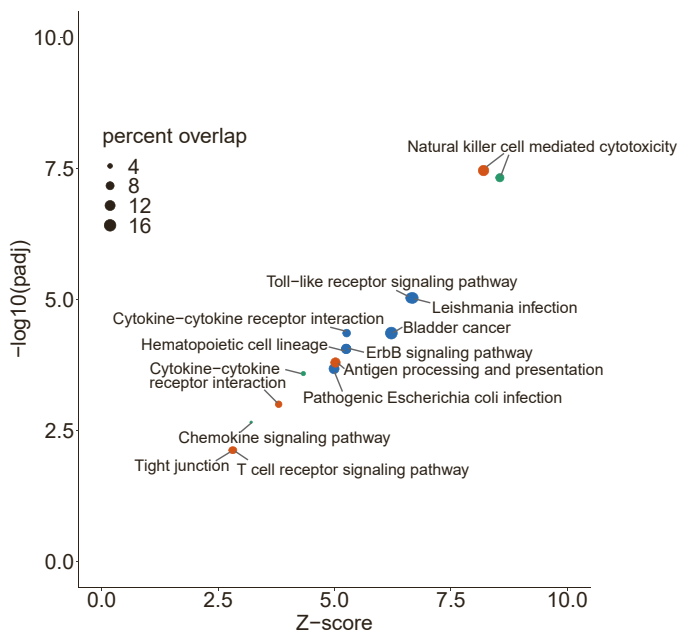
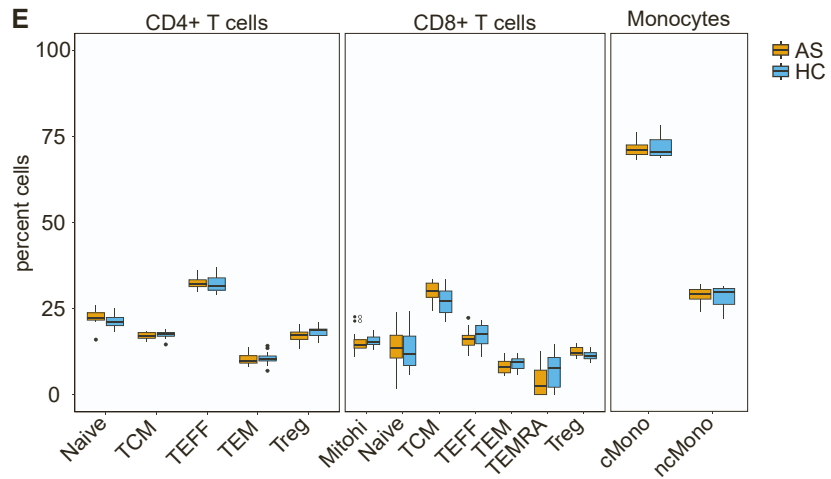


## Supplemental information

### Comprehensive epigenomic profiling reveals

**the extent of disease-specific chromatin states and informs target discovery in ankylosing spondylitis**

**Andrew C. Brown, Carla J. Cohen, Olga Mielczarek, Gabriele Migliorini, F elicie Costantino, Alice Allcock, Connor Davidson, Katherine S. Elliott, Hai Fang, Alicia Lled o Lara, Alice C. Martin, Julie A. Osgood, Anna Sanniti, Giuseppe Scozzafava, Matteo Vecellio, Ping Zhang, Mary Helen Black, Shuwei Li, Dongnhu Truong, Julio Molineros, Trevor Howe, B. Paul Wordsworth, Paul Bowness, and Julian C. Knight**

**A****B****C****D****E**

**Figure S1: Cell-type specific gene expression highlights differential genes and pathways in AS. Related to Figure 1.**

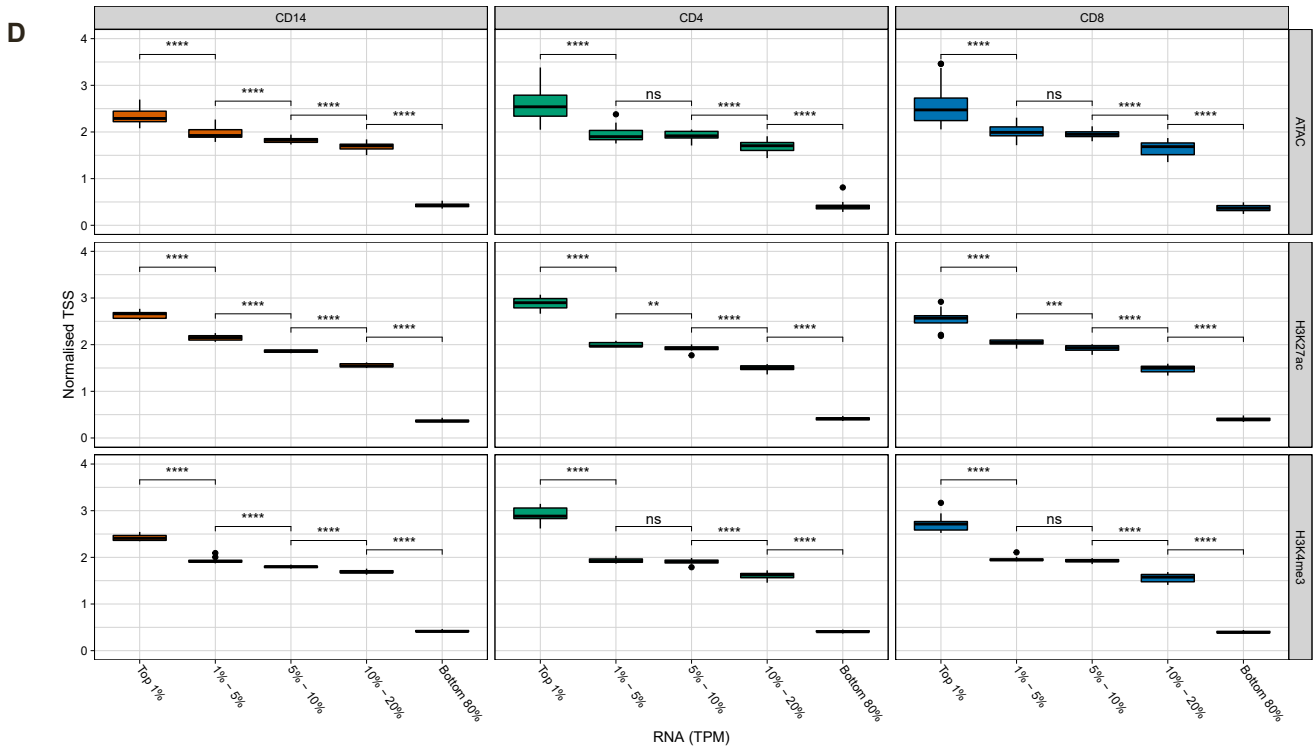
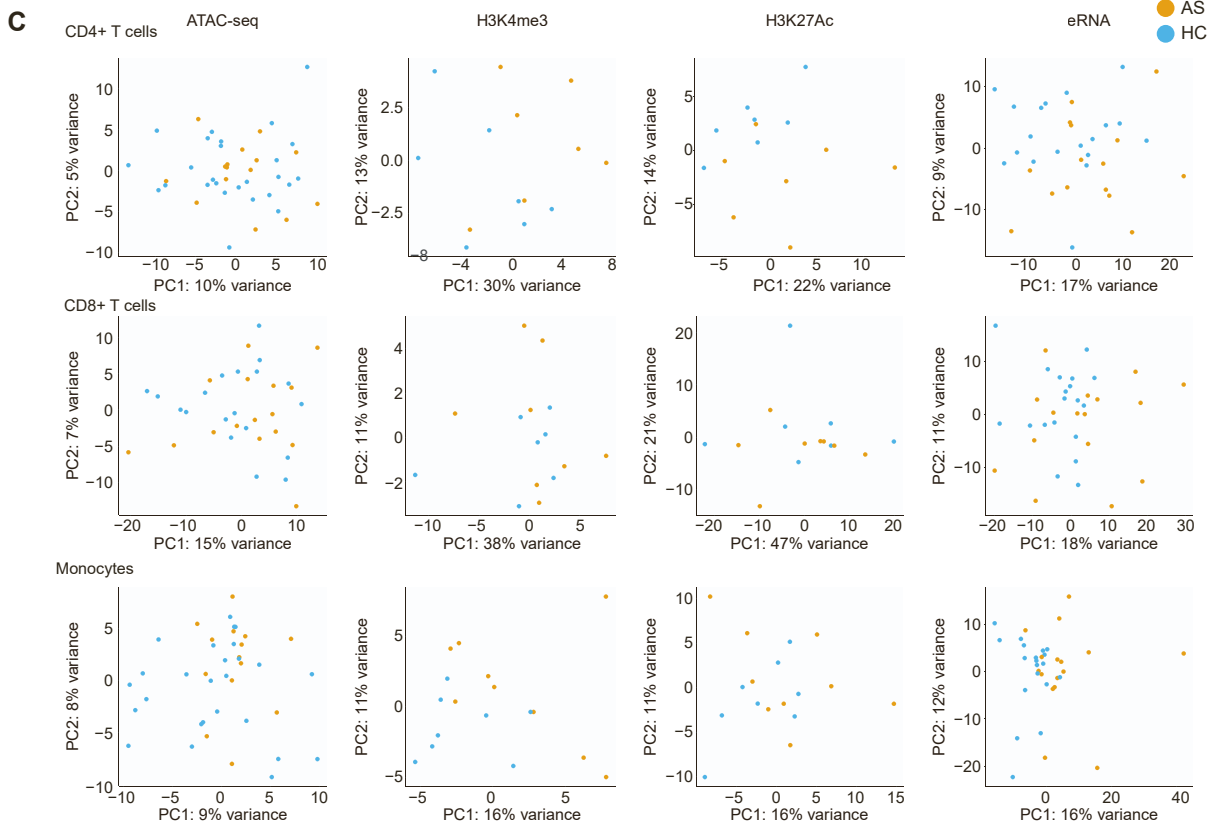
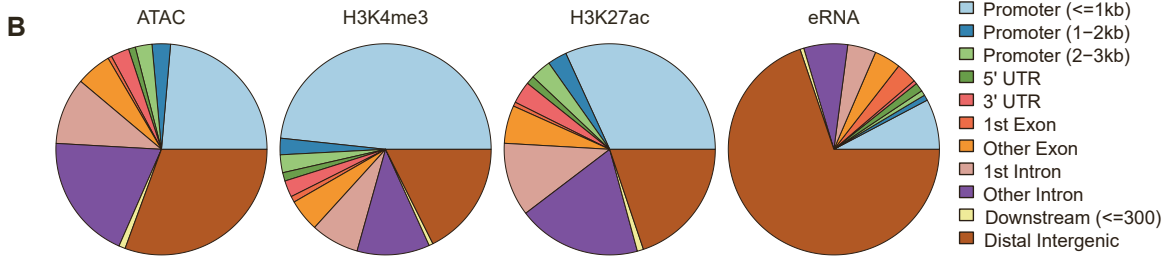
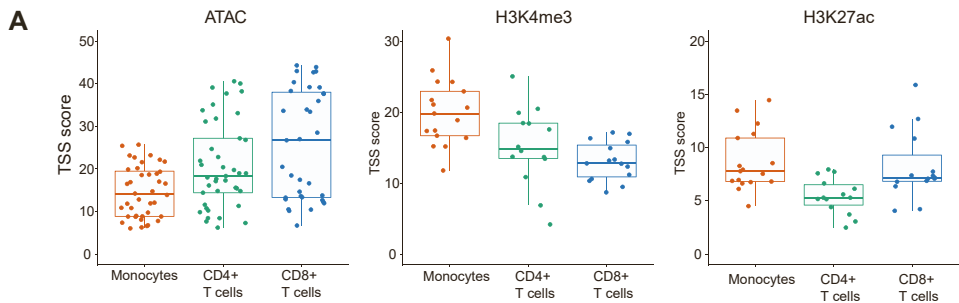
(A) Purity of CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells and monocytes following isolation from PBMC. Cells were stained with antibodies against CD3 in combination with CD4 (left), CD8 (centre) or CD14 (right) and analysed by flow cytometry to demonstrate >95% purity of isolated cell populations.

(B) Examples of cell type-specific gene expression from RNA-seq analysis (TLR4 highly expressed in monocytes and CD3G in T cells).

(C) PCA showing distribution of gene expression in AS patients and HC in each cell type.

(D) Enriched pathways in KEGG database ( $p_{adj} < 0.01$  from XGR output) from significant differentially expressed genes in each cell type. Size of dot represents percentage of genes represented in that pathway, colours represent cell types.

(E) Bulk cell population composition of CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells and monocytes determined via deconvolution of bulk RNA-seq data using CIBERSORTx. T central memory, TCM; T effector, TEFF; T effector memory, TEM; Regulatory T cell, Treg; High mitochondrial content, mitohi; T effector memory expressing CD45RA, TEMRA; classical monocytes, cMono; non-classical monocytes, ncMono.



**Figure S2: Epigenomic marks are cell-type specific and have expected genomic distribution. Related to Figure 2.**

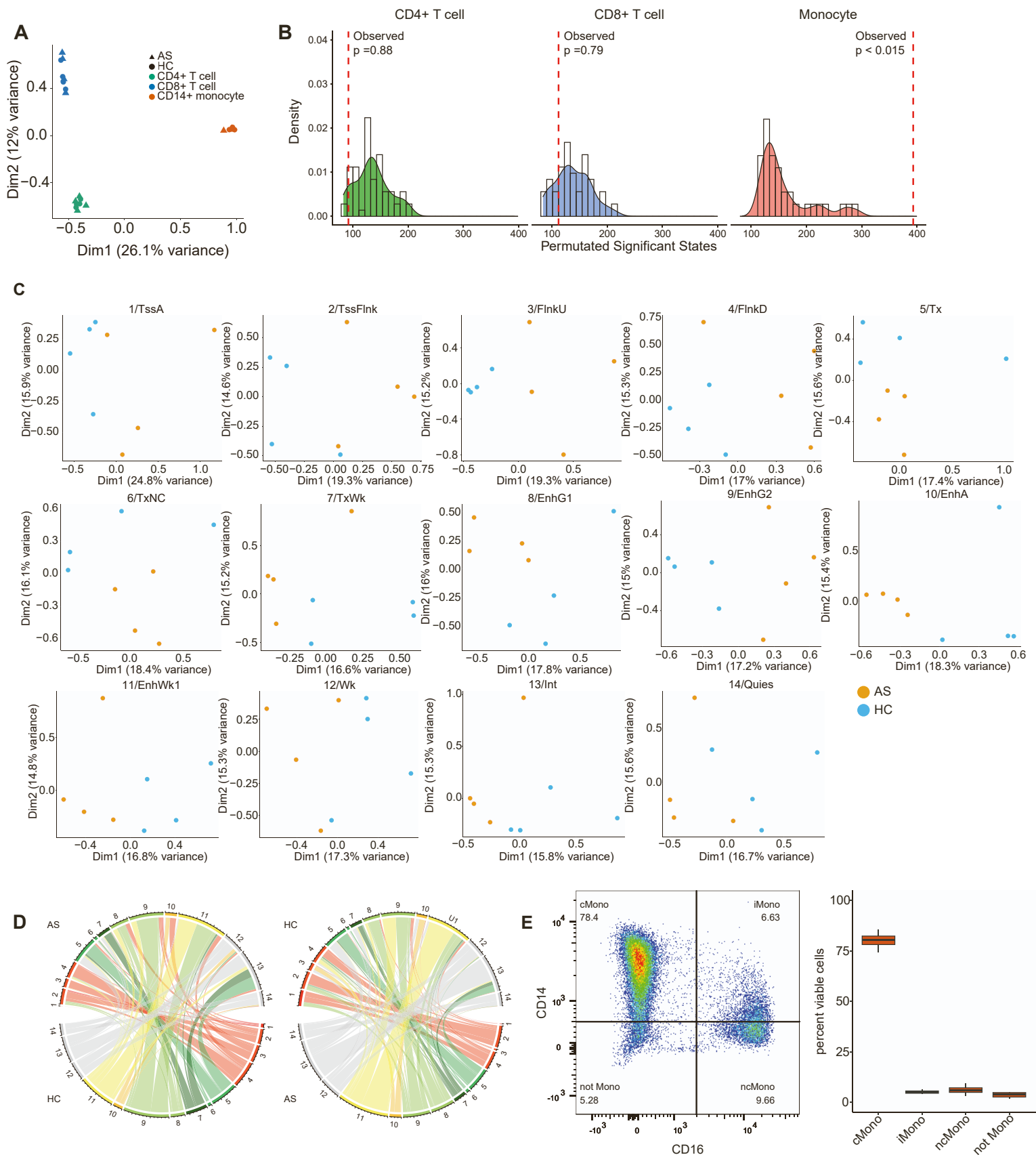
(A) TSS enrichment scores for ATAC-seq and ChIPm (H3K4me3 and H3K27ac) samples in CD4+ T cells, CD8+ T cells and monocytes. Boxes represent the 25th and 75th percentile, the centre line represents the median. Upper and lower whiskers represent values with 1.5x the inter-quartile range.

(B) Genomic distribution of ATAC-seq, ChIP (H3K4me3 and H3K27ac) and eRNA peaks across all three cell types (generated with CHIPseeker).

(C) PCA showing distribution of ATAC-seq, ChIPm (H3K4me3 and H3K27ac) and eRNA peaks in AS patient and HC samples in each cell type (CD4+ T cells, CD8+ T cells and monocytes).

(D) TSS enrichment score of ATAC-seq and ChIPm (H3K4me3 and H3K27ac) peaks per ranked gene expression level of nearest gene. Boxes indicate the 25th and 75th percentile; whiskers indicate 1.5x the inter-quartile range. The centre line represents the median. P values calculated by Student's t-test between groups: \*p < 0.05; \*\*p < 0.01; \*\*\*p < 0.001; \*\*\*\*p < 0.0001.





**Figure S3: ChromHMM signatures are cell-type specific and differentiate AS patients and HC in monocytes. Related to Figure 3.**

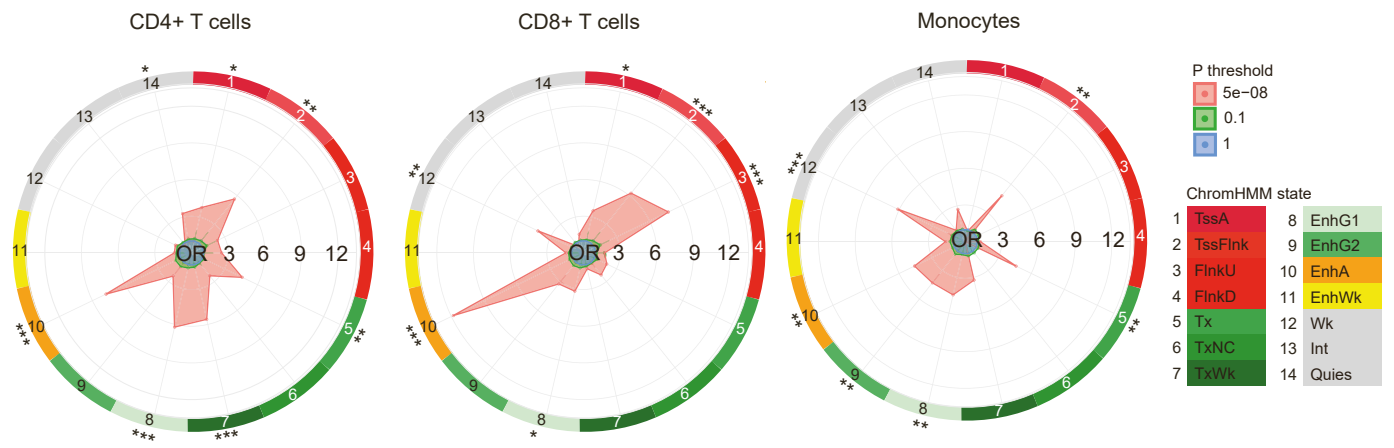
(A) MCA showing distribution of ChromHMM states across three cell types.

(B) Results of permutation analysis of AS-HC differences in ChromHMM state 10 for each cell type. The plot shows the number of significantly different states between AS and HC for each permutation (i.e. density). The experimentally observed frequency is shown with a red dashed line, and p value is indicated for each cell type. Significant p values are also indicated in Figure 3C.

(C) MCA showing distribution of individual ChromHMM states in monocytes, comparing AS patients and HC.

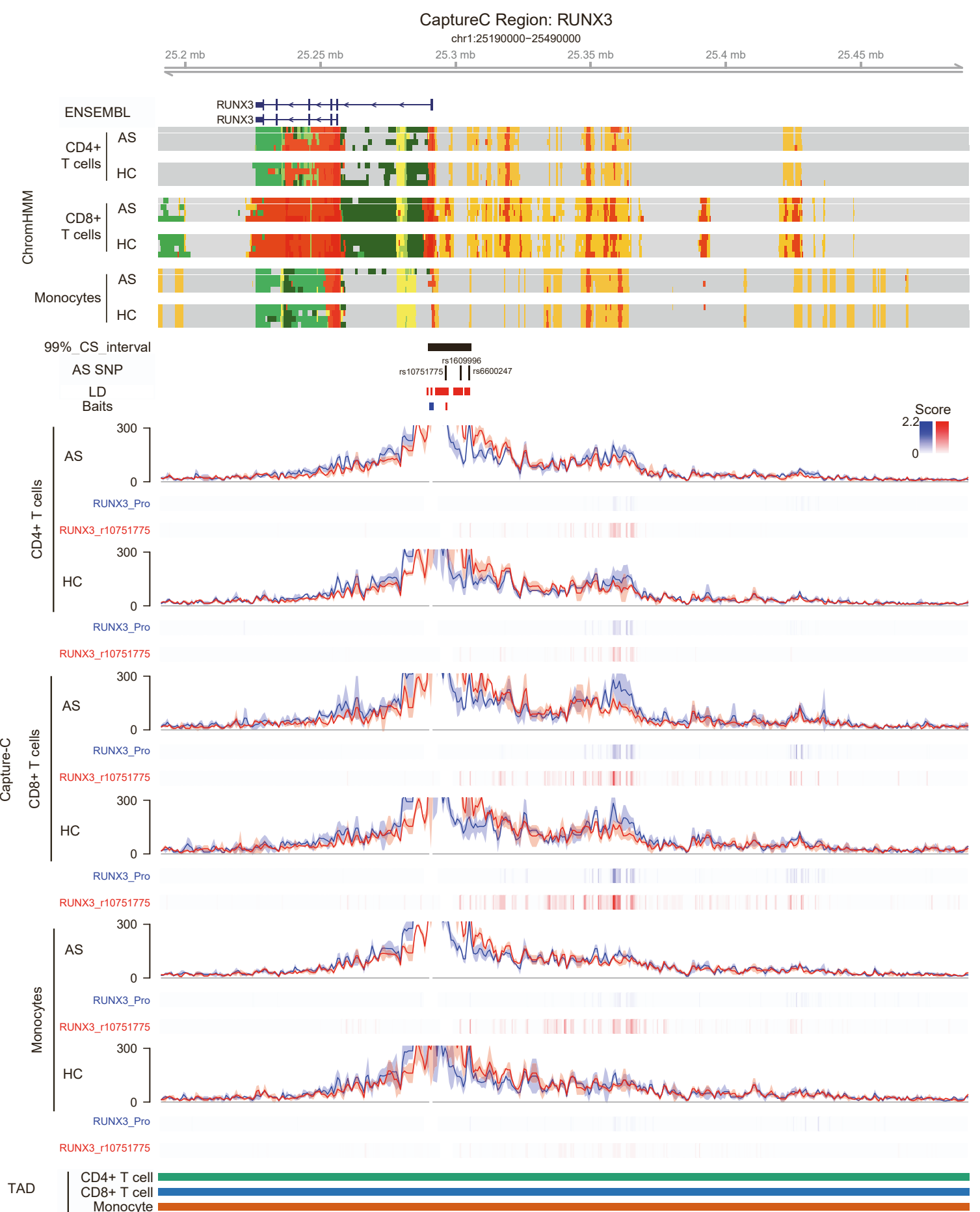
(D) Directionality of state transitions between equivalent genomic fragments in AS patients and HC in monocytes. Circle circumference represents the number of differential ChromHMM fragments (one tick = 200 fragments). Colours are consistent with Figure 3A. Left plot shows fragments present in 4 AS patients (upper half) and the state of that equivalent fragment in HC (bottom half). Right circle shows fragments present in 4 HC (upper half) and the state of the equivalent genomic fragment in AS patients (bottom half).

(E) Monocyte cell composition determined using flow cytometry. Representative plot from one individual (left panel); quantitation



**Figure S4: ChromHMM marks are enriched at AS-associated GWAS loci excluding the MHC in CD4+ T cells, CD8+ T cells and monocytes. Related to Figure 4.**

Associated SNPs were included from Cortes et al9 with association p value thresholds as indicated. ChromHMM state from model in Figure 3. OR, Odds Ratio. \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ .



**Figure S5: Enhancer-gene interactions at the *RUNX3* locus. Related to Figure 5 and Table 1.**

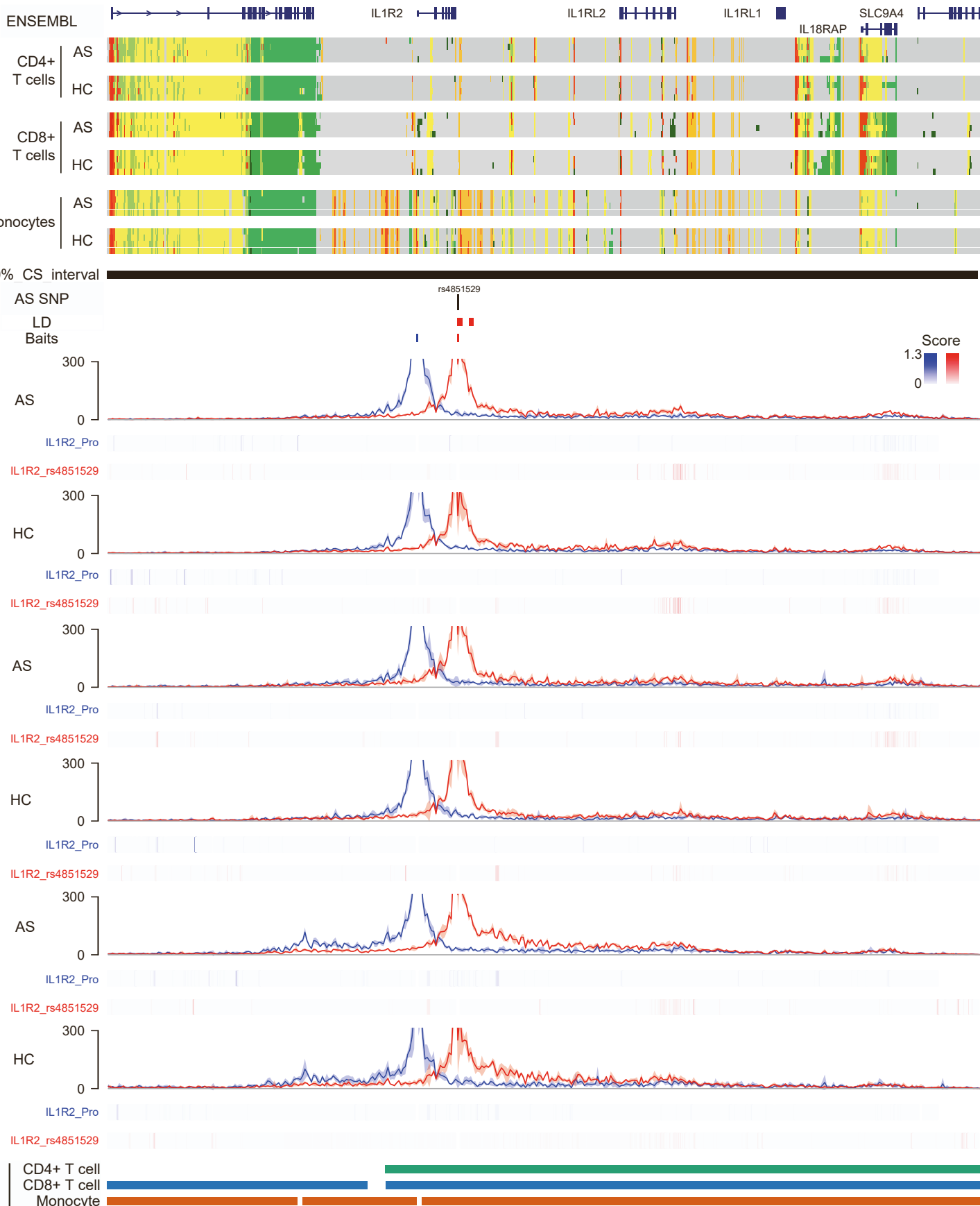
Visualisation of interactions between AS-associated SNPs and the promoter of *RUNX3*.

Ensembl: selected transcripts of Ensembl genes; ChromHMM: data for each cell type from AS patients and HC are shown with colours as in Figure 3B; 99% CS interval: 99% credible set from van de Bunt et al<sup>51</sup>; AS SNP: Position of lead GWAS SNPs from Cortes et al<sup>9</sup> and/or Ellinghaus et al<sup>10</sup> as per Table 1; LD: SNPs in LD ( $r^2 > 0.8$ ) with lead SNP; Baits: Position of Capture-C baits (see also Table S8); Capture-C: Mean interaction count (n=3) and 1 sd shading, with PeakY scores coloured according to bait for AS patients and HC in CD4+ T cells, CD8+ T cells and monocytes; TAD: Topologically Associating Domain for each cell type.

### CaptureC Region: IL18RAP

chr2:102310000-103150000

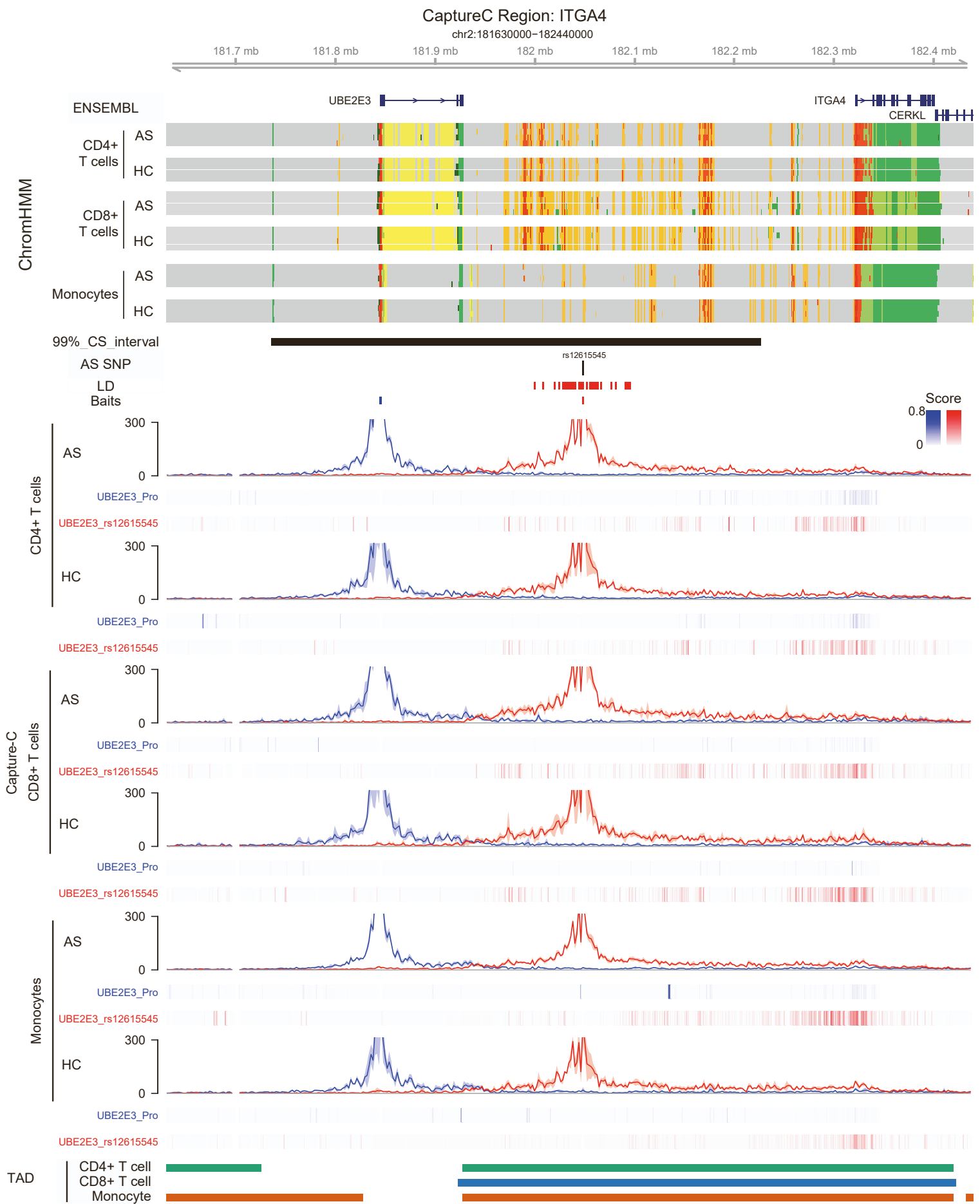
102.4 mb 102.5 mb 102.6 mb 102.7 mb 102.8 mb 102.9 mb 103 mb 103.1 mb



**Figure S6: Enhancer-gene interactions at the *IL18RAP* locus. Related to Figure 5 and Table 1.**

Visualisation of interactions between AS-associated SNPs and the promoter of *IL1RL2* and *IL18RAP*.

Ensembl: selected transcripts of Ensembl genes; ChromHMM: data for each cell type from AS patients and HC are shown with colours as in Figure 3B; 99% CS interval: 99% credible set from van de Bunt et al<sup>51</sup>; AS SNP: Position of lead GWAS SNPs from Cortes et al<sup>9</sup> and/or Ellinghaus et al<sup>10</sup> as per Table 1; LD: SNPs in LD ( $r^2 > 0.8$ ) with lead SNP; Baits: Position of Capture-C baits (see also Table S8); Capture-C: Mean interaction count (n=3) and 1 sd shading, with PeakY scores coloured according to bait for AS patients and HC in CD4+ T cells, CD8+ T cells and monocytes; TAD: Topologically Associating Domain for each cell type.



**Figure S7: Enhancer-gene interactions at the *ITGA4* locus. Related to Figure 5 and Table 1.**

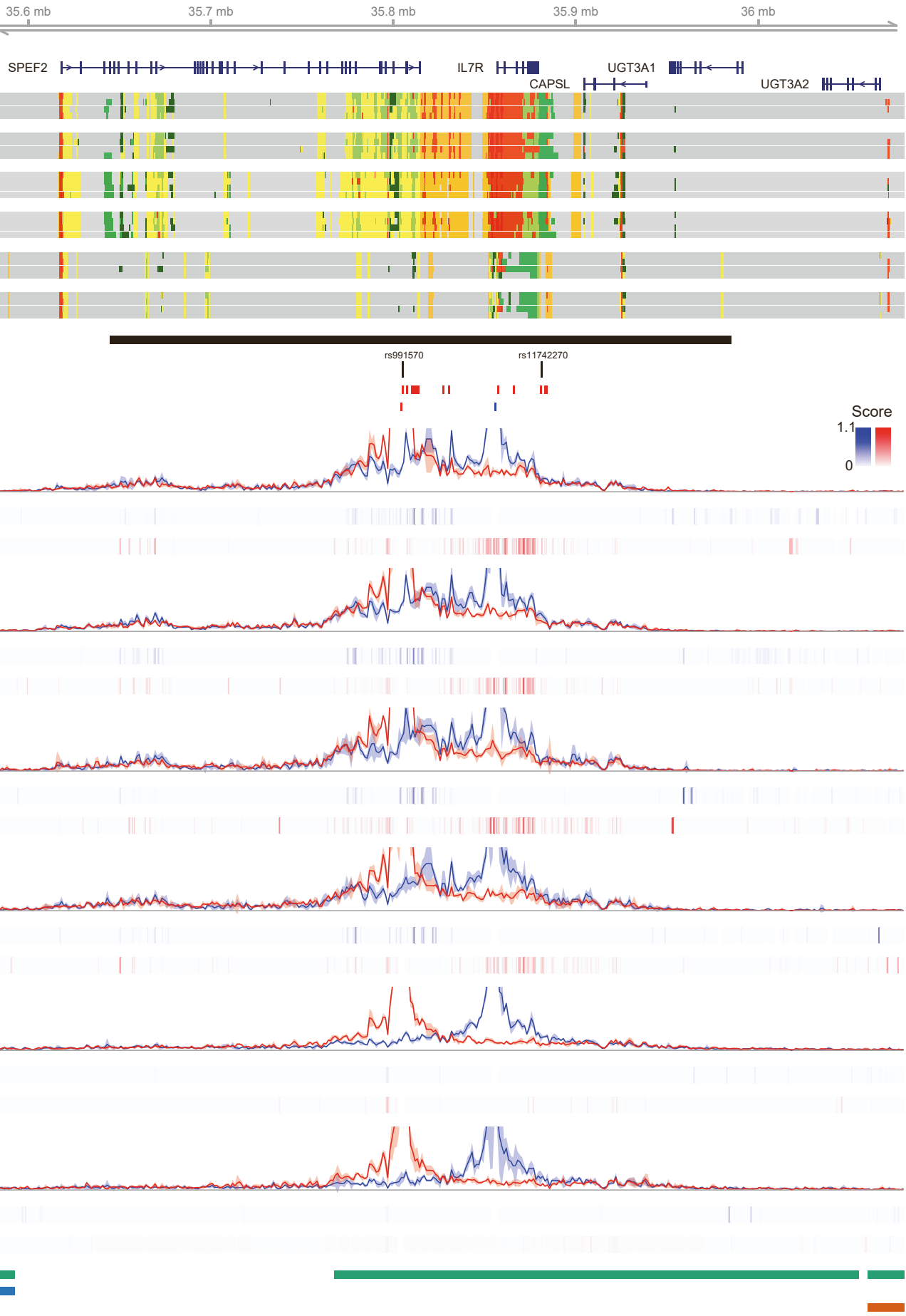
Visualisation of interactions between AS-associated SNPs and the promoter of *ITGA4*.

Ensembl: selected transcripts of Ensembl genes; ChromHMM: data for each cell type from AS patients and HC are shown with colours as in Figure 3B; 99% CS interval: 99% credible set from van de Bunt et al<sup>51</sup>; AS SNP: Position of lead GWAS SNPs from Cortes et al<sup>9</sup> and/or Ellinghaus et al<sup>10</sup> as per Table 1; LD: SNPs in LD ( $r^2 > 0.8$ ) with lead SNP; Baits: Position of Capture-C baits (see also Table S8); Capture-C: Mean interaction count (n=3) and 1 sd shading, with PeakY scores coloured according to bait for AS patients and HC in CD4+ T cells, CD8+ T cells and monocytes; TAD: Topologically Associating Domain for each cell type.



### CaptureC Region: IL7R

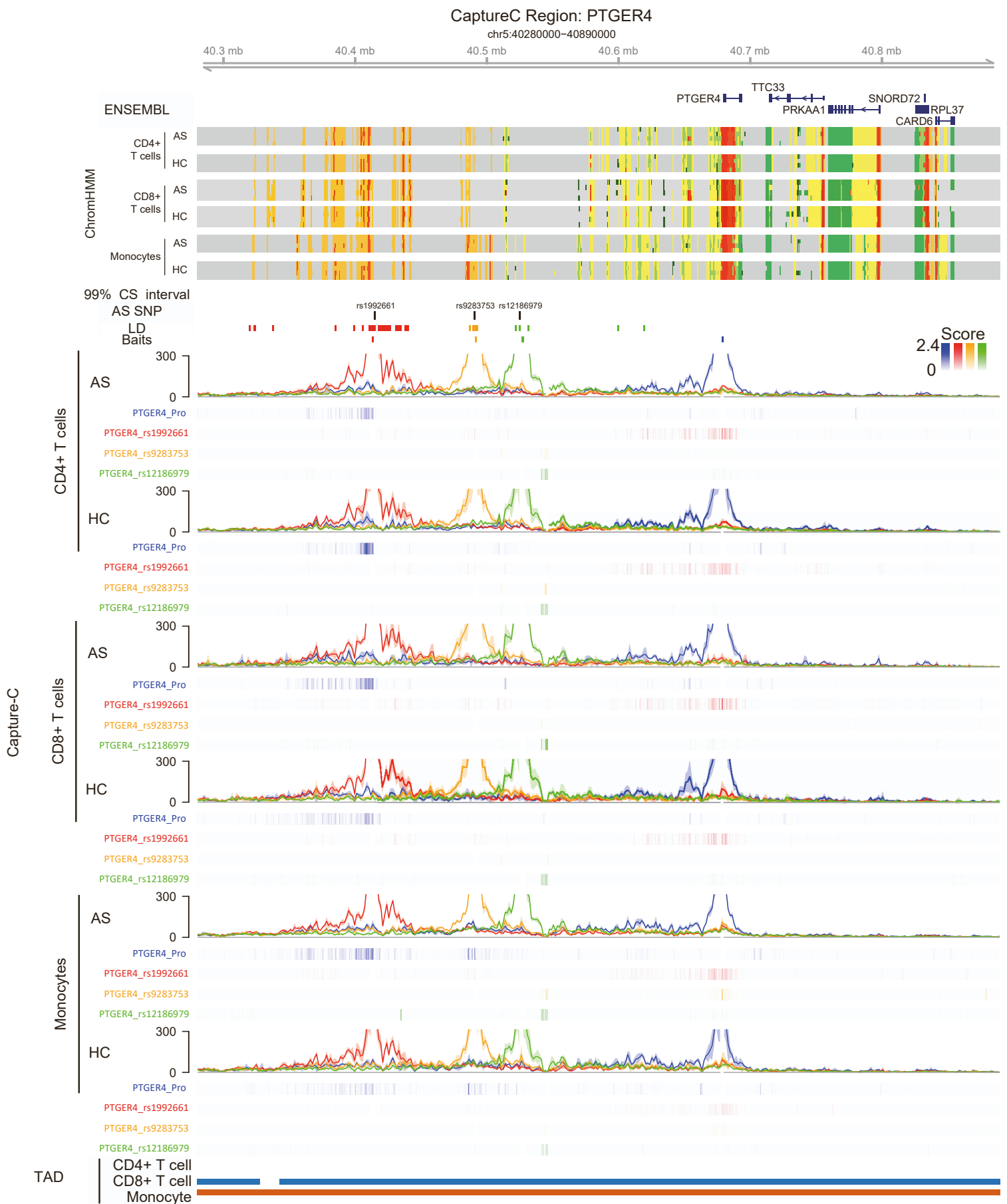
chr5:35580000-36080000



**Figure S8: Enhancer-gene interactions at the *IL7R* locus. Related to Figure 5 and Table 1.**

Visualisation of interactions between AS-associated SNPs and the promoter of *IL7R*.

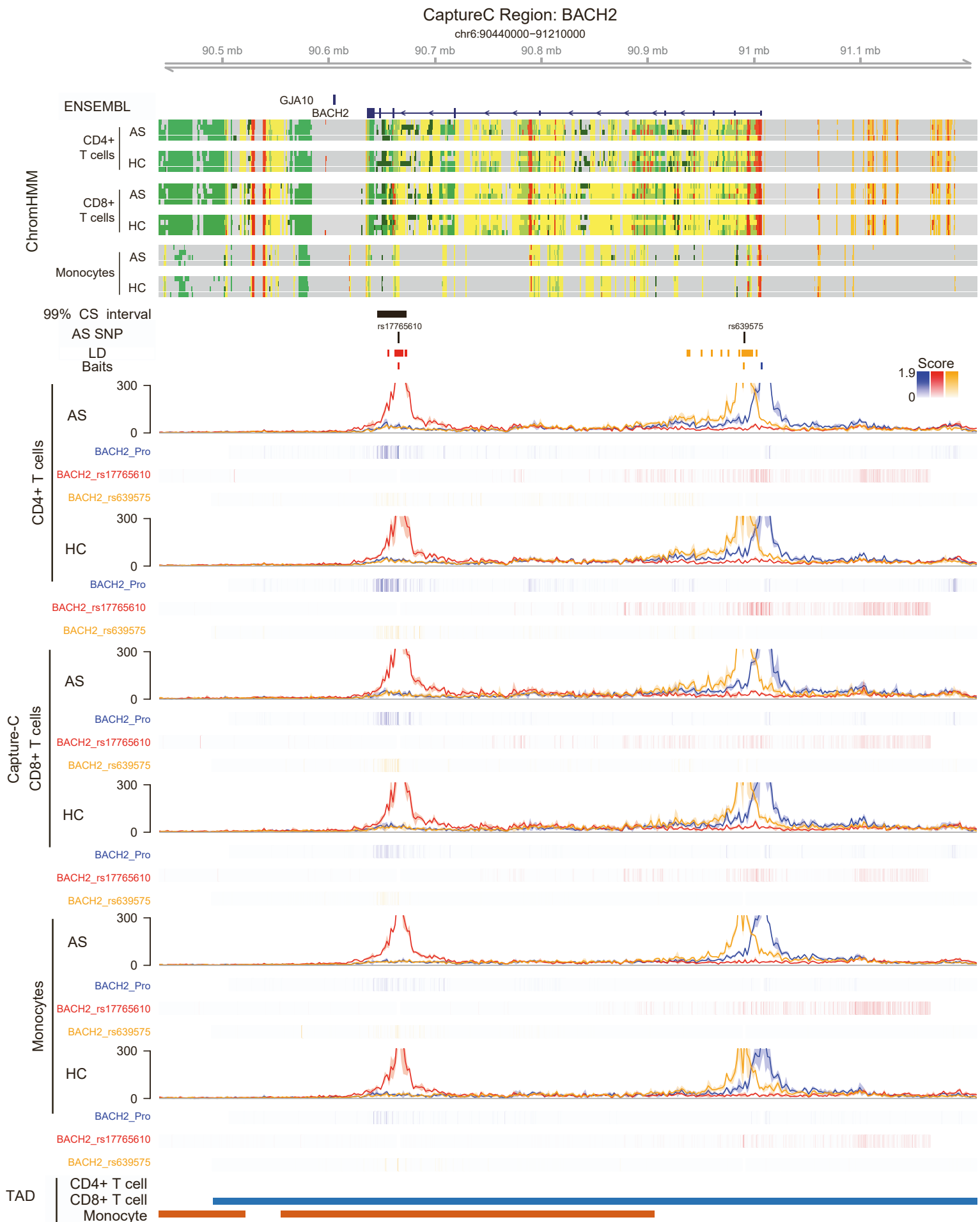
Ensembl: selected transcripts of Ensembl genes; ChromHMM: data for each cell type from AS patients and HC are shown with colours as in Figure 3B; 99% CS interval: 99% credible set from van de Bunt et al<sup>51</sup>; AS SNP: Position of lead GWAS SNPs from Cortes et al<sup>9</sup> and/or Ellinghaus et al<sup>10</sup> as per Table 1; LD: SNPs in LD ( $r^2 > 0.8$ ) with lead SNP; Baits: Position of Capture-C baits (see also Table S8); Capture-C: Mean interaction count (n=3) and 1 sd shading, with PeakY scores coloured according to bait for AS patients and HC in CD4+ T cells, CD8+ T cells and monocytes; TAD: Topologically Associating Domain for each cell type.



**Figure S9: Enhancer-gene interactions at the *PTGER4* locus. Related to Figure 5 and Table 1.**

Visualisation of interactions between AS-associated SNPs and the promoter of *PTGER4*.

Ensembl: selected transcripts of Ensembl genes; ChromHMM: data for each cell type from AS patients and HC are shown with colours as in Figure 3B; 99% CS interval: 99% credible set from van de Bunt et al<sup>51</sup>; AS SNP: Position of lead GWAS SNPs from Cortes et al<sup>9</sup> and/or Ellinghaus et al<sup>10</sup> as per Table 1; LD: SNPs in LD ( $r^2 > 0.8$ ) with lead SNP; Baits: Position of Capture-C baits (see also Table S8); Capture-C: Mean interaction count ( $n=3$ ) and 1 sd shading, with PeakY scores coloured according to bait for AS patients and HC in CD4+ T cells, CD8+ T cells and monocytes; TAD: Topologically Associating Domain for each cell type.



**Figure S10: Enhancer-gene interactions at the *BACH2* locus. Related to Figure 5 and Table 1.**

Visualisation of interactions between AS-associated SNPs and the promoter of *BACH2*.

Ensembl: selected transcripts of Ensembl genes; ChromHMM: data for each cell type from AS patients and HC are shown with colours as in Figure 3B; 99% CS interval: 99% credible set from van de Bunt et al<sup>9</sup>; Lead SNP: Position of lead GWAS SNPs from Cortes et al<sup>9</sup> and/or Ellinghaus et al<sup>10</sup> as per Table 1; LD: SNPs in LD ( $r^2 > 0.8$ ) with lead SNP; Baits: Position of Capture-C baits (see also Table S8); Capture-C: Mean interaction count ( $n=3$ ) and 1 sd shading, with PeakY scores coloured according to bait for AS patients and HC in CD4+ T cells, CD8+ T cells and monocytes; TAD: Topologically Associating Domain for each cell type.

### CaptureC Region: ETS1

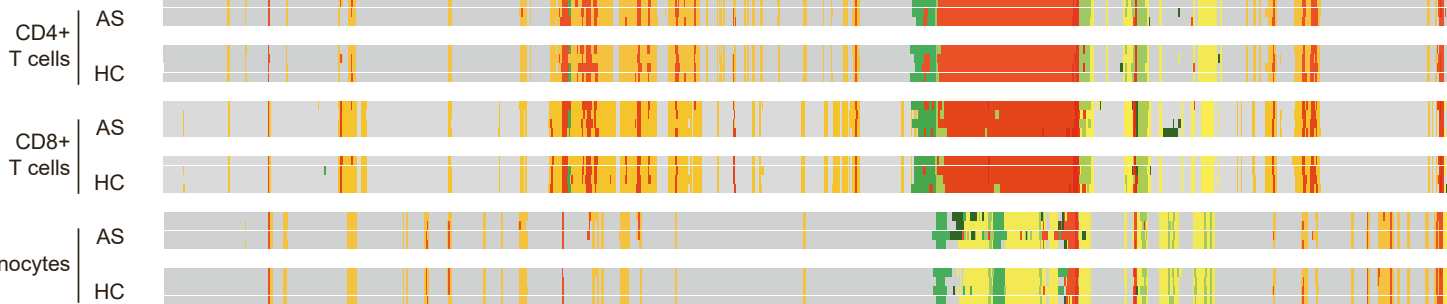
chr11:127980000-128560000



ENSEMBL

ETS1

ChromHMM



99% CS interval

AS SNP

rs7933433

LD

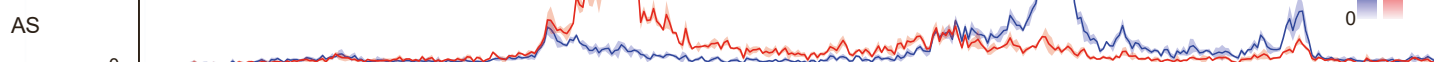
Baits

Score

5.2

0

CD4+ T cells



ETS1\_Pro

ETS1\_rs7933433



ETS1\_Pro

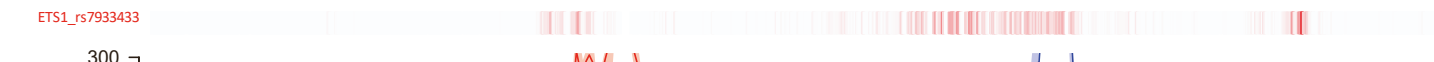
ETS1\_rs7933433

CD8+ T cells



ETS1\_Pro

ETS1\_rs7933433



ETS1\_Pro

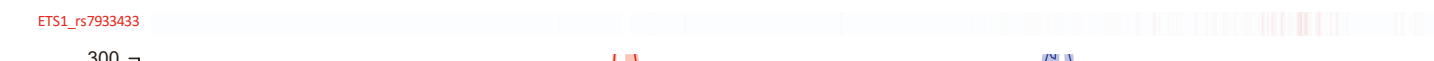
ETS1\_rs7933433

Monocytes



ETS1\_Pro

ETS1\_rs7933433



ETS1\_Pro

ETS1\_rs7933433

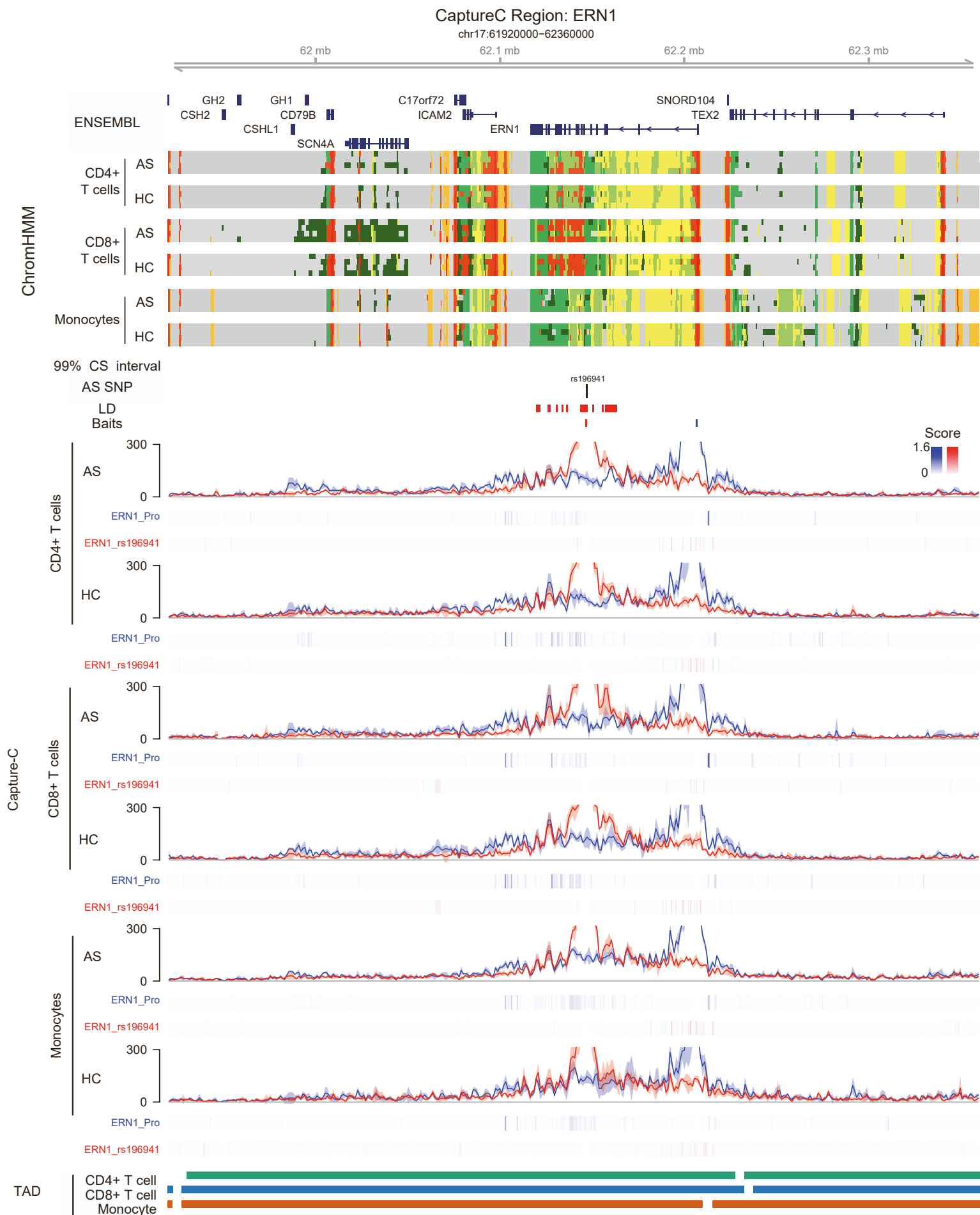
CD4 TAD  
CD8 TAD  
CD14 TAD



**Figure S11: Enhancer-gene interactions at the *ETS1* locus. Related to Figure 5 and Table 1.**

Visualisation of interactions between AS-associated SNPs and the promoter of *ETS1*.

Ensembl: selected transcripts of Ensembl genes; ChromHMM: data for each cell type from AS patients and HC are shown with colours as in Figure 3B; 99% CS interval: 99% credible set from van de Bunt et al<sup>51</sup>; AS SNP: Position of lead GWAS SNPs from Cortes et al<sup>9</sup> and/or Ellinghaus et al<sup>10</sup> as per Table 1; LD: SNPs in LD ( $r^2 > 0.8$ ) with lead SNP; Baits: Position of Capture-C baits (see also Table S8); Capture-C: Mean interaction count ( $n=3$ ) and 1 sd shading, with PeakY scores coloured according to bait for AS patients and HC in CD4+ T cells, CD8+ T cells and monocytes; TAD: Topologically Associating Domain for each cell type.



**Figure S12: Enhancer-gene interactions at the *ERN1* locus. Related to Figure 5 and Table 1.**

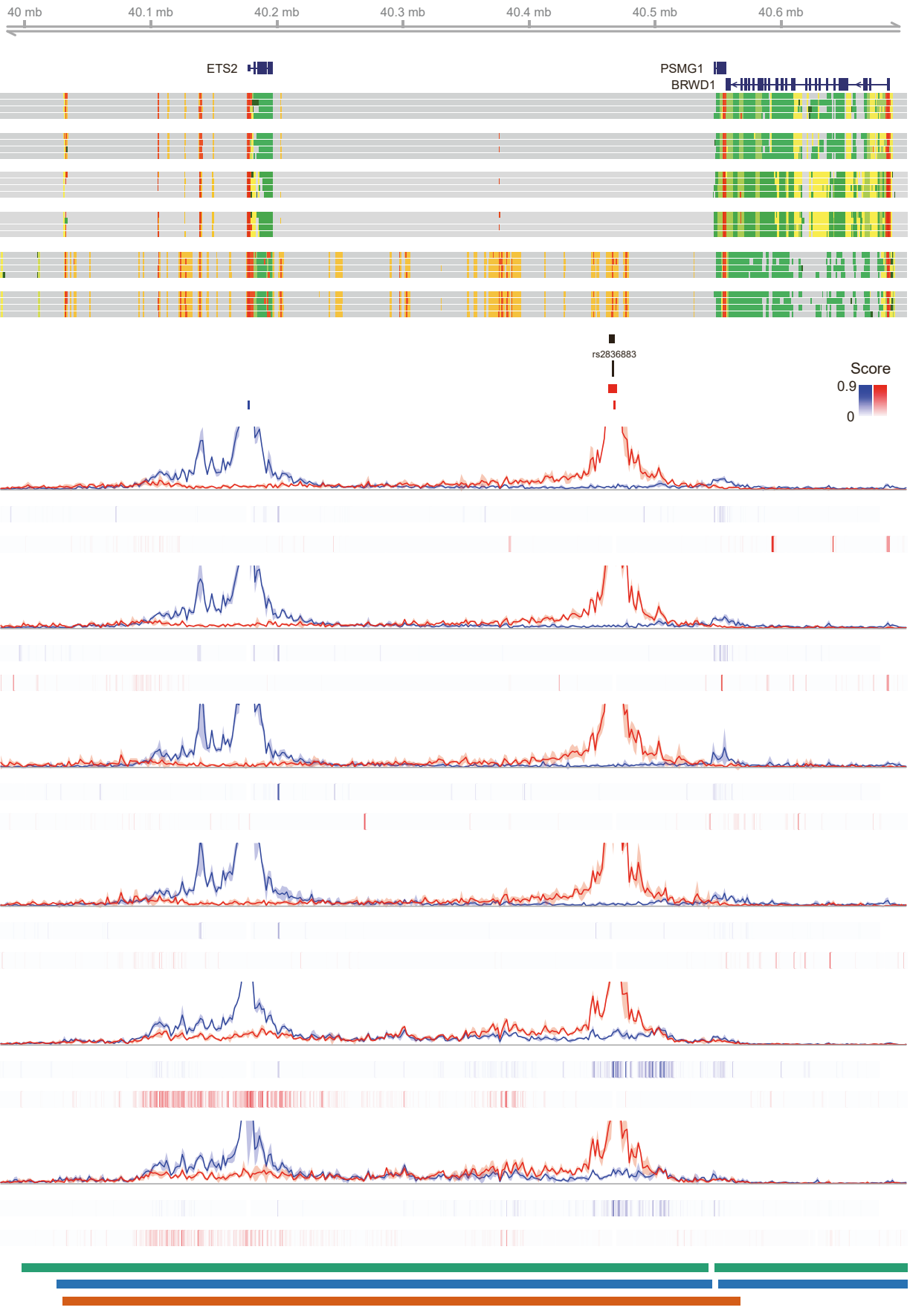
Visualisation of interactions between AS-associated SNPs and the promoter of *ERN1*.

Ensembl: selected transcripts of Ensembl genes; ChromHMM: data for each cell type from AS patients and HC are shown with colours as in Figure 3B; 99% CS interval: 99% credible set from van de Bunt et al<sup>51</sup>; AS SNP: Position of lead GWAS SNPs from Cortes et al<sup>9</sup> and/or Ellinghaus et al<sup>10</sup> as per Table 1; LD: SNPs in LD ( $r^2 > 0.8$ ) with lead SNP; Baits: Position of Capture-C baits (see also Table S8); Capture-C: Mean interaction count ( $n=3$ ) and 1 sd shading, with PeakY scores coloured according to bait for AS patients and HC in CD4+ T cells, CD8+ T cells and monocytes; TAD: Topologically Associating Domain for each cell type.



### CaptureC Region: ETS2

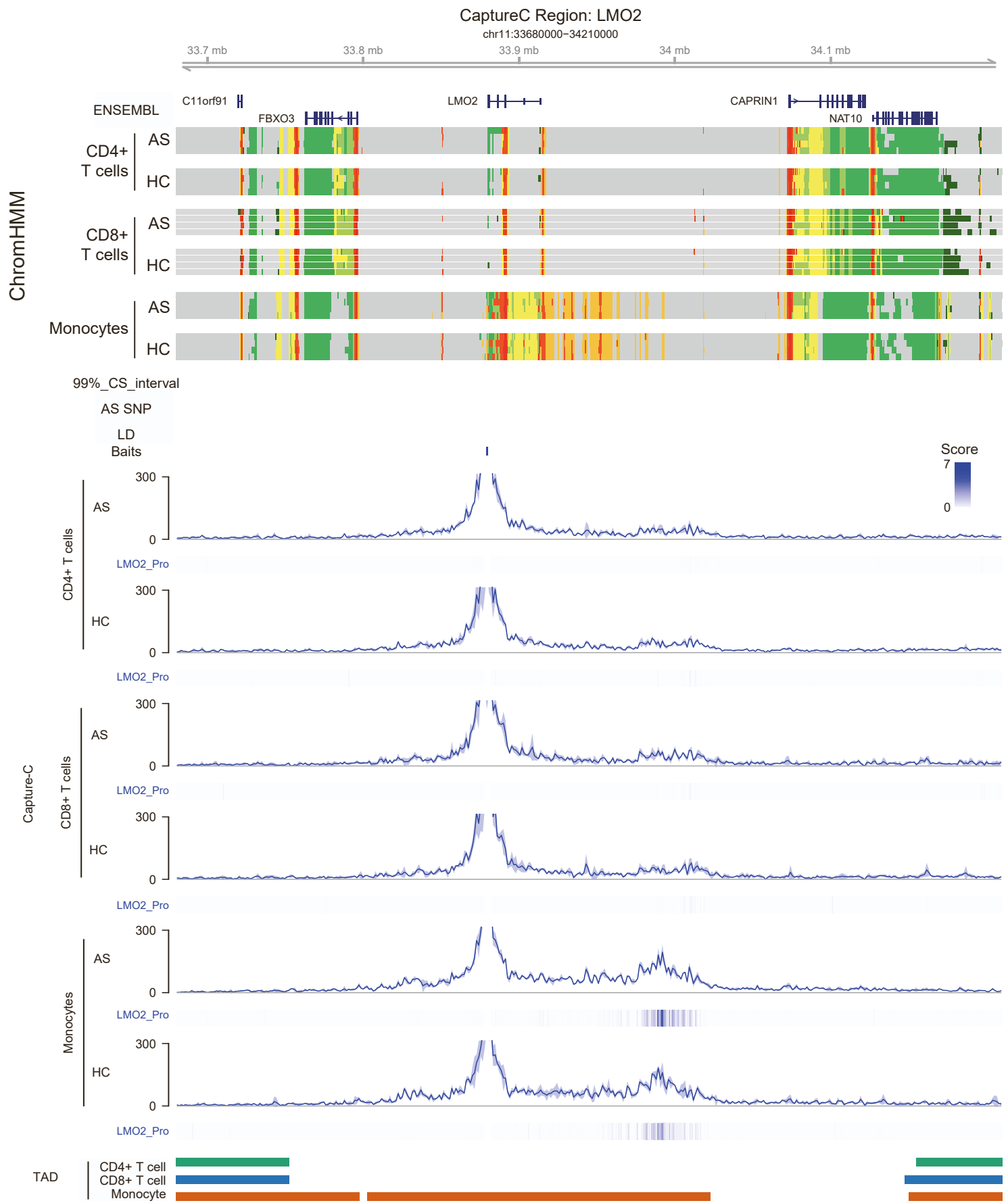
chr21:39980000-40700000



**Figure S13: Enhancer-gene interactions at the *ETS2* locus. Related to Figure 5 and Table 1.**

Visualisation of interactions between AS-associated SNPs and the promoter of *ETS2*.

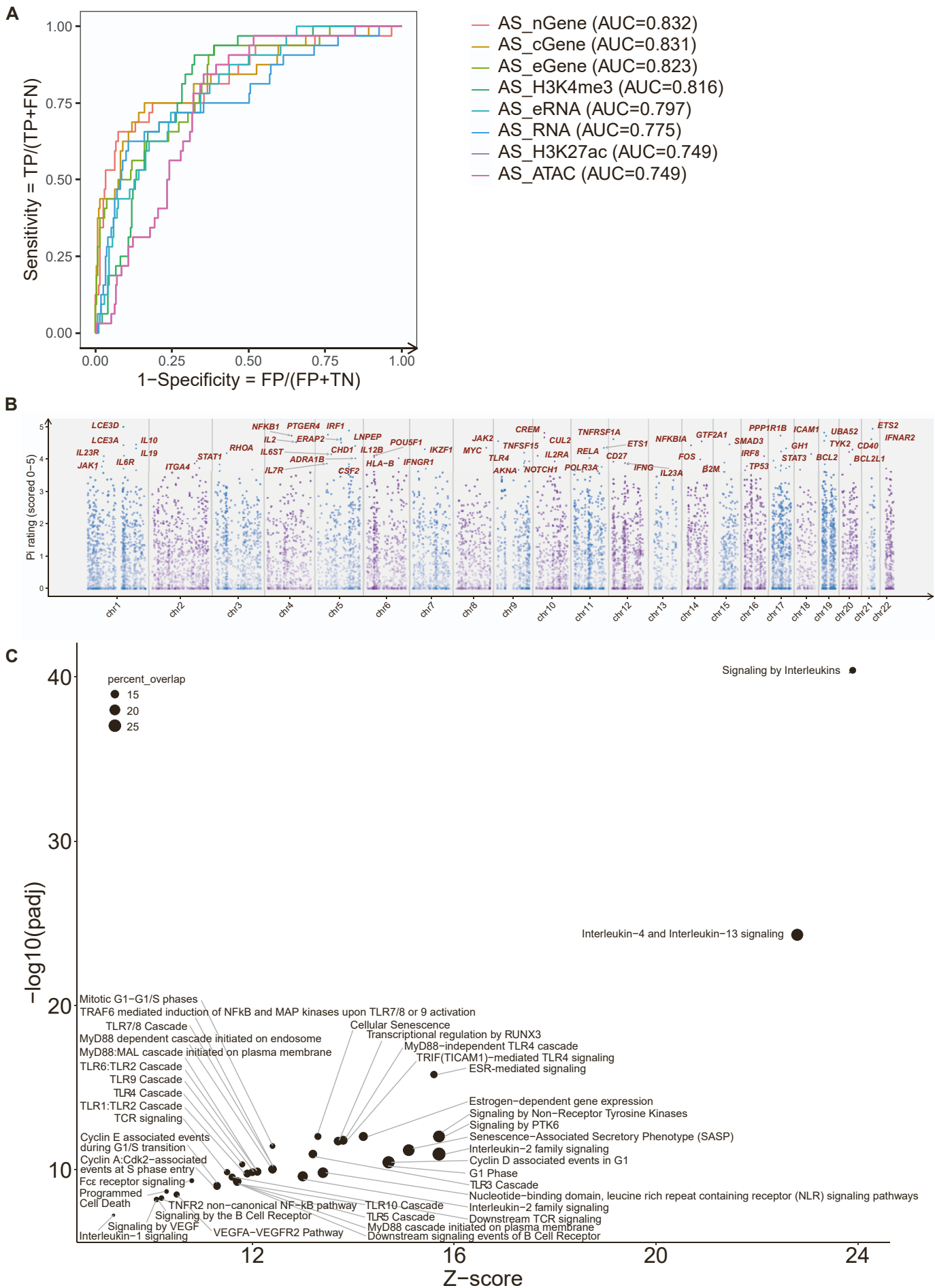
Ensembl: selected transcripts of Ensembl genes; ChromHMM: data for each cell type from AS patients and HC are shown with colours as in Figure 3B; 99% CS interval: 99% credible set from van de Bunt et al<sup>51</sup>; AS SNP: Position of lead GWAS SNPs from Cortes et al<sup>9</sup> and/or Ellinghaus et al<sup>10</sup> as per Table 1; LD: SNPs in LD ( $r^2 > 0.8$ ) with lead SNP; Baits: Position of Capture-C baits (see also Table S8); Capture-C: Mean interaction count ( $n=3$ ) and 1 sd shading, with PeakY scores coloured according to bait for AS patients and HC in CD4+ T cells, CD8+ T cells and monocytes; TAD: Topologically Associating Domain for each cell type.



**Figure S14: Enhancer-gene interactions at the *LMO2* control locus. Related to Figure 5 and Table 1.**

Visualisation of interactions between a known enhancer and the promoter of *LMO2*.

Ensembl: selected transcripts of Ensembl genes; ChromHMM: data for each cell type from AS patients and HC are shown with colours as in Figure 3B; 99% CS interval: 99% credible set from van de Bunt et al<sup>51</sup>; AS SNP: Position of lead GWAS SNPs from Cortes et al<sup>9</sup> and/or Ellinghaus et al<sup>10</sup> as per Table 1; LD: SNPs in LD ( $r^2 > 0.8$ ) with lead SNP; Baits: Position of Capture-C baits (see also Table S8); Capture-C: Mean interaction count ( $n=3$ ) and 1 sd shading, with PeakY scores coloured according to bait for AS patients and HC in CD4+ T cells, CD8+ T cells and monocytes; TAD: Topologically Associating Domain for each cell type.



**Figure S15: Incorporation of disease-specific data improves target discovery. Related to Figure 6.**

(A) Performance of Pi algorithm on the basis of including each predictor type.

(B) Manhattan plot showing top 60 genes ranked by priority rating following inclusion of AS-specific data.

(C) Pathway analysis on the top 1% of Pi ranked genes using the Reactome database ( $p_{adj} < 0.05$ ). Size of dots represents the percentage of overlapping genes in that pathway; only those with  $>10\%$  overlap are shown.

**Table S1: Clinical information for AS patients (AS) and HC. Related to Figure 1 and STAR Methods.**

		<b>AS</b>	<b>HC</b>
Total number		20	35
Age (years)	Mean (range)	45 (28-74)	43 (21-67)
Sex (male)	N (%)	15 (75%)	17 (49%)
HLA-B27 (+)	N (%)	17 (85%)	5 (25%) <sup>b</sup>
Disease duration (months)	Mean (range)	128 (9-456)	N/A
BASDAI (units NRS)	Mean +/- SD	5.9 +/- 1.7	N/A
ASDAS-CRP	Mean +/- SD	3.2 +/- 1.1	N/A
CRP, mg/L	Mean +/- SD	13.5 +/- 22.8	N/A
Axial manifestation of SpA	N (%)	20 (100%)	0 (0%)
Peripheral articular manifestations			
• Peripheral arthritis	N (%)	8 (40%)	0 (0%)
Extra-articular manifestations			
• Psoriasis	N (%)	3 (15%)	0 (0%)
• Uveitis	N (%)	5 (25%)	0 (0%)
• Inflammatory Bowel Disease	N (%)	1 (5%)	0 (0%)
Other comorbidities <sup>a</sup>			
• Cancer	N (%)	1 (5%)	0 (0%)
• Fracture	N (%)	6 (30%)	0 (0%)
• Vertebral fracture	N (%)	1 (5%)	0 (0%)
• Autoimmune Thyroid Disease	N (%)	1 (5%)	0 (0%)
Drugs (current)			
• NSAIDs	N (%)	12 (60%)	0 (0%)
• csDMARDs	N (%)	3 (15%)	0 (0%)
• Glucocorticoids	N (%)	0 (0%)	0 (0%)
Family history			
• AS	N (%)	7 (35%)	Unknown
• Other autoimmune disease	N (%)	7 (35%)	Unknown

<sup>a</sup> No incidences of aortic disease, atlantoaxial subluxation, coeliac disease, dermatomyositis, fibrotic lung disease, haemolytic anaemia, heart block, multiple sclerosis, myasthenia gravis, pemphigus vulgaris, pernicious anaemia, scleroderma, systemic lupus erythematosus, type 1 diabetes, vitiligo.

<sup>b</sup> Not all HC have imputed B27, percentage is calculated on 20 HC that have genotyping. BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; NRS, Numerical Rating Scale; ASDAS, Ankylosing Spondylitis Disease Activity Score; CRP, C-Reactive Protein; SpA,

Spondyloarthropathy; NSAIDs, Non-Steroidal Anti-Inflammatory Drugs; csDMARDs, Conventional Synthetic Disease-Modifying Antirheumatic Drugs.

## Supplemental References

- S1. Cortes A., Hadler J., Cremin K., Pryce K., Harris J., Pointon J.P., Evans D.M., Leo P., Robinson P., Bradbury L.A., et al. (2013). Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nat Genet.* 45, 730–738.
- S2. van de Bunt M., Cortes A., Consortium I., Brown M.A., Morris A.P., McCarthy M.I. (2015). Evaluating the Performance of Fine-Mapping Strategies at Common Variant GWAS Loci. *PLoS Genet.* 11, e1005535. [10.1371/journal.pgen.1005535](https://doi.org/10.1371/journal.pgen.1005535)
- S3. Ellinghaus D., Jostins L., Spain S.L., Cortes A., Bethune J., Han B., Park Y.R., Raychaudhuri S., Pouget J.G., Hübenthal M., et al. (2016). Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. *Nat Genet.* 48, 510. [10.1038/ng.3528](https://doi.org/10.1038/ng.3528)