

A disease-associated gene desert orchestrates macrophage inflammatory responses via ETS2

Authors:

Stankey CT^{1,2*}, Bourges C^{1*}, Turner-Stokes T^{1,2}, Piedade AP¹, Palmer-Jones C³, Papa I¹, Dos Silva dos Santos M⁴, Randzavola LO², Speidel L^{5,6}, Parkes EC¹, Edwards W⁷, Rochford AP³, Murray CD³, MacRae JI⁴, Skoglund P⁵, Wallace C^{7,9}, Cader MZ^{7,8}, Thomas DC², Lee JC^{1,3}.

Affiliations:

¹ Genetic Mechanisms of Disease Laboratory, The Francis Crick Institute, London, UK

² Department of Immunology and Inflammation, Imperial College London, UK

³ Department of Gastroenterology, UCL Institute of Liver and Digestive Diseases, Royal Free Hospital, Pond St, London, UK

⁴ Metabolomics STP, The Francis Crick Institute, London, UK

⁵ Ancient Genomics Laboratory, The Francis Crick Institute, London, UK

⁶ Genetics Institute, University College London, UK

⁷ Cambridge Institute of Therapeutic Immunology and Infectious Disease, Jeffrey Cheah Biomedical Centre, University of Cambridge, Cambridge, UK

⁸ Department of Medicine, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK

⁹ MRC Biostatistics Unit, Cambridge Institute of Public Health, Cambridge, UK

* These authors contributed equally to this work

Correspondence: james.lee@crick.ac.uk

Abstract

Increasing global rates of autoimmune and inflammatory disease present a burgeoning threat to human health¹. This is compounded by the limited efficacy of available treatments¹ and high failure rates during drug development² – underscoring an urgent need to better understand disease mechanisms. Here we show how genetics could address this challenge. By investigating an intergenic haplotype on chr21q22, independently linked to inflammatory bowel disease (IBD), ankylosing spondylitis, primary sclerosing cholangitis and Takayasu’s arteritis³⁻⁶, we discover that the causal gene, *ETS2*, is a master regulator of inflammatory responses in human macrophages and delineate how the risk haplotype increases *ETS2* expression. Genes regulated by *ETS2* were prominently expressed in affected tissues from chr21q22-associated diseases and more enriched for IBD GWAS hits than almost all previously described pathways. Overexpressing *ETS2* in resting macrophages produced an activated effector state that phenocopied intestinal macrophages from IBD⁷, with upregulation of multiple drug targets including TNF α and IL-23. Using a database of cellular signatures⁸, we identify drugs that could modulate this pathway and validate the potent anti-inflammatory activity of one class of small molecules *in vitro* and *ex vivo*. Together, this highlights the potential for common genetic associations to improve both the understanding and treatment of human disease.

Currently, nearly 5% of the world's population are affected by at least one autoimmune or inflammatory disease. These heterogeneous conditions, which range from Crohn's disease and ulcerative colitis (collectively IBD) to psoriasis and rheumatoid arthritis, share a common need for better treatments, but only ~10% of drugs entering clinical development ever become approved therapies². This high failure rate is principally due to a lack of efficacy⁹ or – put another way – because the pathways being targeted are less important than they were assumed to be. Genetics provides a unique opportunity to elucidate disease mechanisms, with hundreds of regions of the human genome now directly linked to the pathogenesis of one or more autoimmune or inflammatory disease¹⁰. Indeed, drugs that target pathways implicated by genetics have a substantially higher chance of becoming approved therapies¹¹.

To fully realise the potential of genetics, however, knowledge of where risk variants lie must first be translated into an understanding of how they contribute to disease¹⁰. This is a formidable challenge since most disease-associated genetic variants do not lie in coding DNA, where effects on protein sequence/structure can be easily predicted, but in the enigmatic non-coding genome where the same DNA sequence can have different biological functions depending on the cell-type and/or external stimuli¹⁰. Most risk variants are thought to affect gene regulation¹², but the need to identify the causal gene(s) – which may lie up to one million bases away – and the causal cell-type(s), which may only express the causal gene under specific conditions, have hindered attempts to discover disease mechanisms. For example, although genome-wide association studies (GWAS) have identified over 240 IBD risk loci³, fewer than 10 have been mechanistically resolved and, to date, none have led to new therapies.

Molecular mechanisms at the pleiotropic chr21q22 locus

Several genetic variants predispose to more than one disease – highlighting both their biological importance and an opportunity to discover shared disease mechanisms. One

notable example is an intergenic region on chr21q22, where the major (risk) allele haplotype has been independently associated with five different inflammatory diseases³⁻⁶. Although the associated locus does not contain any genes, there are several nearby candidates including *PSMG1*, *BRWD1* and *ETS2* (**Fig.1a**), all of which have been proposed to be potentially causal in previous studies^{3-6,13}. The underlying biological mechanisms, however, remain unknown. We hypothesised that this intergenic locus must contain a distal enhancer and – since the associated diseases are all immune-mediated despite affecting different organs – searched for evidence of enhancer activity in disease-relevant immune cell-types. Using H3K27ac ChIP-seq data, which marks active enhancers/promoters, we found that this locus contains a monocyte/macrophage-specific enhancer (**Fig.1a**). Monocytes and monocyte-derived macrophages play a central role in the pathogenesis of many autoimmune and inflammatory diseases, producing cytokines that are often targeted by the most effective therapies¹⁴.

To identify the gene regulated by this enhancer, we first considered publicly-available data from human monocytes, including promoter-capture Hi-C¹⁵ and eQTL datasets¹⁶. We found that the disease-associated locus physically interacts with the promoter of *ETS2*, the most distant of the candidate genes (located 290-kb away) and that the risk haplotype correlates with higher *ETS2* expression, especially after monocyte activation (**Extended Data Fig.1**). Interestingly, however, the disease association was predicted to share a causal variant not with the strong eQTL in activated monocytes, but rather with a weaker eQTL in resting monocytes (**Extended Data Fig.1**). To directly confirm the identity of the target gene, we deleted the 1.85-kb enhancer region in primary human monocytes using CRISPR-Cas9 ribonucleoprotein (RNP) complexes containing gRNAs that flank the enhancer (**Fig.1b**, **Extended Data Fig.2**). These cells were then cultured with combination of inflammatory mediators, including TNF α (a pro-inflammatory cytokine), prostaglandin E2 (a pro-inflammatory lipid) and Pam3CSK4 (a TLR1/2 agonist). This model, termed “TPP”, was designed to mimic a chronic inflammatory environment¹⁷, and better recapitulates the state

of patient-derived monocytes/macrophages than classical M1 or M2 models (ref.18 and **Extended Data Fig.2**). Because flow cytometry antibodies were not available for any of the candidate genes, we used PrimeFlow to measure the dynamics of RNA transcription and found that expression of all 3 genes (*ETS2*, *BRWD1*, *PSMG1*) increased in unedited cells upon exposure to inflammatory stimuli (**Fig.1c**). Deletion of the chr21q22 enhancer did not affect the increase in *BRWD1* and *PSMG1* expression, but the upregulation of *ETS2* was significantly reduced (**Fig.1d**) – confirming that this pleiotropic locus functions as a distal *ETS2* enhancer in monocytes and monocyte-derived macrophages.

We next sought to discover the variant responsible for disease risk at the chr21q22 locus. Unfortunately, statistical fine-mapping – using the largest IBD GWAS to date³ – could not identify the causal variant due to high linkage disequilibrium between the candidate single nucleotide polymorphisms (SNPs) (**Methods, Fig.1e**). We therefore used a high-throughput functional approach to first delineate the active enhancers at the locus, and then determine if any candidate SNPs within these regulatory regions might alter enhancer activity. This method – massively-parallel reporter assay (MPRA) – can simultaneously characterise enhancer activity in thousands of short DNA sequences by coupling each to a uniquely-barcoded reporter gene within an expression vector¹⁹. Genetic sequences that modulate gene expression can be identified by normalising the barcode counts in mRNA extracted from transfected cells to their equivalent counts in the input DNA library. After adapting the MPRA vector for use in primary macrophages (**Methods, Extended Data Fig.3**), we synthesised a pool of overlapping oligonucleotides (oligos) to tile the 2-kb region encompassing all candidate SNPs at chr21q22, and included additional oligos containing either risk or non-risk alleles for every variant. The resulting vector library was transfected into inflammatory macrophages from multiple donors – thus ensuring that a physiological repertoire of transcription factors would be present to interact with the chr21q22 genomic sequences. Using a sliding window analysis to map active enhancers across the tiling sequences, we identified a single 442-bp region of enhancer activity (chr21:40466236-

40466677, hg19; **Fig.1f**) that harboured three (of seven) candidate SNPs. Two of these polymorphisms were transcriptionally inert, but the third (rs2836882) had the strongest expression-modulating effect of any candidate variant, with the risk allele (G) significantly increasing transcription – consistent with the known eQTL (**Fig.1f**). Examining rs2836882 further, we noticed that this SNP lay within an experimentally-confirmed PU.1 ChIP-seq peak in human macrophages (**Fig.1g**). PU.1 is an important myeloid pioneer factor²⁰ that can bind to heterochromatin, initiate nucleosome remodelling – thus enabling other transcription factors to bind – and activate transcription²¹. To determine whether rs2836882 might affect PU.1 binding, we identified two publicly-available macrophage PU.1 ChIP-seq datasets from rs2836882 heterozygotes and used BaalChIP²² to assess for allelic imbalances in PU.1 binding. Despite not lying within a canonical PU.1 binding motif, significant allele-specific PU.1 binding was detected at rs2836882, with over 4-fold greater binding to the risk allele in both datasets (**Fig.1h**). This result was replicated in TPP macrophages from five heterozygous donors by immunoprecipitating PU.1 and genotyping the bound DNA (**Extended Data Fig.4**). Together, this suggests that the rs2836882 risk allele should increase enhancer activity, consistent with the MPRA and eQTL results.

To test whether allele-specific enhancer activity was evident at the endogenous locus, we performed H3K27ac ChIP-seq in inflammatory macrophages from two rs2836882 major allele homozygotes and two minor allele homozygotes. While several nearby enhancer peaks were similar between these donors, the enhancer activity overlying rs2836882 was considerably stronger in major (risk) allele homozygotes (**Fig.1i**), contributing to a ~2.5-fold increase in enhancer activity across the extended chr21q22 locus (**Extended Data Fig.4**). Collectively, these data reveal a genetic mechanism whereby the putative causal variant at chr21q22 – identified via its functional consequences in primary macrophages – promotes binding of a pioneer transcription factor and increases the activity of a long-range *ETS2* enhancer.

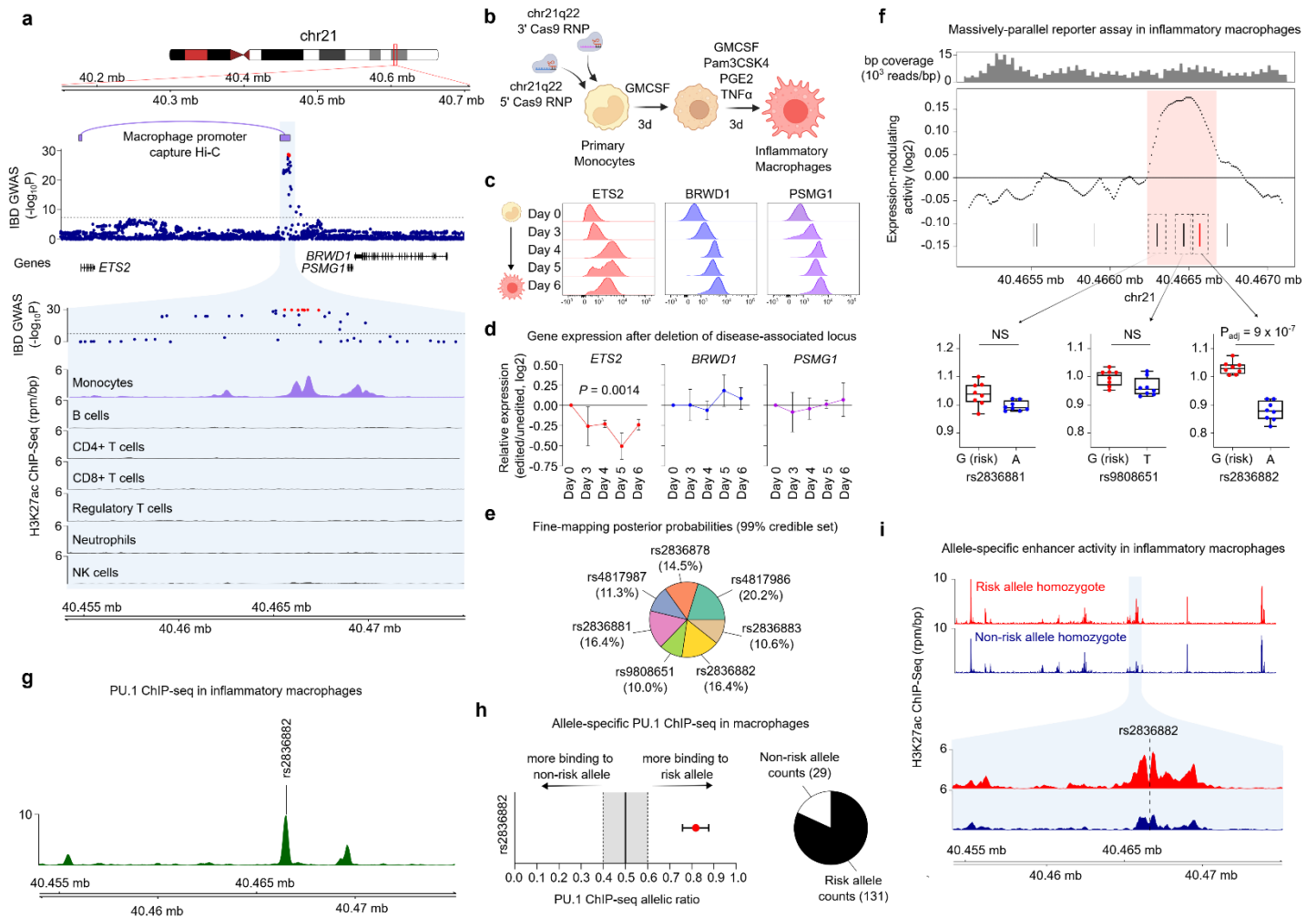


Figure 1. Resolving molecular mechanisms at chr21q22.

a. Annotation of the disease-associated chr21q22 locus depicting the IBD genetic association, physical interactions of the disease-associated haplotype in macrophages (promoter-capture Hi-C), and H3K27ac ChIP-seq data from various immune cell-types. **b.** Schematic of experiment to determine the function of the chr21q22 locus in monocyte-derived macrophages polarised under chronic inflammatory (“TPP”) conditions. **c.** Histograms depicting the expression of *ETS2*, *BRWD1*, and *PSMG1* during inflammatory macrophage polarisation, measured using PrimeFlow RNA assays that quantify RNA by flow cytometry. Data are representative of one of four donors. **d.** Relative *ETS2*, *BRWD1*, and *PSMG1* expression in chr21q22-edited inflammatory macrophages (relative to non-targeting control cells; NTC). Plot shows log₂ fold-change in mean fluorescence intensity (n=4, data represent mean±SEM, two-way ANOVA). **e.** SuSiE fine-mapping posterior probabilities for IBD-associated SNPs at the 21q22 locus (99% credible set). **f.** MPRA at the chr21q22 locus depicting oligonucleotide coverage (top), enhancer activity in inflammatory macrophages (analysed using a sliding window analysis of tiling oligos; middle), and expression-modulating effects of candidate SNPs within the identified enhancer (bottom) (n=8). Shaded region in enhancer activity plot indicates region of significant enhancer activity. **g.** PU.1 ChIP-seq peaks at the chr21q22 locus in macrophages. **h.** BaalChIP analysis of allele-specific PU.1 binding at rs2836882 in two heterozygous macrophage datasets (data represent 95% posterior distribution of allelic binding ratio) **i.** H3K27ac ChIP-seq data from major (top) or minor (bottom) allele homozygotes at the chr21q22 locus. Data are representative of two of four donors.

1 **ETS2 is essential for macrophage inflammatory responses**

2 Having identified a plausible mechanism by which the chr21q22 risk haplotype increases
3 *ETS2* expression, we sought to better understand the role of *ETS2* in
4 monocytes/macrophages. *ETS2* is a member of the ETS family of transcription factors, which
5 has been mainly studied as a proto-oncogene in cancer²³. In contrast, the role of *ETS2* in
6 primary human macrophages has been less clearly defined, with previous studies using either
7 cell-lines or complex mouse models and largely focusing on a limited number of downstream
8 molecules²⁴⁻²⁸. This has led to contradictory reports, with *ETS2* being described as both
9 necessary and redundant for macrophage development^{29,30}, and both pro- and anti-
10 inflammatory²⁴⁻²⁸. To elucidate the specific role of *ETS2* in inflammatory human macrophages
11 – and determine how dysregulated *ETS2* expression might contribute to disease – we first
12 used a CRISPR-Cas9-based loss-of-function approach (**Fig.2a**). Two gRNAs targeting
13 different *ETS2* exons were designed, validated and individually incorporated into Cas9 RNPs
14 for transfection into primary monocytes – thereby minimising the chance that any effect was
15 due to off-target editing. These gRNAs resulted in on-target editing in ~89% and ~79% of total
16 cells respectively (**Extended Data Fig.2**). No differences in cell viability or expression of
17 macrophage markers were observed with either gRNA, suggesting that *ETS2* was not
18 required for inflammatory macrophage survival or differentiation (**Extended Data Fig.2**). In
19 contrast, production of pro-inflammatory cytokines, including IL-6, IL-8 and IL-1 β , was
20 significantly reduced following *ETS2* disruption (**Fig.2b**), whereas IL-10 – an anti-inflammatory
21 cytokine – was less affected. TNF α secretion could not be assessed as it was included in the
22 differentiation culture. We therefore investigated whether *ETS2* was also required for other
23 macrophage effector functions. First, we examined phagocytosis using fluorescently-labelled
24 particles that can be detected by flow cytometry. Similar to pro-inflammatory cytokine
25 production, phagocytosis was significantly impaired following *ETS2* disruption (**Fig.2c**). We
26 next measured extracellular reactive oxygen species (ROS) production – a key effector
27 response that directly contributes to tissue damage in inflammatory disease³¹. Disrupting
28 *ETS2* profoundly reduced the oxidative burst following macrophage activation – an effect that

29 appeared to be due to reduced expression of key NADPH oxidase components (**Fig.2d**,
30 **Extended Data Fig.5**). Together, this suggests that *ETS2* is required for multiple effector
31 functions in inflammatory macrophages.

32 To better understand the molecular basis for these distinct functional effects, we performed
33 RNA-sequencing (RNA-seq) in *ETS2*-edited and unedited inflammatory macrophages from
34 multiple donors. Disruption of *ETS2* led to widespread transcriptional changes, with reduced
35 expression of many inflammatory genes, including several well-known initiators and amplifiers
36 of inflammation (**Fig.2e**). Affected gene classes included cytokines (e.g. *TNFSF10/TRAIL*,
37 *TNFSF13*, *IL1A*, *IL1B*), chemokines (e.g. *CXCL1*, *CXCL3*, *CXCL5*, *CCL2*, *CCL5*), secreted
38 effector molecules (e.g. *S100A8*, *S100A9*, *MMP14*, *MMP9*), cell surface receptors (e.g.
39 *FCGR2A*, *FCGR2C*, *TREM1*), pattern recognition receptors (e.g. *TLR2*, *TLR6*, *NOD2*), and
40 signalling molecules (e.g. *MAP2K*, *GPR84*, *NLRP3*). To better characterise the pathways
41 affected by *ETS2* deletion, we performed gene-set enrichment analysis (GSEA) using the
42 Gene Ontology Biological Pathways dataset. This corroborated the observed functional effects
43 (**Fig.2f**), with the most negatively-enriched pathways (downregulated following *ETS2*
44 disruption) relating to macrophage activation, pro-inflammatory cytokine production,
45 phagocytosis and ROS production. Genes involved in macrophage migration were also
46 significantly downregulated, but gene sets relating to monocyte-to-macrophage differentiation
47 were not significantly affected – consistent with *ETS2* being required for inflammatory effector
48 functions, but not influencing monocyte-to-macrophage development. Although fewer genes
49 increased in expression following *ETS2* deletion (**Fig.2e**), positive enrichment was noted for
50 genes involved in aerobic respiration and oxidative phosphorylation (OXPHOS; **Fig.2f**) –
51 metabolic processes linked to anti-inflammatory macrophage behaviour³². Collectively, these
52 data identify an indispensable role for *ETS2* in a range of macrophage effector functions,
53 which could explain why dysregulated *ETS2* expression is associated with multiple
54 inflammatory diseases. Indeed, deletion of the disease-associated chr21q22 enhancer
55 phenocopied both the transcriptional and functional consequences of disrupting *ETS2* (**Fig.2g**,
56 **Extended Data Fig.5**).

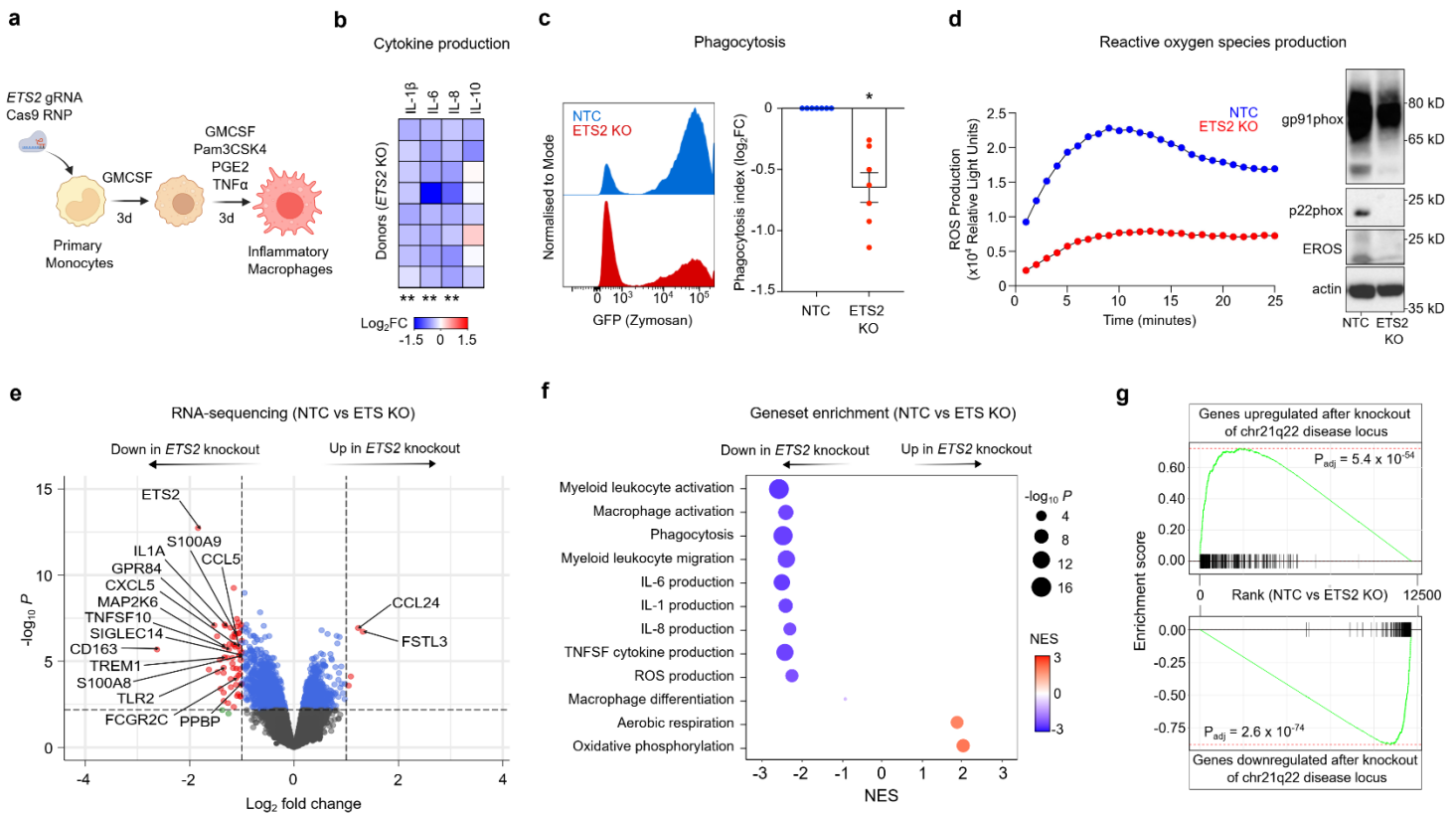


Figure 2. *ETS2* is essential for macrophage inflammatory responses.

a. Schematic of experiment for disrupting *ETS2* in primary monocytes and differentiating monocyte-derived macrophages under chronic inflammatory conditions. **b.** Macrophage cytokine secretion following *ETS2* disruption. Heatmap shows log₂ fold-change of cytokine concentrations in the supernatants of *ETS2*-edited macrophages relative to unedited macrophages transfected with a non-targeting control gRNA-containing RNP (NTC). n=9, Wilcoxon matched-pairs test, two-tailed. **c.** Histogram depicting phagocytosis of fluorescently-labelled zymosan particles by *ETS2*-edited and unedited macrophages (left). Data representative of one of seven donors. Phagocytosis index in *ETS2*-edited and unedited macrophages (calculated as product of proportion and mean fluorescence intensity of phagocytosing cells; right). Plot depicts log₂ fold-change in phagocytosis index for *ETS2*-edited macrophages relative to unedited cells (Wilcoxon signed-rank test, two-tailed; data represent mean±SEM). **d.** Production of ROS by *ETS2*-edited and unedited inflammatory macrophages (measured in relative light units; left). Data representative of one of six donors. Western blot for gp91phox, p22phox, and EROS expression in *ETS2*-edited and unedited macrophages (right). Data representative of one of three donors. **e.** Differentially-expressed genes in *ETS2*-edited versus unedited inflammatory macrophages (limma with voom transformation, n=8). **f.** Gene set enrichment analysis (fGSEA) of differentially-expressed genes between *ETS2*-edited and unedited inflammatory macrophages. Results of selected Gene Ontology Biological Pathways shown. Dot size represents *P*-value and colour denotes normalised enrichment score (NES). **g.** Enrichment of differentially-expressed genes following deletion of the disease-associated chr21q22 locus (upregulated genes, top; downregulated genes, bottom) in *ETS2*-edited versus unedited macrophages. * *P* < 0.05, ** *P* < 0.01.

57 **ETS2 orchestrates macrophage inflammatory responses**

58 Having found that *ETS2* was essential for monocyte-derived macrophage effector functions,
59 we next investigated whether it might also be sufficient to drive them – as would be expected
60 of a master regulator of inflammatory responses. This was particularly important because
61 although loss-of-function approaches can identify a gene’s biological role(s), the chr21q22 risk
62 haplotype increases *ETS2* expression. To do this, we first optimised a method for controlled
63 overexpression of target genes in primary macrophages by transfecting defined amounts of *in*
64 *vitro* transcribed mRNA that was modified to minimise immunogenicity (**Fig.3a, Extended**
65 **Data Fig.3, Methods**). Resting, non-activated (M0) macrophages were then transfected with
66 mRNA encoding *ETS2* or its reverse complement – thereby controlling for quantity, length and
67 purine/pyrimidine composition of the transfected mRNA but with a transcript that would not be
68 translated (**Fig.3b**). After transfection, cells were exposed to low-dose lipopolysaccharide for 6
69 hours to initiate a low-grade inflammatory response that could be amplified if *ETS2* was
70 sufficient to drive inflammatory responses (**Fig.3a**). We first quantified secreted cytokines and
71 found that *ETS2* overexpression increased production of several pro-inflammatory cytokines,
72 although IL-10 was again less affected (**Fig.3c**). To better characterise the consequences of
73 *ETS2* overexpression, we performed RNA-seq and examined the macrophage activation
74 pathways that had required *ETS2*. Strikingly, all of these inflammatory pathways – including
75 macrophage activation, pro-inflammatory cytokine production, ROS production, phagocytosis
76 and migration – were induced in a dose-dependent manner following *ETS2* overexpression,
77 with greater enrichment of every pathway when more *ETS2* mRNA was transfected (**Fig.3d**).
78 This shows that *ETS2* is both necessary and sufficient for inflammatory responses in human
79 macrophages, consistent with being a master regulator of effector functions during chronic
80 inflammation, whose dysregulation is directly linked to human disease.

81

82 **ETS2-regulated genes play a central role in IBD**

83 To understand whether *ETS2* might contribute to the macrophage phenotype observed in
84 inflammatory disease, we compared the transcriptional consequences of overexpressing
85 *ETS2* with a gene signature of intestinal macrophages from Crohn's disease – one of the
86 conditions associated with the chr21q22 locus. Single cell RNA-seq analysis has previously
87 shown that active Crohn's disease is characterised by an expanded population of
88 inflammatory monocyte-derived macrophages that contributes to anti-TNF α resistance⁷. Using
89 the Crohn's disease macrophage signature as a gene set, we found that overexpressing
90 *ETS2* in resting macrophages induced a transcriptional state that closely resembled disease
91 macrophages, with core ("leading edge") enrichment of the majority of genes in the signature,
92 including many that encode therapeutic targets (**Fig.3e**).

93 Given the importance of *ETS2* in macrophage responses, and the fact that *ETS2*
94 overexpression phenocopied a disease-associated inflammatory state, we hypothesised that
95 other genetic associations might also implicate this previously uncharacterised pathway. An
96 important goal of GWAS was to identify key disease pathways¹⁰, but this has proven
97 challenging due to a paucity of confidently-identified causal genes and a limited understanding
98 of how these are affected by genetic variation¹⁰. To better characterise the genetic risk
99 attributable to the macrophage *ETS2* pathway, we focused on IBD since this has far more
100 genetic associations than any other chr21q22-associated disease. Examining the list of
101 commonly downregulated genes following *ETS2* editing ($P_{adj} < 0.05$ for both gRNAs), we
102 identified over 20 IBD risk genes – including many that have been proposed to be causal at
103 their respective loci^{3,33} (**Extended Data Table 1**). These included genes that are thought to
104 affect macrophage biology (e.g. *SP140*, *LACC1/FAMIN*, *CCL2*, *CARD9*, *CXCL5*, *TLR4*,
105 *SLAMF8*, *FCGR2A*) as well as some that are highly expressed in macrophages but not
106 previously linked to specific pathways (e.g. *ADCY7*, *PTPRC*, *TAGAP*, *PTAFR*, *PDLIM5*). To
107 more formally assess the enrichment of an *ETS2*-regulated inflammatory pathway in IBD
108 genetics – and compare this to known disease pathways – we used SNPsea³⁴, an algorithm
109 designed to identify pathways affected by disease loci. 241 IBD-associated genetic loci were
110 tested for enrichment in 7,660 pathways, comprising 7,658 Gene Ontology Biological

111 Pathways and 2 overlapping lists of ETS2-regulated genes (either those downregulated
112 following *ETS2* editing or upregulated following *ETS2* overexpression). Significance of
113 enrichment was empirically computed using 5 million matched null SNP sets, and disease
114 pathways previously implicated by genetics were extracted for comparison. Strikingly, ETS2
115 target genes – however they were defined – were more strongly enriched for IBD-associated
116 loci than almost all previously implicated pathways, with not a single null SNP set showing
117 greater enrichment in either of the ETS2-regulated gene lists. After applying a stringent
118 Bonferroni multiple-testing correction, only ETS2-regulated genes and IBD pathways relating
119 to T cell activation, T-helper 17 cells, autophagy and IL-10 signalling showed significant
120 enrichment (**Fig.3f**). This suggests that ETS2 signalling in macrophages plays a fundamental
121 role in IBD pathogenesis, with stronger genetic enrichment than most previously implicated
122 pathways.

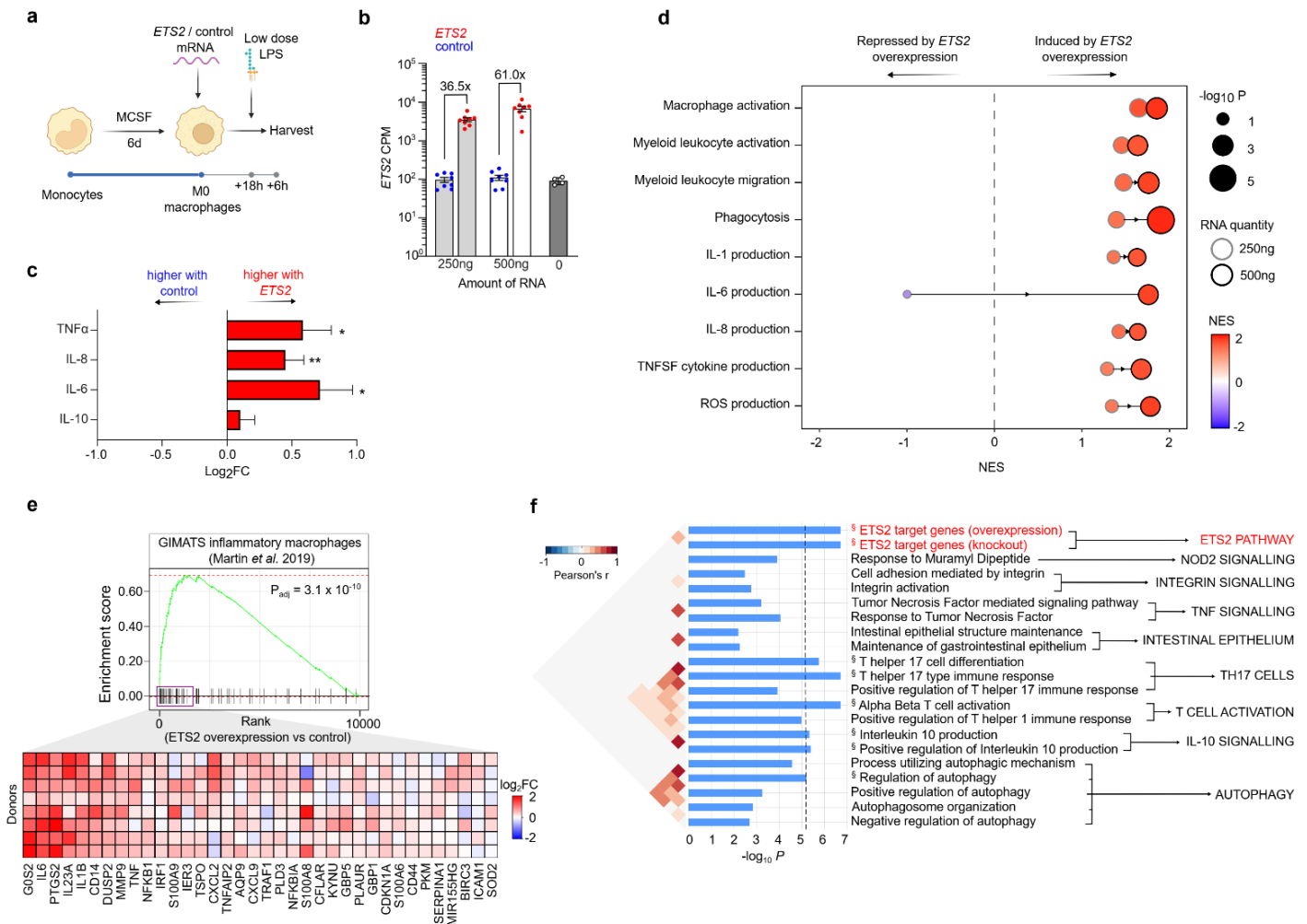


Figure 3. *ETS2* orchestrates macrophage inflammatory responses.

a. Schematic of *ETS2* overexpression experiment. Resting (M0) human macrophages were transfected with pre-defined amounts of *in vitro* transcribed *ETS2* mRNA or control mRNA (*ETS2* reverse complement), activated with low dose LPS (0.5ng/ml), and harvested. n=8. **b.** *ETS2* mRNA levels in macrophages transfected with *ETS2* or control mRNA or untransfected (from separate experiment). **c.** Cytokine secretion following *ETS2* overexpression. Plot shows log₂ fold-change of cytokine concentrations in macrophage supernatants (*ETS2* relative to control) following transfection with 500ng mRNA. **d.** Gene set enrichment analysis (fGSEA) of differentially-expressed genes between *ETS2*-overexpressing and control macrophages. Results shown for the same Gene Ontology Biological Pathways that were negatively enriched following *ETS2* editing. Dot size represents *P*-value, colour denotes normalised enrichment score (NES), and border colour denotes amount of transfected mRNA. **e.** Enrichment of a disease-associated inflammatory macrophage gene signature, derived from single cell RNA-seq of Crohn's disease intestinal biopsies, in *ETS2*-overexpressing macrophages (relative to control; top). Heatmap of leading-edge genes showing log₂ fold-change of gene expression in *ETS2*-overexpressing macrophages relative to control (500ng mRNA; bottom). **f.** SNPsea analysis of enrichment of 241 IBD-associated loci within *ETS2*-regulated genes (red) and pathways previously linked to IBD pathogenesis (black). Significantly enriched pathways (Bonferroni-corrected permutation $P < 0.05$) indicated by §. * $P < 0.05$, ** $P < 0.01$.

123 **ETS2 controls inflammatory responses via transcriptional and metabolic effects**

124 We next sought to understand how *ETS2* controlled such diverse macrophage effector
125 functions. Studying *ETS2* biology is challenging because no ChIP-seq-grade antibodies exist,
126 precluding direct identification of its transcriptional targets. Even the ENCODE project, which
127 performed ChIP-seq for 181 transcription factors, was unable to directly immunoprecipitate
128 *ETS2*³⁵. We therefore first used a “guilt-by-association” approach to identify genes that were
129 co-expressed with *ETS2* across 64 different human macrophage polarisation conditions¹⁷.
130 This identified *PFKFB3* – encoding the rate-limiting enzyme of glycolysis – as the most highly
131 co-expressed gene, with *HIF1A* also highly co-expressed (**Fig.4a**). Together, these genes are
132 known to facilitate a “glycolytic switch” that is required for myeloid inflammatory responses³⁶.
133 We therefore hypothesised that *ETS2* might control inflammatory responses via metabolic
134 reprogramming – a possibility supported by OXPHOS genes being negatively correlated with
135 *ETS2* expression (**Fig.4a**) and upregulated following *ETS2* disruption (**Fig.2f**). To assess the
136 metabolic consequences of disrupting *ETS2*, we quantified label incorporation from ¹³C-
137 glucose in edited and unedited inflammatory macrophages using gas chromatography–mass
138 spectrometry (GC-MS). Widespread modest reductions in labelled and total glucose
139 metabolites were detected following *ETS2* disruption (**Fig.4b, Extended Data Fig.6**). This
140 affected both glycolytic and TCA cycle metabolites, with significant reductions in intracellular
141 and secreted lactate, a hallmark of anaerobic glycolysis, and succinate, an important
142 inflammatory signalling metabolite³⁷. These results would be consistent with glycolytic
143 suppression, with reductions in TCA metabolites being due to reduced flux into TCA and
144 increased consumption by mitochondrial OXPHOS³⁸. To determine whether metabolic
145 changes were responsible for *ETS2*-mediated inflammatory effects, we treated *ETS2*-edited
146 macrophages with roxadustat, a HIF1 α stabiliser that promotes glycolysis via HIF1 α -mediated
147 metabolic reprogramming. This had the predicted effect on genes involved in glycolysis and
148 OXPHOS, but did not rescue the effects of *ETS2* disruption, either transcriptionally or
149 functionally (**Fig.4c, Extended Data Fig.6**). Thus, while disrupting *ETS2* does alter
150 glycometabolism, this does not fully explain the observed differences in inflammation.

151 To try to elucidate the mechanism by which ETS2 controlled such diverse inflammatory
152 effects, we revisited whether we could directly identify ETS2 target genes. Using a range of
153 anti-ETS2 antibodies, we confirmed that none worked for ChIP-seq (data not shown) so
154 investigated whether any might work for Cleavage-Under-Targets-and-Release-Using-
155 Nuclease (CUT&RUN), which does not require formaldehyde fixation. One antibody identified
156 multiple, significantly-enriched genomic regions (peaks) of which 6,560 were reproducibly
157 detected across two biological replicates (Irreproducible Discovery Rate cut-off 0.01) with
158 acceptable quality metrics³⁹ (**Fig.4d**). These peaks were mostly located in active regulatory
159 regions (90% in promoters or active enhancers, **Fig.4e**) and were highly enriched for a
160 canonical ETS2 binding motif (4.02-fold enrichment over global controls, **Fig.4f**) – consistent
161 with being sites of ETS2 binding. After combining the biological replicates to improve peak
162 detection, we identified ETS2 binding at the promoters of many inflammatory genes, including
163 several that are essential for distinct macrophage functions, such as *NCF4* (encoding a key
164 NADPH oxidase component), *NLRP3* (encoding a key inflammasome component), and *TLR4*
165 (encoding a key pattern recognition receptor) (**Fig.4g**). Overall, 48.3% of genes dysregulated
166 following *ETS2* disruption, and 50.3% of genes dysregulated following *ETS2* overexpression,
167 contained an ETS2 binding peak within their core promoter or putative cis-regulatory elements
168 (**Fig.4h**) – consistent with ETS2 directly regulating a range of macrophage inflammatory
169 responses. Notably, these gene targets included *HIF1A*, *PFKFB3*, and other glycolytic genes
170 (e.g. *GPI*, *HK2*, and *HK3*), suggesting that the observed metabolic changes might be directly
171 induced by ETS2, rather than being solely attributable to differences in inflammation.
172 Intriguingly, we also detected ETS2 binding at its own enhancer at chr21q22 (**Fig.4i**). This is
173 consistent with reports that PU.1 and ETS2 can interact synergistically⁴⁰, and would implicate
174 a feed-forward mechanism at the disease-associated locus, where increased *ETS2*
175 expression reinforces *ETS2* enhancer activity. Together, these data implicate ETS2 as a
176 master regulator of monocyte/macrophage responses during chronic inflammation, capable of
177 directing a multifaceted effector programme, and creating a metabolic environment that is
178 permissive for inflammation.

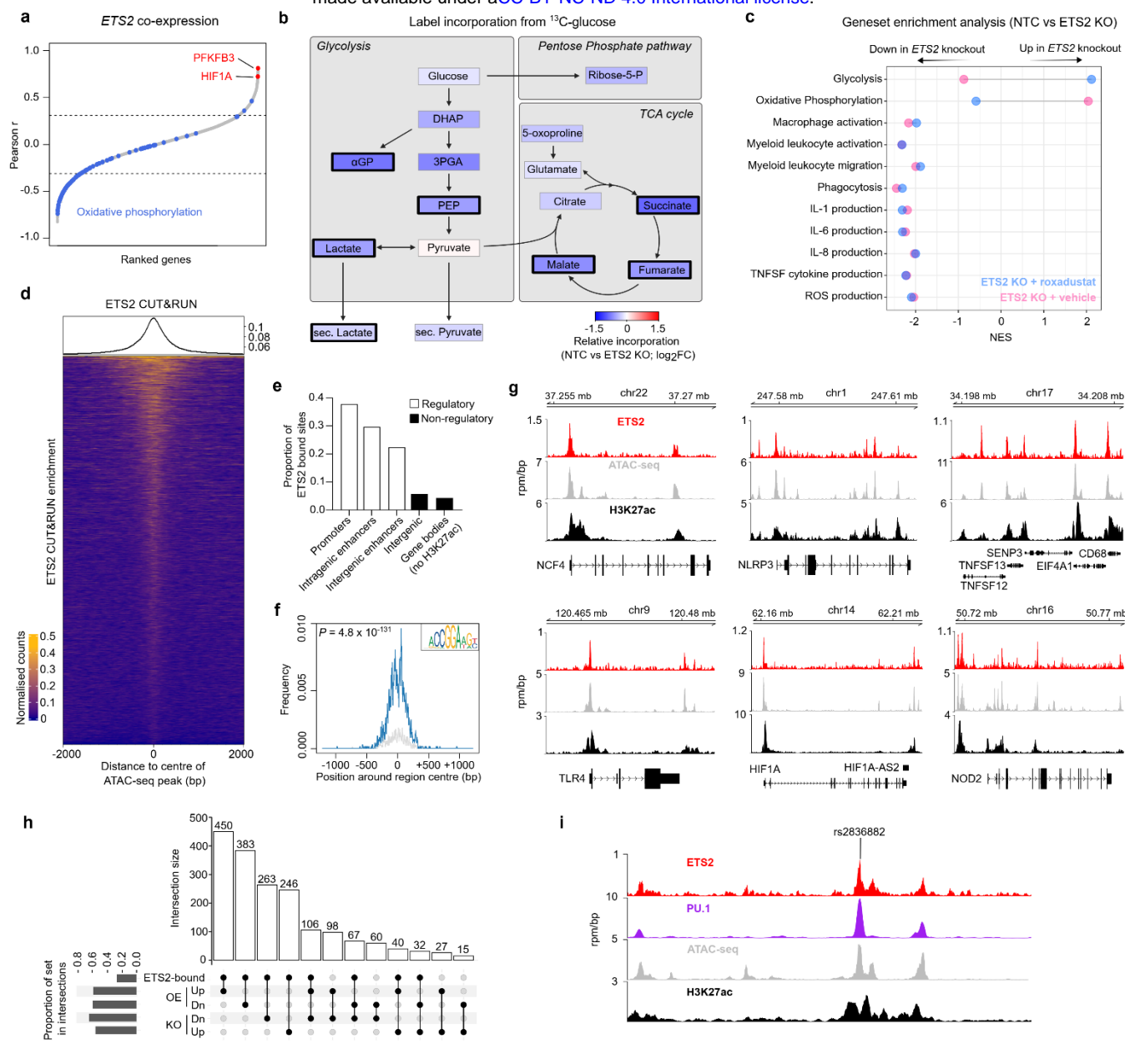


Figure 4. *ETS2* directs macrophage responses via transcriptional and metabolic effects.

a. Genes co-expressed with *ETS2* in 64 monocyte-derived macrophage datasets. Dotted lines equivalent to FDR $P < 0.05$. **b.** Effect of *ETS2* disruption on glucose metabolism. Colour denotes median log₂ fold-change in label incorporation from ^{13}C -glucose in *ETS2*-edited cells relative to unedited cells. Bold black border denotes $P < 0.05$ (Wilcoxon matched-pairs, two-tailed, $n=6$). **c.** Gene set enrichment analysis (fGSEA) of differentially-expressed genes between *ETS2*-edited and unedited TPP macrophages, treated with either roxadustat or vehicle. Results for selected Gene Ontology Biological Pathways shown. **d.** Enrichment heatmap of *ETS2* CUT&RUN peaks (IDR cut-off 0.01, $n=2$) in accessible chromatin from TPP macrophages (4-kb regions centred on ATAC-seq peaks). **e.** Features of *ETS2* binding sites (based on gene coordinates and H3K27ac ChIP-seq in TPP macrophages). **f.** Enrichment of an *ETS2* binding motif in *ETS2* CUT&RUN peaks (hypergeometric P-value). **g.** *ETS2* binding, chromatin accessibility (ATAC-seq), and enhancer activity (H3K27ac) at selected loci. **h.** UpSet plot of intersections between *ETS2* gene lists, including genes with *ETS2* peaks in their core promoters or cis-regulatory elements and significantly up- (Up) or down-regulated (Dn) genes following *ETS2* editing (KO) or overexpression (OE). Vertical bars denote shared genes between lists, indicated by connected dots in lower panel. Horizontal bars denote proportion of gene list within intersections. **i.** *ETS2* binding, chromatin accessibility (ATAC-seq), and enhancer activity (H3K27ac) at the disease-associated chr21q22 locus.

179 **ETS2-driven inflammation is evident in diseased tissue and can be targeted**
180 **pharmacologically.**

181 The strong enrichment of IBD GWAS hits within ETS2-regulated genes led us to hypothesise
182 that the transcriptional footprint of this pathway might be generally detectable in the affected
183 tissues of chr21q22-associated diseases – a possibility that would have important therapeutic
184 implications. Using publicly-available gene expression data from diseases linked to chr21q22
185 – intestinal macrophages from IBD, synovium from ankylosing spondylitis (AS), and liver from
186 PSC – we confirmed that diseased tissue was significantly enriched for ETS2-regulated genes
187 (**Fig.5a, Extended Data Fig.7**). We therefore investigated whether this pathway could be
188 pharmacologically targeted. Specific ETS2 inhibitors do not exist and structural analyses
189 indicate that there is no allosteric inhibitory mechanism that could be easily targeted⁴¹. We
190 therefore used the NIH LINCS database to identify drugs that might modulate ETS2 activity⁸.
191 This repository contains over 30,000 differentially-expressed genelists from cell-lines exposed
192 to over 6,000 small molecules. Using fGSEA, 906 drug signatures were found to mimic the
193 transcriptional effect of disrupting *ETS2* in inflammatory macrophages ($P_{adj} < 0.05$), including
194 several approved treatments for IBD and AS (e.g. JAK inhibitors). Of these candidate
195 therapies, the most common class were MEK inhibitors (**Fig.5b**), which are already licensed
196 for non-inflammatory human diseases (e.g. neurofibromatosis). This result was not due to a
197 single compound, but rather a class effect with multiple MEK1/2 inhibitors downregulating
198 ETS2 target genes (**Fig.5c**). This made biological sense, since MEK1 and MEK2 – together
199 with several other targets identified – are known upstream regulators of ETS-family
200 transcription factors (**Fig.5d**). Indeed, some of these drug classes, including MEK1/2 and
201 HSP90 inhibitors, have been reported to be beneficial in animal colitis models, although this is
202 often a poor indicator of clinical efficacy – with several approved IBD treatments being
203 ineffective in mice and many drugs that improve mouse models being ineffective in human
204 IBD⁴². To determine whether MEK inhibition would abrogate ETS2-driven inflammatory
205 responses in primary human macrophages, we differentiated monocytes under chronic
206 inflammatory conditions and treated them with a selective, non-ATP competitive MEK inhibitor

207 (PD-0325901; **Fig.5e**). We observed potent anti-inflammatory activity that phenocopied the
208 effect of disrupting *ETS2* or deleting the chr21q22 enhancer (**Fig.5f, Extended Data Fig.8**),
209 with downregulation of multiple *ETS2*-regulated pathways (including several drug targets;
210 **Fig.5g**). To further explore the therapeutic potential of targeting *ETS2* signalling⁴², we
211 employed a human gut explant model. Intestinal mucosal biopsies were obtained from
212 patients with active IBD, who were not receiving immunosuppressive or biologic therapies,
213 and cultured with either a MEK inhibitor or a negative or positive control (**Methods**). Release
214 of multiple IBD-associated inflammatory cytokines was significantly reduced by MEK inhibition
215 – to comparable levels observed with infliximab (an anti-TNF α antibody widely used for IBD;
216 **Fig.5h**). Moreover, we confirmed that expression of *ETS2*-regulated genes was reduced
217 (**Fig.5i, Extended Data Fig.8**) and that there was significant improvement in a validated
218 transcriptional inflammation score⁴³ that reflects IBD-associated inflammation and has been
219 shown to reduce upon effective therapy (**Fig.5j**). Together, this shows that targeting an
220 upstream regulator of *ETS2* can abrogate pathological inflammation in a chr21q22-associated
221 disease, and may be useful therapeutically.

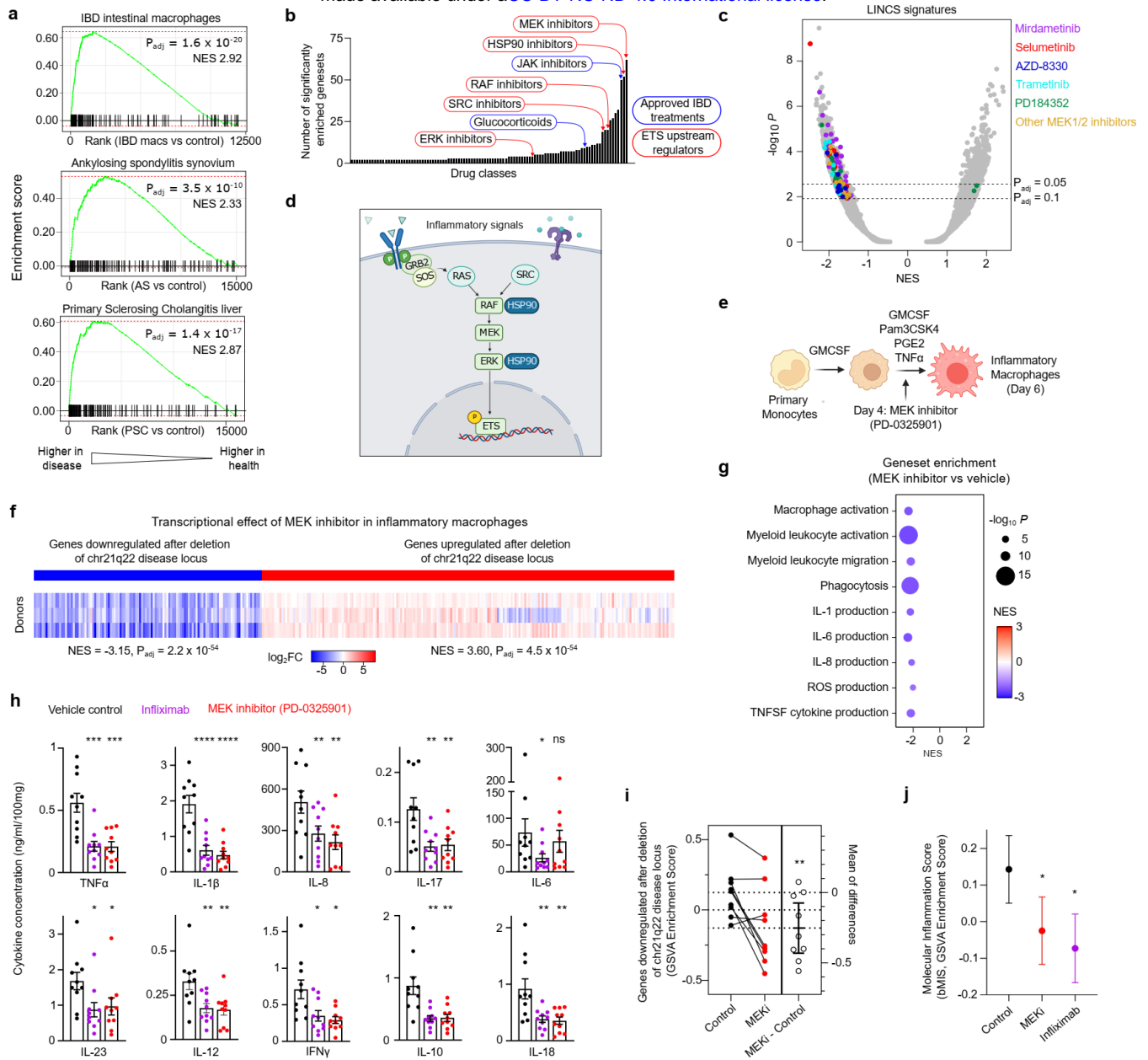


Figure 5. ETS2-driven inflammation is evident in disease and can be therapeutically targeted.

a. Enrichment of chr21q22-regulated genes in IBD intestinal macrophages (top), AS synovium (middle), and PSC liver (bottom). All relative to unaffected tissue. Gene set comprised significantly downregulated genes following chr21q22 deletion. **b.** Candidate drug classes (from NIH LINC1362 database) that phenocopy the transcriptional consequences of *ETS2* disruption. **c.** fGSEA results for NIH LINC1362 drug signatures (FDR P estimated using adaptive multi-level split Monte-Carlo scheme; NES, normalised enrichment score). Significantly enriched MEK inhibitor gene sets coloured by molecule. **d.** Schematic of known upstream regulators of *ETS*-family transcription factors. **e.** Schematic of experiment for treating inflammatory macrophages with a MEK inhibitor (PD-0325901) **f.** Heatmap of relative expression (log₂ fold-change) of chr21q22-regulated genes in inflammatory macrophages following MEK inhibition (compared to vehicle control, n=3). **g.** fGSEA of differentially-expressed genes between MEK inhibitor-treated and control inflammatory macrophages. Results shown for Gene Ontology Biological Pathways that were negatively enriched following *ETS2* editing. Dot size represents P -value and colour denotes NES. **h.** Cytokine secretion from IBD mucosal biopsies cultured for 18 hours with vehicle control, PD-0325901, or infliximab. **i.** Estimation plot of GSEA enrichment scores for chr21q22-downregulated genes in IBD intestinal biopsies following MEK inhibition (MEKi). Error bars indicate 95%CI. **j.** GSEA enrichment scores of a biopsy-derived molecular inflammation score (bMIS). Data in **h** and **i** represent mean \pm SEM. Wilcoxon matched-pairs test, two-tailed, n=10 (h), n=9 (j). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$.

222 Discussion

223 Arguably the greatest challenge in modern genetics is to translate the success of GWAS into a
224 better understanding of human disease. Here, we show how this can provide insights into
225 basic immunology as well as disease mechanisms, with investigation of a single pleiotropic
226 risk locus leading to the discovery of a master regulator of human macrophage responses and
227 a key pathogenic pathway that is potentially druggable. Monocyte-derived macrophages play
228 an important role in the pathogenesis of many inflammatory diseases¹⁴, producing cytokines
229 that are often targeted therapeutically. Blocking individual cytokines, however, is typically only
230 effective in 30-40% of patients⁴⁴ and there is a growing realisation that modulating several
231 cytokines at once may be a better approach⁴⁵. Modulating ETS2 signalling via MEK1/2
232 inhibition has a broad effect on multiple inflammatory cytokines, including TNF α , IL-23, and IL-
233 12 – all targets of existing therapies – and IL-1 β which has previously been linked to
234 treatment-refractory IBD⁴⁶ but is not directly modulated by other small molecules (e.g. JAK
235 inhibitors). However, systemic use of MEK inhibitors may not be an ideal strategy for treating
236 chronic disease due to the physiological roles of MEK in other tissues. Indeed, use of MEK
237 inhibitors for other conditions is currently limited by severe ocular, cardiac, and pulmonary
238 toxicities⁴⁷. Targeting ETS2 directly – for example using PROTACs or molecular glues – or
239 selectively delivering a MEK inhibitor to inflammatory macrophages via antibody-drug
240 conjugates could potentially overcome this toxicity, and provide a safer means of inhibiting
241 ETS2-driven inflammation.

242 These findings emphasise the importance of studying disease mechanisms beyond simple
243 changes in gene expression, especially since the overlap between the chr21q22 risk
244 haplotype and a macrophage eQTL had been noted previously⁵ without any indication as to
245 the importance of the downstream biology. Indeed, it is even possible that *ETS2* is involved in
246 other human pathology, aside from diseases associated with the chr21q22 locus. For
247 example, Down's syndrome – caused by trisomy of chromosome 21 – was recently proposed
248 to be a cytokinopathy⁴⁸, with increased basal expression of multiple inflammatory cytokines,

249 including several that are specifically upregulated following *ETS2* overexpression (e.g. IL-1 β ,
250 TNF α , and IL-6). Whether the additional copy of the *ETS2* gene contributes to this phenotype
251 is unknown, but warrants further study. Relatedly, it is interesting to consider why a
252 polymorphism that increases susceptibility to multiple inflammatory diseases is so common
253 (risk allele frequency ~75% in Europeans and >90% in Africans, **Extended Data Fig.9**). One
254 possibility is that enhanced macrophage effector responses might provide a selective
255 advantage during infection, which would explain why the ancestral risk allele has not been
256 negatively selected. Recent studies have not found evidence for a strong selective sweep at
257 this locus within the past few thousand years^{49,50}, but more ancient selection – or balancing
258 selection maintaining variation – cannot be excluded, especially since rs2836882 appears to
259 be an exceptionally old SNP (conservatively estimated at >500,000 years old; **Extended Data**
260 **Fig.9**) and was even polymorphic between Neanderthals and Denisovans (**Extended Data**
261 **Fig.9**).

262 In summary, using an intergenic GWAS hit as a starting point, we have identified a druggable
263 pathway that is both necessary and sufficient for human macrophage responses during
264 chronic inflammation. Furthermore, we show how genetic dysregulation of this pathway – via
265 perturbation of pioneer factor binding at a critical long-range enhancer – confers susceptibility
266 to multiple inflammatory diseases. This highlights the considerable, yet largely untapped,
267 opportunity to better understand disease biology that is presented by non-coding genetic
268 associations.

269 **References**

- 270 1 Miller, F. W. The increasing prevalence of autoimmunity and autoimmune diseases: an
271 urgent call to action for improved understanding, diagnosis, treatment, and prevention.
272 *Curr Opin Immunol* **80**, 102266, doi:10.1016/j.coi.2022.102266 (2023).
- 273 2 Dowden, H. & Munro, J. Trends in clinical success rates and therapeutic focus. *Nat*
274 *Rev Drug Discov* **18**, 495-496, doi:10.1038/d41573-019-00074-z (2019).
- 275 3 de Lange, K. M. *et al.* Genome-wide association study implicates immune activation of
276 multiple integrin genes in inflammatory bowel disease. *Nat Genet* **49**, 256-261,
277 doi:10.1038/ng.3760 (2017).
- 278 4 International Genetics of Ankylosing Spondylitis, C. *et al.* Identification of multiple risk
279 variants for ankylosing spondylitis through high-density genotyping of immune-related
280 loci. *Nat Genet* **45**, 730-738, doi:10.1038/ng.2667 (2013).
- 281 5 Ji, S. G. *et al.* Genome-wide association study of primary sclerosing cholangitis
282 identifies new risk loci and quantifies the genetic relationship with inflammatory bowel
283 disease. *Nat Genet* **49**, 269-273, doi:10.1038/ng.3745 (2017).
- 284 6 Ortiz-Fernandez, L. *et al.* Identification of susceptibility loci for Takayasu arteritis
285 through a large multi-ancestral genome-wide association study. *Am J Hum Genet* **108**,
286 84-99, doi:10.1016/j.ajhg.2020.11.014 (2021).
- 287 7 Martin, J. C. *et al.* Single-Cell Analysis of Crohn's Disease Lesions Identifies a
288 Pathogenic Cellular Module Associated with Resistance to Anti-TNF Therapy. *Cell*
289 **178**, 1493-1508 e1420, doi:10.1016/j.cell.2019.08.008 (2019).
- 290 8 Stathias, V. *et al.* LINCS Data Portal 2.0: next generation access point for perturbation-
291 response signatures. *Nucleic Acids Res* **48**, D431-D439, doi:10.1093/nar/gkz1023
292 (2020).
- 293 9 Harrison, R. K. Phase II and phase III failures: 2013-2015. *Nat Rev Drug Discov* **15**,
294 817-818, doi:10.1038/nrd.2016.184 (2016).
- 295 10 Claussnitzer, M. *et al.* A brief history of human disease genetics. *Nature* **577**, 179-189,
296 doi:10.1038/s41586-019-1879-7 (2020).
- 297 11 King, E. A., Davis, J. W. & Degner, J. F. Are drug targets with genetic support twice as
298 likely to be approved? Revised estimates of the impact of genetic support for drug
299 mechanisms on the probability of drug approval. *PLoS Genet* **15**, e1008489,
300 doi:10.1371/journal.pgen.1008489 (2019).
- 301 12 Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**,
302 934-947, doi:S0092-8674(13)01227-0 [pii]
303 10.1016/j.cell.2013.09.053 (2013).
- 304 13 Kugathasan, S. *et al.* Loci on 20q13 and 21q22 are associated with pediatric-onset
305 inflammatory bowel disease. *Nat Genet* **40**, 1211-1215, doi:10.1038/ng.203 (2008).
- 306 14 Park, M. D., Silvin, A., Ginhoux, F. & Merad, M. Macrophages in health and disease.
307 *Cell* **185**, 4259-4279, doi:10.1016/j.cell.2022.10.007 (2022).
- 308 15 Javierre, B. M. *et al.* Lineage-Specific Genome Architecture Links Enhancers and Non-
309 coding Disease Variants to Target Gene Promoters. *Cell* **167**, 1369-1384 e1319,
310 doi:10.1016/j.cell.2016.09.037 (2016).
- 311 16 Fairfax, B. P. *et al.* Innate immune activity conditions the effect of regulatory variants
312 upon monocyte gene expression. *Science* **343**, 1246949, doi:343/6175/1246949 [pii]

- 313 10.1126/science.1246949 (2014).
- 314 17 Xue, J. *et al.* Transcriptome-based network analysis reveals a spectrum model of
315 human macrophage activation. *Immunity* **40**, 274-288,
316 doi:10.1016/j.immuni.2014.01.006 (2014).
- 317 18 Kuo, D. *et al.* HBEGF(+) macrophages in rheumatoid arthritis induce fibroblast
318 invasiveness. *Sci Transl Med* **11**, doi:10.1126/scitranslmed.aau8587 (2019).
- 319 19 Melnikov, A. *et al.* Systematic dissection and optimization of inducible enhancers in
320 human cells using a massively parallel reporter assay. *Nat Biotechnol* **30**, 271-277,
321 doi:10.1038/nbt.2137 (2012).
- 322 20 Nerlov, C. & Graf, T. PU.1 induces myeloid lineage commitment in multipotent
323 hematopoietic progenitors. *Genes Dev* **12**, 2403-2412, doi:10.1101/gad.12.15.2403
324 (1998).
- 325 21 Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime
326 cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**,
327 576-589, doi:10.1016/j.molcel.2010.05.004 (2010).
- 328 22 de Santiago, I. *et al.* BaalChIP: Bayesian analysis of allele-specific transcription factor
329 binding in cancer genomes. *Genome Biol* **18**, 39, doi:10.1186/s13059-017-1165-7
330 (2017).
- 331 23 Martinez, L. A. Mutant p53 and ETS2, a Tale of Reciprocity. *Front Oncol* **6**, 35,
332 doi:10.3389/fonc.2016.00035 (2016).
- 333 24 Wei, G. *et al.* Activated Ets2 is required for persistent inflammatory responses in the
334 motheaten viable model. *J Immunol* **173**, 1374-1379, doi:10.4049/jimmunol.173.2.1374
335 (2004).
- 336 25 Zhao, J., Huang, K., Peng, H. Z. & Feng, J. F. Protein C-ets-2 epigenetically
337 suppresses TLRs-induced interleukin 6 production in macrophages. *Biochem Biophys*
338 *Res Commun* **522**, 960-964, doi:10.1016/j.bbrc.2019.11.123 (2020).
- 339 26 Chung, S. W., Chen, Y. H. & Perrella, M. A. Role of Ets-2 in the regulation of heme
340 oxygenase-1 by endotoxin. *J Biol Chem* **280**, 4578-4584, doi:10.1074/jbc.M409125200
341 (2005).
- 342 27 Quinn, S. R. *et al.* The role of Ets2 transcription factor in the induction of microRNA-
343 155 (miR-155) by lipopolysaccharide and its targeting by interleukin-10. *J Biol Chem*
344 **289**, 4316-4325, doi:10.1074/jbc.M113.522730 (2014).
- 345 28 Ma, X. *et al.* Ets2 suppresses inflammatory cytokines through MAPK/NF-kappaB
346 signaling and directly binds to the IL-6 promoter in macrophages. *Aging (Albany NY)*
347 **11**, 10610-10625, doi:10.18632/aging.102480 (2019).
- 348 29 Aperlo, C., Pognonec, P., Stanley, E. R. & Boulukos, K. E. Constitutive c-ets2
349 expression in M1D+ myeloblast leukemic cells induces their differentiation to
350 macrophages. *Mol Cell Biol* **16**, 6851-6858, doi:10.1128/MCB.16.12.6851 (1996).
- 351 30 Henkel, G. W. *et al.* PU.1 but not ets-2 is essential for macrophage development from
352 embryonic stem cells. *Blood* **88**, 2917-2926 (1996).
- 353 31 Mittal, M., Siddiqui, M. R., Tran, K., Reddy, S. P. & Malik, A. B. Reactive oxygen
354 species in inflammation and tissue injury. *Antioxid Redox Signal* **20**, 1126-1167,
355 doi:10.1089/ars.2012.5149 (2014).
- 356 32 Kelly, B. & O'Neill, L. A. Metabolic reprogramming in macrophages and dendritic cells
357 in innate immunity. *Cell Res* **25**, 771-784, doi:10.1038/cr.2015.68 (2015).

- 358 33 Sazonovs, A. *et al.* Large-scale sequencing identifies multiple genes and rare variants
359 associated with Crohn's disease susceptibility. *Nat Genet* **54**, 1275-1283,
360 doi:10.1038/s41588-022-01156-2 (2022).
- 361 34 Slowikowski, K., Hu, X. & Raychaudhuri, S. SNPsea: an algorithm to identify cell types,
362 tissues and pathways affected by risk loci. *Bioinformatics* **30**, 2496-2497,
363 doi:10.1093/bioinformatics/btu326 (2014).
- 364 35 Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome.
365 *Nature* **489**, 57-74, doi:nature11247 [pii]
366 10.1038/nature11247 (2012).
- 367 36 Cramer, T. *et al.* HIF-1alpha is essential for myeloid cell-mediated inflammation. *Cell*
368 **112**, 645-657, doi:10.1016/s0092-8674(03)00154-5 (2003).
- 369 37 Tannahill, G. M. *et al.* Succinate is an inflammatory signal that induces IL-1beta
370 through HIF-1alpha. *Nature* **496**, 238-242, doi:10.1038/nature11986 (2013).
- 371 38 Shiratori, R. *et al.* Glycolytic suppression dramatically changes the intracellular
372 metabolic profile of multiple cancer cell lines in a mitochondrial metabolism-dependent
373 manner. *Sci Rep* **9**, 18699, doi:10.1038/s41598-019-55296-3 (2019).
- 374 39 Landt, S. G. *et al.* ChIP-seq guidelines and practices of the ENCODE and
375 modENCODE consortia. *Genome Res* **22**, 1813-1831, doi:10.1101/gr.136184.111
376 (2012).
- 377 40 Sevilla, L. *et al.* Bcl-XL expression correlates with primary macrophage differentiation,
378 activation of functional competence, and survival and results from synergistic
379 transcriptional activation by Ets2 and PU.1. *J Biol Chem* **276**, 17800-17807,
380 doi:10.1074/jbc.M008270200 (2001).
- 381 41 Newman, J. A., Cooper, C. D., Aitkenhead, H. & Gileadi, O. Structural insights into the
382 autoregulation and cooperativity of the human transcription factor Ets-2. *J Biol Chem*
383 **290**, 8539-8549, doi:10.1074/jbc.M114.619270 (2015).
- 384 42 Koboziev, I., Karlsson, F., Zhang, S. & Grisham, M. B. Pharmacological intervention
385 studies using mouse models of the inflammatory bowel diseases: translating preclinical
386 data into new drug therapies. *Inflamm Bowel Dis* **17**, 1229-1245,
387 doi:10.1002/ibd.21557 (2011).
- 388 43 Argmann, C. *et al.* Biopsy and blood-based molecular biomarker of inflammation in
389 IBD. *Gut*, doi:10.1136/gutjnl-2021-326451 (2022).
- 390 44 Alsoud, D., Verstockt, B., Fiocchi, C. & Vermeire, S. Breaking the therapeutic ceiling in
391 drug development in ulcerative colitis. *Lancet Gastroenterol Hepatol* **6**, 589-595,
392 doi:10.1016/S2468-1253(21)00065-0 (2021).
- 393 45 Feagan, B. G. *et al.* Guselkumab plus golimumab combination therapy versus
394 guselkumab or golimumab monotherapy in patients with ulcerative colitis (VEGA): a
395 randomised, double-blind, controlled, phase 2, proof-of-concept trial. *Lancet*
396 *Gastroenterol Hepatol* **8**, 307-320, doi:10.1016/S2468-1253(22)00427-7 (2023).
- 397 46 Friedrich, M. *et al.* IL-1-driven stromal-neutrophil interactions define a subset of
398 patients with inflammatory bowel disease that does not respond to therapies. *Nat Med*
399 **27**, 1970-1981, doi:10.1038/s41591-021-01520-5 (2021).
- 400 47 Klesse, L. J. *et al.* The Use of MEK Inhibitors in Neurofibromatosis Type 1-Associated
401 Tumors and Management of Toxicities. *Oncologist* **25**, e1109-e1116,
402 doi:10.1634/theoncologist.2020-0069 (2020).

- 403 48 Malle, L. *et al.* Autoimmunity in Down's syndrome via cytokines, CD4 T cells and
404 CD11c(+) B cells. *Nature* **615**, 305-314, doi:10.1038/s41586-023-05736-y (2023).
- 405 49 Irving-Pease, E. K. *et al.* The Selection Landscape and Genetic Legacy of Ancient
406 Eurasians. *bioRxiv*, 2022.2009.2022.509027, doi:10.1101/2022.09.22.509027 (2022).
- 407 50 Le, M. K. *et al.* 1,000 ancient genomes uncover 10,000 years of natural selection in
408 Europe. *bioRxiv*, 2022.2008.2024.505188, doi:10.1101/2022.08.24.505188 (2022).

409 **Methods**

410 **Analysis of existing data relating to chr21q22**

411 We used IBD GWAS summary statistics³ to perform multiple causal variant fine-mapping
412 using susieR⁵¹ with reference minor allele and LD information calculated from 503 European
413 samples from 1000 Genomes phase 3. All R analyses used v.4.2.1. Palindromic SNPs (A/T or
414 C/G) and any SNPs that didn't match by position or alleles were pruned before imputation
415 using the *ssimp* equations reimplemented in R. This did not affect any candidate SNP at
416 chr21q22. We obtained SuSiE fine mapping results for *ETS2* (with identifier
417 ENSG00000157557 or ILMN_1720158) in monocyte datasets from the eQTL Catalog.
418 Colocalisation analysis was performed using coloc v5.2.0⁵² using a posterior probability of H4
419 (PP.H4.abf) > 0.5 to call colocalisation.

420 Raw H3K27ac ChIP-seq data from primary human immune cells were downloaded from Gene
421 Expression Omnibus (GEO series GSE18927 and GSE96014) and processed as described
422 previously⁵³.

423 Processed promoter-capture Hi-C data¹⁵ from 17 primary immune cell-types were downloaded
424 from OSF (<https://osf.io/u8tzp>).

425

426 **Monocyte purification and macrophage differentiation**

427 Leukocyte cones from healthy donors were obtained from NHS Blood and Transplant
428 (Cambridge Blood Donor Centre, Colindale Blood Centre, or Tooting Blood Donor Centre).
429 Peripheral blood mononuclear cells (PBMCs) were isolated by density centrifugation
430 (Histopaque 1077, Sigma) and monocytes were positively selected using CD14 Microbeads
431 (Miltenyi Biotec). Macrophage differentiation was performed using conditions that model
432 chronic inflammation (TPP)¹⁷: 3 days GM-CSF (50ng/ml, Peprotech) followed by 3 days GM-
433 CSF, TNF α (50ng/ml, Peprotech), PGE₂ (1 μ g/ml, Sigma Aldrich), and Pam₃CSK4 (1 μ g/ml,
434 Invivogen). All cultures were performed at 37°C, 5% CO₂ in antibiotic-free RPMI1640 media
435 containing 10% FBS, GlutaMAX, and MEM Non-Essential Amino Acids (all ThermoFisher).
436 Cells were detached using Accutase (Biolegend).

437

438 **Identifying a model of chronic inflammatory macrophages**

439 Human monocyte-derived macrophage gene expression data files (n=299) relating to 64
440 different polarisation conditions were downloaded from Gene Expression Omnibus
441 (GSE47189) and quantile normalised. Data from biological replicates were summarised to the
442 median value for every gene. Gene set variation analysis⁵⁴ (using the GSEA package in R)
443 was performed to identify the polarisation condition that most closely resembled CD14+
444 monocytes/macrophages from active IBD – using disease-associated lists of differentially-
445 expressed⁵⁵.

446

447 **CRISPR-Cas9 editing of primary human monocytes**

448 gRNA sequences were designed using CRISPick (formerly GPP sgRNA Designer) and
449 synthesised by IDT. gRNA sequences: chr21q22 5' gRNA, CCUGGCUGCCUCGCGUUUCC;
450 chr21q22 3' gRNA, CCUCGUCCAACAGAGAGCAA; ETS2 gRNA1,
451 CAGACACAGAAUUACCCCAA; ETS2 gRNA2, UUGCUGCACGGGGUUAACAA. Alt-R
452 CRISPR-Cas9 Negative Control crRNA #1 (IDT) was used as a non-targeting control. Cas9-
453 gRNA ribonucleoproteins were assembled as described previously⁵³ and nucleofected into
454 5×10^6 monocytes in 100 μ L nucleofection buffer (Human Monocyte Nucleofection Kit, Lonza)
455 using a Nucleofector 2b (Lonza, program Y-001). After nucleofection, monocytes were
456 immediately transferred into 5ml of pre-warmed culture media in a 6-well flat-bottomed plate,
457 and differentiated into macrophages under TPP conditions. Editing efficiency was quantified
458 by PCR amplification of the target region in extracted DNA (chr21q22_Fw primer,
459 GGTGGGGAGAGTTCCAAAGG; chr21q22_Rv, TCACCCTTCACCTCTTTGCT;
460 ETS2_g1_Fw, TCCTGAAGGTCCCATGAAAG; ETS2_g1_Rv, TCATTATGGCTCTGGGGTTC;
461 ETS2_g2_Fw, GCGGCACATTCATATCACAC; ETS2_g2_Rv,
462 GCAGAATACCCCAAGCAAAA). Editing efficiency at the chr21q22 locus was measured via
463 quantification of amplified fragments (2100 Bioanalyzer, Agilent) as previously described⁵³.

464 Editing efficiency for individual gRNAs was assessed using the Inference of CRISPR Edits
465 tool⁵⁶ (ICE, Synthego).

466

467 **PrimeFlow RNA Assay**

468 RNA abundance was quantified by PrimeFlow (ThermoFisher) in chr21q22-edited and
469 unedited (NTC) cells on days 0, 3, 4, 5, and 6 of TPP differentiation. Target probes specific for
470 *ETS2* (Alexa Fluor 647), *BRWD1* (Alexa Fluor 568) and *PSMG1* (Alexa Fluor 568) were used
471 according to the manufacturer's instructions. Data were analysed using FlowJo v10 (BD
472 Biosciences).

473

474 **MPRA**

475 Overlapping oligonucleotides containing 114-nt of genomic sequence were designed to tile the
476 region containing chr21q22 candidate SNPs (99% credible set) at 50bp intervals. Six technical
477 replicates were designed for every genomic sequence, each tagged by a unique 11-nt
478 barcode. Additional oligonucleotides were included to test the expression-modulating effect of
479 every candidate SNP in the 99% credible set. Allelic constructs were designed as described
480 previously⁵³ and tagged by 30 unique 11-nt barcodes. Positive and negative controls were
481 included as described previously⁵³. 170-nt oligonucleotides were synthesised as part of a
482 larger MPRA pool (Twist Biosciences) containing the 16-nt universal primer site
483 ACTGGCCGCTTCACTG, 114-nt variable genomic sequence, KpnI and XbaI restriction sites
484 (TGGACCTCTAGA), an 11-nt barcode, and the 17-nt universal primer site
485 AGATCGGAAGAGCGTTCG. Cloning into the MPRA vector was performed as described
486 previously⁵³. A suitable promoter for the MPRA vector (RSV) was identified by testing
487 promoter activities in TPP macrophages. The MPRA vector library was nucleofected into TPP
488 macrophages (5µg vector into 5x10⁶ cells) in 100µl nucleofection buffer (Human Macrophage
489 Nucleofection Kit, Lonza) using a Nucleofector 2b (program Y-011). To ensure adequate
490 barcode representation, a minimum of 2x10⁷ cells were nucleofected for every donor (n=8).
491 After 24 hours, RNA was extracted and sequencing libraries were made from mRNA or DNA

492 input vector as described previously⁵³. Library pools (of 6 samples each) were sequenced on
493 an Illumina HiSeq2500 high output flow-cell (50bp, single-end reads) and data were pre-
494 processed as previously described⁵³. To identify regions of enhancer activity, a paired t-test
495 was performed to identify genomic sequences that enhanced transcription. A sliding window
496 analysis (300-bp window) was then performed across all tiling sequences using the *les*
497 package in R. Expression-modulating variants were identified using QuASAR-MPRA⁵⁷, as
498 described previously⁵³.

499

500 **BaalChIP**

501 Publicly-available PU.1 ChIP-seq datasets from human macrophages were downloaded from
502 GEO, and BAM files were examined (using the IGV genome browser) to identify rs2836882
503 heterozygotes (i.e. files containing both A and G allele reads at chr21:40,466,570; hg19). Two
504 suitable samples were identified (GSM1681423 and GSM1681429) which were used for a
505 Bayesian analysis of allelic imbalances in PU.1 binding, with correction for biases introduced
506 by overdispersion and biases towards the reference allele – implemented in the *BaalChIP*
507 package²² in R.

508

509 **Allele-specific PU.1 ChIP-genotyping**

510 A 100ml blood sample was taken from five healthy individuals who were heterozygous at
511 rs2836882 (assessed via Taqman genotyping, ThermoFisher). All participants provided written
512 informed consent. Ethical approval was provided by the London - Brent Regional Ethics
513 Committee (REC: 21/LO/0682). Monocytes were isolated from PBMC using CD14 Microbeads
514 (Miltenyi Biotec) and differentiated into inflammatory macrophages using TPP conditions¹⁷.
515 Following differentiation, macrophages were detached using Accutase and cross-linked for 10
516 min in fresh media containing 1% formaldehyde. Cross-linking was quenched with glycine for
517 5 min (final concentration 0.125M). Nuclei preparation and shearing were performed as
518 described previously⁵³ with 10 cycles sonication (30s ON/30s OFF, Bioruptor Pico,
519 Diagenode). PU.1 was immunoprecipitated overnight at 4°C using a polyclonal anti-PU.1

520 antibody (1:25; Cell Signaling) with the SimpleChIP Plus kit (Cell Signaling). The ratio of
521 rs2836882 alleles in the PU.1-bound DNA was quantified in duplicate by TaqMan genotyping
522 (assay C___2601507_20). A standard curve was generated using fixed ratios of geneblocks
523 containing either the risk or non-risk allele (200-nt genomic sequence centred on rs2836882;
524 Genewiz).

525

526 **PU.1 MPRA-ChIP-seq**

527 The MPRA vector library was transfected into TPP macrophages from six healthy donors.
528 Assessment of PU.1 binding to SNP alleles was performed as described previously⁵³, with
529 minimal sonication (to remove contaminants while minimising chromatin shearing).
530 Immunoprecipitation was performed overnight at 4°C using a polyclonal anti-PU.1 antibody
531 (1:25; Cell Signaling) with the SimpleChIP Plus kit (Cell Signaling). Sequencing libraries were
532 prepared from isolated plasmids as for MPRA and sequencing on a MiSeq (50bp, single-end
533 reads).

534

535 **H3K27ac ChIP-seq**

536 TPP macrophages from two rs2836882 major allele homozygotes and two minor allele
537 homozygotes were harvested, cross-linked, and quenched as described above. Donors were
538 identified through the NIHR BioResource. H3K27ac ChIP-seq was performed as described
539 previously⁵³ using an anti-H3K27ac antibody (1:250, Abcam) or an isotype control (1:500,
540 rabbit IgG, Abcam). Libraries were sequenced on a HiSeq4000 (50bp, single-end reads). Raw
541 data were processed, QC'd, and analysed as described previously⁵³ (see Code Availability).

542

543 **Assays of macrophage effector functions**

544 *Flow cytometry*

545 Expression of myeloid markers was assessed by flow cytometry (BD LSRFortessa™ X-20).
546 Panel: CD11b PE/Dazzle 594 (BioLegend), CD14 evolve605 (ThermoFisher), CD16 PerCP
547 (BioLegend), CD68 FITC (BioLegend), Live/Dead Fixable Aqua Dead Cell Stain

548 (ThermoFisher), and Fc Receptor Blocking Reagent (Miltenyi). Data were analysed using
549 FlowJo v10 (BD Biosciences).

550

551 *Cytokine quantification*

552 Supernatants were collected on day 6 of TPP macrophage culture and frozen. Cytokine
553 concentrations were quantified in duplicate via electrochemiluminescence using U-PLEX
554 assays (Meso Scale Diagnostics).

555

556 *Phagocytosis*

557 Phagocytosis was assessed using fluorescently-labelled Zymosan particles (Green Zymosan,
558 Abcam) according to the manufacturer's instructions. Cells were seeded at 10^5 cells/well in
559 96-well round bottom plates. Cytochalasin D ($10\mu\text{g/ml}$, ThermoFisher), an inhibitor of
560 cytoskeletal rearrangement, was used as a negative control. Phagocytosis was quantified via
561 flow cytometry, and a phagocytosis index was calculated (proportion of positive cells multiplied
562 by their mean fluorescence intensity).

563

564 *Extracellular ROS production*

565 Extracellular ROS production was quantified using a Diogenes Enhanced Superoxide
566 Detection Kit (National Diagnostics) according to the manufacturer's protocol. Cells were
567 seeded at a density of 10^5 cells/well and pre-stimulated with PMA (200ng/ml , Sigma Aldrich).

568

569 *Western blotting*

570 Western blotting was performed as described previously⁵⁸ using the following primary
571 antibodies: rabbit anti-gp91 phox , rabbit anti-p22 phox (both Santa Cruz), rabbit anti-
572 C17ORF62/EROS (Atlas), rabbit anti-actin (Abcam). Secondary antibody was anti-rabbit IgG-
573 horseradish peroxidase (Cell Signaling). Chemiluminescence was recorded on a ChemiDoc
574 Touch imager (Bio-Rad) following incubation of the membrane with ECL (ThermoFisher) or
575 SuperSignal West Pico PLUS (ThermoFisher) reagent.

576

577 **RNA sequencing**

578 RNA was isolated from macrophage lysates (AllPrep DNA/RNA Micro Kit, Qiagen) and
579 sequencing libraries prepared from 10ng RNA using the SMARTer Stranded Total RNA-Seq
580 Kit v2 - Pico Input Mammalian (Takara) following the manufacturer's instructions. Libraries
581 were sequenced on a NextSeq2000 (50bp, PE reads: CRISPR-based loss-of-function,
582 roxadustat and PD-0325901 experiments) or a NovaSeq6000 (100bp, PE reads:
583 overexpression experiments). Reads were trimmed using Trim Galore (Phred score 24),
584 filtered to remove reads < 20bp, and ribosomal reads were removed using the BBSplit
585 function of BBDMap (BBDMap, sourceforge.net/projects/bbmap/) with the Human ribosomal DNA
586 complete repeating unit (GenBank: U13369.1). Reads were aligned to the human genome
587 (hg38) using HISAT2 (ref.⁵⁹) and converted to BAM files, sorted and indexed using
588 SAMtools⁶⁰. Gene read counts were obtained with the featureCounts program⁶¹ from
589 Rsubread using the GTF annotation file for human genome build GRCh38 (version 102).
590 Differential expression analysis was performed in R using the *limma* package⁶² with the voom
591 transformation and including donor as a covariate.

592

593 **Gene set enrichment analysis**

594 GSEA was performed using the *fgsea*⁶³ package in R. Gene sets were either obtained from
595 Gene Ontology Biological Pathways (downloaded from MSigDB), experimentally-derived
596 based on differential expression analysis, or sourced from published literature⁷. Pathways
597 shown in Figures 2-5 are: GO:0002274, GO:0042116, GO:0097529, GO:0006909,
598 GO:0071706, GO:0032732, GO:0032755, GO:0032757, GO:2000379, GO:0009060,
599 GO:0006119, and GO:0045649. Statistical significance was calculated using the adaptive
600 multilevel split Monte Carlo method.

601

602 ***In vitro* transcription**

603 The cDNA sequence for *ETS2* (NM005329.5) preceded by a Kozak sequence was
604 synthesised and cloned into a TOPO vector. This was linearised and a PCR amplicon of the
605 *ETS2* gene generated, adding a T7 promoter and an AG initiation sequence (Phusion, NEB):
606 Fw primer: GCTAATACGACTCACTATAAGGACAGGCCACCATGAATGATTTTCGGAATC, Rv
607 primer: TCAGTCCTCCGTGTCGG). A reverse complement (control) amplicon was also
608 generated: Fw primer:
609 GCTAATACGACTCACTATAAGGACAGGCCACCTCAGTCCTCCGTGTCGG, Rv primer:
610 GCCACCATGAATGATTTTCGGAATC). These amplicons were used as templates for *in vitro*
611 transcription using the HiScribe T7 mRNA Kit with CleanCap® Reagent AG kit (NEB)
612 according to the manufacturer's instructions, but with substitution of N1-methyl-pseudouridine
613 for uridine and methylcytidine for cytidine (both Stratech) to minimise non-specific cellular
614 activation by the transfected mRNA. mRNA was purified using a MEGAclean Transcription
615 Clean-Up Kit (ThermoFisher) and polyadenylated using an *E. coli* Poly(A) Polymerase (NEB)
616 before further clean-up (MEGAclean), quantification and analysis of product size
617 (NorthernMax®-Gly gel, ThermoFisher). For optimising overexpression conditions, GFP
618 mRNA was produced using the same method: Fw primer
619 GCTAATACGACTCACTATAAGGACAGGCCACCATGGTGAGCAAGGGCGAG, Rv primer
620 TTAATTGTACAGCTCGTCCATGC).

621

622 **mRNA overexpression.**

623 Lipofectamine MessengerMAX (ThermoFisher) was diluted in Opti-MEM (1:75 v/v), vortexed
624 and incubated at room temperature for 10 minutes. IVT mRNA was then diluted in a fixed
625 volume of Opti-MEM (112.5µl per transfection), mixed with an equal volume of diluted
626 Lipofectamine MessengerMAX and incubated for a further 5 minutes at room temperature.
627 The transfection mix was then added dropwise to 2.5×10^6 M0 macrophages (pre-cultured for 6
628 days in a 6-well plate in antibiotic-free RPMI1640 macrophage media containing M-CSF
629 (50ng/ml, Peprotech) – with media change on d3). For GFP overexpression, cells were
630 detached using Accutase 18 hours after transfection and GFP expression was measured by

631 flow cytometry. For *ETS2* / control overexpression, either 250ng or 500ng mRNA was
632 transfected and low dose LPS (0.5ng/ml) was added 18 hours after transfection, and cells
633 detached using Accutase 6 hours later (n=8 donors). Representative *ETS2* expression in
634 untransfected macrophages obtained from previous data (GSE193336).

635

636 **SNPsea**

637 Pathway analysis of 241 IBD-associated GWAS hits³ was performed using SNPsea³⁴. In brief,
638 linkage intervals were defined for every lead SNP based on the furthest correlated SNPs ($r^2 >$
639 0.5 in 1000 Genomes, EUR population) and extended to the nearest recombination hotspots
640 with recombination rate >3 cM/Mb. If no genes were present in this region, the linkage interval
641 was extended up- and down-stream by 500kb. Genes within linkage intervals were tested for
642 enrichment within 7,660 pathways, comprising 7,658 Gene Ontology Biological Pathways and
643 2 lists of *ETS2*-regulated genes (either those significantly downregulated following *ETS2*
644 disruption with gRNA1 or those significantly upregulated following *ETS2* overexpression,
645 based on a consensus list obtained from differential expression analysis including all samples
646 and using donor and mRNA quantity as covariates). The analysis was performed using a
647 single score mode: assuming that only one gene per linkage interval is associated with the
648 pathway. A null distribution of scores for each pathway was performed by sampling random
649 SNP sets matched on the number of linked genes (5,000,000 iterations). A permutation *P*-
650 value was calculated by comparing the enrichment of the IBD-associated gene list with the
651 null distribution. Gene sets relating to the following IBD-associated pathways were extracted
652 for comparison: NOD2 signalling (GO:0032495), Integrin signalling (GO:0033627,
653 GO:0033622), TNF α signalling (GO:0033209, GO:0034612), Intestinal epithelium
654 (GO:0060729, GO:0030277), Th17 cells (GO:0072539, GO:0072538, GO:2000318), T cell
655 activation (GO:0046631, GO:0002827), IL-10 signalling (GO:0032613, GO:0032733), and
656 autophagy (GO:0061919, GO:0010506, GO:0010508, GO:1905037, GO:0010507).

657

658 **ETS2 co-expression**

659 Genes co-expressed with ETS2 across 64 human monocyte-derived macrophage polarisation
660 conditions (normalised data from GSE47189) were identified using the rcorr function in the
661 *Hmisc* package in R.

662

663 **¹³C-glucose GC-MS**

664 *ETS2*-edited or unedited TPP macrophages were generated in triplicate for each donor and
665 on day 6, media was removed, cells washed with PBS, and new media with labelled glucose
666 was added. Labelled media: RPMI1640 Medium, no glucose (ThermoFisher); 10% FBS
667 (ThermoFisher); GlutaMAX (ThermoFisher); ¹³C-labelled glucose (Cambridge Isotope
668 Laboratories). After 24 hours – a time point selected from a time-course to establish steady-
669 state conditions – supernatants were snap-frozen and macrophages detached by scraping.
670 Macrophages were washed three times with ice-cold PBS, counted, re-suspended in 600µl
671 ice-cold chloroform:methanol (2:1, v/v) and sonicated in a waterbath (3 x 8 mins). All
672 extraction steps were performed at 4°C as previously described⁶⁴. Samples were analysed in
673 an Agilent 7890B-7000C GC–MS system. Spitless injection (injection temperature 270°C)
674 onto a DB-5MS (Agilent) was used, using helium as the carrier gas, in electron ionization
675 mode. The initial oven temperature was 70°C (2 min), followed by temperature gradients to
676 295 °C at 12.5 °C per min and to 320 °C at 25 °C per min (held for 3 min). Scan range was m/z
677 50-550. Data analysis was performed using in-house software MANIC (version 3.0), based on
678 the software package GAVIN⁶⁵. Label incorporation was calculated by subtracting the natural
679 abundance of stable isotopes from the observed amounts. Total metabolite abundance was
680 normalised to the internal standard (scyllo-inositol⁶⁴).

681

682 **Roxadustat**

683 *ETS2*-edited or unedited TPP macrophages were generated as described previously. On day
684 5 of culture, cells were detached (Accutase) and re-plated at a density of 10⁵ cells/well in 96-
685 well round bottom plates in TPP media containing Roxadustat (FG-4592, 30µM). After 12
686 hours, cells were harvested for functional assays and RNA-seq as described.

687

688 **CUT&RUN**

689 Pre-cultured TPP macrophages were harvested and processed immediately using the
690 CUT&RUN Assay kit (Cell Signaling) according to the manufacturer's instructions but omitting
691 the use of ConA-coated beads. In brief, 5×10^5 cells per reaction were pelleted, washed, and
692 resuspended in Antibody Binding buffer. Cells were incubated with antibodies: anti-ETS2
693 (1:100, ThermoFisher) or IgG control (1:20, Cell Signaling) for 2h at 4°C. After washing in
694 Digitonin Buffer, cells were incubated with pA/G-MNase for 1h at 4°C. Cells were washed
695 twice in Digitonin Buffer, resuspended in the same buffer and cooled for 5 minutes on ice.
696 Calcium chloride was added to activate pA/G-MNase digestion (30 min, 4°C) before the
697 reaction was stopped and cells incubated at 37°C for 10 min to release cleaved chromatin
698 fragments. Supernatants were collected by centrifugation and DNA extracted using spin
699 columns (Cell Signaling). Library preparation was performed using a protocols.IO protocol
700 ([dx.doi.org/10.17504/protocols.io.bagaibse](https://doi.org/10.17504/protocols.io.bagaibse)) with the NEBNext Ultra II DNA Library Prep Kit.
701 Size selection was performed using AMPure XP beads (Beckman Coulter) and fragment sizes
702 assessed using an Agilent 2100 Bioanalyzer (High Sensitivity DNA kit). Equimolar pools of
703 indexed libraries were sequenced on a NovaSeq6000 (100bp PE reads). Raw data were
704 analysed using guidelines from the Henikoff lab⁶⁶. Briefly, paired-end reads were trimmed
705 using Trim Galore and aligned to the human genome (GRCh37/hg19) using Bowtie2. BAM
706 files were sorted, merged (technical and, where indicated, biological replicates), re-sorted and
707 indexed using SAMtools. Picard was used to mark unmapped reads and SAMtools to remove
708 these reads, re-sort and re-index. Bigwig files were created using the deepTools
709 bamCoverage function. Processed data were initially analysed using the nf-core CUT&RUN
710 pipeline v3.0, using CPM normalisation and default MACS2 parameters for peak calling. This
711 analysis yielded acceptable quality metrics (including an average FRiP score of 0.23) but
712 there were a high number of peaks with low fold enrichment (<4) over control. We therefore
713 applied more stringent parameters for peak calling (--qvalue 0.05 -f BAMPE --keep-dup all -B -
714 -nomodel) and applied an irreproducible discovery rate (IDR; cut-off 0.001) to identify

715 consistent peaks between replicates – implemented in the *idr* package in R (code:
716 https://github.com/JamesLeeLab/chr21q22_manuscript/CUT&RUN/CUTRUN_pipeline.sh).
717 Enrichment of an ETS2 binding motif in consensus IDR peaks was calculated using
718 TFmotifView⁶⁷ using global genomic controls. Overlap between consensus IDR peaks and the
719 core promoter (-250bp to +35bp from TSS) and/or putative cis-regulatory elements of ETS2-
720 regulated genes was assessed using lists of differentially-expressed genes following *ETS2*
721 disruption with gRNA1 or *ETS2* overexpression (based on a consensus across mRNA doses,
722 as described earlier). Putative cis-regulatory elements were defined as shared interactions
723 (CHiCAGO score > 5) in monocyte, M0 and M1 macrophage samples from publicly-available
724 promoter-capture Hi-C data¹⁵.

725

726 **ATAC-seq**

727 ATAC-seq in TPP macrophages was performed using the Omni-ATAC protocol⁶⁸ with the
728 following modifications: cell number was increased to 75,000 cells, cell lysis time was
729 increased to 5 minutes; volume of Tn5 transposase in the transposition mixture was doubled;
730 duration of the transposition step was extended to 40 minutes. Amplified libraries were
731 cleaned using AMPure XP beads (Beckman Coulter) and sequenced on a NovaSeq6000
732 (100bp PE reads). Data were processed as described previously⁶⁹.

733

734 **Chr21q22 disease datasets**

735 Publicly-available raw RNA-seq data from the affected tissues of chr21q22-associated
736 diseases (and controls from the same experiment) were downloaded from GEO: IBD
737 macrophages (GSE123141), primary sclerosing cholangitis liver (GSE159676), ankylosing
738 spondylitis synovium (GSE41038). Reads were trimmed, filtered, and aligned as described
739 earlier. For each disease dataset, a ranked list of genes was obtained by differential
740 expression analysis between cases and controls using *limma* with voom transformation. For
741 IBD macrophages, only IBD samples with active disease were included. fGSEA using ETS2-
742 regulated gene lists was performed as described.

743

744 **LINCS signatures**

745 31,027 lists of down-regulated genes following exposure of a cell line to a small molecule
746 were obtained from the NIH LINCS database⁸ (downloaded in January 2021). These were
747 used as gene sets for fGSEA (as described) with a ranked list of genes obtained by differential
748 expression analysis between *ETS2*-edited and unedited TPP macrophages (gRNA1) – using
749 *limma* with voom transformation and donor as a covariate. Drug classes for gene sets with
750 FDR $P < 0.05$ were manually assigned based on known mechanisms of action.

751

752 **PD-0325901**

753 TPP macrophages were generated as described previously. On day 4 of culture, PD-0325901
754 (0.5 μ M, Sigma) or vehicle (DMSO) were added. Cells were harvested on day 6 and RNA was
755 extracted and sequenced as described.

756

757 **Colonic biopsies**

758 During colonoscopy, intestinal mucosal biopsies (6 per donor) were collected from 10 IBD
759 patients (7 ulcerative colitis, 3 Crohn's disease). All had endoscopically active disease and
760 were not receiving immunosuppressive or biologic therapies. All biopsies were collected from
761 a single inflamed site. All patients provided written informed consent. Ethical approval was
762 provided by the London - Brent Regional Ethics Committee (REC: 21/LO/0682). Biopsies were
763 collected into Opti-MEM and within 1 hour were weighed and placed in pairs onto a transwell
764 insert (ThermoFisher) – designed to create an air-liquid interface⁷⁰ – in a 24-well plate. Each
765 well contained 1ml media and was supplemented with either DMSO (vehicle control), PD-
766 0325901 (0.5 μ M) or infliximab (10 μ g/ml; MSD). Media: Opti-MEM I (Gibco); GlutaMAX
767 (ThermoFisher); 10% FBS (ThermoFisher); MEM Non-Essential Amino Acids (ThermoFisher);
768 1% sodium pyruvate (ThermoFisher); 1% penicillin/streptomycin (ThermoFisher); 50 μ g/ml
769 gentamicin (Merck). After 18 hours, supernatants and biopsies were snap frozen. Supernatant
770 cytokine concentrations were quantified using LEGENDplex Human Inflammation Panel

771 (Biolegend). RNA was extracted from biopsies and libraries prepared as described earlier
772 (n=9, RNA from one donor was too degraded). Sequencing was performed on a NovaSeq
773 6000 (100bp, PE reads). Data were processed as described earlier and GSVA was performed
774 for ETS2-regulated genes and biopsy-derived signatures of IBD-associated inflammation⁴³.

775

776 **Chr21q22 genotypes in archaic humans**

777 Using publicly available genomes from seven Neanderthal individuals⁷¹⁻⁷⁴, one Denisovan
778 individual⁷⁵, and one Neanderthal and Denisovan F1 individual⁷⁶, we called genotypes at the
779 disease-associated chr21q22 candidate SNPs from the respective BAM files using “bcftools
780 mpileup” with base and mapping quality options -q 20 -Q 20 -C 50 and using “bcftools call -m -
781 C alleles”, specifying the two alleles expected at each site in a targets file (-T option). From the
782 resulting vcf file, we extracted the number of reads supporting the reference and alternative
783 alleles stored in the “DP4” field.

784

785 **Inference of Relate genealogy at rs2836882**

786 We used genome-wide genealogies previously inferred for samples of the Simons Genome
787 Diversity Project⁷⁷ dataset (<https://reichdata.hms.harvard.edu/pub/datasets/sgdp/>) using
788 Relate^{78,79}. These genealogies were downloaded from
789 <https://www.dropbox.com/sh/2gjyxe3kqzh932o/AAAQcipCHnySqEB873t9EQjNa?dl=0>. Using
790 the inferred genealogies, the genealogy at rs2836882 (chr21:40466570) was plotted using the
791 TreeView module of Relate.

792

793 **Statistical methodology**

794 Statistical methods used in MPRA analysis, fGSEA, and SNPsea are described above. For
795 other analyses, comparison of continuous variables between paired samples in two groups
796 was performed using a Wilcoxon matched-pairs test for non-parametric data or a paired t-test
797 for parametric data. Comparison against a hypothetical value was performed using a Wilcoxon
798 signed-rank test for non-parametric data or one sample t-test for parametric data. A Shapiro-

799 Wilk test was used to confirm normality. Two-tailed tests were used as standard unless a
800 specific hypothesis was being tested. Sample sizes are provided in respective sections.

801

802 **Code availability**

803 Code to reproduce analyses are available at

804 https://github.com/JamesLeeLab/chr21q22_manuscript and

805 <https://github.com/chr1swallace/ibd-ets2-analysis>.

806

807 **Data availability**

808 The datasets produced in this study, including raw and processed files, have been uploaded
809 to the following databases and will be made publicly available in advance of publication.

810 Gene Expression Omnibus: MPRA (GSE229472), RNA-seq of *ETS2* or chr21q22-edited TPP

811 macrophages (GSE229569), RNA-seq of *ETS2* overexpression (GSE229744), RNA-seq of

812 MEK inhibitor-treated TPP macrophages (GSE229743), H3K27ac ChIP-seq in TPP

813 macrophages (GSE229464), ATAC-seq in TPP macrophages (GSE229624), *ETS2*

814 CUT&RUN (GSE229745), biopsy RNA-seq data (GSE230020).

815 MetaboLights: Metabolomics (MTBLS7665).

816 **Methods references**

- 817 51 Zou, Y., Carbonetto, P., Wang, G. & Stephens, M. Fine-mapping from summary data
818 with the "Sum of Single Effects" model. *PLoS Genet* **18**, e1010299,
819 doi:10.1371/journal.pgen.1010299 (2022).
- 820 52 Wallace, C. A more accurate method for colocalisation analysis allowing for multiple
821 causal variants. *PLoS Genet* **17**, e1009440, doi:10.1371/journal.pgen.1009440 (2021).
- 822 53 Bourges, C. *et al.* Resolving mechanisms of immune-mediated disease in primary CD4
823 T cells. *EMBO Mol Med* **12**, e12112, doi:10.15252/emmm.202012112 (2020).
- 824 54 Hanzelmann, S., Castelo, R. & Guinney, J. GSEA: gene set variation analysis for
825 microarray and RNA-seq data. *BMC Bioinformatics* **14**, 7, doi:10.1186/1471-2105-14-7
826 (2013).
- 827 55 Peters, J. E. *et al.* Insight into Genotype-Phenotype Associations through eQTL
828 Mapping in Multiple Cell Types in Health and Immune-Mediated Disease. *PLoS Genet*
829 **12**, e1005908, doi:10.1371/journal.pgen.1005908 (2016).
- 830 56 Conant, D. *et al.* Inference of CRISPR Edits from Sanger Trace Data. *CRISPR J* **5**,
831 123-130, doi:10.1089/crispr.2021.0113 (2022).
- 832 57 Kalita, C. A. *et al.* QuASAR-MPRA: accurate allele-specific analysis for massively
833 parallel reporter assays. *Bioinformatics* **34**, 787-794, doi:10.1093/bioinformatics/btx598
834 (2018).
- 835 58 Randzavola, L. O. *et al.* EROS is a selective chaperone regulating the phagocyte
836 NADPH oxidase and purinergic signalling. *Elife* **11**, doi:10.7554/eLife.76387 (2022).
- 837 59 Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome
838 alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* **37**, 907-
839 915, doi:10.1038/s41587-019-0201-4 (2019).
- 840 60 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**,
841 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).
- 842 61 Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program
843 for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930,
844 doi:10.1093/bioinformatics/btt656 (2014).
- 845 62 Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing
846 and microarray studies. *Nucleic Acids Res* **43**, e47, doi:10.1093/nar/gkv007 (2015).
- 847 63 Korotkevich, G. *et al.* Fast gene set enrichment analysis. *bioRxiv*, 060012,
848 doi:10.1101/060012 (2021).
- 849 64 Bussi, C. *et al.* Lysosomal damage drives mitochondrial proteome remodelling and
850 reprograms macrophage immunometabolism. *Nat Commun* **13**, 7338,
851 doi:10.1038/s41467-022-34632-8 (2022).
- 852 65 Behrends, V., Tredwell, G. D. & Bundy, J. G. A software complement to AMDIS for
853 processing GC-MS metabolomic data. *Anal Biochem* **415**, 206-208,
854 doi:10.1016/j.ab.2011.04.009 (2011).
- 855 66 Meers, M. P., Bryson, T. D., Henikoff, J. G. & Henikoff, S. Improved CUT&RUN
856 chromatin profiling tools. *Elife* **8**, doi:10.7554/eLife.46314 (2019).
- 857 67 Leporcq, C. *et al.* TFmotifView: a webserver for the visualization of transcription factor
858 motifs in genomic regions. *Nucleic Acids Res* **48**, W208-W217,
859 doi:10.1093/nar/gkaa252 (2020).

- 860 68 Corces, M. R. *et al.* An improved ATAC-seq protocol reduces background and enables
861 interrogation of frozen tissues. *Nat Methods* **14**, 959-962, doi:10.1038/nmeth.4396
862 (2017).
- 863 69 Calderon, D. *et al.* Landscape of stimulation-responsive chromatin across diverse
864 human immune cells. *Nat Genet* **51**, 1494-1505, doi:10.1038/s41588-019-0505-9
865 (2019).
- 866 70 Vadstrup, K. *et al.* Validation and Optimization of an Ex Vivo Assay of Intestinal
867 Mucosal Biopsies in Crohn's Disease: Reflects Inflammation and Drug Effects. *PLoS*
868 *One* **11**, e0155335, doi:10.1371/journal.pone.0155335 (2016).
- 869 71 Prufer, K. *et al.* The complete genome sequence of a Neanderthal from the Altai
870 Mountains. *Nature* **505**, 43-49, doi:10.1038/nature12886 (2014).
- 871 72 Prufer, K. *et al.* A high-coverage Neandertal genome from Vindija Cave in Croatia.
872 *Science* **358**, 655-658, doi:10.1126/science.aao1887 (2017).
- 873 73 Hajdinjak, M. *et al.* Reconstructing the genetic history of late Neanderthals. *Nature*
874 **555**, 652-656, doi:10.1038/nature26151 (2018).
- 875 74 Mafessoni, F. *et al.* A high-coverage Neandertal genome from Chagyrskaya Cave.
876 *Proc Natl Acad Sci U S A* **117**, 15132-15136, doi:10.1073/pnas.2004944117 (2020).
- 877 75 Meyer, M. *et al.* A high-coverage genome sequence from an archaic Denisovan
878 individual. *Science* **338**, 222-226, doi:10.1126/science.1224344 (2012).
- 879 76 Slon, V. *et al.* The genome of the offspring of a Neanderthal mother and a Denisovan
880 father. *Nature* **561**, 113-116, doi:10.1038/s41586-018-0455-x (2018).
- 881 77 Mallick, S. *et al.* The Simons Genome Diversity Project: 300 genomes from 142
882 diverse populations. *Nature* **538**, 201-206, doi:10.1038/nature18964 (2016).
- 883 78 Speidel, L., Forest, M., Shi, S. & Myers, S. R. A method for genome-wide genealogy
884 estimation for thousands of samples. *Nat Genet* **51**, 1321-1329, doi:10.1038/s41588-
885 019-0484-x (2019).
- 886 79 Speidel, L. *et al.* Inferring Population Histories for Ancient Genomes Using Genome-
887 Wide Genealogies. *Mol Biol Evol* **38**, 3497-3511, doi:10.1093/molbev/msab174 (2021).

888 **Acknowledgements**

889 We thank members of the Lee lab and Arthur Kaser for helpful discussions and Gitta
890 Stockinger, Carola Vinuesa, Charlie Swanton, Rickie Patani, and Caetano Reis e Sousa for
891 critical reading of the manuscript. We thank Chris Cheshire, the Francis Crick Institute
892 Advanced Sequencing Facility and Flow Cytometry STP for technical support, Laura Lucaciu
893 for help with patient recruitment, and Sarah Edwards for providing infliximab. We thank NIHR
894 BioResource volunteers for their participation, and acknowledge NIHR BioResource centres,
895 NHS Blood and Transplant, and NHS staff for their contribution. This work was supported by
896 Crohn's and Colitis UK (M2018-3), the Wellcome Trust (Sir Henry Wellcome Fellowship to
897 L.S.: 220457/Z/20/Z, Investigator Award to P.S.: 217223/Z/19/Z, Senior Fellowship to C.W.:
898 WT220788, Clinical Research Career Development Fellowship to M.Z.C.: 222056/Z/20/Z,
899 Wellcome-Beit Prize Clinical Career Development Fellowship to D.C.T.: 206617/A/17/A, and
900 Intermediate Clinical Fellowship to J.C.L.: 105920/Z/14/Z), and the Francis Crick Institute,
901 which receives its core funding from Cancer Research UK (CC2219, FC001595), the UK
902 Medical Research Council (CC2219, FC001595), and the Wellcome Trust (CC2219,
903 FC001595). P.S. is additionally supported by the European Molecular Biology Organisation,
904 the Vallee Foundation, and the European Research Council (852558). C.W. is additionally
905 supported by the Medical Research Council (MC UU 00002/4), GSK, MSD, and the NIHR
906 Cambridge BRC (BRC-1215-20014). D.C.T. is additionally supported by the Sidharth Burman
907 endowment, and J.C.L. is a Lister Institute Prize Fellow. The funders had no role in study
908 design, data collection and analysis, decision to publish, or preparation of the manuscript.
909 Experimental schematics in Figs. 1b, 2a, 3a, 5e created using BioRender. For the purpose of
910 Open Access, the authors have applied a CC BY public copyright licence to any Author
911 Accepted Manuscript version arising from this submission.

912

913 **Author contributions**

914 Conceptualisation, J.I.M., P.S., M.Z.C., C.W., D.C.T., J.C.L.; Methodology, C.T.S., C.B.,
915 M.S.D.S., L.O.R., L.S., J.I.M., C.W., J.C.L.; Software, C.B., M.S.D.S., L.S., J.I.M., C.W.;

916 Investigation, C.T.S., C.B., T.T.S., A.P.P., C.P.J., I.P., M.S.D.S., L.O.R., L.S., E.C.P., W.E.,
917 A.P.R., C.D.M., C.W, J.C.L.; Resources, C.T.S., C.B., M.S.D.S., J.C.L.; Formal analysis,
918 C.T.S., C.B., M.S.D.S., L.S., C.W., J.C.L.; Writing – Original Draft, C.T.S., C.B., J.C.L.; Writing
919 – Review & Editing, all authors; Funding Acquisition, J.C.L.; Supervision, J.I.M., P.S., C.W.,
920 D.C.T., J.C.L.

921

922 **Competing Interests**

923 C.T.S., C.B. and J.C.L. are co-inventors on a patent application relating to this work. C.W.
924 holds a part time position at GSK. GSK had no role in this study.

925

926 **Additional information**

927 **Supplementary information:** results of differential expression analysis in *ETS2*-edited or
928 *ETS2* overexpression experiments are available in Supplementary Tables 1 and 2.

929 **Correspondence and requests for materials** should be addressed to James C. Lee.