

1 **Title: Joint analysis reveals shared autoimmune disease associations**
2 **and identifies common mechanisms**

3

4 **Short title**

5 Joint analysis of autoimmune diseases

6

7 **One sentence summary:**

8 Genetic mapping in autoimmune diseases increases genetic mapping resolution and reveals
9 shared mechanisms

10

11 **Authors:**

12 Matthew R Lincoln¹, Noah Connally¹, Pierre-Paul Axisa¹, Christiane Gasperi¹, Mitja Mitrovic¹,
13 David van Heel², Cisca Wijmenga³, Sebo Withoff³, Iris H Jonkers³, Leonid Padyukov⁴,
14 International Multiple Sclerosis Genetics Consortium, Stephen S Rich^{5,6}, Robert R Graham^{7,8},
15 Patrick M Gaffney⁹, Carl D Langefeld^{10,11}, David A Hafler¹, Sung Chun^{12,13}, Shamil R Sunayev^{12,13},
16 Chris Cotsapas^{1,14}

17

18 **Affiliations:**

19 ¹Department of Neurology, Yale School of Medicine, New Haven, CT

20 ²Blizard Institute, Queen Mary University of London, London, UK

21 ³Department of Genetics, University Medical Center Groningen, University of Groningen,
22 Groningen, The Netherlands

23 ⁴Division of Rheumatology, Department of Medicine, Karolinska Institutet and Karolinska
24 University Hospital, Stockholm, Sweden

25 ⁵Center for Public Health Genomics, University of Virginia, Charlottesville, VA

26 ⁶Department of Public Health Sciences, University of Virginia, Charlottesville, VA

27 ⁷Maze Therapeutics, South San Francisco, CA

28 ⁸Genentech, South San Francisco, CA

29 ⁹Genes and Human Disease Research Program, Oklahoma Medical Research Foundation,
30 Oklahoma City, OK

31 ¹⁰Department of Biostatistics and Data Science, Wake Forest School of Medicine, Winston-
32 Salem, NC

33 ¹¹Center for Precision Medicine, Wake Forest School of Medicine, Winston-Salem, NC

34 ¹²Division of Genetics, Brigham and Women's Hospital, Boston, MA

35 ¹³Department of Biomedical Informatics, Harvard Medical School, Boston, MA

36 ¹⁴Department of Genetics, Yale School of Medicine, New Haven, CT

37 **Abstract:**

38

39 Autoimmune and inflammatory diseases are polygenic disorders of the immune system. Many
40 genomic loci harbor risk alleles for several diseases, but the limited resolution of genetic mapping
41 prevents determining if the same allele is responsible, indicating a shared underlying mechanism.
42 Using a collection of 129,058 cases and controls across six diseases, we show that ~40% of
43 overlapping associations are due to the same allele. We improve fine-mapping resolution for
44 shared alleles two-fold by combining cases and controls across diseases, allowing us to identify
45 more eQTLs driven by the shared alleles. The patterns of sharing indicate widespread shared
46 mechanisms, but not a single global autoimmune mechanism. Our approach can be applied to
47 any set of traits, and is particularly valuable as sample collections become depleted.

48 **Main Text:**

49
50 Autoimmune and inflammatory diseases are a heterogeneous group of disorders, where
51 activation of both the adaptive and innate immune system coupled with loss of self-tolerance
52 leads to target tissue destruction (1). These diseases are heritable, and genome-wide association
53 studies (GWAS) have identified hundreds of susceptibility loci, confirming their polygenic nature
54 (2, 3). Like other complex disease risk traits, heritability is strongly enriched in gene regulatory
55 regions active in specific cell populations (4–6), suggesting risk is mediated to a large extent by
56 altering gene expression in specific cell types under specific conditions. These diseases are also
57 comorbid (7, 8), with dual diagnoses being more frequent in individuals than expected by chance,
58 and multiple diseases aggregating in families (9). We and others have shown that many genetic
59 loci harbor risk variants for multiple autoimmune diseases (10–12), suggesting that comorbidity
60 may be due to shared genetic liability and, hence, shared mechanisms of disease.

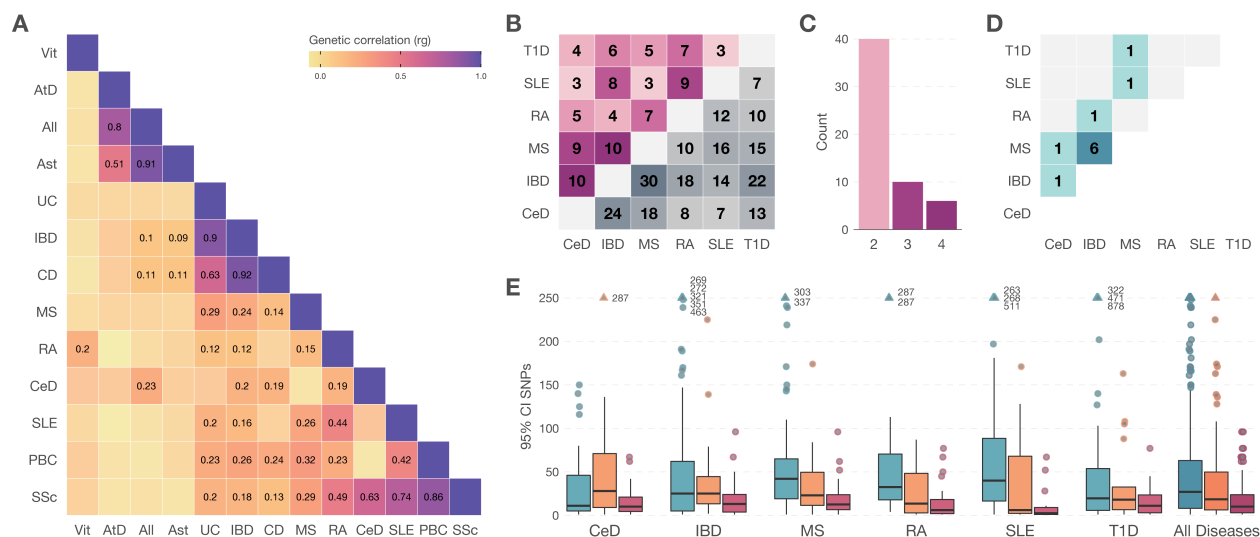
61
62 Instances of pleiotropy, where the same variant influences risk to more than one disease, would
63 by definition point to a shared molecular effect, and thus a shared mechanism. The limited
64 resolution of genetic mapping has made it difficult to distinguish such cases from situations
65 where distinct genetic variants in the same locus mediate risk to different diseases. This limited
66 resolution restricts our ability to uncover shared pathogenic mechanisms, understand why some
67 modulating immune functions can increase risk to one disease whilst decreasing risk to others,
68 or make inferences about the origins of these diseases and their different prevalence rates
69 around the world.

70
71 An important driver of the limited resolution of genetic mapping is disease cohort sample size
72 (13). Currently available disease cohorts, most of which have been extensively studied already,
73 are the result of decades-long international recruitment efforts. Meaningful increases in sample
74 size are thus difficult to envision in the immediate future. An alternative way to increase sample
75 size, and thus genetic mapping resolution, would be to jointly analyze cohorts across diseases.
76 This would also reveal shared pathogenic mechanisms. In conventional meta-analyses of
77 cohorts with the same disease, we assume that any associations are shared across strata; we
78 cannot make this assumption across diseases. It is thus crucial to ensure that the same allele
79 drives risk to two or more diseases, rather than separate alleles in the same genomic locus.

80
81 Here, we first show substantial genome-wide shared heritability between autoimmune and
82 inflammatory diseases (Fig. S1). We then look at 224 instances where genetic associations to
83 multiple diseases occur in the same genomic region, and show that 41.5% of these observed
84 associations are due to pleiotropic variants, with the remainder being due to different alleles in
85 the region. When we combine cases and controls across diseases to map each shared
86 association, we increase fine-mapping resolution two-fold on average. Further, this increase in
87 resolution reveals new disease risk variants that alter gene expression in immune cell subtypes.
88 Comorbidity is widespread between diseases of all organ systems, and sample sizes are limited,
89 so this strategy is widely applicable beyond the immune-mediated diseases. Thus, this approach
90 to careful dissection of shared effects can reveal mechanisms that are common across diseases
91 and pinpoint key genes driving shared biology.

92
93 We first assessed the evidence for genome-wide shared heritability between 19 autoimmune
94 and inflammatory diseases from GWAS summary data. After quality control, we used LD score
95 regression (14) to estimate heritability (h_g^2) for each trait (Fig. S2A). We found that 13/19 diseases
96 had sufficient heritability captured by common variants to make these comparisons meaningful

97 (Z -score > 4) (15), so we restricted our analysis to this subset. We then calculated the proportion
 98 of shared heritability between each pair of diseases, again using LD score regression, which is
 99 robust to sample overlaps between cohorts (15). We found a broad pattern of shared heritability
 100 (Fig. 1), with the strongest overlaps ($0.63 \leq r_g \leq 0.92$) between IBD and its subtypes, Crohn
 101 disease and ulcerative colitis; these are known to share the majority, but not the entirety, of their
 102 genetic architecture (16). We also observed strong correlations between atopic dermatitis,
 103 asthma and allergic traits ($0.51 \leq r_g \leq 0.91$), which may represent a shared basis for atopic
 104 inflammatory disease. We saw a strong correlation between systemic sclerosis and systemic
 105 lupus erythematosus, which were also correlated with primary biliary cirrhosis ($0.42 \leq r_g \leq 0.86$).
 106 In line with our previous findings (10, 17), these results indicate that autoimmune and
 107 inflammatory diseases share a substantial portion of genetic risk factors, even when accounting
 108 for the major histocompatibility locus (MHC), where overlapping haplotypes confer risk to
 109 different autoimmune and inflammatory diseases (18). Overall, this suggests that some
 110 mechanisms are common between sets of diseases, but we find no evidence of universal sharing
 111 indicative of a large core autoimmune susceptibility component (10, 17).
 112



113 **Fig. 1: Joint analysis of shared autoimmune disease risk alleles improves fine-mapping two-fold.** (A)
 114 We find broad genome-wide correlation between association statistics for susceptibility to thirteen
 115 autoimmune and inflammatory diseases for which genome-wide association data were available (vitiligo,
 116 Vit; atopic dermatitis, AtD; allergy, All; asthma, Ast; ulcerative colitis, UC; inflammatory bowel disease,
 117 IBD; Crohn disease, CD; multiple sclerosis, MS; rheumatoid arthritis, RA; celiac disease, CeD; systemic
 118 lupus erythematosus, SLE; primary biliary cirrhosis, PBC; and systemic sclerosis, SSc). (B) This correlation
 119 is reflected in many loci harboring risk alleles to more than one of six diseases with available ImmunoChip
 120 data (lower triangle). Of these 236 pairs of associations, 99 are driven by the same underlying allele (upper
 121 triangle). (C) Risk alleles are mostly shared between two diseases (38 cases), with thirteen shared between
 122 three, and six between four diseases. (D) Twelve shared alleles have opposite effect directions, increasing
 123 risk of one disease and decreasing risk of another. This is most frequent between MS and IBD. (E)
 124 Combining cases and controls across diseases increases fine-mapping resolution for these shared
 125 associations. We assess resolution as the number of variants required to explain 95% of the posterior
 126 probability of association. This credible interval decreases by 58% when combining samples across
 127 diseases (red) compared to using only samples for one disease (orange). Associations that are not shared
 128 across diseases have similar credible interval distributions in individual diseases (blue).
 129
 130

131 While this shared heritability gives an overall impression of the relationship between diseases, it
132 cannot identify specific genetic risk factors—and thus, genes and pathways—shared between
133 diseases. To compare samples from different collections genotyped at different centers, it was
134 important to minimize batch effects by ensuring all samples were profiled on the same platform.
135 We therefore chose six autoimmune and inflammatory diseases with large numbers of samples
136 genotyped on the ImmunoChip (19) (celiac disease, inflammatory bowel disease, multiple
137 sclerosis, rheumatoid arthritis, systemic lupus erythematosus and type 1 diabetes; Fig. S1). This
138 targeted array interrogates variants in 188 known risk loci to saturation, representing only 1.9%
139 of the genome but capturing 38–86% of risk loci that have been identified in the six diseases
140 (Fig. S3A). Using partitioned LD score regression, we confirmed that ImmunoChip regions
141 account for 27.6% (MS) to 46.3% (CeD) of the estimated heritability for five of the six diseases
142 for which GWAS data was available (Fig. S2B). After quality control, removal of population
143 outliers, resolution of duplicate and related samples, and imputation to the 1,000 Genomes
144 reference haplotypes, we analyzed a total of 104,302 SNPs in 188 non-MHC genomic regions
145 for association with disease in 82,630 cases and 104,573 controls (Fig. S1; see Supplementary
146 Text for additional details).

147
148 We first identified associations across the 188 loci in each disease independently by assembling
149 cases and controls into homogeneous population strata and meta-analyzing across these
150 groups. As multiple independent associations at a locus have been described in all diseases, we
151 used stepwise logistic regression followed by fixed-effects meta-analysis to allow for such
152 effects. We found 197 independent associations in 123 different ImmunoChip loci at genome-
153 wide significance ($P < 5 \times 10^{-8}$), and 361 associations at 166 loci with suggestive association
154 evidence ($P < 10^{-5}$, Fig. S3B). Overall, we find some level of support for essentially all known
155 genome-wide significant effects in the ImmunoChip regions.

156
157 We found substantial evidence for multiple independent associations within loci, with 7% (RA)
158 to 30% (IBD) of loci exhibiting more than one independent effect (Fig. S3C). This included three
159 instances of associations that have not been reported before (Fig. S4). In celiac disease, we
160 found suggestive unconditioned associations at two loci: a variant intronic to *ANKS1A* on
161 chromosome 6 (rs12206298; $P = 4.1 \times 10^{-7}$), and a variant intronic to *CTSH* on chromosome 15
162 (rs3784539; $P = 1.3 \times 10^{-5}$). After conditional association, both these associations passed the
163 genome-wide significance threshold ($P = 4.9 \times 10^{-8}$ and 1.1×10^{-8} respectively). We found
164 evidence of a second, independent effect in each locus (rs4713844, $P = 9.9 \times 10^{-8}$; and
165 rs7181033, $P = 8.7 \times 10^{-5}$). Similarly, in IBD, we found that a suggestive association in the
166 *CLEC16A* locus on chromosome 16 (rs7201325, $P = 1.4 \times 10^{-7}$) reached genome-wide
167 significance after conditioning ($P = 1.1 \times 10^{-10}$), with evidence of a secondary, independent effect
168 (rs55773334, $P = 7.6 \times 10^{-5}$). The presence of multiple masked independent effects highlights
169 the need to look carefully at suggestive associations.

170
171 Having ensured we were capturing most of the known associations in ImmunoChip loci for each
172 of the six diseases, we looked for shared effects across diseases, i.e. whether the same variant
173 mediates risk to more than one disease. We found 224 overlapping conditionally independent
174 associations at 98 loci (a lead variant associated to one disease at $P < 10^{-5}$, and a lead variant
175 for another disease $P < 10^{-4}$; both lead variants being in LD $r^2 > 0.5$ with at least one common
176 SNP). Using joint likelihood mapping (JLIM), we found evidence of a shared effect in 93/224
177 (41.5%) such overlaps, involving a total of 56 conditionally independent shared effects spanning
178 53 unique loci (Figure 1B). Of these, 40 effects were shared between two diseases, ten between
179 three, and six between four diseases (Fig. 1C). Unlike previous reports, which could not

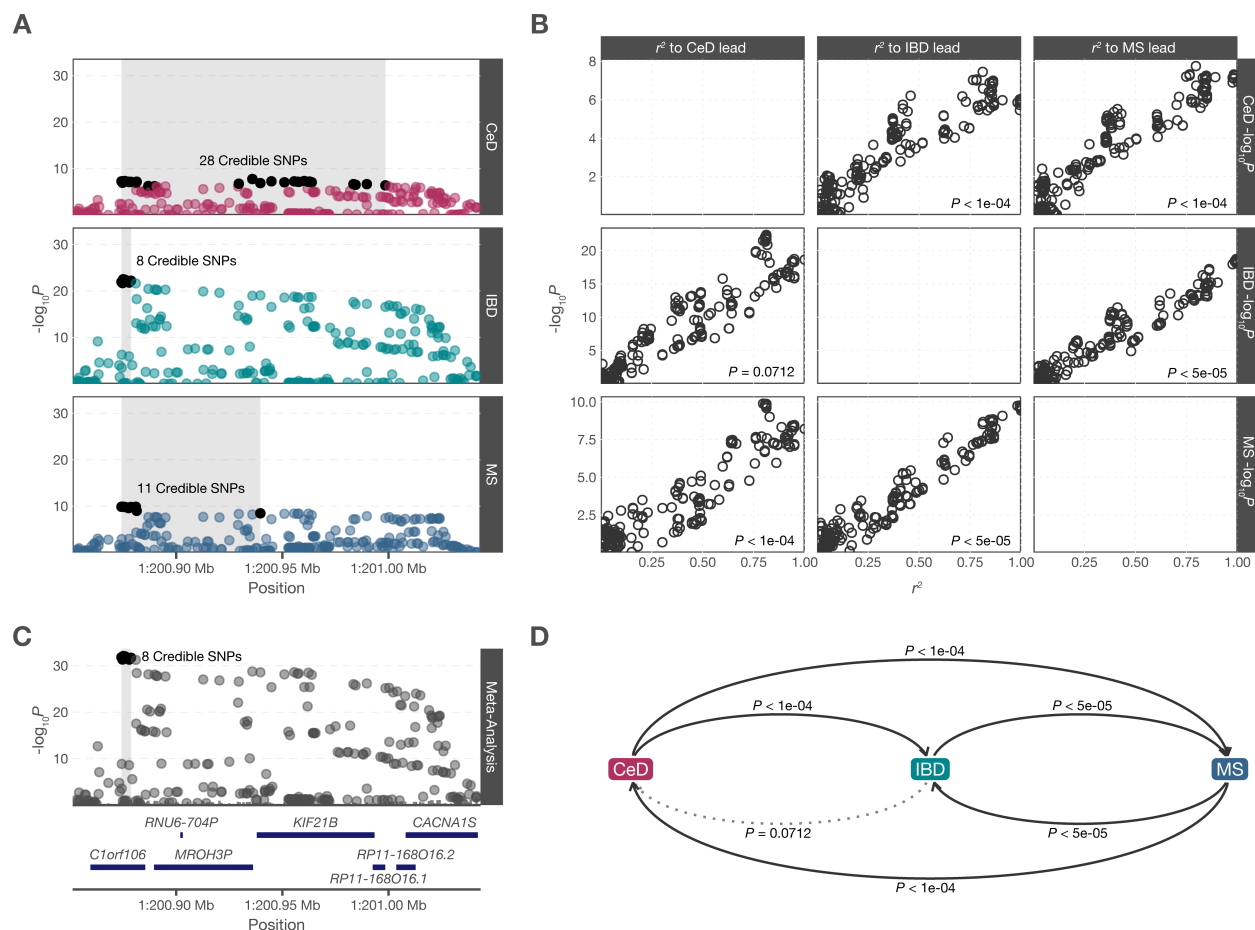
180 distinguish between shared and distinct associations with multiple diseases in a locus, these
181 observations indicate that many mechanisms are shared between autoimmune and inflammatory
182 diseases.

183
184 We found three loci where multiple conditionally independent associations for one disease were
185 shared. In the *STAT4* locus, we found two independent effects each for RA and SLE were shared
186 (Fig. S5A). In the *CD28-CTLA4* locus, one T1D risk association near *CD28* is shared with CeD,
187 whereas another, an intronic variant in *CTLA4* is shared with RA (Fig. S5B). In the *TYK2* locus,
188 one RA risk association is shared with SLE, IBD, and T1D; a second association, localizing to
189 *ICAM3*, is shared with SLE alone (Fig. S5C). Cumulatively, these examples demonstrate that
190 disease-associated alleles in the same locus can have different consequences, and that careful
191 comparisons across diseases can distinguish each effect.

192
193 We next assessed if joint analysis across diseases could improve fine-mapping resolution. For
194 each of the 56 shared associations, we assembled conditionally independent association data
195 across all disease cohorts sharing that association, and combined them with fixed-effects,
196 inverse variance-weighted meta-analysis. In a subset of loci, we saw an unexpected decrease
197 in significance and increase in heterogeneity in the meta-analysis; we found these to be shared
198 associations with opposite effects, where an allele increases risk for one disease and decreases
199 it for the other (Fig. S6-S15). In six of these ten cases, variants with opposing effects were shared
200 between MS and IBD. After inverting the association statistics to account for these effects, our
201 meta-analysis resulted in higher significance for 123/134 (91.8%) associations across all 53 loci
202 harboring a shared effect, demonstrating the potential to bolster association findings with our
203 approach.

204
205 To establish if this increase in sample size provides a meaningful increase in fine-mapping
206 resolution, we used FINEMAP (20) to calculate posterior inclusion probabilities for SNPs at each
207 of the 56 shared effects, both in individual diseases and in the cross-disease meta-analysis. We
208 then calculated 95% credible sets for each disease, both before and after cross-disease meta-
209 analysis. We found a substantial decrease in the mean credible interval size, from 37.2 (s.d. 46.4)
210 to 17.3 (s.d. 20.8), representing an improvement of 54% (Fig. 1E). We saw resolution
211 improvement across the spectrum of initial association evidence, with the largest gains where
212 an effect had relatively weak evidence of association in a disease: for associations below
213 genome-wide significance in a single disease, our resolution increased from a mean of 51.8
214 SNPs to 18.6 SNPs after cross-disease meta-analysis; for associations already above genome-
215 wide significance in a single disease, we saw improvement from a mean of 21.7 SNPs to 15.9
216 SNPs. This is exemplified by a shared association in the *C1orf106* locus on chromosome 1,
217 where credible intervals of 28, 8, and 11 SNPs for CeD, IBD and MS respectively are reduced to
218 eight variants in very tight linkage disequilibrium (minimum $r^2 = 0.976$) on cross-disease meta-
219 analysis (Fig. 2). In this case, there are genome-wide significant associations in each disease
220 independently, but increasing sample size from symmetric equivalent 19,026 (CeD), 53,312 (IBD),
221 35,618 (MS) to a cross-disease meta-analysis 93,001 (symmetric equivalent) increases the
222 resolution for both CeD and MS, identifying a core risk haplotype within *C1orf106*.

223



224
225
226
227
228
229
230
231
232
233
234

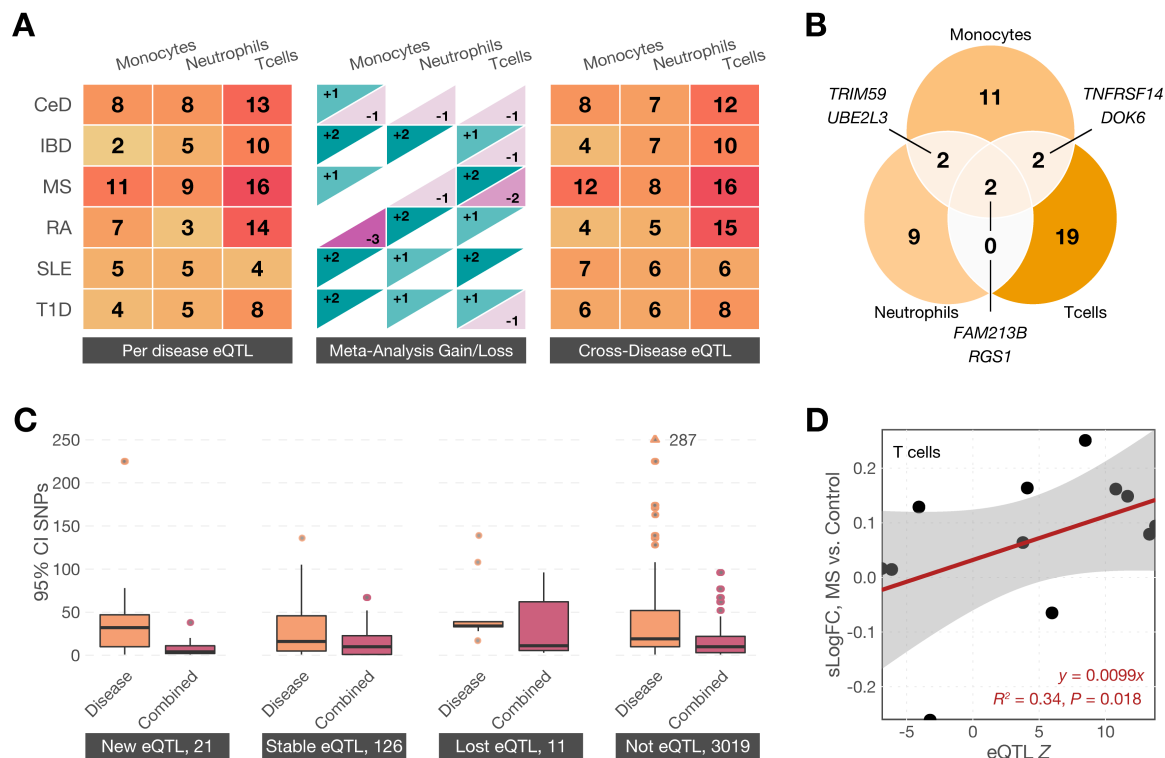
Fig. 2: A shared effect on chromosome 1 can be fine-mapped to eight variants across celiac disease, inflammatory bowel disease, and multiple sclerosis. (A) Overlapping associations in celiac disease, inflammatory bowel disease and multiple sclerosis on chromosome 1, with 95% credible intervals varying both in number of variants and physical span. (B) For each pair of diseases, the strength of association (vertical axis) for the first trait decays in a linear fashion as a function of r^2 to the lead SNP in the second trait, consistent with a shared causal variant. (C) Meta-analyzing across the three diseases gives in a stronger association signal, which can be fine-mapped to a narrow interval within *C1orf106*. (D) We find strong pairwise evidence that the association is shared between all three diseases; JLIM is asymmetric, so we run comparisons in both directions.

235 The ultimate promise of increasing fine-mapping resolution is to increase the interpretability of
236 association signals. We and others have shown that disease risk associations are enriched in
237 non-coding regions with gene regulatory potential (4, 5, 21, 22). We have used the JLIM
238 approach to show that autoimmune disease associations are sometimes shared with expression
239 quantitative trait locus (eQTL) signals, indicating the risk allele also influences gene expression.
240 However, most associations are not shared with an eQTL, nor are they attributable to coding
241 variants. To assess if this is due to limitations in fine-mapping resolution, we looked for shared
242 associations between the 56 shared effects we discovered and *cis*-eQTLs for nearby genes in
243 naïve T cells, monocytes and neutrophils in the BLUEPRINT dataset. We found 137 shared
244 effects between each of 134 shared conditionally independent association signals in a single
245 disease and eQTLs for nearby genes. We then looked for shared effects between the better-
246 powered cross-disease meta-analysis data in each of the 53 loci, and can attribute 21 new
247 disease/eQTL effects to the underlying diseases (Fig. 3; Table S3). Most of the implicated eQTLs

248 are present in only one of the three cell types we interrogated, with T cells providing the largest
 249 number. We also exclude 11/137 disease/eQTL shared effects as no longer relevant because we
 250 do not find evidence of shared association between the cross-disease meta-analysis and eQTL
 251 data. Our gains primarily occur in cases where the cross-disease meta-analysis reduces the
 252 credible interval size (Fig. 3C), indicating that this gain of resolution drives these new
 253 observations.

254
 255 The direction of shared eQTL effects indicate whether we should expect increases or decreases
 256 in expression for those genes to increase disease risk. We reasoned that we might also see the
 257 same direction of effect between cases and controls, where the risk state is magnified. We
 258 therefore looked at single cell RNAseq data derived from T cells collected from a cohort of MS
 259 patients and healthy controls (23). After quality control, we were able to detect twelve genes that
 260 were targets of eQTLs shared with MS risk signals in our analysis. We found a significant pattern
 261 of correlation ($P = 0.018$): when a disease risk allele increased expression of a target gene, we
 262 saw higher expression in cases than in controls, and when it decreased expression we saw lower
 263 levels in cases than in controls (Fig. 3D). This suggests that shared associations do in fact drive
 264 risk-altering changes to gene regulation alter disease risk, and our results are uncovering
 265 pathogenic mechanisms.

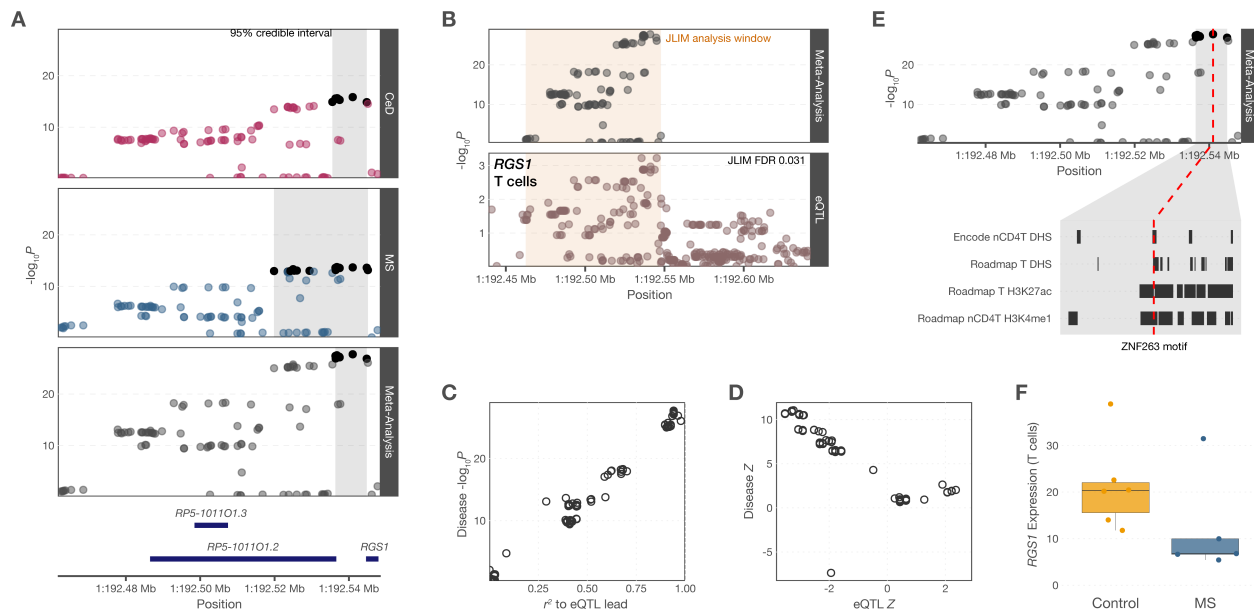
266



267
 268 **Fig. 3: The increased resolution of fine-mapping shared associations across diseases allows**
 269 **identification of more disease-eQTL overlaps.** (A) We looked for shared effects between disease
 270 associations and expression QTLs in loci harboring shared disease effects. When considering each
 271 disease separately, we find 139 significant disease—eQTL overlaps across monocytes, neutrophils and T
 272 cells from the BLUEPRINT consortium (left panel). When comparing eQTLs to cross-disease meta-
 273 analyses, we find new overlaps (blue, middle panel) and no longer find evidence for some eQTLs (red,
 274 middle panel), for a grand total of 157 disease-eQTL overlaps (13% net discovery increase, right panel).
 275 (B) Some of the shared eQTL effects can be detected in multiple tissues, but most are restricted to a single
 276 cell type, indicating substantial effect specificity. (C) We find new eQTL shared effects in loci where the

277 cross-disease meta-analysis decreases the credible interval substantially, suggesting this resolution drives
 278 new discoveries. Disease associations where an eQTL is lost after meta-analysis also have smaller credible
 279 intervals, suggesting these may have been false positive findings due to lack of resolution in individual
 280 disease datasets. (D) The effects of risk-increasing shared alleles on gene expression is mirrored in
 281 expression differences between multiple sclerosis cases and controls. This suggests that risk states
 282 imparted due to small changes in gene expression persist during active disease, and provide validation
 283 that our eQTL discoveries are relevant to pathogenesis.

284
 285 The relative direction of the disease and eQTL associations can also suggest specific
 286 mechanistic hypotheses. This is exemplified by an association in the *RGS1* locus, shared
 287 between celiac disease and MS (Fig. 4). The cross-disease meta-analysis reduces the credible
 288 interval to 10 variants overlapping the promoter region of *RGS1*, which encodes a regulator of
 289 G-protein mediated signaling active in immune cell populations. We find a shared association
 290 between disease risk and *RGS1* expression in CD4 T cells, which is inverted so that lower
 291 expression correlates with higher disease risk. The lead credible interval variant overlaps a region
 292 of accessible chromatin within an active enhancer immediately upstream of the *RGS1* promoter.
 293 Further, this variant is annotated as a binding site for ZNF263 in the JASPAR database, and
 294 position-weight matrix analysis suggests the minor allele abrogates this binding site (24).
 295



296
 297 **Fig. 4: Jointly analyzing an association shared between multiple sclerosis and celiac disease**
 298 **improves fine-mapping resolution and identifies a shared eQTL for *RGS1*.** (A) Overlapping
 299 associations for the two diseases are due to a shared effect (JLIM $P = 5 \times 10^{-5}$ for CeD as primary trait; P
 300 $< 5 \times 10^{-5}$ for MS as primary trait). Meta-analyzing across the two diseases increases the overall
 301 significance and a narrower credible interval (credible interval variants for each panel are in dark grey; the
 302 physical span of the credible interval is shaded grey). The credible interval focuses on the intergenic region
 303 proximal to *RGS1*. (B) This shared association is also shared with an eQTL for *RGS1* in naïve CD4 T cells
 304 (JLIM $P = 0.015$). (C) The lead disease-associated variant lies in a region of accessible chromatin in naïve
 305 CD4 T cells and total T cells. This is marked with H3K27ac in total T cells and with H3K4me1 in naïve CD4
 306 T cells, suggesting this is an active, primed enhancer element. (D) The *RGS1* eQTL lead variant predicts
 307 the disease association P value, further indicating this is a shared effect. (E) Disease and eQTL association
 308 effects are negatively correlated, indicating that disease risk is associated with lower *RGS1* expression.

309 (F) *RGS1* is expressed at lower levels in T cells obtained from MS patients compared to healthy controls,
310 confirming this risk effect direction.

311
312 We have quantified the shared heritability between autoimmune and inflammatory diseases, and
313 demonstrated that we can leverage this to identify genetic variants that alter risk to multiple
314 diseases. This significantly increases fine-mapping resolution, compared to the original genetic
315 mapping studies: the number of effects where a single variant explains 95% of the posterior
316 probability of association increases from 13 to 20 (a 54% increase); for 50% of the posterior
317 probability, we see an increase from 35 to 54 (also 54%). Furthermore, we see an increase in the
318 number of eQTLs, with evidence of sharing an effect with disease risk, from 139 to 154 (11%).
319 Thus, in terms of identifying causal variants and functional interpretation, meta-analyzing across
320 diseases meaningfully increases our ability to interpret genetic associations. This sets the stage
321 for variant-to-function efforts to uncover key pathogenic mechanisms, as we provide high-value
322 targets relevant to multiple diseases.

323
324 This approach can be applied to any set of traits sharing associations; we therefore suggest this
325 is a fruitful avenue to maximize the interpretability of existing genetic studies of human complex
326 traits, especially as shared mechanisms are applicable to multiple conditions. It is particularly
327 valuable as sample collections, particularly of diseases that are difficult to diagnose or not
328 especially common in the population, become depleted. Disease cohorts are often genotyped
329 on different platforms, and the majority of common variants imputed. This can introduce a
330 substantial bias, if cohorts of samples with different diseases have differential genome coverage.
331 We have avoided this in our study by using a common platform, at the expense of not covering
332 the entire genome. These technical hurdles will diminish as genotyping platforms coalesce
333 around a standard set of variants, and as the community shifts to whole-genome sequencing
334 rather than genotyping. We note that biological interpretation of genetic associations, shared or
335 otherwise, is dependent on access to molecular and cellular phenotype studies such as eQTLs,
336 which require profiling a wide array of tissues or cell types under diverse stimuli in order to
337 identify the consequences of disease-associated variants. The BLUEPRINT dataset, which we
338 used here, covers three very different blood cell types, but dozens more exist, in which the
339 variants we have identified could act. This context specificity may be one reason we cannot
340 always assign a cognate eQTL to each well-resolved association (25).

341
342 In terms of understanding the common mechanisms in autoimmunity, we and others have
343 reported that many loci harbor associations to multiple autoimmune diseases. However, these
344 approaches have relied on simple proximity of variants to infer that the underlying mechanisms
345 must be shared. We have, for the first time, quantified the shared heritability between
346 autoimmune and inflammatory diseases, and shown that a substantial proportion of shared loci
347 harbor pleiotropic effects influencing risk to multiple diseases, which represent shared
348 mechanisms. Many loci, however, harbor multiple independent effects, indicative of distinct
349 mechanisms driving risk to different diseases; this is consistent either with the same underlying
350 genes being influenced in different contexts to induce risk for different diseases, or with different
351 genes which happen to be encoded near each other. Previous studies by us and others were
352 not designed with this resolution, and could only identify loci harboring potentially different
353 effects to multiple diseases.

354
355 Our results reveal complex patterns of shared heritability between autoimmune diseases. In
356 particular, we find many opposite effects shared between IBD and MS, where the same allele
357 increases risk for one disease but decreases risk for the other. This is reminiscent of the

358 differential outcomes of anti-TNF α therapies, which are beneficial in IBD but exacerbate MS
359 symptoms (26). Further, it suggests that some disease mechanisms may have an optimum state,
360 and either hypermorphism or hypomorphism are deleterious. Overall, we see no evidence for a
361 substantial component of risk shared across all six diseases, which would be indicative of a pan-
362 autoimmunity mechanism. Our benchmarking suggests this is not due to a lack of power to
363 detect shared effects (27), and our results strongly support independent effects in most loci. As
364 our results argue against a single, shared autoimmune mechanism, they also dispute a single
365 evolutionary origin for autoimmune and inflammatory diseases, which would have resulted in a
366 set of risk alleles driving broad autoimmunity.

367 References and Notes

- 368
- 369 1. M. D. Rosenblum, K. A. Remedios, A. K. Abbas, Mechanisms of human autoimmunity. *J.*
370 *Clin. Invest.* **125**, 2228–2233 (2015).
 - 371 2. International Multiple Sclerosis Genetics Consortium, Multiple sclerosis genomic map
372 implicates peripheral immune cells and microglia in susceptibility. *Science.* **365** (2019),
373 doi:10.1126/science.aav7188.
 - 374 3. K. M. de Lange, L. Moutsianas, J. C. Lee, C. A. Lamb, Y. Luo, N. A. Kennedy, L. Jostins,
375 D. L. Rice, J. Gutierrez-Achury, S.-G. Ji, G. Heap, E. R. Nimmo, C. Edwards, P.
376 Henderson, C. Mowat, J. Sanderson, J. Satsangi, A. Simmons, D. C. Wilson, M.
377 Tremelling, A. Hart, C. G. Mathew, W. G. Newman, M. Parkes, C. W. Lees, H. Uhlig, C.
378 Hawkey, N. J. Prescott, T. Ahmad, J. C. Mansfield, C. A. Anderson, J. C. Barrett,
379 Genome-wide association study implicates immune activation of multiple integrin genes in
380 inflammatory bowel disease. *Nat. Genet.* **49**, 256–261 (2017).
 - 381 4. M. T. Maurano, R. Humbert, E. Rynes, R. E. Thurman, E. Haugen, H. Wang, A. P.
382 Reynolds, R. Sandstrom, H. Qu, J. Brody, A. Shafer, F. Neri, K. Lee, T. Kutayvin, S.
383 Stehling-Sun, A. K. Johnson, T. K. Canfield, E. Giste, M. Diegel, D. Bates, R. S. Hansen,
384 S. Neph, P. J. Sabo, S. Heimfeld, A. Raubitschek, S. Ziegler, C. Cotsapas, N.
385 Sotoodehnia, I. Glass, S. R. Sunyaev, R. Kaul, J. A. Stamatoyannopoulos, Systematic
386 Localization of Common Disease-Associated Variation in Regulatory DNA. *Science.* **337**,
387 1190–1195 (2012).
 - 388 5. A. Gusev, S. H. Lee, G. Trynka, H. Finucane, B. J. Vilhjálmsson, H. Xu, C. Zang, S. Ripke,
389 B. Bulik-Sullivan, E. Stahl, SWE-SCZ Consortium, A. K. Kähler, C. M. Hultman, S. M.
390 Purcell, M. Daly, B. Pasaniuc, P. F. Sullivan, B. M. Neale, N. R. Wray, S. Raychaudhuri, A.
391 L. Price, Schizophrenia Working Group of the Psychiatric Genomics Consortium, S. A.
392 McCarroll, Partitioning heritability of regulatory and cell-type-specific variants across 11
393 common diseases. *Am. J. Hum. Genet.* **95**, 535–552 (2014).
 - 394 6. H. K. Finucane, B. Bulik-Sullivan, A. Gusev, G. Trynka, Y. Reshef, P.-R. Loh, V. Anttila, H.
395 Xu, C. Zang, K. Farh, S. Ripke, F. R. Day, ReproGen Consortium, Schizophrenia Working
396 Group of the Psychiatric Genomics Consortium, RACI Consortium, S. Purcell, E. Stahl, S.
397 Lindstrom, J. R. B. Perry, Y. Okada, S. Raychaudhuri, M. J. Daly, N. Patterson, B. M.
398 Neale, A. L. Price, Partitioning heritability by functional annotation using genome-wide
399 association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
 - 400 7. K. Hemminki, X. Li, K. Sundquist, J. Sundquist, Shared familial aggregation of
401 susceptibility to autoimmune diseases. *Arthritis Rheum.* **60**, 2845–2847 (2009).
 - 402 8. W. W. Eaton, N. R. Rose, A. Kalaydjian, M. G. Pedersen, P. B. Mortensen, Epidemiology
403 of autoimmune diseases in Denmark. *J. Autoimmun.* **29**, 1–9 (2007).
 - 404 9. C.-F. Kuo, M. J. Grainge, A. M. Valdes, L.-C. See, S.-F. Luo, K.-H. Yu, W. Zhang, M.
405 Doherty, Familial Aggregation of Systemic Lupus Erythematosus and Coaggregation of
406 Autoimmune Diseases in Affected Families. *JAMA Intern. Med.* **175**, 1518–1526 (2015).
 - 407 10. C. Cotsapas, B. F. Voight, E. Rossin, K. Lage, B. M. Neale, C. Wallace, G. R. Abecasis, J.
408 C. Barrett, T. Behrens, J. Cho, P. L. De Jager, J. T. Elder, R. R. Graham, P. Gregersen, L.
409 Klareskog, K. A. Siminovitch, D. A. van Heel, C. Wijmenga, J. Worthington, J. A. Todd, D.
410 A. Hafler, S. S. Rich, M. J. Daly, FOCiS Network of Consortia, Pervasive sharing of genetic
411 effects in autoimmune disease. *PLoS Genet.* **7**, e1002254 (2011).
 - 412 11. D. Ellinghaus, L. Jostins, S. L. Spain, A. Cortes, J. Bethune, B. Han, Y. R. Park, S.
413 Raychaudhuri, J. G. Pouget, M. Hübenal, T. Folseraas, Y. Wang, T. Esko, A. Metspalu,
414 H.-J. Westra, L. Franke, T. H. Pers, R. K. Weersma, V. Collij, M. D’Amato, J. Halfvarson, A.
415 B. Jensen, W. Lieb, F. Degenhardt, A. J. Forstner, A. Hofmann, International IBD Genetics

- 416 Consortium (IIBDGC), International Genetics of Ankylosing Spondylitis Consortium (IGAS),
417 International PSC Study Group (IPSCSG), Genetic Analysis of Psoriasis Consortium
418 (GAPC), Psoriasis Association Genetics Extension (PAGE), S. Schreiber, U. Mrowietz, B.
419 D. Juran, K. N. Lazaridis, S. Brunak, A. M. Dale, R. C. Trembath, S. Weidinger, M.
420 Weichenthal, E. Ellinghaus, J. T. Elder, J. N. W. N. Barker, O. A. Andreassen, D. P.
421 McGovern, T. H. Karlsen, J. C. Barrett, M. Parkes, M. A. Brown, A. Franke, Analysis of five
422 chronic inflammatory diseases identifies 27 new associations and highlights disease-
423 specific patterns at shared loci. *Nat. Genet.* **48**, 510–518 (2016).
- 424 12. M. Sirota, M. A. Schaub, S. Batzoglu, W. H. Robinson, A. J. Butte, Autoimmune disease
425 classification by inverse association with SNP alleles. *PLoS Genet.* **5**, e1000792 (2009).
- 426 13. M. van de Bunt, A. Cortes, I. Consortium, M. A. Brown, A. P. Morris, M. I. McCarthy,
427 Evaluating the Performance of Fine-Mapping Strategies at Common Variant GWAS Loci.
428 *PLOS Genet.* **11**, e1005535 (2015).
- 429 14. B. K. Bulik-Sullivan, P.-R. Loh, H. K. Finucane, S. Ripke, J. Yang, Schizophrenia Working
430 Group of the Psychiatric Genomics Consortium, N. Patterson, M. J. Daly, A. L. Price, B.
431 M. Neale, LD Score regression distinguishes confounding from polygenicity in genome-
432 wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
- 433 15. B. Bulik-Sullivan, H. K. Finucane, V. Anttila, A. Gusev, F. R. Day, P.-R. Loh, ReproGen
434 Consortium, Psychiatric Genomics Consortium, Genetic Consortium for Anorexia Nervosa
435 of the Wellcome Trust Case Control Consortium 3, L. Duncan, J. R. B. Perry, N.
436 Patterson, E. B. Robinson, M. J. Daly, A. L. Price, B. M. Neale, An atlas of genetic
437 correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
- 438 16. L. Jostins, S. Ripke, R. K. Weersma, R. H. Duerr, D. P. McGovern, K. Y. Hui, J. C. Lee, L.
439 P. Schumm, Y. Sharma, C. A. Anderson, J. Essers, M. Mitrovic, K. Ning, I. Cleyne, E.
440 Théâtre, S. L. Spain, S. Raychaudhuri, P. Goyette, Z. Wei, C. Abraham, J.-P. Achkar, T.
441 Ahmad, L. Amininejad, A. N. Ananthakrishnan, V. Andersen, J. M. Andrews, L. Baidoo, T.
442 Balschun, P. A. Bampton, A. Bitton, G. Boucher, S. Brand, C. Büning, A. Cohain, S.
443 Cichon, M. D’Amato, D. De Jong, K. L. Devaney, M. Dubinsky, C. Edwards, D. Ellinghaus,
444 L. R. Ferguson, D. Franchimont, K. Fransen, R. Gearry, M. Georges, C. Gieger, J. Glas, T.
445 Haritunians, A. Hart, C. Hawkey, M. Hedl, X. Hu, T. H. Karlsen, L. Kupcinskis, S.
446 Kugathasan, A. Latiano, D. Laukens, I. C. Lawrance, C. W. Lees, E. Louis, G. Mahy, J.
447 Mansfield, A. R. Morgan, C. Mowat, W. Newman, O. Palmieri, C. Y. Ponsioen, U.
448 Potocnik, N. J. Prescott, M. Rigueiro, J. I. Rotter, R. K. Russell, J. D. Sanderson, M.
449 Sans, J. Satsangi, S. Schreiber, L. A. Simms, J. Sventoraityte, S. R. Targan, K. D. Taylor,
450 M. Tremelling, H. W. Verspaget, M. de Vos, C. Wijmenga, D. C. Wilson, J. Winkelmann, R.
451 J. Xavier, S. Zeissig, B. Zhang, C. K. Zhang, H. Zhao, International IBD Genetics
452 Consortium (IIBDGC), M. S. Silverberg, V. Annese, H. Hakonarson, S. R. Brant, G.
453 Radford-Smith, C. G. Mathew, J. D. Rioux, E. E. Schadt, M. J. Daly, A. Franke, M. Parkes,
454 S. Vermeire, J. C. Barrett, J. H. Cho, Host-microbe interactions have shaped the genetic
455 architecture of inflammatory bowel disease. *Nature.* **491**, 119–124 (2012).
- 456 17. A. Zhernakova, S. Withoff, C. Wijmenga, Clinical implications of shared genetics and
457 pathogenesis in autoimmune diseases. *Nat. Rev. Endocrinol.* **9**, 646–659 (2013).
- 458 18. V. Matzaraki, V. Kumar, C. Wijmenga, A. Zhernakova, The MHC locus and genetic
459 susceptibility to autoimmune and infectious diseases. *Genome Biol.* **18**, 76 (2017).
- 460 19. A. Cortes, M. A. Brown, Promise and pitfalls of the Immunochip. *Arthritis Res. Ther.* **13**,
461 101 (2011).
- 462 20. C. Benner, C. C. A. Spencer, A. S. Havulinna, V. Salomaa, S. Ripatti, M. Pirinen,
463 FINEMAP: efficient variable selection using summary data from genome-wide association
464 studies. *Bioinforma. Oxf. Engl.* **32**, 1493–1501 (2016).

- 465 21. K. K.-H. Farh, A. Marson, J. Zhu, M. Kleinewietfeld, W. J. Housley, S. Beik, N. Shores, H.
466 Whitton, R. J. H. Ryan, A. A. Shishkin, M. Hatan, M. J. Carrasco-Alfonso, D. Mayer, C. J.
467 Luckey, N. A. Patsopoulos, P. L. De Jager, V. K. Kuchroo, C. B. Epstein, M. J. Daly, D. A.
468 Hafler, B. E. Bernstein, Genetic and epigenetic fine mapping of causal autoimmune
469 disease variants. *Nature*. **518**, 337–343 (2015).
- 470 22. G. Trynka, C. Sandor, B. Han, H. Xu, B. E. Stranger, X. S. Liu, S. Raychaudhuri, Chromatin
471 marks identify critical cell types for fine mapping complex trait variants. *Nat. Genet.* **45**,
472 124–130 (2013).
- 473 23. J. L. Pappalardo, L. Zhang, M. K. Pecsok, K. Perlman, C. Zografou, K. Raddassi, A.
474 Abulaban, S. Krishnaswamy, J. Antel, D. van Dijk, D. A. Hafler, Transcriptomic and clonal
475 characterization of T cells in the human central nervous system. *Sci. Immunol.* **5** (2020),
476 doi:10.1126/sciimmunol.abb8786.
- 477 24. S. Kumar, G. Ambrosini, P. Bucher, SNP2TFBS - a database of regulatory SNPs affecting
478 predicted transcription factor binding site affinity. *Nucleic Acids Res.* **45**, D139–D144
479 (2017).
- 480 25. B. D. Umans, A. Battle, Y. Gilad, Where Are the Disease-Associated eQTLs? *Trends*
481 *Genet. TIG.* **37**, 109–124 (2021).
- 482 26. The Lenercept Multiple Sclerosis Study Group and The University of British Columbia
483 MS/MRI Analysis Group, TNF neutralization in MS: Results of a randomized, placebo-
484 controlled multicenter study. *Neurology*. **53**, 457–457 (1999).
- 485 27. S. Chun, A. Casparino, N. A. Patsopoulos, D. C. Croteau-Chonka, B. A. Raby, P. L. De
486 Jager, S. R. Sunyaev, C. Cotsapas, Limited statistical evidence for shared genetic effects
487 of eQTLs and autoimmune-disease-associated loci in three major immune-cell types. *Nat.*
488 *Genet.* **49**, 600–605 (2017).
- 489 28. A. Rasmussen, S. Sevier, J. A. Kelly, S. B. Glenn, T. Aberle, C. M. Cooney, A. Grether, E.
490 James, J. Ning, J. Tesiram, J. Morrissey, T. Powe, M. Drexel, W. Daniel, B. Namjou, J. O.
491 Ojwang, K. L. Nguyen, J. W. Cavett, J. L. Te, J. A. James, R. H. Scofield, K. Moser, G. S.
492 Gilkeson, D. L. Kamen, C. W. Carson, A. I. Quintero-del-Rio, M. del Carmen Ballesteros,
493 M. G. Punaro, D. R. Karp, D. J. Wallace, M. Weisman, J. T. Merrill, R. Rivera, M. A. Petri,
494 D. A. Albert, L. R. Espinoza, T. O. Utset, T. S. Shaver, E. Arthur, J.-M. Anaya, G. R.
495 Bruner, J. B. Harley, The lupus family registry and repository. *Rheumatol. Oxf. Engl.* **50**,
496 47–59 (2011).
- 497 29. C. D. Langefeld, H. C. Ainsworth, D. S. Cunninghame Graham, J. A. Kelly, M. E. Comeau,
498 M. C. Marion, T. D. Howard, P. S. Ramos, J. A. Croker, D. L. Morris, J. K. Sandling, J. C.
499 Almlöf, E. M. Acevedo-Vásquez, G. S. Alarcón, A. M. Babini, V. Baca, A. A. Bengtsson, G.
500 A. Berbotto, M. Bijl, E. E. Brown, H. I. Brunner, M. H. Cardiel, L. Catoggio, R. Cervera, J.
501 M. Cucho-Venegas, S. R. Dahlqvist, S. D'Alfonso, B. M. Da Silva, I. de la Rúa Figueroa, A.
502 Doria, J. C. Edberg, E. Endreffy, J. A. Esquivel-Valerio, P. R. Fortin, B. I. Freedman, J.
503 Frostegård, M. A. García, I. G. de la Torre, G. S. Gilkeson, D. D. Gladman, I. Gunnarsson,
504 J. M. Guthridge, J. L. Huggins, J. A. James, C. G. M. Kallenberg, D. L. Kamen, D. R. Karp,
505 K. M. Kaufman, L. C. Kottyan, L. Kovács, H. Laustrop, B. R. Lauwerys, Q.-Z. Li, M. A.
506 Maradiaga-Ceceña, J. Martin, J. M. McCune, D. R. McWilliams, J. T. Merrill, P. Miranda,
507 J. F. Moctezuma, S. K. Nath, T. B. Niewold, L. Orozco, N. Ortego-Centeno, M. Petri, C. A.
508 Pineau, B. A. Pons-Estel, J. Pope, P. Raj, R. Ramsey-Goldman, J. D. Reveille, L. P.
509 Russell, J. M. Sabio, C. A. Aguilar-Salinas, H. R. Scherbarth, R. Scorza, M. F. Seldin, C.
510 Sjöwall, E. Svenungsson, S. D. Thompson, S. M. A. Toloza, L. Truedsson, T. Tusié-Luna,
511 C. Vasconcelos, L. M. Vilá, D. J. Wallace, M. H. Weisman, J. E. Wither, T. Bhangale, J. R.
512 Oksenberg, J. D. Rioux, P. K. Gregersen, A.-C. Syvänen, L. Rönnblom, L. A. Criswell, C.
513 O. Jacob, K. L. Sivils, B. P. Tsao, L. E. Schanberg, T. W. Behrens, E. D. Silverman, M. E.

514 Alarcón-Riquelme, R. P. Kimberly, J. B. Harley, E. K. Wakeland, R. R. Graham, P. M.
515 Gaffney, T. J. Vyse, Transancestral mapping and genetic load in systemic lupus
516 erythematosus. *Nat. Commun.* **8**, 16021 (2017).

517
518

519 **Acknowledgements:**

520 We thank the EAGLE eczema consortium for providing GWAS summary statistics. This
521 research utilizes resources provided by the T1DGC, a collaborative clinical study sponsored by
522 the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute
523 of Allergy and Infectious Diseases, National Human Genome Research Institute, National
524 Institute of Child Health and Human Development, and JDRF and supported by U01
525 DK062418. We thank the RACI consortium for access to RA data and the International IBD
526 Genetics Consortium for access to IBD data. De-identified data were provided from a total of
527 4617 samples (2563 SLE cases and 2054 population controls) in the Lupus Family Registry and
528 Repository collection (28) at the Oklahoma Medical Research Foundation. The SLE Genentech
529 samples were originally genotyped and analyzed as part of a large SLEGEN Consortium
530 ImmunoChip study (29).

531

532 **Funding:** MRL is supported by a Career Transition Fellowship from the Consortium of MS
533 Centers and the National MS Society. CG received a research fellowship from the Deutsche
534 Forschungsgemeinschaft (DFG, German Research Foundation) for this project. She further
535 received funding from the Hans und Klementia Langmatz-Stifung and the Hertie Network of
536 Excellence in Clinical Neuroscience, not related to this study. CW was supported by an ERC
537 Advanced grant (FP/2007- 2013/ERC grant 2012-322698), the NWO Spinoza prize grant (NWO
538 SPI 92-266), a grant from Stiftelsen K. G. Jebsen, and The Netherlands Organ-on-Chip
539 Initiative - an NWO Gravitation project (024.003.001) funded by the Ministry of Education,
540 Culture and Science of the government of The Netherlands. SW was supported by The
541 Netherlands Organ-on-Chip Initiative, an NWO Gravitation project (024.003.001) funded by the
542 Ministry of Education, Culture and Science of the government of The Netherlands. I.H.J. is
543 supported by a Rosalind Franklin Fellowship from the University of Groningen and an NWO
544 VIDI grant (no. 016.171.047). This work is in part supported by NIH/NIAID grant R01 AI122220
545 to CC.

546

547 **Authors contributions:**

548 MRL, NC, CG, and MM curated and analyzed data;
549 DvH, CW, SW, IHJ, LP, IMSGC, SSR, RRG, PMG, CDL, and DAH provided data;
550 MRL and CC wrote and edited the manuscript, with input from all co-authors;
551 SC and SRS designed and implemented analytical methods;
552 CC conceptualized and oversaw the project.

553

554 **Competing interests:** None of the authors have any competing interests.

555

556 **Data and materials availability:** Dataset availability is documented in Table S1. Analysis
557 source code is available from the Cotsapas lab GitHub page.