Chromatin conformation dynamics during CD4+ T cell activation implicates autoimmune disease-associated genes and regulatory elements

Matthew C. Pahl*[1,2], Prabhat Sharma*[1,3], Rajan M. Thomas[1,3], Zachary Thompson[1,3], Zachary Mount[1,3], James Pippin[1,2], Peter A. Morawski[4], Chun Su[1,2], Daniel J. Campbell[4,5], Struan F.A. Grant[1,2,6,7,8], Andrew D. Wells[1,3,9,10]

[1]Center for Spatial and Functional Genomics, The Children's Hospital of Philadelphia
[2]Division of Human Genetics, The Children's Hospital of Philadelphia
[3]Department of Pathology, The Children's Hospital of Philadelphia
[4]Benaroya Research Institute at Virginia Mason, Seattle, WA
[5]Department of Immunology, University of Washington School of Medicine, Seattle, WA
[6]Department of Genetics, Perelman School of Medicine, University of Pennsylvania
[7]Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania
[8]Division of Endocrinology and Diabetes, The Children's Hospital of Philadelphia
[9]Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania
[10]Institute for Immunology, Perelman School of Medicine, University of Pennsylvania

*Co-lead authors

Correspondence:
Andrew D. Wells, PhD
The Children's Hospital of Philadelphia Research Institute
916D Abramson Research Center
3615 Civic Center Boulevard
Philadelphia, PA 19104
Email: wellsa@chop.edu
Phone: 215-590-8710

ABSTRACT

Genome-wide association studies (GWAS) have identified hundreds of genetic loci associated with autoimmune disease, but the causal variants and their contribution to immune dysregulation remain largely unknown. We measured the dynamics of chromosome conformation, chromatin accessibility, and gene transcription across three phases of naive human CD4+ T cell activation and co-localized active *cis*-regulatory elements with the 95% credible set of variants from 15 autoimmune GWAS. Reorganization of chromosome structure placed these elements in direct contact with ~1,200 protein-coding genes, at least one-third of which were dynamically regulated by stimulation. The set of implicated genes is enriched for high-throughput CRISPR screen targets that control multiple aspects of CD4+ T cell activation, and we pharmacologically validated eight novel gene products as potent regulators of T cell proliferation. These maps also allowed the identification and functional validation of a novel stretch of intergenic enhancers for the autoimmune gene *IL2* whose activity is influenced by autoimmune-associated genetic variation. This study represents a powerful strategy and resource for assigning physiologic relevance to autoimmune-risk variants and identify novel genes that control T cell activation and function.

45 **INTRODUCTION**

46 GWAS has revealed hundreds of regions of the human genome robustly associated with
47 autoimmune disease susceptibility. The majority of GWAS variants are located in non-coding
48 regions of the genome, and likely contribute to disease risk by modulating cis-regulatory element
49 (cRE) activity to influence gene expression[1]. Identifying causal variants and their corresponding
50 effector genes molecularly responsible for increased disease risk is critical for identifying targets
51 for downstream molecular study and therapeutic intervention. Identifying and prioritizing disease-
52 causal variants is limited by a lack of genetic fine-mapping and functional studies that define the
53 mechanisms of control over cRE activity and gene expression.

54 CD4+ T cells are key regulators of innate and adaptive immune responses that combat
55 infection by orchestrating the activity of other immune cells. In their quiescent, naïve state, 'helper'
56 T cells traffic between the blood and secondary lymphoid tissues as part of an
57 immunosurveillance program maintained by transcription factors such as KLF2, TOB, and
58 FOXO[2–4]. Upon encountering specific antigen, a cascade of signals activated through the TCR
59 and costimulatory receptors results in large-scale changes in gene expression driven by factors
60 like NFKB, NFAT, IRF4, and STATs, leading to proliferation and differentiation into specialized
61 Th1, Th2, and Th17 subsets capable of participating in protective immunity[5,6]. CD4+ T cells are
62 also key players in the induction and pathogenesis of autoimmunity. The cis-regulatory
63 architecture of CD4+ T cells is enriched for autoimmune disease GWAS variants[7–14], and T cells
64 from autoimmune patients harbor distinct epigenetic and transcriptomic signatures linking
65 dysregulated gene expression with disease pathogenesis. Autoimmune variants may therefore
66 contribute to the breakdown of immune self-tolerance by shifting the balance between
67 autoreactive conventional vs. regulatory CD4+ T cells, by altering cytokine production, or by
68 promoting auto-antigen production[15,16].

69 Chromatin conformation assays allow for identification of putative target genes of autoimmune
70 variant-containing cRE in close physical proximity to gene promoters. Previous work using

71 promoter-capture HiC and HiC in combination with other epigenetic marks have implicated sets
72 of autoimmune variants and effector genes that may participate in T cell activation[17,18]. In addition
73 to chromatin conformation approaches, multiple complementary approaches have been
74 developed to link disease-associated variants to their downstream effector genes. Expression
75 quantitative trait mapping (eQTL) and chromatin co-accessibility data can be used to implicate
76 effector genes through statistical association of genotypes with readouts of expression or other
77 molecular markers[19]. However, these approaches do not account for the relevant 3D structure of
78 the genome in the nucleus, and are highly susceptible to *trans*-effects and other confounding
79 factors.

80 In this study, we define the impact of TCR-CD28 activation on the autoimmune-associated
81 *cis*-regulatory architecture of CD4+ helper T cells, leading to the discovery of a stretch of distal
82 enhancers required for normal expression of the canonical T cell activation gene *IL2* whose
83 activity is influenced by autoimmune risk variants. The set of variant-connected effector genes
84 defined by 3D physical proximity to autoimmune-associated cRE is enriched for genes that
85 regulate T cell activation, as validated pharmacologically in this current study and by CRISPR-
86 based screens reported in orthogonal studies[20–22].

87

88 **RESULTS**

89 **Gene expression dynamics as a function of naïve CD4+ T cell activation**

90 To identify the set of genes and cREs that are affected by and may contribute to CD4+
91 activation, we characterized the dynamics of gene expression, chromatin accessibility, and 3D
92 chromatin conformation in human CD3+CD4+CD45RA+CD45RO- naïve T cells purified directly
93 *ex vivo* and in response to *in vitro* activation through the T cell receptor (TCR) and CD28 for 8 or
94 24 hours using RNA-seq, ATAC-seq, and HiC (**Fig. 1A**). ATAC-seq and RNA-seq libraries were
95 prepared using CD4+ T cells derived from N=3 donors, while HiC libraries were prepared from
96 N=2 donors. We verified the reproducibility of replicate ATAC-seq and RNA-seq libraries with

97 principal component analysis and HiC libraries using distance-controlled, stratum-adjusted
98 correlation coefficient (SCC). Replicate samples were highly correlated and clustered by
99 activation stage (**Fig. S1A-C**). As expected, quiescent naïve CD4+ T cells expressed high levels
100 of *SELL, TCF7, CCR7,* and *IL7R*, and rapidly upregulated *CD69, CD44, HLADR,* and *IL2RA* upon
101 stimulation (**Fig 1B**). Genome-scale gene set variance analysis based on MsigDB hallmark
102 pathways showed that genes involved in KRAS and Hedgehog signaling are actively enriched in
103 quiescent naive CD4+ T cells, while stimulated cell transcriptomes are enriched for genes
104 involved in cell cycle, metabolism, and TNF-, IL-2/STAT5-, IFNg-, Notch-, and MTORC1-mediated
105 signaling pathways (**Fig. 1C**, **Table S1**).

106 To define global gene expression dynamics during the course of CD4+ T cell activation, we
107 performed pairwise comparisons and k-means clustering across the three cell states assayed.
108 This approach identified 4390 differentially expressed genes after 8 hours of stimulation (3289
109 upregulated and 1101 downregulated) and 3611 differentially expressed genes between 8 hours
110 and 24 hours (3015 upregulated and 596 downregulated, **Fig. S1D**) that could be further
111 separated into five clusters based on their distinct trajectories: upregulated early (cluster 1;
112 n=1621 genes), downregulated late (cluster 2; n=1600 genes), monotonically increasing (cluster
113 3; n=2676 genes), downregulated early (cluster 4; n=1628 genes) and upregulated late (cluster
114 5; n=2154 genes, **Fig. 1D**, **Fig. S1E-G**, **Table S2**).

115 Genes upregulated early (cluster 1; *e.g.*, *IL2, IFNG, TNF, IL3, IL8, IL2RA, IL15RA, IL21R,*
116 *ICOS, CD40LG, FASLG, MYC, FOS, FOSB, FOSL1, FOSL2, JUNB, REL, NFKBID, NFKB1,*
117 *RELB, NFKBIA, STAT5A, BATF, BTLA*) are enriched for pathways involved in the unfolded
118 protein response, cytokine signaling, and translation (**Fig 1E**; **Table S3**), and genes
119 downregulated late (cluster 2) are moderately enriched for receptor tyrosine kinase signaling,
120 cytokine signaling, and extracellular matrix organization. Genes monotonically increasing upon
121 activation (cluster 3; *e.g.*, *TBX21, BHLHE40, LTA, IL12RB2, STAT1, SLAMF7, CCND2, CDK4,*
122 *PRMT1, ICAM1, NFKB2, EZH2*) are highly enriched for pathways involved in infectious disease

123 and RNA stability, translation and metabolism, and moderately enriched for pathways involved in
124 the unfolded protein response, cellular responses to stress, regulation of apoptosis, and DNA
125 repair. Genes downregulated early upon stimulation (cluster 4; *e.g.*, *KLF2, IL7R, NOG, PIK3IP1,*
126 *TSC22D3, RORA, IL10RA*) are enriched for pathways involved in inositol phosphate biosynthesis,
127 neutrophil degranulation, and metabolism of nucleotides, and genes upregulated late (cluster 5;
128 *e.g.*, *CDK2, E2F1, CDK1, CCNE1, CCNA2, PCNA, WEE1, CDC6, ORC1, MCM2*) are highly
129 enriched for pathways involved in cell cycle, DNA unwinding, DNA repair, chromosomal
130 maintenance, beta-oxidation of octanoyl-CoA, and cellular response to stress (**Fig 1E**; **Table S3**).
131 These patterns are consistent with known changes in the cellular processes that operate during
132 T cell activation, confirming that *in vitro* stimulation of CD4+ T cells recapitulates gene expression
133 programs known to be engaged during a T cell immune response.
134

135 **Dynamic changes in chromosomal architecture and genome accessibility during naïve**
136 **CD4+ T cell activation**

137 To understand the *cis*-regulatory dynamics underlying the observed activation-induced
138 changes in gene expression, we examined CD4+ T cell nuclear chromosome conformation and
139 chromatin accessibility as a function of stimulation state using HiC and ATAC-seq. The human
140 genome consists of ~3 meters of DNA that is incorporated into chromatin and compacted into the
141 ~500 cubic micron nucleus of a cell in a hierarchically ordered manner. This degree of compaction
142 results in only ~1% of genomic DNA being accessible to the machinery that regulates gene
143 transcription[23], therefore a map of open chromatin regions (OCR) in a cell represents its potential
144 gene regulatory landscape. Open chromatin mapping of human CD4+ T cells at all states
145 identified a consensus *cis*-regulatory landscape of 181,093 reproducible OCR (**Table S4**). Of
146 these, 14% (25,291) exhibited differential accessibility following 8 hours of stimulation
147 (FDR<0.05). Most differentially accessible regions (DAR) became more open (18,887), but some
148 DAR (6,629) showed reduced accessibility (**Fig. 2A**, **Table S5**). The change in accessibility over

149 the next 16 hours of stimulation showed the opposite dynamic, with 6,629 regions exhibiting
150 reduced accessibility, and only 4,417 DAR becoming more open (total of 11,046 DAR, **Fig. 2A**,
151 **Table S5**). These OCR represent the set of putative cRE with dynamic activity during T cell
152 activation.

153 The vast majority of putative cRE are located in intergenic or intronic regions of the genome
154 far from gene promoters, meaning that the specific impact of a given cRE on gene expression
155 cannot be properly interpreted from a one-dimensional map of genomic or epigenomic features.
156 To predict which cRE may regulate which genes in CD4+ T cells across different states of
157 activation, we created three-dimensional maps of cRE-gene proximity in the context of genome
158 structure. The highest order of 3D nuclear genome structure is represented by A/B compartments,
159 which are large chromosomal domains that self-associate into transcriptionally active (A) *vs*.
160 inactive (B) regions[24]. In agreement with prior studies, we find that OCR located in active A
161 compartments exhibit higher average accessibility than those OCR located in less active B
162 compartments (**Fig S2A; Table S6**), a trend observed across all cell states. Likewise, genes
163 located in A compartments show higher average expression than those located in B
164 compartments (**Fig S2B**). A quantitative comparison across cell states showed that 94% of the
165 CD4+ T cell genome remained stably compartmentalized into A (42%) and B (52%), indicating
166 that activation does not cause a major shift in the large-scale organization of the genome within
167 the nucleus. However, ~6% of the CD4+ T cell genome did experience a transient (A>B>A or
168 B>A>B) or stable (A>B>B, A>A>B, B>B>A or B>A>A) shift in A/B compartmentalization in
169 response to activation (**Fig. 2B**). Genes that switched from A to B compartments are enriched for
170 processes involved in fatty acid receptors and metabolism, eicosanoid and leukotriene
171 biosynthesis, Runx and Yap activity, and IL-20 signaling, while genes that switched from B to A
172 compartments are enriched for pathways involved in cell-ECM and cell-cell interactions, vitamin
173 metabolism, and MECP2 activity (**Fig. 2B**), and include activation-induced genes known to
174 regulate T cell function like *EGR2* and *IL12RB2*.

175 Within each A or B compartment, the genome is further organized into topologically
176 associating domains (TADs)[25]. These structures are defined by the fact that genomic regions
177 within them have the potential to interact with each other in 3D, but have low potential to interact
178 with regions outside the TAD. The location of TAD boundaries can influence gene expression by
179 limiting the access of cRE to specific, topologically associated genes. While ~80% of TAD
180 boundaries remained stable across all states, 20% of TAD boundaries (8925) changed as a
181 function of T cell activation (**Fig. 2C**, **Fig. S2C** and **D**, **Table S7**). TAD boundary dynamics were
182 categorized and 2198 boundaries exhibited a change in strength, 2030 boundaries shifted
183 position, and 4697 boundaries exhibited more complex changes such as loss of a boundary
184 resulting in merger of two neighboring TAD, addition of a boundary splitting one TAD into two, or
185 a combination of any of these changes. Genes nearby dynamic TAD boundaries are enriched for
186 pathways involved in RNA metabolism, cellular response to stress, and the activity of PTEN, p53,
187 JAK/STAT, Runx and Hedgehog (**Fig. 2C**).

188 We detected chromatin stripes, which are TAD-like structures that consist of a genomic anchor
189 region that is brought into contact with a larger loop domain via an active extrusion mechanism.
190 Chromatin stripes are contained within and/or overlap TAD regions, and are enriched for active
191 enhancers and super-enhancers[26,27]. We identified 1526 chromatin stripes in quiescent naïve
192 CD4+ T cells, 1676 stripes in 8 hour stimulated cells, and 2028 stripes at 24 hours post-stimulation
193 (**Fig S3A**). Consistent with prior studies in other cell types, chromatin stripes were preferentially
194 located in A compartments, and genes and OCR within stripe regions showed increased
195 expression and chromatin accessibility (**Fig. S3B-D**).

196 Within active topological structures, transcriptional enhancers can regulate the expression of
197 distant genes by looping to physically interact with gene promoters[28,29]. To identify potential
198 regulatory cRE-gene interactions, we identified high confidence loop contacts across all cell states
199 using Fit-HiC (merged 1kb, 2kb, and 4kb resolutions, **Fig. 2D**). This approach detected 933,755
200 loop contacts in quiescent naïve CD4+ T cells, 900,267 loop contacts in 8-hour stimulated cells,

201 and 551,802 contacts in 24-hour stimulated cells (2,099,627 total unique loops). Approximately
202 23% of these loops involved a gene promoter at one end and an OCR at the other, and these
203 promoter-interacting OCR were enriched for enhancer signatures based on flanking histone
204 marks from CD4+ T cells in the epigenome roadmap database[30] (**Fig. S2E**). T cell activation
205 resulted in significant reorganization of the open chromatin-promoter interactome, as 907
206 promoter-OCR exhibited increased contact and 1333 showed decreased contact following 8
207 hours of stimulation (**Fig. 2E**). Continued stimulation over the next 16 hours was associated with
208 an increase in the contact frequency of 41 promoter-OCR pairs, while only 4 pairs exhibited
209 decreased contact (**Fig. 2E**). Activation-induced changes in chromatin architecture and gene
210 expression were highly correlated, as genomic regions exhibiting increased promoter connectivity
211 became more accessible at early stages of stimulation, which was associated with increased gene
212 transcription from connected promoters (**Fig. 2F, Table S8**). The accessibility of promoter-
213 connected OCR and the expression of their associated genes decreased globally from 8 to 24
214 hours of stimulation (**Fig. 2F**). We compared these loop calls to a prior chromatin capture analysis
215 in CD4+ T cells by Burren et al.[17] and found that roughly 40% of stable loops in both stimulated
216 and unstimulated cells were identical in both studies, despite differing in approach (HiC vs.
217 PCHiC), analysis (HiC vs. CHiCAGO), sample (naïve CD4+ vs. total CD4+), timepoint (8 vs. 4
218 hour), and donor individuals (**Fig. S2F**). As expected, unstimulated samples were more similar
219 than activated samples.

220 We next focused on the 5 sets of genes with the dynamic expression patterns defined in **Fig.**
221 **1D**, and identified 57,609 OCR that contact dynamic gene promoters in at least one stage. Most
222 dynamic genes contacted between 1 and ~35 OCR, with a median of 10 OCR per gene, but a
223 handful of dynamic genes were observed in contact with over 100 distinct open chromatin regions
224 (**Fig. S2G**). Similarly, most OCR were connected to a single dynamic gene, but many contacted
225 more than one gene (median 2 genes per OCR), suggesting that most dynamic genes have a
226 complex regulatory architecture. Increased gene expression upon activation correlated with an

227 increase in the accessibility and promoter contact frequency of distant cRE (**Fig**. **S2H**), as
228 exemplified by *GEM* and *IRF4* (**Fig. S4A**,**B**). Conversely, the 3D regulatory architecture of genes
229 like KLF2 and DPEP2, which were downregulated following stimulation, exhibited decreased
230 contact and accessibility (**Fig. S4C**,**D**).
231

232 **Transcription factor footprints enriched in dynamic open chromatin identify regulators of**
233 **T cell activation**

234 To explore what factors drive dynamic changes in the regulatory architecture of the CD4+ T
235 cell genome during activation, we conducted a quantitative footprinting analysis at 1,173,159
236 transcription factor (TF) motifs located in the consensus CD4+ T cell open chromatin landscape
237 using the average accessibility of the region surrounding each motif as a measure of regulatory
238 activity (**Table S9**). Activation of naïve CD4+ T cells resulted in increased chromatin accessibility
239 around bZIP motifs at tens of thousands of genomic regions by 8 hours post-stimulation (**Fig. 3A**
240 and **C**, **Table S9**), which was associated with increased expression of Fos and Jun family
241 members constituting the AP-1 complex, as well as the bZIP factors BATF, BACH1, and BACH2
242 (**Fig. 3A**). The activity of NFE2 and SMAD was increased without increased expression (**Fig. 3A**
243 and **C**), likely due to post-translational regulation of these factors by phosphorylation[31].
244 Conversely, the motifs for a number of TF exhibited significantly reduced accessibility early after
245 stimulation, including those for EGR2 and FOXP3 that are known to negatively regulate T cell
246 activation[32,33] (**Fig. 3A**). By 24 hours post-activation, bZIP activity remained largely unchanged
247 compared to 8 hours (**Fig. 3B**), but a number of factors showed decreased activity. These include
248 several members of the Sp family, the Myc cofactor MAZ that also cooperates with CTCF to
249 regulate chromatin looping[34], KLF2, which controls genes involved in naïve CD4+ T cell
250 quiescence and homing[2,35], NRF1, a factor implicated in age-associated T cell hypofunction[36],
251 and EGR2 and 3, which are known to oppose T cell activation and promote tolerance[32,37,38] (**Fig.**
252 **3B** and **C**).

253 To explore how transcription factor activity may operate *via* the CD4+ T cell open chromatin
254 landscape to regulate distinct programs of dynamic gene expression during TCR/CD28-induced
255 activation, we focused on TF motifs enriched among those OCR specifically contacting promoters
256 of dynamic genes identified by our clustering analysis (**Fig. S5A**). The set of OCR contacting
257 dynamic gene promoters were enriched for the motifs of 89 expressed (TPM>5) transcription
258 factors, as compared to motifs present in the total open chromatin landscape (**Fig. 3D**). The
259 majority of this TF activity was enriched in OCR connected to genes highly upregulated at 24
260 hours post-activation (clusters 3 and 5, **Fig. 3D**), with the exception of CREB3, ELF1, ESRRA,
261 GABPA, RELA, XPB1, ZNF384, and the transcriptional repressor IKZF1 known as a strong
262 negative regulator of T cell activation and differentiation[39–42]. Conversely, motifs for IKZF1,
263 ZNF384, GABPA, ESRRA, and ELF1 were highly enriched in the set of OCR contacting genes
264 down-regulated early after activation (cluster 4, **Fig. 3D**). Motifs for KLF2 and the metabolic gene
265 regulator SREBF1 were likewise enriched in OCR connected to down-regulated genes. Open
266 chromatin regions interacting with genes in cluster 2 are negatively enriched for this set of TF
267 except for CTCF (**Fig. 3D**).
268 Finally, we integrated TF footprint, promoter connectome, and gene co-expression data to
269 construct TF-gene regulatory networks likely operating at each timepoint. The connections
270 between regulatory nodes are based on physical promoter-TF footprint interactions with
271 confidence weighted by gene co-expression (**Fig. 3E**, **Table S10**). Highly co-expressed genes at
272 the core of the unstimulated CD4+ T cell regulatory network encode transcription factors such as
273 KLF2, ETS1, IKZF1, and TCF7[43–45] that are known to be involved in T cell gene silencing,
274 quiescence, homeostasis, and homing. Genes connected to the top factors in this network were
275 enriched for pathways involved in immune signaling, DNA replication and repair, protein secretion,
276 and programmed cell death (**Fig S5B**). Costimulation through the TCR and CD28 induced a set
277 of core network genes active at both time points with known roles in T cell activation and
278 differentiation (*NFKB1, JUNB, MYC, IRF4, STAT5, STAT1, LEF1, ATF4*). Also part of this set is

279 *PLAGL2*, an oncogene in the Wnt pathway that regulates hypoxia-induced genes[46] with no prior
280 defined role in T cell activation. Additional nodes specifically implicated at 8 hours post-activation
281 are HIF1A, the major sensor of cellular hypoxia[47], and XBP1, a major transcriptional mediator of
282 the unfolded ER protein response with defined roles in T cell activation, differentiation, and
283 exhaustion[48,49]. Factors specifically implicated at 24 hours post-activation include E2F1, a
284 transcriptional regulator of both cell cycle and apoptosis in T cells[50,51], BHLHE40, a factor known
285 to control multiple aspects of T cell metabolism and differentiation[52], and the Myc cofactor MAZ
286 that has not been previously studied in the context of T cell function. Genes connected to factors
287 in the activated T cell networks were enriched for pathways involved in cytokine signaling, the
288 interferon response, transcription, cell cycle, DNA replication and repair, and programmed cell
289 death (**Fig S5B**). Together, these data indicate that concurrent but separable stimulation-induced
290 gene programs are the result of the activity of distinct sets of DNA binding factors mobilized by
291 antigen and costimulatory receptor signaling in naïve CD4+ T cells.
292

293 **Identification of autoimmune variants associated with CD4+ T cell cREs and putative**
294 **effector genes**

295 Following our established variant-to-gene (V2G) mapping approach to implicate functional
296 SNPs and their effector genes using the combination of GWAS and chromatin conformation
297 capture data[53–57], we intersected promoter-interacting OCR with autoimmune SNPs from the 95%
298 credible set derived from 15 autoimmune diseases (**Fig. 4A**, **Tables S11** and **S12**). Constraining
299 the GWAS SNPs in this way reduced the credible set size from an average of 14 variants per
300 sentinel to 3 variants per sentinel. To determine whether open chromatin in physical contact with
301 dynamically regulated genes in CD4+ T cells is enriched for autoimmune disease heritability, we
302 performed a partitioned LD score regression analysis. This landscape was enriched for variants
303 associated with susceptibility to inflammatory bowel disease (IBD), ulcerative colitis (UC), type I
304 diabetes (T1D), lupus (SLE), celiac disease (CEL), allergy (ALG), eczema (ECZ), and rheumatoid

305 arthritis (RA), but not for variants associated with psoriasis (PSO) or juvenile idiopathic arthritis
306 (JIA) (**Fig. 4B**, **Table S13**). The OCR connected to genes upregulated early and/or progressively
307 upon activation (clusters 1 and 3) were most strongly enriched for ALG, CEL, IBD/UC, RA and
308 T1D heritability, while SLE and ECZ heritability was most enriched in OCR connected to genes
309 upregulated later post-activation (clusters 3 and 5, **Fig. 4B**). SLE was also the only disease
310 (besides PSO and JIA) that was not enriched in open chromatin connected to down-regulated
311 genes (cluster 4, **Fig. 4B**).

312 The promoter-connected open chromatin landscape for all CD4+ T cell states in this study
313 contains 2606 putatively causal variants linked to ~half of the sentinel signals (423) for the 15
314 diseases analyzed, and are in contact with a total of 1836 genes (**Table S12**). A total of 1151
315 autoimmune variants localized to the promoters of 400 genes (-1500/+500bp from the TSS, **Table**
316 **S14**). These variants were on average 103 kb from the TSS of their connected gene (**Fig. 4C**),
317 and each variant contacted an average of 5 genes (**Fig. 4D**). The majority of contacted SNPs
318 interact with genes in addition to the nearest gene, and ~half of linked SNPs 'skip' the nearest
319 gene to target only distant genes (**Fig. 4E**). Approximately 60% of connected genes were
320 implicated across all timepoints (**Fig. 4F**, **Table S12**), while ~40% (753) are dynamically regulated
321 (clusters 1-5) in response to TCR/CD28 co-stimulation. Examples of SNP-genes pairs that exhibit
322 dynamic accessibility, chromosome contact, and expression in response to T cell activation are
323 the *SIK1*, *PARK7*, *DUSP5*, *CLEC2D*, *TRIP10*, *GPR108* and *IL2* loci (**Fig. 4G**, **Table S12**). The
324 *TRIP10* and *GPR108* promoters were each captured in contact with a high confidence variant
325 rs1077667 (PP>0.99), which is located in an intron of *TNFSF14* and is associated with multiple
326 sclerosis (**Fig. 4H**). The accessibility of this SNP and it's contact with *TRIP10* and *GPR108*
327 increased following activation (**Fig. 4H**). Conversely, the allergy-associated SNP rs7380290 is
328 accessible and contacts the *SIK1* promoter in resting cells, but shows reduced accessibility and
329 promoter connectivity upon activation (**Fig. 4I**). *TRIP10* encodes a cytoskeletal binding protein
330 involved in endocytosis that suppresses glucose uptake in response to insulin signaling[58],

331 *GPR108* encodes an orphan G-protein coupled receptor, and *SIK1* encodes a salt-inducible
332 kinase with roles in cancer, epilepsy, and myeloid signaling[59]. None of these genes have been
333 previously implicated in T cell activation.
334
335 **Functional validation of autoimmune V2G-implicated cis-regulatory elements**
336 This analysis specifically implicated potential dynamic, disease-associated regulatory
337 elements in intergenic space at the *IL2* locus. The *IL2* gene encodes a cytokine with crucial,
338 pleotropic roles in the immune system, and dysregulation of IL-2 and IL-2 receptor signaling leads
339 to immunodeficiency and autoimmune disorders in mice and humans[60–63]. Activation-induced
340 transcription of *IL2* involves an upstream regulatory region (URR) ~375 bp from the TSS that has
341 served as a paradigm of tissue-specific, inducible gene transcription for nearly four decades[64–66].
342 However, the presence of evolutionarily conserved non-coding sequences (CNS) in the ~150 kb
343 of intergenic space 46, 51, 80, 83, 85, 96, 122, and 128 kb upstream of the TSS suggest that
344 additional regulatory elements may have evolved to control *IL2*[67] (**Fig. 5A**). The -51 kb CNS
345 contains a SNP linked to T1D, IBD, PSO, CEL and allergy (rs72669153), while the -85 kb CNS
346 contains a SNP linked to RA (rs6232750) and the -128 kb CNS contains two SNPs linked to T1D,
347 JIA, and SLE (rs1512973 and rs12504008, **Fig. 5A**). In TCR/CD28-stimulated naïve CD4+ T cells,
348 these CNS are remodeled to become highly accessible (**Fig. 5A**), and they loop to interact
349 physically with the *IL2* URR (**Fig. 5B**) at both time points. ChIP-seq analyses in human T cells
350 (ENCODE[68]) show that the URR and all distal CNS except -85 are occupied by TF such as
351 Jun/Fos (AP-1), NFAT, and NFkB that are known regulators of *IL2* (**Fig. S6**), and the -85, -122,
352 and -128 CNS are occupied by additional TF not previously thought to be direct regulators of *IL2*,
353 such as MYC, BCL6, and STAT5 (**Fig. S6**). Recombinant reporter assays in primary activated
354 human CD4+ T cells showed that the -46, -51, -83, and -128 CNS/OCR can enhance transcription
355 from the URR (**Fig. 5C**). To determine whether the native elements contribute to the expression
356 of *IL2*, we targeted each CNS/OCR individually in primary human CD4+ T cells or Jurkat T cells

357 using CRISPR/CAS9 (**Fig. 5D**) and measured IL-2 secretion following TCR/CD28 stimulation
358 (**Fig. 5E**). Deletion of the -46, -51, -83, -85, -122, and -128 kb elements in primary human CD4+
359 T cells each resulted in a ~50% reduction in IL-2 production, while deletion of the -80 kb element
360 had little effect. A very similar pattern of impact was observed when these elements were deleted
361 individually in Jurkat T cells (**Fig. 5E**). The URR has a stronger contribution to IL-2 production
362 than any individual intergenic element, as deletion of the URR almost completely abrogated
363 activation-induced IL-2 production by both primary CD4+ or Jurkat T cells (**Fig. 5E**). To determine
364 whether these intergenic enhancers exist in synergistic epistasis[69] necessary for *IL2*
365 transactivation, we generated Jurkat T cell clones in which the stretch of all 7 elements located -
366 46 kb to -128 kb upstream of *IL2* was deleted using CRISPR/CAS9 (**Fig. 5F**). Despite the URR
367 and 46 kb of upstream sequence being intact in these clones, loss of the 81.3 kb stretch of
368 intergenic enhancers renders these cells incapable of expressing *IL2* at both the mRNA and
369 protein level in response to stimulation (**Fig. 5G**). These results show that the URR is not sufficient
370 for activation-induced expression of *IL2*, and that *IL2* has a previously unappreciated, complex,
371 and autoimmune disease-associated regulatory architecture that was accurately predicted by our
372 3D epigenomic V2G approach. Importantly, we find that the distal *IL2* cRE are highly accessible
373 in quiescent memory T cell subsets (Th1, Th2, Th17, Th1-17, Th22) isolated directly *ex vivo* from
374 human blood, whereas naïve CD4+ T cells and non-IL-2-producing Treg showed little accessibility
375 at these elements (**Fig. S6**). This suggests that stable remodeling of distal *IL2* cRE can persist *in*
376 *vivo* after TCR signals cease, and that this epigenetic imprinting contributes to the immediate,
377 activation-induced production of IL-2 exhibited by memory, but not naïve or regulatory, CD4+ T
378 cells[70–72].
379
380
381 **Impact of autoimmune risk-associated genetic variation on *cis*-regulatory element activity**

382 The chromatin conformation approach used here employs GWAS variants as 'signposts' to
383 identify disease-relevant regulatory elements and connect them to their target genes, but does
384 not *per se* determine the effect of disease-associated genetic variation on enhancer activity or
385 gene expression. Expression quantitative trait loci (eQTL) studies identify variant-gene pairs
386 through statistical associations between gene expression and genotype, offering an orthogonal
387 line of evidence to compare with chromatin conformation capture approaches that identify variant-
388 gene pairs by biophysical association. We compared our variant-gene pairs to those from a recent
389 scRNA-seq-based eQTL study in activated CD4+ T cells[13] and identified 41 (out of 127) pairs
390 implicated by both methods, an overlap 10-fold higher than obtained by random sampling of
391 genes within 500 kb of autoimmune variants (**Fig. 6A, Fig S7A**). We experimentally validated
392 variant effects at the *IL2* locus, where the -128 enhancer defined above contains two SNPs linked
393 to T1D, JIA, and SLE (rs1512973 and rs12504008). Using a recombinant reporter assay in
394 primary activated CD4+ T cells, we confirmed that disease-associated genetic variation influences
395 intergenic *IL2* enhancer activity at the -128 kb element, in that the risk allele contributes
396 significantly less transcriptional activity than the reference allele (**Fig. 5H**).
397 Autoimmune variants are likely to influence disease risk by altering the activity of *cis*-
398 regulatory elements in T cells. Indeed, we identified over 1000 cRE likely impacted by
399 autoimmune disease-associated genetic variation, in that they contain autoimmune risk SNPs
400 that are predicted to decrease or increase transcription factor binding affinity. Overall, 1370
401 autoimmune risk variants in open chromatin were predicted to influence the activity of 495 DNA
402 binding factors (**Table S15**), including PLAG1, PRDM1, BACH2, MYC, TBX21, BHLHE40, LEF1,
403 TCF7, BCL6, IRF, p53, STAT (**Fig. S8A**), and hundreds of NFkB, EGR, KLF, FKH/FOX, and FOS-
404 JUN sites (**Fig. 6B**). For example, the MS-associated SNP rs1077667 is located in an element
405 that exhibits increased accessibility and contact frequency with the promoters of *GPR108* and
406 *TRIP10* upon activation. This variant was predicted to reduce affinity for TP53, TP63, and
407 OCT2/POU2F2 (**Fig. S8B**). Similarly, the T1D SNP rs3024505 (PP = 0.200) connected to the

408 promoters of *IL9* and *FAIM3* (**Fig. S8C**) was predicted to disrupt binding of *MZF1*, while the celiac
409 SNP rs13010713 (PP = 0.154) contacting the *ITGA4* promoter (**Fig. S8D**) likely disrupts a SOX4
410 binding site.
411
412 **Functional validation of V2G-implicated effector gene sets**
413 To determine whether genes identified *via* their physical interaction with autoimmune variants
414 in CD4+ T cell chromatin contact maps tend to be directly involved in T cell activation and function,
415 we compared the set of autoimmune genes implicated by chromatin contacts in this study to sets
416 of genes identified in CRISPR-based screens that control aspects of CD4+ T cell activation like
417 proliferation and expression of the inflammatory genes *IFNG, CTLA4, IL2, IL2RA,* and *TNF*[20–22].
418 The set of all V2G-implicated genes was highly enriched for genes shown to regulate IL-2, IL-2
419 receptor, CTLA-4, and proliferation (**Fig. 6C, Fig. S7B, Table S16**). For example, 202 genes
420 shown to regulate IL-2 production and 166 genes shown to regulate proliferation were also
421 implicated in our autoimmune V2G set (**Table S16**). Genes implicated by V2G in activated CD4+
422 T cells were moderately enriched for genes known to control the production of IFN-G, but at the
423 individual disease level, only genes connected to CRO- and ATD-associated variants were
424 enriched for IFNG regulatory genes (**Fig. 6C, Fig. S7B**). Genes contacting CRO, PSO, RA, SLE,
425 T1D and VIT variants were moderately enriched for TNF regulatory genes, but the set of all V2G
426 genes was not enriched for TNF genes. We also queried the orthologs of our V2G-implicated
427 genes from the international mouse phenotype consortium (IMPC) database and identified 97
428 genes that when knocked out give an immune phenotype and 126 V2G genes that result in a
429 hematopoietic phenotype (**Fig. 6D**, **Fig. S7C**, **Table S16**). This frequency of observed
430 immune/hematopoietic phenotypes represents a significant (adjP<0.05) ~30% enrichment over
431 expected. This gene set was also enriched for mortality, homeostasis/metabolism, growth/body
432 size, skeleton, and embryonic phenotypes (**Fig. 6D**, **Table S16**).
433

434 **Identification of novel regulators of T cell activation**

435 An important application of this V2G approach is the identification of novel regulators of T cell
436 activation and their potential as drug targets, as nearly 20% of implicated gene products have at
437 least one chemical modulator currently available (**Fig 6E**, **Table S17**). As shown above, this
438 approach validates genes that are well-studied regulators of T cell function, however, a significant
439 portion of implicated genes are not well-studied and are not currently known to regulate T cell
440 activation (**Fig. 6F**). We observed a trend that genes expressed more highly in immune tissues
441 (GTEX) have in general been better investigated, and identified several less-studied genes that
442 could encode novel targets, including the kinases GRK6, PTK6, SIK1, and MAP3K11, the G
443 protein-coupled receptors OXER1, GPR183, GPR18, and KISS1R, the acetylcholine receptor
444 CHRNB1, and the *de novo* purine pathway enzyme GART. To determine whether these V2G-
445 implicated genes are novel regulators of T cell activation, we used commercially available
446 pharmacologic modulators in dose-response assays of activation-induced T cell proliferation.
447 Stimulation of T cells in the presence of ligands for CHRNB1, KISS1R and OXER1 did not
448 significantly affect T cell proliferation, however, small molecules targeting GRK6, PTK6,
449 MAP3K11, GPR183, GART, and SIK1 inhibited activation of both human and mouse T cells in
450 the nanomolar to micromolar range (**Fig. 6G**). Together, these data show that maps of dynamic,
451 3D variant-gene chromatin contacts in stimulated CD4+ T cells are able to identify genes with
452 *bona fide,* evolutionarily conserved roles in T cell activation.

453

454 **DISCUSSION**

455 By measuring dynamic changes in chromosome folding, chromatin accessibility, and gene
456 expression in naïve CD4+ T cells as a function of TCR-CD28 costimulation, we identified the
457 putative *cis*-regulatory landscape of autoimmune disease-associated genetic variation, and
458 physically connected these elements to their putative effector genes. We conducted
459 pharmacologic targeting experiments and compared our results with orthogonal eQTL and

460 CRISPR-based studies to validate sets of effector genes with a confluence of evidence supporting
461 their role in CD4+ T cell activation. This study and prior studies show that most cRE and their
462 variants interact with, and therefore have the potential to regulate, more than one gene (median
463 5 in this study), and the majority of cRE/variants 'skip' the nearest gene to interact with distant
464 genes[54,73,74]. For these reasons, chromatin conformation mapping in disease-relevant cell types
465 brings a crucial added dimension to genome-scale 1D data for understanding the complex genetic
466 and epigenetic mechanisms that regulate gene expression, and for predicting novel therapeutic
467 targets.

468 Many of the genes identified in this 3D epigenomic V2G screen have known roles in T cell
469 activation and function. An example is *IL2*, and we used the resulting maps to identify and validate
470 a stretch of previously unknown distal enhancers whose activity is required for *IL2* expression and
471 is influenced by autoimmune genetic variation. Another example is the phosphatase DUSP5 that
472 regulates MAPK signaling during T cell activation[75,76]. However, roles for many of the genes
473 implicated here in T cell activation are not known. For example, one of the top implicated genes,
474 PARK7, is a deglycase studied in the context of Parkinson's disease, but has a recently been
475 shown to modulate regulatory T cell function[77]. The orphan G protein-coupled receptor GPR108
476 is another top gene in our study that has not been studied in T cells, but was identified in a recent
477 CRISPR screen for genes affecting IL-2 levels[22]. Also co-implicated by our study and recent
478 CRISPR screens are the cannabidiol receptor GPR18[78] and the purine biosynthetic enzyme
479 GART[79]. The GPR18 agonist arachidonyl glycine inhibited CD4+ T cell activation above 10 uM,
480 while the GPR18 antagonist O-1918 slightly enhanced T cell activation. Antagonism of GART, an
481 enzyme we previous identified as a V2G effector gene in COVID19 severity[56], with the FDA-
482 approved drug lometrexol inhibited T cell activation in the 10 nM range. Antagonism of GRK6, a
483 member of the G-coupled receptor kinase family associated with insulin secretion and T2D
484 susceptibility[80], and PTK6, an oncogenic kinase studied in the context of cancer[81], led to inhibition
485 of T cell activation in the nM to uM range. Inhibition of MAP3K11, a kinase that facilitates signaling

486 through JNK1 and IkappaB kinase in cancer[82,83], and 25-hydroxycholesterol, a ligand of the G
487 protein-coupled receptor GPR183, inhibited stimulation-induced CD4+ T cell proliferation in the
488 100 nM range. SIK1 is a member of the salt-inducible kinase family that negatively regulates
489 TORC activity[59], and a small molecule SIK1 inhibitor potently antagonized stimulation-induced
490 CD4+ T cell activation in the pM range.

491 Our integration of high-resolution Hi-C, ATAC-seq, RNA-seq and GWAS data in a single
492 immune cell type across multiple activation states identified hundreds of autoimmune variant-
493 gene pairs at ~half of all GWAS loci studied, and application of this technique to additional immune
494 cell types will likely identify effector genes at many of the remaining loci. This study highlights the
495 value of chromosome conformation data as a powerful biological constraint for focusing variant-
496 to-gene mapping efforts[84], and shows that dynamic changes in the spatial conformation of the
497 genome that accompany cell state transitions alter gene expression by exposing promoters to a
498 varying array of *cis*-regulatory elements, transcription factors, and genetic variants.

499

500 **MATERIALS AND METHODS**

501

502 <u>T cell isolation and *in vitro* stimulation</u>

503 Human primary CD4+ T cells were purified from the apheresis products obtained from healthy,
504 screened human donors through University of Pennsylvania Human Immunology Core (HIC).
505 Naïve CD4+ T cells were purified using EasySepTM Human naïve CD4+ T cell isolation kit II
506 (STEM cells Technologies, cat#17555) by immunomagnetic negative selection as per
507 manufacturer's protocol. Isolated untouched, highly purified (93-98%) naïve human CD4 T cells
508 were activated using anti-CD3+ anti-CD28 Dynabeads (1:1) (Thermofisher scientific, cat #
509 11161D) for 8-24 hours. Cells were then used to prepare sequencing libraries. The human
510 leukemic T cell line Jurkat was obtained from ATCC, cloned by limiting dilution, and clones with
511 high activation-induced secretion of IL-2 were selected for further study.

512

513 RNA-seq library generation and sequencing

514 RNA was isolated from ~ 1 million of each cell stage using Trizol Re- agent (Invitrogen), purified
515 using the Directzol RNA Miniprep Kit (Zymo Research), and depleted of contaminating genomic
516 DNA using DNAse I. Purified RNA was checked for quality on a Bioanlayzer 2100 using the Nano
517 RNA Chip and samples with RIN>7 were used for RNA-seq library preparation. RNA samples
518 were depleted of rRNA using QIAseq Fastselect RNA removal kit (Qiagen). Samples were then
519 processed for the preparation of libraries using the SMARTer Stranded Total RNA Sample Prep
520 Kit (Takara Bio USA) according to the manufacturer's instructions. Briefly, the purified first-strand
521 cDNA is amplified into RNA-seq libraries using SeqAmp DNA Polymerase and the Forward and
522 the Reverse PCR Primers from the Illumina Indexing Primer Set HT for Illumina. Quality and
523 quantity of the libraries was assessed using the Agilent 2100 Bioanalyzer system and Qubit
524 fluorometer (Life Technologies). Sequencing was performed on the NovaSeq 6000 platform at
525 the CHOP Center for Spatial and Functional Genomics.

526

527 ATAC-seq library generation and sequencing

528 A total of 50,000 to 100,000 cells were centrifuged at 550 g for 5 min at 4 °C. The cell pellet was
529 washed with cold PBS and resuspended in 50 µL cold lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM
530 NaCl, 3 mM MgCl2, 0.1% NP-40/IGEPAL CA-630) and immediately centrifuged at 550 g for
531 10 min at 4 °C. Nuclei were resuspended in the Nextera transposition reaction mix (25 µL 2x TD
532 Buffer, 2.5 µL Nextera Tn5 transposase (Illumina Cat #FC-121-1030), and 22.5 µL nuclease free
533 H2O) on ice, then incubated for 45 min at 37 °C. The tagmented DNA was then purified using the
534 Qiagen MinElute kit eluted with 10.5 µL Elution Buffer (EB). Ten microliters of purified tagmented
535 DNA was PCR amplified using Nextera primers for 12 cycles to generate each library. PCR
536 reaction was subsequently cleaned up using 1.5x AMPureXP beads (Agencourt), and

537 concentrations were measured by Qubit. Libraries were paired-end sequenced on the Illumina
538 HiSeq 4000 platform (100 bp read length).
539
540 Hi-C library preparation
541 Hi-C library preparation was performed using the Arima-HiC kit (Arima Genomics Inc), according
542 to the manufacturer's protocols. Briefly, cells were crosslinked using formaldehyde. Crosslinked
543 cells were then subject to the Arima-HiC protocol, which utilizes multiple restriction enzymes to
544 digest chromatin. Arima-HiC sequencing libraries were prepared by first shearing purified
545 proximally-ligated DNA and then size-selecting 200-600 bp DNA fragments using AmpureXP
546 beads (Beckman Coulter). The size-selected fragments were then enriched using Enrichment
547 Beads (provided in the Arima-HiC kit), and then converted into Illumina-compatible sequencing
548 libraries with the Swift Accel-NGS 2SPlus DNA Library Kit (Swift, 21024) and Swift 2S Indexing
549 Kit (Swift, 26148). The purified, PCR-amplified DNA underwent standard QC (qPCR, Bioanalyzer,
550 and KAPA Library Quantification [Roche, KK4824]) and was sequenced with unique single
551 indexes on the Illumina NovaSeq 6000 Sequencing System using 200 bp reads.
552
553 ATAC-seq data analysis
554 ATAC-seq peaks were called using the ENCODE ATAC-seq pipeline
555 (https://www.encodeproject.org/atac-seq/). Briefly, pair-end reads from three biological replicates
556 for each cell type were aligned to hg19 genome using bowtie2[85], and duplicate reads were
557 removed from the alignment. Narrow peaks were called independently for each replicate using
558 macs2[86] (-p 0.01 --nomodel --shift -75 --extsize 150 -B --SPMR --keep-dup all --call-summits) and
559 ENCODE blacklist regions (ENCSR636HFF) were removed from peaks in individual replicates.
560 Reproducible peaks, peaks called in at least two replicates, were used to generate a consensus
561 peakset. Signal peaks were normalized using csaw[87] in 10kb bins background regions and low
562 abundance peaks (CPM>1) were excluded from the analysis. Tests for differential accessibility

563 was conducted with the glmQLFit approach implemented in edgeR[88] using the normalization
564 factors calculated by csaw. OCRs with FDR < 0.05 and abs(log_2FC > 1) between stages were
565 considered differentially accessible.
566
567 Hi-C data analysis
568 Paired-end reads from two replicates were pre-processed using the HICUP pipeline v0.7.4[89], with
569 bowtie as aligner and hg19 as the reference genome. The alignment .bam file were parsed to
570 .pairs format using pairtools v0.3.0 (https://github.com/open2c/pairtools) and pairix v0.3.7
571 (https://github.com/4dn-dcic/pairix), and eventually converted to pre-binned Hi-C matrix in .cool
572 format by cooler v0.8.10[90] with multiple resolutions (500bp, 1kbp, 2kbp, 2.5kbp, 4kbp, 5kbp,
573 10kbp, 25kbp, 40kbp, 50kbp, 100kbp, 250kbp, 500kbp, 1Mbp and 2.5Mbp) and normalized with
574 ICE method[91]. Replicate similarity was determined by HiCRep v1.12.2[90] at 10K resolution. For
575 each sample, eigenvectors were determined from an ICE balanced Hi-C matrix with 40kb
576 resolution using cooltools v0.3.2[92] and first principal components were used to determine A/B
577 compartments with GC% of genome region as reference track to determine the sign. Differential
578 TAD comparison was performed using TADcompare with the default settings for each
579 chromosome (v1.4.0)[93]. Finally, for each cell type, significant intra-chromosomal interaction loops
580 were determined under multiple resolutions (1kb, 2kb and 4kb) using the Hi-C loop caller Fit-Hi-
581 C2 v2.0.7[94] (FDR<1e-6) on merged replicates matrix. The consensus chromatin loops within
582 resolution were identified by combining all three stages. These sets of loops were used as
583 consensus for quantitative differential analysis explained below. The final consensus interaction
584 loops for visualization were collected by merging loops from all the resolutions with preference to
585 keep the highest resolution. Quantitative loop differential analysis across cell types was performed
586 on fast lasso normalized interaction frequency (IF) implemented in multiCompareHiC v1.8.0[95] for
587 each chromosome at resolution 1kb, 2kb and 4kb independently. The contacts with zero
588 interaction frequency (IF) among more than 80% of the samples and average IF less than 5 were

589 excluded from differential analysis. The QLF test based on a generalized linear model was
590 performed in cell type-pairwise comparisons, and p-values were corrected with FDR. The final
591 differential loops were identified by overlapping differential IF contacts with consensus interaction
592 loops.
593
594 Bulk RNA-seq data analysis
595 Bulk RNA-seq libraries were sequenced on an Illumina Novaseq 6000 instrument. The pair-end
596 fastq files were mapped to the genome assembly hg19 by STAR (v2.6.0c) independently for each
597 replicate. The GencodeV19 annotation was used for gene feature annotation and the raw read
598 count for gene feature was calculated by htseq-count (v0.6.1)[96] with parameter settings -f bam -r
599 pos -s yes -t exon -m union. The gene features localized on chrM or annotated as rRNAs, small
600 coding RNA, or pseudo genes were removed from the final sample-by-gene read count matrix.
601 Gene set variation analysis was performed using GSVA 1.42.0[97] with the MSigDB hallmark
602 geneset[98], with resulting scores analyzed using limma (limma_3.50.3)[99]. Low abundance peaks
603 (CPM>1) were excluded from the analysis. Testing for differential expression was conducted with
604 the glmQLFit approach implemented in edgeR[88]. Genes with FDR<0.05 and abs($\log_2$FC>1)
605 between stages were considered differentially expressed. Differential genes were then clustered
606 using k-means clustering. The number of clusters was determined using the elbow method on
607 the weighted sum of squares, where was set to k=5. Score for how similar each gene followed
608 the clusters expression pattern was determined by calculating pearson correlation coefficients
609 between each gene in the cluster and the cluster centroid.
610
611 Transcription factor footprinting and motif analysis
612 Transcription factor footprints were called using Regulatory Analysis Toolbox HINT-ATAC
613 (v0.13.0) with pooled ATAC-seq data for each stage and consensus peak calls[100]. The rgt-hint
614 footprinting was run with parameters –atac-seq, --paired-end, and organism=hg19 set. The output

615 footprint coordinates were subsequently matched using rgt-motifanalysis matching with
616 parameters --organism hg19 and –pseudocount 0.8 set. The JASPAR2020 position weight matrix
617 database was used to match footprints[101]. Differential analysis of TF binding across conditions
618 was performed using rgt-hint differential with parameters –organism hg19, --bc, --nc 24 using the
619 motif matched transcription factor footprints. An activity score is then calculated based on the
620 accessibility surrounding the footprint.

621

622 Partitioned heritability LD score regression enrichment analysis

623

624 Partitioned heritability LD score regression[102] (v1.0.0) was used to identify heritability enrichment
625 with GWAS summary statistics and open chromatin regions annotated to genes. The baseline
626 analysis was performed using LDSCORE data
627 (https://data.broadinstitute.org/alkesgroup/LDSCORE) with LD scores, regression weights, and
628 allele frequencies from the 1000G phase 1 data. The summary statistics were obtained from
629 studies as described in the supplemental table and harmonized with the munge_sumstats.py
630 script. Annotations for Partitioned LD score regression were generated using the coordinates of
631 open chromatin regions that contact gene promoters through Hi-C loops for each cell type. Finally,
632 partitioned LD scores were compared to baseline LD scores to measure enrichment fold change
633 and enrichment p-values, which were adjusted with FDR across all comparisons.

634

635 Variant-to-gene mapping using HiC-derived promoter contacts

636 95% credible sets were determined as previously described[103]. Briefly, P values from GWAS
637 summary statistics were converted to Bayes factors. Posterior probabilities were then calculated
638 for each variant. The variants were ordered from highest to lowest posterior probabilities added
639 to the credible set until the cumulative sum of posterior probabilities reached 95%. Variants in the
640 95% credible set were then intersected with the CD4+ T cell promoter interacting region OCRs
641 from the three timepoints using the R package GenomicRanges (v1.46.1)[104].

642

643 Genomic reference and visualizations

644 All analyses were performed using the hg19 reference genome using gencodeV19 as the gene
645 reference. Genomic tracks were visualized with pyGenomeTracks v3.5[105]. HiC matrices depict
646 the log1p(balanced count) from the cooler count matrix. ATAC-seq tracks were generated from
647 bigwig files that were normalized using deeptools[106].

648

649 Colocalized eQTL comparisons

650 Enrichment of sentinel-gene assignments was conducted similarly as described previously[54].
651 Briefly, a null distribution was constructed by randomly selecting genes within 1 Mb of the sentinel
652 compared to the set of colocalized cis-eQTL[13] found in through 10,000 iterations. The observed
653 overlap reports the set of gene identified by both our HiC based approach with the set of
654 colocalized eQTLs. We report the empirical P value of the observed value relative to null
655 distribution.

656

657 International mouse phenotyping consortium comparisons

658 The set of HiC implicated genes were compared to the mouse international phenotyping
659 consortium set of genes with reported phenotypes[107]. We converted the list of V2G implicated
660 genes to mouse homologs using homologene. We tested for enrichment for each phenotype using
661 a one-sided proportion test implicated in R prop.test with type set to “upper”.

662

663 Identification of pharmacological agents

664 We queried the Drug-Gene Interaction Database with the set of V2G implicated genes for
665 chemicals using rDGIdb (v1.20.0)[108]. To identify the number of papers for each gene with at least
666 one drug annotated to target it, we queried pubmed titles and abstracts using the R package
667 RISmed (v2.3.0) with each gene’s name and either “autoimmune” and the list of autoimmune

668 diseases (**Table S17**). A score to approximate expression specificity was computed using the
669 sum GTEX median expression values (v8) for whole blood or spleen divided by other tissues[109].
670
671 Lentiviral-based CRISPR/CAS9 targeting in Jurkat cells
672 LentiCRISPRv2-mCherry vectors encoding gRNA-CAS9 and the fluorescent reporter mCherry
673 were used for Jurkat targeting. CRISPR guide RNAs (sgRNA) targeting human IL-2-21 intergenic
674 -46, -51, -80, -83, -85, -122, -128 CNS regions were designed using http://crispr.tefor.net and
675 cloned into lentiCRISPRv2-mCherry. Empty vector without gRNA insert was used as a control.
676 Below is the list of CRISPR gRNA for causing deletion of human IL2-21 intergenic regions:
677 CNS region CRISPR gRNA
678 -46 CNS gRNA #1 AGGATGCCTACCTCCAAATG
679 gRNA # 2 AGGTGACAACATTTAGTCAG
680 -51 CNS gRNA #1 GGCAACGAAATTCACTGTGA
681 gRNA # 2 ATTCTAACAGGAATCATTCG
682 -80 CNS gRNA # 1 GTTCTACCTATGCCGCATTG
683 gRNA # 2 GAGATTTACTCAGTCCAATG
684 -83 CNS gRNA # 1 GTGACAAGCATGACTCTACA
685 gRNA # 2 GTGATGGTGAATTAAGCTGA
686 -85 CNS gRNA #1 AGGGTTTTCTAGTTACGAGA
687 gRNA # 2 ATGGTTAGTTAGCTCCCAAG
688 -96 CNS gRNA#1 TGGGAAAAACATCTTACCTG
689 gRNA # 2 TGGCCCATGAACCATCAAAG
690 -122 CNS gRNA# 1 GTTATTAATCTAAGCGGAGA
691 gRNA# 2 GGAAGTTAGGCAGTCAATCG
692 -128 CNS gRNA#1 CTTCAATCATTGCATTCCAC
693 gRNA# 2 TGACACCACCCCTGCTTGAG

694 HEK 293T cells were grown in RPMI1640 complete medium (RPMI1640 + 1X P/S, 1x L-Glu, 10%
695 FBS), 37C, 7%CO2. 293T cells were transfected with 10 ug of lenti-CRISPR-V2-CRE construct
696 along with packaging plasmid 6 ug of PsPAX2 (Addgene, Cat #12260) and 3.5 ug of PmD2.G
697 (Addgene, Cat #12259) using Lipofectamine 2000 transfection reagent (Invitrogen cat #
698 11668019). After 6 hours, the transfection medium was replaced with complete culture medium.
699 Transfected cells were incubated at 37C for 48-72 hours in a cell culture incubator. and then the
700 Lentiviral supernatants were harvested and spun at 300g for 5 minutes to remove cellular debris.
701 Lentiviral supernatants were concentrated using Lenti-X™ Concentrator (Takara Bio, Cat #
702 631232) and then centrifuged at 1500g for 30 minutes at 4°C and supernatant was discarded.
703 The lentiviral pellet was resuspended at a ratio of 1:20 of the original volume using RPMI media
704 and concentrated virus supernatant aliquots were prepared and stored until use at -80°C. To
705 achieve high transduction efficiency, the viral supernatant was titrated in Jurkat cells through
706 transduction using various dilutions of the viral supernatants and transduction efficiency was
707 determined by mcherry expression analyzed through flow cytometry. Jurkat cells were seeded in
708 a 24 well plate at 0.5 x106/well in culture media, viral supernatant with 8 ug/mL of polybrene was
709 added to each well. Spinfection was performed for 90 min. at 2500 rpm, and transduced cells
710 were equilibrated at 37C for ~6 hrs, followed by incubation at 37C 5%CO2 for ~72 hours culture.
711 Transduced Jurkat Cells were then harvested and stimulated using PMA (15 ng/mL) + Ionomycin
712 (1 uM) + human anti-CD28 (2 ug/ml), BioXcell cat # BE0248 in 96 well culture plates (TPP, cat #
713 92097) in triplicates. Cell culture supernatants were collected at the end of culture and analyzed
714 for IL-2 by ELISA using a kit (cat # 88-7025-76) purchased from Thermofisher Scientific.
715
716 <u>gRNA-CAS9 RNP-based targeting in primary human CD4+ T cells</u>
717 Primary Human CD4 T cells derived from 5 normal healthy donors were obtained from Human
718 Immunology core (University of Pennsylvania). Alt-R S.p. HiFi Cas9 Nuclease V3 cat # 1081061

719 CAS9 and following list of Alt-R® CRISPR-Cas9 sgRNA targeting IL-2-21 CRE were purchased

720 from Integrated DNA Technologies, USA.

Name	20nt gRNA Seq
46CNS-1	GACGTATATAGTCATCTGAT
46CNS-2	TCTGTGGAGCTGCTGCGTTA
46CNS-3	AGGATGCCTACCTCCAAATG
51CNS-1	ATTCTAACAGGAATCATTCG
51CNS-2	GAGTAAAAAGACGTGTTACC
51CNS-3	GGCAACGAAATTCACTGTGA
80CNS-1	AATGCGGCATAGGTAGAACT
80CNS-2	GAGATTTACTCAGTCCAATG
80CNS-3	TGGTTGTCACAGTAACTAGG
83CNS-1	GTGACAAGCATGACTCTACA
83CNS-2	GTGATGGTGAATTAAGCTGA
83CNS-3	TGGGCTCTGACTCACTTAGA
85CNS-1	GCTAACATTGACTTCTCTAC
85CNS-2	AATCTATGCAAGGGGTGAAT
85CNS-3	GCATCATGATGAAGCTTATC
122CNS-1	GTTATTAATCTAAGCGGAGA
122CNS-2	GGAAGTTAGGCAGTCAATCG
122CNS-3	CACTTTGTGCTCGGATGCTC
128CNS-1	CTTCAATCATTGCATTCCAC
128CNS-2	TCAAGCAGGGGTGGTGTCAA
128CNS-3	TGGTGATTCATCTTTAGCAT
IL-2URR-1	TCCATTCAGTCAGTCTTTGG

IL-2URR-2	TAGTGTCCCAGGTGATTTAG
IL-2URR-3	TAGAGCTATCACCTAAGTGT
B2m-1	GAGTAGCGCGAGCACAGCTA
B2m-2	CACGCGTTTAATATAAGTGG

721

722 Primary Human CD4 T (5-10e6) were incubated with sgRNA and CAS9 protein complex and
723 electroporation was done using P3 Primary Cell 4D-Nucleofector™ X Kit L (Lonza, cat # V4XP-
724 3024) and Lonza 4D nucleofector system. B2M gene was CRISPR targeted as a positive control.
725 As per manufacturer's protocol cells were electroporated pulse code Fl-115 in 100ul cuvette
726 format. After nucleofection, cells were allowed to rest in the complete media for ~2 days. Cells
727 were then harvested, washed with PBS and aliquots of cells were used for further experimentation
728 such as flow staining and cell activation. Primary human CD4 T Cells (0.1e6/well) were seeded
729 in triplicates (for each experimental condition) in 96 well plate format in RPMI complete medium
730 and stimulated using ImmunoCult™ Human CD3/CD28 T Cell Activator (STEM cells
731 Technologies, cat # 10971). Cell culture supernatants were collected at 4h of stimulation and
732 stored at -80C until assayed. IL-2 ELISA was performed using Thermofisher Scientific IL-2 Human
733 Uncoated ELISA Kit with Plates, cat # 88-7025-76.

734

735 Recombinant reporter assays

736 The *IL2* URR was cloned into Xho I and Hind III sites of PGL4 Luc vector (Promega) and then
737 ~500 bp individual sequence representing *IL2-IL21* intergenic CNS at -46, -51, -80, -83, -85, -128
738 were cloned upstream to IL-2 URR at the Xho I site of the pGL4 vector. Primary human CD4 T
739 cells obtained from 5 normal healthy donors were transfected with PGL4-cRE-URR constructs.
740 Briefly, primary human CD4 T cells were activated with anti-CD3+ anti-CD28 Dynabeads (1:1)
741 (Thermofisher scientific, cat # 11161D) + IL-2 overnight and then 1 million aliquots of cells
742 (triplicates) were electroporated using Nucleofector 2b, human T cell Nuclefactor kit (Lonza VPA

743 # 1002, program # T-020) with 2ug of PGL4 firefly vector constructs along with 0.2 ug of PGL4
744 Renilla vector; cells were allowed to rest ON in RPMI + IL-2 and then re-stimulated with plate-
745 bound anti-CD3+ anti-CD28 (2 ug/ml each) for 5 hrs. Dual luciferase assay was performed with
746 cell lysate prepared using Promega dual luciferase assay kit. Cell lysate was prepared in PLB as
747 per manufacturer's protocols and then firefly and renilla luciferase activities were analyzed by
748 spectramax ID5 (Molecular Devices). Firefly luciferase activity was normalized against the internal
749 control renilla luciferase activity.
750
751 Pharmacologic validation of novel T cell activation regulatory genes
752 The drugs lometrexol (GART antagonist), 25-hydroxycholesterol (GPR183 ligand), epinephrine &
753 norepinephrine (CHRNB agonists), and arachidonoyl glycine (GPR18 agonist) were purchased
754 from Sigma. The GPR18 antagonist O-1918 was purchased from ChemCruz. CEP1347
755 (MAP3K11 antagonist), kisspeptin 234 (KISSIR ligand) and the PTK6 antagonist tilfrinib were
756 purchased from Tocris. The GRK6 antagonist GRK-IN-1 was purchased from DC Chemicals. The
757 SIK1 antagonist HG-91-9-1 was purchased from Selleckchem. The OXER1 agonist 5-Oxo-ete
758 was purchased from Cayman chemicals. The drugs were dissolved in DMSO or ethanol as
759 suggested by the vendor and working stocking concentrations were prepared in RPMI1640
760 medium or PBS. The drug effect on T cell proliferation was assayed using murine lymphocytes
761 cultured under TCR and CD28 activation conditions. Spleen and lymph nodes were collected from
762 6-8 weeks old female C57BL/7 mice and single cell lymphocyte cell suspensions were prepared
763 in RPMI 1640 complete medium. 20 million lymphocytes were labeled with CFSE and re-
764 suspended in RPMI 1640 medium. Cells (0.5 million labeled cells/well) were loaded into 48 well
765 culture plate and activated with mouse anti-CD3 and anti-CD28 agonistic antibodies (1μg mL
766 each). Drugs at the indicated concentrations were added in the culture medium and both
767 untreated and drug treated cell cultures were incubated at 37°C for 72 hours in a cell culture

768 incubator. Cells were harvested after 3 days of culture, washed with PBS and then stained with
769 live-dead aqua dye. After washing with FACS buffer (PBS containing 2% FBS), cells were stained
770 with fluorochrome conjugated antibodies CD4-APC, CD8-PB, CD44-Percp Cy5.5 and CD25-
771 BV650. Stained cells were analyzed on a Beckman Coulter Cytoflex S flow cytometer. The
772 division profile of CD4+ CFSE+ T cells were gated on live populations. The flow data was
773 analyzed using Flowjo10 software and the number of divided CD4+ T cells were determined as
774 described previously[110].
775
776 **FIGURE LEGENDS**
777 **Figure 1: Defined gene expression dynamics throughout activation of naïve CD4+ T cells.**
778 (**A**) Diagram of study design: RNA-seq, ATAC-seq, and HiC libraries were prepared from sorted
779 CD4+CD62L^{hi}CD44^{lo} naïve T cells (donor N=3) prior to or following 8hr or 24hr stimulation with
780 anti-CD3/28 beads. cREs are classified to genes allowing for interrogation of the pattern of
781 expression, accessibility, and chromatin structure changes of autoimmune-associated variants.
782 (**B**) Heatmap showing normalized expression of known markers of T cell activation. (**C**) Gene set
783 variance analysis (GSVA, N=3) using log-transformed expression values to MSigDB Hallmark
784 gene set. Red indicates higher activity and blue indicates lower activity. (**D**) K-means clustering
785 of differentially expressed genes into five groups using the elbow and within-cluster-sum of
786 squares methods to select the number of clusters. The centroid of the cluster is depicted as a
787 black line, and members of the cluster are depicted as colored lines. Color indicates the Pearson's
788 correlation coefficient between the gene and the cluster centroid, with red indicating higher
789 correlation, and blue lower correlation. (**E**) Gene set enrichment for each cluster of differential
790 genes using the REACTOME pathway database. Enriched pathways associated with differentially
791 expressed genes in those regions. X-axis indicates ratio of observed pathway members
792 compared to those expected by chance, the Y-axis indicates the -log10(p-value) of enrichment

793 (hypergeometric test), and dot size corresponds to the absolute number of genes observed in the
794 pathway.
795
796 **Figure 2: Chromatin accessibility and architecture changes across three stages of T cell**
797 **activation.** (**A**) Volcano plots depicting the $\log_2$ change in accessibility in all reproducible OCRs
798 (present in 2 of the 3 replicates) compared to the $-\log_{10}$(FDR), significant points are indicated in
799 red (FDR < 0.05; abs($\log_2$FC) > 1). (**B**) Genomic regions were classified by membership in A vs.
800 B compartments at each time point. The pie chart depicts regions with differential A/B
801 compartment assignment in response to activation. Significantly enriched pathways of genes
802 located in variable A/B compartments are shown as bubble plots. Color indicates the pattern of
803 A/B compartment assignment across the three timepoints, and node size indicates the number of
804 genes in that pathway. (**C**) TAD structure was determined at each time point, and regions
805 exhibiting a shift, split/merge, change in strength, or more complex rearrangements are depicted
806 in the pie chart. Enriched pathways associated with differentially expressed genes in those
807 regions are shown. (**D**) Loop calls identified from HiC data for each timepoint called at three
808 resolutions (1kb, 2kb, 4kb bins). (**E**) Differential loop-calls called across all resolutions. (**F**) Density
809 plots of cRE accessibility and gene expression separated by whether contact frequency
810 increased, decreased, or remained stable from the transition from unstimulated to 8 hour
811 activation and 8 hour to 24 hour activation.
812
813 **Figure 3: Transcription factors prediction potential regulators of chromatin status and**
814 **expression changes**. Footprints were annotated to each TF motif by sequence matching to the
815 JASPAR motif database, and the average accessibility of the region surrounding each motif was
816 used as a measure of activity at each time point. Scatterplots depict the change in accessibility
817 for each TF motif (activity score) and $\log_2$FC of TF gene expression between unstimulated and 8
818 hour activation (**A**) or 8 hour and 24 hour activation (**B**). TF were classified as differentially

819 expressed (orange), differentially active (brown), both (red), or neither (grey). Dot size indicates
820 the number of predicted footprints occupied by each motif. (**C**) Average accessibility (normalized
821 by depth and motif number) in a 200bp window surrounding motif footprints for FOSL2::JUNB,
822 Smad2::Smad3, MAZ, and KLF2 from three timepoints (unstim: green; 8hr: blue; 24hr: red). (**D**)
823 Z-score of TF motif enrichment for cRE connected to the five clusters compared to all OCR.
824 Lighter color indicates higher specificity of enrichment for that TF cluster. (**E**) Connections
825 between different TFs based on physical interactions and predicted regulation-based co-
826 expression determined by GENIE3 for the three timepoints. Node color indicates expression
827 (TPM; darker = higher expression, lighter = lower expression), edge color reflects confidence in
828 the interaction called by GENIE3 (darker higher confidence).
829
830 **Figure 4: Variant-to-gene mapping of autoimmune-associated loci implicates genetic**
831 **variants in the control of T cell activation**. (**A**) Identification of genes in contact with open
832 chromatin regions harboring autoimmune-associated SNPs. Sentinel SNPs from 15 autoimmune
833 GWAS (**Table S11**) were used to identify the 95% credible set of proxy SNPs for each lead SNP.
834 SNP locations were integrated with ATAC-seq and HiC data to identify the 95% credible set of
835 accessible SNPs in physical contact with a gene promoter. Genes are further refined based on
836 expression dynamics over the time course of T cell activation. (**B**) Partitioned LD score regression
837 for autoimmune GWAS studies using the OCR in contact with the genes in the five clusters
838 defined by RNA-seq. Circle size = enrichment, color = significance, **FDR<0.01, ***FDR<0.001.
839 (**C**) Log distribution of the 1D distance between each proxy SNP and its interacting gene based
840 on 3D chromatin conformation (median=103 kbp). (**D**) Distribution of the number of genes
841 contacted by each accessible variant (median=5). (**E**) Fraction of open disease-associated
842 variants that interact with no gene promoters (grey), only with the nearest gene promoter (purple),
843 with the nearest gene and a distant gene(s) (orange), and with only a distant gene(s) (blue). (**F**)
844 Number of genes identified by variant-to-gene mapping at each time point. Black = shared in all

845 stages, blue = shared in two stages, red = specific to one stage. (**G**) Set of implicated proxy-genes
846 pairs that are both differentially expressed, display differential accessibility, and chromatin contact
847 across T cell activation. (**H**) Example MS variant rs1077667 (PP=1.0) exhibits increased
848 accessibility and contact with the promoters of the *GPR108* and *TRIP10* at 8 hours post activation,
849 which are increased in expression at this time point. (**I**) Example allergy variant rs7380290
850 interacts with the *SIK1* gene that is upregulated after activation.
851
852 **Figure 5: Functional validation of autoimmune V2G-implicated cRE at the *IL2* locus.** (**A**) The
853 combination of evolutionary conservation (blue), open chromatin (red), and autoimmune disease-
854 associated SNPs at the *IL2* locus identify putative cRE in quiescent vs. 8-hour activated naïve
855 CD4+ T cells. (**B**) Activation-dependent TAD/sub-TAD structure (heatmaps), chromatin
856 remodeling (grey) and promoter-OCR interactions (red) at the *IL2* locus. (**C**) Recombinant reporter
857 assay showing transcriptional activity of the *IL2* URR (+35 to -500 from the TSS) in activated
858 primary CD4+ T cells (N=5 donors) compared to a basal promoter (pGL4, left panel). The right
859 panel depicts transcriptional activity of the CNS regions indicated in **A** cloned upstream of the
860 URR. All regions except the -80 CNS show statistically significant activity relative to the URR
861 alone (N=5, P<0.05, line = median, box = 95/5% range). (**D**) Scheme of CRISPR/CAS9-based
862 deletion of individual *IL2* CNS using flanking gRNAs. (**E**) Activation-induced secretion of IL-2
863 protein by CRISPR-targeted primary CD4+ T cells (N=5 donors, left panel) or Jurkat cells (N=3
864 replicates, right panel) relative to untargeted control (CAS9, no gRNA) cells (****P<0.0001,
865 ***P<0.001, **P<0.01, *P<0.05, line = median, box = 95/5% range). In primary cells, *B2M* gRNAs
866 served as an irrelevant targeted control. (**F**) Scheme of CRISPR/CAS9-based deletion of the 81.3
867 kb region containing all distal *IL2* cRE using flanking gRNAs in Jurkat cells. (**G**) Activation-induced
868 IL-2 protein (left panel) and mRNA (right panel) by control (black) vs. 81.3 kb deleted (red) Jurkat
869 cells (N=3 separate clones). (**H**) Recombinant reporter assay in activated primary CD4+ T cells

870 (N=5 donors) showing transcriptional activity of the reference vs. risk alleles of the *IL2* -128 cRE
871 relative to the URR alone (P<0.0001).
872
873 **Figure 6: Functional validation of autoimmune V2G-implicated genes.** (**A**) Physical variant-
874 gene associations (HiC) are highly enriched for statistical variant-gene associations (eQTL) in
875 activated CD4+ T cells. The histogram depicts the null distribution of shared variant-gene pairs
876 expected at random (~5) while the red line indicates the observed number of variant-gene pairs
877 (41) shared with the 127 eQTL identified by Soskic et al. in a similar CD4+ T cell activation system.
878 (**B**) Prominent TF motifs predicted to be disrupted (blue) or stabilized (red) by promoter-connected
879 autoimmune SNPs. (**C**) 3D chromatin-based V2G genes are enriched for CRISPR-implicated
880 genes that regulate CD4+ T cell activation. Observed enrichment of genes regulating multiple
881 aspects of CD4+ T cell activation (IL-2, IL-2 receptor, CTLA-4, IFNg, TNFa or proliferation) from
882 CRISPR screens by Freimer, Schmidt, and Shifrut among sets of 3D chromatin-implicated genes
883 among individual diseases (green, FDR<0.05) or all diseases (purple, FDR<0.05). (**D**) Enrichment
884 for 3D chromatin-based autoimmune V2G genes among genes with germline knock-out mouse
885 immune (red) and other (black) phenotypes (adjP<0.05, IMPC database). (**E**) Autoimmune V2G-
886 implicated genes with at least one pharmacologic modulator (rDGIdb). (**F**) Comparison of the
887 number of manuscripts retrieved from PubMed related to autoimmune disease for each V2G gene
888 with pharmaceutical agents available (x-axis) with an immune-specific expression score
889 computed using the sum GTEX median expression values (v8) for whole blood or spleen divided
890 by other tissues (y-axis). Genes highlighted in red were selected for functional validation in **G**. (**G**)
891 Dose-dependent impact of the indicated pharmacologic agents targeting the V2G-implicated
892 genes *KISS1R, CHRNB, OXER1, GPR18, GRK6, PTK6, MAP3K11, GPR183, GART,* and *SIK1*
893 on proliferation of anti-CD3/28 activated murine (left) and human (right) CD4+ T cells *in vitro*
894 (N=4).
895

896 **Supplemental Tables**

897 Table S1: Expression and clustering of differentially expressed genes.

898 Table S2: Pathway enrichment of different RNA-seq genes

899 Table S3: Accessible chromatin regions

900 Table S4: Differential accessible open chromatin regions

901 Table S5: Differential contact frequency

902 Table S6: Differential TAD boundaries

903 Table S7: A/B compartment calls

904 Table S8: Stripe calls

905 Table S9: Summary of TF footprinting

906 Table S10: TF target gene pathway enrichment

907 Table S11: List of all GWAS studies included

908 Table S12: Variant to gene mapping across all timepoints

909 Table S13: Variant to gene mapping promoter OCRs

910 Table S14: Partitioned LD score regression output

911 Table S15: V2G genes implicated by orthogonal data

912 Table S16: Motifs predicted to disrupt transcription factor binding sites.

913 Table S17: V2G target gene drug repurposing results

914

915 **REFERENCES**

916 1. Zhang, F. & Lupski, J. R. Non-coding genetic variants in human disease. *Hum Mol Genet* **24**,
917 R102-10 (2015).
918 2. Weinreich, M. A. *et al.* KLF2 Transcription-Factor Deficiency in T Cells Results in Unrestrained
919 Cytokine Production and Upregulation of Bystander Chemokine Receptors. *Immunity* **31**,
920 122–130 (2009).
921 3. Tzachanis, D. *et al.* Tob is a negative regulator of activation that is expressed in anergic and
922 quiescent T cells. *Nat Immunol* **2**, 1174–82 (2001).
923 4. Yusuf, I. & Fruman, D. A. Regulation of quiescence in lymphocytes. *Trends Immunol* **24**, 380–
924 386 (2003).
925 5. Zhu, J., Yamane, H. & Paul, W. E. Differentiation of Effector CD4 T Cell Populations. *Annu Rev*
926 *Immunol* **28**, 445–489 (2010).

927 6. Saravia, J., Chapman, N. M. & Chi, H. Helper T cell differentiation. *Cell Mol Immunol* **16**, 634–
928 643 (2019).
929 7. Farh, K. K.-H. *et al.* Genetic and epigenetic fine mapping of causal autoimmune disease
930 variants. *Nature* **518**, 337 343 (2014).
931 8. Chen, L. *et al.* Genetic Drivers of Epigenetic and Transcriptional Variation in Human Immune
932 Cells. *Cell* **167**, 1398 1414.e24 (2016).
933 9. Soskic, B. *et al.* Chromatin activity at GWAS loci identifies T cell states driving complex immune
934 diseases. *Nat Genet* 1–8 (2019) doi:10.1038/s41588-019-0493-9.
935 10. Robertson, C. C. *et al.* Fine-mapping, trans-ancestral and genomic analyses identify causal
936 variants, cells, genes and drug targets for type 1 diabetes. *Nat Genet* **53**, 962–971 (2021).
937 11. Lu, X. *et al.* Global discovery of lupus genetic risk variant allelic enhancer activity. *Nat*
938 *Commun* **12**, 1611 (2021).
939 12. Chandra, V. *et al.* Promoter-interacting expression quantitative trait loci are enriched for
940 functional genetic variants. *Nat Genet* **53**, 110–119 (2021).
941 13. Soskic, B. *et al.* Immune disease risk variants regulate gene expression dynamics during
942 CD4+ T cell activation. *Nat Genet* 1–10 (2022) doi:10.1038/s41588-022-01066-3.
943 14. Mouri, K. *et al.* Prioritization of autoimmune disease-associated genetic variants that perturb
944 regulatory element activity in T cells. *Nat Genet* 1–10 (2022) doi:10.1038/s41588-022-
945 01056-5.
946 15. Rioux, J. D. & Abbas, A. K. Paths to understanding the genetic basis of autoimmune disease.
947 *Nature* **435**, 584–589 (2005).
948 16. Buckner, J. H. Mechanisms of impaired regulation by CD4(+)CD25(+)FOXP3(+) regulatory T
949 cells in human autoimmune diseases. *Nat Rev Immunol* **10**, 849 859 (2010).
950 17. Burren, O. S. *et al.* Chromosome contacts in activated T cells identify autoimmune disease
951 candidate genes. *Genome Biol* **18**, 165 (2017).
952 18. Yang, J. *et al.* Analysis of chromatin organization and gene expression in T cells identifies
953 functional genes for rheumatoid arthritis. *Nat Commun* **11**, 4402 (2020).
954 19. Bykova, M., Hou, Y., Eng, C. & Cheng, F. Quantitative trait locus (xQTL) approaches identify
955 risk genes and drug targets from human non-coding genomes. *Hum Mol Genet* **31**, R105–
956 R113 (2022).
957 20. Shifrut, E. *et al.* Genome-wide CRISPR Screens in Primary Human T Cells Reveal Key
958 Regulators of Immune Function. *Cell* **175**, 1958-1971.e15 (2018).
959 21. Schmidt, R. *et al.* CRISPR activation and interference screens decode stimulation responses
960 in primary human T cells. *Science* **375**, eabj4008 (2022).
961 22. Freimer, J. W. *et al.* Systematic discovery and perturbation of regulatory genes in human T
962 cells reveals the architecture of immune networks. *Nat Genet* **54**, 1133–1144 (2022).
963 23. Thurman, R. E. *et al.* The accessible chromatin landscape of the human genome. *Nature* **489**,
964 75 (2012).
965 24. Lieberman-Aiden, E. *et al.* Comprehensive Mapping of Long-Range Interactions Reveals
966 Folding Principles of the Human Genome. *Science* **326**, 289–293 (2009).
967 25. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of
968 chromatin interactions. *Nature* **485**, 376 380 (2012).
969 26. Mirny, L. A., Imakaev, M. & Abdennur, N. Two major mechanisms of chromosome
970 organization. *Curr Opin Cell Biol* **58**, 142–152 (2019).
971 27. Yoon, S., Chandra, A. & Vahedi, G. Stripenn detects architectural stripes from chromatin
972 conformation data using computer vision. *Nat Commun* **13**, 1602 (2022).
973 28. Kim, T.-K. & Shiekhattar, R. Architectural and Functional Commonalities between Enhancers
974 and Promoters. *Cell* **162**, 948 959 (2015).
975 29. Hamamoto, K. & Fukaya, T. Molecular architecture of enhancer–promoter interaction. *Curr*
976 *Opin Cell Biol* **74**, 62–70 (2022).

30. Consortium, R. E. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317 330 (2015).
31. Yoon, J.-H. *et al.* Phosphorylation status determines the opposing functions of Smad2/Smad3 as STAT3 cofactors in TH17 differentiation. *Nat Commun* **6**, 7600 (2015).
32. Safford, M. *et al.* Egr-2 and Egr-3 are negative regulators of T cell activation. *Nat Immunol* **6**, 472–480 (2005).
33. Schubert, L. A., Jeffery, E., Zhang, Y., Ramsdell, F. & Ziegler, S. F. Scurfin (FOXP3) Acts as a Repressor of Transcription and Regulates T Cell Activation. *J Biol Chem* **276**, 37672–37679 (2001).
34. Xiao, T., Li, X. & Felsenfeld, G. The Myc-associated zinc finger protein (MAZ) works together with CTCF to control cohesin positioning and genome organization. *Proc National Acad Sci* **118**, e2023127118 (2021).
35. Sebzda, E., Zou, Z., Lee, J. S., Wang, T. & Kahn, M. L. Transcription factor KLF2 regulates the migration of naive T cells by restricting chemokine receptor expression patterns. *Nat Immunol* **9**, 292–300 (2008).
36. Moskowitz, D. M. *et al.* Epigenomics of human CD8 T cell differentiation and aging. *Sci Immunol* **2**, eaag0192 (2017).
37. Harris, J. E. *et al.* Early Growth Response Gene-2, a Zinc-Finger Transcription Factor, Is Required for Full Induction of Clonal Anergy in CD4+ T Cells. *J Immunol* **173**, 7331–7338 (2004).
38. Williams, J. B. *et al.* The EGR2 targets LAG-3 and 4-1BB describe and regulate dysfunctional antigen-specific CD8+ T cells in the tumor microenvironment. *J Exp Med* **214**, 381–400 (2017).
39. Avitahl, N. *et al.* Ikaros Sets Thresholds for T Cell Activation and Regulates Chromosome Propagation. *Immunity* **10**, 333–343 (1999).
40. Bandyopadhyay, S. *et al.* Interleukin 2 gene transcription is regulated by Ikaros-induced changes in histone acetylation in anergic T cells. *Blood* **109**, 2878–86 (2007).
41. Thomas, R. M. *et al.* Ikaros Enforces the Costimulatory Requirement for IL2 Gene Expression and Is Required for Anergy Induction in CD4+ T Lymphocytes. *J Immunol* **179**, 7305–7315 (2007).
42. Thomas, R. M. *et al.* Ikaros Silences T-bet Expression and Interferon-γ Production during T Helper 2 Differentiation. *J Biol Chem* **285**, 2545–2553 (2010).
43. Weber, B. N. *et al.* A critical role for TCF-1 in T-lineage specification and differentiation. *Nature* **476**, 63–68 (2011).
44. Johnson, J. L. *et al.* Lineage-Determining Transcription Factor TCF-1 Initiates the Epigenetic Identity of T Cells. *Immunity* **48**, 243-257.e10 (2018).
45. Wang, W. *et al.* TCF-1 promotes chromatin interactions across topologically associating domains in T cell progenitors. *Nat Immunol* 1–11 (2022) doi:10.1038/s41590-022-01232-z.
46. Furukawa, T. *et al.* Involvement of PLAGL2 in activation of iron deficient- and hypoxia-induced gene expression in mouse cell lines. *Oncogene* **20**, 4718–4727 (2001).
47. Wu, Q. *et al.* Hypoxia-inducible factors: master regulators of hypoxic tumor immune escape. *J Hematol Oncol* **15**, 77 (2022).
48. Fu, W. *et al.* A multiply redundant genetic switch “locks in” the transcriptional signature of regulatory T cells. *Nat Immunol* **13**, 972–980 (2012).
49. Pramanik, J. *et al.* Genome-wide analyses reveal the IRE1a-XBP1 pathway promotes T helper cell differentiation by resolving secretory stress and accelerating proliferation. *Genome Med* **10**, 76 (2018).
50. Gao, X., Tewari, K., Svaren, J. & Suresh, M. Role of cell cycle regulator E2F1 in regulating CD8 T cell responses during acute and chronic viral infection. *Virology* **324**, 567–576 (2004).
51. DeRyckere, D. & DeGregori, J. E2F1 and E2F2 Are Differentially Required for Homeostasis-Driven and Antigen-Induced T Cell Proliferation In Vivo. *J Immunol* **175**, 647–655 (2005).

1028 52. Cook, M. E., Jarjour, N. N., Lin, C.-C. & Edelson, B. T. Transcription Factor Bhlhe40 in
1029 Immunity and Autoimmunity. *Trends Immunol* **41**, 1023–1036 (2020).
1030 53. Chesi, A. *et al.* Genome-scale Capture C promoter interactions implicate effector genes at
1031 GWAS loci for bone mineral density. *Nat Commun* **10**, 1260 (2019).
1032 54. Su, C. *et al.* Mapping effector genes at lupus GWAS loci using promoter Capture-C in follicular
1033 helper T cells. *Nat Commun* **11**, 3294 (2020).
1034 55. Pahl, M. C. *et al.* Cis-regulatory architecture of human ESC-derived hypothalamic neuron
1035 differentiation aids in variant-to-gene mapping of relevant complex traits. *Nat Commun* **12**,
1036 6749 (2021).
1037 56. Pahl, M. C. *et al.* Implicating effector genes at COVID-19 GWAS loci using promoter-focused
1038 Capture-C in disease-relevant immune cell types. *Genome Biol* **23**, 125 (2022).
1039 57. Su, C. *et al.* 3D chromatin maps of the human pancreas reveal lineage-specific regulatory
1040 architecture of T2D risk. *Cell Metab* **34**, 1394-1409.e4 (2022).
1041 58. Feng, Y. *et al.* The Cdc42-interacting Protein-4 (CIP4) Gene Knock-out Mouse Reveals
1042 Delayed and Decreased Endocytosis*. *J Biol Chem* **285**, 4348–4354 (2010).
1043 59. Darling, N. J. & Cohen, P. Nuts and bolts of the salt-inducible kinases (SIKs). *Biochem J* **478**,
1044 1377–1397 (2021).
1045 60. Cerosaletti, K. *et al.* Multiple autoimmune-associated variants confer decreased IL-2R
1046 signaling in CD4+ CD25(hi) T cells of type 1 diabetic and multiple sclerosis patients. *Plos*
1047 *One* **8**, e83811 (2013).
1048 61. Spolski, R., Li, P. & Leonard, W. J. Biology and regulation of IL-2: from molecular mechanisms
1049 to human therapy. *Nat Rev Immunol* **18**, 648 659 (2018).
1050 62. Abbas, A. K., Trotta, E., Simeonov, D. R., Marson, A. & Bluestone, J. A. Revisiting IL-2:
1051 Biology and therapeutic prospects. *Sci Immunol* **3**, eaat1482 (2018).
1052 63. Joosse, M. E. *et al.* Duplication of the IL2RA locus causes excessive IL-2 signaling and may
1053 predispose to very early onset colitis. *Mucosal Immunol* 1–11 (2021) doi:10.1038/s41385-
1054 021-00423-5.
1055 64. Fujita, T., Shibuya, H., Ohashi, T., Yamanishi, K. & Taniguchi, T. Regulation of human
1056 interleukin-2 gene: Functional DNA sequences in the 5′ flanking region for the gene
1057 expression in activated T lymphocytes. *Cell* **46**, 401–407 (1986).
1058 65. Durand, D. B., Bush, M. R., Morgan, J. G., Weiss, A. & Crabtree, G. R. A 275 basepair
1059 fragment at the 5' end of the interleukin 2 gene enhances expression from a heterologous
1060 promoter in response to signals from the T cell antigen receptor. *J Exp Medicine* **165**, 395
1061 407 (1987).
1062 66. Novak, T. J., White, P. M. & Rothenberg, E. V. Regulatory anatomy of the murine interleukin-
1063 2 gene. *Nucleic Acids Res* **18**, 4523–4533 (1990).
1064 67. Mehra, P. & Wells, A. D. Long-Range Transcriptional Control of the Il2 Gene by an Intergenic
1065 Enhancer. *Mol Cell Biol* **35**, 3880–3891 (2015).
1066 68. Consortium, T. E. P. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* **306**,
1067 636–640 (2004).
1068 69. Lin, X. *et al.* Nested epistasis enhancer networks for robust genome regulation. *Science* **377**,
1069 1077–1085 (2022).
1070 70. Hedlund, G., Dohlsten, M., Ericsson, P. O. & Sjögren, H. O. Rapid response to Con A by
1071 CD4+CD45R− rat memory lymphocytes as compared to CD4+CD45R+ lymphocytes. *Cell*
1072 *Immunol* **119**, 317–326 (1989).
1073 71. Picker, L. *et al.* Direct demonstration of cytokine synthesis heterogeneity among human
1074 memory/effector T cells by flow cytometry. *Blood* **86**, 1408–1419 (1995).
1075 72. Sojka, D. K., Bruniquel, D., Schwartz, R. H. & Singh, N. J. IL-2 Secretion by CD4+ T Cells In
1076 Vivo Is Rapid, Transient, and Influenced by TCR-Specific Competition. *J Immunol* **172**,
1077 6136–6143 (2004).

1078 73. Xia, Q. *et al.* The type 2 diabetes presumed causal variant within TCF7L2 resides in an
1079 element that controls the expression of ACSL5. *Diabetologia* **59**, 2360–2368 (2016).
1080 74. Sobreira, D. R. *et al.* Extensive pleiotropism and allelic heterogeneity mediate metabolic
1081 effects of IRX3 and IRX5. *Science* **372**, 1085–1091 (2021).
1082 75. Kovanen, P. E. *et al.* T-cell development and function are modulated by dual specificity
1083 phosphatase DUSP5. *J Biological Chem* **283**, 17362–9 (2008).
1084 76. Chuang, H.-C. & Tan, T.-H. MAP4K Family Kinases and DUSP Family Phosphatases in T-
1085 Cell Signaling and Systemic Lupus Erythematosus. *Cells* **8**, 1433 (2019).
1086 77. Danileviciute, E. *et al.* PARK7/DJ-1 promotes pyruvate dehydrogenase activity and maintains
1087 Treg homeostasis during ageing. *Nat Metabolism* **4**, 589–607 (2022).
1088 78. McHugh, D. *et al.* N-arachidonoyl glycine, an abundant endogenous lipid, potently drives
1089 directed cellular migration through GPR18, the putative abnormal cannabidiol receptor. *Bmc*
1090 *Neurosci* **11**, 44 (2010).
1091 79. Kan, J. L., Jannatipour, M., Taylor, S. M. & Moran, R. G. Mouse cDNAs encoding a
1092 trifunctional protein of de novo purine synthesis and a related single-domain glycinamide
1093 ribonucleotide synthetase. *Gene* **137**, 195–202 (1993).
1094 80. Varney, M. J. *et al.* G protein–coupled receptor kinase 6 (GRK6) regulates insulin processing
1095 and secretion via effects on proinsulin conversion to insulin. *J Biol Chem* **298**, 102421
1096 (2022).
1097 81. Gilic, M. B. & Tyner, A. L. Targeting protein tyrosine kinase 6 in cancer. *Biochimica Et*
1098 *Biophysica Acta Bba - Rev Cancer* **1874**, 188432 (2020).
1099 82. Slattery, M. L., Lundgreen, A. & Wolff, R. K. MAP kinase genes and colon and rectal cancer.
1100 *Carcinogenesis* **33**, 2398–2408 (2012).
1101 83. Knackmuss, U. *et al.* MAP3K11 is a tumor suppressor targeted by the oncomiR miR-125b in
1102 early B cells. *Cell Death Differ* **23**, 242–252 (2016).
1103 84. Hammond, R. K. *et al.* Biological constraints on GWAS SNPs at suggestive significance
1104 thresholds reveal additional BMI loci. *Elife* **10**, e62206 (2021).
1105 85. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**,
1106 357–359 (2012).
1107 86. Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).
1108 87. Lun, A. T. L. & Smyth, G. K. csaw: a Bioconductor package for differential binding analysis of
1109 ChIP-seq data using sliding windows. *Nucleic Acids Res* **44**, e45–e45 (2016).
1110 88. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for
1111 differential expression analysis of digital gene expression data. *Bioinform Oxf Engl* **26**, 139–
1112 40 (2009).
1113 89. Wingett, S. *et al.* HiCUP: pipeline for mapping and processing Hi-C data. *F1000research* **4**,
1114 1310 (2015).
1115 90. Abdennur, N. & Mirny, L. A. Cooler: scalable storage for Hi-C data and other genomically
1116 labeled arrays. *Bioinform Oxf Engl* **36**, 311–316 (2019).
1117 91. Imakaev, M. *et al.* Iterative correction of Hi-C data reveals hallmarks of chromosome
1118 organization. *Nat Methods* **9**, 999 1003 (2012).
1119 92. Nora, E. P. *et al.* Molecular basis of CTCF binding polarity in genome folding. *Nat Commun*
1120 **11**, 5612 (2020).
1121 93. Cresswell, K. G. & Dozmorov, M. G. TADCompare: An R Package for Differential and
1122 Temporal Analysis of Topologically Associated Domains. *Frontiers Genetics* **11**, 158 (2020).
1123 94. Kaul, A., Bhattacharyya, S. & Ay, F. Identifying statistically significant chromatin contacts from
1124 Hi-C data with FitHiC2. *Nat Protoc* **15**, 991–1012 (2020).
1125 95. Stansfield, J. C., Cresswell, K. G., Vladimirov, V. I. & Dozmorov, M. G. HiCcompare: an R-
1126 package for joint normalization and comparison of HI-C datasets. *Bmc Bioinformatics* **19**,
1127 279 (2018).

96. Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinform Oxf Engl* **31**, 166–9 (2014).
97. Hänzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-Seq data. *Bmc Bioinformatics* **14**, 7–7 (2013).
98. Liberzon, A. *et al.* The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Syst* **1**, 417–425 (2015).
99. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47–e47 (2015).
100. Li, Z. *et al.* Identification of transcription factor binding sites using ATAC-seq. *Genome Biol* **20**, 45 (2019).
101. Fornes, O. *et al.* JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* **48**, D87–D92 (2020).
102. Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet* **47**, 1228–35 (2015).
103. Su, C. *et al.* 3D promoter architecture re-organization during iPSC-derived neuronal cell differentiation implicates target genes for neurodevelopmental disorders. *Prog Neurobiol* 102000 (2021) doi:10.1016/j.pneurobio.2021.102000.
104. Lawrence, M. *et al.* Software for Computing and Annotating Genomic Ranges. *Plos Comput Biol* **9**, e1003118 (2013).
105. Lopez-Delisle, L. *et al.* pyGenomeTracks: reproducible plots for multivariate genomic datasets. *Bioinform Oxf Engl* **37**, 422–423 (2020).
106. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**, W187–W191 (2014).
107. Muñoz-Fuentes, V. *et al.* The International Mouse Phenotyping Consortium (IMPC): a functional catalogue of the mammalian genome that informs conservation. *Conservation Genetics Print* **19**, 995–1005 (2017).
108. Wagner, A. H. *et al.* DGIdb 2.0: mining clinically relevant drug–gene interactions. *Nucleic Acids Res* **44**, D1036–D1044 (2016).
109. Consortium, T. Gte. *et al.* The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).
110. Wells, A. D., Gudmundsdottir, H. & Turka, L. A. Following the fate of individual T cells throughout activation and clonal expansion. Signals from T cell receptor and CD28 differentially regulate the induction and duration of a proliferative response. *J Clin Invest* **100**, 3173 3183 (1997).

Figure 1

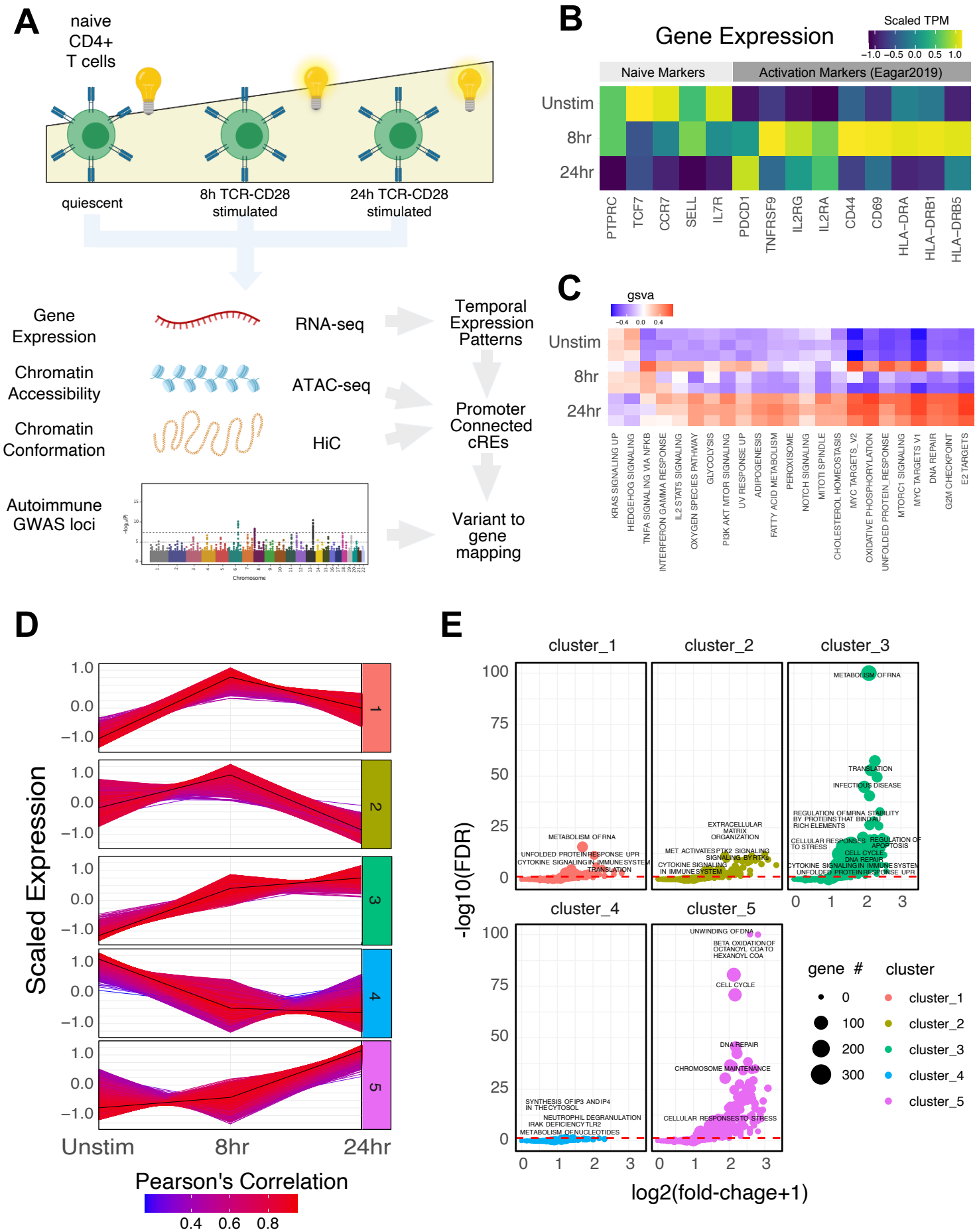
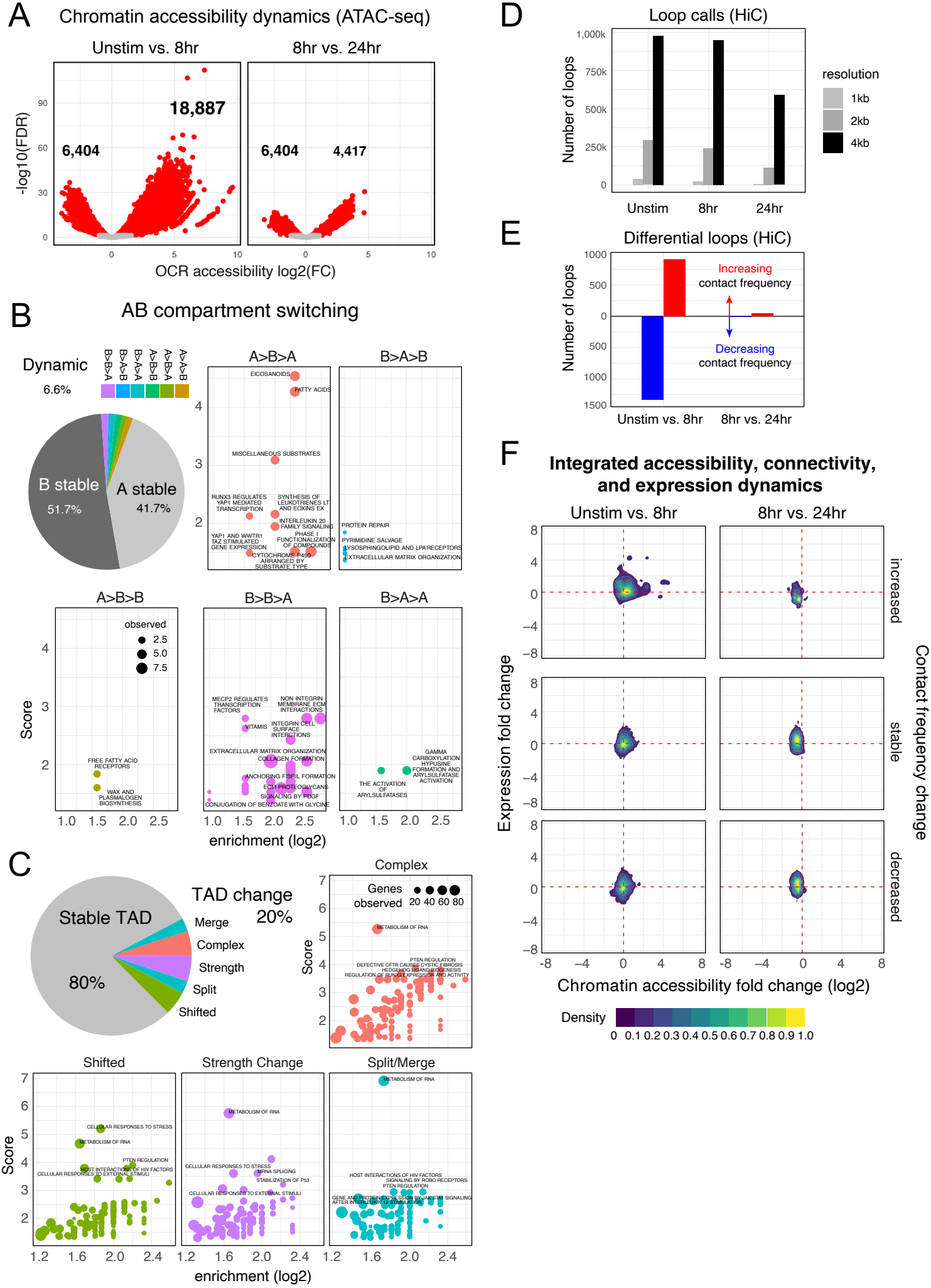

A
naive CD4+ T cells
quiescent
8h TCR-CD28 stimulated
24h TCR-CD28 stimulated
Gene Expression
RNA-seq
Chromatin Accessibility
ATAC-seq
Chromatin Conformation
HiC
Autoimmune GWAS loci
Chromosome
Temporal Expression Patterns
Promoter Connected cREs
Variant to gene mapping
B
Gene Expression
Scaled TPM
−1.0 −0.5 0.0 0.5 1.0
Naive Markers
Activation Markers (Eagar2019)
Unstim
8hr
24hr
PTPRC
TCF7
CCR7
SELL
IL7R
PDCD1
TNFRSF9
IL2RG
IL2RA
CD44
CD69
HLA−DRA
HLA−DRB1
HLA−DRB5
C
gsva
−0.4 0.0 0.4
Unstim
8hr
24hr
KRAS SIGNALING UP
HEDGEHOG SIGNALING
TNFA SIGNALING VIA NFKB
INTERFERON GAMMA RESPONSE
IL2 STAT5 SIGNALING
OXYGEN SPECIES PATHWAY
GLYCOLYSIS
PI3K AKT MTOR SIGNALING
UV RESPONSE UP
ADIPOGENESIS
FATTY ACID METABOLISM
PEROXISOME
NOTCH SIGNALING
MITOTI SPINDLE
CHOLESTEROL HOMEOSTASIS
MYC TARGETS_V2
OXIDATIVE PHOSPHORYLATION
UNFOLDED PROTEIN_RESPONSE
MTORC1 SIGNALING
MYC TARGETS V1
DNA REPAIR
G2M CHECKPOINT
E2 TARGETS
D
Scaled Expression
1.0
0.0
−1.0
1
2
3
4
5
Unstim
8hr
24hr
Pearson's Correlation
0.4 0.6 0.8
E
cluster_1
cluster_2
cluster_3
cluster_4
cluster_5
-log10(FDR)
100
75
50
25
0
METABOLISM OF RNA
UNFOLDED PROTEIN RESPONSE UPR
CYTOKINE SIGNALING IN IMMUNE SYSTEM
TRANSLATION
EXTRACELLULAR MATRIX ORGANIZATION
MET ACTIVATES PTK2 SIGNALING
CYTOKINE SIGNALING IN IMMUNE SYSTEM
TRANSLATION
INFECTIOUS DISEASE
REGULATION OF MRNA STABILITY BY PROTEINS THAT BIND AU RICH ELEMENTS
CELLULAR RESPONSES TO STRESS
REGULATION OF APOPTOSIS
CELL CYCLE
DNA REPAIR
CYTOKINE SIGNALING IN IMMUNE SYSTEM
UNFOLDED PROTEIN RESPONSE UPR
UNWINDING OF DNA
BETA OXIDATION OF OCTANOYL COA TO HEXANOYL COA
CELL CYCLE
DNA REPAIR
CHROMOSOME MAINTENANCE
CELLULAR RESPONSES TO STRESS
SYNTHESIS OF IP3 AND IP4 IN THE CYTOSOL
NEUTROPHIL DEGRANULATION
IRAK DEFICIENCY TLR2
METABOLISM OF NUCLEOTIDES
0 1 2 3
log2(fold-chage+1)
gene #
0
100
200
300
cluster
cluster_1
cluster_2
cluster_3
cluster_4
cluster_5

Figure 2



A
Chromatin accessibility dynamics (ATAC-seq)
Unstim vs. 8hr
8hr vs. 24hr
-log10(FDR)
18,887
6,404
6,404
4,417
OCR accessibility log2(FC)
B
AB compartment switching
Dynamic
6.6%
B>B>A
B>A>B
B>A>A
A>B>B
A>B>A
A>A>B
B stable
51.7%
A stable
41.7%
A>B>A
B>A>B
A>B>B
B>B>A
B>A>A
observed
2.5
5.0
7.5
Score
enrichment (log2)
C
Stable TAD
80%
TAD change
20%
Merge
Complex
Strength
Split
Shifted
Complex
Genes observed
20 40 60 80
Score
Shifted
Strength Change
Split/Merge
Score
enrichment (log2)
D
Loop calls (HiC)
Number of loops
resolution
1kb
2kb
4kb
Unstim
8hr
24hr
E
Differential loops (HiC)
Number of loops
Increasing contact frequency
Decreasing contact frequency
Unstim vs. 8hr
8hr vs. 24hr
F
Integrated accessibility, connectivity, and expression dynamics
Unstim vs. 8hr
8hr vs. 24hr
increased
stable
decreased
Expression fold change
Contact frequency change
Chromatin accessibility fold change (log2)
Density
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0

Figure 3

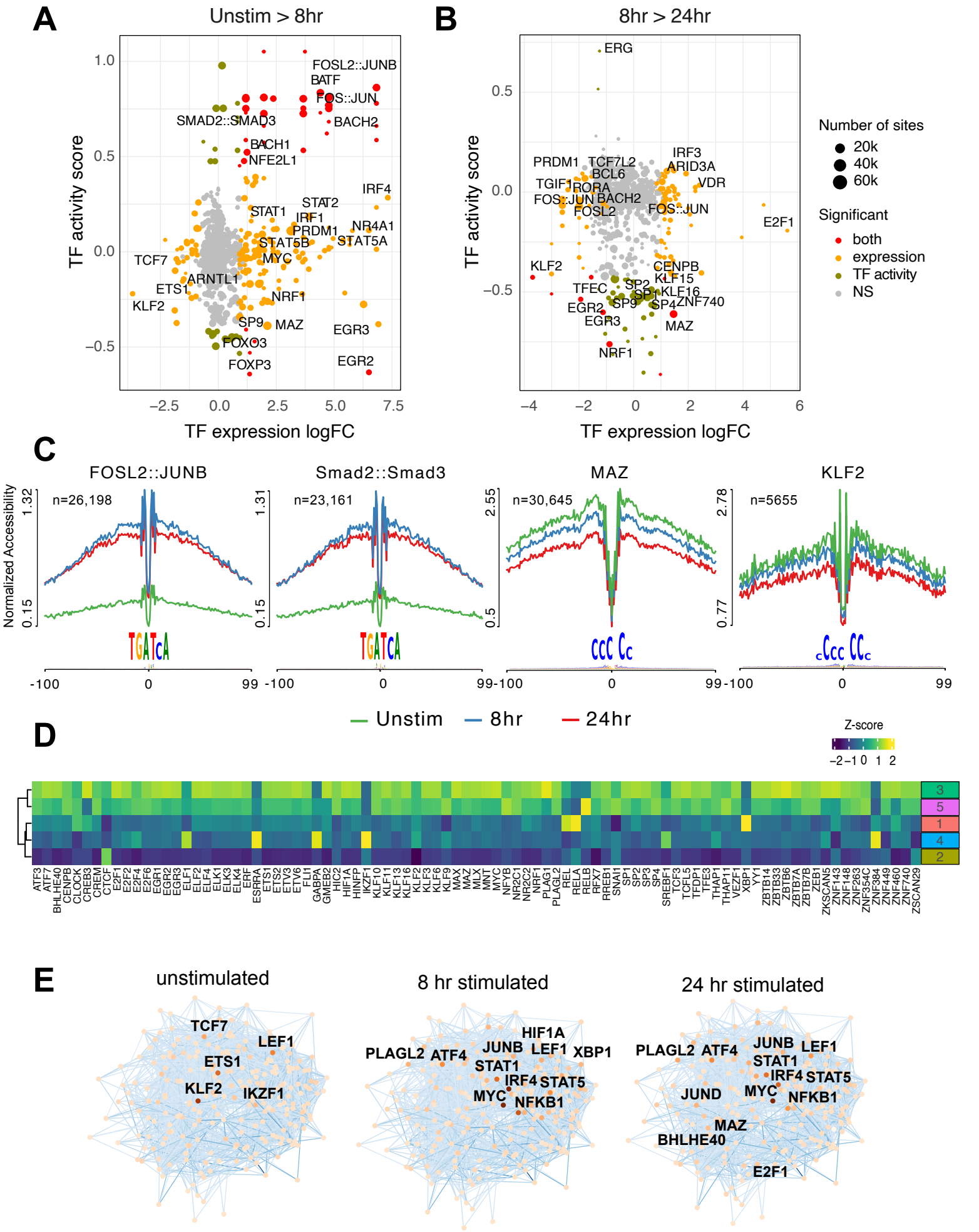

A
Unstim > 8hr
TF activity score
TF expression logFC
FOSL2::JUNB
BATF
FOS::JUN
BACH2
SMAD2::SMAD3
BACH1
NFE2L1
IRF4
STAT2
STAT1
IRF1
PRDM1
NR4A1
STAT5B
STAT5A
TCF7
MYC
ARNTL1
ETS1
NRF1
KLF2
SP9
MAZ
EGR3
FOXO3
FOXP3
EGR2
B
8hr > 24hr
ERG
PRDM1
TCF7L2
BCL6
IRF3
ARID3A
VDR
TGIF1
RORA
FOS::JUN
BACH2
FOSL2
FOS::JUN
E2F1
KLF2
CENPB
KLF15
TFEC
SP2
SP1
KLF16
ZNF740
EGR2
SP9
SP4
EGR3
MAZ
NRF1
Number of sites
20k
40k
60k
Significant
both
expression
TF activity
NS
C
FOSL2::JUNB
Smad2::Smad3
MAZ
KLF2
Normalized Accessibility
n=26,198
n=23,161
n=30,645
n=5655
Unstim
8hr
24hr
D
Z-score
E
unstimulated
8 hr stimulated
24 hr stimulated
TCF7
LEF1
ETS1
KLF2
IKZF1
HIF1A
PLAGL2
ATF4
JUNB
LEF1
XBP1
STAT1
IRF4
STAT5
MYC
NFKB1
JUND
MAZ
BHLHE40
E2F1

Figure 4

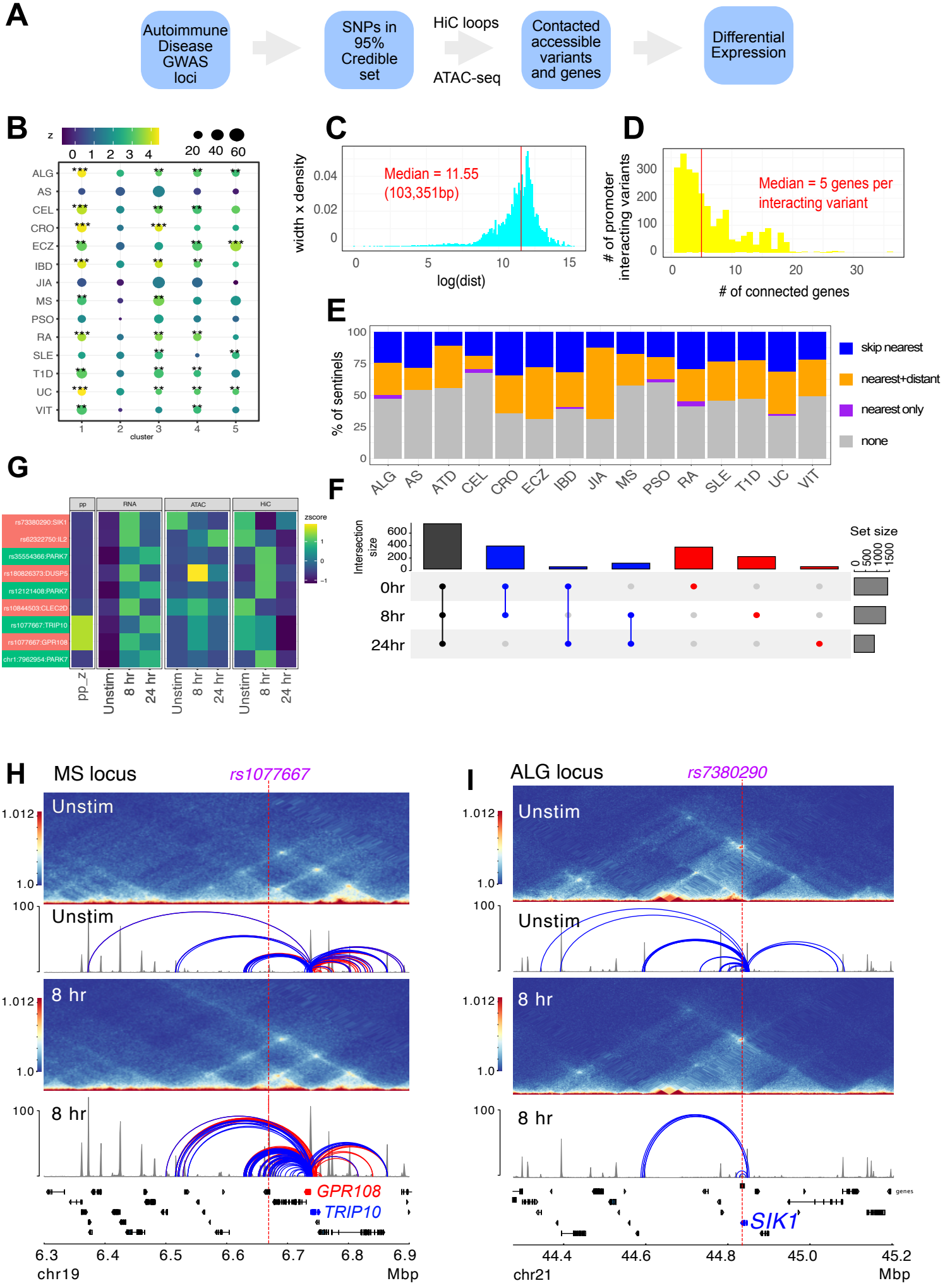

A
Autoimmune Disease GWAS loci
SNPs in 95% Credible set
HiC loops
ATAC-seq
Contacted accessible variants and genes
Differential Expression
B
z
0 1 2 3 4
20 40 60
ALG
AS
CEL
CRO
ECZ
IBD
JIA
MS
PSO
RA
SLE
T1D
UC
VIT
1 2 3 4 5
cluster
C
width x density
Median = 11.55 (103,351bp)
log(dist)
D
# of promoter interacting variants
Median = 5 genes per interacting variant
# of connected genes
E
% of sentinels
skip nearest
nearest+distant
nearest only
none
ALG AS ATD CEL CRO ECZ IBD JIA MS PSO RA SLE T1D UC VIT
F
Intersection size
Set size
0hr
8hr
24hr
G
pp
RNA
ATAC
HiC
zscore
rs73380290:SIK1
rs62322750:IL2
rs35554366:PARK7
rs180826373:DUSP5
rs12121408:PARK7
rs10844503:CLEC2D
rs1077667:TRIP10
rs1077667:GPR108
chr1:7962954:PARK7
pp_z
Unstim
8 hr
24 hr
H
MS locus
rs1077667
Unstim
8 hr
GPR108
TRIP10
chr19
Mbp
I
ALG locus
rs7380290
Unstim
8 hr
SIK1
chr21
Mbp

Figure 5

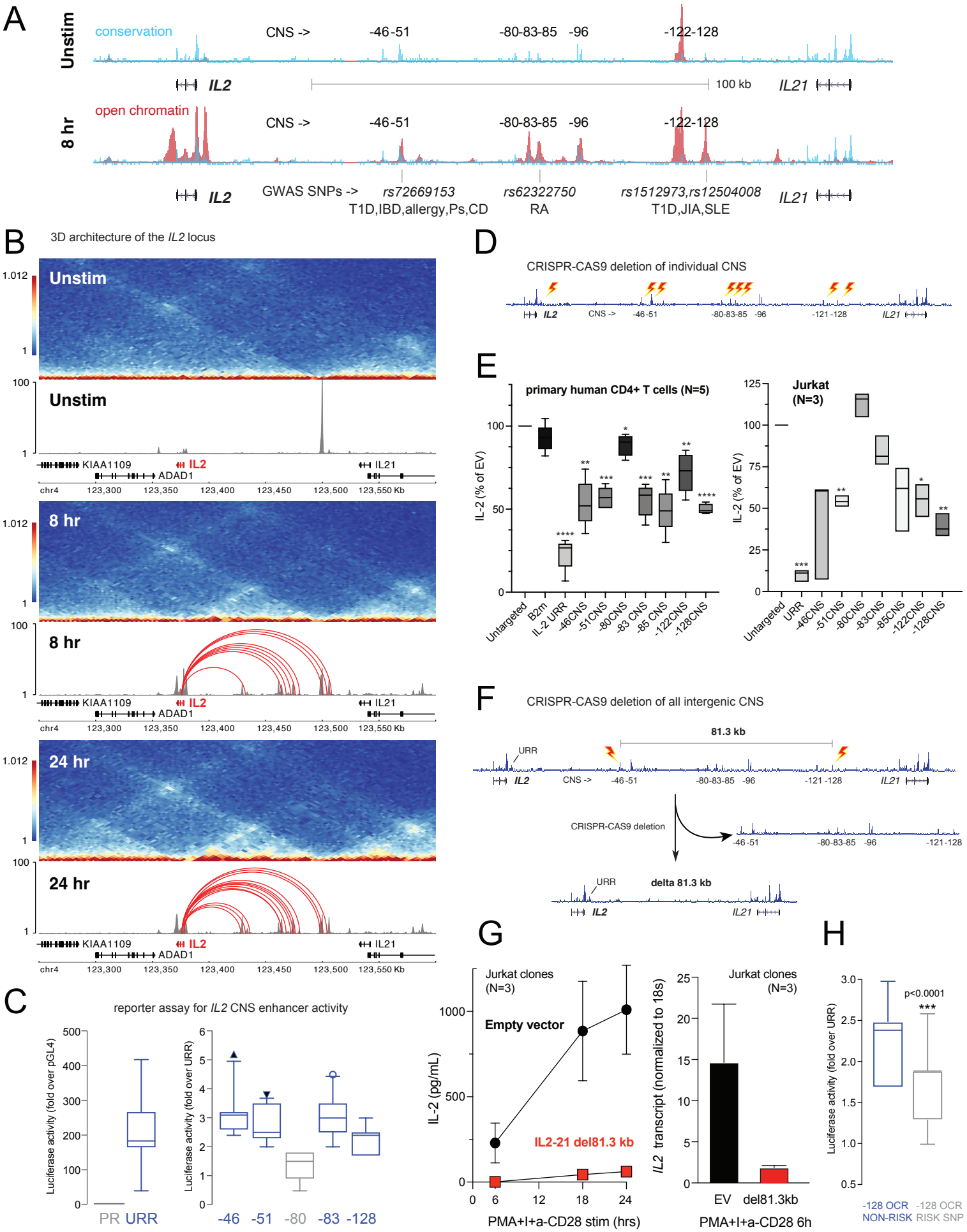

A
Unstim
conservation
CNS ->
-46-51
-80-83-85
-96
-122-128
IL2
100 kb
IL21
8 hr
open chromatin
GWAS SNPs ->
rs72669153
T1D,IBD,allergy,Ps,CD
rs62322750
RA
rs1512973,rs12504008
T1D,JIA,SLE
B
3D architecture of the IL2 locus
Unstim
8 hr
24 hr
KIAA1109
ADAD1
IL21
chr4
123,300
123,350
123,400
123,450
123,500
123,550 Kb
C
reporter assay for IL2 CNS enhancer activity
Luciferase activity (fold over pGL4)
Luciferase activity (fold over URR)
PR
URR
-46
-51
-80
-83
-128
D
CRISPR-CAS9 deletion of individual CNS
E
primary human CD4+ T cells (N=5)
Jurkat (N=3)
IL-2 (% of EV)
Untargeted
B2m
IL-2 URR
-46CNS
-51CNS
-80CNS
-83 CNS
-85 CNS
-122CNS
-128CNS
F
CRISPR-CAS9 deletion of all intergenic CNS
81.3 kb
URR
CRISPR-CAS9 deletion
delta 81.3 kb
G
Jurkat clones (N=3)
Empty vector
IL2-21 del81.3 kb
IL-2 (pg/mL)
PMA+I+a-CD28 stim (hrs)
IL2 transcript (normalized to 18s)
EV
del81.3kb
PMA+I+a-CD28 6h
H
p<0.0001
Luciferase activity (fold over URR)
-128 OCR NON-RISK
-128 OCR RISK SNP

Figure 6

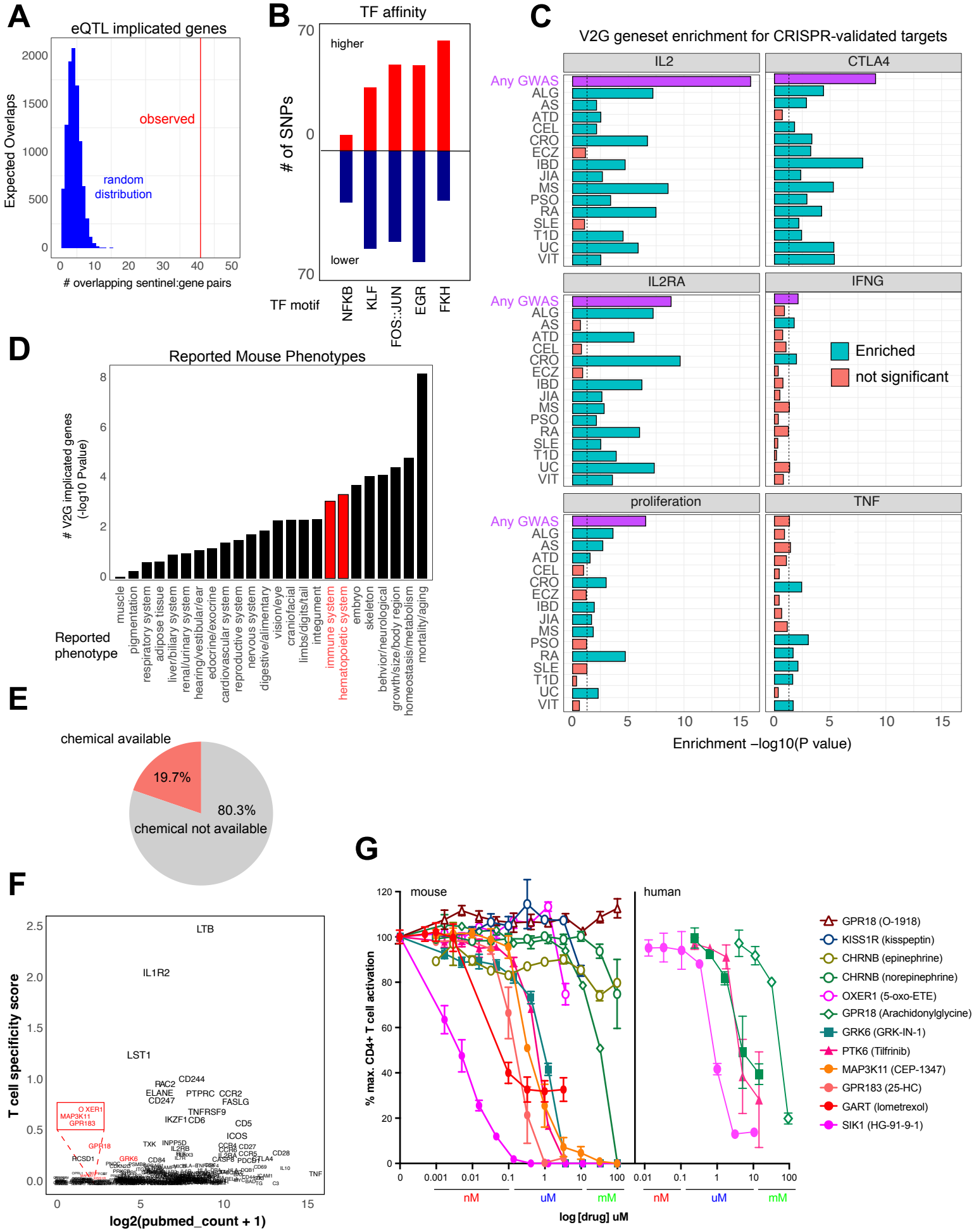

A
eQTL implicated genes
Expected Overlaps
observed
random distribution
overlapping sentinel:gene pairs
B
TF affinity
higher
lower
# of SNPs
TF motif
NFKB
KLF
FOS::JUN
EGR
FKH
C
V2G geneset enrichment for CRISPR-validated targets
IL2
CTLA4
IL2RA
IFNG
proliferation
TNF
Any GWAS
ALG
AS
ATD
CEL
CRO
ECZ
IBD
JIA
MS
PSO
RA
SLE
T1D
UC
VIT
Enriched
not significant
Enrichment –log10(P value)
D
Reported Mouse Phenotypes
V2G implicated genes (-log10 Pvalue)
Reported phenotype
muscle
pigmentation
respiratory system
adipose tissue
liver/biliary system
renal/urinary system
hearing/vestibular/ear
edocrine/exocrine
cardiovascular system
reproductive system
nervous system
digestive/alimentary
vision/eye
craniofacial
limbs/digits/tail
integument
immune system
hematopoietic system
embryo
skeleton
behvior/neurological
growth/size/body region
homeostasis/metabolism
mortality/aging
E
chemical available
19.7%
80.3%
chemical not available
F
T cell specificity score
log2(pubmed_count + 1)
LTB
IL1R2
LST1
CD244
RAC2
ELANE
CD247
PTPRC
CCR2
FASLG
TNFRSF9
IKZF1
CD6
CD5
ICOS
O XER1
MAP3K11
GPR183
GPR18
GRK6
RCSD1
CD84
TNF
G
mouse
human
% max. CD4+ T cell activation
log [drug] uM
nM
uM
mM
GPR18 (O-1918)
KISS1R (kisspeptin)
CHRNB (epinephrine)
CHRNB (norepinephrine)
OXER1 (5-oxo-ETE)
GPR18 (Arachidonylglycine)
GRK6 (GRK-IN-1)
PTK6 (Tilfrinib)
MAP3K11 (CEP-1347)
GPR183 (25-HC)
GART (lometrexol)
SIK1 (HG-91-9-1)

Figure S1: Library sequencing reproducibility and expression clustering. (A) PCA of RNA-seq libraries. (B) PCA of ATAC-seq libraries. (C) Stratified correlation (SCC) of HiC libraries from two donors by stage of activation. (D) Volcano plot of RNA-seq data for CD4+ T cells unstimulated vs. 8 hr stimulation (top) and 8 hr vs. 24 hr stimulation (number of DEG indicated). (E) Elbow plot of the within group sum of squares used to determine optimal cluster number. (F) Genes per expression cluster. (G) Centroids +/- standard deviation for each cluster.

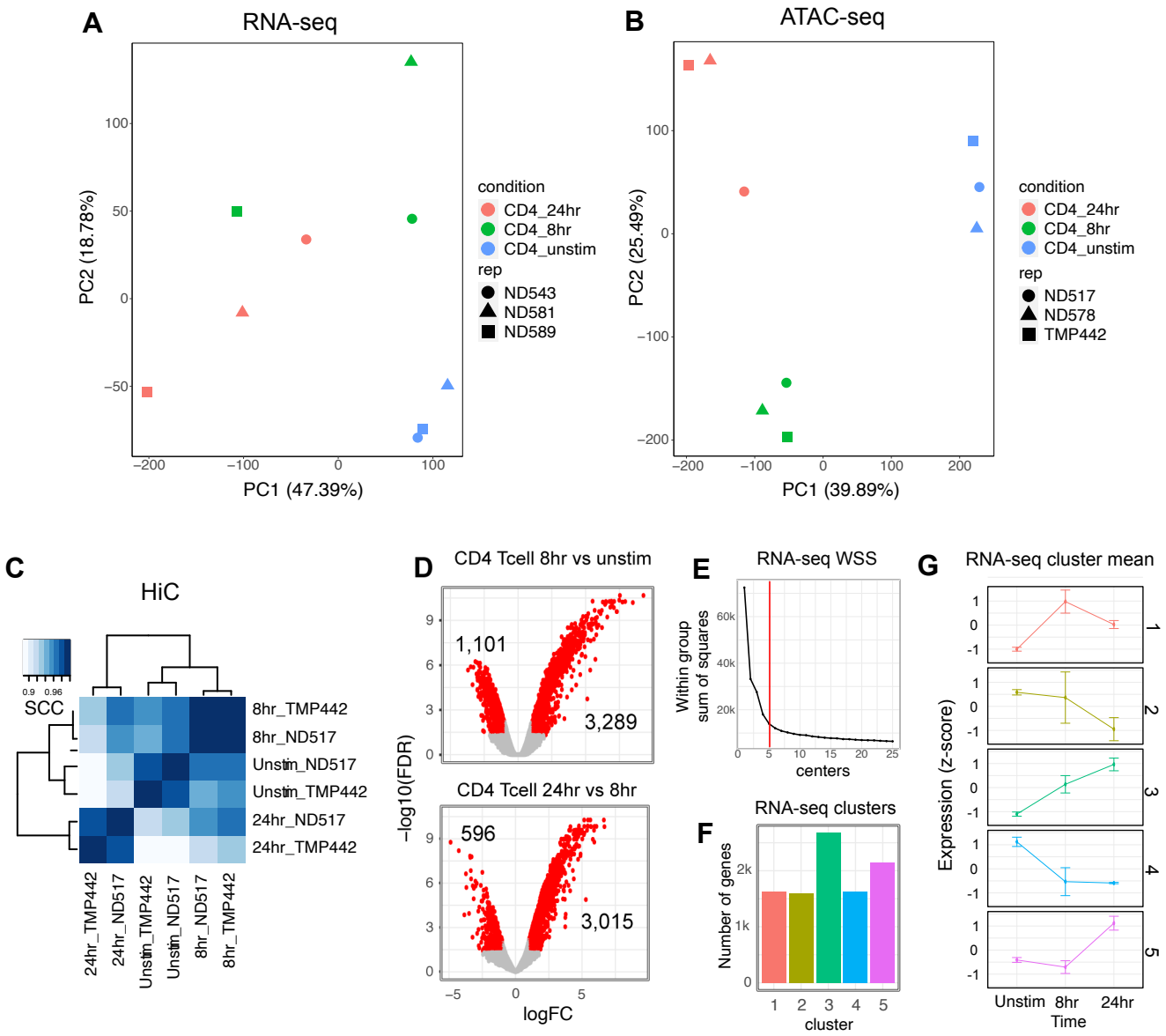


Figure S2: Characterization of CD4+ T cell epigenomic data. (A) Accessibility of OCR and (B) expression of genes located in A vs. B compartments (log2 fkpm). Differential accessibility (C) and expression (D) of cRE-gene pairs in regions exhibiting differential TAD structure. (E) Enrichment of promoter interacting region OCR for annotated regulatory elements in CD4+ T cells from the epigenome roadmap project (E038_15). (F) Overlap of cRE-promoter loops defined in this study to loops defined by Burren et al. in a prior promoter capture Hi-C study. (G) Distribution of OCR contacted per gene (top, median = 8 OCRs per gene). Distribution of genes contacted per OCR (bottom, median = 3). (H) Differential promoter-OCR connectivity by HiC as a function of differential OCR accessibility (central bar = median, boxplot edges = 25-75 quantiles, whiskers = 1.5x inter-quantile range).

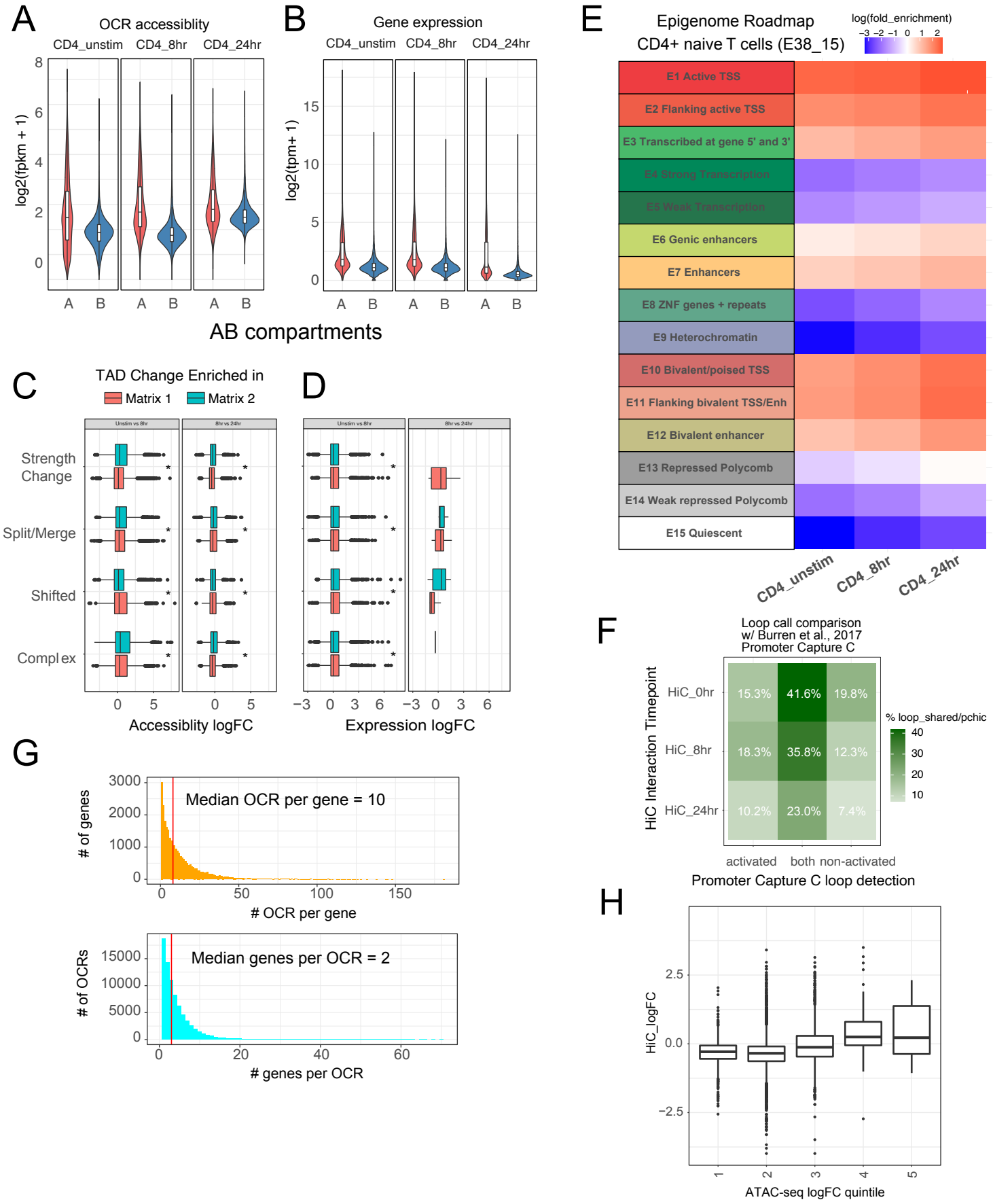


Figure S3: Characterization of chromatin stripes during CD4+ T cell activation. (A) Stripes called per timepoint. (B) Proportion of SNPs annotated in A vs B compartments. Accessibility of OCR (C) and expression of genes (D) located in anchor (small more interactive region of stripe), stripe excluding the anchor (stripe), or outside stripes.

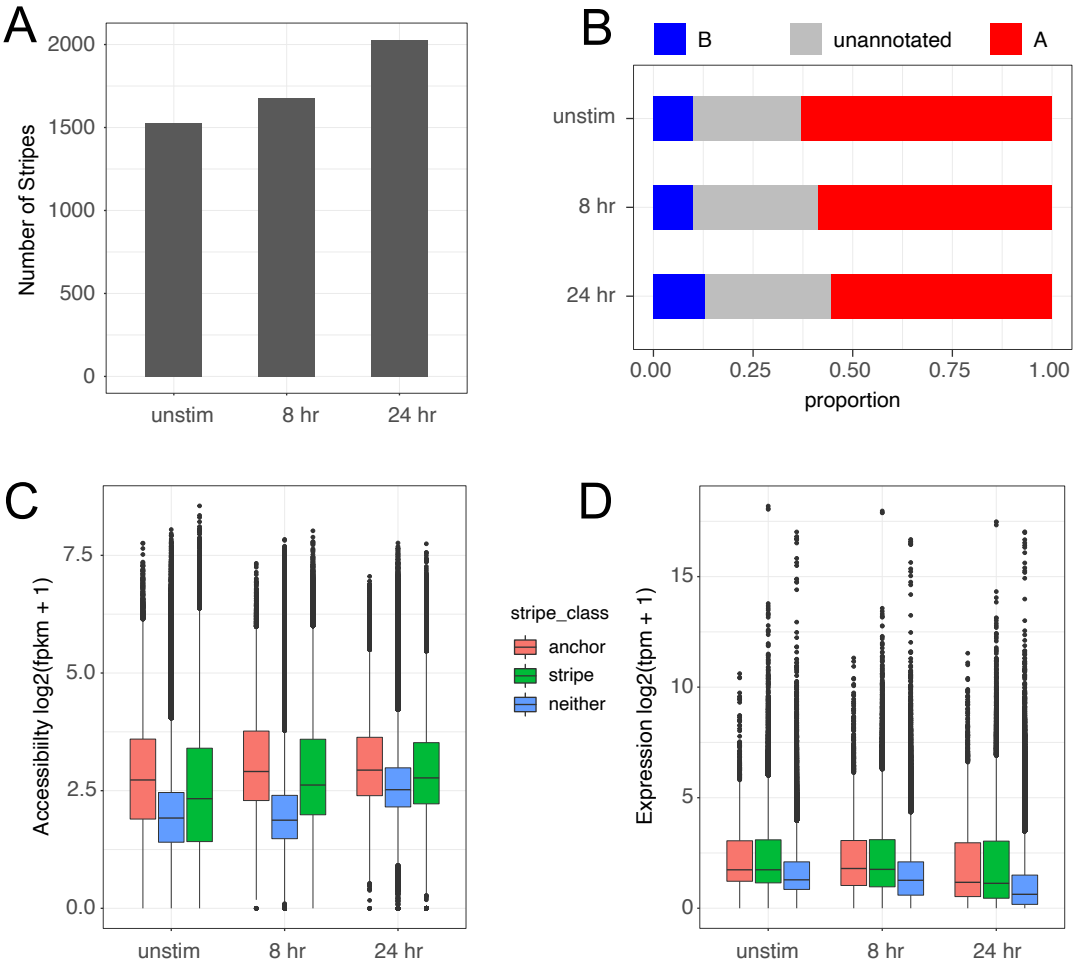


Figure S4: Large-scale epigenetic changes at dynamically expressed genes during CD4+ T cell activation. HiC contact frequency matrix (heatmap), chromatin accessibility (grey), and loop calls (red) for (A) GEM, (B) IRF4, (C) KLF2, and (D) DPEP2. The IRF4 matrix is truncated at by end of the chromosome annotation, and inset shows the expression (TPM) at each time point for each gene.

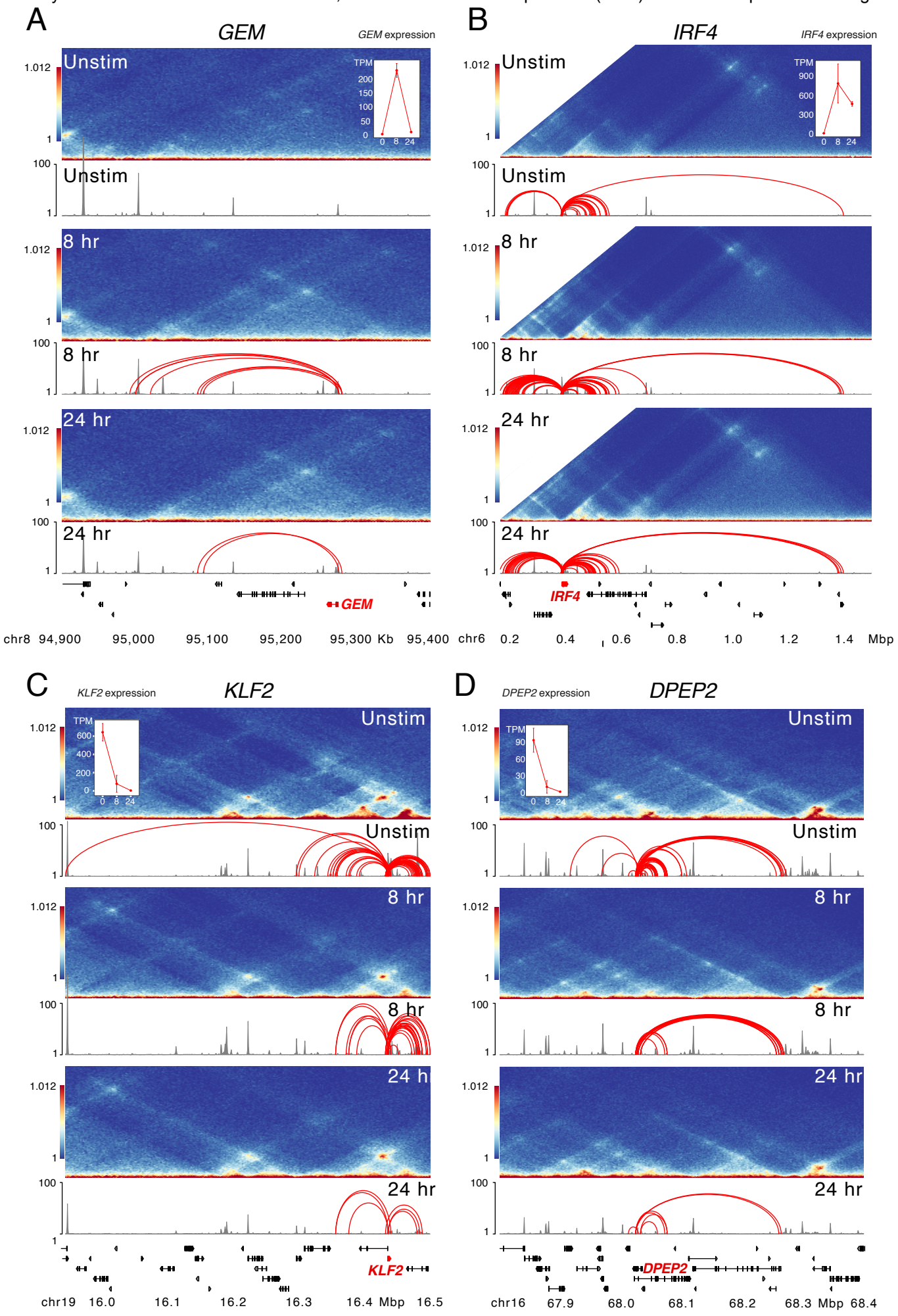


Figure S5: Enrichment of transcription family members and gene regulatory network construction. (A) -log10(P) enrichment of transcription factors annotated by family for each expression cluster. (B) Top five pathways enriched for a subset of TF-regulated gene expression.

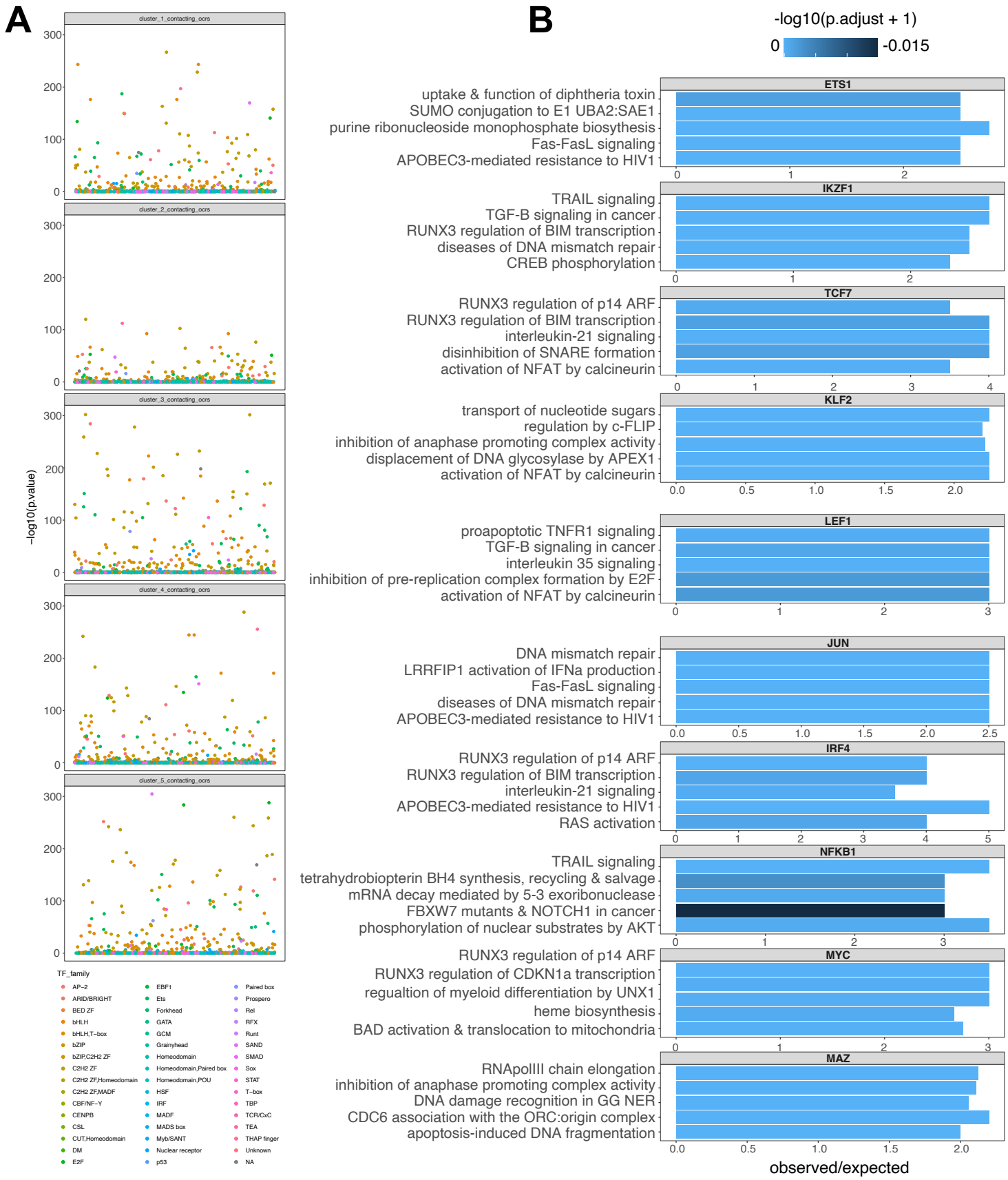


Figure S6: TF occupancy and stability of distal IL2 cRE in CD4+ T cell subsets. Evolutionary conservation (blue), open chromatin (red), and autoimmune disease-associated SNPs at the IL2 locus in quiescent (A) and 8-hour activated (B) naïve CD4+ T cells. (C) Curated TF occupancy at the IL2 URR and distal cRE from the ReMap atlas of regulatory regions.

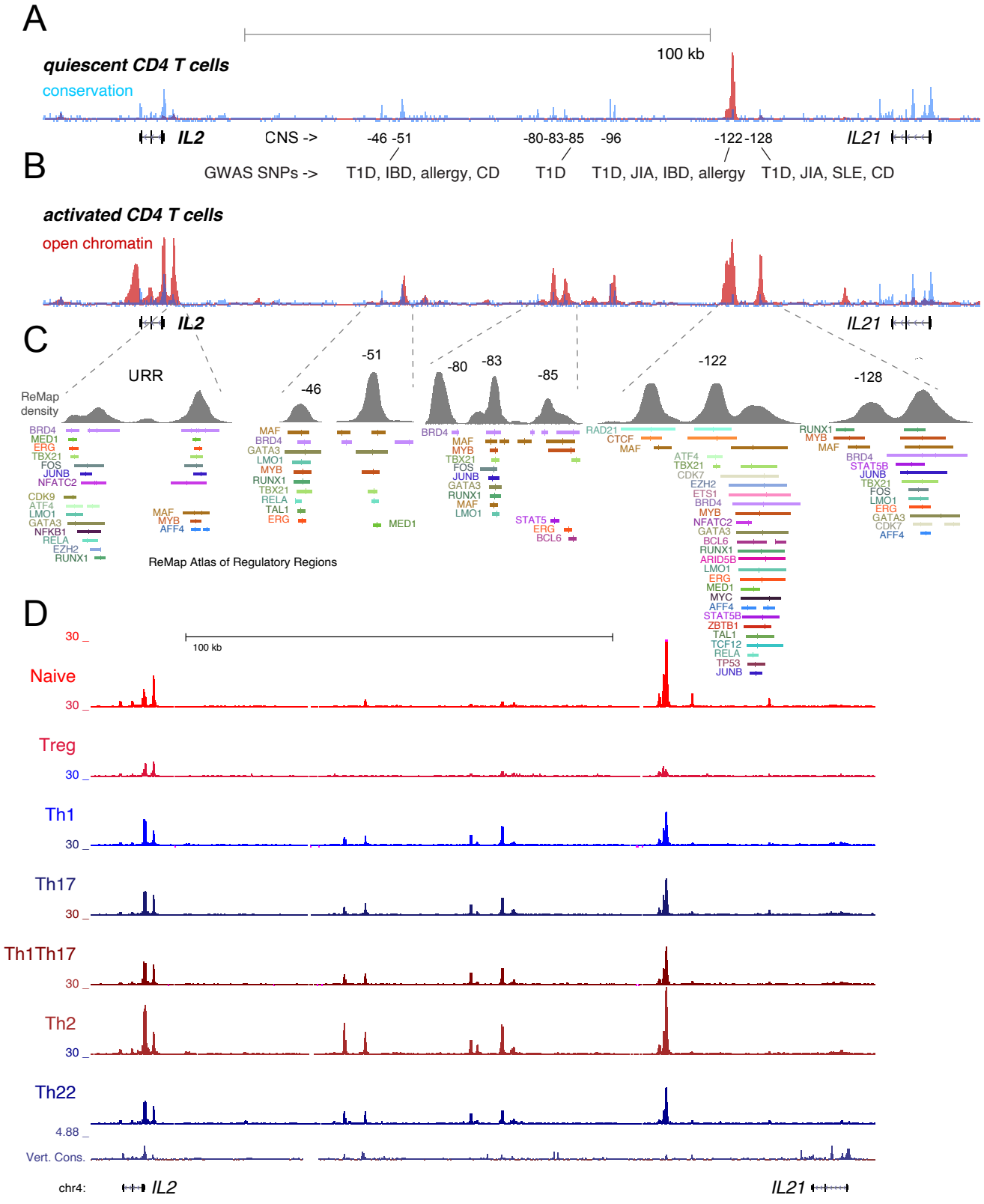


Figure S7: Orthogonal validation of 3D chromatin-based V2G genes. Expression (scaled TPM) of genes implicated by chromatin contacts in (A) prior colocalized CD4+ T cell single-cell eQTL27, (B) prior CRISPRi/a screens, and (C) genes with known immune phenotypes in the international mouse phenotyping consortium. Red = increased, blue = decreased.

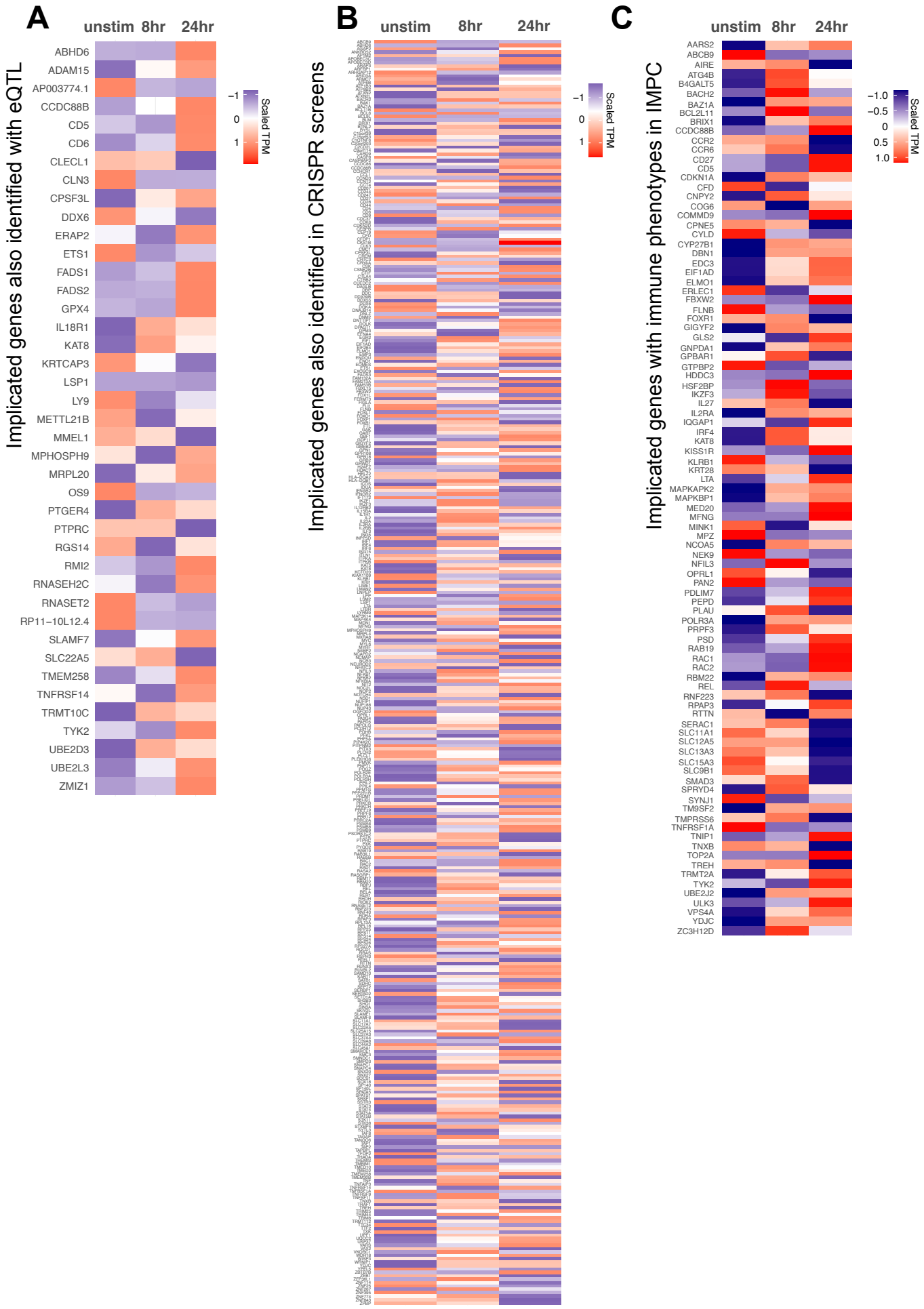
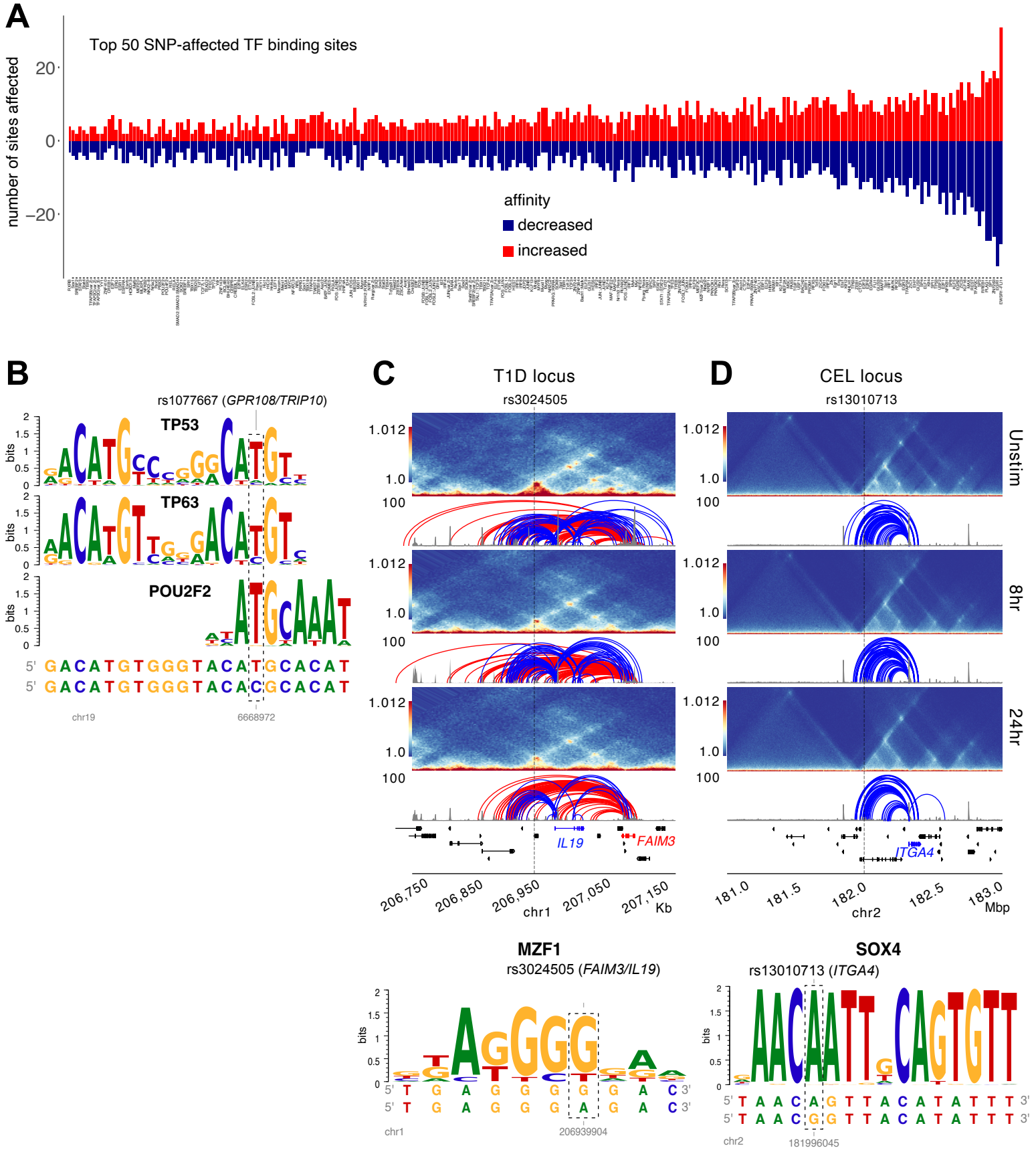



Figure S8: Predicted impact of autoimmune disease-associated genetic variation at V2G-implicated loci. (A) Top 50 TF motifs impacted by autoimmune SNPs (excerpted from Table S15). (B) Genetic variation at the MS rs1077667 SNP is predicted to disrupt TP53, TP63, and POU2F2 (OCT2) binding sites. Dashed box indicates the SNP affected by allelic variation. HiC contact frequency matrix (heatmap), chromatin accessibility (grey), loop calls (blue-IL19/ITGA4, red-FAIM3) and predicted TF motif disruption for the rs3024505-IL19-FAIM3 T1D locus (C) and the rs1301071-ITGA4 celiac locus (D). SNP rs3024505 is predicted to disrupt a MZF1 motif and rs13010713 is predicted to disrupt a SOX4 binding motif.