

1 **Genetic variants associated with cross-disorder and disorder-**
2 **specific risk for psychiatric disorders are enriched at epigenetically**
3 **active sites in peripheral lymphoid cells**

4 **Supplementary Material**

5
6 Mary-Ellen Lynall^{1,2,3,4}, Blagoje Soskic^{5,6}, James Hayhurst⁶, Jeremy Schwartzentruber⁵, Daniel F.
7 Levey^{7,8}, Gita A. Pathak^{7,8}, Renato Polimanti^{7,8}, Joel Gelernter^{7,8,9}, Murray B. Stein^{10,11}, Gosia
8 Trynka^{5,6}, Menna R. Clatworthy^{3,4}, Ed Bullmore^{1,2}

9
10
11 ¹University of Cambridge, Department of Psychiatry, Herchel Smith Building of Brain & Mind
12 Sciences, Cambridge Biomedical Campus, Cambridge CB2 0SZ, UK

13
14 ²Cambridgeshire & Peterborough NHS Foundation Trust, Cambridge, UK

15
16 ³Molecular Immunity Unit, University of Cambridge Department of Medicine, Cambridge, UK

17
18 ⁴Cellular Genetics, Wellcome Sanger Institute, UK

19
20 ⁵Wellcome Sanger Institute, Wellcome Genome Campus, Cambridge, UK

21
22 ⁶Open Targets, Wellcome Genome Campus, Hinxton, UK

23
24 ⁷VA Connecticut Healthcare System, West Haven, CT, USA

25
26 ⁸Yale Univ. School of Medicine, Dept Psychiatry, New Haven, CT, USA

27
28 ⁹Yale Univ. School of Medicine, Depts. Of Genetics and Neuroscience, New Haven, CT, USA

29
30 ¹⁰VA San Diego Healthcare System, San Diego, CA, USA

31
32 ¹¹Department of Psychiatry and School of Public Health, University of California San Diego, La
33 Jolla, CA, USA

36 Contents

- 37 • Table S1: Genome-wide association study details
- 38 • Table S2: Code and data availability
- 39 • Table S3: Statistical comparison of original vs. conditional stratified linkage disequilibrium
- 40 score regression (s-LDSC models) by Z-test
- 41 • Table S4: s-LDSC results for active regulatory annotations for 88 ROADMAP tissues for all
- 42 disorders (statistical tables; separate supplementary excel file)
- 43 • Figure S1: Enrichment of trans-diagnostic risk at active regulatory elements (active
- 44 promoters and enhancers) in 88 tissues from the Roadmap epigenomics consortium
- 45 • Figure S2: Cross-disorder psychiatric risk is enriched in brain and lymphoid immune cell
- 46 active genomic elements
- 47 • Figure S3: Enrichment of genetic risk for multiple psychiatric disorders in immune tissues
- 48 • Figure S4: Psychiatric genetic risk enrichment at active lymphoid enhancers/promoters
- 49 (BLUEPRINT dataset)
- 50 • Figure S5: Psychiatric genetic risk enrichment at activation-dependent T cell
- 51 enhancers/promoters (Soskic immune stimulation dataset)
- 52 • Figure S6: Soskic stimulated immune cell dataset: overlap of H3K27ac peaks implicated by
- 53 different disorders
- 54 • Figure S7: Soskic stimulated immune cell dataset: pathway enrichment for genes nearest to
- 55 peaks both specific to T cells and overlapped by risk variants

56

Table S1

Study	Number cases	Number controls	Number of genome-wide independently significant loci	Download link
Cross-disorder psychiatric risk ¹	162,151	276,846	115	https://pgcdata.med.unc.edu/cross_disorder/pgc_cdg2_meta_no23andMe_oct2019_v2.txt.daner.txt.gz
Depression ²	264,984	581,929	122	dbGaP Study Accession: phs001672.v6.p1
Schizophrenia ³	36,989	113,075	108	https://pgcdata.med.unc.edu/schizophrenia/ckqny.scz2snpres.gz
Bipolar disorder ⁴	20,352	31,358	16	https://www.med.unc.edu/pgc/download-results/ File = daner_PGC_BIP32b_mds7a_0416a
Autism ⁵			2	https://pgcdata.med.unc.edu/autism_spectrum_disorders/iPSYCH-PGC_ASD_Nov2017.gz
ADHD ⁶	19,099	34,194	10	https://pgcdata.med.unc.edu/adhd/adhd_eur_jun2017.gz
Body mass index ⁷	806,834	NA	1023	https://zenodo.org/record/1251813#.X_iGV_S-11TZ File = bmi.giant-ukbb.meta-analysis.combined.23May2018.txt
Alzheimer's disease ⁸	71,880	383,378	25	https://ctg.cncr.nl/documents/p1651/AD_su_mstats_Jansenetal_2019sept.txt.gz
Rheumatoid arthritis ⁹ .	14,361	43,923	48	http://plaza.umin.ac.jp/~yokada/datasource/files/GWASMetaResults/RA_GWASmeta_European_v2.txt.gz

57

58

59

60

61

62

63

Table S1 Genome-wide association study details. Genetic variants associated trans-diagnostically with risk for 8 psychiatric disorders and cis-diagnostically with risks for each of 5 specific psychiatric / neurodevelopmental disorders, and 3 positive control disorders. Loci associated with risk were thresholded at $P < 5 \times 10^{-8}$, then distance-based clumping was used to define independently significant loci (see **Methods**).

Table S2 Code and data availability

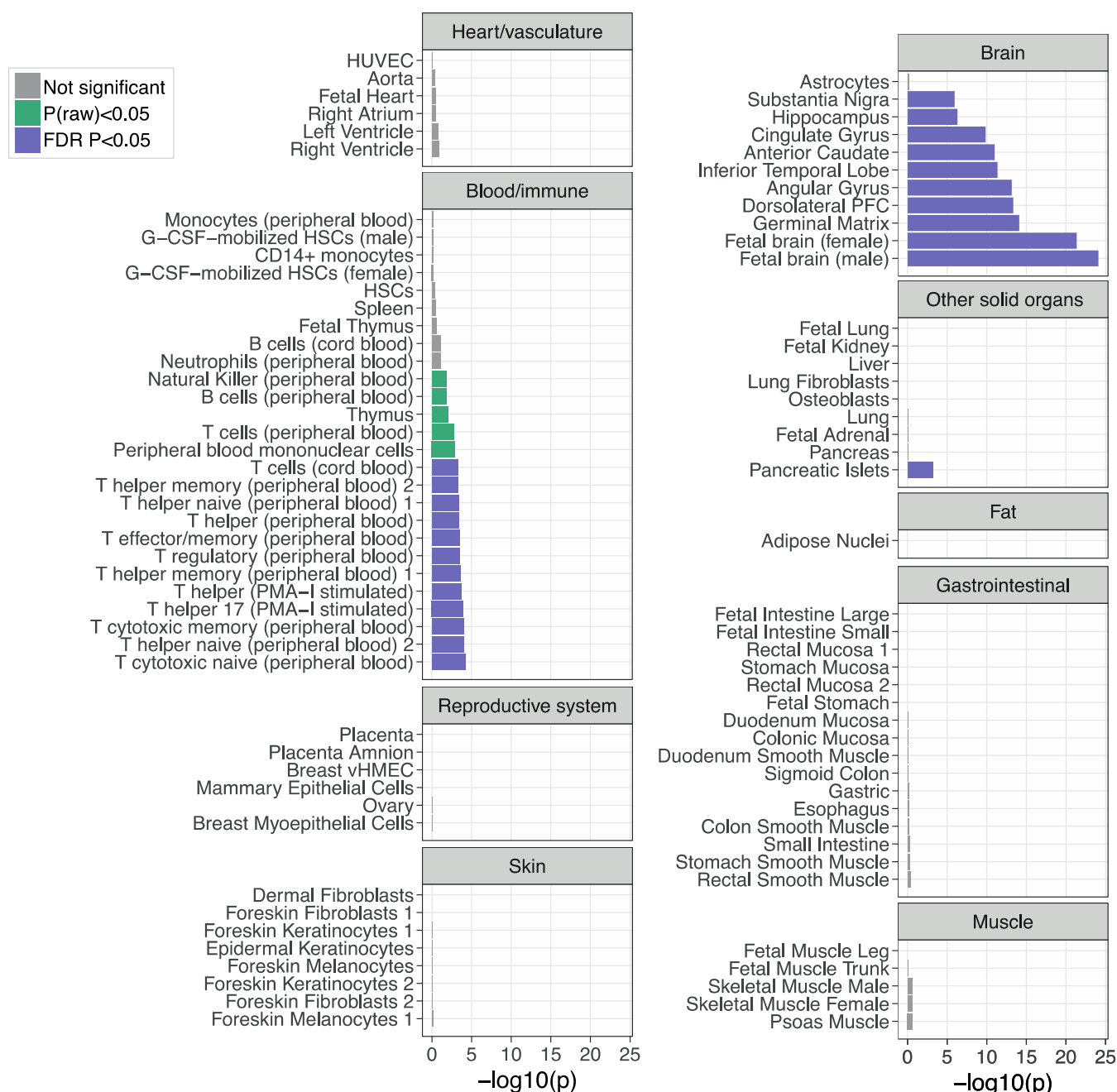
Resource	Availability
Code used to perform this analysis and generate the figures in the paper	https://github.com/maryellenlynall/psychimmgen2021
Summary statistics	See Table S1
Roadmap Epigenomics datasets	http://bx.psu.edu/~yuzhang/Roadmap_ideas/trackDb_test.txt
BLUEPRINT datasets	
Soskic immune stimulation dataset (H3K27ac)	https://www.ebi.ac.uk/ega/studies/EGAS00001002749
IDEAS annotations	http://bx.psu.edu/~yuzhang/Roadmap_ideas/trackDb_test.txt
1000 genomes called against GRCh38	http://ftp.1000genomes.ebi.ac.uk/vol1/ftp/data_collections/1000_genomes_project/release/20190312_biallelic_SNV_and_INDEL/
CHEERS code	https://github.com/trynkaLab/CHEERS
Partitioned LD scores for active regulatory elements in Roadmap tissues	Generated in this analysis; available at https://doi.org/10.5281/zenodo.5153661
GnomAD v2.1.1	https://gnomad.broadinstitute.org

68 **Table S3** Statistical comparison of original vs. conditional s-LDSC models by Z-test (see
 69 **Methods**), to accompany **Figure 1B** and **Figure S2A**

	Original model vs. conditional model including fetal male brain (Fig S2A)	Original model vs. conditional model including fetal female brain (Fig S2A)
Brain Angular Gyrus	z=1.80; p=0.04	z=2.19; p=0.01
Brain Anterior Caudate	z=1.74; p=0.04	z=2.20; p=0.01
Brain Cingulate Gyrus	z=1.75; p=0.04	z=2.19; p=0.01
Brain Germinal Matrix	z=2.61; p=0.005	z=2.92; p=0.002
Brain Hippocampus Middle	z=1.50; p=0.07	z=1.99; p=0.02
Brain Inferior Temporal Lobe	z=1.80; p=0.04	z=2.19; p=0.01
Brain Dorsolateral Prefrontal Cortex	z=1.94; p=0.03	z=2.31; p=0.01
Brain Substantia Nigra	z=1.34; p=0.09	z=1.77; p=0.04
Fetal Brain (female)	z=1.73; p=0.04	NA
Fetal Brain (male)	NA	z=2.77; p=0.003
Immune cell subsets for which trans-risk showed significant enrichment ($q < 0.05$) in original s-LDSC model	Original model vs. conditional model including all 10 significantly enriched brain regions (Fig 1B)	
T cytotoxic naive cells (peripheral blood)	z=0.34; p=0.37	
T helper naive cells (peripheral blood) 2	z=0.20; p=0.42	
T cytotoxic memory cells (peripheral blood)	z=0.21; p=0.42	
T helper 17 cells (PMA-I stimulated)	z=0.19; p=0.42	
T helper cells (PMA-I stimulated)	z=0.19; p=0.42	
T helper memory cells (peripheral blood) 1	z=0.17; p=0.43	
T regulatory cells (peripheral blood)	z=0.20; p=0.42	
T effector/memory (peripheral blood)	z=0.26; p=0.40	
T helper cells (peripheral blood)	z=0.20; p=0.42	
T helper naive cells (peripheral blood) 1	z=0.28; p=0.39	
T helper memory cells (peripheral blood) 2	z=0.18; p=0.43	
T cells (cord blood)	z=0.23; p=0.41	

70
71

72



73

74

75

76

77

78

79

80

81

82

83

84

85

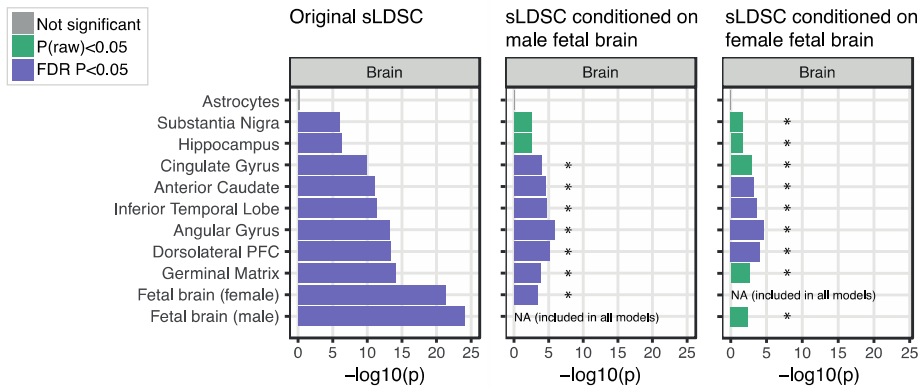
86

87

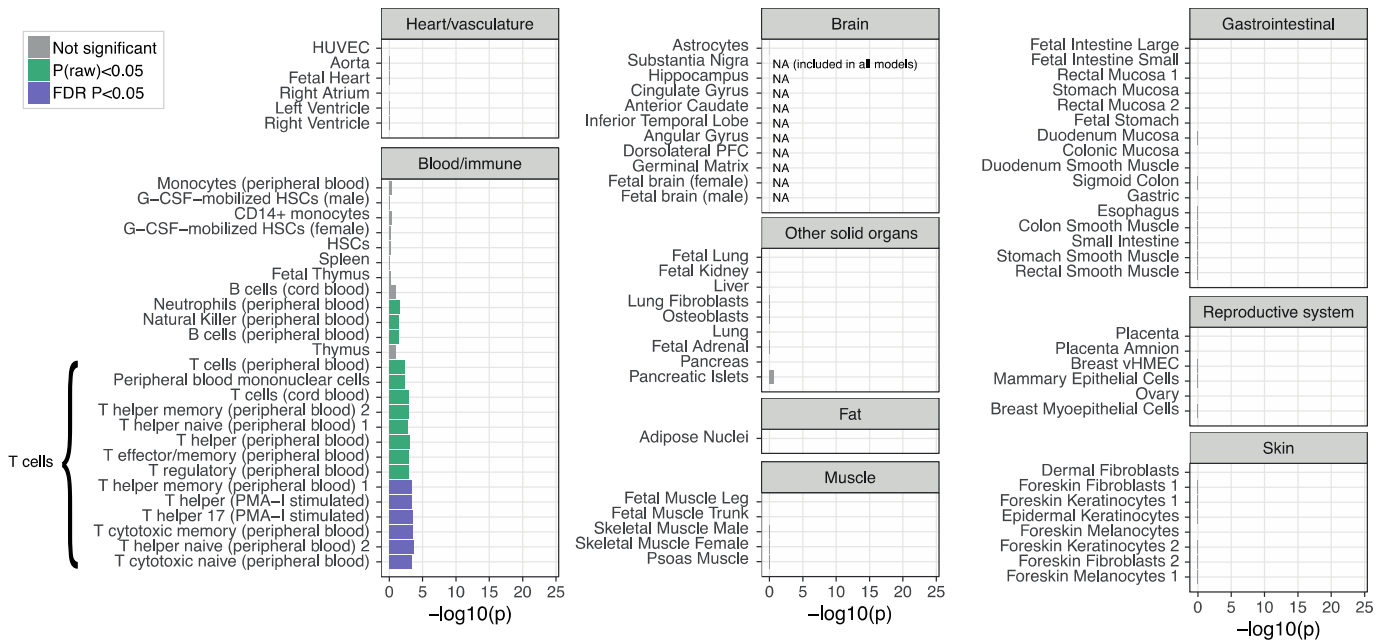
Figure S1 Enrichment of trans-diagnostic risk at active regulatory elements (active promoters and enhancers) in 88 tissues from the Roadmap epigenomics consortium. P -values are shown for the results of stratified linkage disequilibrium score regression (s-LDSC) analysis, taking the union of active elements in a given cell type as the annotation of interest (see **Methods**). The P -values from s-LDSC were used to test the null hypotheses that risk variants were not co-located with epigenetically activated sites more frequently than expected by chance, at two probability thresholds: $P < 0.05$ (green); and Benjamini-Hochberg $P_{FDR} < 0.05$ (purple), to correct for multiple tests across $N=88$ tissues. HUVEC, human umbilical vein endothelial cells; vHMEC, variant human mammary epithelial cells; PFC, prefrontal cortex; HSC, hematopoietic stem cell; PMA-I, phorbol-myristate-acetate and ionomycin. A subset of these results is shown in **Figure 1A**.

Psychiatric risk variant enrichment at epigenetically active sites in lymphoid cells
 Lynall et al SUPPLEMENTARY MATERIAL 4 August 2021

A Cross-disorder genetic risk: brain tissue enrichment controlling for overlap with active brain tissue elements in male or female fetal brain (benchmarking conditional analysis)



B Cross-disorder genetic risk: conditional analysis for all tissues, controlling for overlap with active brain tissue elements in all brain regions significantly enriched in original LDSC model (Fig 1A)



88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99

Figure S2 Cross-disorder psychiatric risk is enriched in brain and lymphoid immune cell active genomic elements (A) Validation of brain-conditioned analysis method: repeat of the analysis in **Figure 1A**, including the active regulatory annotation for fetal male brain as an additional term in the s-LDSC models for other cell types. LHS shows original s-LDSC for brain regions; RHS shows s-LDSC analysis with the addition of fetal male brain annotation to all models. Asterisks indicate those annotations showing significantly decreased enrichment in the conditional compared to the original analysis (two-sample Z -test $P < 0.05$). (B) Brain-conditioned analysis: repeat of the analysis in **Figure 1A**, including the active regulatory annotations for all 10 significantly enriched brain regions as additional terms in the s-LDSC models for all other cell types. A subset of these results is shown in **Figure 1B**.

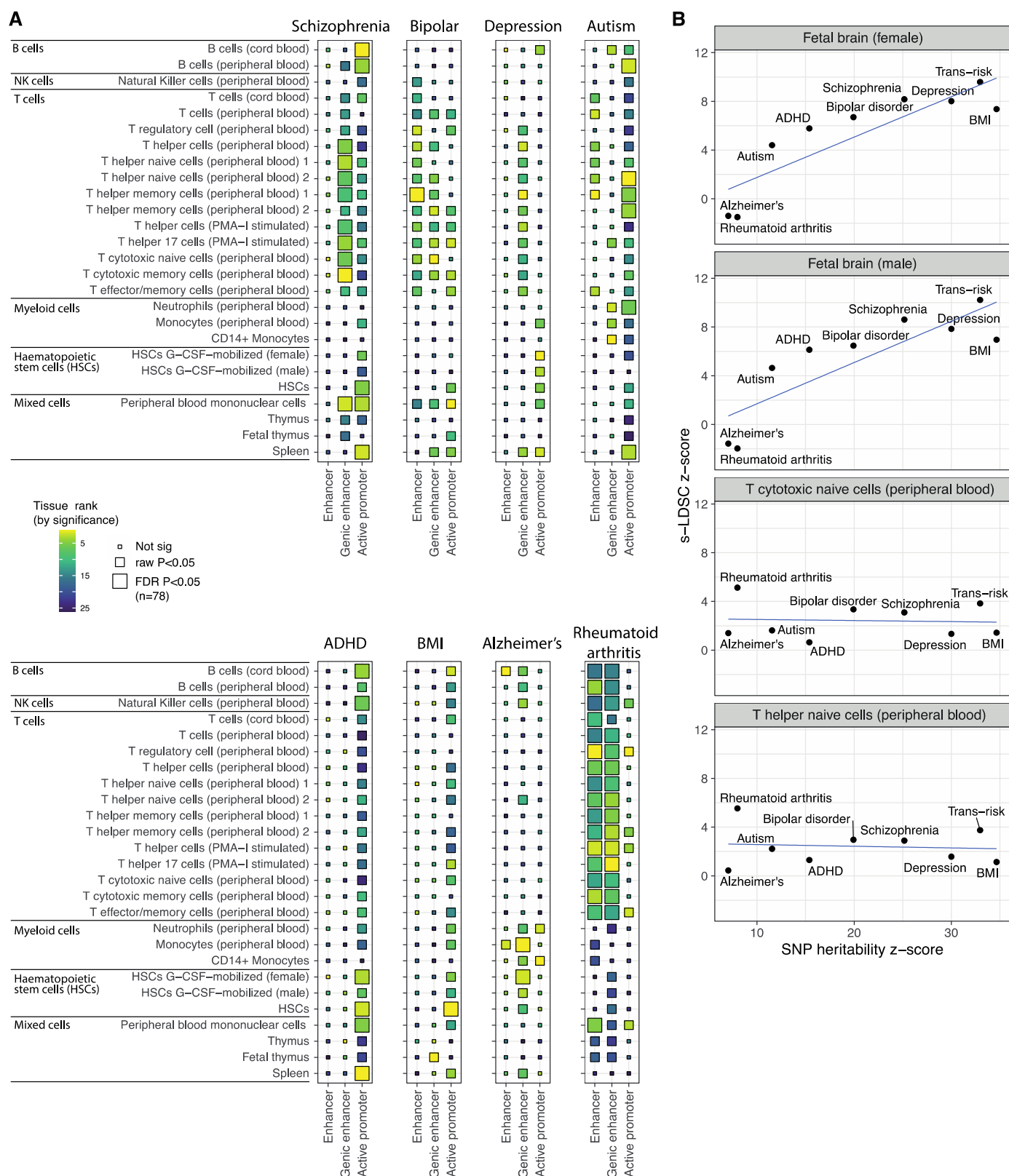
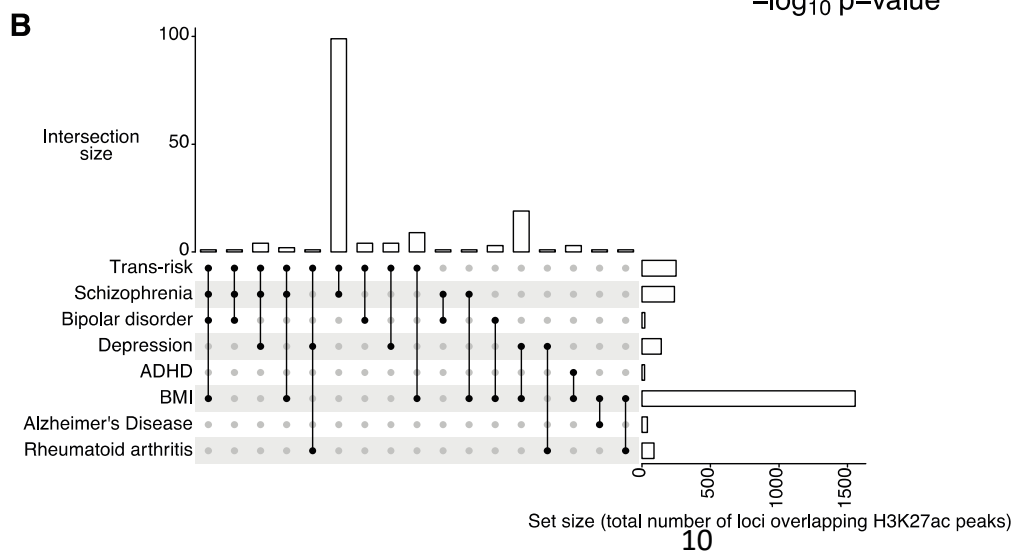
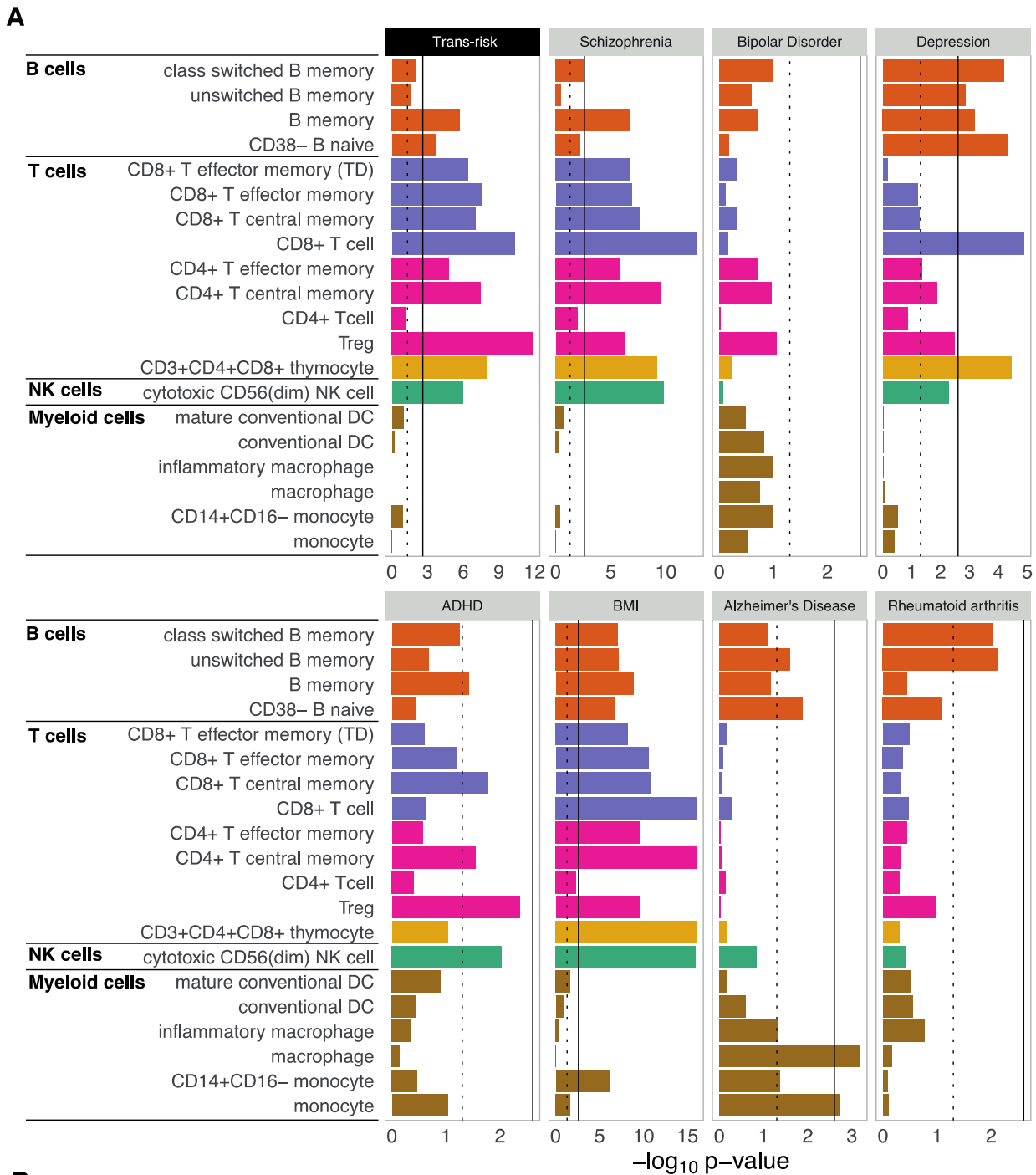


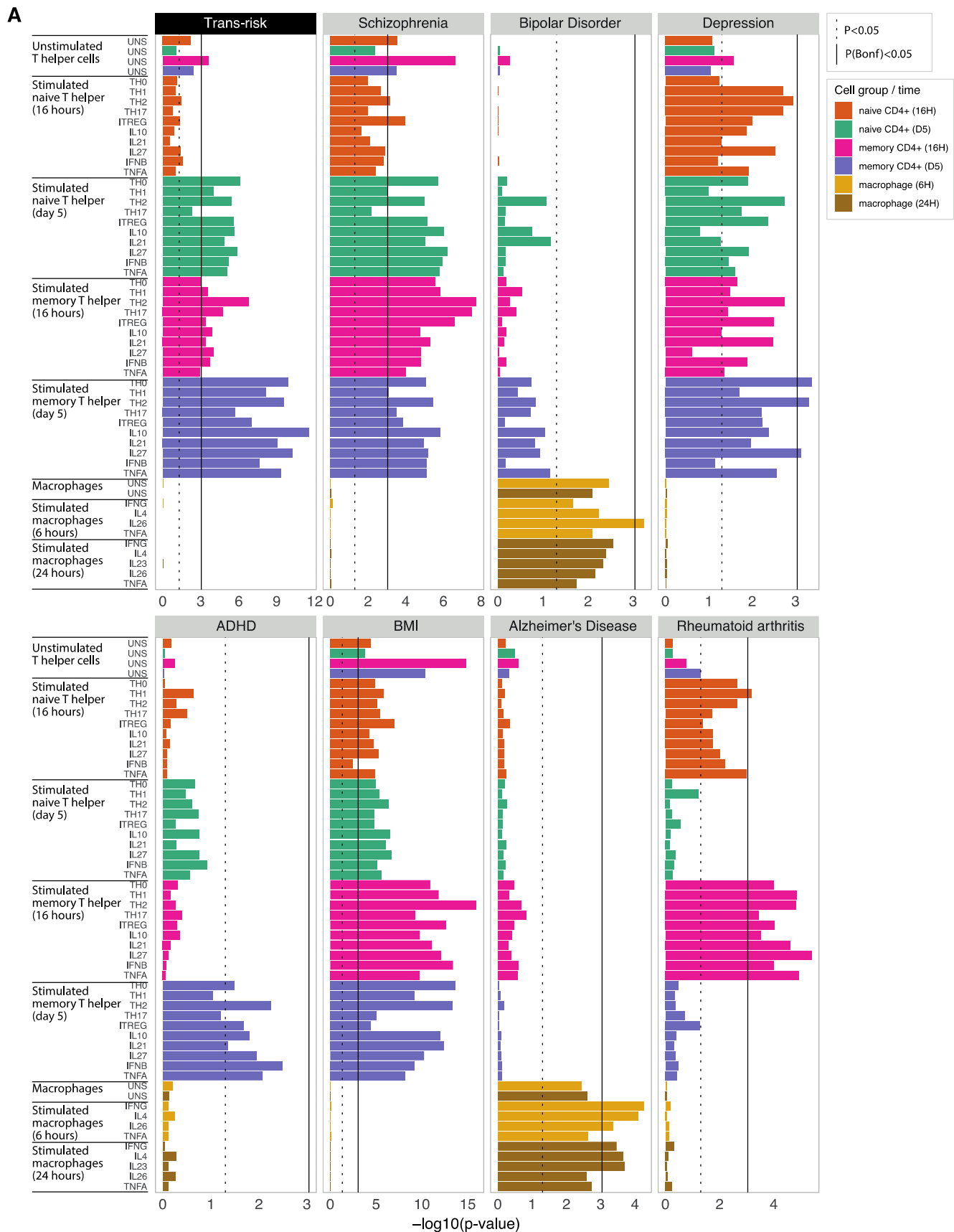
Figure S3 Enrichment of genetic risk for multiple psychiatric disorders in immune tissues
 (A) Enrichment at immune cell enhancers, genic enhancers and active promoters in Roadmap immune tissues. Enrichment at each annotation was calculated using stratified linkage disequilibrium score regression (s-LDSC). The *P*-values from s-LDSC indicate the significance of the coefficient for the cell type specific annotations. Tile size indicates significance with FDR correction across all 78 annotations tested. Tile fill indicates the *P*-value rank (ranking within each

108 annotation across cell types). (B) Correlations between GWAS SNP heritability Z-scores and s-
109 LDSC Z-scores for the top two brain and immune annotations for trans-risk. Spearman's
110 correlations with heritability z-score are as follows: fetal male brain $\rho = 0.87$, $P = 0.005$; fetal
111 female brain $\rho = 0.87$ $P = 0.005$; cytotoxic T cells $\rho = 0$, $P = 1$; helper T cells: $\rho = 0.03$, $P = 0.9$.
112

Psychiatric risk variant enrichment at epigenetically active sites in lymphoid cells
 Lynall et al SUPPLEMENTARY MATERIAL 4 August 2021



114 **Figure S4 Psychiatric genetic risk enrichment at active lymphoid enhancers/promoters**
115 **(BLUEPRINT dataset)** (A) Bar plots show enrichment of genetic risk for each condition at active
116 promoters/enhancers (H3K27ac marks) in unstimulated sorted immune cells. CHEERS was used
117 to detect enrichment of risk loci at cell-type specific H3K27ac peaks by quantifying (for each cell
118 type) the mean cell type specificity score (for that cell type) of peaks overlapping genetic risk
119 variants (see **Methods**). The dotted black line marks the nominal significance threshold, $P < 0.05$;
120 the solid black line marks the Bonferroni-corrected significance threshold, $P_{FDR} < 0.05$. Note
121 differing x-axis scales. (B) Upset plot for all Soskic immune stimulation H3K27ac immune peaks
122 overlapped by risk variants for each disorder, showing counts (vertical bars) of shared peaks,
123 compared to total peak number implicated by each disorder (horizontal bars).
124
125

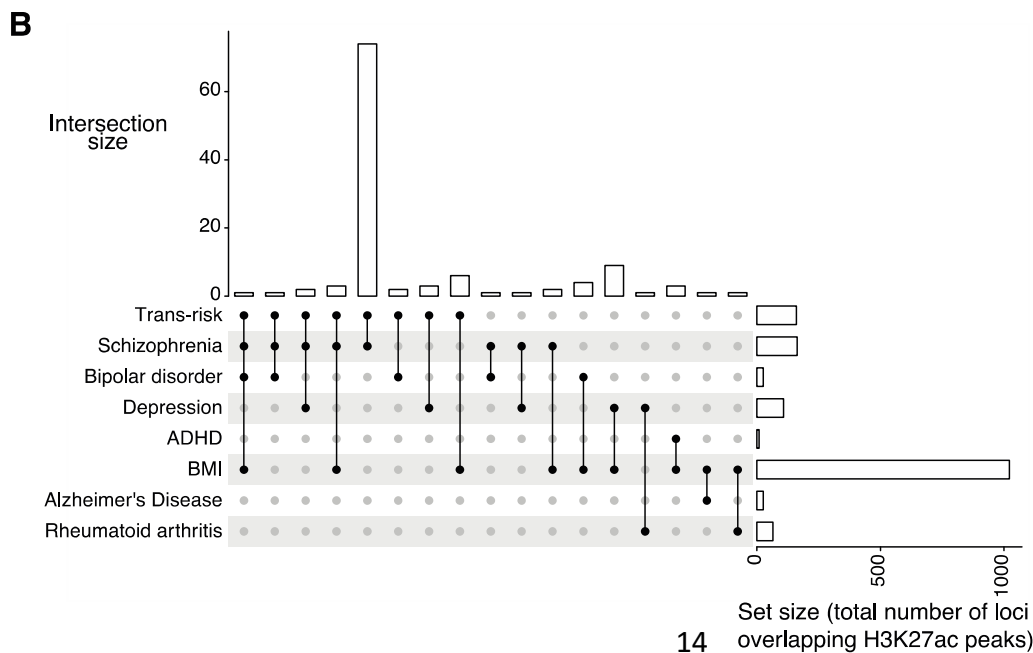
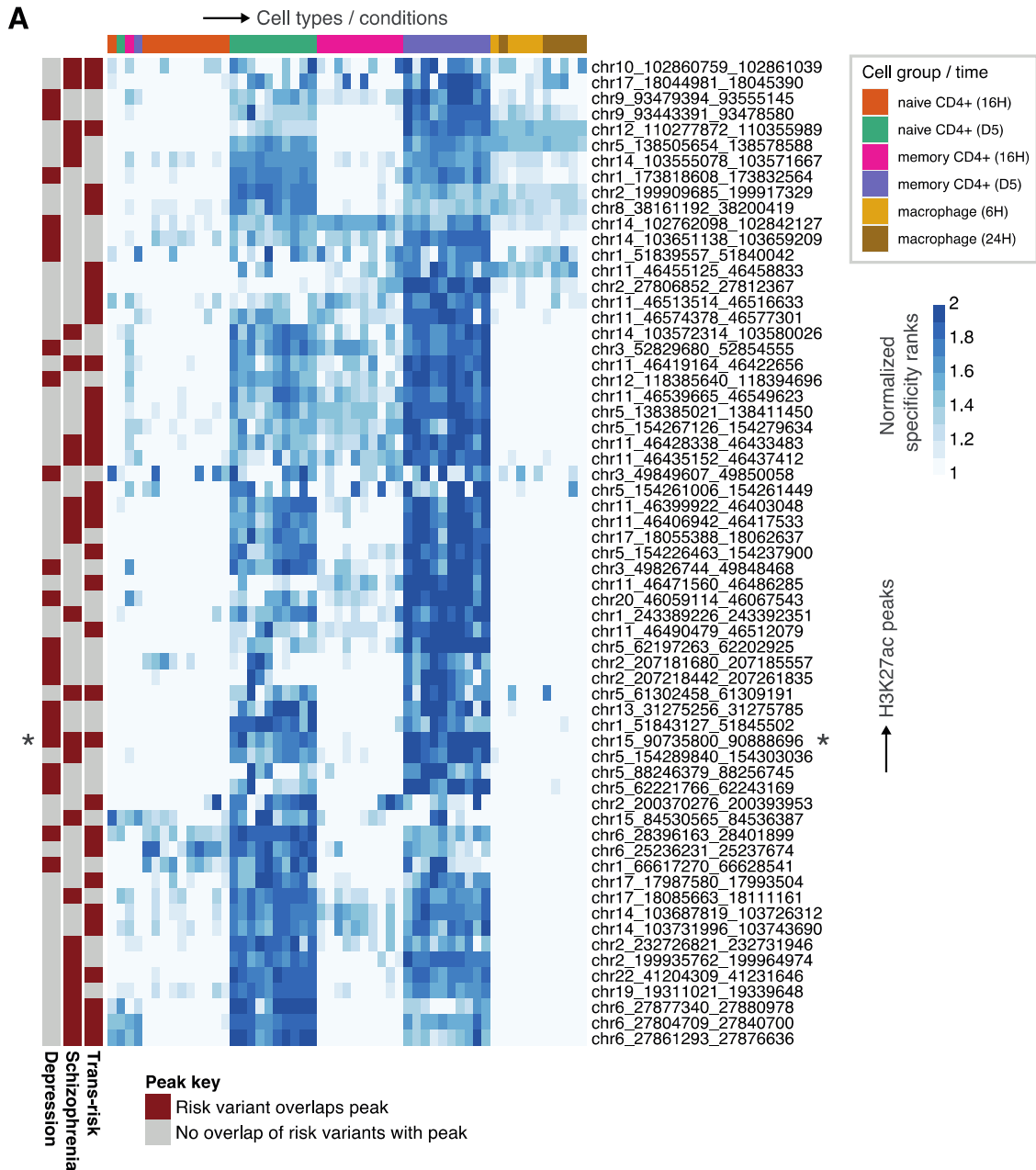


126
 127
 128
 129
 130

Figure S5 Psychiatric genetic risk enrichment at activation-dependent T cell enhancers/promoters (Soskic immune stimulation dataset). Bar plots show enrichment of genetic risk for each condition at active promoters/enhancers (H3K27ac marks) in unstimulated and *ex vivo* stimulated immune cells. Stimulated cells are sorted macrophages, naïve CD4⁺

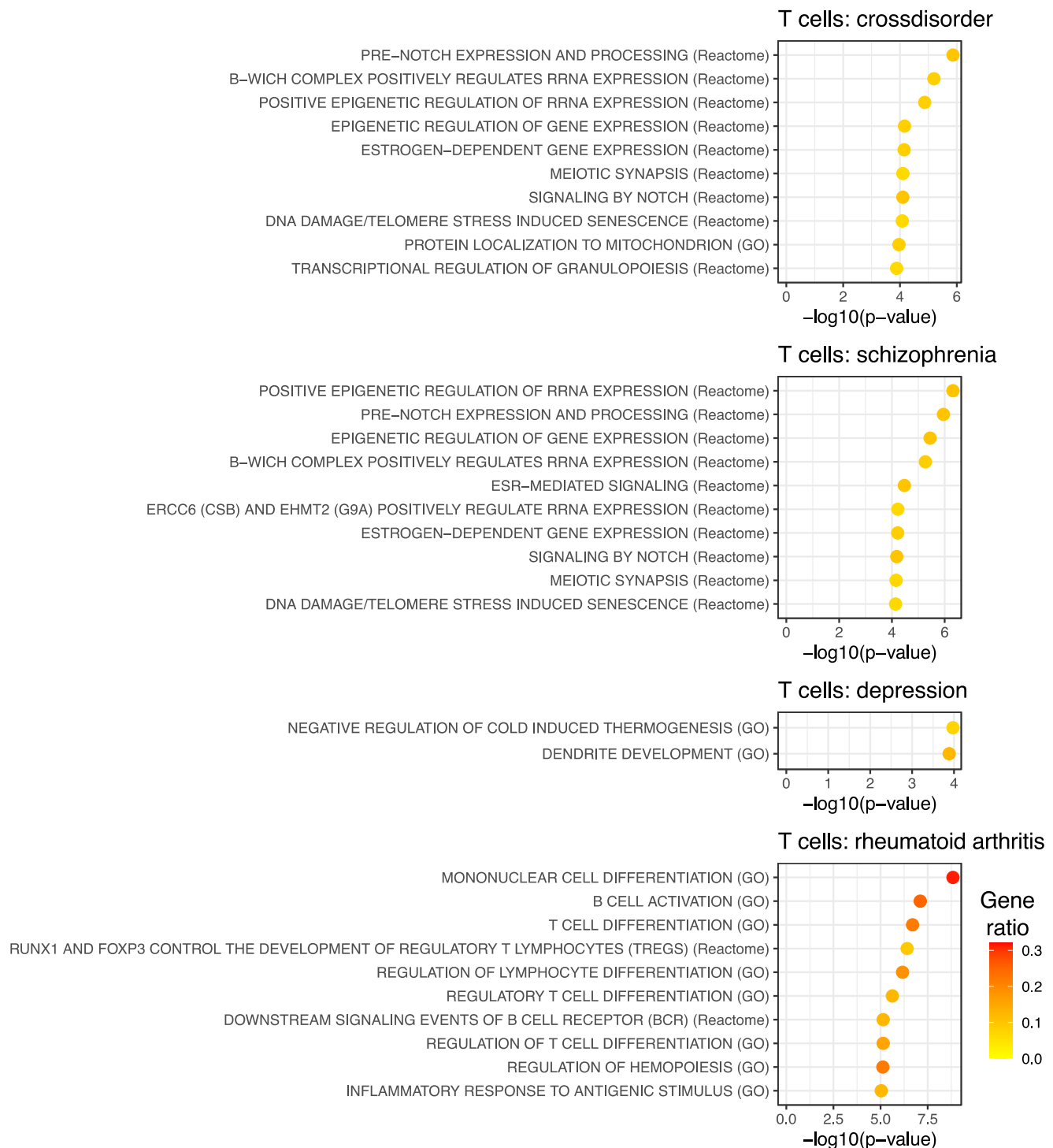
131 (helper) T cells and memory CD4⁺ T cells, assayed at both early and late timepoints (see legend).
132 CHEERS was used to detect enrichment of risk loci at cell-type specific H3K27ac peaks by
133 quantifying, for each cell type, the mean cell type specificity score (for that cell type) of peaks
134 overlapping genetic risk variants (see **Methods**). P-values are reported from a discrete uniform
135 distribution. The dotted black line marks raw $P < 0.05$; the solid black line marks Bonferroni-
136 corrected $P_{\text{FDR}} < 0.05$. Note differing x-axis scales.
137

Psychiatric risk variant enrichment at epigenetically active sites in lymphoid cells
 Lynall et al SUPPLEMENTARY MATERIAL 4 August 2021



139 **Figure S6 Soskic stimulated immune cell dataset: overlap of H3K27ac peaks implicated by**
140 **different disorders** (A) Heatmap shows the subset of peaks with specificity for late-activated
141 naïve and/or memory CD4⁺ T cells which are also overlapped by risk variants for either trans-risk,
142 schizophrenia, or major depressive disorder. Each row corresponds to a H3K27ac peak
143 overlapping a risk variant; each column corresponds to a different cytokine-induced cell state (see
144 legend), ordered as in **Figure 4A**. Blue fill shade represents how specific each peak is to each cell
145 state (specificity rank of the peak normalized to the mean specificity rank of all peaks). Row
146 annotations indicate peaks which overlap (dark red) or do not overlap (grey) risk variants for the
147 disorder indicated. Of the late-activation T cell specific peaks, only 1 (starred *) is overlapped by
148 both schizophrenia and depression risk variants. (B) Upset plot for all Soskic dataset H3K27ac
149 immune peaks overlapped by risk variants for each disorder, showing counts (vertical bars) of
150 shared peak overlaps, compared to total number of peaks implicated by each disorder (horizontal
151 bars).
152

153
 154



155
 156
 157
 158
 159
 160
 161
 162
 163
 164

Figure S7 Soskic stimulated immune cell dataset: pathway enrichment for genes nearest to peaks both specific to T cells and overlapped by risk variants. Results are shown for those disorders which showed enrichment in T cell subset. Peaks which both overlapped a disease-specific risk variant and were highly specific (specificity rank > 0.9) to any of the T cell subsets in the Soskic immune stimulation dataset were selected, and nearest genes identified (see **Methods**). For these genes, over-representation analysis was used to detect pathway enrichment in Reactome and GO Biological Process pathways. Only pathways with $P_{FDR} < 0.05$ are shown, with a maximum of 10 pathways shown per condition. Fill colour indicates gene ratio (number of test genes in the pathway / total number of test genes).

Supplementary references

- 1 Cross-Disorder Group of the Psychiatric Genomics Consortium. Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* **179**, 1469-1482 e1411, doi:10.1016/j.cell.2019.11.020 (2019).
- 2 Levey, D. F. *et al.* Bi-ancestral depression GWAS in the Million Veteran Program and meta-analysis in >1.2 million individuals highlight new therapeutic directions. *Nat Neurosci*, doi:10.1038/s41593-021-00860-2 (2021).
- 3 Ripke, S. *et al.* Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet* **45**, 1150-1159, doi:10.1038/ng.2742 (2013).
- 4 Stahl, E. A. *et al.* Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet* **51**, 793-803, doi:10.1038/s41588-019-0397-8 (2019).
- 5 Grove, J. *et al.* Identification of common genetic risk variants for autism spectrum disorder. *Nat Genet* **51**, 431-444, doi:10.1038/s41588-019-0344-8 (2019).
- 6 Demontis, D. *et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat Genet* **51**, 63-75, doi:10.1038/s41588-018-0269-7 (2019).
- 7 Pulit, S. L. *et al.* Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. *Hum Mol Genet* **28**, 166-174, doi:10.1093/hmg/ddy327 (2019).
- 8 Jansen, I. E. *et al.* Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet* **51**, 404-413, doi:10.1038/s41588-018-0311-9 (2019).
- 9 Okada, Y. *et al.* Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* **506**, 376-381, doi:10.1038/nature12873 (2014).