

Integrative Genetics Analysis of Juvenile Idiopathic Arthritis Identifies Novel Loci

Table S1. The composition of the 7 subtypes in the JIA cohort after quality control filtering. A common set of controls having no history of any existing autoimmune or immune-mediated diseases was used in association analyses.

Subtype Name	Abbreviation	Case	Female %	Control
Enthesitis-Related Arthritis	ENT	161	39.1	9250
Polyarthritis RF Negative	PRFN	284	73.6	9250
Polyarthritis RF Positive	PRFP	56	92.9	9250
Oligoarthritis	OLG	460	76.7	9250
Psoriatic Arthritis	PSO	88	73.9	9250
Systemic Arthritis	SYS	109	56.9	9250
Undifferentiated Arthritis	UND	87	56.3	9250
Juvenile Idiopathic Arthritis	JIA	1245	68.5	9250

The percentage of female among controls is 46.8%.

Table S2. The genome-wide significant loci from hsGWAS that have been reported in literature.

SNP	Chr	Pos	Diseases	reported SNP	r-Chr	r-Pos	A1	Beta	95% CI	P-value	PMID
rs2066363	1	82237577	Pediatric autoimmune diseases	rs2066363	1	82237577				8.00E-11	26301688
rs144844686	2	23457697 0	Pediatric autoimmune diseases	rs3600148 8	2	233276621	C			8.00E-11	26301688
			Crohn's disease	rs3792109	2	233275771	A	1.34	[1.29-1.40]	7.00E-41	21102463
rs7660520	4	18282416 8	Pediatric autoimmune diseases	rs7660520	4	182824168	A			8.00E-11	26301688
			Systemic lupus erythematosus	rs1001895 1	4	183688220	T	1.31	[NR]	1.00E-14	23740775
rs7731626	5	56148856	Pediatric autoimmune diseases	rs7731626	5	56148856	A			1.00E-10	26301688
			Rheumatoid arthritis	rs7731626	5	56148856	A	0.193	[0.15-0.23] unit decrease	8.00E-23	24390342
rs12203592	6	396321	Inflammatory bowel disease	rs7773324-A	6	382559	A	1.06		6.00E-09	26192919
			Crohn's disease	rs7773324-A	6	382559	A	1.08	[1.06-1.11]	1.00E-09	26192919
			Rheumatoid arthritis	rs9378815	6	426155	C	1.09	[1.06-1.12]	2.00E-10	24390342

SNP=index SNP at each locus in our study; Chr=chromosome; Pos=Position on genome-build hg19; Disease=associated disease reported in literature; reported SNP=the index SNP at each locus reported in literature; r-Chr=the chromosome of the reported SNP; r-Pos=the position of the reported SNP; A1, Beta, 95% CI and P-value =the effect allele, the regression coefficient, the 95% confidence interval, the association p-value for the reported SNP; PMID= the pubmed ID of the literature.

Table S3. The association statistics in conditional analyses at loci 3q28, 4q34.3, 7q11.22 and 18q21.1

LOCUS	STEP	TOP SNP	CHR	POS	A1	MAF	BETA	SE	P-value
3q28	1	rs7636581	3	189781195	A	0.12	0.783	0.095	2.32E-16
	2	rs6763934	3	189780459	T	0.11	-1.448	0.193	6.70E-14
4q34.3	1	rs13119493	4	180911259	G	0.048	0.700	0.124	1.48E-08
	2	rs11131944	4	180687498	G	0.16	0.347	0.0707	9.23E-07
7q11.22	1	rs727845	7	67607209	G	0.19	-0.351	0.0632	2.82E-08
	2	rs9654878	7	67610420	T	0.19	1.843	0.235	5.00E-15
18q21.1	1	rs11663074	18	45023793	C	0.17	-0.650	0.0609	1.26E-26
	2	rs11659667	18	45023809	G	0.19	1.760	0.128	4.69E-43

STEP=the order of steps in step-wise conditional analysis; TOP SNP=the most significantly associated SNP at each step of the analysis; CHR=chromosome; POS=Position (bp); A1=alternative allele; MAF=minor allele frequency; BETA=the coefficient in association testing; SE=standard error; P_value=nominal P-value at each step of the conditional association analysis.

Table S4. The 28 loci associated with JIA at genome-wide marginal significance in hsGWAS ($5 \times 10^{-8} < P\text{-value} < 1 \times 10^{-6}$)

SNP	Chr	Pos	Cytoband	Gene	Annotati on	A 1	MAF	hsGWAS _P	JIA_sub
rs55857240	1	5934103	1p36.31	NPHP4	intronic	G	0.052 2	5.08E-07	PRFN,PRFP,OLG,SYS
rs34822619	1	1082356 14	1p13.3	VAV3	intronic	G	0.034 7	2.03E-07	PRFN,OLG,PSO,SYS,UND
rs2714572	2	3569581 3	2p22.3	MIR548AD	657bp 5'	T	0.184	6.09E-07	PRFP,OLG,UND
rs11214422 3	2	5845494 7	2p16.1	FANCL	intronic	T	0.011 4	7.17E-08	PRFP,OLG,PSO,UND
rs18192689 6	2	8387227 1	2p11.2	U6	13kb 5'	A	0.011 7	1.36E-07	ENT,PSO
rs36029722	2	1829424 51	2q31.3	PPP1R1C	intronic	A	0.015 6	8.05E-08	PRFN,PRFP,OLG,PSO,UND
rs79213572	2	2298032 42	2q36.3	PID1	3'	G	0.066 9	1.10E-07	ENT,PRFN,OLG,PSO,SYS,UND
rs14017624 2	3	1761577 7	3p24.3	TBC1D5	intronic	G	0.017 6	1.29E-07	UND
rs6856482	4	4022201 7	4p14	RHOH	intronic	G	0.430	6.35E-07	ENT,PRFN,PRFP,OLG,PSO
rs62322747	4	1234564 57	4q27	IL21	77kb 3'	A	0.322	6.67E-07	PRFN,PRFP,OLG
rs11533402 3	4	1412638 31	4q31.1	RP11-425I 13.3	intronic	T	0.028 4	9.29E-07	PRFP,OLG,PSO,SYS
rs14320084 4	4	1464248 29	4q31.21	SMAD1	intronic	A	0.015 5	7.82E-08	PRFN
rs18915730 4	4	1788762 65	4q34.3	RP11-389E 17.1	intronic	G	0.012 1	6.46E-07	ENT,PRFN,OLG,PSO,UND
rs11511200 2	5	1136878 22	5q22.3	KCNN2	8.8kb 5'	A	0.013 7	7.25E-07	ENT,PRFN,PRFP,OLG,SYS,UN D
rs11242863	6	3570361	6p25.2	RP1-223B1 .1	24kb 3'	A	0.383	9.43E-08	ENT,PRFN,PRFP,OLG,PSO,SYS
rs6918563	6	2562257 1	6p22.2	LRRC16A	1.8kb 3'	G	0.146	6.05E-07	PRFN,OLG,SYS
rs1555688	6	3383724 2	6p21.31	RP3-468B3 .3	3'	G	0.094 0	1.22E-07	ENT,PRFP,OLG,PSO,UND
rs11394194 6	6	1297188 76	6q22.33	LAMA2	intronic	T	0.029 5	4.55E-07	ENT,PRFN,PRFP,SYS,UND
rs15103906 5	6	1489644 77	6q24.3	SASH1	91kb 3'	G	0.058 6	6.35E-07	PRFN,PRFP,OLG,SYS
rs11814828 5	7	4816205 6	7p12.3	UPP1	14kb 3'	G	0.029 1	3.11E-07	ENT,PRFP,OLG,SYS
rs74608517	7	1077830 22	7q31.1	NRCAM	5.1kb 3'	T	0.020 7	5.09E-07	ENT,PRFN,PRFP,OLG,SYS
rs3112343	7	1353869 64	7q33	SLC13A4	intronic	G	0.040 5	2.05E-07	PRFP,OLG,PSO,SYS
rs14232152 1	8	1040899 20	8q22.3	ATP6V1C1	4.6kb 3'	G	0.013 6	5.63E-07	ENT,PRFN,OLG,SYS
rs628920	9	1359620 85	9q34.2	snoU13	6.4kb 5'	C	0.232	2.56E-07	PRFN,OLG,PSO,UND
rs7080552	10	5059369 5	10q11.23	DRGX	intronic	G	0.339	6.03E-07	ENT,PRFN,PRFP,PSO
rs73241417	13	7774115 8	13q22.3	MYCBP2	intronic	A	0.024 1	9.18E-07	SYS
rs2134802	13	8994035 4	13q31.2	LINC00440	7.9kb 3'	C	0.010 9	2.50E-07	SYS
rs16943576	15	9031624 9	15q26.1	MESP2	3'	G	0.050 4	1.65E-07	ENT,PRFN,PRFP,OLG,PSO,UN D

Table S5. The significant association between the GWS SNPs from our hsGWAS and the phenotypes in UK Biobank.

SNP	Phenotype	OR	95% CI	SE	Cases	P-value
rs12203592	Rheumatoid factor > 16 IU/mL	0.921	0.896, 0.947	0.0143	16700	8.792E-09
rs7731626	rheumatoid arthritis	0.886	0.851, 0.921	0.0201	5586	1.468E-09
rs7660520	Psoriasis	0.929	0.887, 0.974	0.0238	5074	1.468E-09
	Psoriasis arthropathy	0.859	0.770, 0.958	0.0558	954	6.50E-03
rs114664970	Chronic sinusitis	1.235	1.052, 1.451	0.0821	4393	1.00E-02

Table S6. The association of the GWS loci with JIA in previously reported immunochip data (Hinks et al., 2013).

hsGWAS				Immunochip					
SNP	CHR	POS	P_adj	SNP	CHR	POS	Pvalue	beta	snp_count
rs2066363	1	82237577	6.82E-11	rs6696329	1	82692386	2.58E-02	0.125	18
rs144844686	2	234576970	7.40E-09	rs11694654	2	234332343	8.46E-04	0.292	487
rs7636581	3	189781195	6.82E-11	rs4398410	3	189344011	3.37E-02	-0.118	11
rs13119493	4	180911259	3.26E-08	rs1458669	4	180747463	4.99E-02	-0.134	14
rs7660520	4	183745321	6.82E-11	rs6841520	4	183553379	3.61E-02	-0.260	12
rs7731626	5	55444683	4.62E-10	rs10065637	5	55438851	9.95E-06	-0.319	77
rs12203592	6	396321	4.62E-09	rs55683615	6	441043	4.05E-04	-0.194	193
rs114664970	6	40127169	1.03E-08	rs7753175	6	39863295	5.98E-02	-0.417	11
rs727845	7	67607209	4.49E-08	rs6953349	7	67728697	1.69E-01	0.255	9
rs7042370	9	12785073	1.39E-09	rs10122431	9	13137049	6.71E-02	-0.106	8
rs117572873	10	91997663	1.52E-08	rs1886409	10	91680365	3.19E-01	0.0583	6
rs17703382	11	30715593	2.76E-09	rs507969	11	30830612	2.47E-02	-0.130	15
rs12795402	11	32255936	4.19E-08	rs2207550	11	32310891	1.23E-02	0.136	11
rs147585949	16	20726695	1.98E-09	rs13339094	16	21097770	5.19E-03	-0.166	5
rs11663074	18	45023793	6.82E-11	rs2571034	18	44578565	2.21E-02	-0.126	13
rs138816451	22	43649657	1.63E-09	rs9623807	22	43704052	5.50E-02	0.168	17

SNP-count: The total number of SNPs at each locus in the immunochip dataset.

Table S7. The causal variants identified at each locus.

Locus	Causal variants	Chr	Pos	A1	Association P-value	Prob
3q28	rs7636581	3	189781195	G	2.32E-16	1
4q34.3	rs11131944	4	180687498	A	8.98E-06	0.307
6p21.2	rs114664970	6	40127169	T	1.10E-18	0.998
7q11.22	rs727845	7	67607209	A	2.82E-08	1
9p23	rs7042370	9	12785073	T	1.20E-11	1
10q23.31	rs146684573	10	92002571	C	4.80E-09	0.232
	rs144590246	10	92000291	G	5.58E-09	0.201
11p13	rs17703382	11	30715593	C	3.11E-25	0.992
16p12.3	rs147585949	16	20726695	G	3.52E-10	0.508
18q21.1	rs11663074	18	45023793	C	1.26E-26	1
22q13.2	rs138816451	22	43649657	G	3.82E-12	1

Chr=chromosome; Pos=position; A1=alternative allele; Association P-value=P-value in the original hsGWAS; Prob=the posterior probability.

Table S8. Significant association between JIA subtypes and imputed classical HLA-A, HLA-C, HLA-B, HLA-DRB1, HLADQB1, HLA-DQA and HLA-DPB1 alleles. The HLA class I and HLA class II alleles were analyzed at a 2-digit level based on the 2×2 table. Since a total of 116 HLA alleles were tested, the significance threshold was set as P-value < 4.31e-04.

Enthesitis-Related Arthritis (ENT)				
SNP	Freq_case	Freq_control	OR	P
HLA_B_27	0.29	0.037	10.56	3.91E-112
HLA_C_01	0.20	0.032	7.38	4.63E-57
HLA_DRB1_08	0.12	0.027	4.74	8.10E-22
HLA_DQB1_04	0.11	0.026	4.56	2.36E-19
HLA_C_02	0.13	0.041	3.58	5.05E-16
HLA_DRB1_01	0.18	0.11	1.76	0.0001138
Polyarthritis RF Negative (PRFN)				
SNP	Freq_case	Freq_control	OR	P
HLA_DQB1_04	0.13	0.026	5.51	1.92E-46
HLA_DRB1_08	0.13	0.027	5.22	2.44E-43
HLA_B_40	0.11	0.054	2.27	5.16E-10
HLA_DRB1_07	0.056	0.14	0.381	4.41E-08
HLA_B_27	0.079	0.037	2.238	2.44E-07
HLA_C_01	0.070	0.032	2.3	4.18E-07
HLA_DQB1_02	0.14	0.22	0.580	7.60E-06
HLA_A_02	0.36	0.28	1.481	8.72E-06
HLA_C_03	0.17	0.11	1.633	1.39E-05
HLA_DPB1_02	0.21	0.15	1.567	1.59E-05
HLA_C_06	0.049	0.098	0.476	9.97E-05
Polyarthritis RF Positive (PRFP)				
SNP	Freq_case	Freq_control	OR	P
HLA_DRB1_04	0.36	0.15	3.063	2.84E-09
HLA_B_08	0.23	0.11	2.557	1.53E-05
HLA_DPB1_26	0.0089	0.00043	20.82	4.55E-05
HLA_C_03	0.23	0.11	2.397	6.12E-05
HLA_DRB1_07	0.0089	0.14	0.0575	9.13E-05
Oligoarthritis (OLG)				
SNP	Freq_case	Freq_control	OR	P
HLA_DRB1_08	0.17	0.027	7.01	3.48E-112
HLA_DQB1_04	0.16	0.026	6.937	5.23E-106
HLA_A_02	0.42	0.28	1.863	5.52E-20
HLA_DPB1_02	0.26	0.15	2.007	1.51E-19
HLA_B_40	0.12	0.054	2.315	1.24E-15
HLA_DRB1_07	0.045	0.14	0.297	1.49E-15
HLA_DRB1_04	0.070	0.15	0.412	3.19E-12
HLA_C_03	0.18	0.11	1.72	7.97E-10
HLA_C_06	0.041	0.098	0.395	9.46E-09
HLA_B_39	0.045	0.019	2.384	1.07E-07
HLA_DRB1_13	0.16	0.11	1.604	2.49E-07
HLA_DQB1_02	0.15	0.22	0.632	9.93E-07
HLA_B_27	0.067	0.037	1.879	2.95E-06
HLA_DQB1_03	0.29	0.36	0.734	2.74E-05
HLA_DRB1_15	0.090	0.14	0.623	4.39E-05
HLA_B_57	0.014	0.038	0.362	0.000173
Psoriatic Arthritis (PSO)				
SNP	Freq_case	Freq_control	OR	P
HLA_DQB1_04	0.085	0.026	3.483	1.21E-06
HLA_DRB1_08	0.080	0.027	3.061	3.01E-05
HLA_DPB1_19	0.028	0.0056	5.223	6.96E-05
HLA_B_27	0.091	0.037	2.601	0.000182
HLA_C_02	0.097	0.041	2.485	0.000259
HLA_A_02	0.40	0.28	1.732	0.000330
Undifferentiated Arthritis (UND)				
SNP	Freq_case	Freq_control	OR	P
HLA_DQB1_04	0.11	0.026	4.582	1.42E-11
HLA_DRB1_08	0.11	0.027	4.341	9.16E-11
HLA_B_27	0.092	0.037	2.634	0.000148

Freq_case= The frequency of HLA alleles in JIA cases of the particular subtype; Freq_control= the frequency of HLA alleles in controls of the cohort; OR= odds ratio.

Table S9. Pathways over-presented among the candidate genes at genome-wide significant loci.

Gene Set	Description	Size	Overlap	Enrichment Ratio	P-Value	FDR	Overlap genes
hsa04659	Th17 cell differentiation	107	4	33.98	2.97E-06	9.67E-04	SMAD2;IRF4;IL1RAP;IL6ST
hsa04550	Signaling pathways regulating pluripotency of stem cells	139	2	13.08	0.00942	1	SMAD2;IL6ST
hsa04060	Cytokine-cytokine receptor interaction	294	2	6.18	0.0388	1	IL1RAP;IL6ST

Size=the total number of genes in each pathway; Overlap=the number of genes overlapped between the input gene list and the pathway genes

Table S10. GO pathways over-presented among the candidate genes at genome-wide significant loci.

Gene Set	Description	Size	Overlap	Enrichment Ratio	P-Value	FDR	Overlap genes
GO:0032616	interleukin-13 production	24	2	94.92	1.94E-04	0.165	IRF4;IL1RAP
GO:0032633	interleukin-4 production	34	2	67	3.93E-04	0.167	IRF4;IL1RAP
GO:0032623	interleukin-2 production	63	2	36.16	0.00135	0.318	IRF4;IL1RAP
GO:0098742	cell-cell adhesion via plasma-membrane adhesion molecules	270	3	12.66	0.00150	0.318	IL1RAP;CLDN1;TENM3
GO:0042107	cytokine metabolic process	106	2	21.49	0.00376	0.639	IRF4;IL1RAP
GO:0001819	positive regulation of cytokine production	418	3	8.17	0.00517	0.733	IRF4;IL1RAP;IL6ST
GO:0050769	positive regulation of neurogenesis	447	3	7.64	0.00624	0.757	KIF20B;TENM3;IL6ST
GO:0006091	generation of precursor metabolites and energy	469	3	7.29	0.00712	0.757	ACSM1;TYRP1;IL6ST
GO:0034341	response to interferon-gamma	192	2	11.86	0.0119	0.943	IRF4;CLDN1
GO:0050803	regulation of synapse structure or activity	221	2	10.31	0.0155	0.943	LRFN2;IL1RAP

Table S11. The candidate genes at genome-wide significant loci which are drug target genes.

SNP	Gene	Drug
rs7660520	DCTD	Cytarabine, gemcitabine
rs114664970	KIF6	pravastatin, atorvastatin, rosuvastatin, simvastatin,
rs144844686	UGT1A8	Mycophenolate mofetil, Diflunisal, Naproxen, Nintedanib, methotrexate, cyclosporine, sirolimus, tacrolimus
	ACKR3	plerixafor
rs144844687	TRPM8	levomenthol
	UGT1A7	Mycophenolate mofetil, Diflunisal, Naproxen, Nintedanib, methotrexate
rs117572873	HTR7	hmg coa reductase inhibitors, atorvastatin, pravastatin, simvastatin
rs147585951	PLK1	Fostamatinib
rs7636581	IL1RAP	IL-1 blockade
rs138816451	SCUBE1	Angiogenesis Inhibitors

Fig. S1. The quantile–quantile (QQ) plot of the hsGWAS of JIA cases vs. the controls.

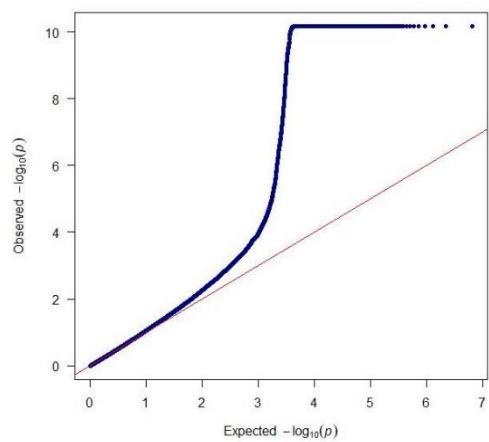


Fig. S2 Regional association plots of genetic loci previously reported to be associated with JIA or other autoimmune diseases.

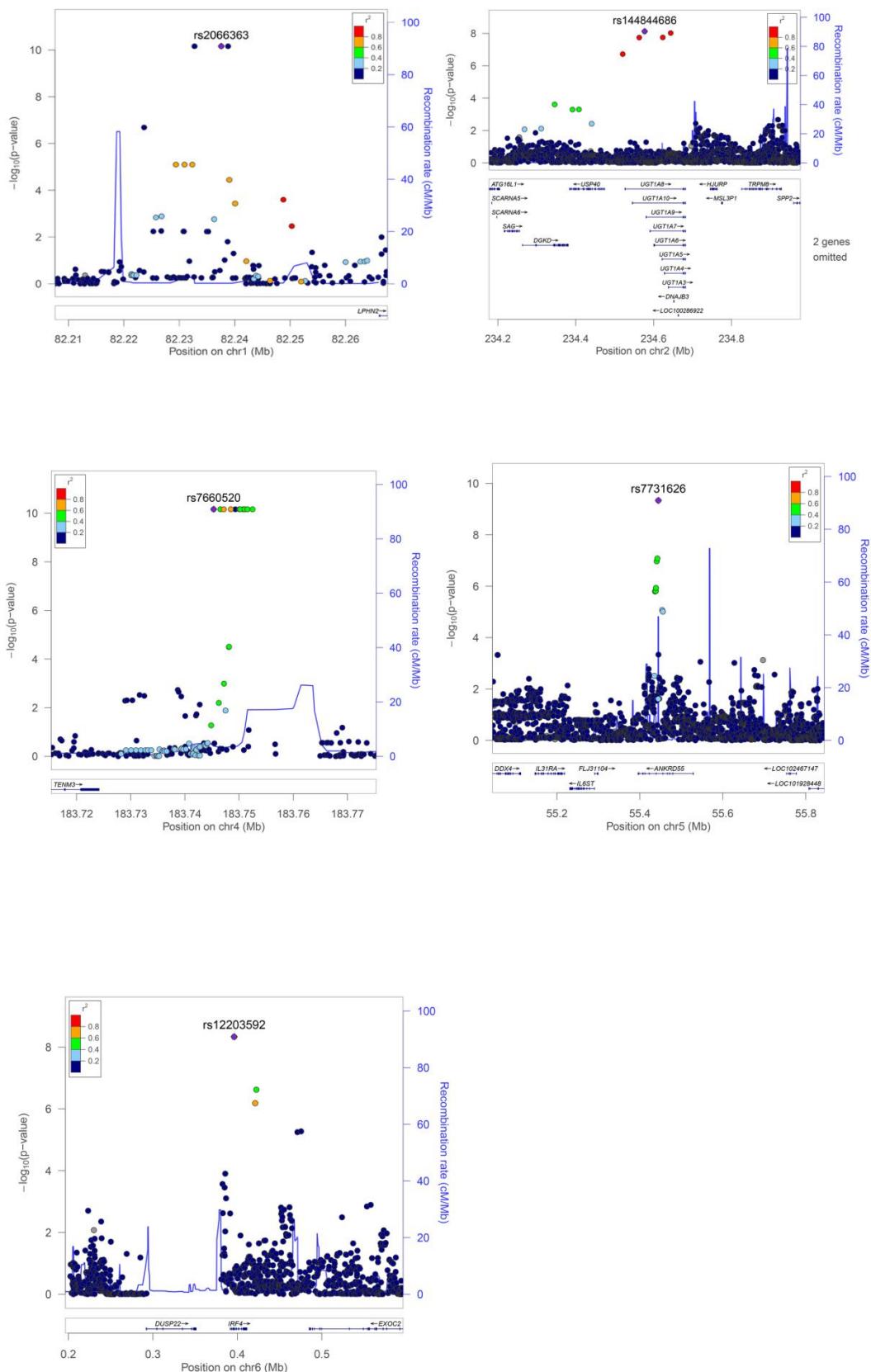


Fig. S3. Heatmap showing the association significance of each genome-wide significant SNP with each individual JIA subtype.

SNP	Chr	Pos	ENT	PRFN	PRFP	OLG	PSO	SYS	UND	P_adj
rs2066363	1	82237577								6.82E-11
rs144844686	2	234576970								7.40E-09
rs7636581	3	189781195								6.82E-11
rs13119493	4	180911259								3.26E-08
rs7660520	4	183745321								6.82E-11
rs7731626	5	55444683								4.62E-10
rs12203592	6	396321								4.62E-09
rs114664970	6	40127169								1.03E-08
rs727845	7	67607209								4.49E-08
rs7042370	9	12785073								1.39E-09
rs117572873	10	91997663								1.52E-08
rs12795402	11	32255936								4.19E-08
rs17703382	11	30715593								2.76E-09
rs147585949	16	20726695								1.98E-09
rs11663074	18	45023793								6.82E-11
rs138816451	22	43649657								1.63E-09

Fig. S4. Annotation of the GWS and GMS JIA loci with hsGWAS.P<1E-06. For each SNP, grey color indicates no annotation in a particular functional category and other colors indicate reported annotations including functional, conserved, regulatory, epigenomic annotation categories.

Abbreviations: cpg: CPG islands; dgv: copy number variable regions; dnase: DNase hypersensitivity I sites; eqtl: expression quantitative trait loci; gad: known genetic associations; gerp: conserved positions; mir: miRNAs; sift: functional mutations in SIFT; tfbs: transcription factor binding sites.



Fig. S5. The violin plot showing the significant correlation between the SNP genotype and the mRNA level of the gene in immune tissues based on GTex data.

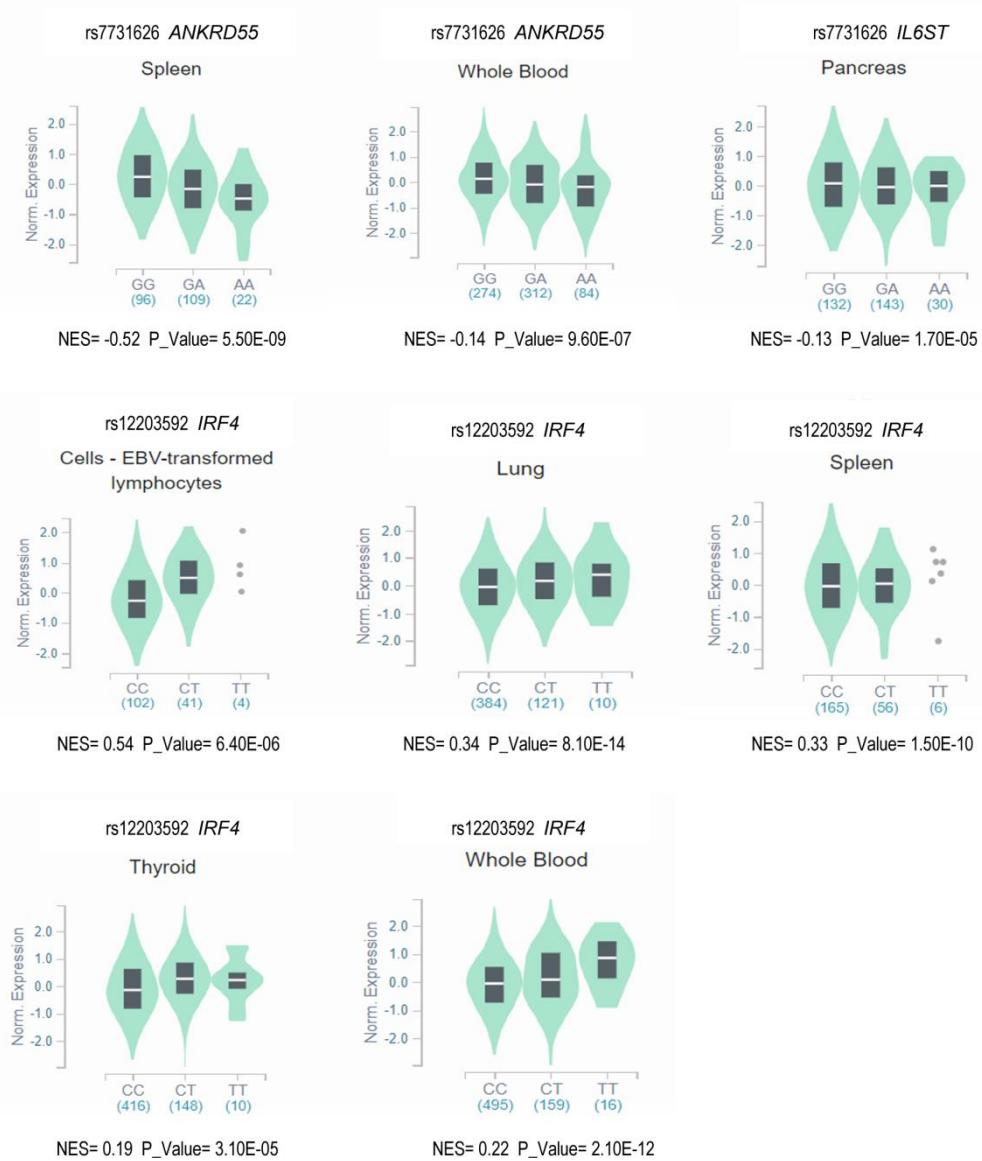
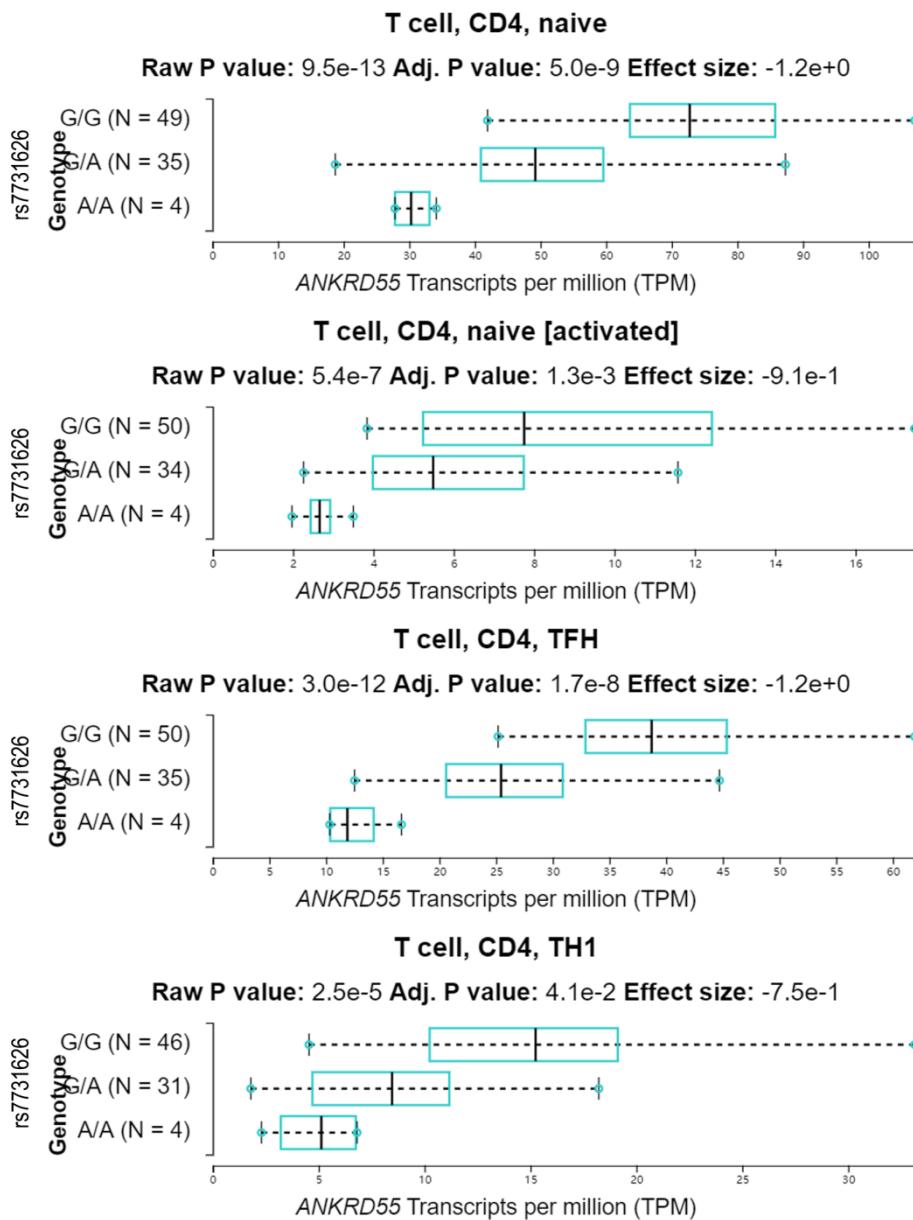
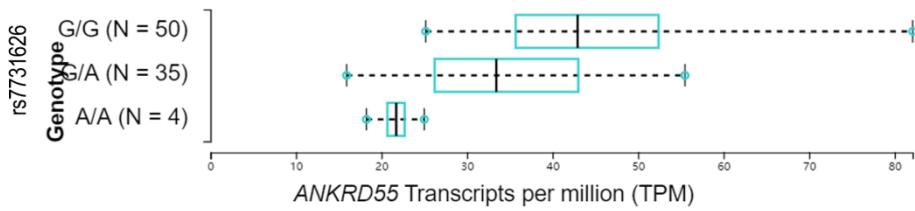


Fig. S6. Box plots showing the significant correlation between the SNP genotype and the mRNA level of the gene in immune cell types based on DICE data.



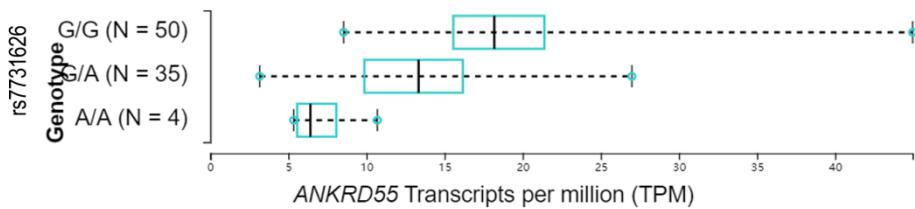
T cell, CD4, TH2

Raw P value: 8.6e-7 Adj. P value: 1.6e-3 Effect size: -8.9e-1



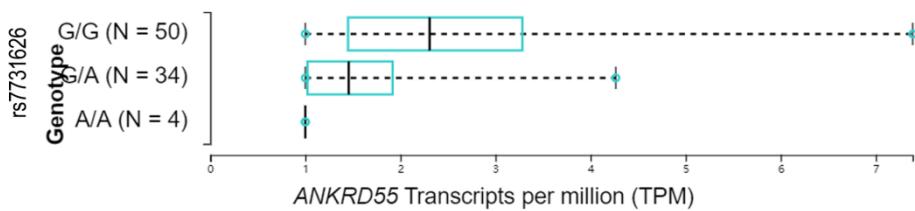
T cell, CD4, TH17

Raw P value: 1.8e-6 Adj. P value: 2.9e-3 Effect size: -8.2e-1



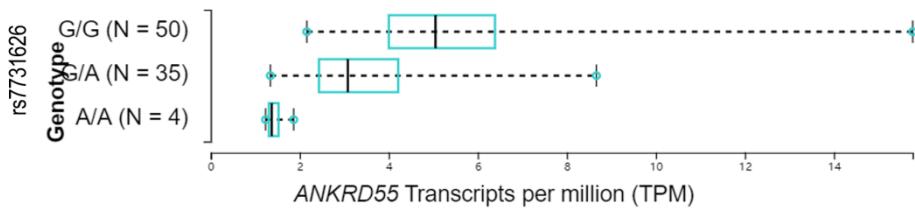
T cell, CD4, TH1/17

Raw P value: 5.3e-5 Adj. P value: 5.2e-2 Effect size: -7.4e-1



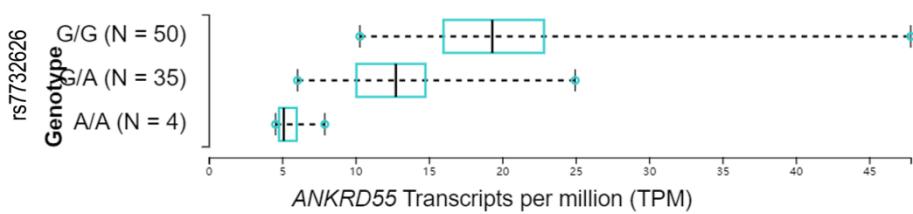
T cell, CD4, memory TREG

Raw P value: 3.0e-8 Adj. P value: 9.7e-5 Effect size: -9.8e-1



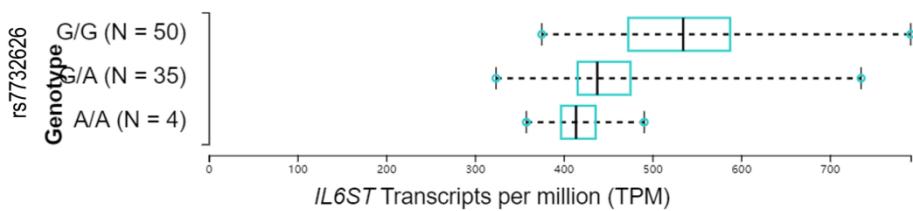
T cell, CD4, naive TREG

Raw P value: 5.3e-14 **Adj. P value:** 3.1e-10 **Effect size:** -1.3e+0



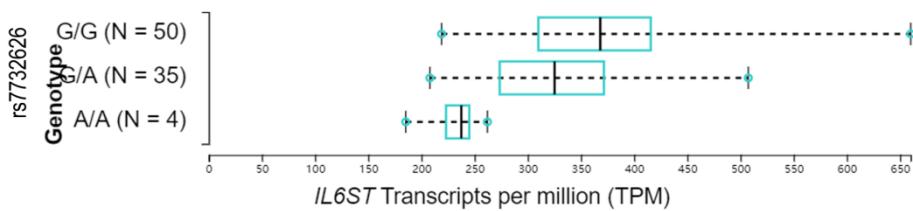
T cell, CD8, naive

Raw P value: 3.0e-5 **Adj. P value:** 3.3e-2 **Effect size:** -7.4e-1



T cell, CD4, TFH

Raw P value: 4.1e-5 **Adj. P value:** 4.5e-2 **Effect size:** -7.6e-1



T cell, CD4, naive TREG

Raw P value: 4.3e-6 **Adj. P value:** 6.4e-3 **Effect size:** -8.1e-1

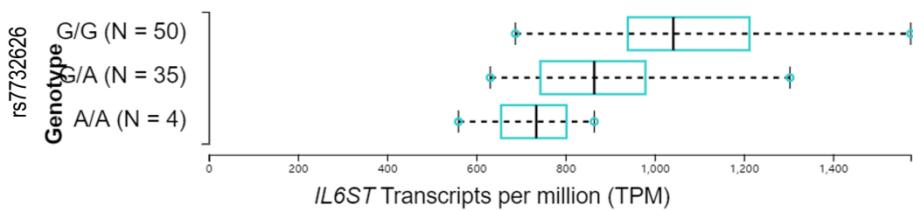
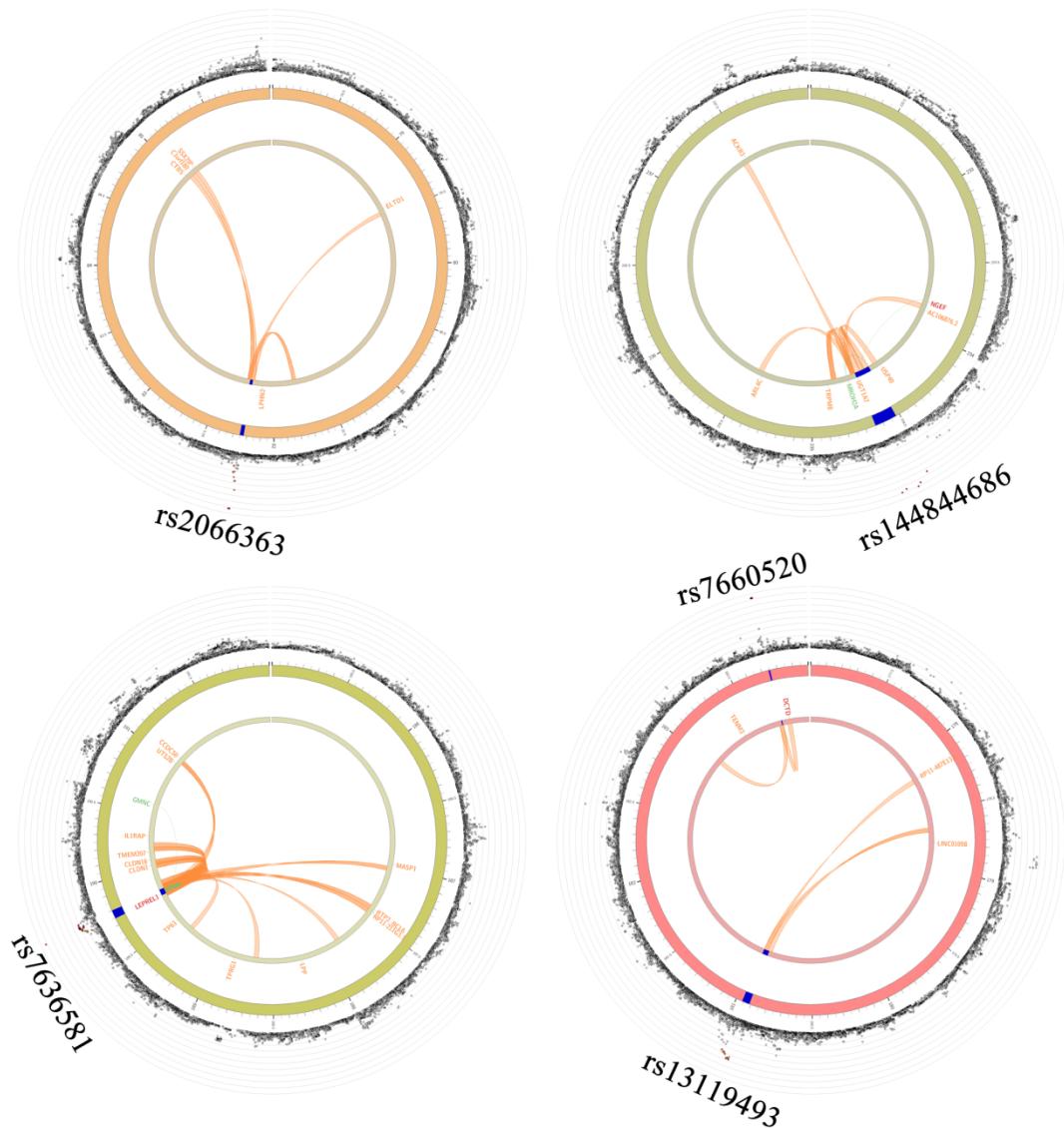
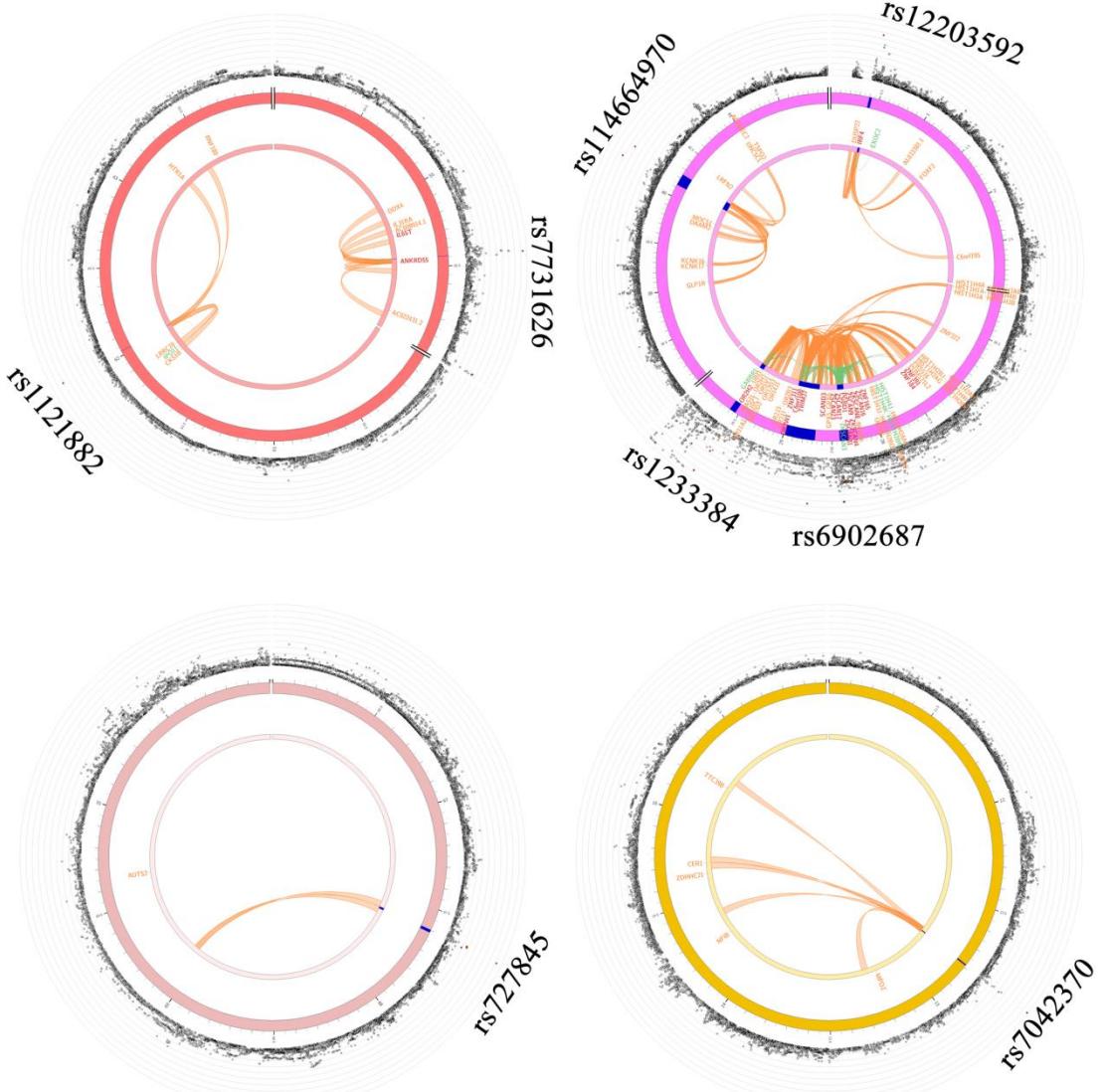
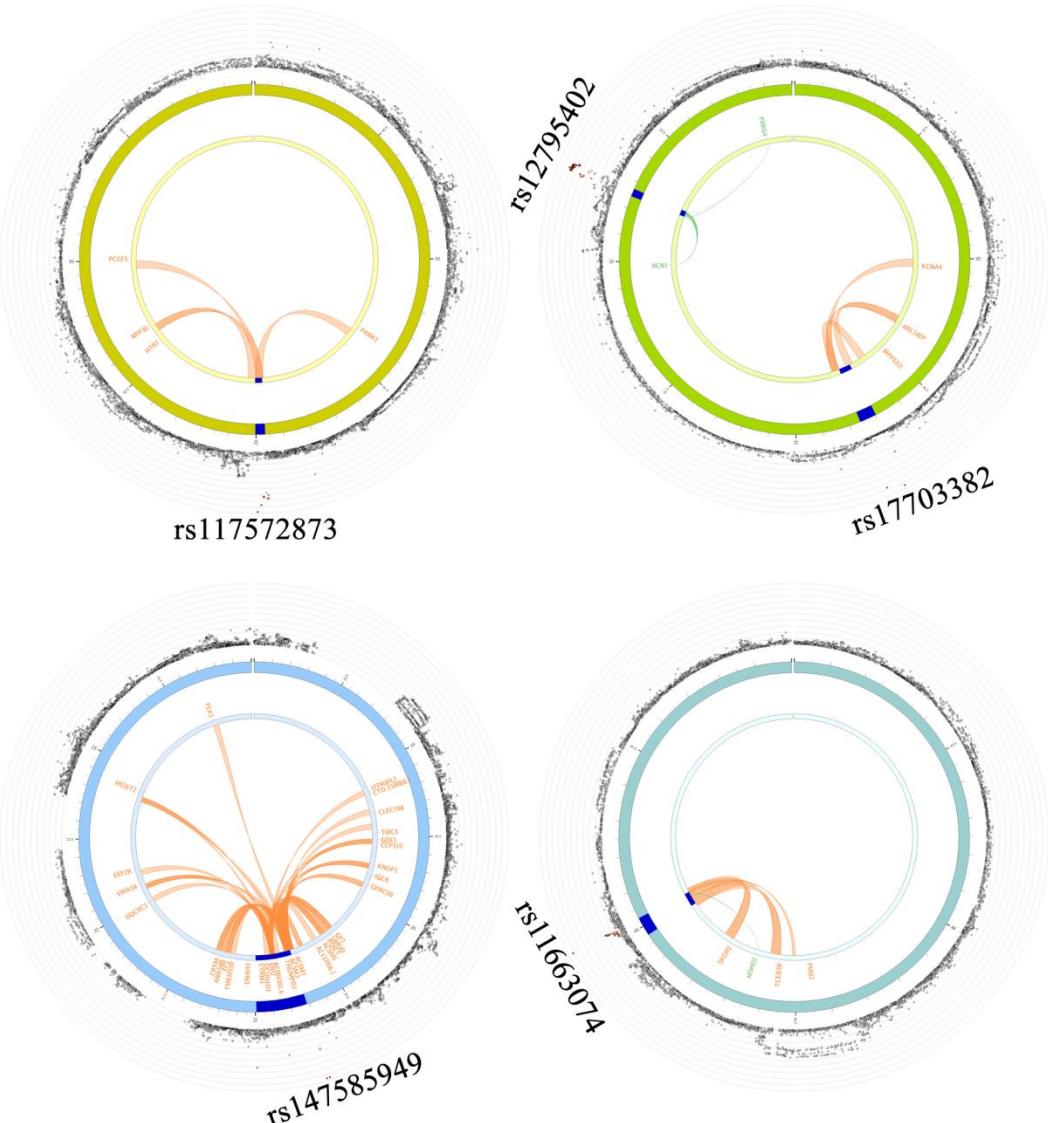


Fig. S7. The circus plots showing the chromatin interactions between each JIA locus and other chromatin regions.







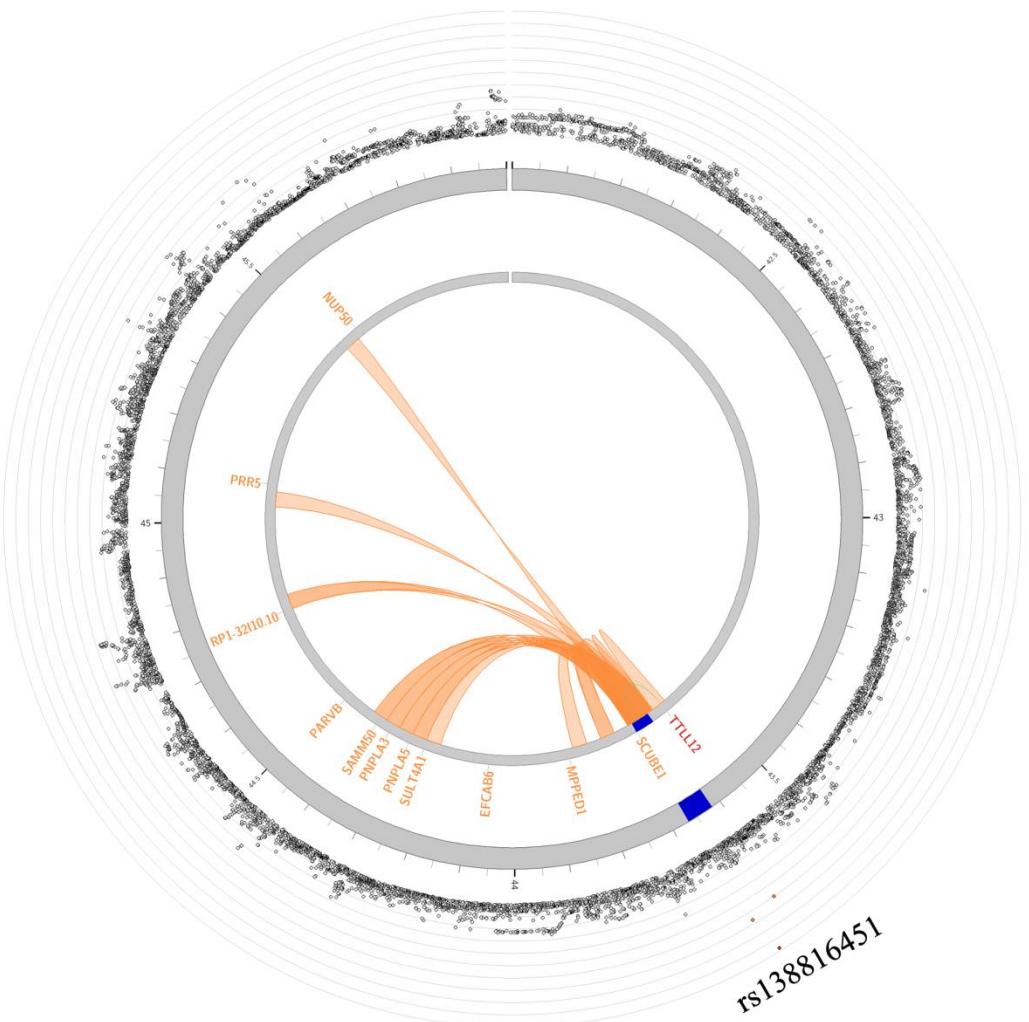
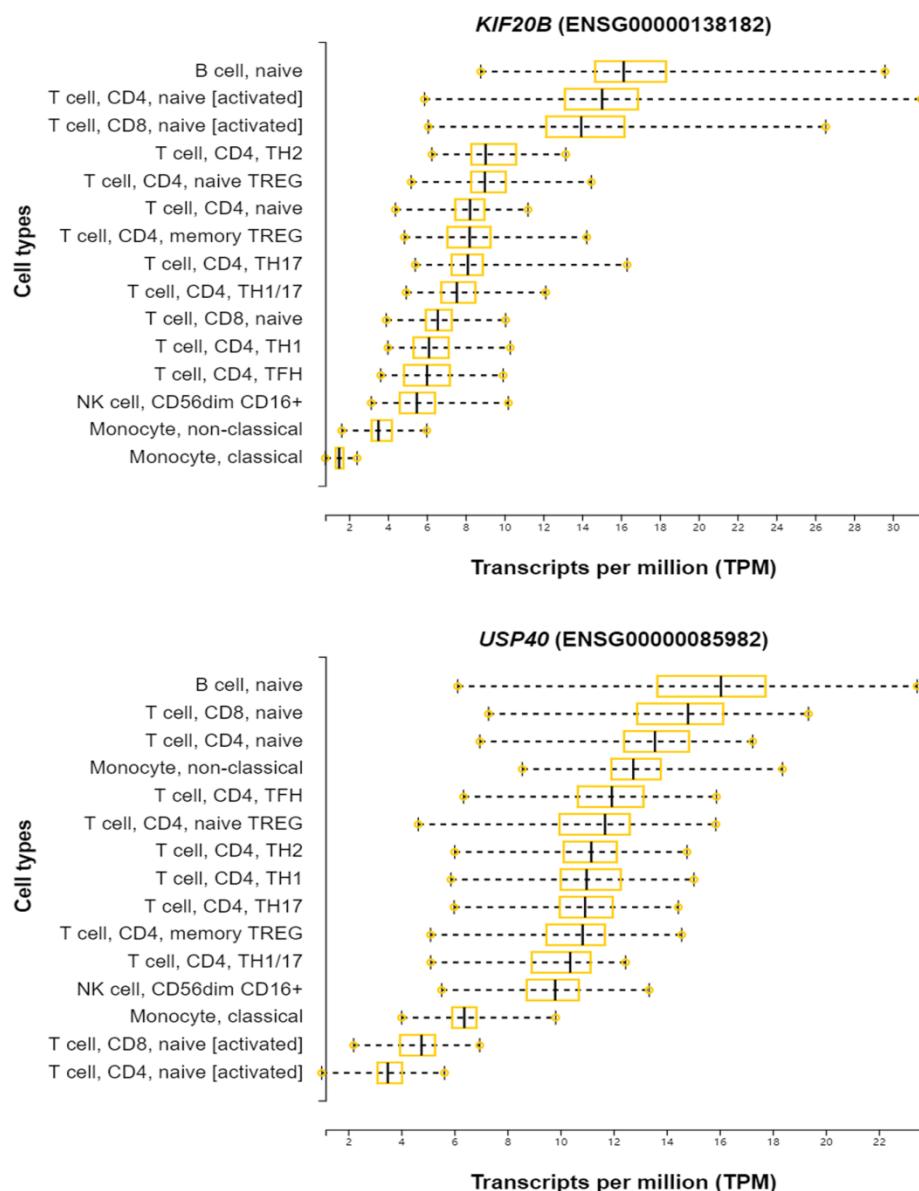
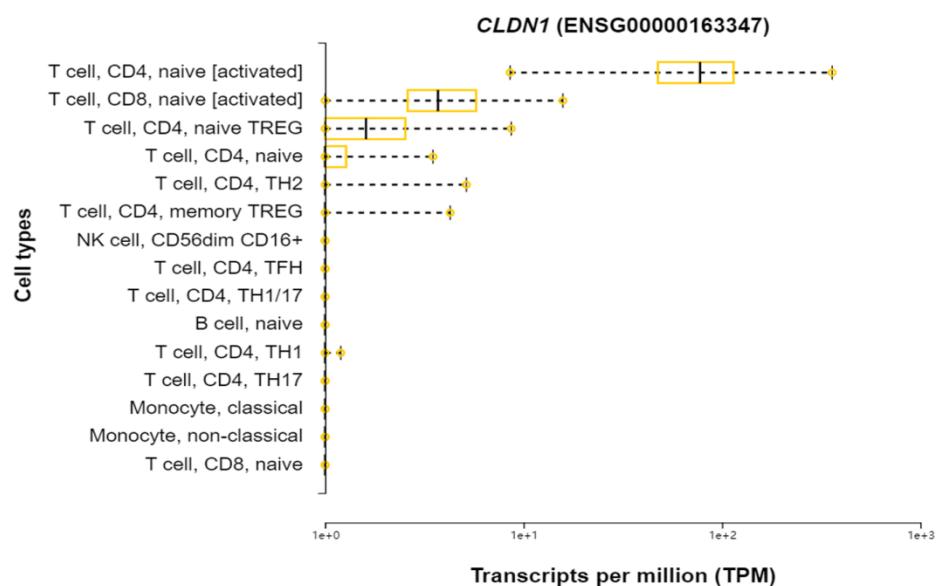
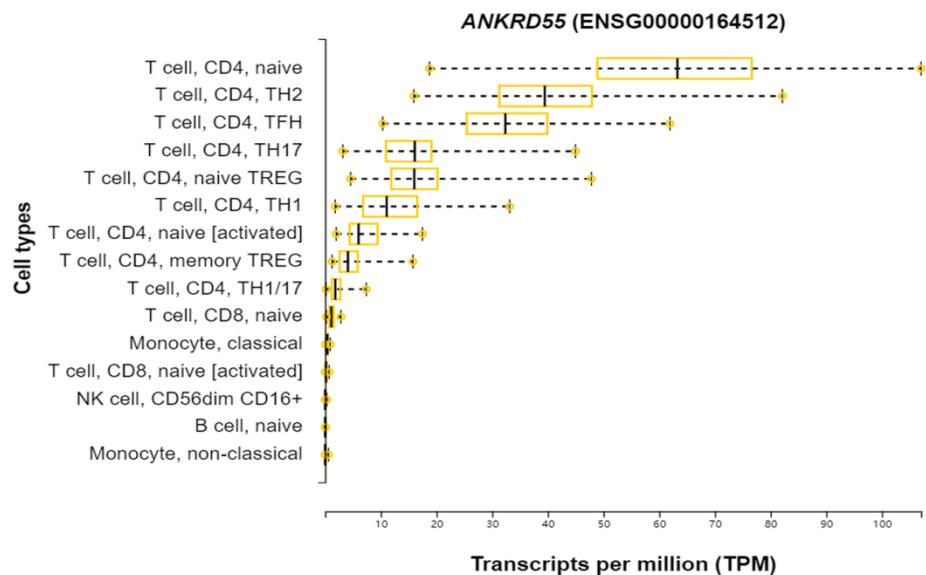
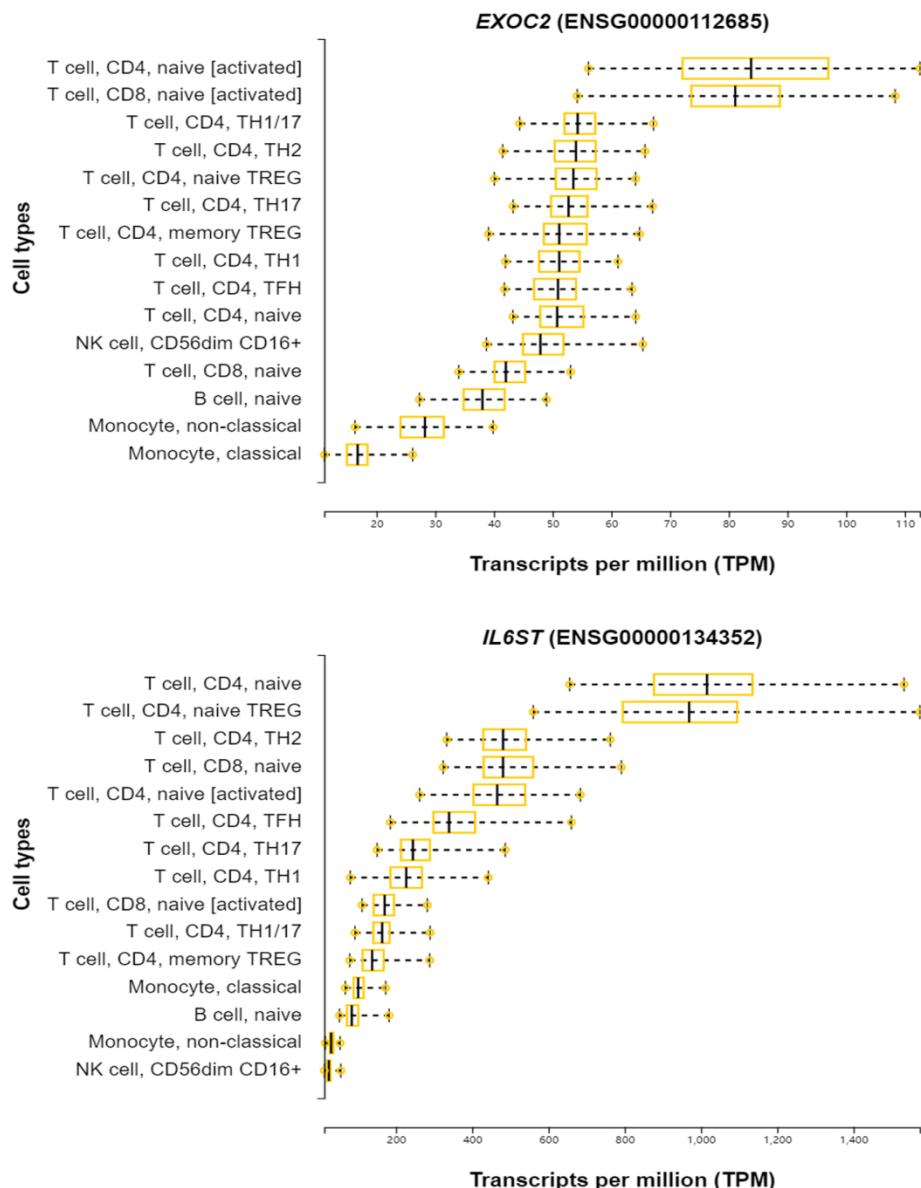
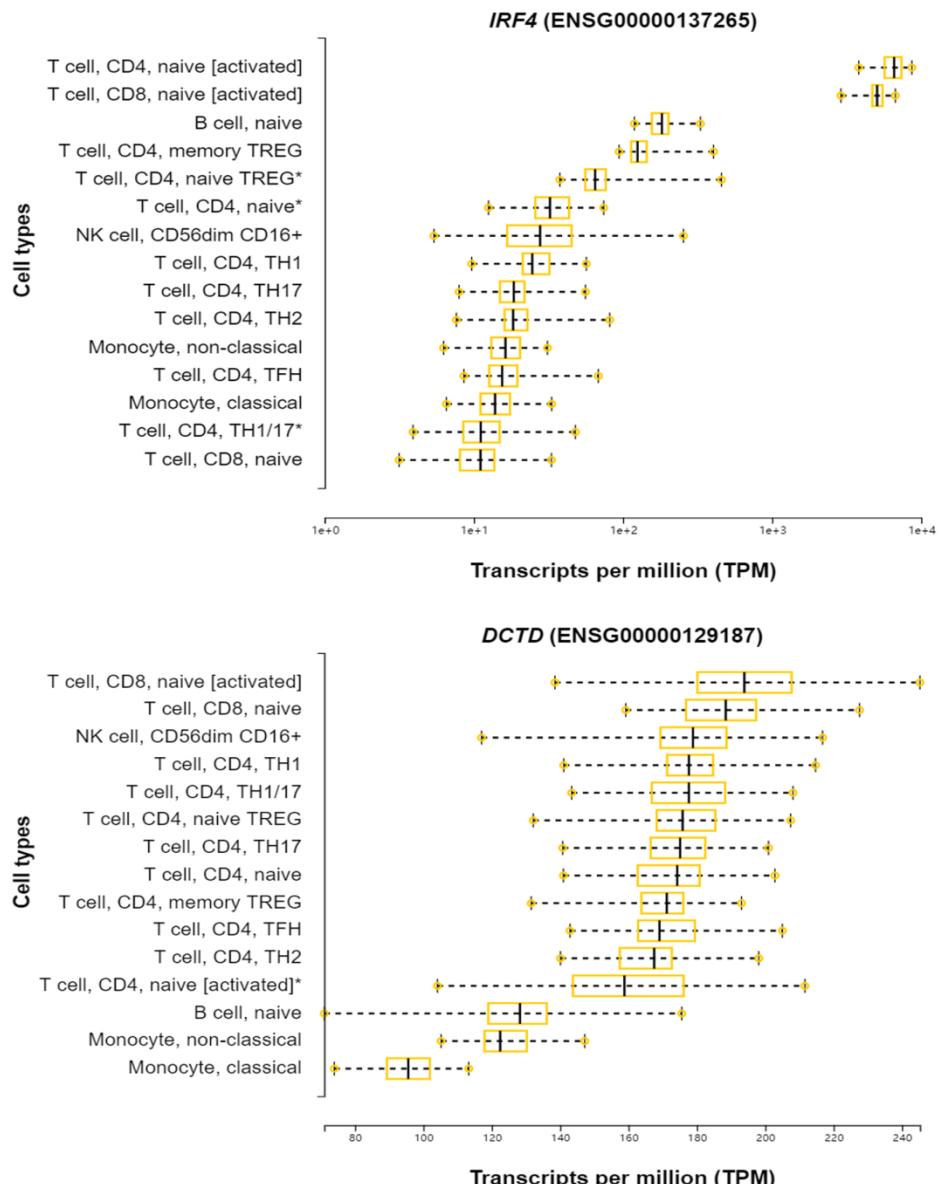


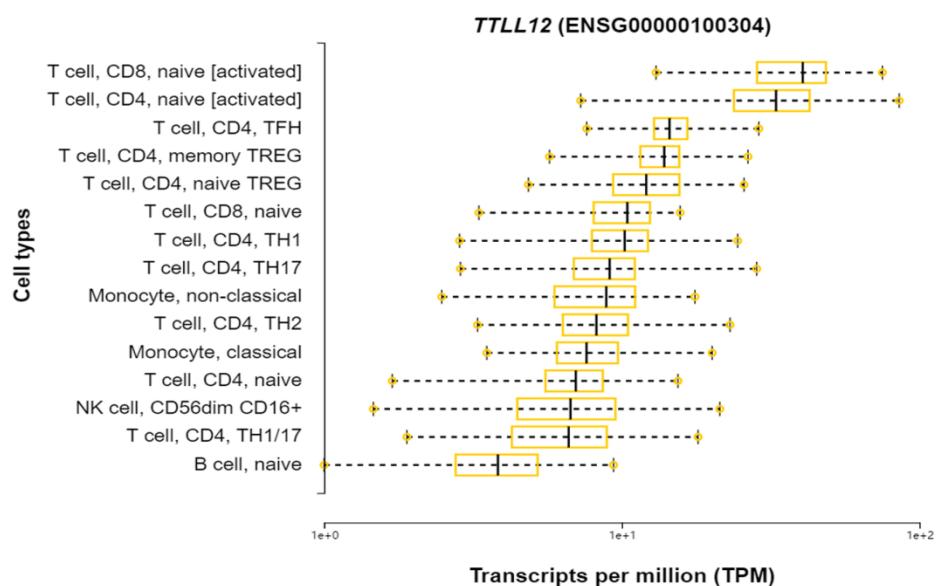
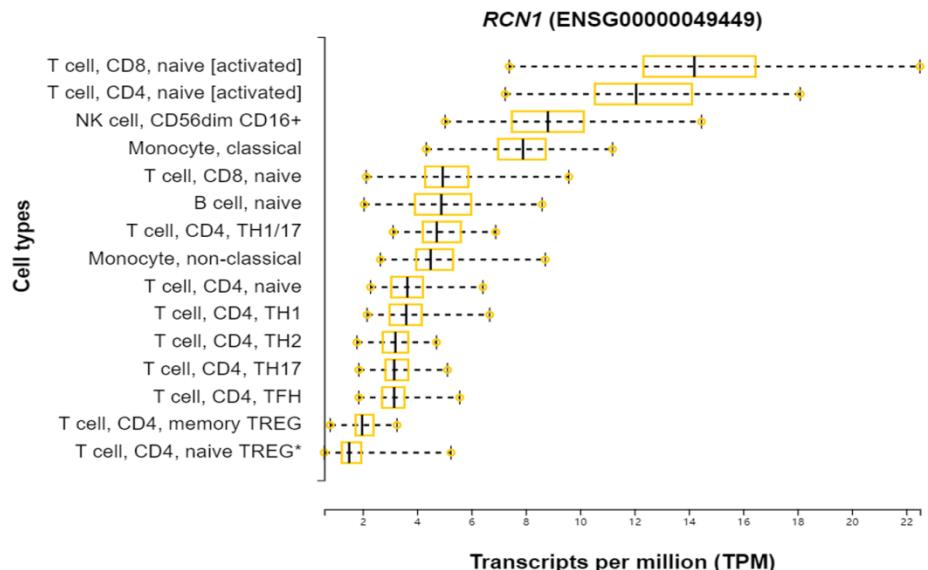
Fig. S8. Candidate genes at genome-wide significant loci are expressed at medium to high level in immune cell types.

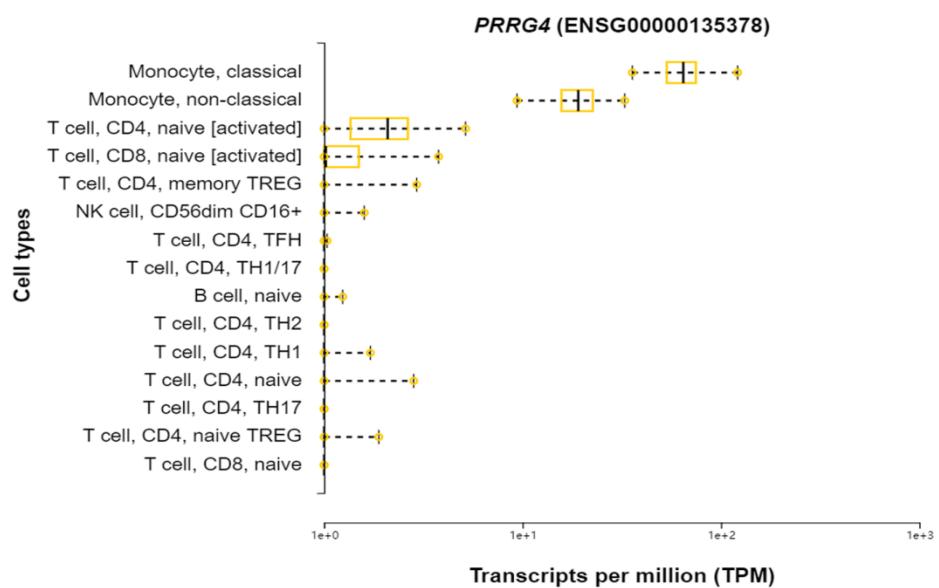
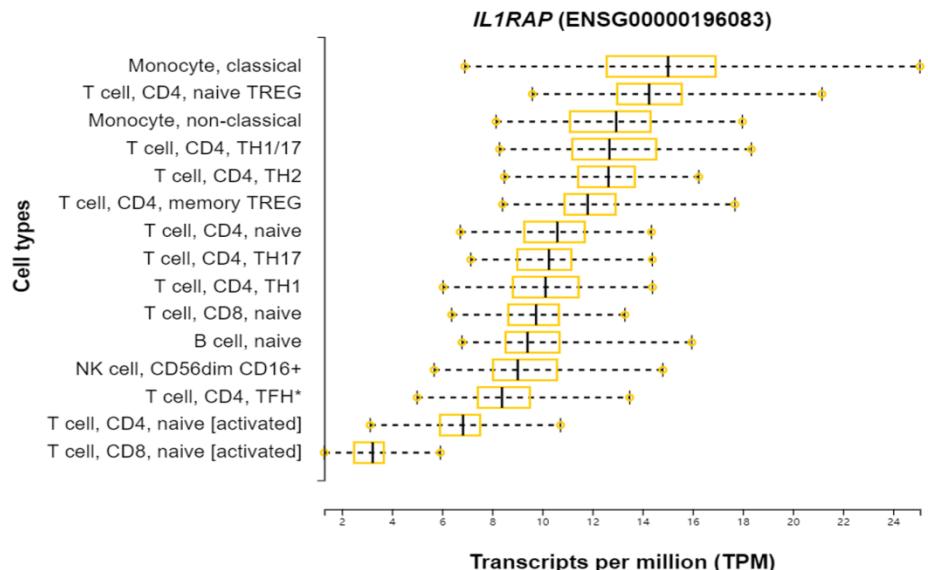












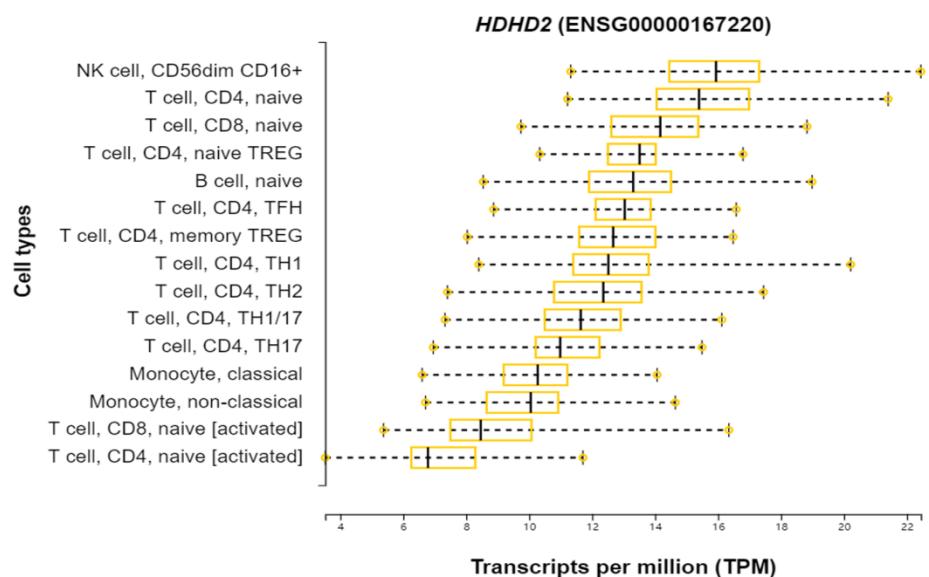
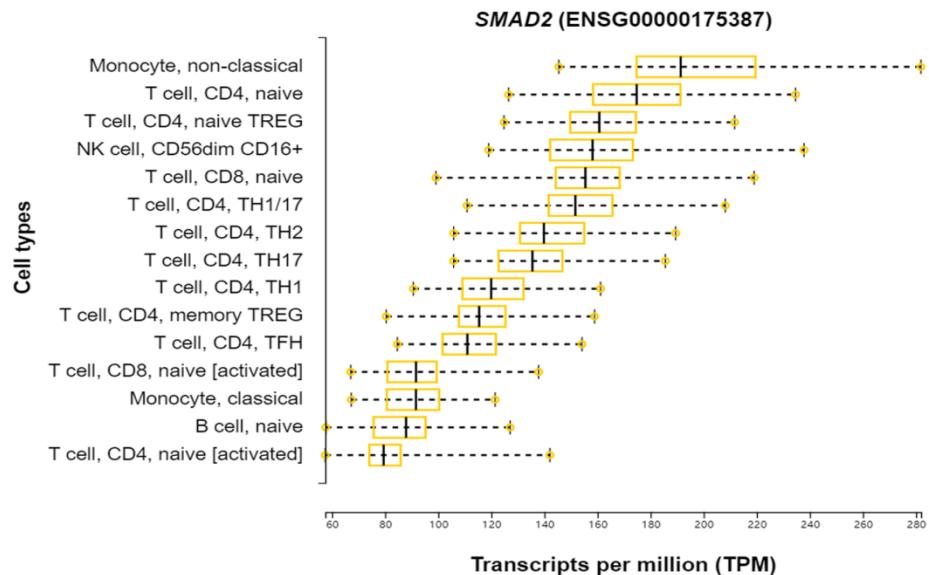


Fig. S9. PPI network constructed based on our JIA hsGWAS results.

