Functional Annotations of Common Disease Markers in Immune Regulatory Genes

Sandeep K. Singh*, Yongjun Huang*, Allen Caobi*,
Manuel A. Barbieri*, Mehmet Tevfik Dorak**

*Dept of Biological Sciences, Florida International University, Miami, FL, U.S.A.
**School of Health Sciences, Liverpool Hope University, Liverpool, U.K.

BACKGROUND

A number of polymorphisms in immune system genes show associations with multiple diseases.

We have started a systematic study to functionally annotate disease-associated immunogenetic polymorphisms.

This is the initial report of our results on selected *PTPN22*, *CTLA4*, *FCGR2A*, and *LTA*/*TNF* SNPs that have shown numerous disease associations.

BACKGROUND

Gene	SNP ID	Popular name	Associations
PTPN22	rs2476601	R620W	Multiple autoimmune diseases: T1D; RA; vitiligo; Crohn disease
CTLA4	rs3087243	CT60	RA; T1D; selective IgA deficiency; Graves disease, Hashimoto thyroiditis; myasthenia gravis
CTLA4	rs231775	+49G>A; T17A	Multiple cancers
FCGR2A	rs1801274	H131R	Tx outcome, SLE, malaria, HIV control, phagocytic function; IBD; ulcerative colitis; Kawasaki disease; SLE
TNF	rs361525	-238G>A	Flucloxacillin-induced liver injury; proxies: RA, psoriasis, HIV control
TNF	rs1800629	-308G>A	SLE; triglyceride levels; proxies: myasthenia gravis, idiopathic membranous nephropathy, RA, T1D & laryngeal squamous cell carcinoma
LTA	rs1799964	-857C>T	Crohn disease; aGVHD; proxies: RA, height, HIV control
LTA	rs909253 (~rs1800629)	IVS1 +90A>G; NcoI	Proxies: Idiopathic membranous nephropathy; RA

Liverpool Hope University EST. 1844

Pheno Scanner

A database of human genotype-phenotype associations

PhenoScanner, the new name for MR Catalogue

METHODS

Functional annotations were done using ANNOVAR, RegulomeDB, SNPnexus, PheGenI, SNiPA, rVarBase, mQTLdb and several expression quantitative trait loci (eQTL) browsers.

We also examined GWAS associations of these SNPs and their proxies on GRASP and PhenoScanner (formerly MR Catalogue).

METHODS

We also examined relationships of HLA region SNPs to HLA haplotypes in our ImmunoChip data on 95 HLA-typed IHWG cell lines.

RESULTS

None of the algorithms predicted a damaging effect of the missense SNPs in *PTPN22*, *CTLA4* or *FCGR2A* on proteins.

Gene	SNPID	SIFT	PP2HVAR	PP2HDIV	Mutation taster	Mutation assesor	LRT	FATHMM	metaSVM	metaLR
PTPN22	rs2476601	1.00, T	0.000, B	0.000, B	0.000, P	.	0.000, N	4.64, T	4.64, T	0.004, T
FCGR2A	rs1801274	0.28, T	0.002, B	0.000, B	0.520, P	1.355, L	0.000, N	2.78, T	2.78, T	0.008, T
CTLA4	rs231775	0.09, T	0.006, B	0.015, B	1.000, P	1.100, L	0.072, N	1.73, T	1.73, T	0.000, T

All were tolerated, benign, polymorphism -as opposed to mutation-, or non-functional.

When assessed by RegulomeDB scores for their regulatory function (including eQTL effects), no SNP in LD with the lead SNPs appeared to have higher functionality.

Liverpool Hope University EST. 1844

RESULTS

PTPN22 **missense SNP is within an H3K27ac histone mark, it alters transcription factor binding sites, and has a RegulomeDB score of 2b (also a CADD score of 16.8) suggesting high regulatory activity (only one SNP in LD, and not in a coding region).**

Thus, its associations with autoimmune disorders (T1D, RA, GD, MG, CD, IgAD, and vitiligo) should not be attributed to its damaging effect on the PTPN22 protein, but rather to the effect on ***PTPN22*** **expression, which appeared to be different in peripheral blood (PB) and PB mononuclear cells (PBMC).**

RESULTS

***TNF* promoter SNP rs361525 had no eQTL effect on *TNF* or *LTA* expression.**

***TNF* rs1800629 showed no eQTL effect on *TNF* or *LTA* in PBMC despite having many other eQTL targets, although a negative eQTL effect on *TNF* and *LTA* was apparent in PB in a large meta-analysis.**

***LTA* rs909253 showed a negative correlation with *TNF* expression in peripheral blood, but not in PBMC.**

RESULTS

	SNP	p.valu	beta	cistra	gene_chro	HGNC
2	rs1800629	1.31E-34	-0.184056629	cis	6	HLA-C
3	rs1800629	8.31E-32	0.081530089	cis	6	VARS2
4	rs1800629	7.47E-08	-0.033836119	cis	6	C2
5	rs1800629	7.02E-07	-0.030614309	cis	6	HLA-DRA
6	rs1800629	1.44E-06	-0.033205168	cis	6	TUBB
7	rs1800629	3.22E-05	-0.024176841	cis	6	MRPS18B
8	rs1800629	4.35E-05	-0.01970222	cis	6	CLIC1
9	rs1800629	5.48E-05	0.015031985	cis	6	APOM
10	rs1800629	6.20E-05	0.018880316	cis	6	NEU1
11	rs1800629	0.0002563	-0.01675752	cis	6	CFB
12	rs1800629	0.0002694	-0.019602256	cis	6	DDAH2
13	rs1800629	0.0002898	0.021688845	cis	6	TUBB
14	rs1800629	0.0005064	-0.024362038	cis	6	FLOT1
15	rs1800629	0.0020118	0.015046126	cis	6	DDR1

SNP	p.value	beta	cistrans	ene_chror	HGNC
rs1800629	2.30E-38	0.2185398	trans	6	BTN3A2
rs1800629	1.27E-24	0.2386027	trans	6	HLA-DQA1
rs1800629	8.69E-21	0.2354405	trans	6	HLA-DQA1
rs1800629	2.49E-19	0.1018103	trans	6	HLA-G
rs1800629	3.34E-19	0.083739	trans	6	BTN3A2
rs1800629	5.71E-15	-0.110709	trans	6	HLA-DPB1
rs1800629	6.53E-15	-0.088106	trans	6	PSMB9
rs1800629	1.26E-14	-0.068734	trans	6	PSMB9
rs1800629	2.58E-14	0.6810286	trans	6	HLA-DRB1
rs1800629	1.86E-10	0.1488955	trans	6	HLA-H
rs1800629	2.53E-10	-0.067704	trans	4	TMEM154
rs1800629	4.55E-10	-0.08317	trans	2	RSAD2
rs1800629	1.60E-09	-0.111784	trans	1	ISG15
rs1800629	3.64E-09	-0.094129	trans	1	IFI44L
rs1800629	6.31E-09	-0.195856	trans	1	IFI44L
rs1800629	1.03E-08	-0.087807	trans	8	LY6E
rs1800629	2.07E-08	-0.027194	trans	20	RNF114
rs1800629	2.58E-08	-0.10658	trans	10	IFIT3
rs1800629	5.47E-08	-0.110155	trans	1	IFI6
rs1800629	5.62E-08	-0.050945	trans	2	SPATS2L
rs1800629	5.79E-08	-0.079711	trans	17	XAF1
rs1800629	6.11E-08	-0.121206	trans	1	IFI44
rs1800629	6.24E-08	-0.077291	trans	12	OASL

Human Molecular Genetics, 2015, Vol. 24, No. 16 4746–4763
doi: 10.1093/hmg/ddv194
Advance Access Publication Date: 27 May 2015
Original Article

ORIGINAL ARTICLE

Dissecting the genetics of the human transcriptome identifies novel trait-related *trans*-eQTLs and corroborates the regulatory relevance of non-protein coding loci†

Holger Kirsten[1,2,7], Hoor Al-Hasani[3,8,10], Lesca Holdt[11], Arnd Gross[1,2], Frank Beutner[2,4], Knut Krohn[5], Katrin Horn[1,2], Peter Ahnert[1,2], Ralph Burkhardt[2,6], Kristin Reiche[3,9,10], Jörg Hackermüller[3,9,10], Markus Löffler[1,2], Daniel Teupser[11], Joachim Thiery[2,6] and Markus Scholz[1,2,*]

Liverpool Hope University EST. 1844

RESULTS

Blood eQTL browser

Query eQTL Results

Or, you can query the *cis*- and *trans*-eQTLs below (examples: rs7807018 or *VWCE*):

Gene or SNP name: rs1800629 Search

Your query: rs1800629

Trans-eQTLs

P-value SNP SNP Chr. SNP Chr. position Probe Probe Chr. Probe Chr. position SNP Alleles Minor Allele Z-score Gene name FDR

No records found.

Cis-eQTLs

P-value	SNP	SNP Chr.	SNP Chr. Position	Probe	Probe Chr.	Probe Chr. position	SNP Alleles	Minor Allele	Z-score	Gene name	FDR
3.1897075995926226E-17	rs18006296		31651010	1440603	6	31556059	G/A	A	-8.44	-	0.00
5.714159755300914E-12	rs18006296		31651010	580452	6	31558934	G/A	A	6.89	-	0.00
1.276417429398821E-7	rs18006296		31651010	2640301	6	31653892	G/A	A	-5.28	TNF	0.00
2.29134823035681E-7	rs18006296		31651010	3450156	6	31745428	G/A	A	5.17	CSNK2B	0.00
0.0013099087919970723	rs18006296		31651010	1030743	6	31650043	G/A	A	-3.21	LTA	0.29

Systematic identification of *trans*-eQTLs as putative drivers of known disease associations

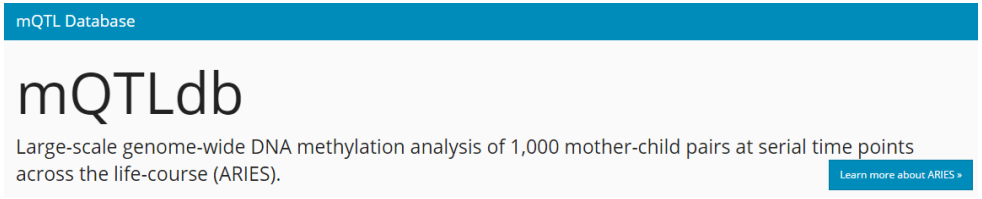
Harm-Jan Westra[1,48], Marjolein J Peters[2,3,48], Tõnu Esko[4,48], Hanieh Yaghootkar[5,48], Claudia Schurmann[6,48], Johannes Kettunen[7,8,48], Mark W Christiansen[9,48], Benjamin P Fairfax[10,11], Katharina Schramm[12,13], Joseph E Powell[14,15], Alexandra Zhernakova[1], Daria V Zhernakova[1], Jan H Veldink[16], Leonard H Van den Berg[16], Juha Karjalainen[1], Sebo Withoff[1], André G Uitterlinden[2,3,17], Albert Hofman[3,17], Fernando Rivadeneira[2,3,17], Peter A C 't Hoen[18], Eva Reinmaa[4], Krista Fischer[4], Mari Nelis[4], Lili Milani[4], David Melzer[19], Luigi Ferrucci[20], Andrew B Singleton[21], Dena G Hernandez[21,22], Michael A Nalls[21], Georg Homuth[6], Matthias Nauck[23], Dörte Radke[24], Uwe Völker[6], Markus Perola[4,9], Veikko Salomaa[9], Jennifer Brody[9], Astrid Suchy-Dicey[25], Sina A Gharib[26], Daniel A Enquobahrie[25], Thomas Lumley[27], Grant W Montgomery[28], Seiko Makino[29], Holger Prokisch[12,13], Christian Herder[24], Michael Roden[24-31], Harald Grallert[32], Thomas Meitinger[12,13,33,34], Konstantin Strauch[35,36], Yang Li[37], Ritsert C Jansen[17], Peter M Visscher[14,15], Julian C Knight[10], Bruce M Psaty[8,38,39], Samuli Ripatti[7,9,39,41], Alexander Teumer[6,42], Timothy M Frayling[5,42], Andres Metspalu[4,42], Joyce B J van Meurs[2,3,43] & Lude Franke[1,43]

Liverpool Hope University EST. 1844

RESULTS

Screening of mQTLdb yielded the most interesting results. The two *TNF/LTA* SNPs were very strong trans-mQTLs for CpG sites on different chromosomes.

Timepoint	SNP	SNP Chr	SNP Pos	CpG	CpG Chr	CpG Pos	beta	Effect Size	p-value	Trans
Adolescence	rs1800629	6	31543031	**cg12441246**	8	1972838	-0.45556	0.07825	1.20E-23	Y
Birth	rs1800629	6	31543031	**cg12441246**	8	1972838	-0.45075	0.08573	1.13E-21	Y
Pregnancy	rs1800629	6	31543031	cg15196058	4	130934318	-0.3795	0.01958	1.16E-18	Y
Childhood	rs1800629	6	31543031	cg15196058	4	130934318	-0.34496	0.03597	8.36E-17	Y
Childhood	rs1800629	6	31543031	**cg12441246**	8	1972838	-0.36969	0.07965	8.73E-17	Y
Childhood	rs909253	6	31540313	cg26036029	6	32552443	0.33895	0.04938	3.06E-16	Y
Adolescence	rs1800629	6	31543031	cg15196058	4	130934318	-0.37722	0.01947	9.18E-16	Y
Childhood	rs909253	6	31540313	cg20240154	6	31464980	-0.33101	0.02266	2.04E-15	N
Pregnancy	rs1800629	6	31543031	**cg12441246**	8	1972838	-0.36474	0.05744	2.39E-15	Y
Middle Age	rs1800629	6	31543031	cg15196058	4	130934318	-0.34467	0.02655	2.59E-15	Y
Birth	rs909253	6	31540313	cg10536999	7	26193109	-0.34467	0.0311	6.00E-15	Y
Birth	rs1800629	6	31543031	cg15196058	4	130934318	-0.37785	0.02851	9.89E-15	Y
Adolescence	rs1800629	6	31543031	cg12736254	6	32557419	0.43479	0.03736	2.06E-14	Y

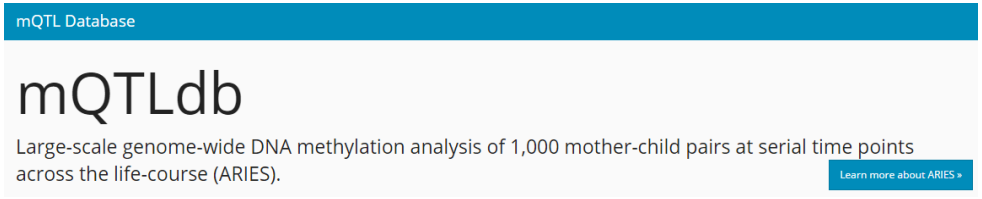


Liverpool Hope University EST. 1844

RESULTS

Screening of mQTLdb yielded the most interesting results. The two *TNF/LTA* SNPs were very strong trans-mQTLs for CpG sites on different chromosomes.

Timepoint	SNP	SNP Chr	SNP Pos	CpG	CpG Chr	CpG Pos	Target Genes
Adolescence	rs1800629	6	31543031	cg12441246	8	1972838	KBTBD11/LOC101928058
Birth	rs1800629	6	31543031	cg12441246	8	1972838	KBTBD11/LOC101928058
Pregnancy	rs1800629	6	31543031	cg15196058	4	130934318	LOC105377421-LOC105377422 (nearest gene)
Childhood	rs1800629	6	31543031	cg15196058	4	130934318	LOC105377421-LOC105377422 (nearest gene)
Childhood	rs1800629	6	31543031	cg12441246	8	1972838	KBTBD11/LOC101928058
Childhood	rs909253	6	31540313	cg26036029	6	32552443	HLA-DRB6 (nearest gene)
Adolescence	rs1800629	6	31543031	cg15196058	4	130934318	LOC105377421-LOC105377422
Childhood	rs909253	6	31540313	cg20240154	6	31464980	HCP5
Pregnancy	rs1800629	6	31543031	cg12441246	8	1972838	KBTBD11/LOC101928058
Middle Age	rs1800629	6	31543031	cg15196058	4	130934318	LOC105377421-LOC105377422
Birth	rs909253	6	31540313	cg10536999	7	26193109	HNRNPA2B1
Birth	rs1800629	6	31543031	cg15196058	4	130934318	LOC105377421-LOC105377422
Adolescence	rs1800629	6	31543031	cg12736254	6	32557419	HLA-DRB6



Liverpool Hope University EST. 1844

RESULTS

In ImmunoChip data, *TNF* rs1800629 (-308A) was present in ancestral haplotype (AH) 8.1 and AH 58.1.
TNF rs361525 (-238A) was present in AH 18.2, 18.3 and 57.1.

IHWG-ID	CELL LINE	ST	AH	TNF rs1800629_308	HLA-A 1	HLA-A 2	HLA-B 1	HLA-B 2	HLA-C 1	HLA-C 2	HLA-DRB1 1	HLA-DRB1 2	HLA-DQA1 1	HLA-DQA1 2	HLA-DQB1 1	HLA-DQB1 2
9022	COX	52	8.1	2	A*0101	A*0101	B*0801	B*0801	C*0701	C*0701	DRB1*0301	DRB1*0301	DQA1*05010	DQA1*05010	DQB1*0201	DQB1*0201
9063	WT47	52	.	2	A*32010	A*32010	B*4402	B*4402	C*0501	C*0501	DRB1*1302	DRB1*1302	DQA1*01020	DQA1*01020	DQB1*0604	DQB1*0604
9070	LUY	08	.	2	A*02:01	A*02:01	B*51:01	B*51:01	C*14:02	C*14:02	DRB1*08:03	DRB1*08:03	DQA1*04:01	DQA1*06:01	DQB1*03:01	DQB1*03:01
9088	PF04015	52	8.1	2	A*0101	A*0101	B*0801	B*0801	C*0701	C*0701	DRB1*0301	DRB1*0301	DQA1*05010	DQA1*05010	DQB1*0201	DQB1*0201
9157	HAU, ML	52	58.1	2	A*33	A*33	B*5801	B*58	C*0302	C*0302	DRB1*0301	DRB1*0301	DQA1*0501	DQA1*0501	DQB1*0201	DQB1*0201

Liverpool Hope University EST. 1844

CONCLUSIONS

PTPN22 rs2476601 : eQTL

CTLA4 rs3087243 : eQTL

CTLA4 rs231775 : eQTL

FCGR2A rs1801274 : eQTL

LTA/*TNF* rs361525 : eQTL

LTA/*TNF* rs1800629 : eQTL, trans-meQTL

LTA/*TNF* rs909253 : eQTL, trans-meQTL

CONCLUSIONS

> Even missense SNPs are more likely to modify disease susceptibility via their effects on gene expression rather than on protein structure

> *TNF/LTA* SNPs frequently associated with disease risk actually target CpG sites on different chromosomes rather than their own host genes

> Assumptions on regulatory functions based on proximity to a gene may not be verified experimentally

YOUR FUTURE
STARTS WITH HOPE

Liverpool Hope University EST. 1844