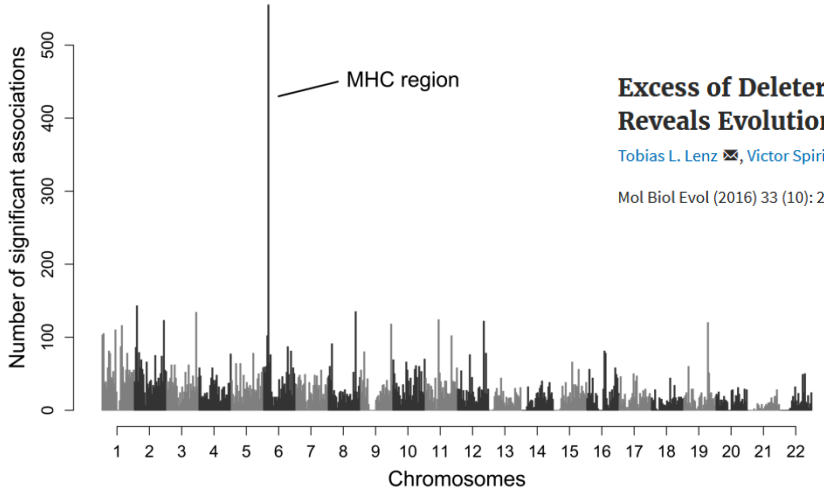
SNPs Correlating with Functional HLA Region Epitopes and Supertypes Offer Insight into GWAS Associations

Umut Ozbek[1], Amy E. Kennedy[2], Sandeep K. Singh[3], M. Alejandro Barbieri[3], Ioanna Konidari[4], Jabob L. McCauley[4], Mehmet Tevfik Dorak[5]

[1] Icahn School of Medicine at Mount Sinai, New York, NY, USA
[2] National Cancer Institute, Bethesda, MD, USA
[3] Florida International University, Miami, FL, USA
[4] University of Miami, John P. Hussman Institute for Human Genomics, Miami, FL, USA
[5] School of Health Sciences, Liverpool Hope University, Liverpool, UK

BACKGROUND

Single-nucleotide polymorphisms (SNPs) across the HLA region often reveal statistically significant associations in genome-wide association studies (GWAS), but most of these associated SNPs do not correspond to classical HLA alleles, even in immune-mediated disorders.



Excess of Deleterious Mutations around HLA Genes Reveals Evolutionary Cost of Balancing Selection

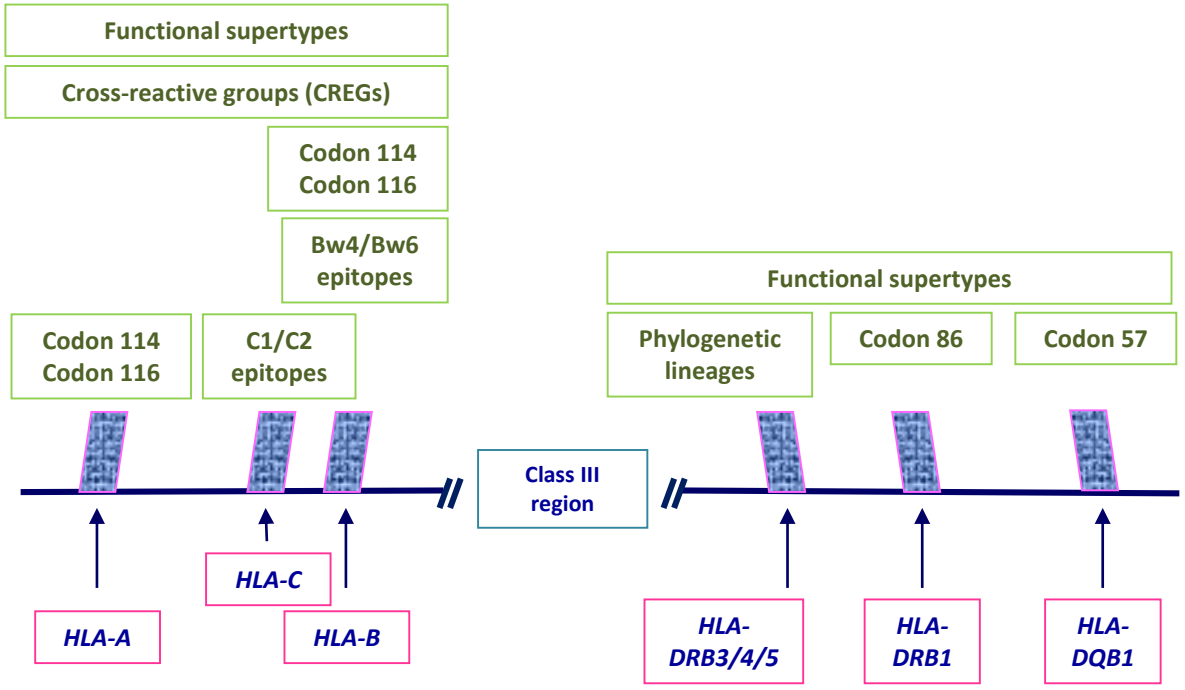
Tobias L. Lenz ✉, Victor Spirin, Daniel M. Jordan, Shamil R. Sunyaev ✉

Mol Biol Evol (2016) 33 (10): 2555-2564. **DOI:** https://doi.org/10.1093/molbev/msw127

Another level of functional HLA region variation (epitopes/supertypes) is relevant in disease associations, but not attracted much attention in GWAS.

Liverpool Hope University EST. 1844

BACKGROUND



Functional HLA polymorphisms

BACKGROUND

HLA-Bw4 and Bw6 Epitopes

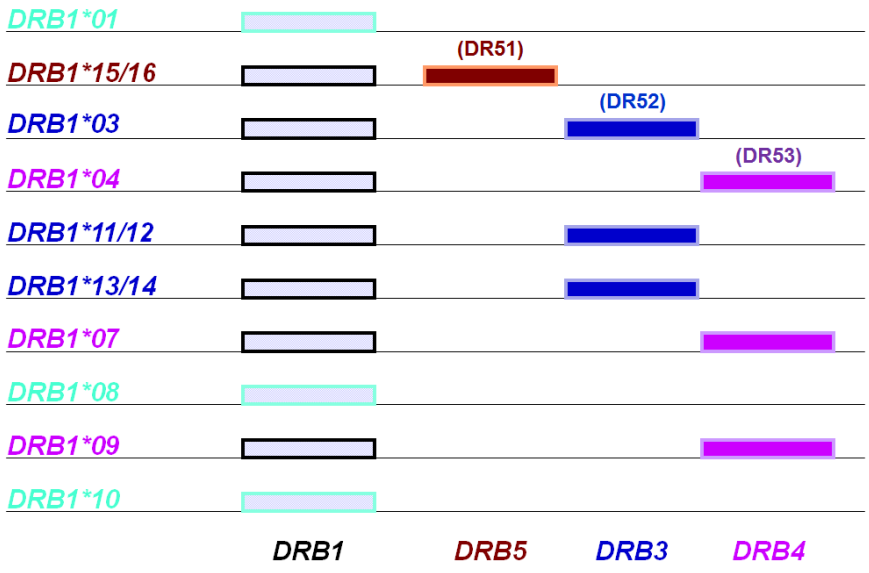
HLA-Bw4

B5, B5102, B5103, B13, B17, B27, B37, B38(16), B44(12), B47, 49(21), B51/52(5), B53, B57/58(17), B59, B63/77(15)

HLA-Bw6

B7, B703, B8, B14, B18, B22, B2708, B35, B39(16), B3901, B3902, B40, B4005, B41, B42, B45(12), B46, B48, B50(21), B54/55/56(22), B60(40), B61(40), B62(15), B64/65(14), B67, B70, B71/72(70), B73, B75/76(15), B78, B81/82

HLA-DR51, DR52, DR53 Supertypes



Liverpool Hope University EST. 1844

AIM

As the initial step of a systematic search for proxies for HLA region functional polymorphisms rather than individual HLA alleles,
we examined correlations between SNPs and

- **common HLA epitopes (Bw4/Bw6; C1/C2)**
- **genetic supertypes (DR51/DR52/DR53)**

in a panel of 95 HLA-typed IHWG cell lines

METHODS

Immunochip

(8,045 SNPs from 95 IHWG cell lines that passed quality control)

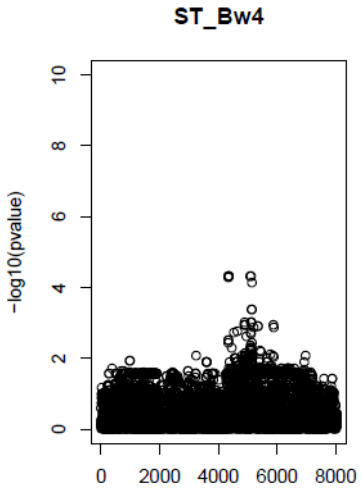
Proxy SNPs were searched for each epitope/supertype

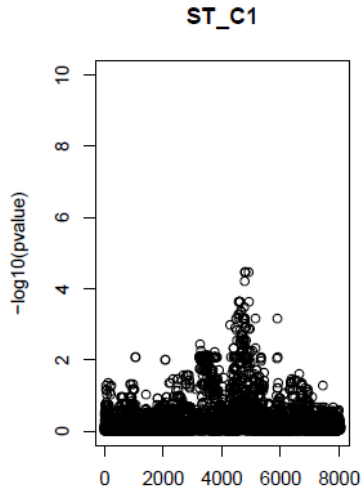
(r>0.50 (P<0.0001) by simple correlation tests for one-to-one correlation, and haplotype associations for multi-SNP correlation)

GRASP and PhenoScanner

(GWAS associations of SNPs that were most significantly correlated with HLA epitopes/supertypes)

RESULTS: HLA-Bw4/Bw6





Epitope	SNP	r	Associations
HLA-Bw4	rs2442719	- 0.570	Psoriasis (*P* = 3E-13); HIV control (*P* = 3E-11); triglycerides (*P* = 1E-08); T1D (*P* = 3E-08)
HLA-Bw4	rs9264942	0.521	**HIV-1 control (*P* = 3E-35); CD (*P* = 5E-28)**
HLA-Bw4	rs4947248	0.519	-
HLA-C1	rs2524084	0.546	RA (*P* = 5E-53)
HLA-C1	rs28397285	0.541	MG [4] (*P* = 5E-108); RA [5] (*P* = 5E-30); Triglycerides [6] (*P* = 1E-11)
HLA-C1	rs9366769 rs4122189	0.535	MG [4] (*P* = 5E-108); RA [5] (*P* = 5E-30); Triglycerides [6] (*P* = 1E-11)
HLA-C1	rs6921663	0.526	MG [4] (*P* = 5E-108); RA [5] (*P* = 5E-30); Triglycerides [6] (*P* = 1E-11)
HLA-C1	rs2524076	0.524	RA (*P* = 1E-11)
HLA-C1	rs12211087	- 0.514	**Psoriasis [1] (*P* = 5E-723)**; RA [2] (*P* = 2E-17); CD [3] (*P* = 3E-18)
HLA-C1	rs887464	0.513	MG (*P* = 7E-58); T1D (*P* = 8E-41); RA (*P* = 8E-30); idiopathic membranous nephropathy (*P* = 6E-14)

Liverpool Hope University EST. 1844

RESULTS: HLA-Bw4/Bw6

The HLA-C region SNP rs9264942 correlated with Bw4.

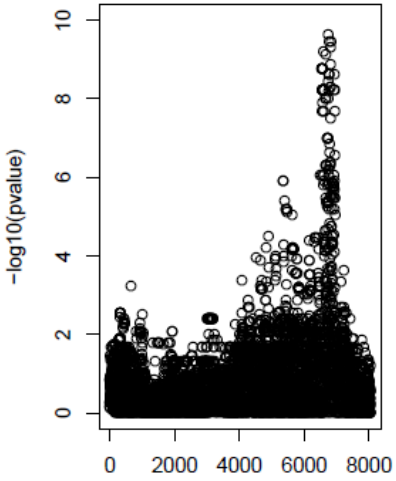
This SNP is the top GWAS hit for HIV-1 control (*P* = 3E-35) and implicated in HLA-C expression levels. Our result offers an alternative explanation that rs9264942 is a proxy for Bw4 (a known marker for HIV-1 control).

rs9264942 is also the top GWAS marker for Crohn disease in HLA class I region.

RESULTS: HLA-C1/C2

The strongest proxy for C1/C2, rs12211087, is the top genome-wide risk marker for psoriasis via its proxy (r^2=1) rs4406273 (*P* = 5E-723). However, this is due to the strict correlation of rs12211087 with *HLA-C*0602*.

RESULTS: HLA-DR51/52/53



Supertype	SNP	r	Associations
HLA-DR51	rs9270986	1.000	MS [1] (risk; *P* = 4E-225) and eQTL for *HLA-DRB5* with sex interaction (*P* = 4E-09)
HLA-DR51	rs1966002	0.988	RA (protective; *P* = 1E-250)
HLA-DR52	rs9271850	0.849	RA (*P* = 5E-204); cervical cancer (*P* = 1E-11); schizophrenia (*P* = 1E-08)
HLA-DR52	rs2097432	0.833	SS [2] (*P* = 1E-186); RA (*P* = 1E-123); eQTL for *HLA-DQA1* [2] (*P* = 3E-39) and *HLA-DRB1* [2] (*P* = 2E-36)
HLA-DR53	rs3104389	0.976	RA (*P* = 1E-250); UC (*P* = 5E-46)
HLA-DR53	rs9271488	0.976	RA (*P* = 1E-250); UC (*P* = 4E-41); sarcoidosis (*P* = 1E-08)
HLA-DR53	rs2133035	0.975	RA (*P* = 1E-250); UC (*P* = 5E-47)
HLA-DR53	rs9271574	0.974	RA (*P* = 1E-250); UC (*P* = 3E-43); sarcoidosis (*P* = 1E-08)

RESULTS: HLA-DR52

We identified three SNPs (rs2097432, rs9271613, rs9271850) as DR52 proxies, and their GWAS associations included the top genome-wide risk marker for systemic sclerosis (via proxy rs3129763; *P* = 9E-187).

The proxies for DR52 were the strongest eQTLs for HLA class II genes (*P* = 3E-189 for *HLA-DQA1*). Using two SNPs together increased the strength of the correlation (for rs3129887-A & rs2097432-G, r=0.94; *P* = 1E-38).

RESULTS: HLA-DR53

Four SNPs showed strong correlations (r>0.97) with DR53 (rs2133035, rs9271574, rs9271488, rs3104389) and were among the top GWAS hits for rheumatoid arthritis, ulcerative colitis and sarcoidosis, and strongest eQTLs for *HLA-DQA2*, *-DQA1*, *-DRB6* and *-DRB1*.

RESULTS: HLA-DR51

Two SNPs showed absolute correlations ($r \geq 0.99$) with DR51 (rs9270986, rs1966002). One of these (rs9270986) is a known sex-interacting eQTL for *HLA-DRB5*.

HLA-DR51 proxy SNPs are the strongest eQTLs for HLA class II genes.

RESULTS: HLA-DR51

rs9270986 is a known sex-interacting eQTL for *HLA-DRB5*.

rs9270986 as an eQTL for *HLA-DRB1*

Real time engine for expression Quantitative Trait Loci

MRCE		Intercept					SNP1											
Probe names	Gene names	Effect	SE	DF	t-value	p-value	SNP names	Effect	SE	DF	t-value	p-value	AL1	AL2	FREQ1	Rsq	Chr	Position
GI_4504410-S	HLA-DRB1	-0.19	0.055	253.0	-3.4	0.000709	rs9270986	0.544	0.0829	241.0	6.57	3.07e-10	C	A	0.1575	0.9212	6	41316

MRCA		Intercept					SNP1											
Probe names	Gene names	Effect	SE	DF	t-value	p-value	SNP names	Effect	SE	DF	t-value	p-value	AL1	AL2	FREQ1	Rsq	Chr	Position
208306_x_at	HLA-DRB1	-0.40	0.13	136.0	-3.1	0.00221	rs9270986	0.240	0.0728	45.0	3.30	0.00189	C	A	0.844	0.9585	6	41316
208306_x_at	HLA-DRB1	-0.23	0.12	117.0	-1.9	0.0581	rs9270986	0.131	0.0682	28.0	1.93	0.0643	C	A	0.844	0.9585	6	41316

MRCE		Intercept					SNP1											
Probe names	Gene names	Effect	SE	DF	t-value	p-value	SNP names	Effect	SE	DF	t-value	p-value	AL1	AL2	FREQ1	Rsq	Chr	Position
GI_4504410-S	HLA-DRB1	-0.19	0.066	172.0	-2.9	0.00486	rs9270986	0.568	0.108	72.0	5.27	0.00000139	C	A	0.1575	0.9212	6	41316
GI_4504410-S	HLA-DRB1	-0.19	0.079	174.0	-2.4	0.0171	rs9270986	0.521	0.120	74.0	4.33	0.0000461	C	A	0.1575	0.9212	6	41316

Liverpool Hope University EST. 1844

rs9270986 is a stronger eQTL for *HLA-DRB1* in males

RESULTS

Strongest and always positive eQTL effects on HLA class II genes are observed on HLA-DR51 haplotypes
The only negative eQTL effects were on DR53 haplotypes

HLA-DR Superytpe-representing SNPs as eQTLs in Peripheral Blood [a,b,c]

Supertype	Target gene	Effect size (beta)	Statistically most significant correlation
DR51	*HLA-DRB1*	0.879 to 0.941	rs9270986 (*P* = 2E-92; beta = 0.941)
DR52		No beta values	rs2097432, rs9271613 (*P* = 5E-09; no beta)
DR53		- 0.391 to +0.392	rs9271574 (*P* = 1E-14; beta=-0.391); rs3104389/rs2133035 (*P* = 1E-14; beta=+0.392)
DR51	*HLA-DQB1*	0.914 to 0.976	rs9270986 (*P* = 3E-30; beta=0.974)
DR52		0.442 to 0.505	rs17211580 (*P* = 1E-09; beta=0.505)
DR53		- 0.655 to +0.655	rs3104389 (*P* = 2E-18; beta=0.655)
DR51	*HLA-DQA1*	0.651 to 0.773	rs9270986 (*P* = 2E-25; beta=0.773)
DR52		0.398 to 0.657	rs2097432 (*P* = 3E-32; no beta)
DR53		- 0.392 to +0.392	rs9271488 (*P* = 2E-29; no beta)

[a] Analysis restricted to the results with effect sizes (beta values) and P values < 5E-08
[b] Results are from directly supertype-representing SNPs or their proxies with r^2>0.80
[c] Insufficient data for *HLA-DRA*

Liverpool Hope University EST. 1844

CONCLUSIONS

Our results suggest that the interpretation of HLA region SNP associations can be improved by taking into account additional levels of functional variation within the HLA region

Our result also point out the importance of evolutionarily important HLA supertypical lineages in HLA class II gene expression levels

YOUR FUTURE
STARTS WITH HOPE

Liverpool Hope University EST. 1844