Disease Associations of HLA Region SNPs

Mehmet Tevfik DORAK, MD PhD

School of Life Sciences, Pharmacy & Chemistry

Kingston University London

29 October 2023

Kingston University London

Outline

What is HLA?

Unique features of HLA region for medical biologists/geneticists

HLA region polymorphisms and disease

The meaning of HLA region associations with disease

A specific HLA region association with cancer

Conclusions

Kingston University London

HLA

Human MHC

HLA: Human leukocyte antigens
MHC: Major histocompatibility complex

First discovered in mice

The human version discovered in the 1950s

The first disease association was with leukemia in mice

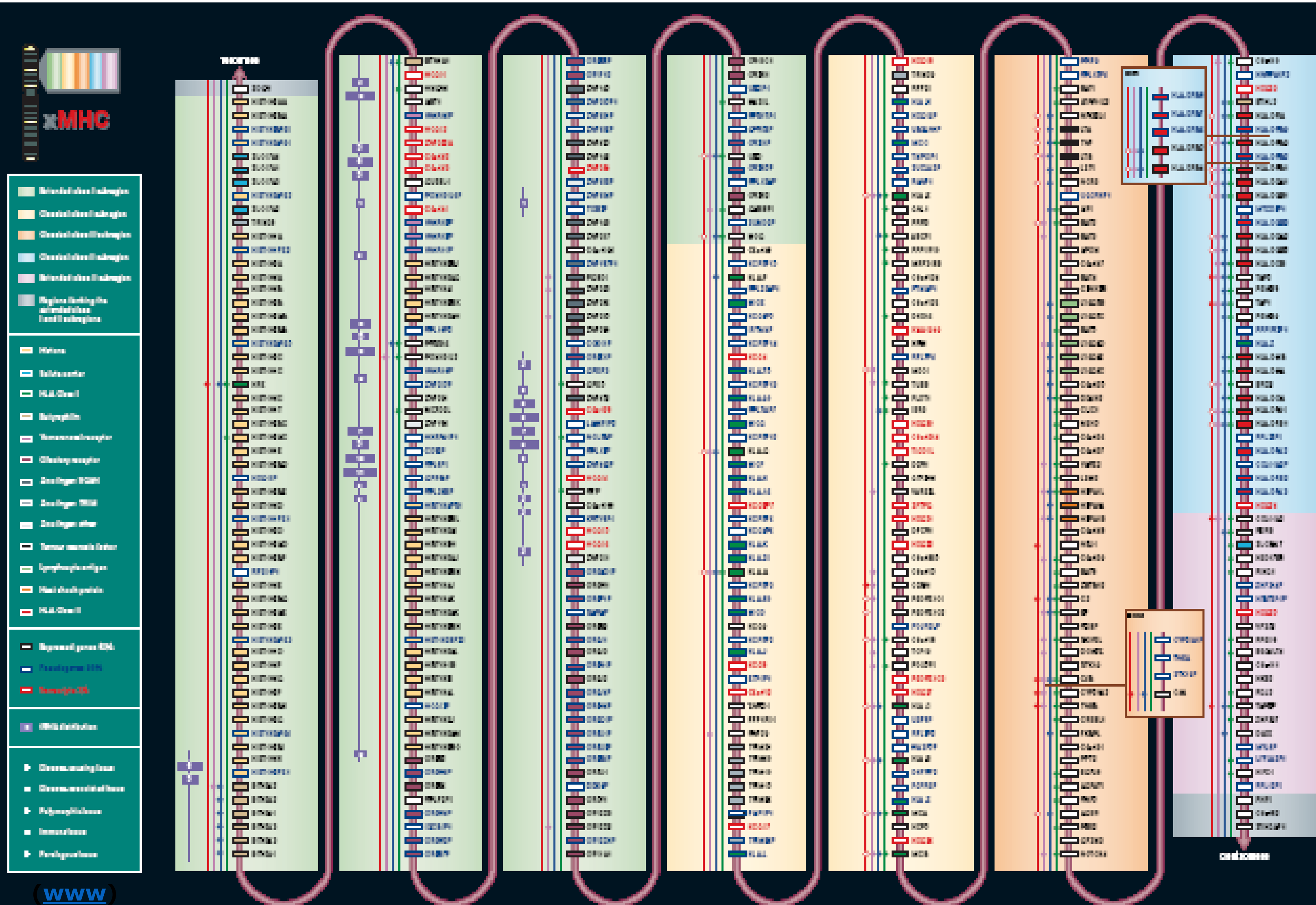
Human studies also confirmed associations with leukemia, other cancers and many other diseases

nature REVIEWS GENETICS

Gene map of the extended human MHC

Roger Horton, Laurens Wilming, Vikki Rand, Ruth C. Lovering, Elspeth A. Bruford, Varsha K. Khodiyar, Michael J. Lush, Sue Povey, C. Conover Talbot Jr., Mathew W. Wright, Hester M. Wain, John Trowsdale, Andreas Ziegler and Stephan Beck.

Abbott Molecular Diagnostics



HLA

Genes & SNPs

Most gene-dense region in the genome

Table 1. Human Genome Top 10 Gene-Dense Regions

GoldenPath location	Region	%GC	% repeats	Genes/Mb	Comments
chr6:31250001–32500000	HLAC–HLADRB3	47	47	48.8	Includes MHC class III region
chr6:25500001–26500000	FLJ20048–BTN2A3	41	43	44.0	Includes histone families
chr12:6250001–7250000	FLJ10665–PXR1	46	41	43.1	Includes CD4, complement 1
chr17:39000001–40000000	KRT23–ACLY	46	44	43.0	Includes keratin families
chr19:53250001–55000000	ELSPBP1–TCBAP0758	52	57	42.3	Includes CD37
chr16:250001–1500000	DKFZP761D0211–KIAA0683	60	28	40.8	GC rich
chr11:250001–1500000	AP2A2–HCCA2	53	36	40.2	Gap in sequence; includes IRF7, TOLLIP
chr17:7000001–8000000	ASGR1–PER1	51	43	39.0	Includes TNSF12, 13; CD68; TP53
chrX:150500001–151500000	DUSP9–GAB3	53	43	39.0	Includes G6PD; IRAK1
chr19:59250001–60250000	OSCAR–RDH13	49	53	36.0	Includes KIR, ILT, LILR families

Using a window offset of 250 kb, the number of genes per megabase and GC content were calculated as described in Figure 1. If a region appeared in the top 20 hits more than once (e.g., chr16:250001–250000 and chr16:5000001–1,500000), the regions were combined. "Region" indicates the outermost genes within the GoldenPath span.

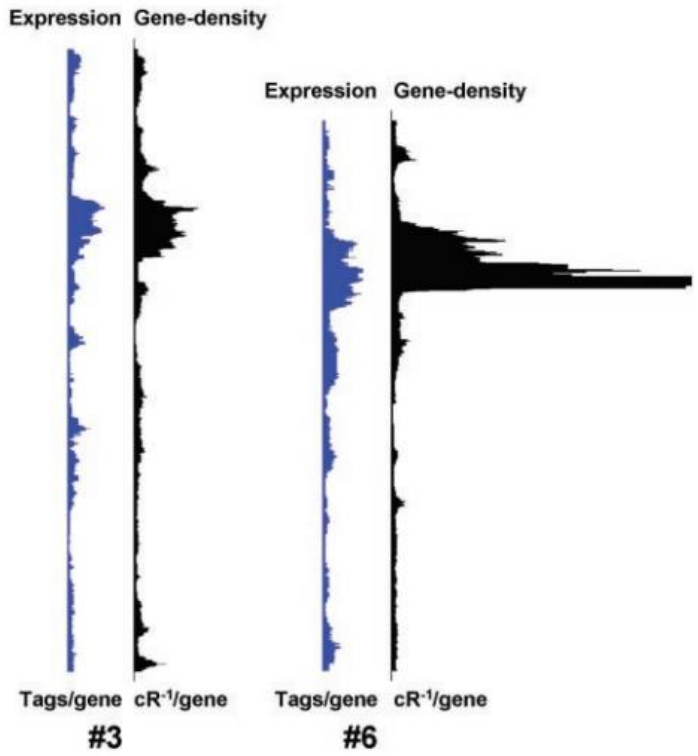
Analysis of the Gene-Dense Major Histocompatibility Complex Class III Region and Its Comparison to Mouse

Tao Xie,[1,4,7] Lee Rowen,[1,7] Begoña Aguado,[2,5] Mary Ellen Ahearn,[3] Anup Madan,[1,6] Shizhen Qin,[1] R. Duncan Campbell,[2] and Leroy Hood[1,8]

HLA

Genes & SNPs

Fig. 4. Comparison of median gene expression levels and gene density for chromosomes 3 and 6. The left diagrams of each chromosome show the expression levels as a moving median with a window size of 39 UniGene clusters. The right diagram of each chromosome shows gene density. For each UniGene cluster, we calculated the average distance between adjacent clusters in a window of 39 adjacent UniGene clusters. The inverse of this value is shown (inverse centirays per gene).



The Human Transcriptome Map: Clustering of Highly Expressed Genes in Chromosomal Domains

Huib Caron,[1,2] Barbera van Schaik,[1,3] Merlijn van der Mee,[3] Frank Baas,[4] Gregory Riggins,[6] Peter van Sluis,[1] Marie-Christine Hermus,[1] Ronald van Asperen,[1] Kathy Boon,[1] P. A. Voûte,[2] Siem Heisterkamp,[5] Antoine van Kampen,[3] Rogier Versteeg[1]

The chromosomal position of human genes is rapidly being established. We integrated these mapping data with genome-wide messenger RNA expression profiles as provided by SAGE (serial analysis of gene expression). Over 2.45 million SAGE transcript tags, including 160,000 tags of neuroblastomas, are presently known for 12 tissue types. We developed algorithms to assign these tags to UniGene clusters and their chromosomal position. The resulting Human Transcriptome Map generates gene expression profiles for any chromosomal region in 12 normal and pathologic tissue types. The map reveals a clustering of highly expressed genes to specific chromosomal regions. It provides a tool to search for genes that are overexpressed or silenced in cancer.

HLA

Genes & SNPs

Genome (3.2Gb)	xHLA Region (25.7 to 33.4Mb)	Comparison
Total No of Genes **60155**	**Total No of Genes** **674**	...
Protein-coding genes **19881**	**Protein-coding genes** **453**	32.66 vs 67.21% *P*<0.0001
Non-coding RNA Genes **25411**	**Non-coding RNA Genes** **54**	42.63 vs 8.01% *P*<0.0001
Long non-coding RNA genes **15877**	**Long non-coding RNA genes** **13**	1.93 vs 26.39% *P*<0.0001
Small non-coding RNA genes **9534**	**Small non-coding RNA genes** **7**	1.04 vs 15.85% *P*<0.0001
Pseudogenes **14467**	**Pseudogenes** **172**	24.03 vs 25.52% *P* = 0.37

xHLA makes up 0.24% of the genome, but contains 0.40% of all SNPs in the human genome

A systematic analysis of the gene and variation content of the extended HLA region

Ertan KANBUR, Mustafa DOGAN, Mehmet Tevfik DORAK

Presented at EFI 2017

Kingston University London

HLA

HLA Region and Disease Associations

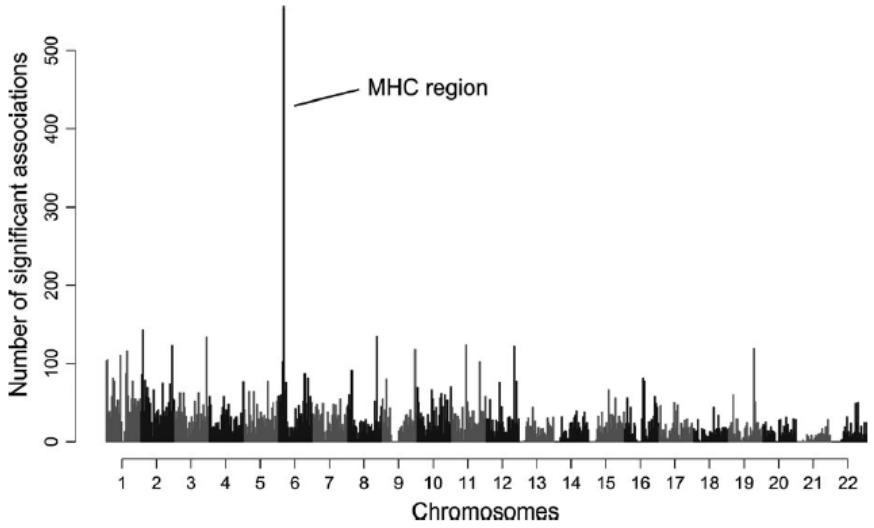


FIGURE 1 Number of significant GWAS associations along the genome. The chromosomal location of significant trait associations from GWAS (*N* = 18,682) is shown for all autosomes. Data from NHGRI GWAS catalog. Reproduced from "Lenz TL, Spirin V, Jordan DM, Sunyaev SR. Excess of Deleterious Mutations around HLA Genes Reveals Evolutionary Cost of Balancing Selection. Mol Biol Evol 2016;33(10):2555-64. https://doi.org/10.1093/molbev/msw127" by permission of Oxford University Press on behalf of the Society for Molecular Biology and Evolution

including:

- **Schizophrenia**
- **Alzheimer disease**
- **Parkinson disease**
- **Lung cancer**
- **Hodgkin lymphoma**

Received: 9 March 2017 | Revised: 16 June 2017 | Accepted: 20 July 2017

DOI: 10.1111/iji.12332

REVIEW

WILEY INTERNATIONAL JOURNAL OF IMMUNOGENETICS

What has GWAS done for HLA and disease associations?

A. E. Kennedy[1] | U. Ozbek[2,3] | M. T. Dorak[4]

HLA

A Recent HLA Association with Somatic Mutation Frequency

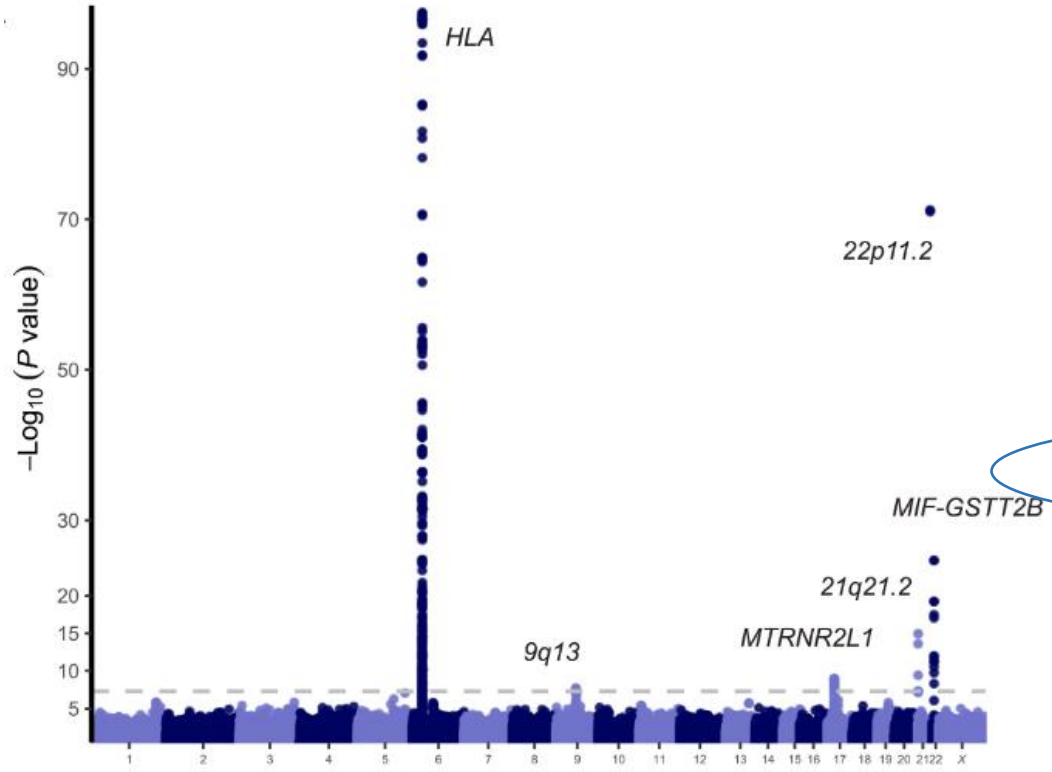


Fig. 3A. Germline determinants of recurrent non-missense somatic mutations (RNMSM) burden. Manhattan plot from GWAS of RNMSM burden, computed using SAIGE. Germline variants included had a minor allele count ≥ 600 and were distinct from the set of RNMSMs.

Germline genetic basis of RNMSMs

We performed a multi-ancestry genome-wide association study (GWAS) of the RNMSM burden using the scalable and accurate implementation of generalized mixed model (SAIGE) (*33*) and identified five genome-wide significant loci, all of which collectively highlight the influence of immune function on RNMSM burden (Fig. 3A). The strongest signal was observed in the *HLA* region at rs9271735, which is a common (MAF 28%) variant and is 2.5 kb upstream of the transcription start site (TSS) of *HLA-DQA1*, whose protein plays an essential role in antigen presentation. The A allele was associated with a 0.09 SD increase in RNMSM burden (P value = 3.8×10^{-98}). Given that there is extensive linkage disequilibrium at the *HLA* locus, this variant likely tags a specific *HLA* haplotype. To ensure that this signal was not a consequence of population stratification, we then performed both European and African ancestry–specific GWAS (Methods; fig. S3). We observed that rs9271735 was genome-wide significant in both ancestry-specific GWAS (African ancestry, P value = 4.9×10^{-18}; European ancestry, P value = 5.1×10^{-40}), indicating that the association is unlikely to be a consequence of population stratification. Consistent with the possible role of the adaptive immune system in surveillance of HSCs for excessive mutation, a recent report showed that HSCs in humans are antigen-presenting cells (*34*).

The genetic determinants of recurrent somatic mutations in 43,693 blood genomes

JOSHUA S. WEINSTOCK, CECELIA A. LAURIE, JAI G. BROOME, KENT D. TAYLOR, XIUQING GUO, ALAN R. SHULDINER, JEFFREY R. O'CONNELL, JOSHUA P. LEWIS, ERIC BOERWINKLE, [...], AND NHLBI TRANS-OMICS FOR PRECISION MEDICINE (TOPMED) CONSORTIUM +78 authors Authors Info & Affiliations

SCIENCE ADVANCES · 26 Apr 2023 · Vol 9, Issue 17 · DOI: 10.1126/sciadv.abm4945

Kingston University London

Unique Features of HLA

Heritability of DNA methylation

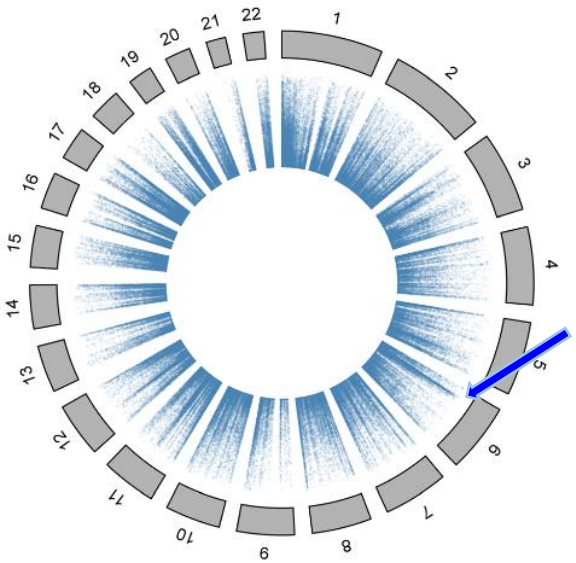


Figure 2 Distribution of genetic heritability estimates across the genome. The MHC region, which had the highest estimates of genetic heritability is clearly visible on chromosome 6. Telomeric regions show an increased density of probes with high genetic heritability, although this is primarily due to higher numbers of probes in these regions.

McRae *et al. Genome Biology* 2014, **15**:R73
http://genomebiology.com/2014/15/5/R73

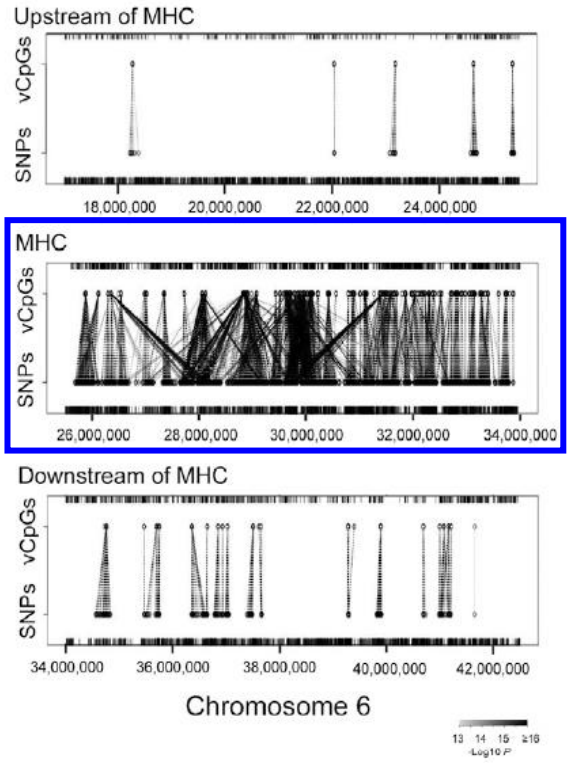
Genome Biology

RESEARCH Open Access

Contribution of genetic variation to transgenerational inheritance of DNA methylation

Allan F McRae[1,2*], Joseph E Powell[1,2], Anjali K Henders[3], Lisa Bowdler[3], Gibran Hemani[1,2], Sonia Shah[1,2], Jodie N Painter[3], Nicholas G Martin[3], Peter M Visscher[1,2†] and Grant W Montgomery[3†]

meQTL density



(E) Associations between CpG sites and SNPs upstream of (top panel), within (middle panel), or downstream of (bottom panel) the major histocompatibility complex (MHC) region. Each dashed line represents a significant association, and the shades of black indicate significance of the associations.

GeMes, Clusters of DNA Methylation under Genetic Control, Can Inform Genetic and Epigenetic Analysis of Disease

Yun Liu,[1,2,9] Xin Li,[1,2,9] Martin J. Aryee,[1,3] Tomas J. Ekström,[4,5] Leonid Padyukov,[4,6] Lars Klareskog,[4,6] Amy Vandiver,[1,2] Ann Zenobia Moore,[7] Toshiko Tanaka,[7] Luigi Ferrucci,[7] M. Daniele Fallin,[1,8,*] and Andrew P. Feinberg[1,2,*]

trans-eQTL density

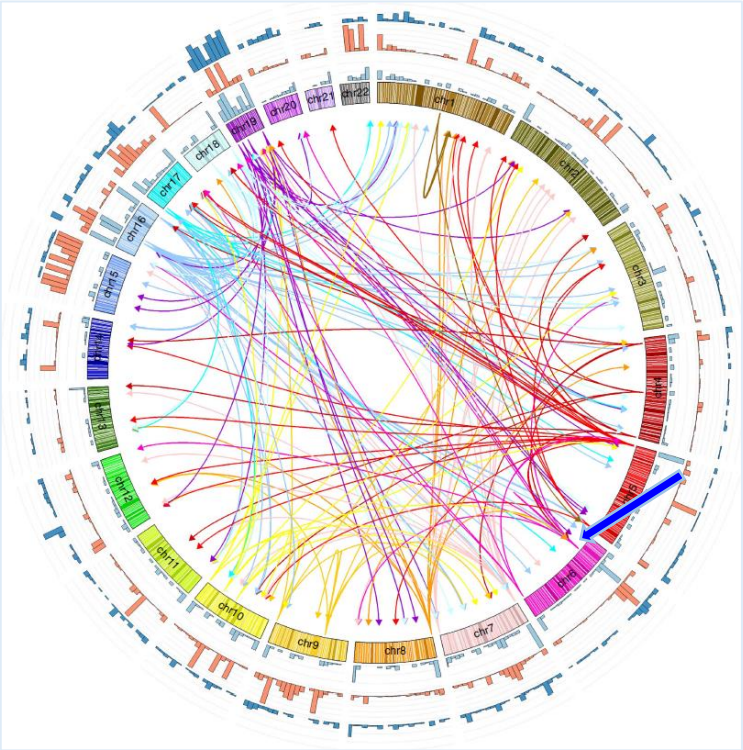


Figure 1 | Enrichment of features in regions harbouring SNPs involved in distal SNP–CpG associations. Outer histograms: number of SNPs involved in distal SNP–CpG associations (light blue), calculated in 7.5 Mb bins; number of piRNA sequences (orange); number of transcription factors (dark blue). Inner links: SNP regions associated with four or more CpG sites. Arrows are pointing from SNPs to the CpG sites they are associated with, and are coloured according to the chromosomes where the SNPs reside.

Long-range epigenetic regulation is conferred by genetic variation located at thousands of independent loci

Mathieu Lemire[1], Syed H.E. Zaidi[1], Maria Ban[2], Bing Ge[3], Dylan Aïssi[4,5,6], Marine Germain[4,5,6], Irfahan Kassam[7], Mike Wang[1], Brent W. Zanke[8], France Gagnon[7], Pierre-Emmanuel Morange[9,10,11], David-Alexandre Trégouët[4,5,6], Philip S. Wells[8], Stephen Sawcer[2], Steven Gallinger[12,13], Tomi Pastinen[3] & Thomas J. Hudson[1,14,15]

Kingston University London

Unique Features of HLA

trans-meQTLs

	P Value	SNPName	SNPChr	SNPChrPos	ProbeName	ProbeChr	ProbeCenterChrPos	CisTrans	SNPType	Allele	OverallZScore
2	3.27167E-310	rs9268923	6	32432835	cg10154826	6	17601018	trans	C/T	T	61.2544774
3	3.27167E-310	rs2395185	6	32433167	cg10154826	6	17601018	trans	G/T	T	61.2544774
4	3.27167E-310	rs9268853	6	32429643	cg10154826	6	17601018	trans	T/C	C	61.1446653
5	3.27167E-310	rs477515	6	32569691	cg10154826	6	17601018	trans	G/A	A	58.2171575
6	3.27167E-310	rs4239020	17	80176641	cg07393940	7	158741793	trans	C/T	C	-55.4574242
7	3.27167E-310	rs12342831	9	33124872	cg20290983	6	43655494	trans	T/C	C	55.1041246
8	3.27167E-310	rs10813951	9	33128021	cg20290983	6	43655494	trans	A/G	G	55.0625727
9	3.27167E-310	rs12342831	9	33124872	cg04842962	6	43655513	trans	T/C	C	54.9486903
10	3.27167E-310	rs10813951	9	33128021	cg04842962	6	43655513	trans	A/G	G	54.9126278
11	3.27167E-310	rs3780486	9	33139453	cg20290983	6	43655494	trans	C/T	T	54.832052
12	3.27167E-310	rs3780486	9	33139453	cg04842962	6	43655513	trans	C/T	T	54.7699307
13	3.27167E-310	rs10813957	9	33153527	cg20290983	6	43655494	trans	G/T	T	54.7260295
14	3.27167E-310	rs10813957	9	33153527	cg04842962	6	43655513	trans	G/T	T	54.6814431
15	3.27167E-310	rs5756504	22	37467270	cg13737042	9	37806390	trans	C/T	T	-50.6708043
16	3.27167E-310	rs5756506	22	37467392	cg13737042	9	37806390	trans	G/C	C	-50.6605533
17	3.27167E-310	rs62103177	18	77624479	cg05926928	17	57297748	trans	G/A	A	49.4548764
18	3.27167E-310	rs8321	6	30032522	cg01620082	3	125678431	trans	A/C	C	-46.813694
19	3.27167E-310	rs2074977	19	3434028	cg08382705	11	45687343	trans	C/A	C	-46.6429193
20	3.27167E-310	rs8321	6	30032522	cg06606381	12	133084921	trans	A/C	C	-46.2606567

BIOS QTL browser

This web page accompanies the manuscripts titled ***'Disease variants alter transcription factor levels and methylation of their binding sites'***, by *Bonder & Luijk et al* and ***'Unbiased identification of regulatory modifiers of genetic risk factors'***, by *Zhernakova et al*, both have been submitted to Nature Genetics.
For further questions, contact the corresponding author: lude@ludesign.nl

Download meQTL results

You can download the full *cis*- and *trans*-meQTL and eQTMs, detected at a false-discovery rate of 0.05:
Cis-meQTLs
Cis-eQTMs
Trans-meQTLs

Disease variants alter transcription factor levels and methylation of their binding sites

Marc Jan Bonder[1,22], René Luijk[2,22], Daria V Zhernakova[1], Matthijs Moed[2], Patrick Deelen[1,3], Martijn Vermaat[4], Maarten van Iterson[2], Freerk van Dijk[1,3], Michiel van Galen[3], Jan Bot[5], Roderick C Slieker[2], P Mila Jhamai[6], Michael Verbiest[3], H Eka D Suchiman[2], Marijn Verkerk[6], Ruud van der Breggen[2], Jeroen van Rooij[6], Nico Lakenberg[2], Wibowo Arindrarto[7], Szymon M Kielbasa[8], Iris Jonkers[1], Peter van 't Hof[8], Irene Nooren[5], Marian Beekman[2], Joris Deelen[2], Diana van Heemst[9], Alexandra Zhernakova[1], Ettje F Tigchelaar[1], Morris A Swertz[1,3], Albert Hofman[10], André G Uitterlinden[6], René Pool[11], Jenny van Dongen[11], Jouke J Hottenga[11], Coen D A Stehouwer[12,13], Carla J H van der Kallen[12,13], Casper G Schalkwijk[12,13], Leonard H van den Berg[14], Erik W van Zwet[7], Hailiang Mei[8], Yang Li[1], Mathieu Lemire[15], Thomas J Hudson[15–17], the BIOS Consortium[18], P Eline Slagboom[2], Cisca Wijmenga[1], Jan H Veldink[14], Marleen M J van Greevenbroek[12,13], Cornelia M van Duijn[19], Dorret I Boomsma[11], Aaron Isaacs[13,19,20], Rick Jansen[21], Joyce B J van Meurs[6], Peter A C 't Hoen[4,23], Lude Franke[1,23] & Bastiaan T Heijmans[2,23]

Sex-differential in trait associations

NATURE GENETICS ARTICLES

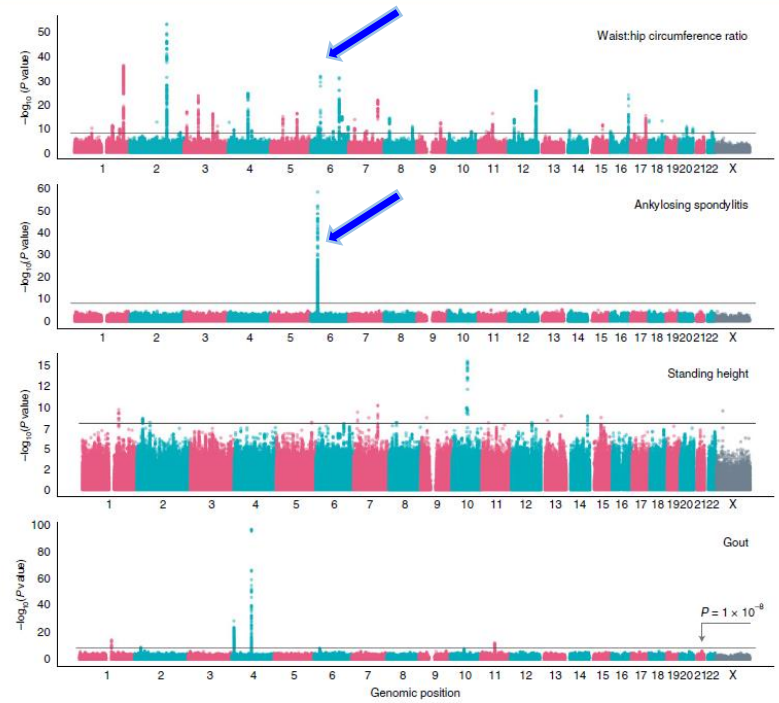


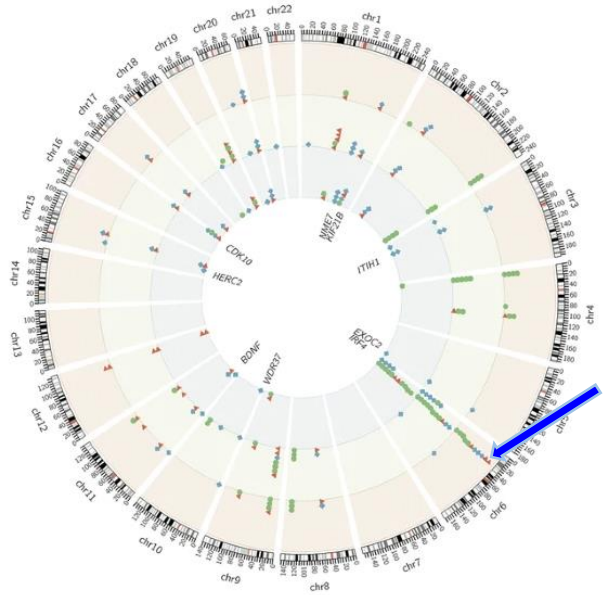
Fig. 3 | Manhattan plots for traits with most lead sdSNPs. The *x* axis corresponds to the genomic position in the genome and the *y* axis to the $-\log_{10}(P$ value) of the two-sided Student's *t*-test (Methods), for which the null hypothesis is that there is no difference between the sexes. Each point corresponds to a genetic variant. Points that go above the statistical significance line at $-\log_{10}(P) = 1 \times 10^{-8}$ are considered to be sdSNPs. Traits represented include: waist:hip circumference ratio, ankylosing spondylitis, standing height and gout.

Sex differences in genetic architecture in the UK Biobank

Elena Bernabeu[1], Oriol Canela-Xandri[2], Konrad Rawlik[1], Andrea Talenti[1], James Prendergast[1] and Albert Tenesa[1,2]

Kingston University London

Unique Features of HLA



Overview of PheWAS associations in the genome after functional annotation. **a** This matrix shows the number of functional SNPs for their respective phenotype. **b** The *Circos plot* showing the PheWAS associations in different types of functional data. *Red triangles* represent the associations in the GWAS Catalog only, *green circles* represent GWAS Catalog associations replicated by PheWAS ($P < 0.05$), and *blue diamonds* represent new phenotype associations identified by PheWAS ($P < 4.6 \times 10^{-6}$ or FDR < 0.1)

including:

- **Schizophrenia**
- **Alzheimer disease**
- **Parkinson disease**
- **Lung cancer**
- **Hodgkin lymphoma**

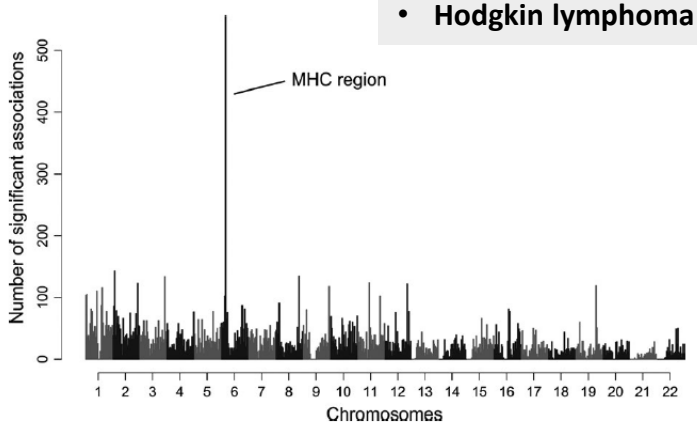


FIGURE 1 Number of significant GWAS associations along the genome. The chromosomal location of significant trait associations from GWAS (*N* = 18,682) is shown for all autosomes. Data from NHGRI GWAS catalog. Reproduced from "Lenz TL, Spirin V, Jordan DM, Sunyaev SR. Excess of Deleterious Mutations around HLA Genes Reveals Evolutionary Cost of Balancing Selection. Mol Biol Evol 2016;33(10):2555-64. https://doi.org/10.1093/molbev/msw127" by permission of Oxford University Press on behalf of the Society for Molecular Biology and Evolution

…. despite already showing the highest number of disease associations, the true extent of the involvement of the MHC region in disease genetics may not have been uncovered.

Zhao et al. Genome Medicine (2018) 10:7
DOI 10.1186/s13073-018-0513-x

Genome Medicine

RESEARCH Open Access

An integrative functional genomics framework for effective identification of novel regulatory variants in genome–phenome studies

Junfei Zhao[1†], Feixiong Cheng[2,3†], Peilin Jia[1], Nancy Cox[4,5], Joshua C. Denny[5,6] and Zhongming Zhao[1,7*]

Received: 9 March 2017 | Revised: 16 June 2017 | Accepted: 20 July 2017
DOI: 10.1111/iji.12332

REVIEW

WILEY INTERNATIONAL JOURNAL OF IMMUNOGENETICS

What has GWAS done for HLA and disease associations?

A. E. Kennedy[1] | U. Ozbek[2,3] | M. T. Dorak[4]

Unique Features of HLA

Heritability of Gene Expression

Heritability and genomics of gene expression in peripheral blood

Fred A Wright[1–3,13], Patrick F Sullivan[4,13], Andrew I Brooks[5], Fei Zou[6], Wei Sun[6], Kai Xia[6], Vered Madar[6], Rick Jansen[7], Wonil Chung[6], Yi-Hui Zhou[1,2], Abdel Abdellaoui[8], Sandra Batista[9], Casey Butler[9], Guanhua Chen[6], Ting-Huei Chen[6], David D'Ambrosio[10], Paul Gallins[4], Min Jin Ha[6], Jouke Jan Hottenga[8], Shunping Huang[9], Mathijs Kattenberg[8], Jaspreet Kochar[10], Christel M Middeldorp[8], Ani Qu[10], Andrey Shabalin[11], Jay Tischfield[5], Laura Todd[4], Jung-Ying Tzeng[1,2], Gerard van Grootheest[7], Jacqueline M Vink[8], Qi Wang[10], Wei Wang[12], Weibo Wang[9], Gonneke Willemsen[8], Johannes H Smit[7], Eco J de Geus[8], Zhaoyu Yin[6], Brenda W J H Penninx[7] & Dorret I Boomsma[8]

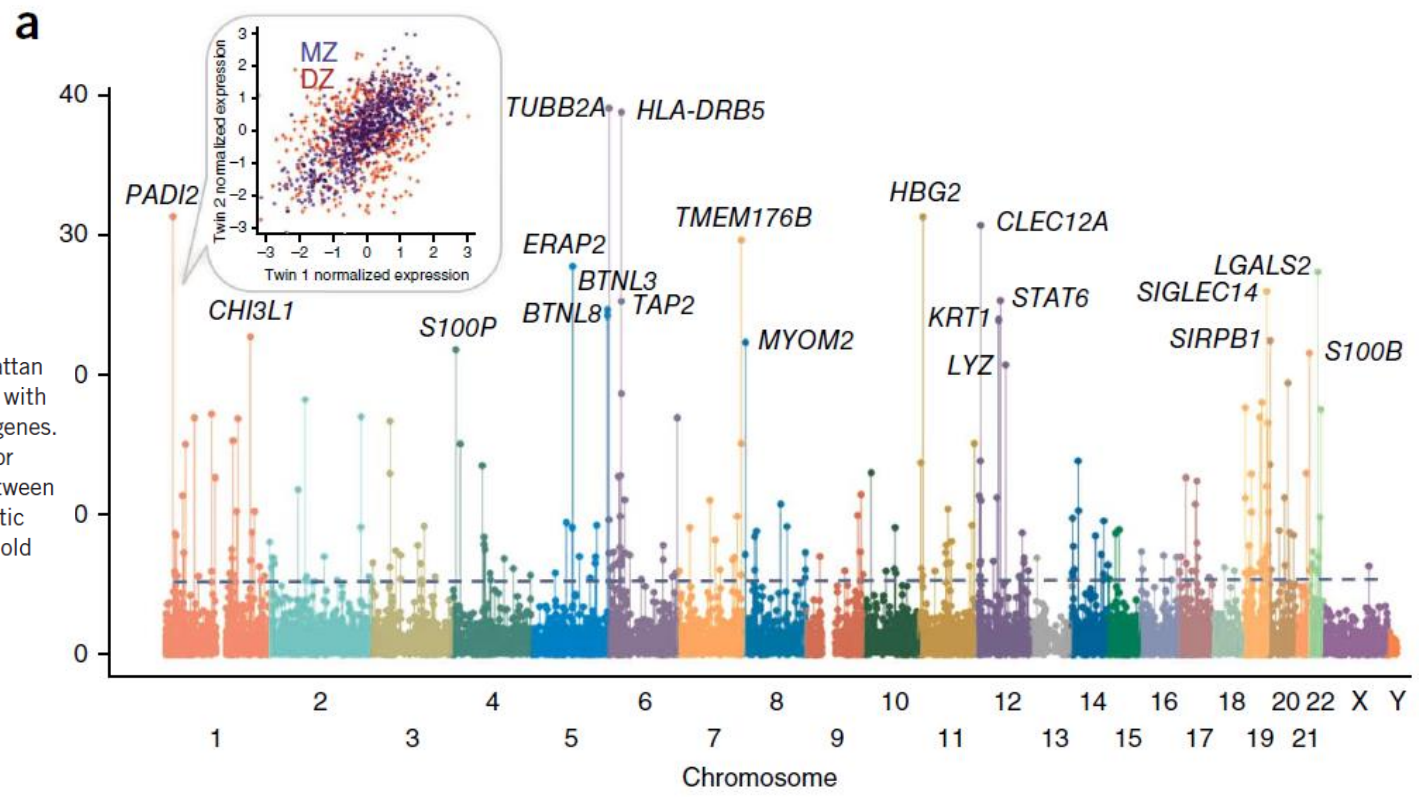


Figure 1 Transcriptome-wide estimates of heritability based on 2,752 twins. (**a**) Manhattan plot of heritability *P* values for the transcript with the highest h^2 estimate for each of 18,392 genes. The inset (*PADI2*) shows that the evidence for heritability is based on higher correlation between monozygotic pairs (MZ) than between dizygotic pairs (DZ). The dashed line marks the threshold for genes with $q < 0.05$.

Unique Features of HLA

Heritability of Gene Expression

Table 3 Details of the 10 most heritable probes that do not contain any annotated SNPs

Probe ID	Chr	CpG position	h^2	Genomic context	Type	HIL	#CpG	GWAS SNP	SNP position	*P* value
cg15671450	6	29895116	0.934	Upstream (HCG4B)	II	HC	1	rs111482415	29923140	4.8×10^{-78}
cg01903420	13	27295928	0.933	Intergenic	II	IC	2	rs1374010	27295317	3.0×10^{-105}
cg03168497	17	48586147	0.932	Intronic (MYCBPAP)	II	HC	4	rs73351675	48585554	8.1×10^{-84}
cg11064039	7	766100	0.932	Intronic (PRKAR1B)	I	HC	3	rs11763218	852281	8.8×10^{-58}
cg24372256	21	43528868	0.931	Intronic (UMODL1)	II	IC	1	rs34212454	43529216	2.9×10^{-101}
cg26764761	16	87682142	0.927	Intronic (JPH2)	I	IC	7	rs748554	87682775	1.4×10^{-107}
cg16761754	14	105127242	0.927	Intergeneic	I	IC	3	rs4075355	105125512	1.8×10^{-77}
cg21358336	17	6558440	0.927	Upstream (MIR4520B)/ Downstream (MIR4520A)	II	ICshore	1	rs2040847	6558011	1.3×10^{-91}
cg04118610	4	62707027	0.926	Intronic (LPHN3)	II	LC	2	rs10021525	62707476	2.1×10^{-105}
cg08164151	12	131118432	0.925	Intergeneic	II	IC	3	rs10848167	131123623	2.9×10^{-101}

Genomic context of probes was annotated using ANNOVAR [24], with a probe being upstream or downstream defined as being within 2 KB of the transcription start site or transcription end site, respectively.
GWAS SNP = Most significant SNP from GWAS; HIL = 'HIL' classification of CpG [22]; Type = Illumina HumanMethylation450 assay probe type.

McRae *et al. Genome Biology* 2014, **15**:R73
http://genomebiology.com/2014/15/5/R73

Genome Biology

RESEARCH Open Access

Contribution of genetic variation to transgenerational inheritance of DNA methylation

Allan F McRae[1,2*], Joseph E Powell[1,2], Anjali K Henders[3], Lisa Bowdler[3], Gibran Hemani[1,2], Sonia Shah[1,2], Jodie N Painter[3], Nicholas G Martin[3], Peter M Visscher[1,2†] and Grant W Montgomery[3†]

Unique Features of HLA

Heritability of Gene Expression

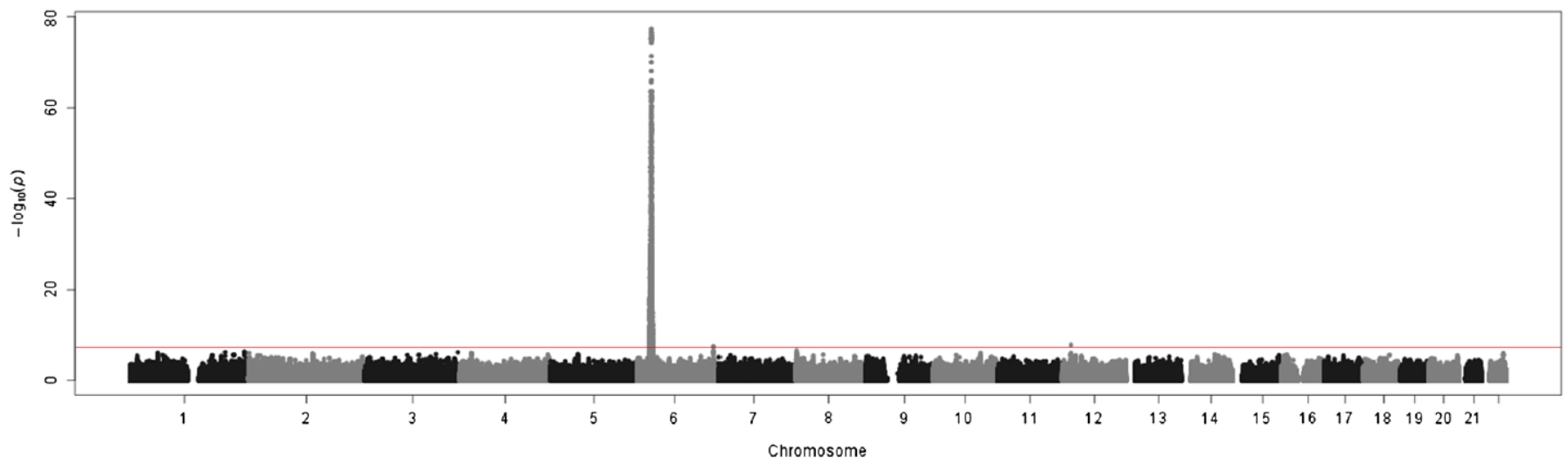


Figure 3 Manhattan plot of the genome-wide association *P* values for methylation probe cg15671450. The genome-wide significance level of 5×10^{-8} is indicated by the horizontal line. A highly significant effect is observed *cis* to the methylation probe on chromosome 6.

McRae *et al. Genome Biology* 2014, **15**:R73
http://genomebiology.com/2014/15/5/R73

Genome Biology

RESEARCH Open Access

Contribution of genetic variation to transgenerational inheritance of DNA methylation

Allan F McRae[1,2*], Joseph E Powell[1,2], Anjali K Henders[3], Lisa Bowdler[3], Gibran Hemani[1,2], Sonia Shah[1,2], Jodie N Painter[3], Nicholas G Martin[3], Peter M Visscher[1,2†] and Grant W Montgomery[3†]

Unique Features of HLA

How Can HLA be Involved in So Many Phenotypes?
trans-eQTLs in xMHC

Genetics of gene expression in primary immune cells identifies cell type–specific master regulators and roles of HLA alleles

Benjamin P Fairfax[1], Seiko Makino[1], Jayachandran Radhakrishnan[1], Katharine Plant[1], Stephen Leslie[2], Alexander Dilthey[3], Peter Ellis[4], Cordelia Langford[4], Fredrik O Vannberg[1,5] & Julian C Knight[1]

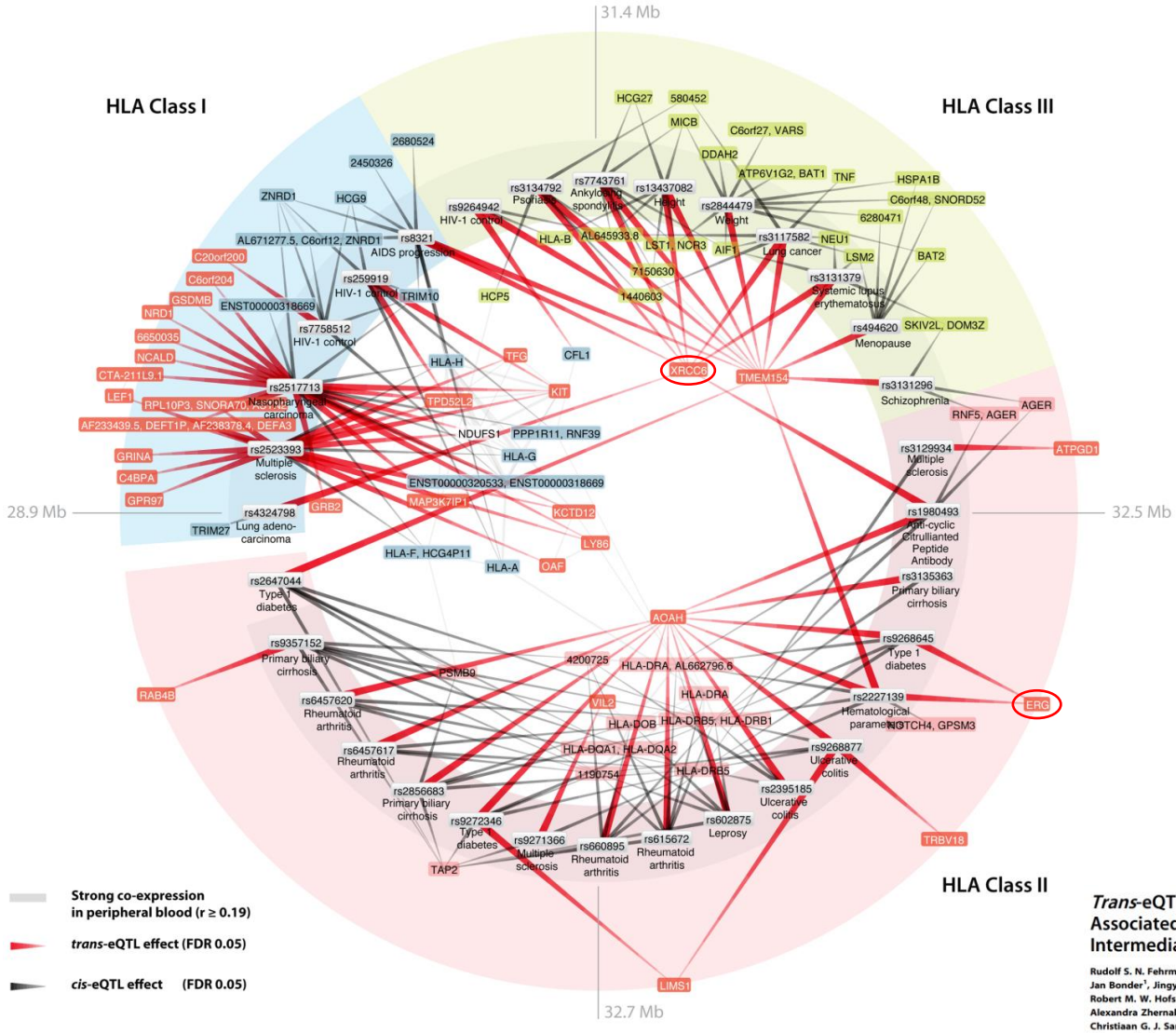
***Trans*-acting genetic variants have a substantial, albeit poorly characterized, role in the heritable determination of gene expression. Using paired purified primary monocytes and B cells, we identify new predominantly cell type–specific *cis* and *trans* expression quantitative trait loci (eQTLs), including multi-locus *trans* associations to *LYZ* and *KLF4* in monocytes and B cells, respectively. Additionally, we observe a B cell–specific *trans* association of rs11171739 at 12q13.2, a known autoimmune disease locus, with *IP6K2* ($P = 5.8 \times 10^{-15}$), *PRIC285* ($P = 3.0 \times 10^{-10}$) and an upstream region of *CDKN1A* ($P = 2 \times 10^{-52}$), suggesting roles for cell cycle regulation and peroxisome proliferator-activated receptor γ (PPARγ) signaling in autoimmune pathogenesis. We also find that specific human leukocyte antigen (HLA) alleles form *trans* associations with the expression of *AOAH* and *ARHGAP24* in monocytes but not in B cells. In summary, we show that mapping gene expression in defined primary cell populations identifies new cell type–specific *trans*-regulated networks and provides insights into the genetic basis of disease susceptibility.**

HLA-DR53 family members have strong genome-wide trans-eQTL effects in monocytes

Kingston University London

Unique Features of HLA

How Can HLA be Involved in So Many Phenotypes?
trans-eQTLs in xMHC



***Trans*-eQTLs Reveal That Independent Genetic Variants Associated with a Complex Phenotype Converge on Intermediate Genes, with a Major Role for the HLA**

Rudolf S. N. Fehrmann[1], Ritsert C. Jansen[2,*], Jan H. Veldink[3,*], Harm-Jan Westra[1,*], Danny Arends[2], Marc Jan Bonder[1], Jingyuan Fu[1], Patrick Deelen[1], Harry J. M. Groen[4], Asia Smolonska[1], Rinse K. Weersma[1,5], Robert M. W. Hofstra[1], Wim A. Buurman[6], Sander Rensen[6], Marcel G. M. Wolfs[7], Mathieu Platteel[7], Alexandra Zhernakova[8], Clara C. Elbers[9], Eleanora M. Festen[1], Gosia Trynka[1], Marten H. Hofker[7], Christiaan G. J. Saris[3], Roel A. Ophoff[3,10,11], Leonard H. van den Berg[3], David A. van Heel[12], Cisca Wijmenga[1], Gerard J. te Meerman[1*], Lude Franke[1,12,*]

Kingston University London

Unique Features of HLA

How Can HLA be Involved in So Many Phenotypes?
trans-meQTLs in xMHC

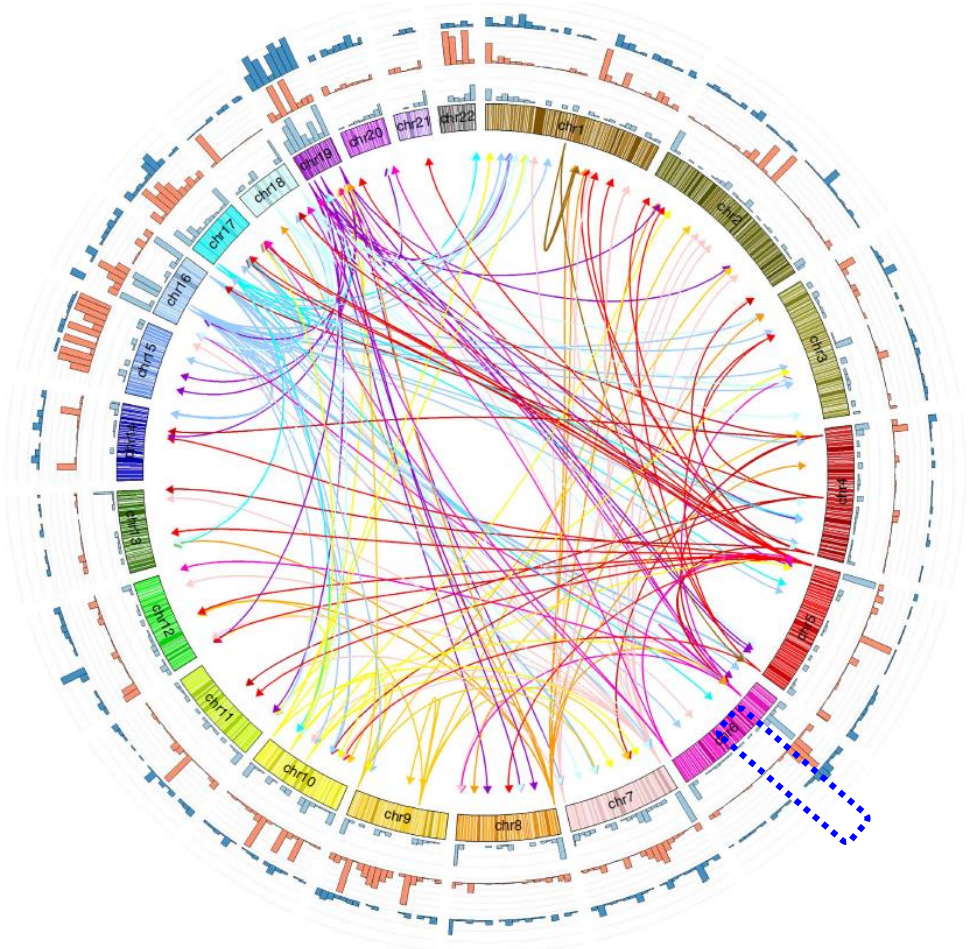


Figure 1 | Enrichment of features in regions harbouring SNPs involved in distal SNP–CpG associations. Outer histograms: number of SNPs involved in distal SNP–CpG associations (light blue), calculated in 7.5 Mb bins; number of piRNA sequences (orange); number of transcription factors (dark blue). Inner links: SNP regions associated with four or more CpG sites. Arrows are pointing from SNPs to the CpG sites they are associated with, and are coloured according to the chromosomes where the SNPs reside.

Long-range epigenetic regulation is conferred by genetic variation located at thousands of independent loci

Mathieu Lemire[1], Syed H.E. Zaidi[1], Maria Ban[2], Bing Ge[3], Dylan Aïssi[4,5,6], Marine Germain[4,5,6], Irfahan Kassam[7], Mike Wang[1], Brent W. Zanke[8], France Gagnon[7], Pierre-Emmanuel Morange[9,10,11], David-Alexandre Trégouët[4,5,6], Philip S. Wells[8], Stephen Sawcer[2], Steven Gallinger[12,13], Tomi Pastinen[3] & Thomas J. Hudson[1,14,15]

Kingston University London

Unique Features of HLA

How Can HLA be Involved in So Many Phenotypes?
Gene Networks

Table 1. Regions showing top scores in the genome.

region	criteria	Closest gene	Distance to closest gene	Description of closest gene	2nd closest gene	Description of 2nd closest gene
chr2:91959344-91968231	high number of components	GGT8P	inside gene	pseudogene		
chr6:33037767-33038449	high number of components	HLA-DPA1/ HLA-DPB1	inside gene	Homo sapiens major histocompatibility complex, class II		
chr:7203189-7420319641	high number of components	ITGB8	50,684 bp	integrin	HLA-DPA1/	major histocompatibility
					HLA-DPB1	complex, class II
				praja ring finger 2,		
chr5:108634323-108635534	high average degree	PJA2	34,876 bp	E3 ubiquitin protein ligase	AK021888	unknown function
chr8:25935936-25937929	high average degree	EBF2	inside gene	early B-cell factor 2		
				Homo sapiens		
chr6:32507854-32508257	high average path length	HLA-DRB1	inside gene	major histocompatibility complex, class II		
				Homo sapiens		major
chr6:32568909-32569343	high average path length	HLA-DRB5	11,297 bp	major histocompatibility complex, class II	HLA-DQA1	histocompatibility complex, class II
				Homo sapiens		
chr6:32611264-32611586	high average path length	HLA-DQA1	inside gene	major histocompatibility complex, class II		
chr3:36921415-36921688	high number of vertices	TRANK1	inside gene	tetratricopeptide repeat and ankyrin repeat Containing 1		
chr4:9176678-9178624	high number of vertices	C9JJH3	33,759 bp	Deubiquitinating enzyme	LOC650293	transmembrane helix receptor
chr8:35105546-35106981	high number of vertices	UNC5D	inside gene	receptor of netrin involved in nervous system		
chr4:9200148-9202368	few components, but large number of vertices	USP17L10	10,015 bp	Deubiquitinating enzyme		
chr6:31357915-31358747	few components, but large number of vertices	MICA	8,814 bp	MHC class I polypeptide-related sequence A	HLA-B	major histocompatibility complex, class I
chr6:31455010-31456012	few components, but large number of vertices	MICB	6,646 bp	MHC class I polypeptide-related Sequence B	uc003ntm.3	HLA complex Group 26 (non-protein coding)

doi:10.1371/journal.pone.0099424.t001

OPEN ACCESS Freely available online

PLOS ONE

Human Genome Variation and the Concept of Genotype Networks

Giovanni Marco Dall'Olio[1]*¤, Jaume Bertranpetit[1], Andreas Wagner[2,3,4], Hafid Laayouni[1,5]

1 Institut de Biologia Evolutiva, CSIC-Universitat Pompeu Fabra, Barcelona, Catalonia, Spain, **2** Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Zürich, Switzerland, **3** The Swiss Institute of Bioinformatics, Lausanne, Switzerland, **4** The Santa Fe Institute, Santa Fe, New Mexico, United States of America, **5** Universitat Autonòma de Barcelona, Barcelona, Spain

Abstract

Genotype networks are a concept used in systems biology to study sets of genotypes having the same phenotype, and the ability of these to bring forth novel phenotypes. In the past they have been applied to determine the genetic heterogeneity, and stability to mutations, of systems such as metabolic networks and RNA folds. Recently, they have been the base for reconciling the neutralist and selectionist views on evolution. Here, we adapted this concept to the study of population genetics data. Specifically, we applied genotype networks to the human 1000 genomes dataset, and analyzed networks composed of short haplotypes of Single Nucleotide Variants (SNV). The result is a scan of how properties related to genetic heterogeneity and stability to mutations are distributed along the human genome. We found that genes involved in acquired immunity, such as some HLA and MHC genes, tend to have the most heterogeneous and connected networks, and that coding regions tend to be more heterogeneous and stable to mutations than non-coding regions. We also found, using coalescent simulations, that regions under selection have more extended and connected networks. The application of the concept of genotype networks can provide a new opportunity to understand the evolutionary processes that shaped our genome. Learning how the genotype space of each region of our genome has been explored during the evolutionary history of the human species can lead to a better understanding on how selective pressures and neutral factors have shaped genetic diversity within populations and among individuals. Combined with the availability of larger datasets of sequencing data, genotype networks represent a new approach to the study of human genetic diversity that looks to the whole genome, and goes beyond the classical division between selection and neutrality methods.

Unique Features of HLA

How Can HLA be Involved in So Many Phenotypes?
Linkage Disequilibrium

Preliminary evidence of an association between HLA-*DPB1*0201* and childhood common ALL supports an infectious aetiology
Leukemia 1995;9(3):440-3

Evidence that an *HLA-DQA1-DQB1* haplotype influences susceptibility to childhood common ALL in boys provides further support for an infection-related aetiology
Br J Cancer 1998;78(5):561-5

Why not LD?

Unique Features of HLA

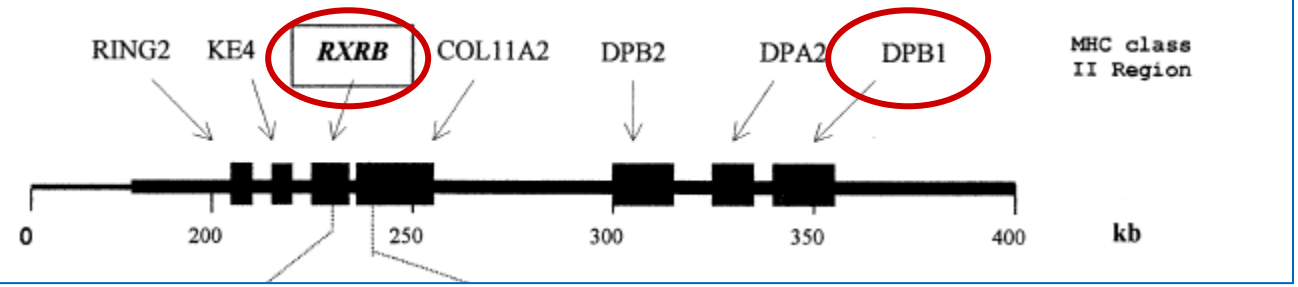
How Can HLA be Involved in So Many Phenotypes?
Linkage Disequilibrium

Linkage Disequilibrium Between HLA-DPB1 Alleles and Retinoid X Receptor β Haplotypes

Human Immunology 63, 771–778 (2002)

Published by Elsevier Science Inc.

Rajsbaum, 2002 (www)



Linkage disequilibrium 'confounding by locus' has to be ruled out before attributing a direct causal role to any genetic association

Unique Features of HLA

How Can HLA be Involved in So Many Phenotypes?
Linkage Disequilibrium

HLA-B47 association with congenital adrenal hyperplasia is due to deletion of *CYP21A2* on HLA-B47DR7 haplotype

HLA-B14 association with late-onset adrenal hyperplasia is due to an exon 7 missense mutation (V281L) in *CYP21A2* on HLA-B14DR1 haplotype

An HLA association was found in congenital adrenal hyperplasia, but it was not thought to be an immune disease

Cancer Susceptibility

HLA Associations with Cancer Susceptibility

Patterns of Human Leukocyte Antigen Class I and Class II Associations and Cancer

Zhiwei Liu[1], Andriy Derkach[1,2], Kelly J. Yu[1], Meredith Yeager[3,4], Yu-Sun Chang[5,6], Chien-Jen Chen[7], Ulf Gyllensten[8], Qing Lan[1], Mei-Hsuan Lee[9], James D. McKay[10], Nathaniel Rothman[1], Hwai-I Yang[7,9,11], Allan Hildesheim[1], and Ruth M. Pfeiffer[1]

Associations for most cancers with infectious etiology or of hematopoietic origin were driven by multiple HLA regions, suggesting that both **cytotoxic and helper T-cell responses** are important.

A Genome-Wide Association Study Identified Novel Genetic Susceptibility Loci for Oral Cancer in Taiwan

Da-Tian Bau [1,2,3], Ting-Yuan Liu [4], Chia-Wen Tsai [1,2], Wen-Shin Chang [1,2], Jian Gu [5], Jai-Sing Yang [4], Liang-Chun Shih [1,2] and Fuu-Jen Tsai [6,7,*]

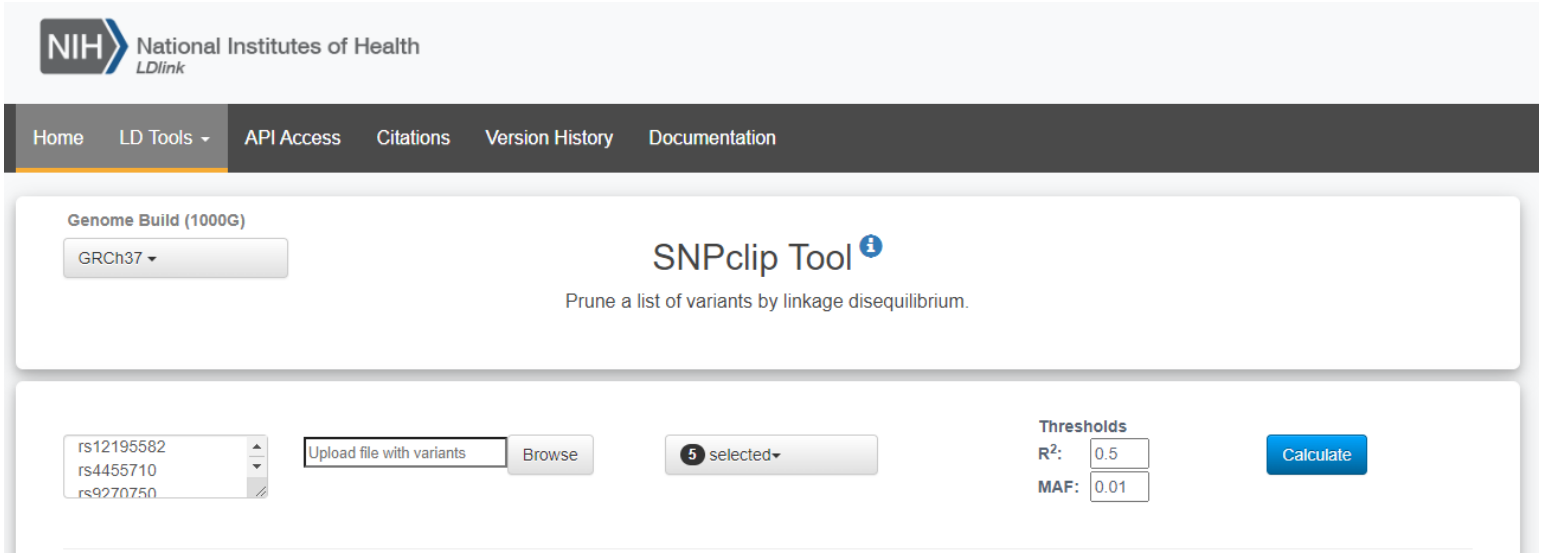
We confirmed two previously reported loci on the 6p21.33 (HLA-B) and 6p21.32 (HLA-DQ gene cluster) loci, highlighting the importance of the human leukocyte antigen and, hence, the **immunologic mechanisms** in oral carcinogenesis.

Kingston University London

HLA Polymorphisms and Disease

Descriptive Statistics

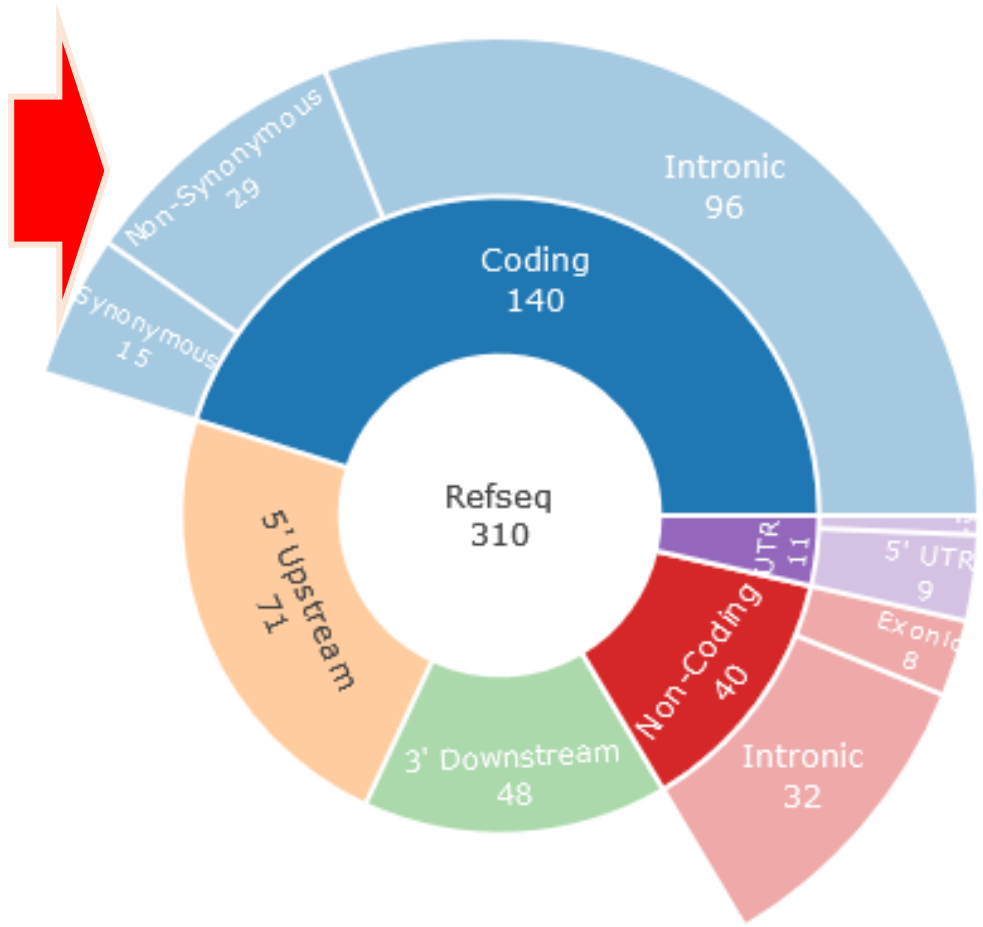
555 889 associations in the whole catalog
13 555 xMHC associations
612 cancer-associations with xMHC SNPs
473 unique cancer-associated xMHC SNPs
After LD pruning with SNPclip [r^2 = 0.60; all populations]: 284 left



Kingston University London

HLA Polymorphisms and Disease

Descriptive Statistics

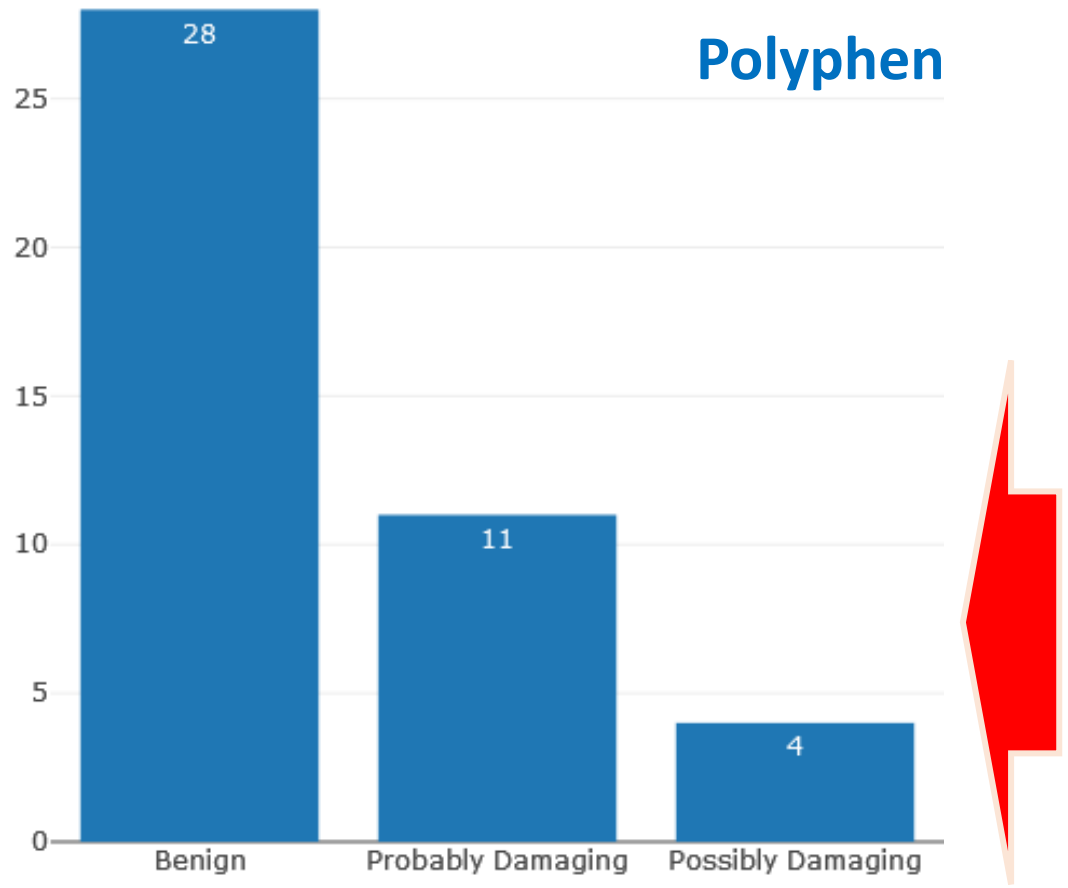


Minority of these disease-associated SNPs are in coding regions; therefore, not directly determining the HLA type

Kingston University London

HLA Polymorphisms and Disease

Descriptive Statistics



Majority of these coding region disease-associated SNPs do not have a deleterious effect on protein function

Kingston University London

HLA Polymorphisms and Disease

Descriptive Statistics

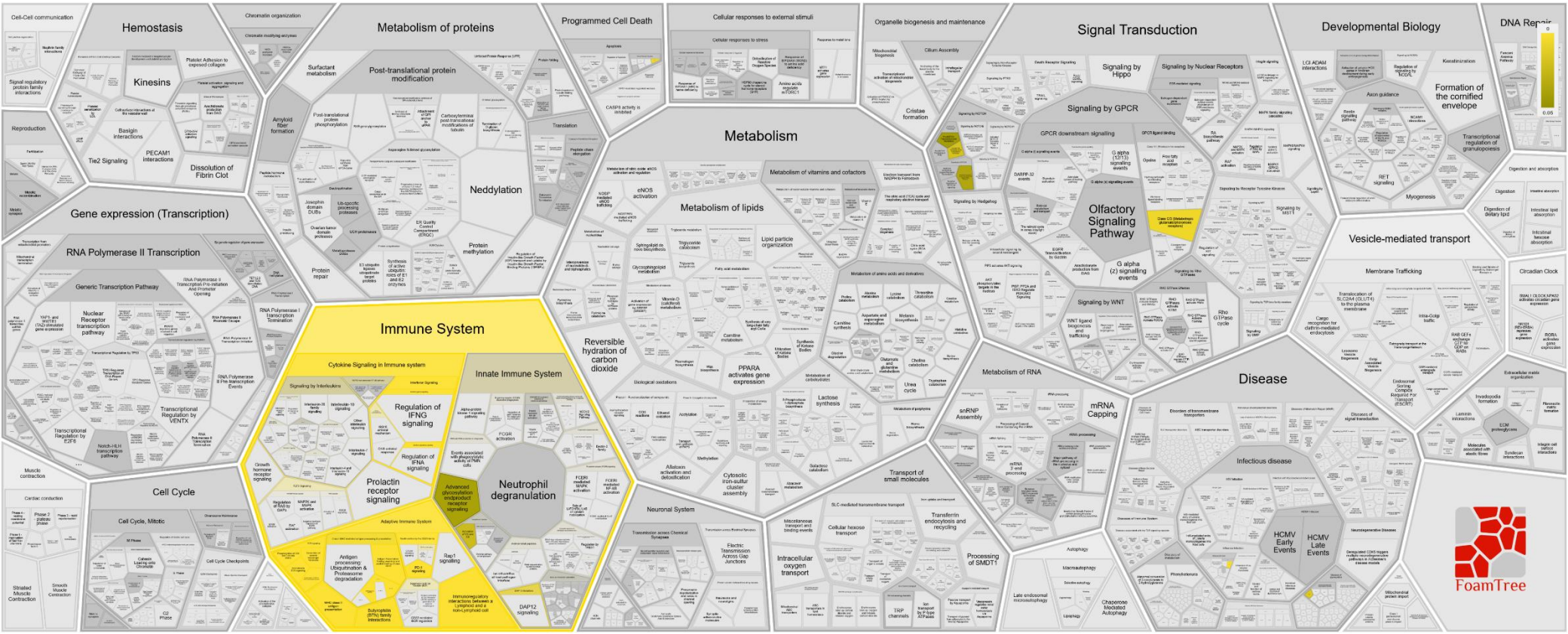
AGER	GRM4	LTA	PPP1R10	TRIM15
AIF1	HCG14	LTB	PRRC2A	TRIM31
ANKS1A	HCG17	MAS1LP1	PRSS16	TRIM31-AS1
APOM	HCG20	MCCD1P1	PSORS1C1	TSBP1
ATAT1	HCG27	**MDC1**	PSORS1C2	TSBP1-AS1
BAG6	HCG4P8	MICA	PTMAP1	TUBB
BAK1	HCG9	MICB	RF02219	UHRF1BP1
BTN2A1	HCP5	MICD	RNA5SP206	USP8P1
BTNL2	HIST1H1E	MIR6891	RNU6-283P	VPS52
C2	HIST1H2AC	MPIG6B	RPL3P2	ZKSCAN3
C6orf106	HIST1H2AJ	**MSH5**	RPS18	ZNF184
C6orf15	HIST1H2BD	**MSH5-SAPCD1**	SAPCD1	ZNRD1
CARMIL1	HIST1H2BM	MUC21	SAPCD1-AS1	ZNRD1ASP
CCHCR1	HIST1H3H	MUC22	SKIV2L	ZSCAN31
CDSN	HIST1H4H	NCR3	SLC17A3	
COL11A2P1	**HSPA1A**	NEU1	SLC17A4	
DDX39BP1	**HSPA1B**	**NOTCH4**	SPDEF	
DHFRP2	HTATSF1P1	OR2H2	TAP2	
GABBR1	LINC00243	OR5V1	**TNF**	
GGNBP1	LINC01149	**PBX2**	TNXB	
	LINC02571	PGBD1		

Directly cancer-related non-HLA gens genes are implicated by these cancer-associated SNPs

Kingston University London

HLA Polymorphisms and Disease

Reactome (Pathway) Analysis

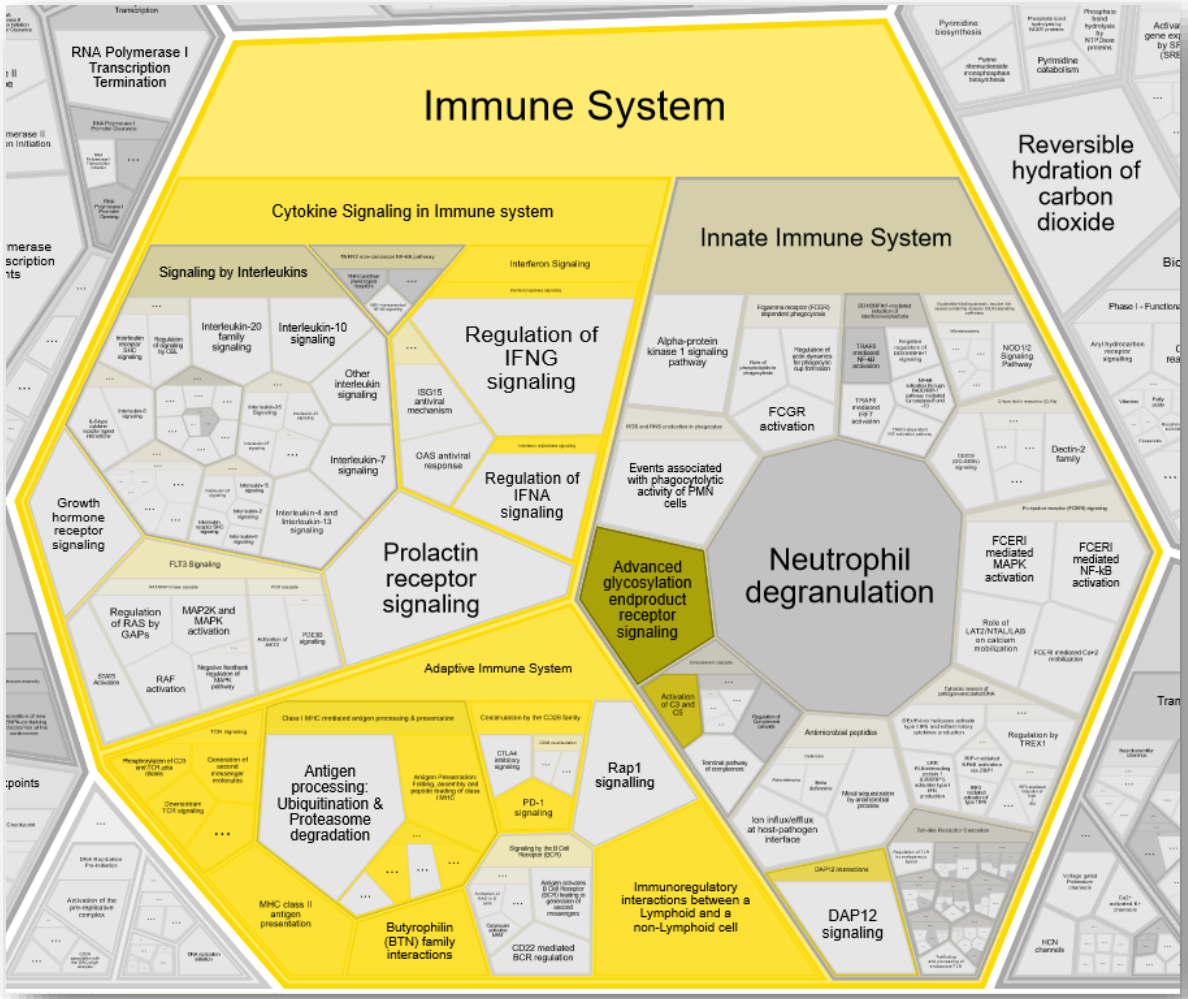


Pathway analysis based on nearest gene algorithm implicate immune system pathways including adaptive immune response, but not exclusively

Kingston University London

HLA Polymorphisms and Disease

Reactome (Pathway) Analysis



Kingston University London

Meaning of HLA Region Associations

Different Scenarios

Associations with HLA antigens

Associations with HLA lineages

Associations with non-HLA genes

Kingston University London

Meaning of HLA Region Associations

Associations with HLA Antigens

Table 5. List of HLA alleles and their tag SNPs that are described in the GWAS catalog[a]

SNP	*HLA allele*	r^2	*Function of SNP*	*Genes*	*Reported trait of SNP*	*Analyzed ethnicity*	*Reference*
rs2860580	HLA-A*11:01	0.87	Unknown	—	Nasopharyngeal carcinoma	Southern Chinese descent	31
rs9263739	HLA-C*12:02, HLA-B*52:01	0.94	Intron	*CCHCR1*	Ulcerative colitis	Japanese	11
rs1265112	HLA-C*04:01	0.83	Intron	*CCHCR1*	Nevirapine-induced rash	HIV-infected Thai	32
rs4418214	HLA-B*13:01, HLA-DRB1*12:02	1.0	Unknown	—	HIV-1 control	Caucasian	33
rs2255221	HLA-C*14:03, HLA-B*44:03	1.0	Missense	*HCP5*	HIV-1 control	African-American	33
rs10484561	HLA-DQB1*05:01	1.0	Unknown	—	Follicular lymphoma	Caucasian	34
rs10484561	HLA-DRB1*01:01	0.92	Unknown	—	Follicular lymphoma	Caucasian	34
rs11752643	HLA-DRB1*13:02	0.87	Unknown	—	Coronary heart disease	Japanese	12
rs11752643	HLA-DQB1*06:04	0.93	Unknown	—	Coronary heart disease	Japanese	12
rs2281388	HLA-DPB1*05:01	1.0	Unknown	—	Graves disease	Chinese Han	14

Abbreviations: GWAS, genome-wide association study; HIV, human immunodeficiency virus; HLA, human leukocyte antigen; SNP, single-nucleotide polymorphism. [a]The GWAS catalog is based on the Table Browser of the UCSC Genome Bioinformatics database (GRCh37/hg19).

Genes and Immunity (2012) **13,** 543–548

www.nature.com/gene

ORIGINAL ARTICLE

HLA and SNP haplotype mapping in the Japanese population

H Kitajima, M Sonoda and K Yamamoto

Kingston University London

Meaning of HLA Region Associations

Associations with HLA Lineages/Haplotypes

Most SNPs were not exclusive to HLA alleles/haplotypes/lineages, but the SNPs associated with lymphoid malignancies, nasopharyngeal cancer, lung cancer, and prostate cancer showed some correlations.

HLA-DRA rs2395185 ~ HLA-DRB4 lineage

(Hodgkin lymphoma, lung cancer)

BAG6 rs3117582 ~ HLA-A1-B8-DR3 (lung cancer)

BTNL2 rs28362675 ~ HLA-B52-DR15 (prostate cancer)

*** * ***

HLA-DPB1 rs2281389 ~ HLA-DPB1*0301 (Hodgkin lymphoma)

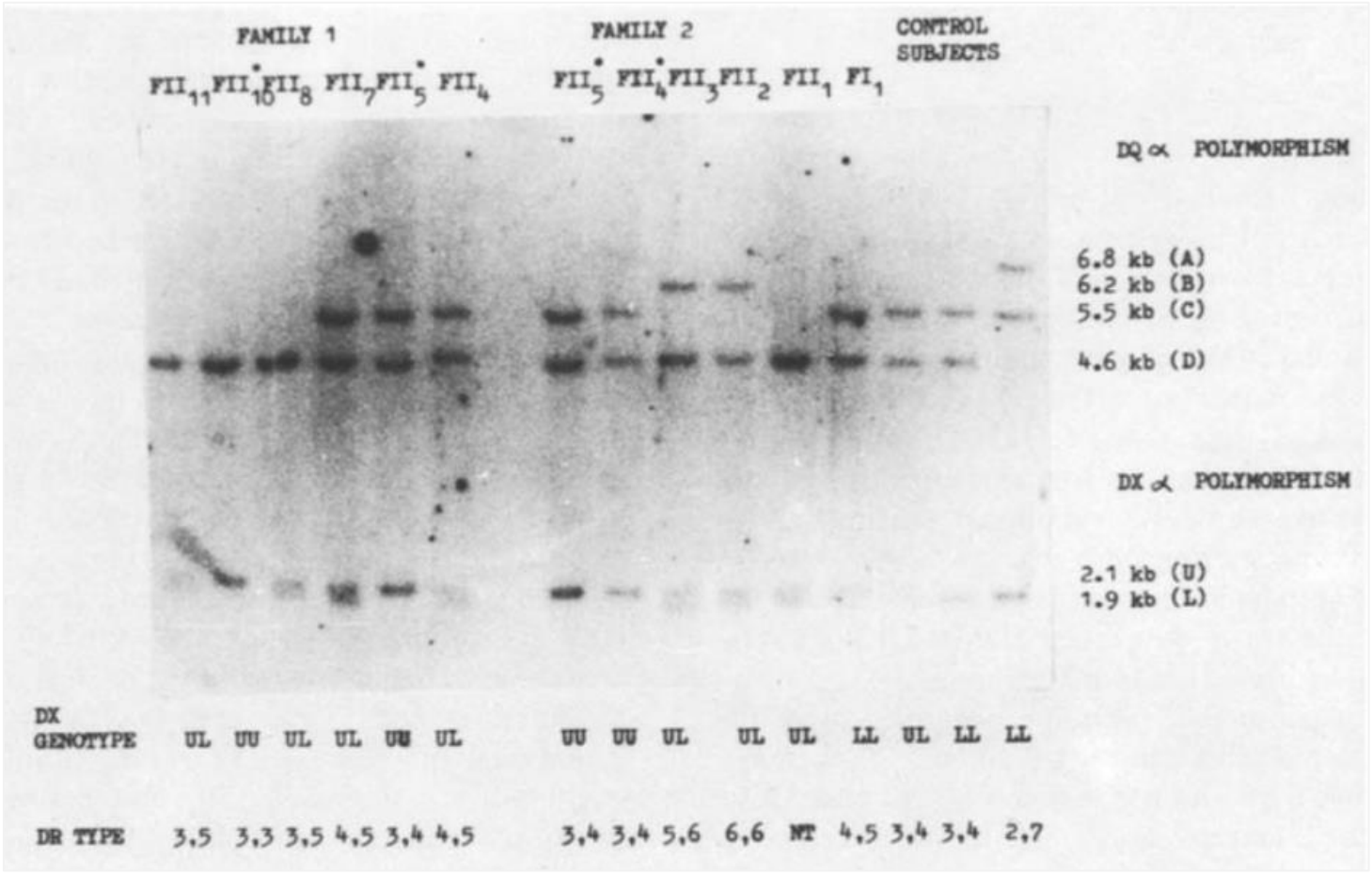
Correlations of Complex Disease-associated HLA Region SNPs with HLA Alleles

Amy E. Kennedy, Sandeep K. Singh, Malaroviyam Samikkannu, M. Tevfik Dorak
Department of Environmental and Occupational Health, Robert Stempel College of Public Health and Social Work, Florida International University, Miami 33199, USA

Kingston University London

HLA-DRB4 (DR53) Lineage

An Ancestral HLA Lineage



A DR3-related DX$_\alpha$ gene polymorphism strongly associates with insulin-dependent diabetes mellitus

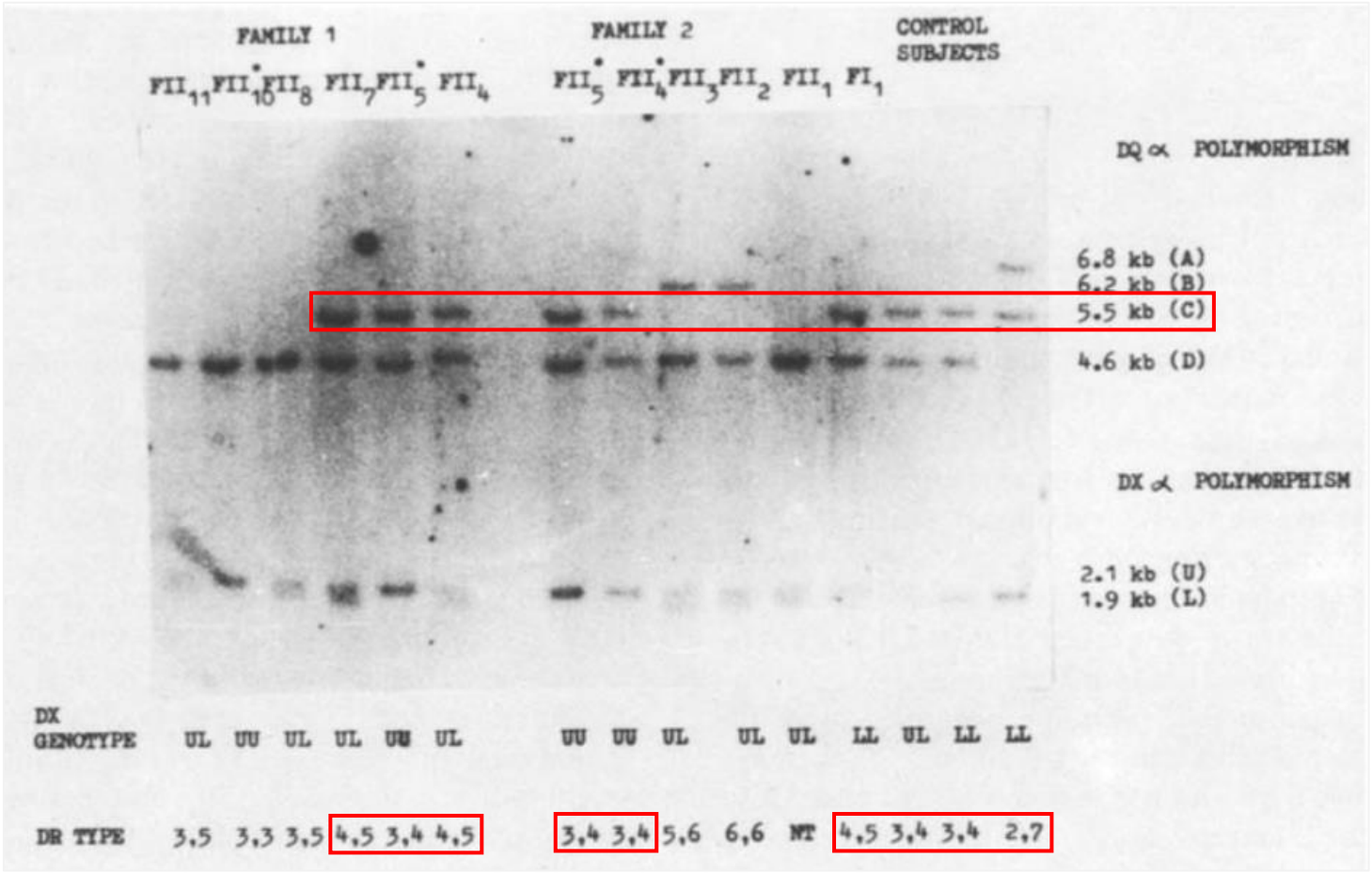
G. A. Hitman, J. Sachs, P. Cassell, J. Awad, G. F. Bottazzo, A. C. Tarn, G. Schwartz, J. P. Monson & H. Festenstein

Immunogenetics **23**, 47–51 (1986) | Cite this article

Kingston University London

HLA-DRB4 (DR53) Lineage

An Ancestral HLA Lineage



A DR3-related DX$_\alpha$ gene polymorphism strongly associates with insulin-dependent diabetes mellitus

G. A. Hitman, J. Sachs, P. Cassell, J. Awad, G. F. Bottazzo, A. C. Tarn, G. Schwartz, J. P. Monson & H. Festenstein

Immunogenetics **23**, 47–51 (1986) | Cite this article

Kingston University London

HLA-DRB4 (DR53) Lineage

An Ancestral HLA Lineage

HLA-DR53 Fact File

M.Tevfik Dorak, M.D., Ph.D.

Search for a Leukemia Susceptibility Gene in the HLA Complex

Chronic Myeloid Leukemia (CML): Scotland and Turkey

Chronic Lymphoid Leukemia (CLL): Germany

Childhood Acute Lymphoblastic Leukemia (ALL): Wales, Scotland and Turkey

All showed some form of HLA-DRB4 (DR53) association

HLA-DRB4 (DR53) Lineage

An Ancestral HLA Lineage

IMMUNOBIOLOGY

Unravelling an HLA-DR Association in Childhood Acute Lymphoblastic Leukemia

By M. Tevfik Dorak, Tom Lawson, Helmut K.G. Machulla, Chris Darke, Ken I. Mills, and Alan K. Burnett

Genetic and environmental factors play an interactive role in the development of childhood acute lymphoblastic leukemia (ALL). Since the demonstration of a major histocompatibility complex (MHC) influence on mouse leukemia in 1964, an HLA association has been considered as a possible genetic risk factor. Despite extensive efforts, however, no strong evidence comparable to the H-2^k influence on mouse leukemia has been shown. The number of negative serological studies resulted in a loss of interest and consequently, no molecular HLA-DR association study has been published to date. We reconsidered the HLA-DR association in childhood ALL in 114 patients from a single center and 325 local newborn controls by polymerase chain reaction (PCR) analysis of the HLA-DRB1/3/4/5 loci. With conventional analysis, there was a moderate allelic association with the most common allele in the HLA-DR53 group, HLA-DRB1*04, in the whole group that was stronger in males (P = .0005, odds ratio = 2.9). When the other expressed HLA-DRB loci were examined, homozygosity for HLA-DRB4*01, encoding the HLA-DR53 specificity, was increased in patients (21.1% *v* 8.3%; odds ratio = 2.9, P = .0005). Consideration of gender showed that all of these associations were reflections of a male-specific increase in homozygosity for HLA-DRB4*01 (32.8% *v* 4.0%; odds ratio = 11.7, 95% confidence interval [CI] = 4.9 to 28.0; $P = 3 \times 10^{-8}$). This highly significant result provided the long-suspected evidence for the HLA-DR influence on the development of childhood ALL while confirming the recessive nature of the MHC influence on human leukemogenesis as in experimental models. The cross-reactivity between HLA-DR53 and H-$2E^k$, extensive mimicry of the immunodominant epitope of HLA-DR53 by several carcinogenic viruses, and the extra amount of DNA in the vicinity of the HLA-DRB4 gene argue for the case that HLA-DRB4*01 may be one of the genetic risk factors for childhood ALL.

Blood, Vol 94, No 2 (July 15), 1999: pp 694-700

Dorak et al, *Blood* 1999 (www)

HLA-DRB4 (DR53) Lineage

An Ancestral HLA Lineage

Despite the obvious effect modification by sex, and recessive nature of most MHC associations in childhood leukemia, old habits are maintained, and most studies only compare all cases with all controls, and using only one genetic model. And, find nothing!

Kingston University London

HLA-DRB4 (DR53) Lineage

An Ancestral HLA Lineage

Comment on:

Genome-Wide Association Study of Classical Hodgkin Lymphoma and Epstein–Barr Virus Status–Defined Subgroups FREE

Kevin Y. Urayama, Ruth F. Jarrett, Henrik Hjalgrim, Arjan Diepstra, Yoichiro Kamatani, Amelie Chabrier, Valerie Gaborieau, Anne Boland, Alexandra Nieters, Nikolaus Becker ... Show more

JNCI: Journal of the National Cancer Institute, Volume 104, Issue 3, 8 February 2012, Pages 240–253, https://doi.org/10.1093/jnci/djr516

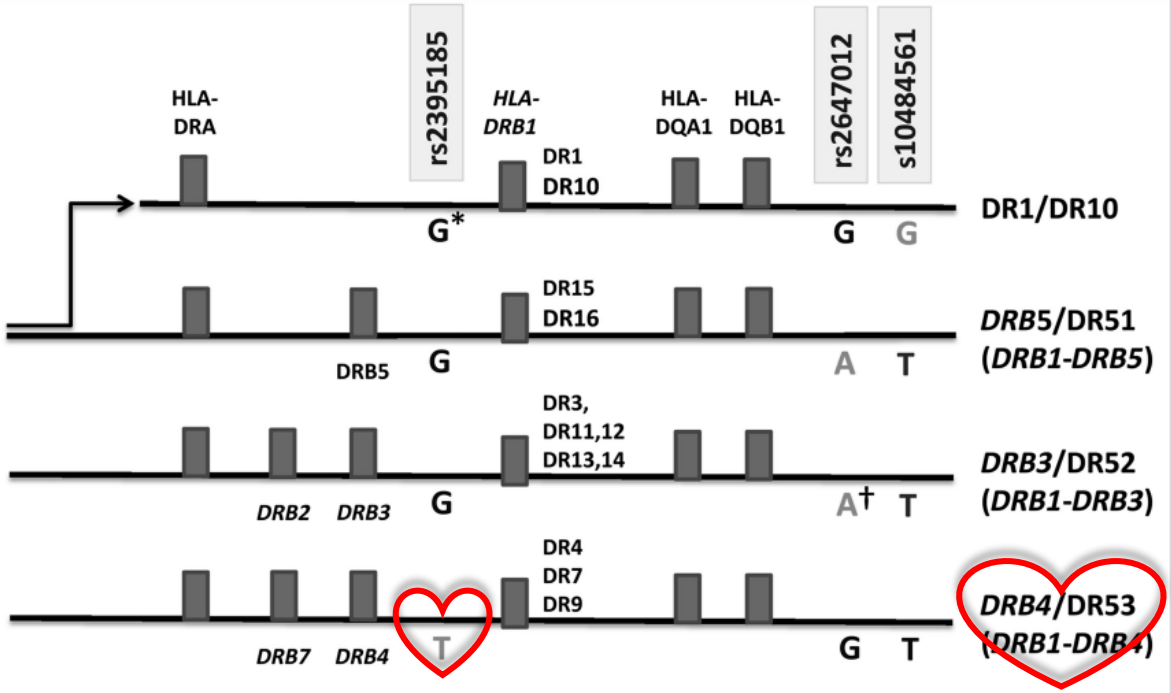


Figure 1. Schematic illustration of HLA-DR/DQ subregion structure and allelic composition in ancestral lineages. The DR1/DR10 lineage is closely related to the *DRB5*/DR51 lineage, but lost the *DRB5* gene. The only lineage not shown is the DR8 lineage, which is believed to have separated from the *DRB3*/DR52 lineage and lost the *DRB3* gene. The *DRB1* types belonging to each lineage are shown at the broad specificity level corresponding to serological specificities next to the *HLA-DRB1* gene. Lineage designations are given on the right. The figure is not to scale. Besides the *DRB4*/DR53 lineage, the rare *DRB1**0102 haplotype carries allele T of rs2395185 (**asterisk**). The majority but not all of the *DRB3*/DR52 haplotypes carry the variant allele G (**dagger**). On the rare *DRB1**08 haplotypes that are related to the *DRB3*/DR52 lineage but do not carry the *DRB3* gene, the allele at this single-nucleotide polymorphism is the common allele A.

Kennedy, Singh & Dorak, *JNCI* 1999 (www)

Kingston University London

HLA-DRB4 (DR53) Lineage

An Ancestral HLA Lineage

***HLA-DRB4* family of haplotypes are functionally suboptimal in their antigen presentation role**

***HLA-DRB4* family of haplotypes express *DRB1* and *DRB4* genes at a lower level**

Louis, 1994 (www); Vincent, 1996 (www) & 1997 (www)

The lowest cumulative expression of *HLA-DRB* genes (even lower in homozygotes) may result in:
- DRα-DQβ mixed isotype heterodimer formation
- Aberrant T-cell response and autoimmune reactions

Charron, 1984 (www); Lotteau, 1987 (www) & 1989 (www); Matsunaga, 1990 (www); Spencer 1989 (www) & 1992 (www) & 1993 (www)

Kingston University London

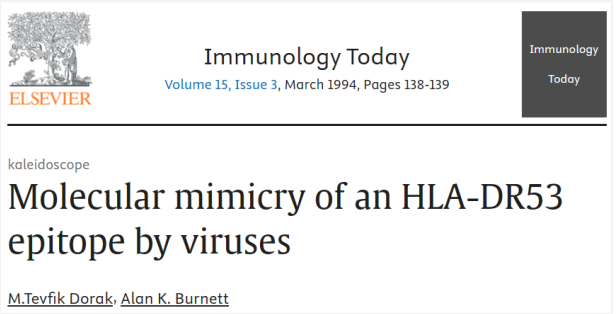
HLA-DRB4 (DR53) Lineage

An Ancestral HLA Lineage

HLA-DRB4 HVR3 epitope (LLERRRAE) is mimicked in its entirety by adenovirus and EBV

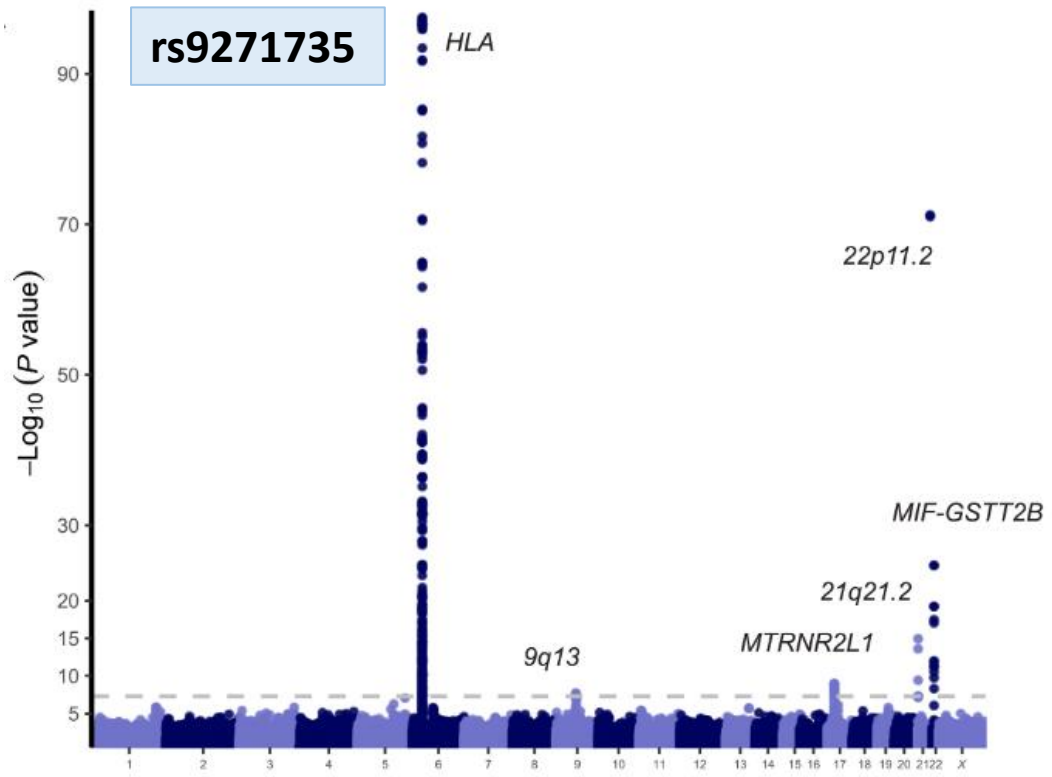
Table 1. Epitope sharing of HLA-DR53 by non-HLA proteins

Amino acid sequence*	Protein
D L L E R R R A E V	HLA-DR53 antigen, HV3 epitope
R L L E R R R A A A	Adenovirus type 5, 14.7 kDa protein
R L L E R R K A A S	Adenovirus type 2, 14.7 kDa protein
G L L E K R R A A E	EBV, large tegument protein
P L L E R R R G P L	HIV, zinc-finger protein 40
K A R R R R R A E K	HTLV-1, protein B
Y L L K R R R K A I	HPV-35, L2 minor capsid protein



HLA-DRB4 (DR53) Lineage

An Ancestral HLA Lineage



rs9271735 represents the HLA-DR53 lineage, which is the protective marker for Alzheimer disease

		rs2395185 chr6:32433167 G	rs2395185 chr6:32433167 T	
rs9271735 chr6:32593649	A	32	308	340 (0.338)
	G	651	15	666 (0.662)
		683 (0.679)	323 (0.321)	1006

Haplotypes		Statistics	
G_G:	651 (0.647)	D':	0.9299
A_T:	308 (0.306)	R^2:	0.8009
A_G:	32 (0.032)	Chi-sq:	805.7553
G_T:	15 (0.015)	p-value:	<0.0001

The genetic determinants of recurrent somatic mutations in 43,693 blood genomes

JOSHUA S. WEINSTOCK, CECELIA A. LAURIE, JAI G. BROOME, KENT D. TAYLOR, XIUQING GUO, ALAN R. SHULDINER, JEFFREY R. O'CONNELL, JOSHUA P. LEWIS, ERIC BOERWINKLE, [...], AND NHLBI TRANS-OMICS FOR PRECISION MEDICINE (TOPMED) CONSORTIUM +78 authors Authors Info & Affiliations

SCIENCE ADVANCES · 26 Apr 2023 · Vol 9, Issue 17 · DOI: 10.1126/sciadv.abm4945

Kingston University London

HLA-DRB4 (DR53) Lineage

The strongest trans-meQTLs in the genome are in xMHC

trans-meQTLs

	P Value	SNPName	SNPChr	SNPChrPos	ProbeName	ProbeChr	ProbeCenterChrPos	CisTrans	SNPType	Allele	OverallZScore
2	3.27167E-310	rs9268923	6	32432835	cg10154826	6	17601018	trans	C/T	T	61.2544774
3	3.27167E-310	rs2395185	6	32433167	cg10154826	6	17601018	trans	G/T	T	61.2544774
4	3.27167E-310	rs9268853	6	32429643	cg10154826	6	17601018	trans	T/C	C	61.1446653
5	3.27167E-310	rs477515	6	32569691	cg10154826	6	17601018	trans	G/A	A	58.2171575
6	3.27167E-310	rs4239020	17	80176641	cg07393940	7					
7	3.27167E-310	rs12342831	9	33124872	cg20290983	6					
8	3.27167E-310	rs10813951	9	33128021	cg20290983	6					
9	3.27167E-310	rs12342831	9	33124872	cg04842962	6					
10	3.27167E-310	rs10813951	9	33128021	cg04842962	6					
11	3.27167E-310	rs3780486	9	33139453	cg20290983	6					
12	3.27167E-310	rs3780486	9	33139453	cg04842962	6					
13	3.27167E-310	rs10813957	9	33153527	cg20290983	6					
14	3.27167E-310	rs10813957	9	33153527	cg04842962	6					
15	3.27167E-310	rs5756504	22	37467270	cg13737042	9					
16	3.27167E-310	rs5756506	22	37467392	cg13737042	9					
17	3.27167E-310	rs62103177	18	77624479	cg05926928	17					
18	3.27167E-310	rs8321	6	30032522	cg01620082	3	125678431	trans	A/C	C	-46.813694
19	3.27167E-310	rs2074977	19	3434028	cg08382705	11	45687343	trans	C/A	C	-46.6429193
20	3.27167E-310	rs8321	6	30032522	cg06606381	12	133084921	trans	A/C	C	-46.2606567

RS Number	Position (GRCh37)	Allele Frequencies	Haplotypes			
rs9268853	chr6:32429643	T=0.679, C=0.321	T	C	C	T
rs9268923	chr6:32432835	C=0.679, T=0.321	C	T	T	C
rs2395185	chr6:32433167	G=0.679, T=0.321	G	T	T	G
rs477515	chr6:32569691	G=0.685, A=0.315	G	A	G	A
		Haplotype Count	667	301	22	16
		Haplotype Frequency	0.663	0.2992	0.0219	0.0159

HLA-DR53 family members have strong genome-wide trans-meQTL effects in monocytes

BIOS QTL browser

This web page accompanies the manuscripts titled *'Disease variants alter transcription factor levels and methylation of their binding sites'*, by *Bonder & Luijk et al* and *'Unbiased identification of regulatory modifiers of genetic risk factors'*, by *Zhernakova et al*, both have been submitted to Nature Genetics.
For further questions, contact the corresponding author: lude@ludesign.nl

Download meQTL results

You can download the full *cis*- and *trans*-meQTL and eQTMs, detected at a false-discovery rate of 0.05:
Cis-meQTLs
Cis-eQTMs
Trans-meQTLs

Kingston University London

Disease variants alter transcription factor levels and methylation of their binding sites

Marc Jan Bonder[1,22], René Luijk[2,22], Daria V Zhernakova[1], Matthijs Moed[2], Patrick Deelen[1,3], Martijn Vermaat[4], Maarten van Iterson[2], Freerk van Dijk[1,3], Michiel van Galen[3], Jan Bot[5], Roderick C Slieker[2], P Mila Jhamai[6], Michael Verbiest[3], H Eka D Suchiman[2], Marijn Verkerk[6], Ruud van der Breggen[2], Jeroen van Rooij[6], Nico Lakenberg[2], Wibowo Arindrarto[7], Szymon M Kielbasa[8], Iris Jonkers[1], Peter van 't Hof[8], Irene Nooren[5], Marian Beekman[2], Joris Deelen[2], Diana van Heemst[9], Alexandra Zhernakova[1], Ettje F Tigchelaar[1], Morris A Swertz[1,3], Albert Hofman[10], André G Uitterlinden[6], René Pool[11], Jenny van Dongen[11], Jouke J Hottenga[11], Coen D A Stehouwer[12,13], Carla J H van der Kallen[12,13], Casper G Schalkwijk[12,13], Leonard H van den Berg[14], Erik W van Zwet[7], Hailiang Mei[8], Yang Li[1], Mathieu Lemire[15], Thomas J Hudson[15–17], the BIOS Consortium[18], P Eline Slagboom[2], Cisca Wijmenga[1], Jan H Veldink[14], Marleen M J van Greevenbroek[12,13], Cornelia M van Duijn[19], Dorret I Boomsma[11], Aaron Isaacs[13,19,20], Rick Jansen[21], Joyce B J van Meurs[6], Peter A C 't Hoen[4,23], Lude Franke[1,23] & Bastiaan T Heijmans[2,23]

HLA-DRB4 (DR53) Lineage

The strongest trans-eQTLs in the genome are in xMHC

Abstracts from the NCRI Cancer Conference

B137: ERG and XRCC6 expression levels are influenced by trans-eQTLs exclusively located in the HLA region

Izabela Stasik[1],Ken Mills[2],Mehmet Dorak[1]

[1]School of Health Sciences, Liverpool Hope University, Liverpool, UK,[2]Centre for Cancer Research and Cell Biology, Queens University Belfast, Belfast, UK

Results

We found 24 trans-eQTLs for XRCC6 and 20 for ERG at the genome-wide statistical thresholds, all mapped to the HLA complex in chromosome 6p21.3. Ten of those were trans-eQTLs for both XRCC6 and ERG and their effects on these genes were opposite. Surprisingly, XRCC6 and ERG did not have any cis-eQTLs in their neighborhood. Some of the trans-eQTLs are GWAS top hits for Hodgkin and non-Hodgkin lymphomas, as well as some other cancers. Most SNPs also correlated with expression levels of HLA genes, most notably HCG22, a gene encoding lincRNA (long intergenic non-coding RNA). None of the eQTLs was in a transcription factor or microRNA binding site. Five SNPs mapped to yet uncharacterised genes, one of which is a ncRNA gene, and another is a microRNA gene (miR-6891).

Conclusion

The extra-ordinary effect of HLA region trans-eQTLs on cancer-related genes XRCC6 and ERG may be mediated by ncRNA genes in the HLA complex. Regardless of the mechanism, our observations highlight the role played by the HLA region variants in cancer susceptibility.

An Unexpected Function of HLA Class II

HLA-DR Molecules and Signal Transduction

Intracellular MHC class II molecules promote TLR-triggered innate immune responses by maintaining activation of the kinase Btk

Xingguang Liu[1], Zhenzhen Zhan[1], Dong Li[2], Li Xu[1], Feng Ma[2], Peng Zhang[1], Hangping Yao[2] & Xuetao Cao[1,3]

The molecular mechanisms involved in the full activation of innate immunity achieved through Toll-like receptors (TLRs) remain to be fully elucidated. In addition to their classical antigen-presenting function, major histocompatibility complex (MHC) class II molecules might mediate reverse signaling. Here we report that deficiency in MHC class II attenuated the TLR-triggered production of proinflammatory cytokines and type I interferon in macrophages and dendritic cells, which protected mice from endotoxin shock. Intracellular MHC class II molecules interacted with the tyrosine kinase Btk via the costimulatory molecule CD40 and maintained Btk activation, but cell surface MHC class II molecules did not. Then, Btk interacted with the adaptor molecules MyD88 and TRIF and thereby promoted TLR signaling. Therefore, intracellular MHC class II molecules can act as adaptors, promoting full activation of TLR-triggered innate immune responses.

NATURE IMMUNOLOGY VOLUME 12 NUMBER 5 MAY 2011

An unexpected role for MHC class II

Ghada S Hassan & Walid Mourad

In addition to their classical function in antigen presentation and their lesser known ability to act as signal transducers, major histocompatibility complex class II molecules are now shown to promote Toll-like receptor signaling. This intriguing role requires intracellular association with the kinase Btk and the costimulatory molecule CD40.

NATURE IMMUNOLOGY VOLUME 12 NUMBER 5 MAY 2011

Kingston University London

An Unexpected Function of HLA Class II

HLA-DR Molecules and Signal Transduction

TLR-induced activation of Btk – Role for endosomal MHC class II molecules revealed

Joan Ní Gabhann[1], Caroline A Jefferies[1]

Cell Research **21**, 998–1001 (2011)

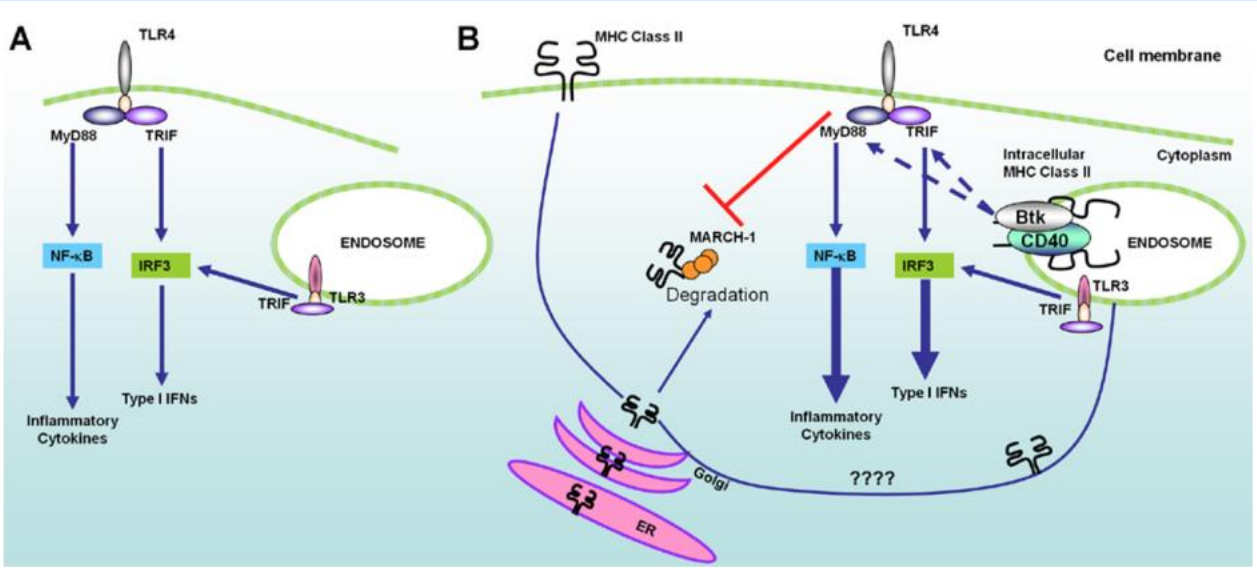


Figure 1 Schematic of crosstalk between TLR pathway and MHC molecules. **(A)** The traditional view of TLR signaling which holds that TLR ligation initiates several signaling cascades within the cell that drive the production of type I IFNs and pro-inflammatory cytokines that are essential for both innate and adaptive immune responses resulting in pathogen clearance. **(B)** Recent evidence has shown that a second pathway involving crosstalk between the TLR pathway and MHC pathway is initiated. This crosstalk results in the formation the Btk-CD40-MHC class II signaling complex at the endosome that enhances TLR-driven cytokine production. Ligation of TLRs also results in the inhibition of MARCH-1-mediated ubiquitination and degradation of MHC class II molecules, potentially increasing the pool of intracellular MHC class II molecules available to form the MHC-Btk signalosome at the endosome.

Kingston University London

An Unexpected Function of HLA Class II

HLA-DR Molecules and Signal Transduction

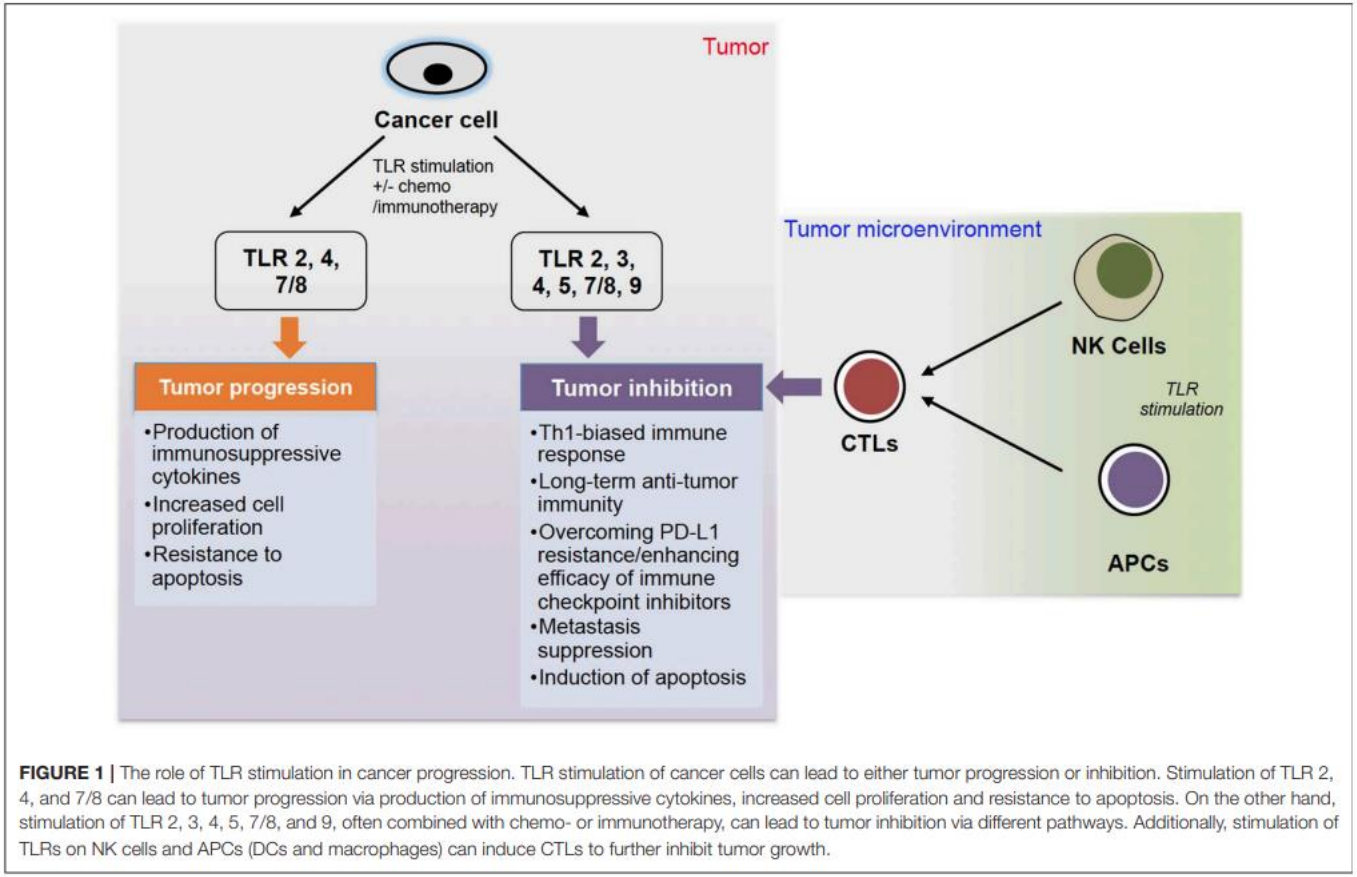


FIGURE 1 | The role of TLR stimulation in cancer progression. TLR stimulation of cancer cells can lead to either tumor progression or inhibition. Stimulation of TLR 2, 4, and 7/8 can lead to tumor progression via production of immunosuppressive cytokines, increased cell proliferation and resistance to apoptosis. On the other hand, stimulation of TLR 2, 3, 4, 5, 7/8, and 9, often combined with chemo- or immunotherapy, can lead to tumor inhibition via different pathways. Additionally, stimulation of TLRs on NK cells and APCs (DCs and macrophages) can induce CTLs to further inhibit tumor growth.

The Role of TLRs in Anti-cancer Immunity and Tumor Rejection

Zuzanna Urban-Wojciuk[1†], Mohd M. Khan[2,3†‡], Benjamin L. Oyler[3‡], Robin Fåhraeus[1,4,5,6], Natalia Marek-Trzonkowska[1,7], Aleksandra Nita-Lazar[2], Ted R. Hupp[1,6,8*] and David R. Goodlett[1,9*]*

Kingston University London

Alzheimer/Parkinson Disease

HLA Associations with AD/PD Susceptibility

PNAS RESEARCH ARTICLE | GENETICS

Multiancestry analysis of the HLA locus in Alzheimer's and Parkinson's diseases uncovers a shared adaptive immune response mediated by *HLA-DRB1*04* subtypes

Yann Le Guen, Guo Luo, Aditya Ambati, +146, and Emmanuel Mignot Authors Info & Affiliations

Contributed by Emmanuel Mignot; received March 2, 2023; accepted May 18, 2023; reviewed by Florence Clavaguera and Michel Goedert

August 29, 2023 120 (36) e2302720120 https://doi.org/10.1073/pnas.2302720120

Kingston University London

Alzheimer/Parkinson Disease

HLA Associations with AD/PD Susceptibility

PNAS RESEARCH ARTICLE | GENETICS

Table 1. *HLA-DRB1* alleles *HLA-DRB104:04 and *HLA-DRB1**04:01 are associated with a decreased risk of Parkinson's and ADs**

		PD				AD				PD + AD		
	HLA	FreqC	N	OR	*P*-val	FreqC	N	OR	*P*-val	OR	*P*-val	p-het
HLA Alleles	*HLA-DRB1**04:01	0.196	1,484,656	0.92[0.89; 0.95]	2.4E-08	0.191	486,478	0.93[0.9; 0.96]	6.4E-07	0.92[0.91; 0.94]	8.9E-14	0.56
	*HLA-DRB1**04:02	0.019	1,474,730	0.92[0.85; 0.99]	0.02	0.022	155,846	1.00[0.91; 1.10]	0.99	0.95[0.89; 1.01]	0.07	0.17
	*HLA-DRB1**04:03	0.012	980,868	0.89[0.81; 0.97]	0.01	0.072	7,587	1.09[0.91; 1.30]	0.34	0.93[0.85; 1.01]	0.07	0.04
	*HLA-DRB1**04:04	0.074	1,475,574	0.84[0.80; 0.88]	1.5E-11	0.073	476,236	0.86[0.82; 0.90]	8.9E-12	0.85[0.82; 0.88]	9.3E-22	0.60
	*HLA-DRB1**04:05	0.013	1,507,057	1.00[0.95; 1.05]	0.86	0.026	169,080	0.98[0.91; 1.06]	0.62	0.99[0.95; 1.03]	0.68	0.75
	*HLA-DRB1**04:06	0.046	32,327	0.95[0.82; 1.09]	0.46	0.058	7,587	0.95[0.78; 1.15]	0.60	0.95[0.85; 1.06]	0.37	0.99
	*HLA-DRB1**04:07	0.021	526,189	0.79[0.69; 0.91]	7.3E-04	0.019	474,840	0.88[0.81; 0.96]	4.5E-03	0.86[0.79; 0.92]	2.7E-05	0.18
	*HLA-DRB1**04:10	0.041	4,853	0.91[0.67; 1.25]	0.57	0.031	7,985	1.23[0.94; 1.59]	0.13	1.08[0.89; 1.32]	0.42	0.16
	*HLA-DRB1**01:01	0.181	1,485,033	1.05[1.01; 1.09]	7.0E-03	0.186	487,120	1.07[1.04; 1.10]	3.9E-07	1.06[1.04; 1.09]	1.3E-08	0.44
	*HLA-DQB1**03:02	0.191	1,501,065	0.91[0.88; 0.93]	2.6E-14	0.190	510,130	0.89[0.86; 0.91]	1.2E-19	0.90[0.88; 0.91]	4.7E-32	0.23
	*HLA-DQA1**03:01	0.177	1,507,147	0.89[0.87; 0.91]	2.5E-20	0.186	507,263	0.89[0.87; 0.91]	1.8E-19	0.89[0.88; 0.91]	3.E-37	0.79
	*HLA-DRB1**15:01	0.272	1,507,057	1.06[1.03; 1.10]	5.0E-04	0.263	485,383	1.02[0.99; 1.04]	0.13	1.03[1.01; 1.05]	1.1E-03	0.054

Table 2. *HLA-DRB1* H13/H33 amino acid is associated with reduced tau and neurofibrillary tangles and to a lesser extent with reduced Amyloid-β or neuritic plaques, when testing their association with AD neuropathology and CSF biomarkers

rs601945 is the SNP that shows a protective association with both AD and PD.

Multiancestry analysis of the HLA locus in Alzheimer's and Parkinson's diseases uncovers a shared adaptive immune response mediated by *HLA-DRB1*04* subtypes

Yann Le Guen, Guo Luo, Aditya Ambati, +146, and Emmanuel Mignot Authors Info & Affiliations

Kingston University London

Alzheimer/Parkinson Disease

HLA Associations with AD/PD Susceptibility

HLA-DR4 (the main allele within the DR53 family) is associated with decreased risk (represented by rs601945)
vs
HLA-DR15 (the main allele within the DR51 family) is associated with increased risk

Most HLA-cancer associations, including the first HLA and childhood leukemia and lung cancer risk, concern the HLA-DR53 lineage (presumably due to low level of expression of the alleles of this lineage)
and
almost none of the representatives of HLA-DR51 lineage (which shows the highest level of HLA-DR expression levels)

Kingston University London

An Unexpected Function of HLA Class II

HLA-DR Molecules and Signal Transduction

Journal of Neuroimmunology 348 (2020) 577362

Contents lists available at ScienceDirect

Journal of Neuroimmunology

journal homepage: www.elsevier.com/locate/jneuroim

Review Article

Toll-like receptors in Alzheimer's disease

Sara Momtazmanesh[a,b], George Perry[c], Nima Rezaei[d,e,f,*]

[a] School of Medicine, Tehran University of Medical Sciences, Tehran, Iran
[b] Neuroimmunology Research Association (NIRA), Universal Scientific Education and Research Network (USERN), Tehran, Iran
[c] Department of Biology and Neurosciences Institute, University of Texas at San Antonio (UTSA), San Antonio, TX, USA
[d] Research Center for Immunodeficiencies, Pediatrics Center of Excellence, Children's Medical Center, Tehran University of Medical Sciences, Tehran, Iran
[e] Department of Immunology, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran
[f] Network of Immunity in Infection, Malignancy and Autoimmunity (NIIMA), Universal Scientific Education and Research Network (USERN), Tehran, Iran

ARTICLE INFO

Keywords:
Alzheimer disease
Toll-like receptor
Microglia
Amyloid-beta
Inflammation
Single nucleotide polymorphisms

ABSTRACT

Neuroinflammation and microglial dysfunction are key contributors to the development of Alzheimer's disease (AD). Toll-like receptors (TLRs) are transmembrane proteins primarily involved in immune responses and expressed by several immune and non-immune cells within the central nervous system. Signaling of TLRs affects the core of AD changes, including synaptic plasticity, microglial activity, tau phosphorylation, and inflammatory responses. We reviewed the activity, expression, potential applications, and genetic polymorphisms of TLRs in AD. Activation of TLRs has shown both destructive and protective effects. Several genetic polymorphisms of TLRs have been also recognized as protective or risk factors for AD. We concluded that TLRs are one of the major components of AD pathogenesis, particularly in the early stages of the disease, which can provide novel therapeutic options.

Conclusions

The HLA region is biomedically one of the most important parts of the human genome

Its analysis requires special attention

Difficulties with analysis of HLA region data does not justify its exclusion from whole genome data analysis

HLA region associations means much more than immune response modifications

HLA region polymorphisms have genome-wide effects

Kingston University London

ACKNOWLEDGEMENTS

I am grateful to all who have taught, helped and supported in one way or another at

Glasgow Leukaemia Research Laboratory
Glasgow Tissue Typing Laboratory
Glasgow Royal Hospital for Sick Children
University of Wales College of Medicine
Welsh Transplantation & Immunogenetics Laboratory
HLA Lab, Martin Luther University, Halle, Germany
St Jude Children's Hospital, HLA Laboratory
University of Tennessee at Memphis
University of Alabama at Birmingham
Newcastle University (U.K.)
School of Clinical Medical Sciences (Child Health)
HUMIGEN LLC, The Institute for Genetic Immunology, Hamilton, NJ
Florida International University, Miami, FL
&
All Funding Bodies supported this research

ACKNOWLEDGEMENTS

FIU | Robert Stempel College of Public Health and Social Work