Research Paper

MULTIPLE SCLEROSIS JOURNAL | MSJ

Methylation differences at the *HLA-DRB1* locus in CD4+ T-Cells are associated with multiple sclerosis

Multiple Sclerosis Journal
2014, Vol. 20(8) 1033–1041

DOI: 10.1177/1352458513516529
msj.sagepub.com
SAGE

MC Graves[1], M Benton[2,3], RA Lea[1,2], M Boyle[4], L Tajouri[5], D Macartney-Coxson[3], RJ Scott[1,6] and J Lechner-Scott[1,7]

Abstract

Background: Multiple sclerosis (MS) is thought to be caused by T-cell mediated autoimmune dysfunction. Risk of developing MS is influenced by environmental and genetic factors. Modifiable differences in DNA methylation are recognized as epigenetic contributors to MS risk and may provide a valuable link between environmental exposure and inherited genetic systems.

Objectives and methods: To identify methylation changes associated with MS, we performed a genome-wide DNA methylation analysis of CD4+ T cells from 30 patients with relapsing–remitting MS and 28 healthy controls using Illumina 450K methylation arrays.

Results: A striking differential methylation signal was observed at chr. 6p21, with a peak signal at *HLA-DRB1*. After prioritisation, we identified a panel of 74 CpGs associated with MS in this cohort. Most notably we found evidence of a major effect CpG island in *DRB1* in MS cases ($p_{FDR} < 3 \times 10^{-3}$). In addition, we found 55 non-HLA CpGs that exhibited differential methylation, many of which localise to genes previously linked to MS.

Conclusions: Our findings provide the first evidence for association of DNA methylation at *HLA-DRB1* in relation to MS risk. Further studies are now warranted to validate and understand how these findings are involved in MS pathology.

Keywords

Epigenetics, multiple sclerosis, immunology, genetics, methylation

Date received: 3 May 2013; revised: 27 October 2013; accepted: 30 October 2013

Introduction

Multiple sclerosis (MS) is a neurological disease that stems from lymphocyte-mediated inflammation causing demyelination and axonal degeneration. The underlying pathogenesis of MS remains unclear although T lymphocytes (particularly CD4+ T-cells) have long been considered to play a pivotal role in orchestrating the self-reactive immune responses in MS patients.[1–3] Risk of developing MS is influenced both by genetic factors and environmental exposures. The primary environmental exposures known to modify MS risk are sunlight and related vitamin D levels, Epstein-Barr virus (EBV) and smoking.[4,5] Genetic susceptibility to MS is partly explained by haplotypic variation in the major histocompatibility complex (MHC) on chromosome (chr.) 6p21.[6] A recent meta-analysis of MS genome-wide association studies (GWAS) showed that seven separate loci in this region were associated with MS.[7] Not surprisingly, the largest effect single-nucleotide polymorphism (SNP) represented the well-described human leukocyte antigen haplotype (*HLA-DRB1**1501) in

[1]Centre for Information-Based Medicine, Hunter Medical Research Institute, Australia
[2]Griffith Health Institute, Griffith University, Australia
[3]Institute of Environmental Science and Research, New Zealand
[4]Department of Immunology, Division of Medicine, John Hunter Hospital, Australia
[5]Faculty of Health Sciences and Medicine, Bond University, Australia
[6]Division of Molecular Genetics, Hunter Area Pathology Service, Australia
[7]Department of Neurology, Division of Medicine, John Hunter Hospital, Australia

Corresponding author:
Jeannette Lechner-Scott, Department of Neurology, John Hunter Hospital, Locked Bag 1, Hunter Region Mail Centre, NSW 2310, Australia.
Email: Jeannette.lechner-scott@hnehealth.nsw.gov.au

the MHC region.[7] Despite the significant contribution of HLA and the more modest effect of the non-HLA loci, there remains a large proportion of unexplained heritability in terms of MS risk.[6]

Epigenetic mechanisms can cause a change in gene expression without differences in DNA sequence and can be inherited, established in utero or acquired throughout the lifespan.[8] Importantly, epigenetic mechanisms can be modified by environmental factors providing a potential link between external exposures and inherited genetic systems.[5] Recently a number of reviews have described the importance of epigenetic mechanisms in MS pathology.[4,5,9] DNA methylation at CpG dinucleotides is one important epigenetic mechanism and array-based technologies now exist to rapidly and reliably scan the human genome for variation in methylation levels. Genome-wide DNA methylation patterns have been shown to vary widely among lymphocyte subtypes.[10] To search for epigenetics markers of MS, Baranzini et al. measured the genome-wide methylation patterns in CD4+ T cells from three MS-discordant monozygotic twin pairs and found only a small fraction of differences among these twin pairs. However, small sample size and high stringency thresholds may have precluded detection of more subtle differences in methylation profiles in relation to MS.[11]

In an effort to identify differentially methylated CpGs associated with MS we performed a genome-wide DNA methylation study of CD4+ T-cells from MS patients and healthy controls. Identifying epigenetic loci associated with MS should inform questions of disease pathology and could also reveal potentially modifiable targets for environmental exposures or new drug design as well as highlight important candidates for blood-based biomarkers of MS risk and treatment response.

Results

Genome-wide methylation analysis

We recruited 30 relapsing–remitting MS (RRMS) patients and 28 healthy controls for this study. For all 58 subjects CD4+ T cells were isolated and genomic DNA was extracted and hybridised to Illumina 450K arrays. Raw fluorescence data were processed and 460,691 probes representing CpG sites spanning the genome and that passed QC were analysed by comparing methylation levels (β values) between MS cases and controls. Figure 1 shows the genome-wide distribution of differential methylation scores (Δ_{meth}) for all CpG sites tested. As expected, there were no CpGs that were completely different in terms of methylation state between cases and controls. However, observation of the genome-wide plot revealed a distinct differential methylation signal on chr. 6. Closer examination of chr. 6 showed that the signal localises to the MHC region on 6p21 (Figure 2). Not surprisingly, the peak signal within the MHC region maps to the well-described MS susceptibility gene *HLA-DRB1*. We also noted some more modest differential methylation peaks near *HLA-A* and *HLA-C* (Figure 3).

CpG prioritisation analysis

We proceeded to conduct a stepwise prioritisation strategy to extract the most robust CpG loci associated with MS. First we compared the entire group of cases and controls and extracted all CpGs that met both of our sets criteria of i) $p < 0.05$ and ii) $\Delta_{meth} = \pm 0.1$. Based on these thresholds, a total of 556 CpGs were extracted. In an effort to filter out potential effects of gender and drug treatment, we performed a sub-group analysis of the methylation statistics (see Supplementary methods). This process reduced the number of associated CpG sites down to a core panel of 74 (See Supplementary Table 1). Of these 74 CpGs, 59% showed decreased methylation in the case group compared to the control group and 25% represented the *MHC* region. The top ranked CpG site was located in the *HLA-DRB1* gene. At this site we observed approximately 37% decreased methylation in the MS case group compared to the control group ($\Delta_{meth} = -0.375$, $p_{FDR} = 4.31 \times 10^{-4}$). Of the 74 CpGs, 19 mapped to the MHC region on chr. 6 (Table 1). Interestingly, 10 of these CpGs were located within *HLA-DRB1* with eight clustered within a 565bp CpG island. Collectively these eight adjacent CpGs showed ~26% decreased methylation in cases compared to controls and may represent a major effect hypomethylation locus for MS. Differential methylation (both increased and decreased) was also observed at CpGs in *HLA-DRB5, HLA-DRB6* and *HLA-DQB1* (Table 1).

The effect of HLA-DRB1*1501 haplotype on DRB1 methylation and MS

Focusing on the *HLA-DRB1* gene, we genotyped a tagging SNP of the major MS risk haplotype (*HLA-DRB1*1501*) in all cases and controls. The observed genotype frequencies for *HLA-DRB1**1501 +/+, +/– and –/– were 0.10, 0.66, 0.24 and 0.03, 0.29, 0.68 for case and control groups, respectively. As expected, there was a much higher prevalence of risk haplotype carriers in the MS group compared to controls (76% vs 32%, respectively) and this difference was statistically significant even in this small sample (odds ratio (OR) =6.6, $p = 0.0005$). In the healthy control group, all 10 (of 74) MS-associated CpG sites within the *DRB1* gene showed statistically significant association with the *DRB1*1501* haplotype ($p < 0.0001$). Of the 10 CpGs, eight had decreased and two had increased methylation in the risk haplotype carriers compared to non-haplotype carriers in the control group, which is consistent with the direction of the MS associations shown in Table 1. These results suggest that the *DRB1* haplotype may influence the association we observed between the methylation level at *DRB1* CpGs and

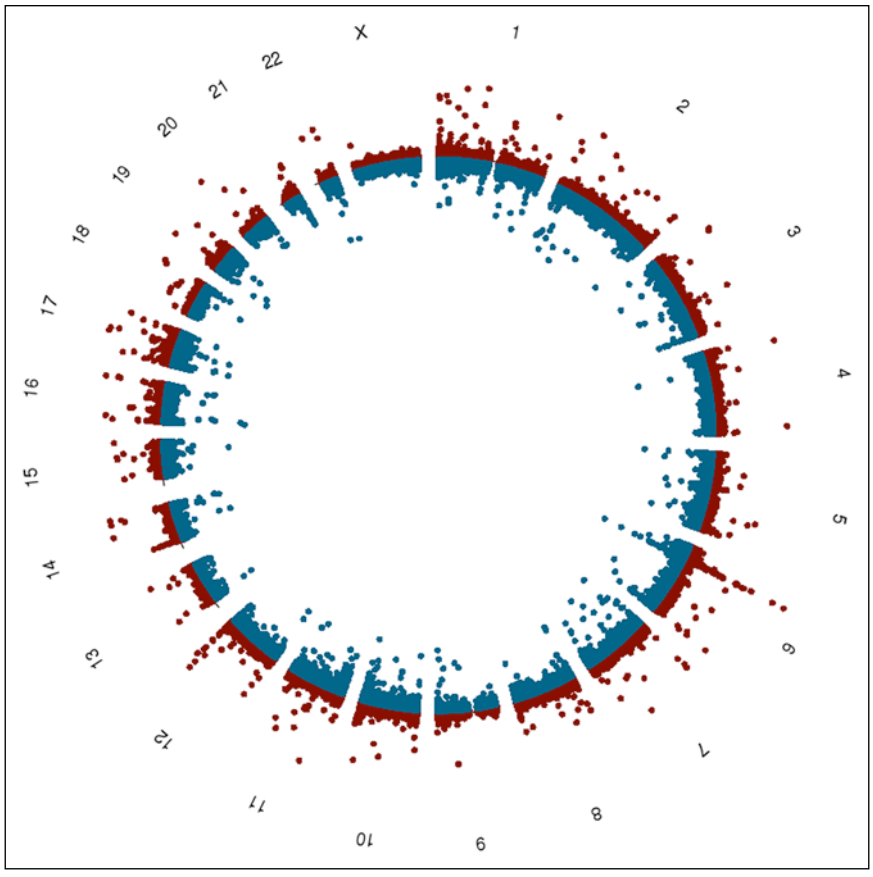

Figure 1. A genome-wide differential methylation plot. Data points outside the circle represent increased methylation in multiple sclerosis (MS) patients compared to controls (i.e. Δ_{meth}), whereas points inside the circle represent decreased methylation in the MS group. For colour plates go to: http://msj.sagepub.com/

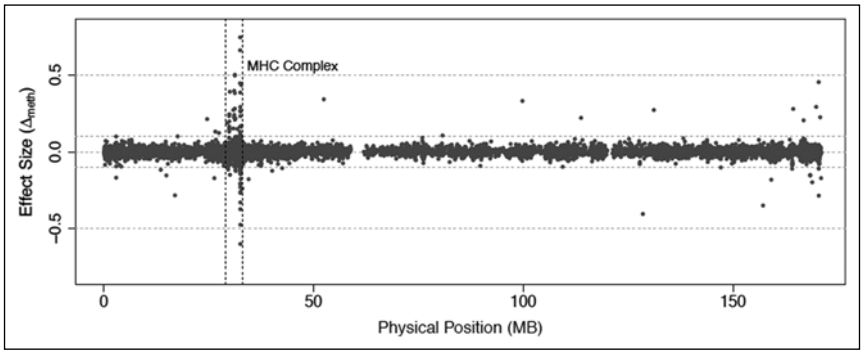

Figure 2. Differential methylation plot of chromosome (chr.) 6. Points above the line represent increased methylation in multiple sclerosis (MS) cases vs controls. A major peak (or methylation hotspot) is shown at the major histocompatibility complex (*MHC*) region on chr. 6p21. For colour plates go to: http://msj.sagepub.com/

MS risk. MS cases were not included in this sub-analysis because we were attempting to control for MS disease state when exploring the natural association between DRB1 genotype and methylation level.

After stratifying the cohort into only those subjects that carried the risk haplotype ($n = 31$ in total, 22 cases and nine controls), the MS associations remained for most of the *DRB1* CpG sites shown in Table 1 ($p < 0.05$) (See Supplementary Table 2). This implies that the hypomethylation signal we observed for these *DRB1* sites in MS patients is not entirely due to the *DRB1*1501* haplotype. More accurate testing of the effect modification of *DRB1*

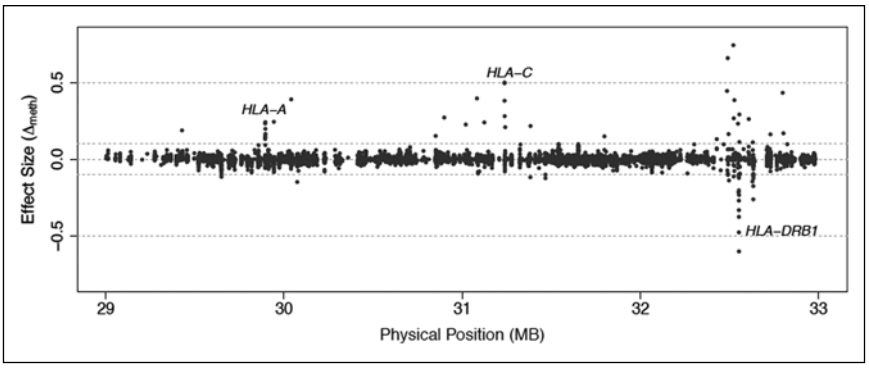

Figure 3. Differential methylation plot of all CpG sites spanning the major histocompatibility complex (*MHC*) region. Points above the line represent increased methylation in cases vs controls. A major peak was observed at human leukocyte antigen *(HLA) DRB1*. For colour plates go to: http://msj.sagepub.com/

Table 1. MS-associated CpGs within the MHC region.

Probe ID	Chr.	Cytoband	Position	Gene	Median (case)	Median (cont)	Δ_{meth}	*p* value	p_{FDR}
cg04985482	6	6p21.33	31382065	MICA	0.720962697	0.83825157	–0.117	0.009931067	0.017090674
cg11964823	6	6p21.33	31466726	MICB	0.301236465	0.412541254	–0.111	0.007002286	0.015702097
cg15837308	6	6p21.32	32427751		0.668098145	0.539321582	0.129	0.000412184	0.004357373
cg17369694	6	6p21.32	32485396	HLA-DRB5	0.553006501	0.08258889	0.47	0.018992279	0.023820825
cg01341801	6	6p21.32	32489203	HLA-DRB5	0.734269503	0.06594094	0.668	0.009931067	0.017090674
cg00440797	6	6p21.32	32493873	HLA-DRB5	0.402351504	0.238417111	0.164	0.046067146	0.046067146
cg21698879	6	6p21.32	32496326	HLA-DRB5	0.726314026	0.864969004	–0.139	0.002267586	0.011985811
cg11752699	6	6p21.32	32526669	HLA-DRB6	0.647165288	0.244310476	0.403	0.018992279	0.023820825
cg19575208	6	6p21.32	32551888	HLA-DRB1	0.611323918	0.814965836	–0.204	0.018992279	0.023820825
cg08845336	6	6p21.32	32551891	HLA-DRB1	0.585127768	0.866561696	–0.281	0.018992279	0.023820825
cg17316649	6	6p21.32	32552016	HLA-DRB1	0.640458075	0.768886044	–0.128	0.003355676	0.013795556
cg06032479	6	6p21.32	32552026	HLA-DRB1	0.650044687	0.75184766	–0.102	0.007002286	0.015702097
cg08578320	6	6p21.32	32552039	HLA-DRB1	0.425172532	0.800291811	–0.375	0.00000582	0.000430702
cg09139047	6	6p21.32	32552042	HLA-DRB1	0.365492952	0.846477585	–0.481	0.000162158	0.002999923
cg15982117	6	6p21.32	32552106	HLA-DRB1	0.410006789	0.745645398	–0.336	0.000162158	0.002999923
cg10632894	6	6p21.32	32552453	HLA-DRB1	0.634886055	0.862430721	–0.228	0.0000595	0.0022027
cg17416722	6	6p21.32	32554385	HLA-DRB1	0.377399997	0.070294588	0.307	0.001017522	0.007529664
cg24147543	6	6p21.32	32554481	HLA-DRB1	0.474777756	0.337005181	0.138	0.000644142	0.005958312
cg13423887	6	6p21.32	32632694	HLA-DQB1	0.276417103	0.531170806	–0.255	0.00152245	0.00938844

MS: multiple sclerosis; MHC: major histocompatibility complex; Δ_{meth}: differential methylation score; FDR: false discovery rate; ID: identification; Chr.: chromosome; HLA: human leukocyte antigen.

haplotype on methylation state in relation to MS risk will need to be determined with larger samples.

Differential methylation at CpGs outside of the MHC region

Of the 74 CpGs showing MS association after filtering, 55 CpGs resided at genomic regions outside of *MHC* on chr. 6p21 (Table 2). To explore the importance of these results, we collected annotation on all the CpG loci including cytoband and gene function data as well as previous MS and/or autoimmune disease association information. Nineteen of the 55 CpGs were found to be intergenic or mapped to genes of unknown function. Interestingly, of the remaining 36 CpGs, 30 mapped to genes that had previously been reported to have some relationship to MS. Nine CpGs, all showing increased methylation in MS patients, mapped to a 20 kb region within the T-cell antigen receptor (α-subunit) gene (TCRα) on chr. 14q11.2. Alleles of this gene have been previously associated with MS.[12] Moreover, a CpG on chr. 12p13.3 maps to the immune regulatory gene *SHP-1*, which has recently been shown to have hypermethylation in the promoter in relation to MS.[13] Two CpGs at chr. 16p13.3 map to the interleukin (IL)32 gene for which there is some prior evidence of allelic association to MS.[14]

Table 2. MS-associated CpGs outside of the MHC region.

Probe ID	Chr.	Cytoband	Position	Gene	Median (case)	Median (cont)	Δ_{meth}	*p* value	p_{FDR}
cg05392448	1	1p36.33	2266933	*MORN1*	0.896731115	0.734593262	0.162	0.034687695	0.036153372
cg12454169	2	2p23.1	30669597	*LCLAT1*	0.229812356	0.479496137	–0.25	0.007002286	0.015702097
cg11236515	2	2p13.1	74213762		0.297364947	0.13431677	0.163	0.001017522	0.007529664
cg04722349	2	2q21.2	133090251		0.689794453	0.565885889	0.124	0.009931067	0.017090674
cg06109482	2	2q37.1	232531087		0.718806772	0.827429359	–0.109	0.034687695	0.036153372
cg27051683	2	2q37.3	242802069	*PDCD1*	0.17966704	0.071497522	0.108	0.009931067	0.017090674
cg17322655	2	2q37.3	242802127	*PDCD1*	0.155577706	0.052761168	0.103	0.009931067	0.017090674
cg25702651	3	3q39	192675515		0.53256991	0.657108432	–0.125	0.026114596	0.029730463
cg27353899	3	3q39	195488922	*MUC4*	0.213800121	0.368438739	–0.155	0.009931067	0.017090674
cg13752114	3	3q39	195489708	*MUC4*	0.759478656	0.929965347	–0.17	0.034687695	0.036153372
cg17401179	5	5p15.33	400517	*AHRR*	0.48931918	0.630434783	–0.141	0.026114596	0.029730463
cg15909443	5	5p15.1	17381798		0.509454773	0.849598402	–0.34	0.018992279	0.023820825
cg18828365	5	5q14.1	78281819	*ARSB*	0.170947592	0.282753772	–0.112	0.004849646	0.01380284
cg24166916	5	5q14.1	78282669	*ARSB*	0.0597167	0.358376615	–0.299	0.003355676	0.013795556
cg12949141	5	5q31.1	134259545	*PCBD2*	0.51593311	0.634574028	–0.119	0.00026813	0.00396832
cg17386240	5	5q31.1	135384080	*TGFBI*	0.85529306	0.49288067	0.362	0.009931067	0.017090674
cg19650706	5	5q31.3	140594406	*PCDHB13*	0.761168306	0.506092538	0.255	0.007002286	0.015702097
cg23497306	5	5q31.3	140626788	*PCDHB15*	0.570650421	0.690388723	–0.12	0.003355676	0.013795556
cg24246628	6	6q27	168435914	*KIF25*	0.358390824	0.508558167	–0.15	0.013634908	0.021020483
cg08476511	6	6q27	168435923	*KIF25*	0.513822425	0.657107076	–0.143	0.009931067	0.017090674
cg17529386	6	6q27	170452270		0.620649223	0.756284153	–0.136	0.002267586	0.01198581 1
cg14957240	8	8p21.3	19457346	*CSGALNACT1*	0.812053688	0.674359389	0.138	0.009931067	0.017090674
cg14424376	9	9q22.32	98513318		0.468056065	0.114490562	0.354	0.018992279	0.023820825
cg11270017	10	10p15.3	1401818	*ADARB2*	0.403264663	0.298643017	0.105	0.004849646	0.01380284
cg17723206	10	10q25.3	118931160		0.889813907	0.499228571	0.391	0.034687695	0.036153372
cg21653586	11	11p15.4	10530636		0.312101683	0.537393447	–0.225	0.00152245	0.00938844
cg06588556	11	11p15.1	18477299	*LDHAL6A*	0.54937171	0.713258786	–0.164	0.013634908	0.021020483
cg23468927	11	11q13.1	67206263	*CORO1B*	0.258168344	0.111195504	0.147	0.026114596	0.029730463
cg12044599	11	11q13.1	67206308	*CORO1B*	0.340944971	0.154076644	0.187	0.018992279	0.023820825
cg15243034	11	11q14.1	77907656	*USP35*	0.604170589	0.735601597	–0.131	0.004849646	0.01380284
cg04757806	11	11q21	94278595	*FUT4*	0.826549837	0.932040093	–0.105	0.013634908	0.021020483
cg23764766	12	12p13.3	1459744	*ERC1*	0.610679896	0.930022241	–0.319	0.003355676	0.013795556

(Continued)

Table 2. (Continued)

Probe ID	Chr.	Cytoband	Position	Gene	Median (case)	Median (cont)	Δ_{meth}	*p* value	p_{FDR}
cg07374109	12	12q24.33	131645153		0.429269841	0.575413463	–0.146	0.026114596	0.029730463
cg16130238	14	14q11.2	22958402	*TCRA*	0.416198482	0.29958354	0.117	0.013634908	0.021020483
cg08198540	14	14q11.2	22963867	*TCRA*	0.339142349	0.14354067	0.196	0.004849646	0.01380284
cg20507742	14	14q11.2	22966633	*TCRA*	0.336249351	0.23480688	0.101	0.034687695	0.036153372
cg16302916	14	14q11.2	22969695	*TCRA*	0.319515195	0.215678024	0.104	0.007002286	0.015702097
cg24018633	14	14q11.2	22971114	*TCRA*	0.537136766	0.37737874	0.16	0.013634908	0.021020483
cg09953583	14	14q11.2	22974144	*TCRA*	0.351900961	0.246338806	0.106	0.009931067	0.017090674
cg02867735	14	14q11.2	22975384	*TCRA*	0.256100596	0.140564866	0.116	0.018992279	0.023820825
cg24065807	14	14q11.2	22978729	*TCRA*	0.289611525	0.181935929	0.108	0.004849646	0.01380284
cg20481655	14	14q11.2	22979162	*TCRA*	0.273572425	0.150166899	0.123	0.034687695	0.036153372
cg21543270	14	14q32.33	105840286	*PACS2*	0.607855904	0.80818158	–0.2	0.046067146	0.046067146
cg18350391	16	16p13.3	3115552	*IL32*	0.336448012	0.225957087	0.11	0.004849646	0.01380284
cg08978665	16	16p13.3	3115707	*IL32*	0.439208538	0.325142354	0.114	0.018992279	0.023820825
cg03666441	17	17p13.1	7254671	*KCTD11*	0.273557417	0.398402451	–0.125	0.007002286	0.015702097
cg05164926	17	17p13.1	7255624	*KCTD11*	0.396285015	0.5036796	–0.107	0.018992279	0.023820825
cg19001909	17	17q11.2	26205940	*C17orf108*	0.260233077	0.154529861	0.106	0.000412184	0.004357373
cg27060340	17	17q21.31	43502999	*ARHGAP27*	0.658361958	0.787145167	–0.129	0.004849646	0.01380284
cg17911788	17	17q21.31	44343683		0.296815999	0.451124504	–0.154	0.018992279	0.023820825
cg21028142	17	17q25.3	79581711	*NPLOC4*	0.602179481	0.483884753	0.118	0.046067146	0.046067146
cg10819238	19	19p13.3	1155184	*SBNO2*	0.83448977	0.938234103	–0.104	0.004849646	0.01380284
cg26988138	19	19p13.3	2639424	*GNG7*	0.230168544	0.360757621	–0.131	0.026114596	0.029730463
cg10296238	21	21q22.3	47605174	*C21orf56*	0.367803593	0.535242434	–0.167	0.026114596	0.029730463
cg15911859	22	22q13.31	45810043	*RIBC2*	0.231158482	0.379240438	–0.148	0.007002286	0.015702097

MS: multiple sclerosis; MHC: major histocompatibility complex; Chr.: chromosome; Δ_{meth}: differential methylation score; FDR: false discovery rate; ID: identification.

None of the six major MS genes outside of *HLA* identified in the meta-GWAS by Patsopoulos et al. (2011) i.e. *CD58, THADA, EOMES, MLANA, IL2RA, STAT3*[7] revealed any significant CpG island methylation changes as judged by the assay used for this study.

We also noted some hits to genes not previously linked to MS but which may have relevance to disease pathology. Two sites at chr. 5q14.1 showing decreased methylation in MS represent the *ARSB* gene, which is involved in oxidative stress and has been associated with hippocampal atrophy.[15] Also, two CpGs showing decreased methylation in MS map to the *KCTD11* gene on chr. 17p13.1, which plays a role in neuronal differentiation.[16]

Gene-Set Enrichment Analysis (GSEA)

The 74 CpG sites identified in the methylation analysis were located within 38 different genes across the genome, with a major cluster at the MHC region. We conducted exploratory analysis to assess the biological relevance of this complete gene set in terms of MS pathology. Specifically, we performed GSEA using the WebGestalt engine to identify putative pathways, diseases and cytobands associated with the 38-gene set. Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis revealed a significant alignment of the gene set to the "Antigen processing and presentation pathway" ($p = 0.001$). Alignment analysis of the gene set against disease association databases yielded statistically significant alignments to several autoimmune diseases including MS ($p = 0.005$). Cytoband analysis reinforced the alignment of the gene set to the HLA gene region on chr. 6p21 ($p = 1 \times 10^{-5}$), whilst also identifying 5q31 as being significantly enriched ($p < 0.05$). Interestingly, the 5q31 region contains an IL gene cluster and has been previously associated with MS in a large GWAS.[17]

Discussion

In this study we report results of a genome-wide association of methylation levels in CD4+ cells to MS susceptibility. The most noteworthy finding of this study was the striking differential methylation signal observed at CpG sites across the MHC complex, especially the hypomethylation of eight tightly clustered sites in *HLA-DRB1*. This is the first evidence of differential methylation at *HLA* being associated with MS, and the hypomethylation at *DRB1* may represent a major effect epigenetic locus for MS. At this stage the relevance of hypomethylation at this CpG island, which resides in the gene body of *DRB1*, can only be speculated but might involve aberrant over-expression of alternately spliced antigenic transcripts initiating a self-reactive immune response. Interestingly, Alcina et al. recently reported that the major MS risk variant *DRB1**1501 associates with high expression of *DRB1* in relation to disease, which is plausibly related to the hypomethylation results we have identified in *DRB1*.[18] In support of this notion, additional preliminary studies in our laboratory indicate a tendency for increased *DRB1* transcript level in a subset of these MS cases vs controls and that the hypomethylation at our eight key CpG sites in *DRB1* tends to correlate with an increase in transcript level at *DRB1* (results not shown).

We also provided evidence that DNA methylation at the CpG island in *DRB1* may be partially dependent on the *DRB1**1501 haplotype, which suggests that the effects of genetic variation may be exerted via epigenetic changes at this MS risk region. In support of this, Chao et al. examined 1792 families affected with MS and showed distorted transmission of *DRB1**1501 haplotypes from the maternal lineage in affected female offspring indicating parent-of-origin effects at the *HLA-DRB1* locus.[19] They postulated in 2011 that methylation in this region[20] might be responsible for the parent-of-origin effect in MS[21] and so for the increase in female/male ration seen over the last century in Canada.[22] Handel et al. has previously examined DNA methylation at CpGs across the *HLA-DRB1* and *HLA-DRB5* haplotype and found no association with MS.[23] However, there were some notable differences in their study design. First, they examined methylation in peripheral blood mononuclear cells (PBMCs) – or mixed blood cell samples. Second, they compared different MS severity sub-types rather than having an unaffected control group for comparison. By addressing these design features, our study had substantially improved power to detect methylation effects at this HLA locus. Ramagopalan et al. has examined methylation of the promoter of the MHC class II transactivator gene (*MHC2TA*) in PBMCs from MS twins and found no evidence of methylation differences, but they too could not rule out cell-specific methylation as playing a role in MS risk.[24]

We also identified differential methylation hits at genomic regions outside of HLA (and MHC). Many of these hits are at genes previously implicated in MS risk, which adds credibility to the findings. Most notably, nine CpGs all showing increased methylation in MS patients mapped to a 20 kb region within the *TCR*α on chr. 14q11.2. Antigen-specific T-cell receptors play a key role in immune recognition. Alleles of *TCR*α have been previously associated with MS.[12] Our results now suggest that CpG hypermethylation at this gene, perhaps having a knock-down effect on receptor expression, might also be involved in immune dysregulation in MS. In addition, we also identified loci at 5q14.1 and 17p13.1 that harbour novel candidate genes for involvement in MS pathology (*ARSB* and *KCTD11*, respectively).

One important consideration of our study is that the patients tested were being treated with a variety of medications at the time of recruitment. In particular, eight patients were being treated with fingolimod. This drug is a sphingosine-1-phosphate (S1P) receptor modulator whose primary effect is to prevent CD4+ lymphocytes exiting lymphoid tissue.[25] Whilst we attempted to control for the effects of

fingolimod in our analysis, future studies would benefit from testing either treatment-naïve patients or focusing on treatment-specific sub-groups.

Larger studies should also factor in the known environmental risk factors (vitamin D, EBV and smoking) as well as allow for more controlled testing of gender effects. Other clinical subtypes and severities for MS should also be examined. A major challenge for future studies will be to discern whether epigenetic variation associated with MS precedes disease onset or occurs during disease progression.

Using a T-cell specific approach, we have shown genome-wide differences in methylation levels between MS patients and controls. These novel findings provide convincing evidence for association of epigenetic variation at *HLA-DRB1* in relation to MS and highlight the potential importance of several non-HLA regions. Overall, this study provides a firm foundation for further research into this emerging area of MS genetics and demonstrates that larger-scale studies are now warranted both to validate these findings and to detect smaller effect differential methylation loci.

Methods

Subject recruitment

This was a case-control study involving genome-wide analysis of methylation levels in CD4+ T cells collected from 30 MS patients and 28 healthy controls. The project was conducted within the Hunter Medical Research Institute (HMRI). Each MS participant was invited to be a part of the study and written and verbal consent was provided. Ethics was sort and approved by the Hunter New England Ethics Committee before the commencement of the study. The MS patients were diagnosed by the McDonald criteria as having RRMS. The MS case group consisted of 26 females and four males ranging in age from 18 to 55 years. MS patients were being (or had been) treated with disease-modifying medications including interferons, glatiramer acetate, natalizumab and fingolimod. Only one female patient (who was newly diagnosed) was treatment naïve. The control group was composed of healthy (non-MS) volunteers recruited from the HMRI and the Red Cross Blood Bank (RCBB) blood donors between 2011 and 2012. The control group consisted of 15 females and 13 males ranging in age from 18 to 55 years.

Blood sample processing and DNA methylation analysis

Each MS participant donated ~50 ml peripheral blood and each control provided ~50 ml of buffy coat. All specimens were processed within 18 hours of collection. Plasma was removed after collection and PBMCs were isolated using standard Ficoll gradient procedures. Total CD4+ T cells were extracted from the PBMC population using EasySep negative magnetic separation according to the manufacturer's instructions (StemCell, Canada). Purity was assessed with a fluorochrome-conjugated anti-CD4 antibody (fluorescein isothiocyanate (FITC) anti-CD4, catalog# 10403) on the fluorescence-activated cell scanning (FACS) Diva (as per manufacturer's instructions). The CD4 purity was 90 ±5%. DNA was extracted using the QIAamp DNA micro kit (Qiagen, USA) and bisulphite converted using the MethylEasy Xceed kit according to the manufacturer's instructions. Converted DNA was then applied to the Illumina Infinium Human450K BeadChip arrays (service provided by the Australian Genome Research Facility). These arrays cover 96% of the CpG sites in the genome and measure methylation at an average coverage of 17 CpG sites per gene distributed across the promoter, 5'UTR, first exon, gene body and 3'UTR.

Data analysis

We designed an in-house data analysis pipeline that uses a combination of R/Bioconductor and custom scripts. Briefly, Illumina 450K raw intensity data (idat files) were parsed into the Bioconductor MINFI package.[26] Methylation data were background-corrected and control-normalised according to MINFI routines. Data were cleaned by removing (failed) CpG probes for which the intensity of both the methylated and unmethylated probes was <1000 units across all samples. We selected a threshold of 1000 units based on the profile of the available negative control probes. Y chromosome probes were also filtered out. All probe sequences were mapped to the human genome (build Hg19) using BOWTIE2[27] to identify potential hybridisation issues. In total, 33,457 CpG probes were identified to align multiple times to the human genome and were filtered out of subsequent analysis.

Measures of methylation level (β values) were produced for each probe and ranged from 0 (completely unmethylated) to 1 (completely methylated). To identify differentially methylated CpG sites, we calculated the difference in median β values by subtracting $\text{median}_{\text{cont}}$ from $\text{median}_{\text{case}}$ to produce a Δ_{meth} score – a measure of differential methylation ranging from −1 to 1. The Δ_{meth} score can be broadly interpreted as percentage up or down methylation in cases compared to controls (or effect size). The two-sample Kolmogorov-Smirnov test (K-S test) was used to test whether Δ_{meth} was statistically significant as implemented in *R*.[28] The K-S was chosen (over the *F* test comparison of means) because of the marked variation in the distribution of the β values among probes. We conducted a stepwise filtration process to select the core CpG probe set most likely to be associated with MS (see supplementary information for details).

The relative importance of the core probes was assessed based on ranking of *p* values adjusted for a false discovery rate (FDR) at the 0.05 level (p_{FDR}) the Benjamini-Hochberg

procedure.[29] Given the relatively modest sample size, this study is limited to the detection of major effect differential methylation loci only. Gene set enrichment analysis was performed using the WebGstalt webserver (http://bioinfo.vanderbilt.edu/webgestalt, version and date). The probes in the Illumina manifest file contained gene annotation used for GSEA.

Conflict of interest

Dr Lechner-Scott's institution receives non directed funding as well as honoraria for presentations and membership on advisory boards from Sanofi Aventis, Biogen Idec, Bayer Health Care, Merck Serono and Novartis Australia.

Funding

The study was supported by the John Hunter Charitable Trust. Rod Lea is funded by a fellowship from the Multiple Sclerosis Research Australia.

References

1. Jersild C, Svejgaard A and Fog T. HL-A antigens and multiple sclerosis. *Lancet* 1972; 3: 1240–1241.
2. Friese MA and Fugger L. Pathogenic CD8(+) T cells in multiple sclerosis. *Ann Neurol* 2009; 66: 132–141.
3. Chitnis T. The role of CD4 T cells in the pathogenesis of multiple sclerosis. *Int Rev Neurobiol* 2007; 79: 43–72.
4. Koch MW, Metz LM and Kovalchuk O. Epigenetics and miRNAs in the diagnosis and treatment of multiple sclerosis. *Nat Rev Neurol* 2013; 9: 35–43.
5. Burrell AM, Handel AE, Ramagopalan SV, et al. Epigenetic mechanisms in multiple sclerosis and the major histocompatibility complex (MHC). *Discov Med* 2011; 11: 187–196.
6. Watson CT, Disanto G, Breden F, et al. Estimating the proportion of variation in susceptibility to multiple sclerosis captured by common SNPs. *Sci Rep* 2012; 2: 770.
7. Patsopoulos NA, Bayer Pharma MS Genetics Working Group, Steering Committees of Studies Evaluating IFNβ-1b and a CCR1-Antagonist, et al. Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. *Ann Neurol* 2011; 70: 897–912.
8. Jones PA and Laird PW. Cancer epigenetics comes of age. *Nat Genet* 1999; 21: 163–167.
9. Huynh JL and Casaccia P. Epigenetic mechanisms in multiple sclerosis: Implications for pathogenesis and treatment. *Lancet Neurol* 2013; 12: 195–206.
10. Reinius LE, Acevedo N, Joerink M, et al. Differential DNA methylation in purified human blood cells: Implications for cell lineage and studies on disease susceptibility. *PLoS One* 2012; 7: e41361.
11. Baranzini SE, Mudge J, van Velkinburgh JC, et al. Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. *Nature* 2010; 464: 1351–1356.
12. Oksenberg JR, Sherritt M, Begovich AB, et al. T-cell receptor V alpha and C alpha alleles associated with multiple and myasthenia gravis. *Proc Natl Acad Sci U S A* 1989; 86: 988–992.
13. Kumagai C, Kalman B, Middleton FA, et al. Increased promoter methylation of the immune regulatory gene SHP-1 in leukocytes of multiple sclerosis subjects. *J Neuroimmunol* 2012; 246: 51–57.
14. Vandenbroeck K, Alvarez J, Swaminathan B, et al. A cytokine gene screen uncovers *SOCS1* as genetic risk factor for multiple sclerosis. *Genes Immun* 2012; 13: 21–28.
15. Potkin SG, Guffanti G, Lakatos A, et al. Hippocampal atrophy as a quantitative trait in a genome-wide association study identifying novel susceptibility genes for Alzheimer's disease. *PLoS One* 2009; 4: e6501.
16. Di Marcotullio L, Ferretti E, De Smaele E, et al. REN(KCTD11) is a suppressor of Hedgehog signaling and is deleted in human medulloblastoma. *Proc Natl Acad Sci U S A* 2004; 101: 10833–10838.
17. Sawcer S, Hellenthal G, Pirinen M, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* 2011; 476: 214–219.
18. Alcina A, Abad-Grau Mdel M, Fedetz M, et al. Multiple sclerosis risk variant *HLA-DRB1**1501 associates with high expression of *DRB1* gene in different human populations. *PLoS One* 2012; 7: e29819.
19. Chao MJ, Bernardo MC, Lincoln MR, et al. HLA class I alleles tag *HLA-DRB1**1501 haplotypes for differential risk in multiple sclerosis susceptibility. *Proc Natl Acad Sci U S A* 2008; 105: 13069–13074.
20. Chao MJ, Ramagopalan SV, Herrera BM, et al. MHC transmission: Insights into gender bias in MS susceptibility. *Neurology* 2011; 76: 242–246.
21. Chao MJ, Herrera BM, Ramagopalan SV, et al. Parent-of-origin effects at the major histocompatibility complex in multiple sclerosis. *Hum Mol Genet* 2010; 19: 3679–3689.
22. Orton SM, Herrera BM, Yee IM, et al. Sex ratio of multiple sclerosis in Canada: A longitudinal study. *Lancet Neurol* 2006; 5: 932–936.
23. Handel AE, De Luca GC, Morahan J, et al. No evidence for an effect of DNA methylation on multiple sclerosis severity at *HLA-DRB1**15 or *HLA-DRB5*. *J Neuroimmunol* 2010; 223: 120–123.
24. Ramagopalan SV, Dyment DA, Morrison KM, et al. Methylation of class II transactivator gene promoter IV is not associated with susceptibility to multiple sclerosis. *BMC Med Genet* 2008; 9: 63.
25. Killestein J, Rudick RA and Polman CH. Oral treatment for multiple sclerosis. *Lancet Neurol* 2011; 10: 1026–1034.
26. Hansen KD and Aryee M. minfi: Analyze Illumina's 450k methylation arrays. *Bioconductor* 2012.
27. Liu Y and Schmidt B. Long read alignment based on maximal exact match seeds. *Bioinformatics* 2012; 28: i318–i324.
28. R-Development-Core-Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (2010); (http://www.r-project.org/ date accessed: 2/3/2012).
29. Benjamini Y and Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Ser B* 1995; 57: 289–300.