



OPEN

## Epigenetic differences at the *HTR2A* locus in progressive multiple sclerosis patients

Vicki E. Maltby<sup>1,2,14</sup>, Rodney A. Lea<sup>2,3,14</sup>, Sean Burnard<sup>1,2,4</sup>, Alexandre Xavier<sup>1,2,4</sup>, Thao Van Cao<sup>3</sup>, Nicole White<sup>3</sup>, Daniel Kennedy<sup>3</sup>, Kira Groen<sup>1,2</sup>, Katherine A. Sanders<sup>2,5</sup>, Rebecca Seeto<sup>1,6</sup>, Samara Bray<sup>1,2</sup>, Melissa Gresle<sup>7,8,10</sup>, Louise Laverick<sup>7,8</sup>, Helmut Butzkueven<sup>9,10</sup>, Rodney J. Scott<sup>4,11,12</sup> & Jeannette Lechner-Scott<sup>1,2,13</sup>✉

The pathology of progressive multiple sclerosis (MS) is poorly understood. We have previously assessed DNA methylation in the CD4<sup>+</sup> T cells of relapsing–remitting (RR) MS patients compared to healthy controls and identified differentially methylated regions (DMRs) in *HLA-DRB1* and *RNF39*. This study aimed to investigate the DNA methylation profiles of the CD4<sup>+</sup> T cells of progressive MS patients. DNA methylation was measured in two separate case/control cohorts using the Illumina 450K/EPIC arrays and data was analysed with the Chip Analysis Methylation Pipeline (ChAMP). Single nucleotide polymorphisms (SNPs) were assessed using the Illumina Human OmniExpress24 arrays and analysed using PLINK. Expression was assessed using the Illumina HT12 array and analysed in R using a combination of Limma and Illuminaio. We identified three DMRs at *HTR2A*, *SLC17A9* and *HDAC4* that were consistent across both cohorts. The DMR at *HTR2A* is located within the bounds of a haplotype block; however, the DMR remained significant after accounting for SNPs in the region. No expression changes were detected in any DMRs. *HTR2A* is differentially methylated in progressive MS independent of genotype. This differential methylation is not evident in RRMS, making it a potential biomarker of progressive disease.

Multiple Sclerosis (MS) is an inflammatory and neurodegenerative disease of the central nervous system (CNS). The majority of patients are initially diagnosed with relapsing–remitting MS (RRMS), which is thought to involve migration of self-reactive immune cells into the CNS<sup>1</sup>. These cells initiate an immune response, which results in localized demyelination and corresponding neurological symptoms<sup>2</sup>. Over a period of time, patients may develop progressive MS, which is characterized by compounding neurodegeneration and a decreased ability to repair damage, resulting in increasing disability. This phase can be active with superimposed relapses or non-active without relapses<sup>1,3</sup>. The role of inflammation in progression is less clear and it is presumed that both immunological and neurobiological mechanisms are responsible for the accumulating damage<sup>1</sup>.

The underlying pathogenesis of MS remains unclear but the prevailing hypothesis is that MS develops due to a combination of genetic predisposition and environmental factors. Epigenetics can influence the genome without changing the DNA sequence. Environmental exposures, such as smoking, have been demonstrated to influence epigenetic patterns<sup>4,5</sup>. One such epigenetic change is DNA methylation, which refers to the addition of a methyl group to a CpG dinucleotide. We, and others, have previously assessed the genome-wide DNA methylation profile

<sup>1</sup>School of Medicine and Public Health, University of Newcastle, Callaghan, NSW 2308, Australia. <sup>2</sup>Centre for Brain and Mental Health, Hunter Medical Research Institute, New Lambton Heights, NSW 2305, Australia. <sup>3</sup>Institute of Health and Biomedical Innovations, Genomics Research Centre, Queensland University of Technology, Kelvin Grove, QLD 4059, Australia. <sup>4</sup>School of Biomedical Sciences and Pharmacy, University of Newcastle, Callaghan, NSW 2308, Australia. <sup>5</sup>Centre for Anatomical and Human Sciences, Hull York Medical School, Hull, UK. <sup>6</sup>School of Environmental and Life Sciences, University of Newcastle, Callaghan, NSW 2308, Australia. <sup>7</sup>Department of Medicine, University of Melbourne, Melbourne, VIC, Australia. <sup>8</sup>Royal Melbourne Hospital, Melbourne, VIC, Australia. <sup>9</sup>Alfred Hospital, Melbourne, VIC, Australia. <sup>10</sup>MS and Neuroimmunology Unit, Central Clinical School, Monash University, Melbourne, VIC, Australia. <sup>11</sup>Division of Molecular Genetics, Pathology North, John Hunter Hospital, New Lambton Heights, NSW 2305, Australia. <sup>12</sup>Centre for Cancer Research, Hunter Medical Research Institute, New Lambton Heights, NSW 2305, Australia. <sup>13</sup>Department of Neurology, John Hunter Hospital, New Lambton Heights, NSW 2305, Australia. <sup>14</sup>These authors contributed equally: Vicki E. Maltby and Rodney A. Lea. ✉email: Jeannette.lechner-scott@hnehealth.nsw.gov.au

	HC discovery cohort	MS discovery cohort	HC replication cohort	MS replication cohort
Methylation arrays (N)	16	23	12	12
SNP arrays (N)	13	22	0	0
Expression arrays (N)	16	22	0	0
Female	16	23	12	12
Age in years (mean $\pm$ SD)	58 $\pm$ 9.5	57.9 $\pm$ 9.5	61.0 $\pm$ 8.6	60.2 $\pm$ 9.0
EDSS (mean $\pm$ SD)	N/A	6.5 $\pm$ 1.3	N/A	6.1 $\pm$ 1.1
Disease duration in years (mean $\pm$ SD)	N/A	24.5 $\pm$ 10.3	N/A	26.5 $\pm$ 12.5
Progression duration in years (mean $\pm$ SD)	N/A	8.8 $\pm$ 5.6	N/A	5.9 $\pm$ 4.1

**Table 1.** Patient characteristics. *SD* standard deviation, *HC* healthy controls, *EDSS* expanded disability status score, *N/A* not applicable.

in the CD4<sup>+</sup> and CD8<sup>+</sup> T cells and CD19<sup>+</sup> B cells of RRMS patients and compared to that of healthy controls<sup>6–12</sup>. In CD4<sup>+</sup> T cells, we found a striking methylation signal located at the MHC locus with peaks at *HLA-DRB1* and *RNF39* present in RRMS patients compared to healthy controls<sup>8,11</sup>. We did not identify either of these DMRs in either CD8<sup>+</sup> T cells or CD19<sup>+</sup> B cells of MS patients<sup>6,7</sup>. A more recent study investigated the interplay between DNA methylation and genetic background in the *DRB1* region and found that MS patients who were carriers of the high risk *HLA-DRB1 15\*01* allele had significantly lower DNA methylation at *HLA-DRB1* compared to those who were not carriers. This was the case in monocytes, B cells and both CD4<sup>+</sup> and CD8<sup>+</sup> T cells. They further demonstrated a function relationship between genetic background, lowered methylation and increased expression in this region. Thus, the results seen in CD4<sup>+</sup> T cells may have been due to genetic background<sup>13</sup>.

Our understanding of RRMS has improved markedly. This has translated into new treatment options that target T and B lymphocytes that are also effective in reducing disability progression (reviewed in<sup>14</sup>). Unfortunately, our knowledge of factors that drive MS progression and resulting disability is lacking. A better understanding of disease progression is needed to track disease activity and progression accurately for patients. This is important because disease progression and accompanying disability may not be targeted by current therapies.

Identifying epigenetic loci associated with MS, particularly progressive MS, could reveal potentially important biomarkers of disease course as well as modifiable targets for new drug design.

## Methods

**Ethics approval and consent to participate.** The Hunter New England Health Research Ethics Committee and the University of Newcastle Ethics Committees approved this study (2019/ETH12346 and H-505-0607 respectively), and methods were carried out in accordance with institutional guidelines on human subject experiments. Written and informed consent was obtained from all patient and control subjects.

**Sample collection.** Whole blood was collected from a discovery cohort of 23 female secondary progressive MS patients (hereafter referred to as progressive MS) and 16 age- and gender-matched healthy controls and an independent validation cohort of 11 female progressive MS patients and 12 age- and gender-matched healthy controls (Table 1). All patients were diagnosed with MS according to the 2017 McDonald criteria<sup>15</sup>. Progressive disease stage was evaluated by the Neurology Staff Specialists, who are EDSS certified, at the John Hunter Hospital or Royal Melbourne Hospital. As there is currently no “gold standard” for defining progression, we have used the most recent recommendations for progression set by the MSBase Study Group<sup>16</sup>. To meet these criteria progressive patients must: (i) have a minimum expanded disability status scale (EDSS) score of 4, (ii) have had disability progression by 1 EDSS step for patients with EDSS  $\leq$  5.5 or 0.5 for patients with EDSS  $\geq$  6 in the last year, (iii) have progression in the absence of relapse, (iv) have confirmed progression over  $\geq$  3 months<sup>16</sup>. All patients had not had MS-specific therapy or steroid treatment for a minimum period of 6 months prior to collection.

**Sample preparation.** Peripheral blood mononuclear cells (PBMCs) were isolated from whole blood by density-gradient centrifugation on Lymphoprep (StemCell Technologies, Canada). CD4<sup>+</sup> T cells were isolated using negative selection magnetic separation kits (StemCell Technologies, Canada) according to the manufacturer's protocols. Purity was assessed using a FITC-conjugated anti-CD4 antibody (anti-human CD4 antibody, clone OTK4, catalog #60016F1, StemCell Technologies, Canada) on a FACSCanto II flow cytometer (BD Biosciences, USA) and analyzed using the FACSDiva software (BD Biosciences, USA). All samples met a minimum purity threshold of  $\geq$  90%.

Extraction of genomic DNA was performed using the QiaAMP DNA micro kit (Qiagen, USA). Processing of the discovery cohort DNA for the Infinium HumanMethylation450 arrays including bisulphite conversion, was done at the Centre for Clinical Genomics (University of Queensland). Processing of the validation cohort samples for Infinium MethylationEPIC arrays was done at the Australian Genome Research Facility (AGRF). The AGRF also performed the Illumina HumanOmniExpress-24 arrays, on the samples with sufficient DNA remaining. Exact numbers of individuals used for each assay are listed in Table 1.

RNA was extracted from samples with sufficient PBMCs using the QiaAMP RNA/miRNA kit (Qiagen, Netherlands). Quality control was assessed using the Qubit fluorometer (quantification) and the Agilent bioanalyzer, using the small and nano RNA kits. Processing of the RNA for Illumina Human HT12 arrays (hereafter referred to as HT12 arrays) was performed by the Centre for Clinical Genomics (University of Queensland).

**DNA methylation analysis.** Raw methylation data was analysed primarily using the ChAMP package in R, with the support of several auxiliary scripts (version 2.16.2)<sup>17</sup>. For quality control (QC) assessments, raw methylation data were used to produce beta density plots, multiple dimensional scaling (MDS) plots and sex check plots, each containing assigned phenotype labels (Suppl. Files 1a). To adjust for background and dye bias from different arrays (450 K and EPIC), the raw data was normalized using the single-sample Noob method (ssNoob), which was applied using the “preprocessNoob” function in the Minfi package (version 1.32.0). The ssNoob script was designed to perform background correction for Illumina Methylation data generated for different arrays (ie. 450K and EPIC)<sup>18</sup>. Following normalisation, the QC metrics were re-checked to confirm an improvement in the beta distributions (Suppl. Files 1b). The next step involved removal of bad probes. This included probes with a detection P-value > 0.01, beadcount < 3 in at least 5% of samples, NoCG, SNPs, MultiHit and probes on XY chromosomes. For this step the Minfi and WateRmelon packages (version 1.30.0) were required to generate the necessary data derived from the ssNoob analysis to run the “champ.filter” function<sup>19</sup>. Batch effect analysis was then performed for both cohorts combined using the “champ.SVD” function. The harmonised data set was then adjusted for significant batch effects using the “champ.runCombat” function (Suppl. Files 2).

**Differentially methylated positions and regions.** The final harmonised beta value matrix was then moved forward to analysis of differentially methylated positions (DMPs) and differentially methylated regions (DMRs). We implemented Champ.DMP and Champ.DMR(DMRCate) to identify DMPs and DMRs respectively. The default false discovery rate of 0.05 was used. Given the modest sample size we also chose to apply selection criteria based on both unadjusted statistical significance and effect size. For this, DMPs were selected if a CpG yielded both an unadjusted P-value less than 0.05 and a  $\Delta_{\text{meth}} \geq 0.1$  (representing 10% differential methylation). For DMRs we focused only on regions which had  $\geq 2$  DMPs that mapped to the same gene, were separated by < 1500 bp distance, and where adjacent CpGs yielded  $\Delta_{\text{meth}}$  in a consistent direction. This approach has been successfully applied in previous studies to identify DMPs in small samples<sup>6–8,11</sup>. All data was processed and analyzed in Rstudio (version 1.1.456)<sup>20</sup> and R (version 3.5.2)<sup>21</sup>.

**Single nucleotide polymorphism (SNP) analysis.** Genomic DNA was extracted from whole blood and genotyped using the Illumina HumanOmniExpress-24 array. PLINK (v1.9)<sup>22</sup> was used for quality control procedures and SNP frequency statistics. Samples were excluded for (1) sex inconsistencies or (2) call rate < 95%. Individual SNPs were removed if (1) the genotype call rate was below 98%, (2) the minor-allele frequency was less than 0.05 and (3) there was significant deviation from the Hardy–Weinberg permutation test ( $P < 1 \times 10^{-5}$ ). SNPs were extracted at least  $\pm 5$  Kb upstream and downstream of HTR2A, from 47,400,677 bp to 47,477,087 bp (hg19). The progressive MS cohort included 22 cases and 13 controls and the RRMS cohort included 24 cases and 17 controls. A total of 48 variants passed QC and filtering. Fisher’s Exact testing was performed to assess SNP frequencies. Haploview (v.4.2) was used to assess haplotypes<sup>23</sup>.

**Methylation quantitative trait loci (QTL) analysis.** Genetic changes can mediate the association between DNA methylation and disease via mQTLs. To explore the possibility of this type of association for key DMRs, we compared mean methylation of CpGs among genotype groups for index SNPs using ANOVA. To assess the conjoint effects of CpG methylation and SNP genotype on MS outcome we used binary logistic regression models.

**Transcriptome analysis.** RNA was extracted for gene expression analysis and was measured using Illumina Human HT12 v4 arrays. The programme R (v.3.6.3)<sup>21</sup> was used with the R packages limma<sup>24</sup> and illuminaio<sup>25</sup> and following the Kebschull et al.<sup>26</sup> pipeline to pre-process, quality control and normalise the idat files. After normalisation, the HTR2A probe value for all samples was extracted and compared.

## Results

Following the QC analysis, the resultant methylation datasets were comprised of 407,400 CpGs for the discovery cohort and 743,477 CpGs for the replication cohort. Of these 378,188 CpGs were common to both cohorts. Batch effect analysis conducted using the SVD function in ChaMP indicated significant methylation clustering by “slide” and “cohort” on methylation profiles ( $P < 1 \times 10^{-5}$ ). These factors were diminished to negligible effects after adjustment using the ComBat normalisation function (Suppl. Files 2).

**Differentially methylated positions (DMPs).** We analysed methylation data from the CD4<sup>+</sup> T cells obtained from our discovery cohort of 23 female progressive MS patients and 16 female healthy controls. ChampDMPs did not reveal any DMPs that were statistically significant at an FDR of 0.05. This is not surprising given the modest sample size used here. For this reason, we also employed a selection algorithm based on an unadjusted P-value < 0.05 and  $\Delta_{\text{meth}}$ . These steps resulted in 72 DMPs that mapped to 34 genes and 23 unannotated genomic locations (Table S1). Of these, 74% were hypermethylated and 26% were hypomethylated with majority located within the gene body or intragenic regions.

CpG	CHR	MAPINFO	Mean control	Mean MS	$\Delta_{\text{meth}}$	P-value	Gene	CpGs in DMR	Feature	cgi
cg02267270	6	37,616,410	0.42	0.53	0.11	0.01	MDGA1	4	Body	Island
cg14926196	6	37,616,482	0.33	0.51	0.18	0.00	MDGA1	4	Body	Island
cg20053110	6	37,617,864	0.53	0.64	0.11	0.04	MDGA1	4	Body	Island
cg24796644	6	37,617,956	0.57	0.70	0.13	0.01	MDGA1	4	Body	Island
cg23299919	7	157,406,096	0.19	0.30	0.11	0.01	PTPRN2	3	Body	Island
cg06715136	7	158,046,025	0.73	0.83	0.11	0.02	PTPRN2	3	Body	Opensea
cg09756125	7	158,250,978	0.66	0.53	-0.14	0.02	PTPRN2	3	Body	Shore
*cg18200810	13	47,472,200	0.60	0.48	-0.11	0.02	HTR2A	3	TSS1500	Opensea
*cg24118521	13	47,472,330	0.74	0.58	-0.16	0.02	HTR2A	3	TSS1500	Opensea
*cg23881368	13	47,472,343	0.68	0.55	-0.13	0.01	HTR2A	3	TSS1500	Opensea
cg07075026	17	47,091,521	0.48	0.38	-0.10	0.01	IGF2BP1	3	Body	Island
cg10950924	17	47,092,072	0.56	0.46	-0.10	0.00	IGF2BP1	3	Body	Island
cg18128536	17	47,092,178	0.59	0.48	-0.11	0.00	IGF2BP1	3	Body	Island
cg12099423	20	61,590,751	0.26	0.36	0.11	0.00	SLC17A9	3	Body	Shore
*cg17221813	20	61,590,823	0.38	0.51	0.13	0.00	SLC17A9	3	Body	Island
*cg19142181	20	61,591,066	0.37	0.54	0.17	0.00	SLC17A9	3	Body	Island
*cg07673080	2	240,241,154	0.89	0.76	-0.12	0.00	HDAC4	2	Body	Opensea
*cg10639368	2	240,241,218	0.80	0.70	-0.11	0.00	HDAC4	2	Body	Opensea
cg18850127	7	39,170,497	0.65	0.51	-0.14	0.01	POU6F2	2	Body	Opensea
cg20302533	7	39,170,763	0.56	0.42	-0.14	0.00	POU6F2	2	Body	Opensea
cg06238316	19	21,265,364	0.22	0.37	0.15	0.00	ZNF714	2	5'UTR	Island
cg09352518	19	21,265,421	0.24	0.44	0.20	0.00	ZNF714	2	5'UTR	Island
cg17099656	22	47,135,171	0.17	0.30	0.13	0.00	CERK	2	TSS1500	Opensea
cg16154810	22	47,135,258	0.35	0.55	0.20	0.00	CERK	2	TSS1500	Opensea

**Table 2.** DMRs in discovery cohort.

We separately analysed the validation cohort of 11 progressive female MS patients and 12 healthy female controls and identified 65 DMPs which mapped to 32 genes and 22 unannotated genomic locations (Table S2). Again, there was a tendency towards hypermethylation, with 68% of DMPs being hypermethylated and the majority of DMPs located within the gene body or intragenic region.

To investigate the level of similarity between the two cohorts, we compared the lists of DMPs identified in both MS cohorts. This revealed 12 DMPs common to both cohorts, which spanned three genes (*HTR2A*, *SLC17A9* and *HDAC4*) and four unannotated regions on chromosomes 12, 16, 5, and 3 (Tables S1, S2 asterisk). All DMPs were altered in a consistent direction (five hypomethylated and seven hypermethylated).

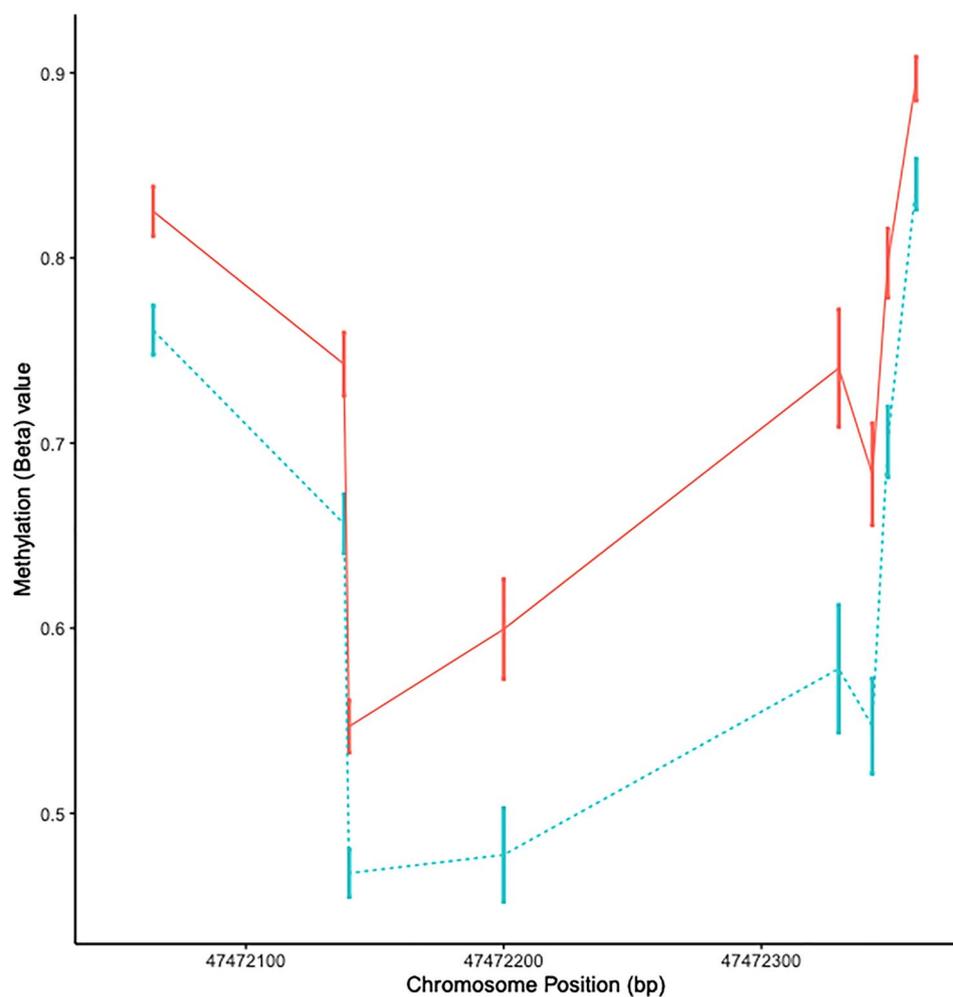
**Differentially methylated regions (DMRs).** Results of DMRCate did not reveal any significant DMRs at an FDR of 0.05. We therefore focused only on regions which had  $\geq 2$  DMPs located within a 1500 bp distance, where adjacent CpGs yielded  $\Delta_{\text{meth}}$  in a consistent direction. In the discovery cohort, this identified 9 DMRs (Table 2). The largest DMR was found at *MDGA1*, which contained four sites located in the gene body within 1546 base pairs. These sites are contained within a CpG island and have a  $\Delta_{\text{meth}}$  ranging from 0.11 to 0.18. In the validation cohort, the largest DMR was located in the *HTR2A* gene, within the transcription start site. This DMR contains 5 DMPs located within 211 base pairs with a  $\Delta_{\text{meth}}$  ranging from -0.10 to 0.17 (Table 3). On comparison between the two cohorts, three DMRs (located in *HTR2A*, *HDAC4* and *SLC17A9*) were replicated and had  $\Delta_{\text{meth}}$  in a consistent direction (Table 2, 3 asterisk).

DNA methylation is most likely to exert an effect on expression within the transcriptional start site (TSS). The only DMR identified both in the TSS and in both cohorts was the DMR at *HTR2A*. Therefore, to gain a more comprehensive profile of the *HTR2A* DMR we joined the two cohorts together and re-analysed all CpGs in *HTR2A* for differential methylation between cases and controls. In this analysis there were eight CpGs exhibiting > 5% hypomethylation in the cases compared to controls ( $P < 0.005$ ) (Table S3). These 8 DMPs are located within a 283 base pair region in the TSS1500 of the gene (Fig. 1). The strongest association was observed for cg09798090 ( $\Delta_{\text{meth}} = -0.08$ ,  $P = 0.00009$ ). There was a strong positive correlation among all 8 CpGs in this cohort supporting this as a single DMR ( $r \geq 0.85$ ,  $P < 0.0001$ ) (Suppl. Fig. S3). The probe cg09798090 was selected as the index (or tagging) CpG for subsequent analyses.

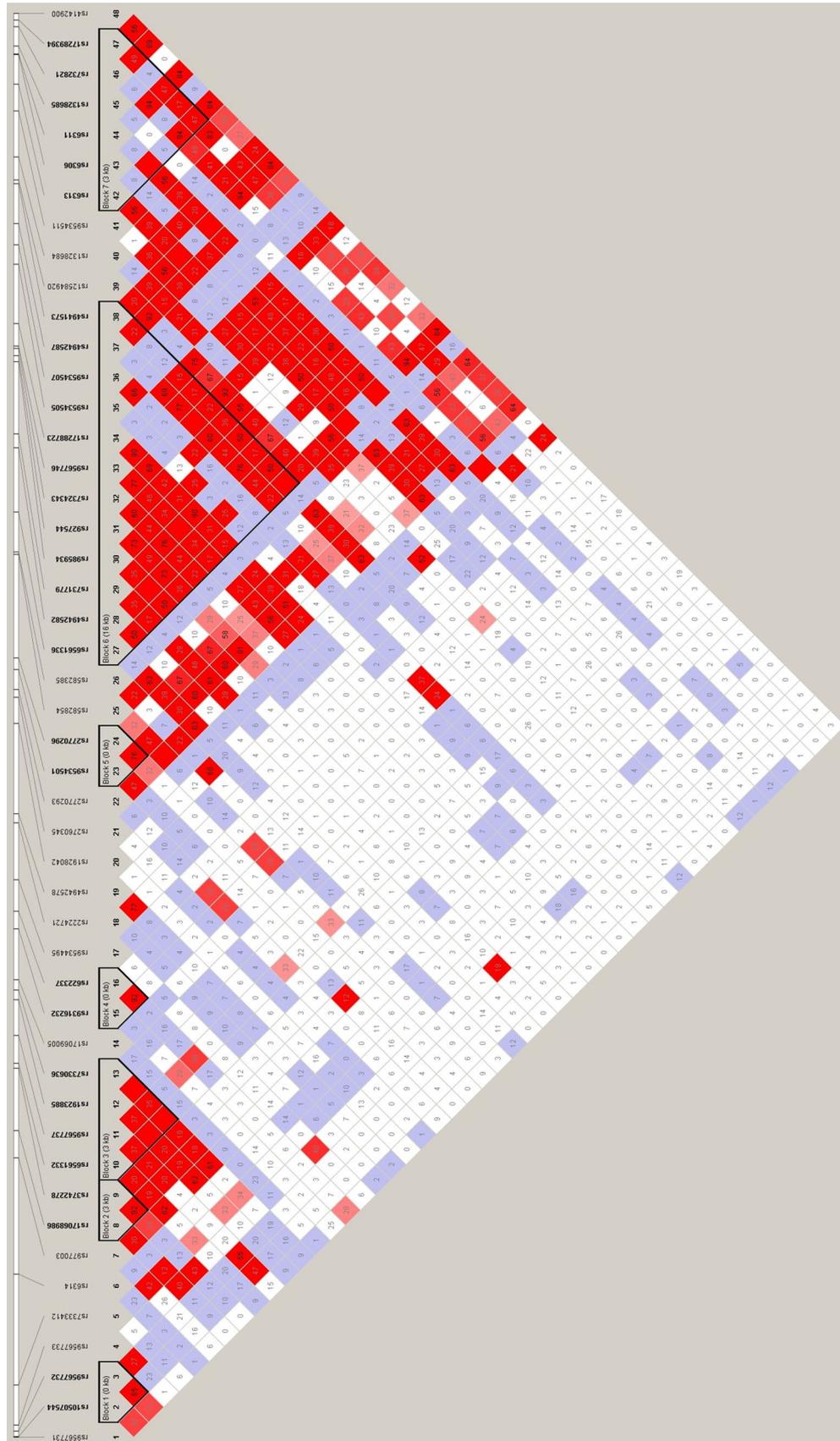
**The *HTR2A* DMR is contained within the bounds of a haplotype block.** A previous study, focused on schizophrenia, demonstrated that both gene expression and methylation at *HTR2A* are likely under genetic control<sup>27</sup>. To investigate if this is the case in our cohort, we performed a haplotype block analysis using Haploview. The HumanOmniExpress-24 arrays contain 50 SNPs for the *HTR2A* gene. We identified 7 haplotype blocks in this cohort (Fig. 2). The DMR is located within block 7 between rs1328685 and rs732821 and is tagged by

CpG	CHR	MAPINFO	Mean control	Mean MS	deltaBeta	P.Value	gene	DMR	feature	cgi
cg06020661	13	47,472,138	0.75	0.65	-0.10	0.01	HTR2A	5	TSS1500	Opensea
*cg18200810	13	47,472,200	0.60	0.47	-0.14	0.03	HTR2A	5	TSS1500	Opensea
*cg24118521	13	47,472,330	0.74	0.57	-0.17	0.03	HTR2A	5	TSS1500	Opensea
*cg23881368	13	47,472,343	0.69	0.54	-0.15	0.02	HTR2A	5	TSS1500	Opensea
cg05506829	13	47,472,349	0.80	0.69	-0.11	0.02	HTR2A	5	TSS1500	Opensea
cg26371957	12	739,280	0.67	0.80	0.13	0.03	NINJ2	4	Body	Opensea
cg14911689	12	739,980	0.28	0.50	0.22	0.02	NINJ2	4	Body	Opensea
cg26654770	12	740,100	0.28	0.55	0.27	0.02	NINJ2	4	Body	Opensea
cg01201512	12	740,338	0.37	0.60	0.23	0.04	NINJ2	4	Body	Opensea
*cg07673080	2	2.4E+08	0.88	0.76	-0.11	0.01	HDAC4	2	Body	Opensea
*cg10639368	2	2.4E+08	0.80	0.69	-0.10	0.02	HDAC4	2	Body	Opensea
cg11938672	6	1.7E+08	0.71	0.59	-0.13	0.01	WDR27	2	Body	Opensea
cg18322025	6	1.7E+08	0.63	0.47	-0.16	0.01	WDR27	2	Body	Shelf
cg18020065	13	1.15E+08	0.34	0.47	0.13	0.01	RASA3	2	Body	Opensea
cg16820615	13	1.15E+08	0.73	0.62	-0.11	0.00	RASA3	2	Body	Opensea
*cg17221813	20	61,590,823	0.39	0.50	0.11	0.03	SLC17A9	2	Body	Island
*cg19142181	20	61,591,066	0.39	0.54	0.15	0.05	SLC17A9	2	Body	Island

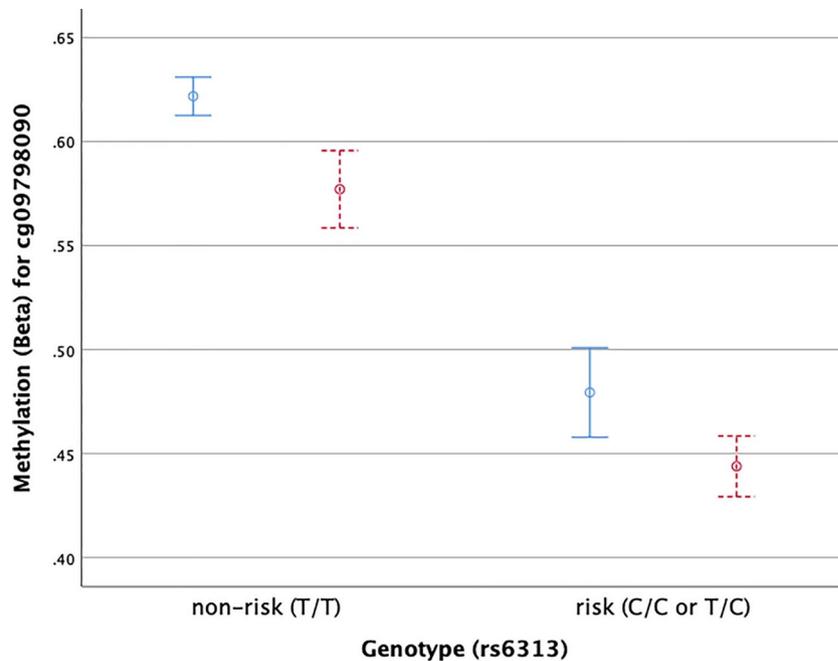
**Table 3.** DMRs in validation cohort.



**Figure 1.** HTR2A DMR plot. The mean beta value at the eight *HTR2A* CpG sites for controls (solid red line) and SPMS patients (dotted blue line) located within the transcription start site. Error bars show standard error of the mean.



**Figure 2.** Haplotype block at HTR2A. Haplotype chart showing the 7 haplotype blocks at HTR2A. Blocks are indicated by a bold black outline. SNPs contained within the block are in bold and the length of the block is indicated (in base pairs) in parentheses. This plot provides information on both the  $r^2$  and  $D'$  values for better clarification of any linkage disequilibrium. The number within each diamond represents the  $r^2$  value between each diagonally corresponding SNP pairs and the colour of each diamond represents the strength/value of  $D'$  (white to bright red/0 < 100).



**Figure 3.** HTR2A mQTL. Mean delta beta values at cp09798090 for SNP rs6313 in either the non-risk allele (T/T) or risk allele (C/C or T/C). Healthy controls (HC) are shown in blue, SPMS are shown in red. Error bars are standard error of the mean.

rs6313, previously linked to a number of other inflammatory and neurological diseases<sup>28,29</sup>. Even in this modestly sized sample there was evidence of association between rs6313 and progressive MS. In particular, patients were more likely to carry the risk allele (C) compared to controls (OR 5.3,  $P = 0.028$ ).

To determine if this DMR also resulted in expression level changes, we examined the *HTR2A* probe of the HT12 arrays between HC and MS patients. The HT12 arrays only have one probe for *HTR2A*, which is lowly expressed in both HC and MS patients (according to manufacturer's manifest/product information). There were no identified changes between HC and progressive MS patients.

**Methylation quantitative trait loci (mQTL) analysis.** The effect of genetic variants on disease phenotypes can be mediated by DNA methylation via mQTLs. To assess this in our cohort we compared the average methylation of the index CpG (cg09798090) across the genotype groups for the index SNP (rs6313). Figure 3 shows a highly significant difference in methylation levels between risk allele carriers and non-carriers ( $P < 0.0001$ ). Figure 3 also indicates that the hypomethylation in progressive patients may occur independently of genotype. To explore this, we performed a logistic regression with phenotype as the outcome and SNP (rs6313) and CpG (cg09798090) as joint predictors. Results of this model showed that the DMR (as represented by cg09798090) remained significantly associated with MS after accounting for rs6313 genotype ( $P = 0.04$ ).

**HTR2A is not differentially methylated in RRMS samples.** In order to identify if this mQTL is specific to progressive MS, we returned to our original RRMS cohort<sup>6</sup> and reanalysed the data using the same methods as above. There was no evidence for differential methylation in *HTR2A*. Similarly, there were no statistically significant differences between the risk allele frequency in the HC and the RRMS populations (Suppl. Table S4).

## Discussion

In this study we report the results of an EWAS of methylation levels in the CD4<sup>+</sup> T cells of female progressive MS patients compared to age and gender-matched healthy controls. We identified a DMR upstream of *HTR2A* in the TSS, which is located within the bounds of a haplotype block. Neither the DMR, nor the haplotype block are associated with RRMS and therefore appear to be progressive MS specific.

*HTR2A* codes for the serotonin 2A receptor. Serotonin is a neurotransmitter that acts in the gastrointestinal system, the blood and the brain. SNPs located within *HTR2A* have been associated with numerous neuropsychiatric disorders (including schizophrenia and obsessive-compulsive disorder) as well as autoimmune disease (rheumatoid arthritis)<sup>29,30</sup>. Although whether these SNPs contribute to disease has been recently challenged<sup>31</sup>. The most well studied SNP, rs6311 is known to be in near complete linkage disequilibrium with rs6313 and rs732821<sup>32</sup>. It is therefore not surprising that two of the three SNPs we identified in this study were rs6311 and 6313.

Expression of *HTR2A* is known to be controlled in the brain by functional polymorphisms, particularly rs6313 (102T/C allele) and rs6311 (-14378A/G allele). However, there is also evidence that low-level expression occurs in the peripheral lymphocytes<sup>27,33</sup>; although, there seems to be discrepancy over whether this expression is mono- or biallelic. In this study, we assessed the expression of *HTR2A* in our discovery cohort and found

no differences between healthy controls and MS patients. This could be due to low expression in CD4<sup>+</sup> T cells. Previous studies assessed a mixed population of PBMCs, therefore, it is possible that expression was driven primarily by a cell population other than CD4<sup>+</sup> T cells<sup>27,33</sup>. Alternately, the HT12 arrays only have a single probe for *HTR2A*. Studies, have suggested that there are alternative splice variants for *HTR2A*, which are not detected by the single probe available on the HT12 array. Future studies using next-generation sequencing technologies will be beneficial to answer these questions. Additional studies, which also evaluate other cell types will also be beneficial to determine if the methylation changes in CD4<sup>+</sup> T cells are reflective of changes in other cells. This would be particularly interesting in CNS resident cells.

The two genetic variants (rs6311 and rs6313) have also been shown to affect methylation status of *HTR2A* in other conditions such as schizophrenia and autism<sup>34</sup>. The minor A allele of rs6311 changes the CpG site to a CpA site, thus reducing the methylation in the promoter region<sup>35,36</sup>. This is also the case for the HLA-DRB1\*1501 allele in MS and for the rs6311 SNP in other diseases. Our logistic regression analysis suggests that the DMR at *HTR2A* exists regardless of rs6311 genotype. Although significance is marginal, these results suggest that methylation at the *HTR2A* gene are not completely driven by genotype and may be influenced by other factors in MS. Furthermore, comparison to our previous data sets, suggests that the DMR is progressive MS specific and supports the notion that *HTR2A* genotype is not the sole regulator of methylation at this locus in MS.

The link between MS and *HTR2A* is unclear. However, Malinova et al. reviewed the literature on serotonin and MS and point to a strong link between serotonin (5-HT) and gut microbiome<sup>37</sup>. Platelets normally take up 5HT in the gut, transport it to the peripheral blood, where it is secreted and transforms CD4<sup>+</sup> T cells into Th1 class cells<sup>37</sup>. In MS patients, the secretion fails to happen, thereby linking 5HT to disease activation.

Although we chose to focus on the *HTR2A* gene, due to its location within the TSS, we did also identify two other DMRs as part of this study in *HDAC4* and *SLC17A9*. Histone Deacetylase 4 is a class IIa histone deacetylase which is ubiquitously expressed throughout the body<sup>38</sup>. HDACs have a well-established roll in cell growth and differentiation. HDAC4 may play a role in neuronal plasticity, and there is some evidence that HDAC inhibitors lessen the neuropathy of mice with experimental models of MS (EAE)<sup>39</sup>. But a direct link between HDAC4 and MS in humans has not been described.

Solute Carrier Family 17 member 9A (SLC179A) codes for a vesicular nucleotide transporter that plays a role in storage and secretion of adenosine triphosphate (ATP). One study investigating normal appearing white matter (NAWM) of MS patients' brain tissue identified that SLC17A9 was hypomethylated in patients compared to healthy controls. This corresponded with an upregulation of the RNA transcript<sup>40</sup>. The second study, investigated the role of hypoxia in the pathophysiology of periplaque demyelinated lesions (PDL) in MS by comparing the PDL to NAWM<sup>41</sup>. They found that SLC179A was a hypoxia-inducible factor (HIF-1a) target gene that was upregulated in PDLs, suggesting that it may play a role in hypoxia induced demyelination<sup>41</sup>.

One limitation of the study is that the patients had a varied and complex treatment history. We cannot exclude that treatment is affecting our results and, given the small sample numbers, we are also not able to perform sub-cohort analysis. Additionally, two different arrays were used for this study, the 450K and EPIC arrays. The EPIC arrays contain nearly twice as many probes as the original 450K arrays, so we are unable to validate the findings from the extra probes. However, future larger studies with the EPIC arrays will help to further validate our findings.

Finally, there are currently no biomarkers of progression, and progressive MS is diagnosed retrospectively<sup>16</sup>. The absence of the *HTR2A* DMR in our previous cohort of RRMS patients suggests that this DMR may be specific to progressive disease, making it a potential biomarker of progression.

## Data availability

The dataset supporting the conclusions of this article is included in the supplemental files. For access to raw datasets contact Rodney Lea.

Received: 1 March 2019; Accepted: 22 November 2020

Published online: 17 December 2020

## References

1. Segal, B. M. Stage-specific immune dysregulation in multiple sclerosis. *J. Interferon Cytokine Res.* **34**, 633–640. <https://doi.org/10.1089/jir.2014.0025> (2014).
2. Broux, B., Stinissen, P. & Hellings, N. Which immune cells matter? The immunopathogenesis of multiple sclerosis. *Crit. Rev. Immunol.* **33**, 283–306 (2013).
3. Palmer, A. J., Colman, S., O'Leary, B., Taylor, B. V. & Simmons, R. D. The economic impact of multiple sclerosis in Australia in 2010. *Mult. Scler.* **19**, 1640–1646. <https://doi.org/10.1177/1352458513488230> (2013).
4. Zhu, H. et al. A genome-wide methylation study of severe vitamin D deficiency in African American adolescents. *J. Pediatr.* **162**, 1004–1009. <https://doi.org/10.1016/j.jpeds.2012.10.059> (2013).
5. Wan, E. S. et al. Cigarette smoking behaviors and time since quitting are associated with differential DNA methylation across the human genome. *Hum. Mol. Genet.* **21**, 3073–3082. <https://doi.org/10.1093/hmg/dds135> (2012).
6. Maltby, V. E. et al. Genome-wide DNA methylation profiling of CD8<sup>+</sup> T cells shows a distinct epigenetic signature to CD4<sup>+</sup> T cells in multiple sclerosis patients. *Clin. Epigenet.* **7**, 118. <https://doi.org/10.1186/s13148-015-0152-7> (2015).
7. Maltby, V. E. et al. A hypermethylated region in the promoter of the lymphotoxin-alpha gene in CD19<sup>+</sup> B cells is associated with multiple sclerosis. *Mult. Scler. J.* **8**, 17418 (2018).
8. Maltby, V. E. et al. Differential methylation at MHC in CD4<sup>+</sup> T cells is associated with multiple sclerosis independently of HLA-DRB1. *Clin. Epigenet.* **9**, 71. <https://doi.org/10.1186/s13148-017-0371-1> (2017).
9. Baranzini, S. E. et al. Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. *Nature* **464**, 1351–1356. <https://doi.org/10.1038/nature08990> (2010).
10. Bos, S. D. et al. Genome-wide DNA methylation profiles indicate CD8<sup>+</sup> T cell hypermethylation in multiple sclerosis. *PLoS ONE* **10**, e0117403. <https://doi.org/10.1371/journal.pone.0117403> (2015).

11. Graves, M. *et al.* Methylation differences at the HLA-DRB1 locus in CD4+ T-Cells are associated with multiple sclerosis. *Mult. Scler.* **20**, 1033–1041. <https://doi.org/10.1177/1352458513516529> (2013).
12. Ruhrmann, S. *et al.* Hypermethylation of MIR21 in CD4+ T cells from patients with relapsing-remitting multiple sclerosis associates with lower miRNA-21 levels and concomitant up-regulation of its target genes. *Mult. Scler.* **24**, 1288–1300. <https://doi.org/10.1177/1352458517721356> (2018).
13. Kular, L. *et al.* DNA methylation as a mediator of HLA-DRB1\*15:01 and a protective variant in multiple sclerosis. *Nat. Commun.* **9**, 2397. <https://doi.org/10.1038/s41467-018-04732-5> (2018).
14. Tintore, M., Vidal-Jordana, A. & Sastre-Garriga, J. Treatment of multiple sclerosis—success from bench to bedside. *Nat. Rev. Neurol.* **15**, 53–58. <https://doi.org/10.1038/s41582-018-0082-z> (2019).
15. Thompson, A. J. *et al.* Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol.* [https://doi.org/10.1016/S1474-4422\(17\)30470-2](https://doi.org/10.1016/S1474-4422(17)30470-2) (2017).
16. Lorscheider, J. *et al.* Defining secondary progressive multiple sclerosis. *Brain* **139**, 2395–2405. <https://doi.org/10.1093/brain/aww173> (2016).
17. Tian, Y. *et al.* ChAMP: updated methylation analysis pipeline for Illumina BeadChips. *Bioinformatics* **33**, 3982–3984 (2017).
18. Aryee, M. J. *et al.* Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363–1369 (2014).
19. Pidsley, R. *et al.* A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics* **14**, 293 (2013).
20. RStudio Team. *RStudio: Integrated Development for R* (RStudio Inc., Boston, 2018).
21. R: *A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2018).
22. Chen, Y. A. *et al.* Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics* **8**, 203–209. <https://doi.org/10.4161/epi.23470> (2013).
23. Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265. <https://doi.org/10.1093/bioinformatics/bth457> (2005).
24. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47. <https://doi.org/10.1093/nar/gkv007> (2015).
25. Smith, M. L., Baggerly, K. A., Bengtsson, H., Ritchie, M. E. & Hansen, K. D. illuminaio: an open source IDAT parsing tool for Illumina microarrays. *F1000Res* **2**, 264. <https://doi.org/10.12688/f1000research.2-264.v1> (2013).
26. Kobschull, M., Fittler, M. J., Demmer, R. T. & Papananou, P. N. Differential expression and functional analysis of high-throughput -omics data using open source tools. *Methods Mol. Biol.* **1537**, 327–345. [https://doi.org/10.1007/978-1-4939-6685-1\\_19](https://doi.org/10.1007/978-1-4939-6685-1_19) (2017).
27. Fukuda, Y. *et al.* Monoallelic and unequal allelic expression of the HTR2A gene in human brain and peripheral lymphocytes. *Biol. Psychiatry* **60**, 1331–1335. <https://doi.org/10.1016/j.biopsych.2006.06.024> (2006).
28. Petit, A. C. *et al.* Converging translational evidence for the involvement of the serotonin 2A receptor gene in major depressive disorder. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **54**, 76–82. <https://doi.org/10.1016/j.pnpbp.2014.04.013> (2014).
29. Williams, J. *et al.* Association between schizophrenia and T102C polymorphism of the 5-hydroxytryptamine type 2a-receptor gene. European Multicentre Association Study of Schizophrenia Group. *Lancet* **347**, 1294–1296. [https://doi.org/10.1016/s0140-6736\(96\)90939-3](https://doi.org/10.1016/s0140-6736(96)90939-3) (1996).
30. Mattina, G. F., Samaan, Z., Hall, G. B. & Steiner, M. The association of HTR2A polymorphisms with obsessive-compulsive disorder and its subtypes: a meta-analysis. *J. Affect. Disord.* **275**, 278–289. <https://doi.org/10.1016/j.jad.2020.06.016> (2020).
31. Spies, M. *et al.* Common HTR2A variants and 5-HTTLPR are not associated with human in vivo serotonin 2A receptor levels. *Hum. Brain Mapp.* <https://doi.org/10.1002/hbm.25138> (2020).
32. Spurlock, G. *et al.* A family based association study of T102C polymorphism in 5HT2A and schizophrenia plus identification of new polymorphisms in the promoter. *Mol. Psychiatry* **3**, 42–49. <https://doi.org/10.1038/sj.mp.4000342> (1998).
33. Guhathakurta, S. *et al.* Analysis of serotonin receptor 2A gene (HTR2A): association study with autism spectrum disorder in the Indian population and investigation of the gene expression in peripheral blood leukocytes. *Neurochem. Int.* **55**, 754–759. <https://doi.org/10.1016/j.neuint.2009.07.008> (2009).
34. Hranilovic, D., Blazevic, S., Stefulj, J. & Zill, P. DNA methylation analysis of HTR2A regulatory region in leukocytes of autistic subjects. *Autism Res.* **9**, 204–209. <https://doi.org/10.1002/aur.1519> (2016).
35. Parsons, M. J., D'Souza, U. M., Arranz, M. J., Kerwin, R. W. & Makoff, A. J. The -1438A/G polymorphism in the 5-hydroxytryptamine type 2A receptor gene affects promoter activity. *Biol. Psychiatry* **56**, 406–410. <https://doi.org/10.1016/j.biopsych.2004.06.020> (2004).
36. Turecki, G. *et al.* Prediction of level of serotonin 2A receptor binding by serotonin receptor 2A genetic variation in postmortem brain samples from subjects who did or did not commit suicide. *Am. J. Psychiatry* **156**, 1456–1458. <https://doi.org/10.1176/ajp.156.9.1456> (1999).
37. Malinova, T. S., Dijkstra, C. D. & de Vries, H. E. Serotonin: a mediator of the gut-brain axis in multiple sclerosis. *Mult. Scler.* **24**, 1144–1150. <https://doi.org/10.1177/1352458517739975> (2018).
38. Fischer, A., Sananbenesi, F., Mungenast, A. & Tsai, L. H. Targeting the correct HDAC(s) to treat cognitive disorders. *Trends Pharmacol. Sci.* **31**, 605–617. <https://doi.org/10.1016/j.tips.2010.09.003> (2010).
39. Camelo, S. *et al.* Transcriptional therapy with the histone deacetylase inhibitor trichostatin A ameliorates experimental autoimmune encephalomyelitis. *J. Neuroimmunol.* **164**, 10–21. <https://doi.org/10.1016/j.jneuroim.2005.02.022> (2005).
40. Huynh, J. L. *et al.* Epigenome-wide differences in pathology-free regions of multiple sclerosis-affected brains. *Nat. Neurosci.* **17**, 121–130. <https://doi.org/10.1038/nn.3588> (2014).
41. Satoh, J., Asahina, N., Kitano, S. & Kino, Y. Bioinformatics data mining approach indicates the expression of chromatin immunoprecipitation followed by deep sequencing (ChIP-Seq)-based hypoxia-inducible factor-1 $\alpha$  target genes in periplaque lesions of multiple sclerosis. *Clin. Exp. Neuroimmunol.* **6**, 159 (2015).

## Acknowledgements

We would like to thank the MS patients and clinical staff at the John Hunter Hospital MS clinic who participated in this study. We also acknowledge the Analytical Biomolecular Research Facility at the University of Newcastle for flow cytometry support and the Australian Genome Research Facility for performing the bisulfite conversions and hybridizations to the Illumina EPIC arrays.

## Author contributions

V.E.M. contributed to original study design, preformed experiments, was involved in the interpretation of the data, wrote the manuscript and revised all versions of the manuscript. R.A.L. performed data analysis, supervised data analysis, was involved in the interpretation of the data, and critically reviewed the manuscript. K.G., K.A.S., R.S. and S.B. performed experiments and critically reviewed the manuscript. N.W., D.K., A.X., T.V.C. and S.B. performed analysis and critically reviewed the manuscript. M.G., L.L. and H.B. assessed and processed patient samples and critically reviewed the manuscript. R.J.S. and J.L.S. contributed to the original study design,

critically edited the manuscript and are responsible for the infrastructure in which the study was conducted. J.L.S. supervised all aspects of the study.

### Funding

This study was supported by a grant from Multiple Sclerosis Research Australia (13-037, 14-027). RAL, VEM, RS and KAS are supported by fellowships from Multiple Sclerosis Research Australia. KAS was supported by a scholarship from the Trish foundation. VEM was supported by a fellowship from the Canadian Institutes of Health Research (320577).

### Competing interests

VEM: has received honoraria for presentations from Biogen and Merck Serono. HB: serves on steering committees and scientific advisory boards for Merck, Biogen, Novartis and Roche. Has received conference travel support from Merck. Institution has received honoraria for speaker engagements for Merck, Biogen, Roche and Novartis. Institution has received research support from Biogen, Roche, Merck, Novartis, NHMRC, MRFF, Trish Foundation, MSRA. Receives personal compensation for serving on the Brain health initiative Steering Committee. JLS'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, Teva and Novartis Australia. The other authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-020-78809-x>.

**Correspondence** and requests for materials should be addressed to J.L.-S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020