



OPEN

# Replication analysis of variants associated with multiple sclerosis risk

Mohammad Dashti<sup>1</sup>, Khadijah Ateyah<sup>2</sup>, Raed Alroughani<sup>3</sup> & Rabeah Al-Temaimi<sup>4</sup>✉

Multiple Sclerosis (MS) is a complex chronic neurodegenerative disorder resulting from an autoimmune reaction against myelin. So far, many genetic variants have been reported to associate with MS risk however their association is inconsistent across different populations. Here we investigated the association of the most consistently reported genetic MS risk variants in the Kuwaiti MS population in a case-control study designs. Of the 94 reported MS risk variants four variants showed MS risk association in Arabs exome analysis (*EVI5* rs11808092  $p = 0.0002$ ; *TNFRSF1A* rs1800693  $p = 0.00003$ ; *MTHFR* rs1801131  $p = 0.038$ ; and *CD58* rs1414273  $p = 0.00007$ ). Replication analysis in Kuwaiti MS cases and healthy controls confirmed *EVI5* rs11808092A (OR: 1.6, 95%CI: 1.19–2.16,  $p = 0.002$ ) and *MTHFR* rs1801131G (OR: 1.79, 95%CI: 1.3–2.36,  $p = 0.001$ ) as MS risk genetic factors, while *TNFRSF1A* rs1800693C had a marginal MS risk association (OR: 1.36, 95%CI: 1.04–1.78,  $p = 0.025$ ) in the Kuwaiti population. *CD58* rs1414273 did not sustain risk association ( $p = 0.37$ ). In conclusion, *EVI5* rs11808092A, *TNFRSF1A* rs1800693C and *MTHFR* rs1801131G are MS risk factors in the Kuwaiti population. Further investigations into their roles in MS pathogenesis and progression are merited.

Multiple Sclerosis (MS) is a complex, chronic, neurodegenerative autoimmune disorder that affects the central nervous system (CNS). MS manifests in the CNS through formation of inflammatory lesions in the white matter of the brain and spinal cord. The lesions represent an endpoint of repeated autoimmune attacks against endogenous myelin-associated antigens which lead to neuroaxonal degeneration, and death of myelin forming cells; the oligodendrocytes. This pathological process results in demyelinated plaques or astrocytic scars throughout the CNS perturbing neural networks<sup>1</sup>. It is estimated that 2,221,188 individuals are living with MS worldwide, and global MS incidence is twice as many in females than in males<sup>2</sup>. In Kuwait, MS prevalence has increased from 31.15 per 100,000 individual in 2009, to 104.88 per 100,000 individual in 2018 with a projected trend of further increases in the future<sup>3,4</sup>. MS has been recognized and described 151 years ago but its underlying cause remains unknown. However, there is mounting evidence of several environmental and genetic risk factors associating with the disease<sup>5,6</sup>. Environmental and genetic contributing factors to MS incidence are not exclusive but are thought to be synergistic as none of these factors fully explain causation alone, nor do they consistently associate with MS risk across different populations<sup>7</sup>.

Among the strongest genetic susceptibility associations with MS is the major histocompatibility complex (MHC) genes region<sup>8</sup>. The most significant associations were seen with polymorphisms in human leukocyte antigen (HLA) HLA-DR2 alleles, however their association was found variable across different MS populations<sup>9,10</sup>. More MS risk associations were found with other HLA loci such as HLA-DR3, -DR4, -DQ2, -DQ6, and -DQ8 allelic variants<sup>11</sup>. However, the associations found with HLA haplotypes are thought to explain only 25–60% of MS genetics. A large number of linkage analyses, genome wide association studies (GWAS) and candidate gene studies have been conducted to discover the missing genetics of MS. Association studies from Kuwait reporting MS genetic risk factors are sparse, and include positive association of polymorphisms in HLA class-II (DR4, DQ6, DQ7 and DQ8) genes<sup>12</sup>, nitric oxide synthase (NOS) genes<sup>13</sup>, vitamin D receptor gene<sup>14</sup>, and leptin gene with MS risk<sup>15</sup>. However, currently there are approximately more than 200 MS genetic risk factors reported so far and replication studies from Kuwait are lacking<sup>16</sup>. Interestingly, the ~200 MS genetic risk factors include different susceptibility markers and modifier genes with inconsistent associations across different MS populations emphasizing

<sup>1</sup>Genetics and Bioinformatics department, Dasman Diabetes Institute, Sharq, Kuwait City, Kuwait. <sup>2</sup>Undergraduate medical program, Faculty of Medicine, Kuwait University, Jabriya, Kuwait. <sup>3</sup>Neurology Clinic, Amiri Hospital, Kuwait City, Kuwait. <sup>4</sup>Human Genetics Unit, Department of Pathology, Faculty of Medicine, Kuwait University, Jabriya, Kuwait. ✉e-mail: [rabeah@hsc.edu.kw](mailto:rabeah@hsc.edu.kw)

Criteria	MS patients (n = 113)	Healthy Arab controls (n = 460)
Sex (%)		
Male	40 (35.4)	229 (49.8)
Female	73 (64.6)	231 (50.2)
Age in years (Mean $\pm$ SD <sup>a</sup> )	32.9 $\pm$ 9.4	40.7 $\pm$ 12.6
MS type <sup>b</sup> (%)		—
RRMS	16 (14.2)	
SPMS	93 (82.3)	
PPMS	4 (3.5)	
Disease duration in years (Mean $\pm$ SD)	6.77 $\pm$ 5.73	—
EDSS <sup>c</sup> (Median, IQR <sup>d</sup> )	2 (1.5–3.5)	—

**Table 1.** Demographics and clinical characteristics of Kuwaiti MS (n = 113) and healthy control Arabs (n = 460) included in the investigative exome analysis. <sup>a</sup>SD: Standard deviation; <sup>b</sup>RRMS: Relapsing-remitting MS, SPMS: Secondary progressive MS, PPMS: Primary progressive MS; <sup>c</sup>EDSS: Expanded disability status scale; <sup>d</sup>IQR: Interquartile range.

Gene	SNP	Allele frequency in Kuwaiti MS cases (%)	Allele frequency in Arab healthy controls (%)	P-value
<i>CD58</i>	rs1414273	G: 158 (76) A: 50 (24)	G: 722 (87.4) A: 104 (12.6)	0.00007
<i>EVI5</i>	rs11808092	C: 93 (63.3) A: 54 (36.7)	C: 687 (77.9) A: 195 (22.1)	0.00024
<i>MTHFR</i>	rs1801131	T: 120 (57.1) G: 90 (42.9)	T: 574 (65.1) G: 308 (34.9)	0.038
<i>TNFRSF1A</i>	rs1800693	T: 127 (60.5) C: 83 (39.5)	T: 631 (75.5) C: 205 (24.5)	0.00002

**Table 2.** Statistically significant exome analysis results of reported variants associated with MS risk in this study's case-control exome cohorts. Variant detection frequency was variable across the two exome cohorts and computed allele frequencies reflect the proportion of alleles divided by the total number of exomes in which the variant was detected.

ethnic, geo-epidemiological, and environmental factors associating with MS risk<sup>17</sup>. To confirm the association of reported MS risk genetic factors, replication studies should be conducted across different MS populations to understand the influence of ethnic and geo-epidemiological factors on MS risk. Here, we report our findings from a replication study investigating non-HLA reported MS risk variants in a sampled Kuwaiti MS population. Our objective was to assess their association in a semi-ethnically homogeneous Kuwaiti cohort as these genetic risk factors were identified from multi-ethnic MS populations.

## Results

**Reported MS variants in arab exomes.** The demographics and clinical characteristics of Kuwaiti MS patients and healthy Arab controls included in the exome analysis are shown in Table 1. Of the 96 selected variants 87 (90.6%) had variable detection frequency in SureSelect V5 library only, and 9 (10.4%) were not covered by our exome library. The final list of reported MS variants that had consistent detection across all exomes included 18 (20.7%) variants of acceptable detection frequencies across the two cohorts (Supplementary Table 1). Four variants had statistically significant different allele frequencies in MS exomes compared to healthy control exomes (Table 2). These variants included; Ecotropic Viral Integration Site 5 (*EVI5*) rs11808092 (OR: 2.04, 95%CI: 1.4–2.9), TNF Receptor Superfamily Member 1 A (*TNFRSF1A*) rs1800693 (OR: 2.01, 95%CI: 1.45–2.75), Lymphocyte Function-Associated Antigen 3 (*CD58*) rs1414273 (OR: 2.2, 95%CI: 1.5–3.2), and a nominally significant and not clearly robust Methylenetetrahydrofolate Reductase (*MTHFR*) variant rs1801131 (OR: 1.4, 95%CI: 1.02–1.9). However, genotype frequencies were significantly different among the two cohorts for only three variants; *TNFRSF1A* rs1800693 ( $p = 0.0001$ ), *MTHFR* rs1801131 ( $p = 0.024$ ), and *CD58* rs1414273 ( $p = 0.0001$ ). Since these variants showed risk association in Kuwaiti MS patients when compared to Arab healthy controls we further investigated their association with MS in an exclusively Kuwaiti nationality case-control population sample.

**Replication analysis of MS variants in kuwaiti samples.** The replication Kuwaiti population sample included 170 Kuwaiti MS patients and 311 healthy Kuwaitis. Replication cohorts' demographics and clinical characteristics are shown in Table 3. Allelic and genotype frequencies for *TNFRSF1A* rs1800693, *EVI5* rs11808092, *CD58* rs1414273 and *MTHFR* rs1801131 are shown in Table 4. *TNFRSF1A* rs1800693, *EVI5* rs11808092 and *MTHFR* rs1801131 genotype frequencies in healthy Kuwaiti controls were in Hardy-Weinberg equilibrium, except for *CD58* rs1414273 for which genotype frequencies were significantly different than those expected in European populations ( $p < 0.0001$ ). *TNFRSF1A* rs1800693C allele did not sustain significant association with

Criteria	Kuwaiti MS patients (n = 170)	Kuwaiti healthy controls (n = 311)
<b>Sex (%)</b>		
Male	58 (34.1)	118 (37.9)
Female	112 (65.9)	193 (62.1)
Age in years (Mean $\pm$ SD)	33 $\pm$ 8.73	30.7 $\pm$ 8.47
MS age of onset in years (Mean $\pm$ SD)	27.06 $\pm$ 7.68	
MS type (%)		—
RRMS	148 (87)	
SPMS	14 (8.2)	
PPMS	4 (2.4)	
Benign MS	4 (2.4)	
Disease duration in years (Mean $\pm$ SD)	5.6 $\pm$ 5.17	—
EDSS (Median, IQR)	1.5 (1–3)	—

**Table 3.** Demographics and clinical characteristics of the replication study's Kuwaiti case-control cohorts and their allelic and genotype frequencies for the associated MS risk variants.

Variant	MS cohort	Healthy controls	P-value
	(n = 170)	(n = 311)	
<i>CD58</i> rs1414273			
A/G	73 (21.5)/267 (78.5)	137 (22)/485 (78)	0.8
AA	7 (4.1)	25 (8)	0.11
AG	59 (34.7)	87 (28)	
GG	104 (61.2)	199 (64)	
<i>EVI5</i> rs11808092			
A / C	110 (32.4)/230 (67.6)	141 (22.6)/481 (77.3)	0.002
AA	20 (11.7)	15 (4.8)	0.003
AC	70 (41.2)	111 (35.7)	
CC	80 (47.1)	185 (59.5)	
<i>MTHFR</i> rs1801131			
G / T	149 (43.8)/191 (56.2)	189 (30.4)/433 (69.6)	0.00003
GG	31 (18.2)	26 (8.4)	0.0001
GT	87 (51.2)	137 (44)	
TT	52 (30.6)	148 (47.6)	
<i>TNFRSF1A</i> rs1800693			
C / T	143 (42)/197 (58)	216 (34.7)/406 (65.3)	0.025
CC	27 (15.9)	42 (13.5)	0.03
CT	89 (52.3)	132 (42.4)	
TT	54 (31.8)	137 (44.1)	

**Table 4.** Replication cohorts' allelic and genotype distributions [n (%)] of the exome analysis resultant MS risk variants.

MS risk (OR: 1.36, 95%CI: 1.04–1.78,  $p = 0.025$ ) following adjustment for multiple testing. *TNFRSF1A* rs1800693 genotype distribution differed between the two cohorts ( $\beta = -0.106$ , 95%CI:  $-0.14$ – $(-0.009)$ ,  $p = 0.027$ ) after adjusting for sex and age albeit did not reach adjusted statistical significance. However, rs1800693CC+CT genotypes significantly associated with MS risk in an autosomal dominant inheritance model (OR: 1.7, 95%CI: 1.14–2.5,  $p = 0.008$ ). No sex-specific MS risk or EDSS association was found for rs1800693, whereas genotype CC marginally associated with a younger age of MS onset ( $\beta = -0.169$ , 95%CI:  $-0.33$ – $(-0.001)$ ,  $p = 0.04$ ) after adjusting for sex and disease duration.

*EVI5* rs11808092A allele showed significant MS risk association in the Kuwaiti population sample (OR: 1.6, 95%CI: 1.19–2.16,  $p = 0.002$ ). *EVI5* rs11808092 genotype distribution differed between the two cohorts ( $\beta = -0.195$ , 95%CI:  $-0.24$ – $(-0.06)$ ,  $p = 0.001$ ) when adjusted for age and sex. In addition, *EVI5* rs11808092 genotype association with female MS risk was stronger than in males ( $p = 0.007$  in females vs.  $p = 0.042$  in males). *EVI5* rs11808092 associated significantly with MS risk in both; autosomal dominant (OR: 1.65, 95%CI: 1.1–2.4,  $p = 0.009$ ) and recessive (OR: 2.63, 95%CI: 1.32–5.2,  $p = 0.009$ ) inheritance models. *MTHFR* rs1801131G allele significantly associated with MS risk (OR: 1.79, 95%CI: 1.3–2.36,  $p = 0.00003$ ). *MTHFR* rs1801131 genotype

distribution differed between the two cohorts ( $\beta = -0.23$ , 95%CI:  $-0.25$ -  $(-0.09)$ ,  $p = 0.0001$ ) after adjusting for sex and age. *MTHFR* rs1801131 GG + GT strongly associated with MS risk in an autosomal dominant inheritance model (OR: 2.06, 95%CI: 1.39–3.1,  $p = 0.0003$ ). Lastly, *CD58* rs1414273A allele did not show any MS risk association in the replication Kuwaiti only population sample. *CD58* rs1414273 genotypes were not significantly different with or without adjusting for sex ( $p = 0.69$  and  $0.38$ , respectively). *CD58* rs1414273 genotypes did not associate with MS risk in any of the inheritance models applied. None of the assessed variants associated with EDSS with or without adjusting for disease duration or/and sex.

## Discussion

There are currently more than 200 reported MS genetic risk factors stemming from linkage analysis studies, GWAS studies, functional candidate gene approaches, exome sequencing association studies, and animal model and post mortem MS brain candidate gene studies. Only a handful of these genetic factors sustained MS risk association across different MS populations, and were linked to MS pathogenesis which provided valuable potential therapeutic targets. Here, we focused on reported non-HLA MS risk variants that might be captured in exome libraries available for clinical research. We noticed that exome coverage is not consistent across samples, and different libraries have variable exome coverage as well. Four candidate genetic factors were found to have statistically significant MS risk association in the Kuwaiti MS exome cohort when compared to Arab healthy controls. However, the Kuwaiti only replication cohorts; which adjusted for possible genetic background variation/bias in the exome study, highlighted the influence of ethnicity and genetic background in false positive results of genetic studies.

*TNFRSF1A* rs1800693, *EVI5* rs11808092, and *MTHFR* rs1801131 have been reported to associate with MS risk in mostly European Caucasian populations (Supplementary Table 2). However, the effect size of these associations is variable and is relatively small depending largely on the number of samples analyzed. *TNFRSF1A* is a membrane-bound and soluble receptor for tumor necrosis factor-alpha (TNF $\alpha$ ) that plays a role in cellular survival, apoptosis and inflammation in the immune and nervous systems<sup>18</sup>. *TNFRSF1A* rs1800693 is an intronic variant that has been shown to impact the splicing of *TNFRSF1A* mRNA resulting in a novel isoform that blocks the action of TNF $\alpha$ <sup>19</sup>. Rs1800693 association with MS risk has been shown to have consistently a small effect size in different MS populations<sup>20</sup>. It is possible that the marginal association of rs1800693 variant with an earlier age of MS onset seen in our study is the reason for the small effect size of this variant reported in other MS populations where younger age of MS onset cases were considerably well represented. Therefore, it is probable that due to the low representation of MS cases with younger age of MS onset in our replication MS cohort (17% with MS age of onset  $\leq 20$  years) rs1800693 association with MS risk did not reach statistical significance following adjustment for multiple testing.

*EVI5* rs11808092 variant lies in the 3'-end intron of *EVI5* gene that shares a similarity to an enhancer element and has been shown to act as a strong enhancer element on the promoter of an adjacent gene (*GFI1*) implicated in MS risk<sup>21</sup>. Nevertheless, *EVI5* rs11808092 as well as other variants in *EVI5* have been reported to associate with MS risk and clinical characteristics<sup>22,23</sup>. On the other hand, *MTHFR* rs1801131 had a nominal association with MS risk in our exome analysis, but later showed a robust association with MS risk in the Kuwaiti only cohorts. *MTHFR* is an enzyme involved in the conversion of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate and the missense variant rs1801131 (E429A) has been shown to cause decreased *MTHFR* enzymatic function and is associated with impaired folate metabolism and mild increases in homocysteine levels (hyperhomocysteinemia) that is reported to occur in MS patients<sup>24,25</sup>. Sustained elevated homocysteine levels associate with cardiovascular disturbances that are thought to predispose to MS pathogenesis and progression<sup>26</sup>.

Lastly, *CD58* was first identified as an MS susceptibility factor by the association of its rs12044852 variant in a study of African American MS patients and further supported by evidence of altered expression in MS<sup>27,28</sup>. A follow-up investigation provided two other *CD58* variants (rs2300747, rs1335532) associating with altered *CD58* expression in MS<sup>29</sup>. Structural analysis of *CD58* coding region revealed rs1414273 to be in strong linkage disequilibrium with the *CD58* MS-associated SNP rs1335532<sup>30</sup>. A recent report highlighted the role of this variant in processing of the *CD58* transcript and associated microRNA in MS pathology<sup>31</sup>. Our exome analysis did not detect any of the other *CD58* variants except for rs1414273 and initially an association with MS risk was found in the exome analysis. However, the association was later lost in the Kuwaiti only replication analysis. It should be noted that rs1414273A is the minor allele in Kuwaitis and has a lower frequency in Kuwaitis (0.2) compared to ExAc database (0.5) and 1000 genome database (0.43) frequencies. Therefore, it is plausible that rs1414273 allelic distribution is well-retained in the majority of Kuwait's genetic background, and rs1414273 reported linkage to rs1335532 may not be applicable to the Kuwaiti genetic background. Moreover, Qatar's patriotization laws are very flexible with an estimated 11.6% Arab Qataris contributing to the largely immigrant 2,723,003 Qatari population<sup>32</sup>. Therefore, the association of rs1414273 with MS risk in our exome study was probably influenced by a non-uniform genetic background in the Qatari controls. Nevertheless, our replication study findings support the lack of rs1414273 association with MS risk in the Kuwaiti MS population but does not refute other *CD58* variants association with MS risk or *CD58* altered expression in MS.

In conclusion, we have analyzed the association of a set of reported MS genetic risk variants in a semi-homogenous exome case-control study, and confirmed the results in a replication study of a homogenous genetic background. Population genetic analyses should strive to maintain genetic background uniformity to accurately ascertain population specific genetic risk susceptibilities. Our findings support the existing evidence on the roles of these genes in MS risk and pathogenesis providing further evidence into MS etiology and potential therapeutic targets.

## Methods

**Patient selection.** Blood samples of 283 Kuwaiti MS patients were collected at Kuwait's Dasman Diabetes Institute (DDI), and 311 Kuwaiti healthy control volunteers were collected by word of mouth and social networks from the Kuwaiti population. Collection criteria for MS patients included; a complete clinical MS disease profile (demographics, age of onset, type of MS, disease duration, expanded disability status scale (EDSS) score, and history of MS treatments), being a Kuwaiti citizen, and agreement to provide a 4 mL blood sample. Exclusion criteria included; having an EDSS score of 0, and a disease duration of  $\leq 1$  year. Healthy controls' exclusion criteria included; being a non-Kuwaiti expatriate, having a family history of MS, and having a diagnosis of any complex disorder. All study protocols were approved by Kuwait's Health Sciences Center's Joint Committee for The Protection of Human Subjects and DDI's ethical review committee both of which adhere to declaration of Helsinki's Ethical Principles for Medical Research Involving Human Subjects' guidelines. All study information and protocols were fully explained to all participants prior to procurement of their informed written consent. Of the 283 MS patients collected, 113 were prioritized as most informative for exome sequencing to include equally distributed subgroups relevant to each of the following characteristics; EDSS score, disease duration, age group, and male to female ratio of  $\sim 1:2$ . Whereas 56 healthy control Kuwaiti samples were age and sex matched to selected MS samples and prioritized for exome sequencing.

**DNA extraction.** Blood samples were collected from all 283 MS patients, while healthy controls' sample collection varied between blood and saliva samples. Blood samples were centrifuged at 2,500xg for 10 minutes at room temperature and three fractions were retrieved. Plasma fractions were stored at  $-80^{\circ}\text{C}$ , and white blood cells fractions (buffy coat) were subjected to DNA extraction using QIAamp DNA blood mini kit (Qiagen, CA, USA) according to manufacturer's standard protocol. For saliva samples the same kit was used with minor modifications. For every 1 mL saliva sample 4 mL of sterile phosphate buffer saline (PBS) were added, and samples were centrifuged at  $1,800 \times g$  for 5 minutes at room temperature. Resultant pellet was suspended in 180  $\mu\text{l}$  of PBS and 20  $\mu\text{l}$  of kit provided proteinase K was added. The following steps were according to manufacturer's standard kit protocol. DNA quality and quantity were assessed by a NanoDrop spectrophotometer.

**Exome sequencing.** Exome sequencing of 113 Kuwaiti MS patients and 56 healthy Kuwaiti controls was performed on an Illumina HiSeq2000 platform using SureSelectXT v5 library preparation with target coverage of 50X (Illumina, CA, USA). Illumina raw paired-end reads, captured with Agilent exome library, were mapped to the human reference assembly (hg19) using Burrows wheeler aligner (BWA)<sup>33</sup>. Sequence Alignment Map data with average of 50X coverage were then compiled and converted to a single compressed binary file, for each strain, using SAMtools<sup>34</sup>. Picard software (<http://picard.sourceforge.net>) was used to flag PCR duplicate reads. Genome Analysis Toolkit (GATK) was used to (a) local realignment of a read that overlapped with an INDEL (b) recalibration of base quality (c) variants calling<sup>35,36</sup>.

**MS variant selection.** A literature search was conducted using the following keywords "Multiple sclerosis genetics, Multiple sclerosis gene, multiple sclerosis polymorphism, multiple sclerosis variant, multiple sclerosis GWAS, multiple sclerosis exome sequencing, multiple sclerosis genome sequencing." In addition, variants included in the MSgene database (MSgene.org) were also retrieved. Collected variants were filtered according to the following criteria; being a single nucleotide variant, being within intronic gene sequence, being encoded in the nuclear genome, and being an MS risk factor. A total of 96 genetic MS risk factors were selected and used for mining exome data of MS patients and healthy controls (Supplementary Table 1). Two healthy control datasets were retrieved from publically available databases that are believed to share genetic background with Kuwaitis; Qatari and middle-eastern exome sequences. Selected Qatari and Middle Eastern (non-diabetic) individuals next generation sequencing data were obtained from the national center for biotechnology information (NCBI) sequence read archive with accessions; SRP060765, SRP061943 and SRP061463. Collectively, the case-control cohorts included 113 MS exomes and 460 healthy control exomes. A Fisher exact test was carried out for each variant, and variants were selected based on p-value of significance  $< 0.05$ .

**Variant genotyping in replication cohorts.** DNA from a replicate MS cohort of 170 MS patients and 311 healthy controls was used to genotype polymorphisms with p-values  $< 0.05$ . Genotyping was performed using Taqman genotyping assays (Life Technologies, CA, USA) according to manufacturer's standard protocols and analyzed using ABI 7500 Fast Real-time PCR system (Life technologies, CA, USA). Genotype allelic discrimination was determined by SDS v1.4.1 software (Applied Biosystems, CA, USA).

**Statistical analysis.** Hardy-Weinberg equilibrium was assessed in the Kuwaiti healthy control cohort for the four replicated variants using allele frequencies of European ancestry as the majority of detected MS risk variants were reported from European populations. For genotype analysis Fisher's exact-test, chi-square test, and linear and logistic regression analyses were used. All statistical analyses were performed using SPSS v.25 (IBM, NY, USA). For the replication analysis, a multiple testing Bonferroni-adjusted p-value  $\leq 0.0125$  was considered significant.

Received: 8 November 2019; Accepted: 15 April 2020;  
Published online: 30 April 2020

## References

- Noseworthy, J. H., Lucchinetti, C., Rodriguez, M. & Weinschenker, B. G. Multiple sclerosis. *The New England journal of medicine* **343**, 938–952, <https://doi.org/10.1056/NEJM200009283431307> (2000).
- Collaborators, G. B. D. M. S. Global, regional, and national burden of multiple sclerosis 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet. Neurology* **18**, 269–285, [https://doi.org/10.1016/S1474-4422\(18\)30443-5](https://doi.org/10.1016/S1474-4422(18)30443-5) (2019).
- Alroughani, R. *et al.* The prevalence of multiple sclerosis continues to increase in Kuwait. *Multiple sclerosis and related disorders* **32**, 74–76, <https://doi.org/10.1016/j.msard.2019.04.033> (2019).
- Alshubaili, A. F., Alramzy, K., Ayyad, Y. M. & Gerish, Y. Epidemiology of multiple sclerosis in Kuwait: new trends in incidence and prevalence. *European neurology* **53**, 125–131, <https://doi.org/10.1159/000085556> (2005).
- Didonna, A. & Oksenberg, J. R. In *Multiple Sclerosis: Perspectives in Treatment and Pathogenesis* (eds. I. S. Zagon & P. J. McLaughlin) (2017).
- Michel, L. Environmental factors in the development of multiple sclerosis. *Revue neurologique* **174**, 372–377, <https://doi.org/10.1016/j.neurol.2018.03.010> (2018).
- Olsson, T., Barcellos, L. F. & Alfredsson, L. Interactions between genetic, lifestyle and environmental risk factors for multiple sclerosis. *Nature reviews. Neurology* **13**, 25–36, <https://doi.org/10.1038/nrneurol.2016.187> (2017).
- Haines, J. L. *et al.* Linkage of the MHC to familial multiple sclerosis suggests genetic heterogeneity. The Multiple Sclerosis Genetics Group. *Human molecular genetics* **7**, 1229–1234 (1998).
- Barcellos, L. F. *et al.* Heterogeneity at the HLA-DRB1 locus and risk for multiple sclerosis. *Human molecular genetics* **15**, 2813–2824, <https://doi.org/10.1093/hmg/ddl223> (2006).
- Schmidt, H., Williamson, D. & Ashley-Koch, A. HLA-DR15 haplotype and multiple sclerosis: a HuGE review. *American journal of epidemiology* **165**, 1097–1109, <https://doi.org/10.1093/aje/kwk118> (2007).
- Zivadinov, R. *et al.* HLA-DRB1\*1501, -DQB1\*0301, -DQB1\*0302, -DQB1\*0602, and -DQB1\*0603 alleles are associated with more severe disease outcome on MRI in patients with multiple sclerosis. *International review of neurobiology* **79**, 521–535, [https://doi.org/10.1016/S0074-7742\(07\)79023-2](https://doi.org/10.1016/S0074-7742(07)79023-2) (2007).
- Al-Shammri, S., Nelson, R. F., Al-Muzairi, I. & Akanji, A. O. HLA determinants of susceptibility to multiple sclerosis in an Arabian Gulf population. *Multiple sclerosis* **10**, 381–386, <https://doi.org/10.1191/1352458504ms10650a> (2004).
- AlFadhli, S., Mohammed, E. M. & Al Shubaili, A. Association analysis of nitric oxide synthases: NOS1, NOS2A and NOS3 genes, with multiple sclerosis. *Annals of human biology* **40**, 368–375, <https://doi.org/10.3109/03014460.2013.786756> (2013).
- Al-Temaimi, R. A., Al-Enezi, A., Al-Serri, A., Alroughani, R. & Al-Mulla, F. The Association of Vitamin D Receptor Polymorphisms with Multiple Sclerosis in a Case-Control Study from Kuwait. *PLoS one* **10**, e0142265, <https://doi.org/10.1371/journal.pone.0142265> (2015).
- Dashti, M., Alroughani, R., Jacob, S. & Al-Temaimi, R. Leptin rs7799039 polymorphism is associated with multiple sclerosis risk in Kuwait. *Multiple sclerosis and related disorders* **36**, 101409, <https://doi.org/10.1016/j.msard.2019.101409> (2019).
- Cotsapas, C. & Mitrovic, M. Genome-wide association studies of multiple sclerosis. *Clinical & translational immunology* **7**, e1018, <https://doi.org/10.1002/cti2.1018> (2018).
- Baranzini, S. E. & Oksenberg, J. R. The Genetics of Multiple Sclerosis: From 0 to 200 in 50 Years. *Trends in genetics: TIG* **33**, 960–970, <https://doi.org/10.1016/j.tig.2017.09.004> (2017).
- Varfolomeev, E. & Vucic, D. Intracellular regulation of TNF activity in health and disease. *Cytokine* **101**, 26–32, <https://doi.org/10.1016/j.cyto.2016.08.035> (2018).
- Gregory, A. P. *et al.* TNF receptor 1 genetic risk mirrors outcome of anti-TNF therapy in multiple sclerosis. *Nature* **488**, 508–511, <https://doi.org/10.1038/nature11307> (2012).
- International Multiple Sclerosis Genetics, C. The genetic association of variants in CD6, TNFRSF1A and IRF8 to multiple sclerosis: a multicenter case-control study. *PLoS one* **6**, e18813, <https://doi.org/10.1371/journal.pone.0018813> (2011).
- Martin, D. *et al.* Genome-wide CTCF distribution in vertebrates defines equivalent sites that aid the identification of disease-associated genes. *Nature structural & molecular biology* **18**, 708–714, <https://doi.org/10.1038/nsmb.2059> (2011).
- Mowry, E. M. *et al.* Multiple sclerosis susceptibility genes: associations with relapse severity and recovery. *PLoS one* **8**, e75416, <https://doi.org/10.1371/journal.pone.0075416> (2013).
- Didonna, A. *et al.* A non-synonymous single-nucleotide polymorphism associated with multiple sclerosis risk affects the EVI5 interactome. *Human molecular genetics* **24**, 7151–7158, <https://doi.org/10.1093/hmg/ddv412> (2015).
- Russo, C. *et al.* Hyperhomocysteinemia is associated with cognitive impairment in multiple sclerosis. *Journal of neurology* **255**, 64–69, <https://doi.org/10.1007/s00415-007-0668-7> (2008).
- Oliveira, S. R. *et al.* Elevated plasma homocysteine levels are associated with disability progression in patients with multiple sclerosis. *Metabolic brain disease* **33**, 1393–1399, <https://doi.org/10.1007/s11011-018-0224-4> (2018).
- Spencer, J. I., Bell, J. S. & DeLuca, G. C. Vascular pathology in multiple sclerosis: reframing pathogenesis around the blood-brain barrier. *Journal of neurology, neurosurgery, and psychiatry* **89**, 42–52, <https://doi.org/10.1136/jnnp-2017-316011> (2018).
- Reich, D. *et al.* A whole-genome admixture scan finds a candidate locus for multiple sclerosis susceptibility. *Nature genetics* **37**, 1113–1118, <https://doi.org/10.1038/ng1646> (2005).
- Arthur, A. T. *et al.* Genes implicated in multiple sclerosis pathogenesis from consilience of genotyping and expression profiles in relapse and remission. *BMC medical genetics* **9**, 17, <https://doi.org/10.1186/1471-2350-9-17> (2008).
- Handel, A. E. *et al.* The effect of single nucleotide polymorphisms from genome wide association studies in multiple sclerosis on gene expression. *PLoS one* **5**, e10142, <https://doi.org/10.1371/journal.pone.0010142> (2010).
- Hecker, M., Fitzner, B., Blaschke, J., Blaschke, P. & Zettl, U. K. Susceptibility variants in the CD58 gene locus point to a role of microRNA-548ac in the pathogenesis of multiple sclerosis. *Mutation research. Reviews in mutation research* **763**, 161–167, <https://doi.org/10.1016/j.mrrev.2014.10.002> (2015).
- Hecker, M. *et al.* A genetic variant associated with multiple sclerosis inversely affects the expression of CD58 and microRNA-548ac from the same gene. *PLoS genetics* **15**, e1007961, <https://doi.org/10.1371/journal.pgen.1007961> (2019).
- Rowold, D. J., Luis, J. R., Terreros, M. C., Herrera, R. J. & Mitochondrial, D. N. A. gene flow indicates preferred usage of the Levant Corridor over the Horn of Africa passageway. *Journal of human genetics* **52**, 436–447, <https://doi.org/10.1007/s10038-007-0132-7> (2007).
- Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760, <https://doi.org/10.1093/bioinformatics/btp324> (2009).
- Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079, <https://doi.org/10.1093/bioinformatics/btp352> (2009).
- McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* **20**, 1297–1303, <https://doi.org/10.1101/gr.107524.110> (2010).
- DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature genetics* **43**, 491–498, <https://doi.org/10.1038/ng.806> (2011).

## Acknowledgements

This work was funded by KFAS grant 2012–1302–02.

### Author contributions

M.D. performed the bioinformatics analyses, K.A. performed genotyping of selected variants, R.A.R. provided patients clinical data, and R.A.T. designed the study, collected samples and performed DNA extractions, genetic analyses and drafted the manuscript. All authors approved the final manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-64432-3>.

**Correspondence** and requests for materials should be addressed to R.A.-T.

**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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020