

ORIGINAL ARTICLE

Gene-expression signatures: biomarkers toward diagnosing multiple sclerosis

JT Tossberg¹, PS Crooke², MA Henderson³, S Sriram⁴, D Mrelashvili⁵, S Chitnis⁶, C Polman⁷, S Vosslamber⁸, CL Verweij⁸, NJ Olsen⁹ and TM Aune^{3,10}

¹Research Department, ArthroChip, LLC, Franklin, TN, USA; ²Department of Mathematics, Vanderbilt University, Nashville, TN, USA; ³Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN, USA; ⁴Department of Neurology, Vanderbilt University School of Medicine, Nashville, TN, USA; ⁵Department of Neurology, University of South Carolina, Charleston, SC, USA; ⁶Department of Neurology, University of Texas Southwestern Medical Center, Dallas, TX, USA; ⁷Department of Neurology, VU University Medical Center, Amsterdam, The Netherlands; ⁸Department of Pathology, VU University Medical Center, Amsterdam, The Netherlands; ⁹Department of Medicine, Penn State Hershey Medical Center, Hershey, PA, USA and ¹⁰Department of Microbiology and Immunology, Vanderbilt University School of Medicine, Nashville, TN, USA

Identification of biomarkers contributing to disease diagnosis, classification or prognosis could be of considerable utility. For example, primary methods to diagnose multiple sclerosis (MS) include magnetic resonance imaging and detection of immunological abnormalities in cerebrospinal fluid. We determined whether gene-expression differences in blood discriminated MS subjects from comparator groups, and identified panels of ratios that performed with varying degrees of accuracy depending upon complexity of comparator groups. High levels of overall accuracy were achieved by comparing MS with homogeneous comparator groups. Overall accuracy was compromised when MS was compared with a heterogeneous comparator group. Results, validated in independent cohorts, indicate that gene-expression differences in blood accurately exclude or include a diagnosis of MS and suggest that these approaches may provide clinically useful prediction of MS.

Genes and Immunity (2012) 13, 146–154; doi:10.1038/gene.2011.66; published online 22 September 2011

Keywords: genomics; multiple sclerosis; diagnosis

Introduction

A cornerstone in diagnosing clinically definite multiple sclerosis (MS) is magnetic resonance imaging detection of brain lesions disseminated in time and space.^{1–3} Laboratory findings include cerebrospinal fluid immunological abnormalities.^{4,5} Criteria, including 2001 McDonald and revised 2005 McDonald classifications, are employed to standardize diagnosis of MS. However, few of the most common clinical features of MS are unique to this disease. A significant limitation in these approaches is the need for dissemination in time requiring detection of new lesions in follow-up scans, thus, potentially delaying diagnosis and onset of therapies that may retard disease progression and onset of disability.

In a search for optimal diagnostics to permit effective triage of patients into specialty care, we have considered that exclusionary tests may have great utility. As MS is relatively rare, the possibility of excluding from further consideration those individuals who do not carry specific

disease markers would greatly reduce the numbers requiring further evaluation by specialists. An exclusionary test would also expedite referral to a specialist of those individuals that may require further evaluation and serve to reassure both patients and providers for those who do not. To be useful, an exclusionary test must have a high degree of accuracy so that individuals in need of further evaluation do not escape into the excluded pool. Patients who fall in the non-excluded pool would not have a specific diagnostic label attached, but would have a high likelihood that an autoimmune condition of some kind is present. At this point, the referring physician could reasonably make a decision regarding which condition is most likely and refer the patient to the appropriate specialist. By excluding patients who need no further evaluation, the medical system would deliver care efficiently to those who do. Optimization of health care delivery to improve outcomes is a major focus of the health care reform movement, and paradigm-shifting exclusionary diagnostics would contribute to this goal.

With the advent of array-based technologies, the possibility that large-scale screening of DNA variants, differences in RNA expression, or differences in protein expression either at affected tissue sites or common sources (blood, plasma) could provide clinically useful information has attracted much interest. In general, it appears that analyses of DNA variants, identified thus

Correspondence: Professor TM Aune, Department of Medicine, Vanderbilt University School of Medicine, MCN T3219, 1161 21st Avenue South, Nashville, TN 37232-2681, USA.

E-mail: tom.aune@vanderbilt.edu

Received 13 June 2011; revised 10 August 2011; accepted 15 August 2011; published online 22 September 2011

far, either singly or combined, are limited in their ability to provide clinically useful prediction of disease.⁶ In contrast, studies have demonstrated the potential utility of RNA or protein-expression profiles to segregate subjects with a given disease from either healthy control subjects or subjects with other diseases.^{7–18} Using this approach, we previously identified a panel of genes whose expression levels varied among different autoimmune diseases.^{19–21} We hypothesized that expression profiles may provide a method to aid in diagnosis of autoimmune diseases, such as MS, and have performed such analyses. Limitations in our previous studies included relatively small numbers of genes included in the analysis thus reducing discriminatory power, small study size, lack of geographical heterogeneity, and lack of sufficient subjects with other inflammatory and non-inflammatory neurological diseases and disorders in the study cohort.²² These limitations have been addressed in the current study to include a significantly larger number of genes in the analysis, extended methods of analyses and significantly larger sample cohorts drawn from various US and European sites.

Results

Gene-expression patterns in distinct neurological diseases

We measured expression patterns of a common set of genes assayed using a common platform in control subjects and subjects with different neurological conditions, including autoimmune diseases. Genes for analysis were selected from prior microarray studies. Expression levels of individual genes were determined by quantitative reverse-transcription PCR by normalization to GAPDH-expression levels. We employed a heatmap to depict those genes differentially expressed in individual disease cohorts relative to the control cohort, $P < 0.05$ (after Bonferroni's correction for multiple testing) (Figure 1, red = overexpressed gene, green = underexpressed gene). Ratios of expression levels of individual genes in the indicated disease cohort relative to the control cohort were calculated and depicted within each colored box. Each disease exhibited an underlying unique pattern of gene expression. However, these profiles were sufficiently overlapping to prohibit accurate discrimination of one disease from another disease using the expression profile alone. For example, *LLGL2*, *RANGAP1*, *ACTB* and *POU6F1* were underexpressed in four, three, four and four of five different conditions, respectively. In contrast, other genes, for example, *ANAPC1* in Parkinson's disease, *EXT2* and *FOS* in transverse myelitis (TM),

HRAS in neuromyelitis optica (NMO), were only differentially expressed in a single disease cohort. Overall, individual genes were either overexpressed, for example, *B2M*, *CD55*, *PMAIP1*, or underexpressed, for example, *LLGL2*, *RANGAP1*, *ACTB*, across multiple disease cohorts. Thus, each gene was differentially expressed in at least one disease cohort relative to the CTRL cohort. However, each individual disease cohort did not possess a unique expression profile distinguishing it from all other disease cohorts.

Discrimination of MS from homogeneous comparator groups: identification of an optimum panel of gene-expression ratios. Healthy control subjects, subjects with MS, subjects with other inflammatory neurological disorders (OND-I) and subjects with neurological disorders typically considered non-inflammatory (OND-NI) were recruited from multiple US and European sites (Table 1 and Supplementary Table 1). Demographical characteristics of the different disease groups, MS, OND-I or OND-NI were matched to the CTRL cohort (Table 2). Subjects with MS included subjects with clinically isolated syndrome (CIS), newly diagnosed MS subjects who were treatment naïve and subjects with established disease (>1 -year duration) on different therapies. Expression levels of test and control genes in blood were determined by quantitative reverse-transcription PCR; Supplementary Table 2). We employed a search algorithm to identify those ratios of gene-expression levels in which the greatest number of subjects in the test group possessed a ratio value greater than the highest ratio value in the comparator group. We employed a second algorithm to perform permutation testing of one subject group to identify the optimum set of discriminatory ratios. We reasoned that examination of expression levels of ratios of genes rather than individual genes would serve the following purposes. First, calculation of ratios normalized for differences in mRNA or cDNA template quantity and quality among different samples. Second, they obviated the need for inclusion of a 'housekeeping' gene in the analysis, and the assumption that expression levels of 'housekeeping' genes did not vary among different subject populations. Third, comparisons of ratios or combinations of ratios may more accurately identify cellular phenotypes that may contribute to disease. For example, a ratio containing one gene in the numerator that is overexpressed in the test group relative to the comparator group and one gene in the denominator that is underexpressed in the test group relative to the comparator group should produce a greater ratio value difference between individuals in the two groups

	ABR	ACTB	ACTR1A	ANAPC1	APOBEC3F	B2M	CD55	CDH1	CHEK2	CSF3R	EPNK2	EXT2	FOS	GAPDH	GATA3	GNB5	HRAS	JUN	KRAS	LLGL2	NFAS	OAS1	PMAIP1	POU6F1	RANGAP1	SC65	SP1B	TGFBR2	TP53	TXK
MS		0.48	0.62			2.31			0.53	0.51	0.54			1.05						0.54	0.48	2.24	1.87	0.36		0.44				
NMO	0.28	0.43	0.65		0.38	2.04	1.87	0.41	0.61	0.51				1.01	0.55	0.56	0.42		0.51	0.53	0.52		1.86	0.41	0.57	0.47	0.39	0.54	0.37	
TM	0.55	0.64	0.62		0.55			0.43	0.55	0.69	0.55	0.69	0.67	0.44	1.02	0.57	0.53		0.41	0.61	0.64	0.48			0.63	0.56	0.61		0.41	0.51
Parkinsons				0.61							0.61													0.62			0.52			
Alzheimer's	0.19	0.29	0.38			1.79			0.52	0.46	0.27	0.31			0.97	0.39	0.27	0.28	0.53					0.44		0.43	0.51	0.30		

Figure 1 Gene-expression profiles across multiple autoimmune diseases. Expression levels of 44 target genes were determined by quantitative reverse-transcription PCR and normalized to expression of GAPDH. Expression levels of 31 genes are shown; expression levels of the remainder were not statistically different between CTRL and any disease cohort. Genes are identified that showed statistically significant ($P < 0.05$ after Bonferroni's correction) increased (green boxes) or decreased (red boxes) expression in individual disease cohorts relative to CTRL subjects. Numerical expression ratios (disease group average/CTRL average) are displayed within the colored boxes.

Table 1 Characteristics of subjects

	Sites		
	Nashville	US ^a	Europe ^b
Multiple sclerosis (total)	84	81	80
CIS	14		10
Treatment naïve	30	4	55
Established disease (on medicines)	40	77	15
OND-I (total)	1	85	
Acute disseminated encephalomyelitis		4	
Bell's palsy		3	
CNS lupus		2	
Guillaine barre		4	
Myasthenia gravis		3	
Neuromyelitis optica		27	
Optic neuritis	1	1	
Transverse myelitis		41	
OND-NI (total)	1	128	
Alzheimer's		6	
Cerebral ataxia		2	
Cerebral bleed		2	
Cervical radiculopathy		6	
Drug-induced movement disorder		1	
Dystonia		1	
Epilepsy	1	4	
Essential tremor		9	
Huntington's disease		1	
Hydrocephalus		1	
Median neuropathy		2	
Meningioma		1	
Migraine		30	
Parkinson's	3	0	
Peripheral neuropathy		1	
Pseudotumor		3	
Restless leg syndrome		1	
Seizures		6	
Spasmodic torticollis		1	
Stroke		18	
Tourette's syndrome		1	
Transient ischemia		1	
Controls	48	61	

Abbreviations: CIS, clinically isolated syndrome; OND-I, other neurologic disorders-inflammatory; OND-NI, other neurologic disorders-non-inflammatory.

^aSix additional sites in United States: MA, MD, NY, SC, AZ, TX, CA, samples from sites in MS, MD, NY, AZ and CA were obtained through the Accelerated Cure Project.

^bDenmark, Netherlands.

than a single-expression value. We employed a point system to award one point to a subject if a ratio value of the test subject was greater than the ratio values of all subjects in the comparator group (Supplementary Figure 1).

We applied this approach to determine how accurately it would distinguish subjects with MS from healthy control subjects. First, we identified ratios capable of discriminating MS subjects from control subjects. The single ratio with the greatest discriminatory power was ANAPC1/CHEK2 (Figure 2a). In all, 50% of MS subjects achieved a ratio value higher than all the CTRL subjects and were awarded one point. Second, we eliminated those ratios that were identified in <20% of MS subjects.

Table 2 Demographic characteristics of the different subject populations

	Age	P*	Gender (% F)	P	Ethnicity (% C/AA/As/H)	P
MS	43 ± 10	NS	76	NS	80/20/0/0	NS
OND-NI	46 ± 10	NS	67	NS	68/26/3/1	NS
OND-I	46 ± 10	NS	68	NS	67/33/0/0	NS
CTRL	41 ± 11		77		71/22/3/3	

Abbreviations: AA, African American; As, Asian; C, Caucasian; H, Hispanic; MS, multiple sclerosis; OND-I, other neurologic disorders-inflammatory; OND-NI, other neurologic disorders-non-inflammatory.

*P calculated by Student's *t*-test (age) or Fisher's exact test, NS: P > 0.05.

Third, as many ratios identified the same MS subjects, we performed another reduction to preserve only one ratio with this characteristic. A total of eight ratios remained after this minimization process (Figure 2b). Using the point system, the combination of these eight ratios positively identified 97% of MS subjects and eliminated 100% of CTRL subjects (Figure 2c). The score distribution was 0–6 for MS subjects and 0 for CTRL subjects (Figure 2d).

Discrimination of MS from homogeneous comparator groups: validation and analysis

Our analyses depended upon determination of multiple ratios, which may create Type 1 errors. Various methods are available to correct for false discovery rates. Rather than relying upon these methods, which all make underlying assumptions, we performed a second evaluation using an independent cohort of 40 new MS subjects and 40 new CTRL subjects to validate results obtained from the initial training set. These subjects were recruited separately and the PCR analyses were performed separately. We used the same ratio values defined from the original CTRL and MS test set to award points to subjects in the validation cohort. All 40 controls were awarded a score of 0, whereas 4% of MS subjects received a score of 0. The remaining 96% of MS subjects achieved a score of 1–6 and the distribution of scores was similar to that observed in the training set (Figure 2e). Taken together, this demonstrates that results obtained in the training set can be replicated in an independent cohort of CTRL and MS subjects.

We applied the point system to OND-I and OND-NI subjects. In contrast to CTRL subjects, 90% of OND-I and 59% of OND-NI subjects scored ≥1 (Figure 2f). We compared scores among subjects with CIS, with newly diagnosed MS not yet on medications, and with established MS on different medications. Scores did not differ significantly among these three groups (Figure 2g). We also compared scores within the MS group as a function of geographical origin. Scores also did not vary significantly among MS subjects from different geographical sites (Figure 2h). Thus, subjects with CIS or subjects after their initial diagnosis of MS had a similar mean score to subjects with established MS on therapies. However, a high percentage of subjects with other neurological conditions, especially inflammatory neurological

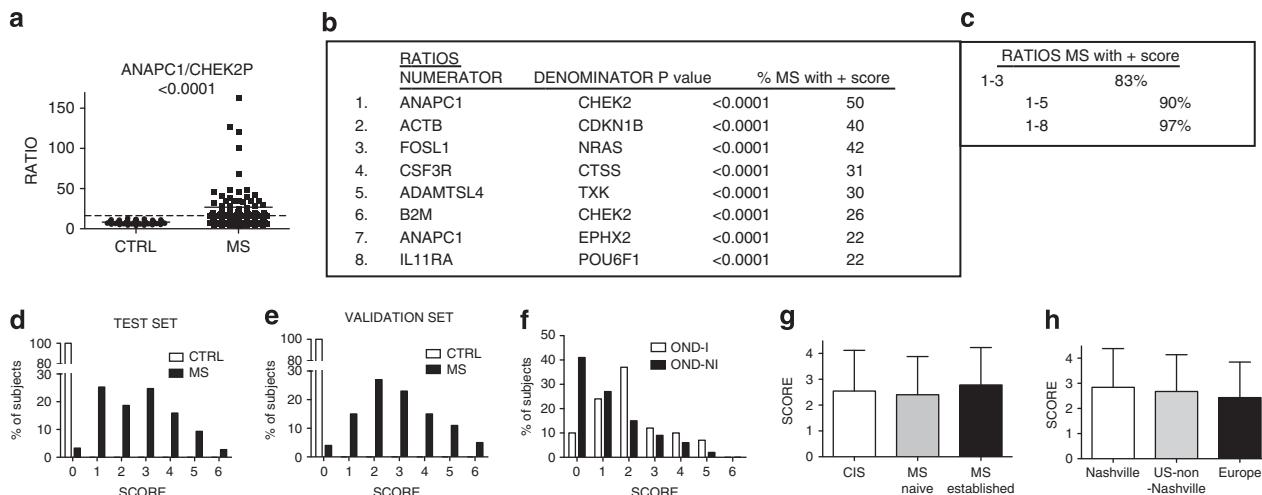


Figure 2 Discrimination between MS and CTRL subjects with an eight ratio scoring system. (a) Performance of the single ratio, ANAPC1/CHEK2 to discriminate MS and CTRL subjects. (b) Genes making up eight unique discriminatory ratios. *P*-values compare expression levels of ratios between MS and CTRL subjects. (c) Increased sensitivity with increasing numbers of ratios. (d) Score distributions among subjects using eight ratios. (e) Validation of results by analyzing 40 new MS subjects and 40 new CTRL subjects. (f) Score distribution between OND-I and OND-NI subjects. (g) Mean scores \pm s.d. among subjects with CIS, initial diagnosis of MS and established MS. *P*-value is not significant among groups. (h) Mean scores \pm s.d. among MS subjects from different geographical locations. *P*-value is not significant among groups.

conditions, also scored >0 in this analysis. Given its extremely high specificity and relatively low sensitivity, this test has greater application to exclude an individual from the diagnosis of MS rather than to establish a diagnosis of MS.

Further, we were able to obtain follow-up clinical information on eight CIS subjects >2 years after the initial consent and blood draw. Of these subjects, the seven CIS subjects who achieved a score >0 in the analysis now have documented MS. The one CIS subject who achieved a score of 0 does not have a documented case of MS.

NMO and TM are inflammatory neurological diseases that scored positive in our analysis. Therefore, we determined whether we could employ a similar approach to discriminate MS from TM and MS from NMO. We identified a series of ratios that, when combined using the point system, were able to discriminate TM from MS and NMO from MS with similar overall accuracy to the MS and CTRL comparisons (Figure 3). Thus, using our approach, it was possible to distinguish MS from TM and MS from NMO with a similar degree of accuracy as obtained for the comparison of MS to CTRL. However, as each disease possessed a unique signature, it was necessary to employ separate combinations of ratios to accurately distinguish MS from NMO and MS from TM.

Above results demonstrate that it is possible to distinguish MS from either a control cohort or even a related inflammatory disease cohort if the disease cohort is a single disease. Next, we asked whether we could discriminate MS from Parkinson's disease, a disorder typically considered non-inflammatory. To test this hypothesis we determined whether subjects with Parkinson's disease ($N=24$) segregated from MS ($N=182$) and CTRL ($N=109$) using the ratio and point system. We identified 10 ratios capable of discriminating 97% of MS subjects from 100% of Parkinson's subjects and 9 ratios capable of discriminating 88% of Parkinson's

	MS vs. TM	MS vs. NMO
1.	FOS/GAPDH	TP53/FOS
2.	TGFBR2/ASL	APOBEC3F/FOS
3.	TGFBR2/IL11RA	TP53/GAPDH
4.	APOBEC3F/ACTR1A	TP53/IL11RA
5.	CSF3R/ASL	CSF3R/GAPDH
6.	JUN/EXT2	ACTR1A/ASL
7.	TP53/CDKN1B	ABR/FOS
8.	JUN/GAPDH	APOBEC3F/LLGL2
9.	TP53/LLGL2	APOBEC3F/GAPDH
10.	JUN/ANAPC1	TBP/CD55

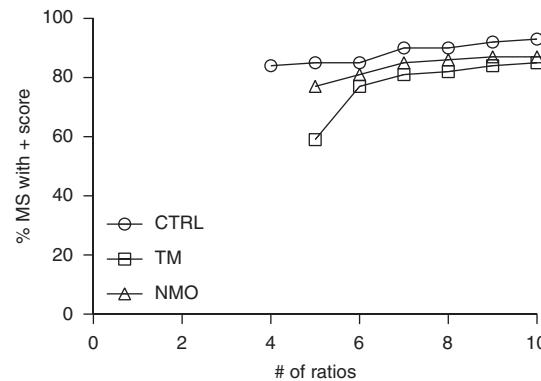


Figure 3 Discrimination of MS subjects from subjects with inflammatory neurological diseases, TM or NMO. Most discriminatory gene-expression ratios were identified that segregate MS subjects from TM and NMO subjects (CTRL is included for reference). The point system was applied to combine ratio performance into a single discriminator.

patients from 100% of CTRL subjects (Figure 4). We interpret these results to demonstrate that subjects with Parkinson's disease express unique gene-expression signatures in blood distinguishing them from CTRL and MS subjects.

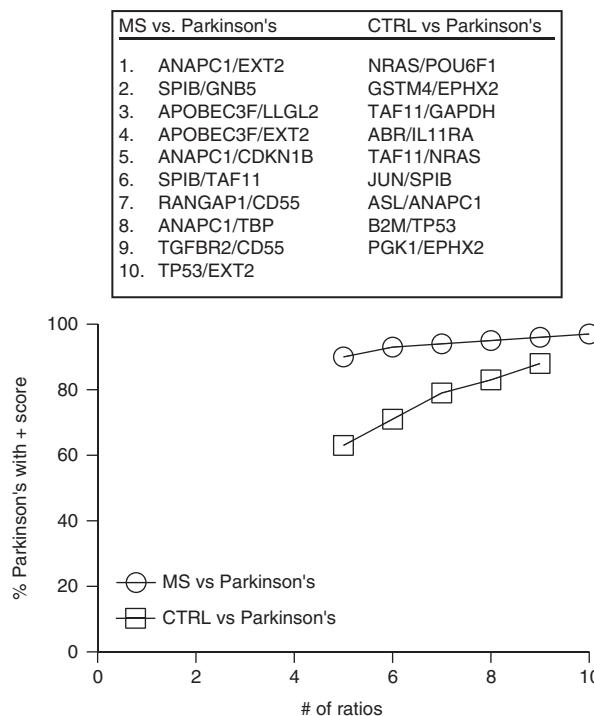


Figure 4 Discrimination of subjects with Parkinson's disease from MS and CTRL. Most discriminatory gene-expression ratios were identified that segregate Parkinson's disease subjects from MS subjects and CTRL subjects. Using the point system, we determined the % of Parkinson's subjects with a score >0 , *y* axis, relative to the number of ratios, *x* axis, for the different comparator groups.

Discrimination of MS from heterogeneous comparator groups
Next, we determined whether we could distinguish MS from more heterogeneous groups of subjects. To do so, we combined subjects with neurological conditions typically considered as inflammatory (OND-I in Supplementary Table S1) into one group. We also combined subjects with neurological conditions typically considered non-inflammatory (OND-NI, OND in Supplementary Table S1) into a second group. We produced a third group consisting of CTRL + OND-I + OND-NI subjects (ALL). We determined the 15 best ratios using permutation testing for each comparison. Overall, comparison of MS to these heterogeneous comparator groups resulted in a marked reduction in overall discrimination ability (Figure 5). We conclude that a binary comparison such as this exhibits much reduced accuracy as the heterogeneity of the comparator group is increased.

Discrimination of MS from OND-I: identification of optimum panels of gene-expression ratios

For additional analysis, we combined OND-I into one group of non-MS inflammatory neurological disorders and investigated the ability of our approach to discriminate this combination of diseases from MS. We relaxed conditions somewhat to identify ratios with the ability to detect zero or one non-MS subjects. Our best results were obtained with 10 ratios (Figure 6a). The combination of which identified 86% of MS subjects with a score >0 and only 8% of OND-I subjects with a score >0 (Figure 6b). Scores ranged from 0 to 7 for MS subjects and 0 to 1 for OND-I subjects (Figure 6c).

MS vs OND-I	MS vs. OND-NI	MS vs ALL
1. APOBEC3F/TAF11	APOBEC3F/CHEK2	APOBEC3F/TAF11
2. APOBEC3F/LLGL2	CSF3R/CDKN1B	GNB5/TAF11
3. RANGAP1/CDKN1B	APOBEC3F/POU6F1	APOBEC3F/LLGL2
4. APOBEC3F/ASL	APOBEC3F/TP53	RANGAP1/CDKN1B
5. ACTB/ASL	OAS1/LLGL2	OAS1/TAF11
6. APOBEC3F/IL11RA	CSF3R/CD55	APOBEC3F/GAPDH
7. APOBEC3F/TGFBR2	ACTB/TAF11	ACTR1A/TAF11
8. JUN/TAF11	APOBEC3F/TAF11	GSTM4/CTSS
9. APOBEC3F/HRAS	CSF3R/TP53	OAS1/GAPDH
10. IL11RA/CTSS	CDH1/NRAS	OAS1/BRCA1
11. SPIB/CD55	BRCA1/TAF11	CSF3R/CDKN1B
12. TBP/CTSS	GNB5/EXT2	ACTB/TAF11
13. APOBEC3F/NRAS	ANAPC1/EXT2	KRAS/TAF11
14. SPIB/CTSS	CSF3R/RANGAP1	APOBEC3F/CDKN1B
15. EPHX2/EXT2	RANGAP1/CDKN1B	APOBEC3F/ASL

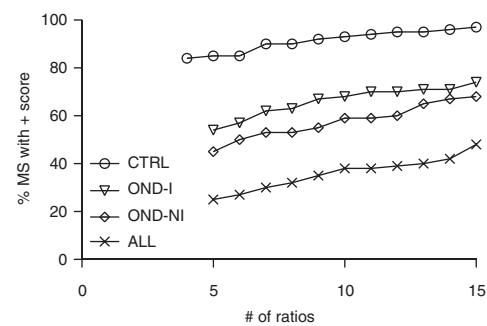


Figure 5 Discrimination of MS subjects from heterogeneous comparator groups. We identified the top 15 gene-expression ratios with the greatest ability to discriminate MS from OND-I, OND-NI or ALL (OND-I, OND-NI and CTRL). Using the point system, we determined the % of MS subjects with a score >0 , *y* axis, relative to the number of ratios, *x* axis, for the different comparator groups (CTRL is included for reference).

Discrimination of MS from OND-I: validation and analysis

We performed additional analyses with 40 new MS subjects and 40 new OND-I subjects (20 NMO and 20 TM) not included in the training set. In the validation set, 88% of MS subjects achieved a score ≥ 1 and 12% of OND-I subjects achieved a score of 1 (Figure 6d), which was similar to the score distribution observed in the training set. We determined mean scores among subjects with CIS, subjects with newly diagnosed MS before the onset of therapies and subjects with established MS on therapies using the 10 ratios identified above. Mean scores were significantly higher in the CIS and MS-naïve groups than in the MS group with established disease (Figure 6e). We also determined mean scores based upon geographical origins of MS subjects. Subjects from Nashville and Europe had mean scores significantly greater than the US subjects from locations other than Nashville (Figure 6f). These results are consistent with results comparing CIS, MS naïve and MS established. The majority of subjects from US sites outside Nashville had established MS and were on therapies (76 of 80 subjects), whereas all European subjects were either CIS or newly diagnosed MS subjects not yet on therapies ($N = 101$). The Nashville site also provided more samples with established disease ($N = 37$) compared with CIS or treatment-naïve MS ($N = 16$) ($P < 0.0001$, χ^2 -test for independence among three geographical locations). The distribution of scores in the CIS and newly diagnosed MS group was also higher than that found in the established MS group. In all, $>50\%$ of subjects with established MS achieved scores of 0 or 1 while 48% of CIS and newly diagnosed MS subjects achieved scores ≥ 3 (Figure 6g).

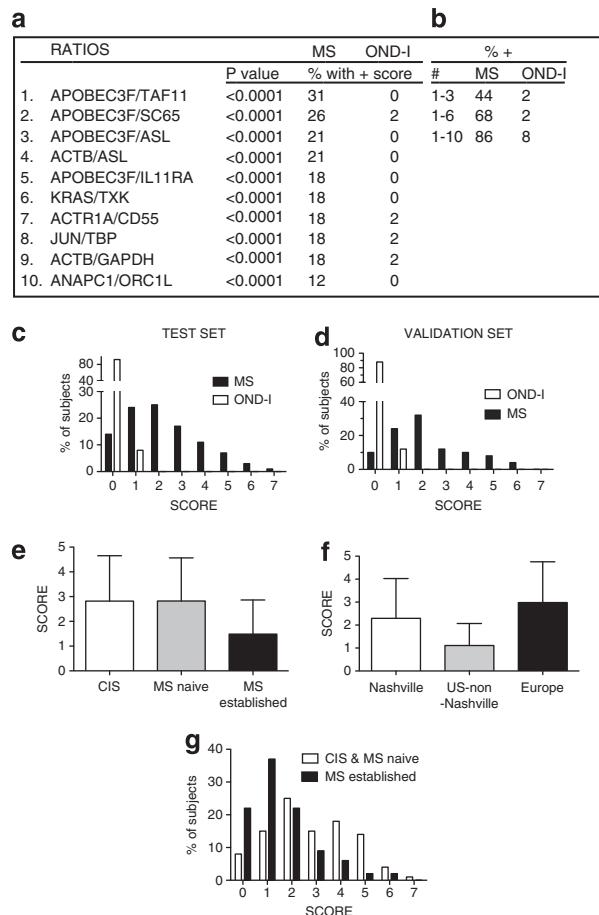


Figure 6 Discrimination between MS and OND-I subjects using 10 gene-expression ratios. (a) Genes making up 10 unique discriminatory ratios. *P*-values compare individual ratio values between MS and OND-I subjects. (b) Increasing number of ratios increases sensitivity or ability to discriminate between MS and OND-I subjects. (c) The score distribution in MS and OND-I subjects using 10 ratios. (d) Validation of results by analyzing 40 new MS subjects and 40 new OND-I subjects (20 TM + 20 NMO). (e) Mean scores \pm s.d. among subjects with CIS, initial diagnosis of MS and established MS. *P* is not significant for CIS versus MS naïve, 0.03 for CIS versus established MS and <0.0001 for MS naïve versus established MS. (f) Mean scores \pm s.d. among subjects based upon geographical sites. *P* is not significant for Nashville versus Europe, <0.0001 for Nashville versus US non-Nashville and <0.0001 for Europe versus US non-Nashville. (g) Score distributions between (CIS and MS naïve) and established MS.

Thus, subjects with CIS, newly diagnosed MS and established MS from different geographical sites can be distinguished from subjects with OND-I with reasonable accuracy based upon gene-expression profiles in whole blood.

Discrimination of MS from OND-NI: identification of optimum panels of gene-expression ratios

Next, we compared gene-expression differences between MS and OND-NI subjects, which included Parkinson's disease, essential tremors, migraines and strokes. We employed the same search strategy used to compare MS and OND-I subjects, and identified 10 expression ratios to construct the point system. ABOBEC3F, CSF3R and ANAPC1 were each in the numerators of two ratios and TAF11 was in the denominator of two ratios. Each ratio

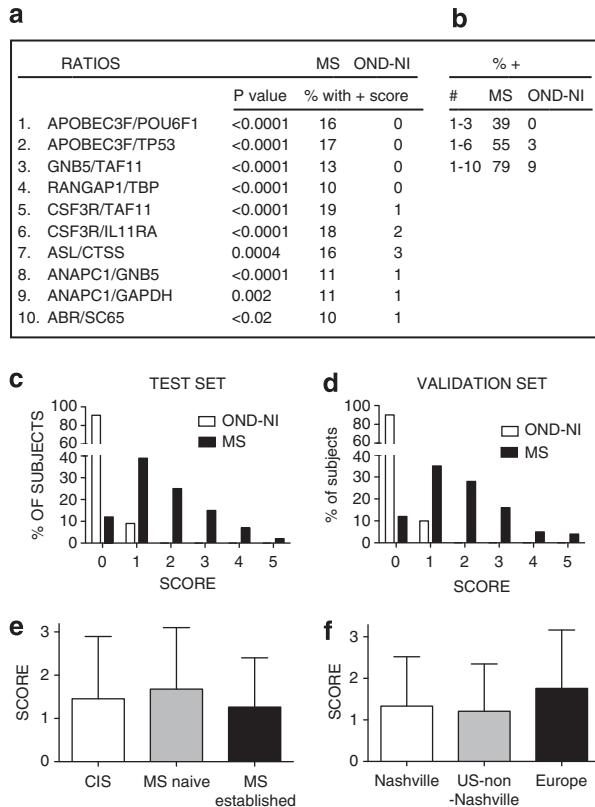


Figure 7 Discrimination between MS and OND-NI subjects using 10 gene-expression ratios. (a) Identification of genes making up the 10 unique discriminatory ratios. *P*-values compare individual ratio values between MS and OND-NI subjects. (b) Increasing the number of gene-expression ratio increases the ability to discriminate between MS and OND-NI subjects. (c) Score distribution using 10 ratios in the training set. (d) Validation of results by analyzing 40 new MS subjects and 40 new OND-NI subjects. (e) Mean scores \pm s.d. among subjects with CIS, initial diagnosis of MS and established MS. *P*-values were not significant among any of the comparisons. (f) Mean scores \pm s.d. among subjects based upon geographical sites. *P*-values were not significant for any of the comparisons.

alone detected $>10\%$ of MS subjects relative to OND-NI subjects (Figure 7a). Combining ratios using the point system improved overall ability to discriminate MS subjects from OND-NI subjects (Figure 7b). Using the point system, 79% of MS subjects achieved a score ≥ 1 and 91% of OND-NI subjects achieved a score of 0, 9% achieved a score of 1 (Figure 7c).

Discrimination of MS from OND-NI: validation and analysis

We performed additional analyses with 40 new MS subjects and 40 new OND-NI subjects not included in the training set as outlined above. In the validation set, 88% of MS subjects achieved a score ≥ 1 (Figure 7d), which was a similar frequency to that observed in the training set, and 90% of OND-NI subjects achieved a score of 0, 10% achieved a score of 1. As above, we determined mean scores of subjects with CIS, newly diagnosed MS and established MS, and these were not statistically different among the three MS groups (Figure 7e). Similarly, mean scores of MS subjects from different geographical sites were not statistically different (Figure 7f). Using the point system, $\sim 80\%$ of MS subjects achieved a score ≥ 1 and

9% of OND-NI subjects achieved a score = 1 in the test set. These results demonstrate that expression in whole blood of a different set of gene ratios discriminated subjects with MS from subjects with OND-NI with reasonable accuracy.

All comparisons in these analyses were binary. Therefore, exclusion of a specific disorder by the analysis may be more accurate than inclusion of a specific disorder (see flow chart, Supplementary Figure 2). Thus, a score of 0 in the MS versus CTRL test decreased the probability that a subject had MS. A second analysis comparing MS with OND-I and MS with OND-NI would be interpreted similarly. Scores of 0 decreased the probability of MS and favored the probability of OND-I or OND-NI, respectively. Finally, specific inflammatory neurological disorders, NMO or TM, were distinguished from MS with high degrees of accuracy. Thus, results from this single platform can be analyzed in a tiered approach to provide meaningful disease classification.

Discussion

Although our focus was on MS, and other inflammatory and non-inflammatory neurological disorders, our results support the notion that this approach could be applicable to an array of diseases. First, discrimination between MS and healthy controls or subjects with individual diseases can be achieved with a relatively high degree of accuracy. However, subjects with OND-I and OND-NI also scored positive in MS-CTRL comparisons. As such, this single comparison has greater utility as an exclusionary test rather than a test of MS inclusion. Second, it is possible to discriminate MS from groups of diseases, such as inflammatory or non-inflammatory neurological diseases, and validate results in independent cohorts, although overall accuracy is somewhat compromised. Third, discrimination of MS from a diverse comparator group including CTRL, OND-I and OND-NI causes a further reduction in overall accuracy. Nevertheless, a score > 0 in this analysis is highly predictive of the presence of MS. Fourth, it is possible to identify small numbers of ratios with high degrees of discriminatory power whose accuracy can be validated in independent cohorts analyzed separately.

One interpretation of our results is that many individual diseases express unique but overlapping gene-expression signatures in whole blood. Given the attention paid to analyses of autoimmune diseases, it is not surprising that inflammatory neurological diseases such as NMO and TM also express unique gene-expression signatures. Perhaps somewhat surprising is that Parkinson's disease, a disorder typically considered non-inflammatory, also possesses a unique gene-expression signature distinguishing it from both CTRL and MS. Implications may be that the immune system can sense specific neurological damage caused by Parkinson's via responses to cytokine mediators, adhesion molecules, neurotransmitters or other mediators read by immune cells. Alternatively, genetic risk factors associated with Parkinson's disease may contribute to altered gene-expression signatures by either direct or indirect mechanisms.

Mechanisms underlying gene-expression differences among study groups or relationships to MS disease

mechanism are not altogether clear. However, defects in DNA damage repair, cellular responses to DNA damage, and regulation of cell cycle progression and arrest are common properties of lymphocytes in certain autoimmune diseases, including MS, and *ANAPC1*, *CHEK2*, *CDKN1B*, *ACTB*, *FOSL1*, *LLGL2* and *NRAS* encode for proteins having key roles in these fundamental cellular processes.^{23–27} These genes are highly represented in the ratios used to distinguish MS from comparator groups. Genes, such as *ADAMTSL4*, *B2M*, *IL11RA*, *TXK* and *POU6F1*, encode for proteins having key functions in regulating cells of both innate and adaptive arms of the immune system.^{28,29} As such, alterations in expression of these genes may contribute to pathogenesis of MS or may represent an altered response by the immune system to MS pathogenesis.

Limitations to our study include selection of patients with pre-existing diagnoses of CIS and MS, as this may not completely represent patients in the general population in whom these tests may be performed. Our follow-up analysis of CIS patients supports the idea that initial scores > 0 will correlate with progression to MS. Future longitudinal studies are planned to better evaluate the utility of these tests in this setting. Further, our binary analysis is also predicated on the fact that MS is best represented by a single set of gene-expression ratios and this may not be the case. Additional analyses, such as analyses of gene-expression ratios in multidimensional space, will address this possibility. We identified several different combinations of gene-expression ratios, which performed equivalently in their ability to discriminate among subject groups. In conclusion, these minimally invasive and relatively inexpensive tests may have utility to either exclude the diagnosis of MS or to contribute to establishing a diagnosis of MS.

Patients and methods

Patients

Blood samples in PAXgene tubes were obtained from patients with a) CIS, b) an initial diagnosis of MS before the onset of therapy and c) established relapsing-remitting MS on medication. Blood samples were also obtained from healthy control subjects (CTRL) and subjects with different inflammatory (OND-I) or non-inflammatory (OND-NI) neurological conditions. MS samples were obtained from a total of nine different sites in the United States and Europe. Samples from subjects with OND-I and OND-NI were obtained from seven sites in the United States. CTRL samples were obtained from three US sites. Inclusion criteria for MS and other neurological conditions were diagnosis by a neurologist using established methods and ability to provide informed consent, thus providing an unbiased study cohort. Age, race and gender were not statistically different among the different study groups. Time of the blood draw, for example, morning/afternoon clinics, was also not statistically different among the different study groups. Relevant institutional review board approval from all participating sites was obtained.

Procedures

Total RNA, purified using Qiagen's isolation kits by standard protocols, was reverse-transcribed using

SuperScript III (Invitrogen, Carlsbad, CA, USA). A Taq-Man Low Density Array was designed to analyze expression levels of 44 genes previously identified from our microarray analysis and of four 'housekeeping' genes in 300 ng cDNA per sample. Patient diagnosis was blinded for all experimental procedures. Relative expression levels were determined directly from the observed threshold cycle (C_T), the cycle number at which fluorescence generated within reactions crosses an assigned threshold reflecting the point where sufficient amplicons have accumulated to be statistically significant above baseline. Linear expression values were determined using the formula, $2^{(40-C_T)}$.

Identification of discriminatory gene-expression ratios

A computational algorithm was designed to identify the most discriminatory combinations of ratios.²² All possible gene-expression ratios were computed (for example, *ACTR1A/BRCA1*, *TAF11/ACTR1A*, and so on). To analyze individual results, we used $R_{i,j}^{\text{control}}$ to denote the i th ratio for the j th control and let $R_{i,k}^{\text{MS}}$ denote the i th ratio for the k th MS patient. Here, $j = 1, \dots, N_{\text{control}}$ and $k = 1, \dots, N_{\text{MS}}$, where N_{control} equals the total number of controls and N_{MS} equals the total number of MS patients in the data set. The second largest member of each data set of ratios was calculated first by $\{R_{i,1}^{\text{control}}, R_{i,2}^{\text{control}}, \dots, R_{i,N_{\text{control}}}^{\text{control}}\}$, and designated $R_i^{(2)}$. This was then applied to the MS data set $\{R_{i,1}^{\text{MS}}, R_{i,2}^{\text{MS}}, \dots, R_{i,N_{\text{MS}}}^{\text{MS}}\}$. We used C_i to designate the number of MS set of ratios larger than $R_i^{(2)}$ such that $0 \leq C_i \leq N_{\text{MS}}$. This process was repeated for each possible ratio. The ratio that produced the largest C_i was selected as the discriminator of the two sets. This process was repeated using all possible ratios. Although more than one optimal ratio could be identified for each number of components queried, we have presented only one discriminator for each combination. Ratios were included only if >20% of subjects within the MS group had expression values greater than all subjects in the CTRL group. A scoring system was developed to combine multiple ratios. To do so, subjects were assigned one point for each ratio in which their expression value was higher than the highest expression value within the CTRL subject group. By this approach, it was also possible to relax search criteria by setting cutoffs to the second highest expression ratio, third highest expression ratio, and so on, of the comparator subject group. Using these relaxed criteria, an individual was awarded one point if the value of their expression ratio was higher than the second or third, and so on, highest expression value of individuals in the comparator group, respectively. These combined ratios established a score discriminating the MS group from comparator groups.

Search algorithm for best ratios

Let D denote the set of 44 gene-expression levels associated with the disease group and C denote the set of gene-expression levels associated with the control group. For example, when D is the set of MS patients, then D is a set of 182 44-tuples; if C is associated with the controls, then C is a set of 51 44-tuples. The algorithm that searches for the 'best' set of gene ratios is the following:

- 80% of the control group was randomly selected and compared with the disease group in the following manner. Gene-expression level ratios were formed for elements in D and C . For each ratio, the number of elements in the disease group that were larger than the largest ratio in the control group was computed. The top 500 ratios that separate elements in D and C were saved. This calculation was repeated 200 times resulting in a set of 200 subsets of ratios (each subset having 500 ratios).
- The 500 subsets were processed to identify the smallest number of ratio, $R = \{r_1, r_2, \dots, r_n\}$, that produced the maximum of separation of D and C . Associated with each of the ratios in R , there were threshold values, $T = \{t_1, t_2, \dots, t_n\}$, which corresponded to the highest value in the control group for each of the ratios in R .
- For each member of the disease group D , the ratios in R were computed, $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$. If $\alpha_i \geq t_i$, then we assigned the ratio a 1; otherwise, it was assigned a 0. In this way, we generated an n -tuple of 1's and 0's for each member of D . For example, if $n = 6$, then a typical 6-tuple would be {1,1,0,0,1,0}. This meant that this individual in the disease group would have three ratios that exceeded the corresponding ratios in the control group.
- Lastly, the percentage of members in the disease group that had nonzero n -tuples was calculated. The larger the percentage, the better the separation of D and C .

Statistical analysis

The Welch's-corrected *T*-test not assuming equal variances was used to calculate *P*-values in two-way comparisons. The χ^2 -test for independence was used to calculate *P*-values in three-way comparisons. The Bonferroni's method was employed to correct for multiple testing.³⁰

Conflict of interest

TMA and NJO are co-owners of ArthroChip.

Acknowledgements

This work was supported by the US National Institutes of Health Grant AI053984.

References

- 1 Swanton JK, Rovira A, Tintore M, Altmann DR, Barkhof F, Filippi M et al. MRI criteria for multiple sclerosis in patients presenting with clinically isolated syndromes: a multicentre retrospective study. *Lancet Neurol* 2007; **6**: 664–665.
- 2 Polman CH, Reingold SC, Edan G, Filippi M, Hartung HP, Kappos L et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the 'McDonald Criteria'. *Ann Neurol* 2005; **58**: 840–846.
- 3 McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the international panel on the diagnosis of multiple sclerosis. *Ann Neurol* 2001; **50**: 121–127.
- 4 Awad A, Hemmer B, Hartung HP, Kieseier B, Bennett JL, Stuve O. Analyses of cerebrospinal fluid in the diagnosis

- and monitoring of multiple sclerosis. *J Neuroimmunol* 2009; **219**: 1–7.
- 5 Link H, Huang Y-M. Oligoclonal bands in multiple sclerosis cerebrospinal fluid: an update on methodology and clinical usefulness. *J Neuroimmunol* 2006; **180**: 17–28.
- 6 Consortium TWTCC.. Genome-wide association study of 14 000 cases of seven common diseases and 3000 shared controls. *Nature* 2007; **447**: 661–678.
- 7 Axtell RC, de Jong BA, Boniface K, van der Voort LF, Bhat R, De Sarno P et al. T helper type 1 and 17 cells determine efficacy of interferon-beta in multiple sclerosis and experimental encephalomyelitis. *Nat Med* 2010; **16**: 406–412.
- 8 Keller A, Leidinger P, Lange J, Borries A, Schroers H, Scheffler M et al. Multiple sclerosis: microRNA expression profiles accurately differentiate patients with relapsing-remitting disease from healthy controls. *PLoS One* 2009; **13**: e7440.
- 9 Harris VK, Sadiq SA. Disease biomarkers in multiple sclerosis: potential for use in therapeutic decision making. *Mol Diagn Ther* 2009; **13**: 225–244.
- 10 Quintana FJ, Farez MF, Viglietta V, Iglesias AH, Merbl Y, Izquierdo G et al. Antigen microarrays identify unique serum autoantibody signatures in clinical and pathologic subtypes of multiple sclerosis. *Proc Natl Acad Sci USA* 2008; **105**: 18889–18894.
- 11 Kostka D, Spang R. Microarray based diagnosis profits from better documentation of gene expression signatures. *PLoS Comput Biol* 2008; **4**: e22.
- 12 Ray S, Britschgi M, Herbert C, Takeda-Uchimura Y, Boxer A, Blennow K et al. Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins. *Nat Med* 2007; **13**: 1359–1362.
- 13 Quackenbush J. Microarray analysis and tumor classification. *N Engl J Med* 2006; **354**: 2463–2472.
- 14 Hofman P. DNA Microarrays. *Nephron Physiol* 2005; **99**: 85–89.
- 15 Gregersen PK, Brehrens TW. Fine mapping the phenotype in autoimmune disease: the promise and pitfalls of DNA microarray technologies. *Genes Immun* 2003; **4**: 175–176.
- 16 Bompuzzi R, Ringnér M, Kim S, Bittner ML, Khan J, Chen Y et al. Gene expression profile in multiple sclerosis patients and healthy controls: identifying pathways relevant to disease. *Hum Mol Genet* 2003; **12**: 2191–2199.
- 17 Brynedal B, Khademi M, Wallström E, Hillert J, Olsson T, Duvefelt K. Gene expression profiling in multiple sclerosis: a disease of the central nervous system, but with relapses triggered in the periphery? *Neurobiol Dis* 2010; **37**: 613–621.
- 18 Harris VK, Sadiq SA. Disease biomarkers: potential for use in therapeutic decision making. *Mol Diagn Ther* 2009; **13**: 225–244.
- 19 Maas K, Chan S, Parker J, Slater A, Moore J, Olsen N et al. Cutting edge: molecular portrait of human autoimmune disease. *J Immunol* 2002; **169**: 5–9.
- 20 Liu Z, Maas K, Aune TM. Identification of gene expression signatures in autoimmune disease without the influence of familial resemblance. *Hum Mol Genet* 2006; **15**: 501–509.
- 21 Maas K, Chen H, Shyr Y, Olsen NJ, Aune T. Shared gene expression profiles in individuals with autoimmune disease and unaffected first-degree relatives of individuals with autoimmune disease. *Hum Mol Genet* 2005; **14**: 1305–1314.
- 22 Fossey SC, Vnencak-Jones CL, Olsen NJ, Sriram S, Garrison G, Deng X et al. Identification of molecular biomarkers for multiple sclerosis. *J Mol Diagn* 2007; **9**: 197–204.
- 23 Weyand CM, Fujii H, Shao L, Goronzy JJ. Rejuvenating the immune system in rheumatoid arthritis. *Nat Rev Rheumatol* 2009; **5**: 583–588.
- 24 Shao L, Fujii H, Colmegna I, Oishi H, Goronzy JJ, Weyand CM. Deficiency of the DNA repair enzyme ATM in rheumatoid arthritis. *J Exp Med* 2009; **206**: 1435–1449.
- 25 Deng X, Ljunggren-Rose A, Maas K, Sriram S. Defective ATM-p53-mediated apoptotic pathway in multiple sclerosis. *Ann Neurol* 2005; **58**: 577–584.
- 26 Maas K, Westfall M, Pietenpol J, Olsen NJ, Aune T. Reduced p53 in peripheral blood mononuclear cells from patients with rheumatoid arthritis is associated with loss of radiation-induced apoptosis. *Arthritis Rheum* 2005; **52**: 1047–1057.
- 27 Butowski N. Immunostimulants for malignant gliomas. *Neurosurg Clin N Am* 2010; **21**: 53–65.
- 28 Readinger JA, Mueller KL, Venegas AM, Horai R, Schwartzberg PL. Tec kinases regulate T-lymphocyte development and function: new insights into the roles of Itk and Rlk/Txk. *Immunol Rev* 2009; **228**: 93–114.
- 29 Buchner DA, Meisler MH. TSRC1, a widely expressed gene containing seven thrombospondin type I repeats. *Gene* 2003; **307**: 23–30.
- 30 Abdi H. *The Bonferroni and Sidak Corrections for Multiple Comparisons*. Sage: Thousand Oaks, CA, 2007.

Supplementary Information accompanies the paper on Genes and Immunity website (<http://www.nature.com/gene>)