Attention to time-of-day variability improves the reproducibility of gene expression patterns in multiple sclerosis

Huang, Suihong; Wu, Tan; Lau, Alexander Y.; Au, Cheryl; Huang, Hao; Wang, Xin; Kim, Jin Young

Published in:
iScience

Published: 19/11/2021

Document Version:
Final Published version, also known as Publisher’s PDF, Publisher’s Final version or Version of Record

License:
CC BY-NC-ND

Publication record in CityU Scholars:
Go to record

Published version (DOI):
10.1016/j.isci.2021.103247

Publication details:

Citing this paper
Please note that where the full-text provided on CityU Scholars is the Post-print version (also known as Accepted Author Manuscript, Peer-reviewed or Author Final version), it may differ from the Final Published version. When citing, ensure that you check and use the publisher's definitive version for pagination and other details.

General rights
Copyright for the publications made accessible via the CityU Scholars portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights. Users may not further distribute the material or use it for any profit-making activity or commercial gain.

Publisher permission
Permission for previously published items are in accordance with publisher's copyright policies sourced from the SHERPA RoMEO database. Links to full text versions (either Published or Post-print) are only available if corresponding publishers allow open access.

Take down policy
Contact lbscholars@cityu.edu.hk if you believe that this document breaches copyright and provide us with details. We will remove access to the work immediately and investigate your claim.
Attention to time-of-day variability improves the reproducibility of gene expression patterns in multiple sclerosis

Suihong Huang, Tan Wu, Alexander Y. Lau, Cheryl Au, Hao Huang, Xin Wang, Jin Young Kim

xin.wang@cityu.edu.hk (X.W.) jinykim@cityu.edu.hk (J.Y.K.)

Highlights
Times of day affect gene expression patterns in patients with RRMS in relapse
Transcriptome profiles in Relapse are changed from day to night
In Relapse, immune response-related genes change the expression at night

Attention to time-of-day variability improves the reproducibility of gene expression patterns in multiple sclerosis

Suihong Huang,1,5 Tan Wu,2,5 Alexander Y. Lau,3 Cheryl Au,3 Hao Huang,1 Xin Wang,2,4,* and Jin Young Kim2,4,6,*

SUMMARY
Low reproducibility in gene expression profiles has been observed in transcriptome studies, and this often limits applying findings to clinical practice. Here, we show time-of-day effects on gene expression and analytical schemes to increase the reproducibility in expression patterns. We recruited patients with relapsing-remitting multiple sclerosis (RRMS) and healthy subjects and collected blood from individuals twice a day, day (2 pm) and night (9 pm). RNA sequencing analyses found that gene expression in RRMS in relapse (Relapse) is significantly changed at night compared with either Relapse at day or RRMS in remission (Remission). Gene set overrepresentation analysis demonstrated that gene sets significantly changed in Relapse at night are enriched to immune responses related to MS pathology. In those gene sets, 68 genes are significantly changed expression in Relapse at night compared with Relapse at day and Remission. This supports that times of sample collections should be standardized to obtain reproducible gene expression patterns.

INTRODUCTION
Biomarkers are measurable indicators that diagnose disease onsets, courses, and activities as well as monitor responsiveness to treatments (Artomov, 2019; Mayeux, 2004). However, very few of them are clinically applicable because they have not been validated by independent studies or independent patient cohorts. Low reproducibility of biomarkers is caused by ambiguous experimental designs that do not address heterogeneity in patients’ phenotypic expression, a time gap from sample collection to processing, and times of sample collections. Among these criteria, especially times of sample collections have been less appreciated than other factors. A recent study found that more than half of known disease-associated genes and drug targets showed circadian rhythmic expression (Zhang et al., 2014). This suggests that many biomarkers and disease-associated genes are expressed at particular times of the day. Thus, collecting samples at the right time point of the day is important for reproducible results.

Gene expression studies in central nervous system (CNS) diseases also have low reproducibility problems in independent studies and patient cohorts (Chang and Kim, 2020). To improve this, we pay attention to time-of-day variability. However, an additional limitation has to be considered to study this in CNS diseases—relatively hard to obtain samples from patients. Although cerebrospinal fluid (CSF) or brain tissues have been used for gene expression studies, their accessibilities are less efficient. This means that taking samples multiple times a day from individual participants is not applicable. To overcome this problem, we started this study with patients with multiple sclerosis (MS). MS is a chronic autoimmune, demyelinating, and neurodegenerative disease in the CNS (Goldenberg, 2012). Since significant immunological activations are detected in the CNS and peripheral immune systems in MS (Palanichamy et al., 2014; Stern et al., 2014), blood has been used as specimens to study MS. MS is classified into three groups: relapsing-remitting MS (RRMS), secondary progressive MS (SPMS), and primary progressive MS (PPMS). Around 85% of patients with MS are initially diagnosed as RRMS characterized by attacks of inflammation against own myelin to induce demyelination (relapse) (Dyment et al., 2004; Ebers, 2008; Olsson et al., 2017). This is followed by partial or complete recovery periods (remission) (Goldenberg, 2012; Lublin et al., 2014). Hence, most proposed biomarkers for MS are immune response related, such as serum antibodies (e.g., anti-myelin oligodendrocyte glycoprotein and myelin basic protein antibodies) (Lalive et al., 2006; Reindl et al., 2013), and serum cytokines, chemokines, and their receptors (e.g., TNF-α and IFN-γ) (Katsavos
and Anagnostouli, 2013; Martins et al., 2011). This suggests that the status of peripheral immune cells in the patients can represent the disease phases.

The main purpose of this study is to provide an approach to obtain reproducible gene expression patterns by considering time-of-day variability. With the understanding characteristics of MS, we will focus on gene expression profiles in RRMS to show how times of day determine mRNA levels according to the disease phases. Also, experimental and analytical schemes will be suggested to increase reproducibility.

RESULTS

Time-of-day effects on gene expression profiles in RRMS

To test time-of-day effects on gene expression, we recruited patients with RRMS in either relapse or remission and healthy controls (HC) who had no history of MS or neurological diseases. Note that a study neurologist determined clinical courses, and all participants’ information is presented in Table S1. Among seven patients with RRMS, four were in remission (Remission), two were having acute relapses (Relapse), and one was in remission but recently had a relapse (Recent-relapse). Blood was collected from all participants twice a day, 2 pm (day) and 9 pm (night). To obtain the transcriptome profiles in blood, RNA sequencing analyses were conducted. RNA sequencing reads were aligned to the human reference genome after quality check, and the uniquely mapped reads were annotated, followed by gene expression calculation based on logCPM (log2 counts per million). We first compared global gene expression profiles between day and night in each group, HC, Remission, and Relapse. There was no obvious difference in HC and Remission, but Relapse showed differences between the two time points (Figure 1A). We then asked if global gene

Figure 1. Time-of-day effects on gene expression profiles in Relapse

Times of day affect gene expression profiles in patients with RRMS in relapse.
(A) Heatmap of transcriptomes: comparisons between day (2 pm) and night (9 pm) in healthy control (HC), remitting-RRMS (Remission), and relapsing-RRMS (Relapse), respectively. Red indicates high expression (log2 counts per million), and blue indicates low expression.
(B) Heatmap showing comparisons between the three groups at day and night, respectively.
(C) Principal component analysis (PCA) of the gene expression profiles of all patients with RRMS and healthy control. The top three principal components accounted for 65.8% of the total variation. HC1-3, Healthy control; MS1-4, Remission; MS5-6, Relapse.
See also Table S2.
expression patterns show any differences between groups at each time point. At day, Remission and Relapse showed a similar pattern, and this was distinguished from HC. However, at night, Relapse showed different patterns from both Remission and HC (Figure 1B). This shows that the gene expression profiles are changed in patients with RRMS compared with healthy subjects, but the differences are the most significant in Relapse at night.

We next performed principal component analysis (PCA) and observed that the gene expression patterns in Relapse at night (Relapse-N) are distinguished from all other groups (Figure 1C). As shown in the space of the first three principal components, Relapse-N was clearly separated from the other five groups—Relapse at day (Relapse-D), Remission at day (Remission-D) and night (Remission-N), and HC at day (HC-D) and night (HC-N). However, there was no clear separation between the five groups. Taken together, these analyses show that the gene expression profiles in RRMS are mostly changed in Relapse at night, and its pattern is uniquely separated from all other groups, including Relapse-D.

Comparison schemes to obtain reproducible gene expression patterns

Figure 1 presents two points. One, the gene expression in Relapse shows differences between day and night, which are not observed in HC and Remission. Two, Relapse shows different gene expression profiles with Remission at night, but not at the day. Based on these, we designed comparison schemes to obtain differentially expressed genes (DEGs) unique in Relapse-N, which are also related to MS pathology. The first step will be identifying DEGs between day and night in Relapse (Group 1 in Figure 2). Then, any genes

![Figure 2. Comparison schemes to obtain reproducible gene expression patterns](image-url)
also showing differences between day and night in HC or Remission will be removed to obtain unique DEGs in Relapse. Gene set overrepresentation analysis (GSOA) based on biological processes from the Gene Ontology (GO) database will be performed to obtain a landscape of biological functions. Among the GO terms, the processes related to MS pathology will be selected to obtain DEGs in those processes as group 1. The second step will identify DEGs between Relapse and Remission at night (group 2 in Figure 2). Although no clear differences were observed between the two groups at day (Figure 1B), any DEGs from the day will be excluded to obtain DEGs unique in Relapse-N. Then, DEGs will go through the same process with group 1 to screen MS pathology-associated DEGs in Relapse-N as group 2. Group 1 and group 2 will be compared with screen overlapping DEGs by the two groups. These DEGs have the following meanings. First, their expression in Relapse-N will show significant differences with Relapse-D, but this difference will not be observed in Remission and HC. Second, they can be specific markers for Relapse compared with Remission. Third, they can be indicators of MS pathology in RRMS.

DEGs between day and night in Relapse are associated with immune response and gene expression process

According to the scheme in Figure 2, we first identified DEGs between day and night in HC, Remission, and Relapse, respectively. Paired designs by edgeR package were employed to perform analyses in small samples of subjects. Note that analysis was performed on paired 2 pm and 9 pm from subjects, which removed potential baseline differences influenced by ages, genders, disease-modifying therapies, and other interindividual variations. At the log fold change cutoff of \( \pm 1 \) and false discovery rate (FDR) cutoff of <0.05, we identified two up-regulated DEGs in Remission but no DEG in HC. We expected to identify a few DEGs in HC because circadian genes in the blood might present diurnal expression patterns at 2 pm and 9 pm. However, no DEG was found, suggesting that 2 pm and 9 pm may not be the best time points to observe robust diurnal patterns of circadian genes in the blood from healthy subjects. This is supported by a study that examined the expression of ten circadian genes in blood mononuclear cells. The authors reported that the well-known circadian gene Per1/2/3 expression in healthy subjects was peaked early in the morning, around the transition from sleep to wakefulness (Kusanagi et al., 2008). Since our study aims to test time-of-day effects on gene expression rather than identifying circadian genes, we continued the analysis in Relapse. A total of 1,910 DEGs were identified, and 777 and 1,133 genes were significantly up- and down-regulated at night, respectively (Figures 3A and 3B; Tables S2 and S3). We then checked whether 1,910 DEGs in Relapse include 2 DEGs in Remission and found NID1 in both groups. With 1,909 DEGs (excluding NID1), GSOA was performed based on biological processes from the GO database. A total of 15 significantly enriched gene sets were identified, and 13 gene sets were enriched into two modules, immune responses and transcription/translation (Figures 3C and 3D). This supports that a large group of genes in Relapse change their expression between day (2 pm) and night (9 pm).

The heatmap based on logCPM in Figure 3B presents variant gene expression levels between two patients in Relapse. We interpreted this as interindividual variations on gene expression. This suggests that paired comparisons between day and night from subjects are more convincing and reproducible. To examine this, differences in the gene expression between day and night in individual Relapse were converted to fold changes, respectively. Then, the relationship of the fold changes from two individuals was visualized in a scatter plot in Figure 3E. The positive correlation (PCC = 0.66) indicates that two individuals show similar fold changes in the transcriptome, regardless of relative gene expression levels. Thus, DEGs identified from the paired comparisons will show fewer false-negative or false-positive results and exclude potential biases influenced by gender, age, and other factors, thereby producing more convincing and reproducible results.

DEGs between Relapse and Remission at night are associated with immune response and gene expression process

We next identified Relapse-unique DEGs in RRMS at different times of the day. At day, three up-regulated and one down-regulated DEGs were identified in Relapse compared with Remission, meaning similar gene expression profiles between the two groups. As expected, more DEGs were identified at night: among a total of 1,337 DEGs, 830 and 507 genes were significantly up- and down-regulated in Relapse, respectively (Figures 4A and 4B; Tables S2 and S4). Note that no DEG was overlapped between day and night. To obtain a landscape of biological functions in Relapse-N, GSOA was performed with the same parameters as Figure 3, and an enrichment map was visualized in Figure 4C. A total of 21 functionally related gene sets were clustered into two modules, immune responses and transcription/translation. Among them, 13 were related to immune responses, including innate immune response and inflammatory response (Figure 4D).
Taken together, a large group of genes associated with immune responses significantly change their expression in Relapse at night.

**MS pathology-associated DEGs in Relapse at night**

We next screened MS pathology-associated DEGs from groups 1 and 2 (Figure 2). In Figures 3 and 4, the gene sets were clustered into the same modules, immune responses and transcription/translation.
Transcription and translation regulate gene expression processes in most cells and thereby are less specific for the disease-associated processes. MS is an autoimmune disease; thus, activated immune responses in patients with MS are related to the pathophysiological processes, such as inflammation, demyelination, and axonal damage (Hemmer et al., 2015; Mayo et al., 2012). The innate immune cells, such as microglia in the CNS and infiltrating macrophages from the peripheral immune system, directly induce neuroinflammation to damage myelin and axons (Hirose et al., 2020; Mayo et al., 2012). Thus, we focused on the immune response module for further analysis. A total of 176 and 191 DEGs in the module from group 1 (Figure S1) and group 2 (Figure S2), respectively, were compared, and 68 DEGs were found in both groups (Figure 5A). Among them, 39 and 29 DEGs were significantly up- and down-regulated in Relapse-N compared with Relapse-D and Remission-D/N (Figure 5B). This means that the 68 DEGs satisfy the following: (1) Their expression in Relapse shows significant differences between day and night but not in Remission and HC. Thus, the 68 DEGs can be subjected for paired comparisons between day and night from subjects to minimize potential interindividual variations. (2) Their expression is changed in Relapse compared with Remission at night, meaning they are indicators of the disease phases. (3) Altered expression of the 68 DEGs in Relapse is associated with MS pathology.
Four up-regulated DEGs in Figure 5B, IL18R1, TPST1, TLR2, and CXCR4, were selected to confirm whether our analysis results are consistent with experimental data by quantitative RT-PCR. Note that four genes were selected based on the expression levels calculated by the logCPM values and fold changes between the groups. We confirmed that mRNA levels of the four DEGs were significantly increased in Relapse-N compared with Relapse-D as well as all Remission and HC groups (Figure 5C). Besides, each gene’s fold changes by paired comparison between day and night from patients in Relapse (MS5 and MS6) were similar, regardless of the relative expression levels. For example, the relative mRNA level of TLR2 in MS6 was around 2-fold higher than MS5 at day and night, respectively. However, its fold change between day and night was 11-fold in both
patients (Figure 5C). This supports that DEGs identified by the comparison scheme in this study produce reproducible results, regardless of comparison styles, either paired or not.

Another patient with RRMS was in remission but experienced a relapse 8 weeks before participating in this study (Recent-relapse). We compared the logCPM values of the 68 DEGs in this subject with other groups (Figure 5B). At night, most 68 DEGs in Recent-relapse showed similar patterns with Remission. However, several genes, such as IL18R1, TPS1T1, IL1R1, and IL18RAP, remained at similar levels with Relapse-N (Figure 5B). We interpret this pattern as Recent-relapse is in a transition state from Relapse to Remission in gene expression. This suggests that the identified DEGs in this study change their expression across the RRMS phases.

**DISCUSSION**

This study was started with the concept of circadian rhythms in a large group of gene expression (Chowdhury et al., 2020). Thus, we hypothesized that samples should be collected at the right time points when genes are expressed, and times of sample collection should be consistent to obtain reproducible patterns. To test this, we recruited patients with RRMS and designed experimental and analytical schemes to identify genes, of which expression is specifically and significantly changed only in Relapse at a particular time of the day. This gives advantages that identified genes could be differentially expressed between different disease phases or on paired times of sample collections from subjects. Our approach identified 68 DEGs that show the following: (1) The expression between day and night is changed only in Relapse. (2) The expression in Relapse shows significant differences with Remission at night, but not at day. (3) They are related to MS pathology. Thus, altered expression levels of the identified DEGs at night represent RRMS patients’ phase as Relapse. Besides, the identified DEGs will show reproducible expression patterns in independent studies and patient cohorts if times of sample collections are consistent. In addition to this, since the identified DEGs can be subjected to paired comparisons between different samples from individuals, this will minimize interindividual variations.

We identified 68 immune response-associated DEGs in Relapse and then asked whether they include previously proposed MS biomarkers. Previous studies proposed biomarkers in peripheral blood serum of patients with MS, such as TNF-α and IFN-γ (Katsavos and Anagnostouli, 2013; Martins et al., 2011). We found only a few of them in Figure 5B, such as TNF and several cytokine-associated genes (e.g., TNFSF8 and IFI16). However, many previous studies did not clarify times of sample collections. So, if they collected samples at different time points with this study, such as early morning, most biomarkers might be excluded from the first step of our approach. This suggests that a standardized time of sample collection is important to obtain reproducible results.

We focus on time-of-day variability as a factor to increase reproducibility in gene expression. For this, gene expression was examined in whole blood collected at different times of the day. In the MS field, recent studies systematically analyzed gene expression profiles in blood, CSF, and isolated blood and CSF immune cells (Chen et al., 2021; Gresle et al., 2020; Ramesh et al., 2020). A recent study by Gresle et al. investigated MS risk single-nucleotide polymorphisms (SNPs), which regulate expression of the MS risk allele, in isolated immune cells from blood collected between 7:30 am and 11 am. This study minimized time-of-day effects and studied cell type variability in genotype-phenotype interaction. The authors found a small number of cell type-specific SNPs in five immune cell types, determining individual susceptibility to MS (Gresle et al., 2020). Combining our findings, if time-of-day variability is examined in specific cell types or cell type variability is examined at various times of the day, this will help identify genes that produce more robust, specific, and reproducible results. Another study by Ramesh et al. also showed similar results. They performed systematic gene expression analysis on paired CSF and blood from subjects with RRMS. They identified DEGs by paired comparisons between CSF and blood or between isolated CSF and blood immune cells from subjects, meaning sample type variability in gene expression. Besides, in patients with RRMS, the CSF has a remarkably different cellular profile from the blood (Ramesh et al., 2020). Taken all together, gene expression profiles vary based on examining sample/cell types and times of sample collections. Thus, considering these factors together will increase specificity and reproducibility in gene expression even more.

We observed significantly changed gene expression profiles in Relapse at night than during the day. Besides, differentially expressed genes in Relapse at night were associated with immune responses that
are related to MS pathology. Although further studies are needed to explain why gene expression is more changed at night, one possible explanation is that genes, of which products are involved in immune responses, gain circadian rhythmicity in Relapse, and circadian systems increase their expression at night. The gaining circadian rhythmicity in disease-implicated pathways/genes was reported in pathological conditions of the brain (Huang et al., 2020, 2021; Seney et al., 2019). Time-of-death analysis in postmortem samples from patients with schizophrenia revealed that a group of genes gained circadian rhythmicity in the patients, which did not oscillate in control subjects. The pathways related to these genes were oxidative phosphorylation and mitochondrial dysfunction, which are implicated in schizophrenia. Furthermore, the authors observed differential expression of these genes only in subjects who died during the night but not in subjects who died during the day (Seney et al., 2019). This suggests that they are not rhythmic in control subjects but become rhythmic in schizophrenia, resulting in induced expression at night.

We here present preliminary evidence that times of day affect gene expression patterns with small samples of subjects—three healthy subjects and seven patients with RRMS (four in remission, two in relapse, and one in recent relapse). To screen MS biomarkers with this approach, further studies can apply the proposed scheme to larger cohorts of patients with RRMS or other MS types. It will give the following advantages: (1) It will identify reproducible and convincing biomarker candidates for MS with higher validation rates in independent patient cohorts. (2) Applying this approach to long-term studies with patients with RRMS will help track the disease phases and understand how RRMS is converted to SPMS. Finally, our approach can also be applied to other neurological diseases in the CNS or human diseases to increase efficiencies of biomarker screenings.

Limitations of the study
We chose two time points to collect blood, 2 pm and 9 pm. Collecting blood multiple times a day from individual participants is one of the critical points in this study. Paired comparisons between times of sample collection from subjects minimize interindividual variations and any effects from age, gender, and other variations. However, there were limitations to collect blood multiple times from an individual. First, the clinical research ethics approved by the institute limited the number of blood collections/day from a subject to minimize stress. Second, since most patients with RRMS were not hospitalized, all participants needed to visit the hospital multiple times a day for blood collection. Thus, we first chose two time points of the day, one during the day and the other at night, to test any differences in gene expression profiles. Since we here show the potential of our approach, further examinations at different times of day will be worth testing.

STAR METHODS
Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Material availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Sample collection and RNA isolation
  - RNA-sequencing analysis
  - Differential gene expression analysis
  - Functional analysis
  - Quantitative RT-PCR
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION
Supplemental information can be found online at https://doi.org/10.1016/j.isci.2021.103247.

ACKNOWLEDGMENTS
We thank Drs. Patrizia Casaccia and Ye He (CUNY Advanced Science Research Center) and Drs. Liang Zhang and Bruce Ransom (City University of Hong Kong) for critical comments on the manuscript and
helpful discussion. This study was partially supported by grants from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 21106716, 11103418, and 11101017 to J.Y.K.), a grant from Shenzhen Science and Technology Fund—Basic Research Subject Layout Program (Project No. JCYJ2017081803115939 to J.Y.K.), the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 21101115, 11102317, 11103718, 11103619, R4017-18, C4041-17GF to X.W.), a grant from Guangdong Basic and Applied Basic Research Foundation (Project No. 2019B030302012 to X.W.), a grant supported by the Young Scientists Fund of the National Natural Science Foundation of China (Project No. 81802384 to X.W.), and the Croucher Foundation, Lui Che Woo Institute of Innovative Medicine, and Gerald Choa Neuroscience Centre.

AUTHOR CONTRIBUTIONS

The majority of the experiments in the article was conducted by S.H. The majority of the analyses in the article was performed by T.W. A.Y.L. determined the clinical courses of all participants. A.Y.L. and C.A. collected blood from all participants. H.H. analyzed a part of RNA sequencing data. X.W. designed and supervised the analyses. J.Y.K. designed and supervised the project and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 19, 2021
Revised: August 30, 2021
Accepted: October 6, 2021
Published: November 19, 2021

REFERENCES


STAR METHODS

KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biological samples</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human whole blood samples</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>Critical commercial assays</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAXgene Blood RNA kit</td>
<td>QIAGEN</td>
<td>Cat# 762174</td>
</tr>
<tr>
<td>PrimeScript RT Master Mix</td>
<td>Takara</td>
<td>Cat# RR036</td>
</tr>
<tr>
<td>SYBR Green PCR master mix</td>
<td>Takara</td>
<td>Cat# RR420</td>
</tr>
<tr>
<td>Deposited data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw and analyzed data</td>
<td>This paper</td>
<td>GEO: GSE150337</td>
</tr>
<tr>
<td>Human reference genome (hg38)</td>
<td>UCSC</td>
<td><a href="https://genome.ucsc.edu/">https://genome.ucsc.edu/</a> RRID:SCR_005780</td>
</tr>
<tr>
<td>Oligonucleotides</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primers for quantitative RT-PCR</td>
<td>This paper</td>
<td>STAR Methods</td>
</tr>
<tr>
<td>Software and algorithms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FastQC (version 0.11.8)</td>
<td>Andrew, S. 2010</td>
<td>RRID:SCR_014583</td>
</tr>
<tr>
<td>‘edgeR’ package</td>
<td>Robinson et al., 2010</td>
<td>RRID:SCR_012802</td>
</tr>
<tr>
<td>EnrichmentMap (v.3.2.1)</td>
<td>Cytoscape</td>
<td><a href="http://apps.cytoscape.org/apps/enrichmentmap">http://apps.cytoscape.org/apps/enrichmentmap</a> RRID:SCR_003032</td>
</tr>
<tr>
<td>STAR (version 2.7.1a)</td>
<td>Dobin et al., 2013</td>
<td><a href="http://code.google.com/p/ma-star/">http://code.google.com/p/ma-star/</a> RRID:SCR_015899</td>
</tr>
<tr>
<td>Gene Ontology (GO) database</td>
<td>Ashburner et al., 2000</td>
<td>Gene Ontology Consortium, RRID:SCR_017505</td>
</tr>
<tr>
<td>HTSanalyzeR package</td>
<td>Wang et al., 2011</td>
<td><a href="http://www.bioconductor.org">www.bioconductor.org</a></td>
</tr>
<tr>
<td>QuantStudioTM 3 Real-Time Polymerase Chain Reaction System</td>
<td>Applied Biosystems</td>
<td>RRID:SCR_018712</td>
</tr>
<tr>
<td>Prism8</td>
<td>GraphPad Software</td>
<td><a href="http://www.graphpad.com/">http://www.graphpad.com/</a> RRID:SCR_002798</td>
</tr>
</tbody>
</table>

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Jin Young Kim (jinykim@cityu.edu.hk).

Material availability
This study did not generate new unique reagents.

Data and code availability
- RNA-sequencing data have been deposited at GEO and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- This paper does not report original codes.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Seven RRMS patients and three healthy controls (HC) who had no history of neurological or psychiatric diseases were recruited in this study: a total of 10 unrelated Asian individuals of either sex participated. HC’s and patients’ information is summarized in Table S1. Multiple sclerosis was diagnosed by the updated McDonald criteria, and the clinical course of individual RRMS patients was classified by the study neurologist (AL) together with MRI features. RRMS patients were either receiving disease-modulating therapeutics
or treatment naïve (Table S1). The study was approved by City University of Hong Kong – the Human Subjects Ethics Sub-Committee (Ref. No. 3-5-201601_02) and The Joint Chinese University of Hong Kong – New Territories East Cluster Clinical Research Ethics Committee (CREC Ref. No. 2016.304). All patients gave written informed consent. Prior to the blood donation, all participants were instructed to follow regular daily schedules for at least three days—three meals at similar times of the day and slept at least 7 hours at night. All samples were collected at Prince of Wales Hospital, Chinese University of Hong Kong, between December 2016 and February 2017 to minimize the seasonal effects.

METHOD DETAILS

Sample collection and RNA isolation

5 ml of whole blood was collected in the PAXgene Blood RNA tubes (QIAGEN) and kept at room temperature until the next day. In the morning, total RNAs were extracted from whole blood using the PAXgene Blood RNA Kit (QIAGEN) following the manufacturer’s protocol. Briefly, the PAXgene Blood RNA tubes containing blood were centrifuged at 2,600 × g for 20 min, and the supernatants were discarded. The pellets were re-suspended in 4 ml of RNase-free water by vortex until they were completely dissolved. After centrifugation at 2,600 × g for 20 min, the supernatants were discarded, and the pellets were re-suspended in 350 µl of the resuspension buffer by vortex. After incubation with proteinase K, the lysate was passed through the PAXgene Shredder spin column. The flow-through was mixed with ethanol and then passed through the PAXgene RNA spin column. The column was washed and incubated with DNase I to remove DNA. After washing, RNA was eluted for further experiments.

RNA-sequencing analysis

cDNA libraries were prepared from total RNA and sequenced on the BGIsq500 (Illumina) platforms (Beijing Genomics Institute, China). Raw sequencing reads were checked for quality using FastQC (Andrew, S. 2010; version 0.11.8), and subsequently mapped to the human reference genome (hg38) using STAR (Dobin et al., 2013) (version 2.5.3a) with default parameters. Uniquely mapped reads were retained for quantification of gene expression by read counts, calculated using ‘quantMode TranscriptomeSAM GeneCounts’ option in STAR. Gene counts were further converted to log2 count-per-million (logCPM) values for hierarchical clustering analysis and principal component analysis (PCA).

Differential gene expression analysis

Differential gene expression analysis was performed using ‘edgeR’ package (Robinson et al., 2010). More specifically, genes that are lowly expressed across all samples were excluded by the function ‘filterByExpr’ with default parameters. Gene counts were normalized using the trimmed mean of M values (TMM) method, and linear models were generated for differential expression analyses. To identify differential genes between RRMS phases, comparisons were made between Remission and Relapse, at day and night, respectively. To identify differential genes between different time-points of the day, comparisons were made between day and night after fitting linear models with a paired design for HC, Remission, and Relapse, respectively. In all comparisons, differentially expressed genes were determined by |log2 fold-change| > 1 and false discovery rate (FDR) adjusted p value using the Benjamini-Hochberg (BH) method <0.05.

Functional analysis

In order to interpret biological functions associated with DEGs identified from the comparisons of Remission versus Relapse at night and Relapse-D versus Relapse-N, gene set overrepresentation analysis (GSOA) was performed based on Biological Process (BP) gene sets in the Gene Ontology (GO) database (Ashburner et al., 2000) using HTSanalyzeR package (Wang et al., 2011). The probabilities of overrepresentation of differentially expressed genes in specific gene sets were evaluated using hypergeometric tests. Significantly enriched gene sets (FDR <0.05) were clustered and visualized using AutoAnnotate (Kucera et al., 2016) (Version 1.3.2) app in Cytoscape (Shannon et al., 2003) (v.3.7.2) with Markov Cluster (MCL) clustering method to summarize and interpret functional clusters of GO terms.

Quantitative RT-PCR

cDNAs were prepared from total RNA extracted from RRMS patients and healthy controls. 0.5 μg RNA was used for reverse transcription (RT) reaction using PrimeScript RT Master Mix (Takara). To compare mRNA levels, quantitative RT-PCR was performed using TB Green Premix EX Taq I (Takara) in Applied Biosystems
QuantStudio™ 3 Real-Time Polymerase Chain Reaction System following the manufacturer’s instruction. Data were analyzed by the 2-ΔΔCt method. The gene expression was normalized to a control gene, Ribosomal Protein-encoded Gene 13 (RPS13). Sequences of the primers used in this study:

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward sequence (5’→3’)</th>
<th>Reverse sequence (5’→3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPS13</td>
<td>AAGTACGTGTTTGTGACAGGCA</td>
<td>CGGTGAATCCGCTCTCTATTTAG</td>
</tr>
<tr>
<td>IL18R1</td>
<td>CCTTGACCTTTGGGTGCTTA</td>
<td>CTCATGTGCAAGTGACACCA</td>
</tr>
<tr>
<td>TLR2</td>
<td>TTATCCACGACACGAATACACAG</td>
<td>AGGCATCGGTAGAGTCATCAA</td>
</tr>
<tr>
<td>TPST1</td>
<td>TTTCTAGATTTAACAAATGCC</td>
<td>AGCAGATCTCCACTTTTGCAA</td>
</tr>
<tr>
<td>CXCR4</td>
<td>GGGCAATGGATTTGGTCATCCT</td>
<td>TGCAGGCTGTACTTTGGCC</td>
</tr>
</tbody>
</table>

QUANTIFICATION AND STATISTICAL ANALYSIS

The Benjamini-Hochberg (BH) method was used to adjust for multiple hypotheses testing. Adjusted p < 0.05 was considered significant. All statistical analyses were performed using R (version 4.0.3, www.r-project.org). Bar plot of qPCR result was generated using GraphPad Prism 8 and was presented as mean ± SD.