Single-cell transcriptomics of resected human traumatic brain injury tissues reveals acute activation of endogenous retroviruses in oligodendrocytes

Raquel Garza[1], Yogita Sharma[1], Diahann Atacho[1], Sami Abu Hamdeh[2], Marie Jönsson[1], Martin Ingelsson[3,4], Patric Jern[5], Molly Gale Hammell[6], Elisabet Englund[7], Johan Jakobsson[1, †] & Niklas Marklund [8, †]

Affiliations:

[1] Laboratory of Molecular Neurogenetics, Department of Experimental Medical Science, Wallenberg Neuroscience Center and Lund Stem Cell Center, Lund University, 221 84 Lund, Sweden
[2] Department of Neuroscience, Neurosurgery, Uppsala University, Uppsala, Sweden.
[3] Department of Public Health/Geriatrics, Uppsala University, Uppsala, Sweden
[4] Krembil Brain Institute, University Health Network and Department of Medicine and Tanz Centre for Research in Neurodegenerative Diseases, University of Toronto, Toronto, Ontario, Canada
[5] Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden.
[6] Simons Center for Quantitative Biology, Cold Spring Harbor Laboratory, New York, USA
[7]Department of Clinical Sciences Lund, Division of Pathology, Lund University, Sweden
[8]Department of Clinical Sciences Lund, Neurosurgery, Lund University, Skåne University Hospital, Lund, Sweden

[†] Co-shared senior authorship

Running title: ERV activation in human TBI

Correspondence to:

Niklas Marklund, Department of Clinical Sciences Lund, Neurosurgery, Lund University, Skåne University Hospital, Skåne University Hospital EA-blocket plan 4, Klinikgatan 17A, 222 42 Lund, Sweden. +46 725 95 02 61. niklas.marklund@med.lu.se

Johan Jakobsson, Lund Stem Cell Center, Department of Experimental Medical Sciences, Laboratory of Molecular Neurogenetics, Lund University, Sölvegatan 17, BMC A11, 221 84 Lund, Sweden. +46 709 28 64 43. johan.jakobsson@med.lu.se

Abstract

Traumatic brain injury (TBI) is a leading cause of persistent functional brain impairment and results in a robust, but poorly understood, neuroinflammatory response that contributes to the long-term pathology. Here, we used single-nuclei RNA-sequencing to study transcriptomic changes in different cell populations from human brain tissue obtained acutely after severe, life-threatening TBI. We found a unique transcriptional response in several cell types, including the activation of an interferon response in oligodendrocytes coupled with the transcriptional activation of MHC-class I and class II related genes. Thus, oligodendrocytes undergo a transformation to an immune-like cell state immediately after TBI, indicating an important role for these cells in the initiation of neuroinflammation. Notably, the activation of immune-related genes correlated with the expression of endogenous retroviruses in oligodendrocytes, linking these ancient viral sequences to neuroinflammation. In summary, this work provides a unique insight into the initiating events of the neuroinflammatory response in TBI, which has new therapeutic implications.

Introduction

Traumatic brain injury (TBI) caused by, for example, motor vehicle accidents, violence or falls is a leading cause of mortality and persisting morbidity. The impact to the head results in immediate death to neuronal and glial cells, as well as injury to axons and the microvasculature. This initial primary brain injury is then substantially exacerbated by a poorly understood and progressive cascade of secondary events that may ultimately result in severe long-term consequences. This chronic phase of the injury is often characterized by persistent white matter atrophy, linked to neuroinflammation[1,2], and an increased risk of neurodegenerative disorders, such as Alzheimer's or Parkinson's disease [3,4].

Clinical and experimental evidence suggests that a robust neuroinflammatory response plays a key role in the subsequent development of posttraumatic disability and the increased risk of chronic neurodegenerative disorders [2,5-7]. The neuroinflammatory cascade after TBI is complex and involves the activation of resident microglia as well the recruitment of peripheral immune cells. In addition, other resident brain cells such as astrocytes and oligodendrocytes may be directly involved in modulating or driving the neuroinflammatory response, although their role remains poorly understood [8-10].

Myelinating oligodendrocytes are essential for saltatory nerve conduction in the central nervous system and are involved in the metabolic support of neurons and the modulation of neuronal excitability [11]. There is also emerging evidence that brain oligodendrocytes are morphologically and transcriptionally heterogeneous, and that their functional role may be altered in disease [12,13]. The neuroinflammatory response occurring after TBI seems to be associated with oligodendrocyte

vulnerability [9], and persistent neuroinflammation is often found in atrophied white matter tracts of long-term TBI survivors [2]. These observations argue for an important contribution of the neuroinflammatory response to white matter injury. Notably, oligodendrocytes have recently been shown to activate the transcription of immune-response genes in certain disease contexts [14,15], suggesting that these cells may play a direct role in the neuroinflammatory process by adopting an immune-like cell state. Despite these observations, the role of oligodendrocytes in the neuroinflammatory response occurring after TBI remains poorly understood.

While the presence of neuroinflammation after TBI, and in other neurodegenerative conditions, has been firmly established, the underlying cause of this sterile inflammatory response remains largely unknown. We and others have recently found that transcriptional activation of endogenous retroviruses (ERVs) can trigger or boost a neuroinflammatory response in mouse models [16,17]. ERVs are remnants of old retrovirus infections that have entered our germline and make up about 8% of our genome. Recent evidence suggests ERVs can be transcriptionally activated in certain neurological disease states [16]. This aberrant expression of ERVs results in the formation of double-stranded RNAs, reverse-transcribed DNA molecules, and TE-derived peptides that induce a "viral mimicry", where cells of the central nervous system respond by activating an interferon response as if they were infected [18-20]. This event results in a downstream trigger of inflammation that may participate in the respective disease processes. If and how ERVs are activated in human brain disorders and after insult, such as in TBI, remains poorly documented.

In this study, we performed single nuclei RNA sequencing (snRNA-seq) on different cell populations in fresh-frozen brain tissue, which had been surgically removed due to life-threatening TBI. We found that TBI resulted in a unique transcriptional response in several cell types in the

injured human brain tissue, including the activation of cell-cycle-related genes in microglia as well as alterations in synaptic gene expression in excitatory neurons. Notably, we detected clear evidence of an interferon response in oligodendrocytes that was linked to the transcriptional activation of the major histocompatibility complexes (MHC) class I and II, as well as regulators and related genes from both classes. These results demonstrate that oligodendrocytes undergo a transformation to an immune-like cell state after TBI and suggest a key role for these cells in the initiation of neuroinflammation following such an insult. Moreover, the activation of immune-related genes correlated with the expression of ERVs in oligodendrocytes, implicating a role for these ancient viral sequences in human neurodegeneration. These results not only provide novel insights into initiating events of the neuroinflammatory response following severe TBI, but also opens up new therapeutic avenues.

Results

Samples and experimental design

To investigate changes in cell-type composition and cell-type-specific transcriptional responses after severe, acute TBI, we performed single-nuclei RNA sequencing (snRNA-seq) from fresh-frozen human brain tissue (Fig 1a). We recruited 12 severe TBI patients, defined as post-resuscitation Glasgow Coma Scale score ≤ 8. Detailed demographic and clinical characteristics are shown in Table 1. The age (mean ± standard deviation) of the patients (10 males, 2 females) was 49.5 ± 18.2 years. In these patients, decompressive surgery was a life-saving measure to remove injured and swollen space-occupying brain tissue causing marked mass effect or increased intracranial pressure (ICP) refractory to conservative, medical neurointensive care treatment. The

injured and contused brain regions (typically the injured part of a temporal or frontal lobe) were surgically removed between 4 hours and 8 days after injury [21]. As control tissue, we used five fresh-frozen *post-mortem* samples from the frontal and temporal lobe obtained from three non-neurological deaths of patients aged 69, 75 and 87 years (Table 1).

We isolated nuclei from frozen human brain tissue using ultracentrifugation and FACS, and performed snRNA-seq using the 10X Genomics pipeline. The tissue obtained from TBI patients contained a high degree of blood and damaged cells, which greatly influenced the quality of sequencing data. Therefore, we used strict thresholds to validate the quality of each sequenced cell: only cells with between 1,000 and 12,000 detected genes were retained for further analysis. After quality control (QC), we kept high-quality sequencing data from the 12 TBI samples, with a total of 6,806 nuclei and a mean of 2,361 genes per nucleus. From the five control samples we obtained 8,304 nuclei with a mean of 3,389 genes per nucleus (Fig 1a).

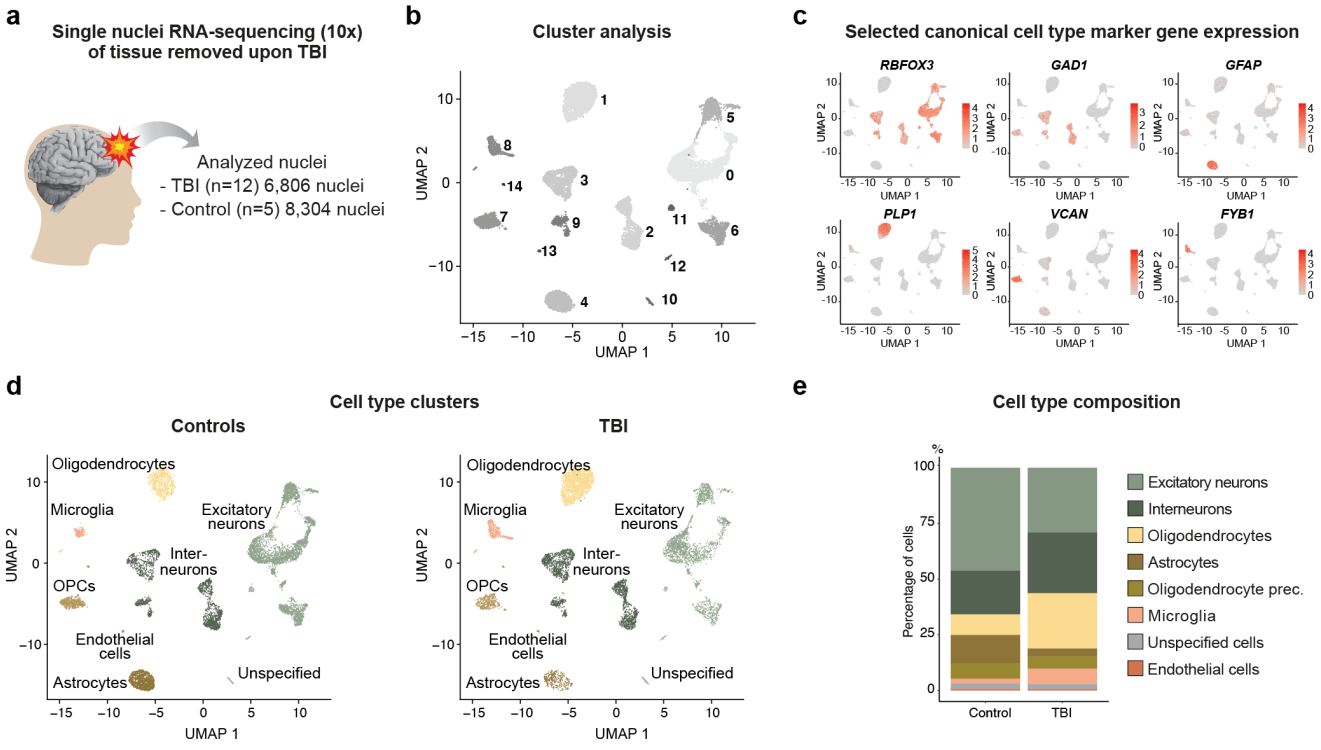


Figure 1 a) Schematic of experimental approach. Brain tissue was removed after severe TBI followed by single-nuclei RNA-seq. Numbers indicate the number of nuclei recovered after quality control per condition. b) UMAP labelled with the nuclei clusters identified (clusters 0-14). c) Projection of gene expression of canonical gene markers to identify cell types. d) UMAP labelled by characterized cell types, split by condition. e) Bar plot showing cell type ratio per condition.

Sample	Region	Age	Sex	Injury	Time post injury (h)	GOSE
TBI 1	Temporal lobe	22	M	MVA	9	3
TBI 2	Frontal lobe	72	M	Fall	4	7
TBI 3	Frontal lobe	42	M	MVA	4	8
TBI 6	Temporal lobe	74	M	Fall	4	4
TBI 7	Temporal lobe	58	M	Fall	9	7
TBI 8	Temporal lobe	49	M	Fall	84	1
TBI 10	Temporal lobe	65	M	Fall	180	3
TBI 11	Frontal/parietal lobe	25	M	SPR	24	3
TBI 16	Temporal lobe	52	M	Fall	42	4
TBI 19	Frontal lobe	49	F	HOB	57	-
TBI 20	Frontal lobe	62	F	MVA	192	8
TBI 21	Frontal lobe	24	M	Fall	106	6
CTL 501	Frontal lobe	87	M	-	-	-
CTL 501	Temporal lobe	87	M	-	-	-
CTL 502	Temporal lobe	75	M	-	-	-
CTL 529	Frontal lobe	69	M	-	-	-
CTL 529	Temporal lobe	69	M	-	-	-

Table 1 Patients' and controls' demographics, and region of sample collection. MVA = motor vehicle accident, SPR = sports related, HOB = hit by object. GOSE = Glasgow Outcome Scale extended

Cell type composition

Using the snRNA-seq data, we performed an unbiased clustering using Seurat to identify and quantify the different cell types present in the injured brain tissue. We detected 15 different clusters and used canonical marker gene expression to identify the different cell types (Fig 1b-e). We identified three clusters of excitatory neurons and four clusters of inhibitory neurons, as well as clusters with oligodendrocytes, oligodendrocyte precursor cells (OPCs), astrocytes, endothelial cells, and microglia/macrophages (Fig 1d). All clusters were represented in both TBI and control samples, and the cellular composition was similar between samples obtained from temporal and frontal brain regions (Fig 1d-e and Suppl Fig 1a-b).

The most abundant cell type in the control samples was excitatory neurons, which accounted for almost 50% of the cells, followed by interneurons (20%) and astrocytes (13%) (Fig 1e). Oligodendrocytes made up 9% of the cells while microglia represented 2%. In the TBI samples we found a reduction in the proportion of excitatory neurons, only 29% of the cells. In contrast, the proportions of both oligodendrocytes and microglia were higher (25% and 7%, respectively). The increase in oligodendrocytes could potentially be explained by the presence of white matter tissue in the TBI samples, which is difficult to control for in this clinical setting. However, the reduced numbers of excitatory neurons as well as the increase in microglia are likely to be due to the direct, acute consequences of the injury (Fig. 1e).

Oligodendrocytes display an interferon response after acute TBI

The neuroinflammatory response is thought to be a key mechanism in the subsequent pathological processes following TBI. However, how this process starts and what cell types are involved in humans is unknown. To investigate transcriptional changes linked to inflammation, we analyzed cell-type-specific transcriptional alterations after TBI. We found that each cell type displayed a distinct transcriptional response to TBI. For example, in excitatory neurons we found that genes linked to synaptic functions, such as NPTX2 and HOMER1, were upregulated: gene set enrichment analysis (GSEA) confirmed that genes linked to cell communication and synaptic signaling were dysregulated in excitatory neurons (Suppl Fig 2d). In microglia we found that genes linked to the cell cycle, such as MKI67 and TOP2A, were upregulated, and GO analysis confirmed a transcriptional response linked to cell proliferation (Suppl Fig 2e). In line with this observation, a global analysis of cell-cycle-related genes (function '*CellCycleScoring*', Seurat) confirmed that

we detected microglial cells that were in a proliferative state in TBI samples and not in the controls (Suppl Fig 2f), which is in line with what is expected in healthy human brain. No evidence of proliferation was detected in any other cell types after TBI. Thus, these results demonstrate that we detected neuronal dysfunction and the initiation of an inflammatory response upon TBI using the snRNA-seq approach.

When investigating transcriptomic changes in the other cell types, we found that oligodendrocytes displayed a unique transcriptional response linked to neuroinflammation following severe TBI. We found that many genes involved in an interferon response and viral infection, including STAT1 and STAT2, CD47, IRF1, IFI16, and PARP14, were specifically activated in oligodendrocytes (Fig 2a-c). GSEA confirmed that genes related to terms such as innate immune response, response to interferon-gamma, and response to cytokine were significantly enriched among upregulated genes in TBI oligodendrocytes (Fig 2d). These genes were found to be especially enriched in oligodendrocytes of samples collected a short time after the injury (up to 4 hours) (Suppl Fig 2a). These observations indicate that oligodendrocytes play a key role in the initiation of inflammation following TBI via the activation of an interferon response. Interestingly, the upregulation of inflammation-related genes was also found in OPCs, and GSEA confirmed the enrichment among upregulated genes of GO terms related to innate immunity and defense response (Suppl Fig 2b-c).

Notably, many of the genes upregulated in TBI oligodendrocytes are MHC class I genes as well as regulators from class I and class II (Fig 2f). For example, TBI oligodendrocytes expressed NLRC5 and CIITA (Fig 2e-f), the major regulators of genes in the MHC class I and class II respectively [22], as well as PSMB8, PSMB9, TAP1, and TAP2 which process MHC class I molecules and peptides, and MYO1E and CD74 which help process and guide the expression of

MHC class II molecules [23] (Fig 2e). These results indicate that oligodendrocytes undergo a transcriptional switch following TBI which turns on MHC-related genes normally expressed in immune cells, hereby transforming to an immune-like cell state.

We confirmed these observations using immunohistochemical staining for the interferon response genes STAT1 and CD47 on tissue sections from TBI patients. Using confocal microscopy, we found evidence for STAT1 and CD47 protein expression in OLIG2-expressing oligodendrocytes in TBI tissue (Fig 2g).

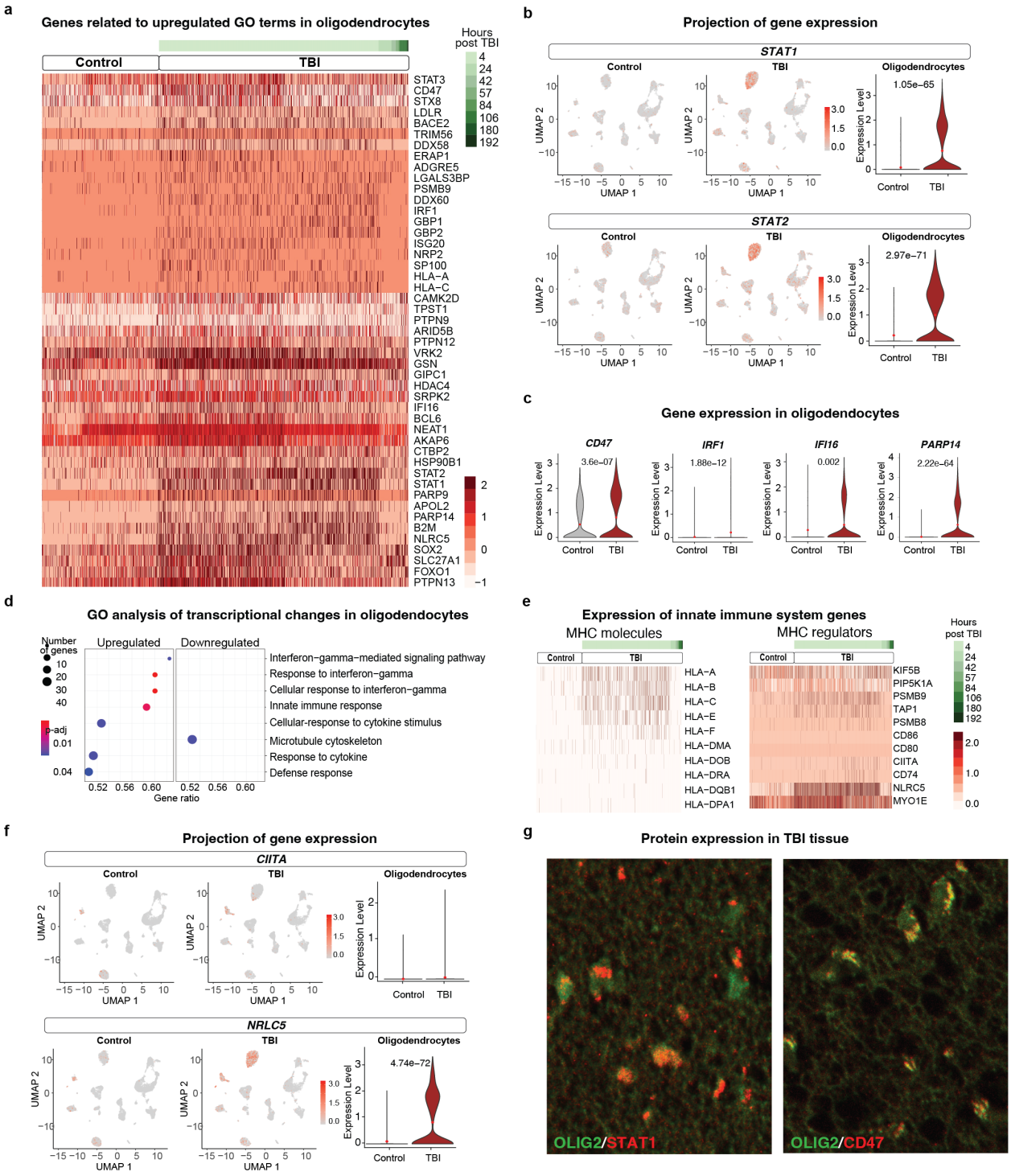


Figure 2 a) Heatmap of genes from activated GO terms (Interferon-gamma-mediated signaling pathway, Response to interferon-gamma, Cellular response to interferon-gamma, Innate immune response, Cellular-response to cytokine stimulus, Response to cytokine, and Defense response) in oligodendrocytes. b) Top panel showing STAT1 expression projected onto control (left) and TBI (middle) UMAPs; violin plot (right) showing expression per condition in oligodendrocytes. Bottom panel showing STAT2 expression projected onto control (left) and TBI (middle) UMAPs; violin plot (right) showing expression per condition in

oligodendrocytes. c) Violin plots showing the expression of selected genes related to defense, interferon response, or inflammation in oligodendrocytes. d) GSEA of differentially-expressed genes (TBI vs control; padj < 0.01) in oligodendrocytes. e) Expression of detected molecules (top) of MHC class I and class II and selected related genes (bottom). f) Top panel showing the expression of the key regulator of MHC class I molecules, CIITA, projected onto control (left) and TBI (middle) UMAPs; violin plot (right) showing expression per condition in oligodendrocytes. Bottom panel showing the expression of the key regulator of MHC class II molecules, NLRC5, projected onto control (left) and TBI (middle) UMAPs; violin plot (right) showing NLRC5 expression per condition in oligodendrocytes. g) Immunohistochemical staining of STAT1 and CD47 proteins in OLIG2-positive cells in TBI samples.

Activation of endogenous retroviruses in oligodendrocytes after TBI

The transcriptional response in oligodendrocytes after TBI is reminiscent to that which occurs after viral infection. In this respect, the transcriptional activation of endogenous retroviruses (ERVs) (Fig 3a) has been linked to an interferon response. The aberrant expression of ERVs results in the formation of double-stranded RNAs, reverse-transcribed DNA molecules, and ERV-derived peptides that induce a "viral mimicry", where cells respond as though infected and this triggers or boosts inflammation [18-20].

To investigate if ERVs, or any other transposable elements (TEs), were activated following TBI, we used an in-house bioinformatical pipeline [17] allowing the analysis of TE expression from 10X single-nuclei RNA-seq data set (Fig 3b). In brief, this method uses the cell clusters determined based on the gene expression. Then, by back-tracing the reads from cells forming each cluster, it is possible to analyze the expression of ERVs and other TEs using the specialized software *TEtranscripts* [24] in distinct cell populations. This approach greatly increases the sensitivity of the analysis and enables quantitative estimation of ERV expression at single-cell-type resolution.

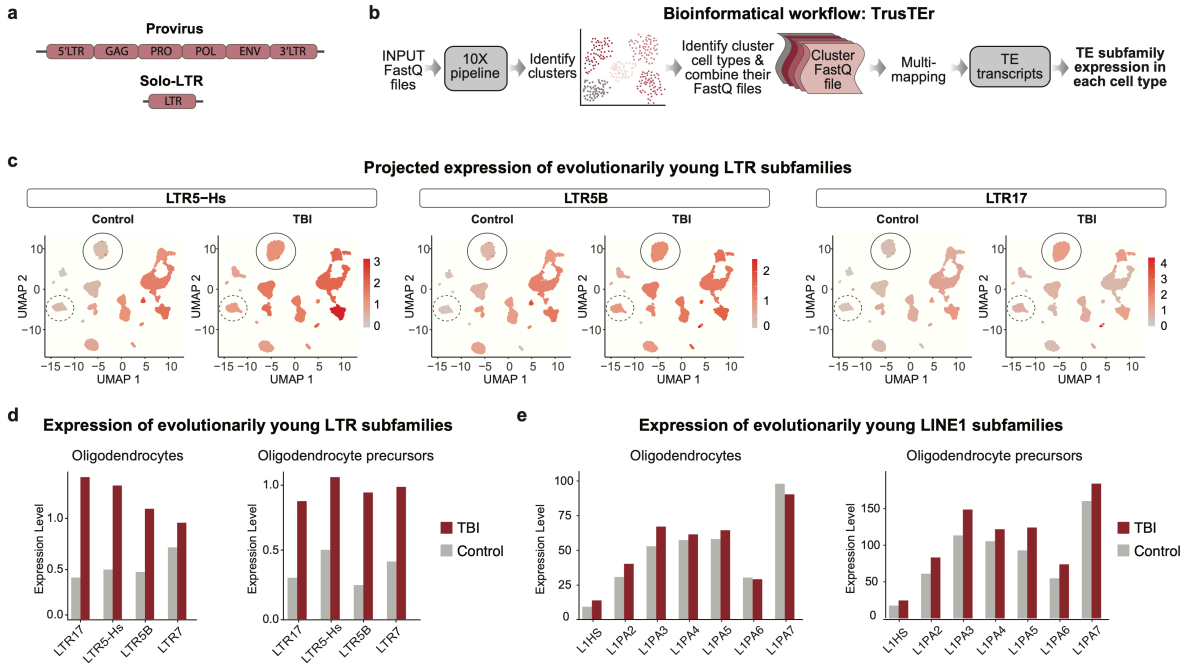


Figure 3 a) Schematic of the structure of an ERV provirus (top) and solo LTR (bottom). b) Schematic of bioinformatical approach to quantify TE subfamilies per cell cluster. c) Split UMAPs (control and TBI) with projected expression of LTR subfamilies per cluster. Oligodendrocytes circled, and OPCs circled in dotted line. d) Expression of evolutionary young LTR subfamilies in oligodendrocytes and OPCs. e) Expression of evolutionary young L1 subfamilies in oligodendrocytes and OPCs.

When using this bioinformatical approach, we found that several ERV (long terminal repeat – LTR) subfamilies were transcriptionally activated in oligodendrocytes and OPCs. This response was especially noticeable in ERV subfamilies such as HERV-K (LTR5-Hs & LTR5B, fold change of 2.34-2.7 in oligodendrocytes, 2-3.7 in OPCs), HERV-W (LTR17, fold change 3.41 in oligodendrocytes, 2.8 in OPCs) and HERV-H (LTR7 fold change 1.34 in oligodendrocytes, 2.3 in OPCs) which are evolutionary young ERVs identified in primates (Fig 3c-d). We found no transcriptional activation of other families of transposable elements such as LINE-1s (Fig 3e), and no evidence for the activation of ERVs or other TEs in other cell types, except for microglia which also displayed some evidence of ERV activation (Suppl Fig 3a-b).

Many of the ERV families that were upregulated in TBI oligodendrocytes include provirus insertions with the potential to be transcribed and translated to produce ERV-derived peptides (top panel Fig 3a). Such ERVs have previously been linked to viral mimicry and the induction of an interferon response [19]. However, since 10X libraries display a 3' bias, it is not possible to distinguish between ERVs and solo LTR fragments that are present in large numbers in the human genome due to recombination events between the two endogenous provirus LTRs during primate evolution (Fig 3a). Thus, to investigate which unique loci the upregulated ERV expression in TBI oligodendrocytes originates from, we performed deep 2×150 bp paired-end, strand-specific bulk RNA-seq of four of the TBI samples with the highest ratio of oligodendrocytes, as well as the control samples (Fig 4a and Suppl Fig 1a). Using a unique-mapping approach, we quantified ERV predictions identified using RetroTector [25] and found 13 significantly upregulated ERV loci in the TBI samples (*DESeq2*, p-value < 0.05, log2FC > 1) (Fig 4b) which range in length from 6.2 kbp to 13.5 kbp, where transcription was found across the ERV sequence (Fig 4c). Six of these upregulated ERVs retained only the 5' LTR, while the remaining seven retained both the 5' and 3' LTRs (Fig 4d). Furthermore, the predicted structures of these elements show that 12 out of the 13 upregulated ERVs still contain at least one ORF of the ERV-genes (*gag, pro, pol* and *env*) (Fig 4e). These data demonstrate that ERVs, with the potential to induce viral mimicry, are transcriptionally activated in oligodendrocytes within hours after TBI.

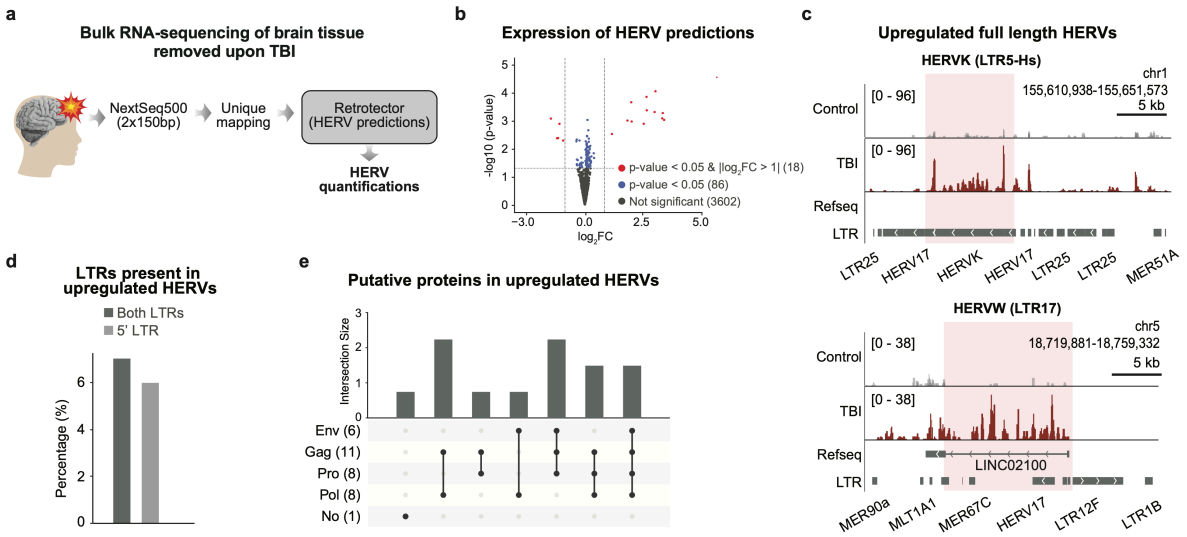


Figure 4 a) Schematic for bulk RNA-seq and bioinformatical approach. b) Volcano plot of *DESeq2* results for all ERV predictions. P-value < 0.05, log2FC > 1 highlighted in red; p-value < 0.05 highlighted in blue. c) Tracks per condition with two examples of upregulated HERVs (top - HERVK, bottom - HERVW). d) Bar plots showing the number of upregulated ERV predictions with both LTRs or only the 5' LTR. e) Upset plot showing putative proteins contained in the upregulated ERVs.

Discussion

Neuroinflammation is a hallmark of acute and chronic neurodegenerative states, including TBI, and is therefore a promising route for urgently-needed disease modifying therapies. However, information on how the neuroinflammatory response starts is scarce, and the molecular events needed to initiate this process, and in which cell types, has not been established. This lack of basic understanding hampers the development of new treatment strategies. In this study, we used snRNA-seq to perform a detailed analysis of human tissue samples obtained in conjunction with acute surgery for TBI. Our results have provided two unique insights into the start of a human neuroinflammatory response. First, we found that oligodendrocytes play an unanticipated key role in this process by initiating an interferon response and adopting an immune-like cell state. Secondly, we found that the activation of an interferon response in oligodendrocytes correlates with the transcriptional activation of ERVs.

Neuronal cell death is common in acute, severe TBI due to hemorrhages, increased intracranial pressure, and energy metabolic failure. In addition, there is a distinct inflammatory response, associated with white matter abnormalities, that persists from the acute post-injury phase for many years post-injury [2,6,26,27]. Prior studies have established that there are numerous contributors to the neuroinflammatory response after TBI, including microglial activation, infiltration of systemic immune cells, and the subsequent release of pro-inflammatory cytokines [28,29]. Our observations indicate that oligodendrocytes also seem to play an important role in the initiation of this process. Oligodendrocytes are vulnerable to the TBI-induced neuroinflammation as well as to excitotoxicity and reactive oxygen species formation [30,31], and loss of mature oligodendrocytes is observed at an early stage following rodent [32-34] and human TBI [35]; treatment with an anti-inflammatory antibody neutralizing interleukin-1β, a key pro-inflammatory cytokine, attenuates this post-injury loss of oligodendrocytes in rodent models [9]. In the present study, we found that oligodendrocytes activate the expression of interferon response genes as well as genes related to MHC class I and class II. Some of these genes affect the expression, processing, and guidance of MHC molecules, which are crucial to immune cell functions. Our results suggest that, following TBI, oligodendrocytes adopt a different cell state characterized by the expression of immune genes. These results are reminiscent of what has been observed in multiple sclerosis, where oligodendrocytes initiate a similar transcriptional program and adopt different functional properties, such as cell phagocytosis [14,15]. Future investigations into the role of oligodendrocytes in the inflammatory response, and the underlying mechanisms linked to their cell fate change, could lead to new treatment opportunities.

Almost 10% of the human genome is made up of ERVs, a consequence of their colonization of our germline throughout evolution [36]. In adult tissues, including the brain, ERVs are normally transcriptionally silenced via epigenetic mechanisms, including DNA and histone methylation [16]. However, there is emerging evidence, both from animal models and the analysis of human material, indicating that ERVs can be transcriptionally activated in the diseased brain and that this correlates with neuroinflammation. ERV expression has been found to be elevated in the cerebrospinal fluid and in *post-mortem* brain biopsies from patients with multiple sclerosis, amyotrophic lateral sclerosis, and Alzheimer's and Parkinson's disease [16,37-45]. In experimental drosophila or mouse models there is causative evidence linking upregulation of ERVs and other transposable elements to neuroinflammation and neurodegeneration [17,46]. However, accurate estimation of ERV expression is challenging: their repetitive nature complicates the bioinformatical analysis and many of these observations remain controversial. Furthermore, most of the clinical studies use end-stage *post-mortem* material where the inflammatory process has been ongoing for decades or are limited to the study of biofluids such as cerebrospinal fluid or plasma. These limitations have made it challenging to conclusively link ERV expression to the initiation of the neuroinflammatory response. Our results now provide direct evidence that ERV proviruses are transcriptionally activated at the start of human neuroinflammation. Notably, we find that ERVs are specifically activated in oligodendrocytes and OPCs, the cell types involved in the interferon response. These results provide direct clinical evidence linking the induction of an immune response and the expression of ERVs.

The exact molecular mechanisms underlying the TBI-induced transcriptional activation of ERVs remain to be established, but it is likely that epigenetic remodeling plays a role in this process: we

have previously observed alterations in DNA-methylation patterns following human TBI [21]. From our results it is also not possible to say whether ERVs are driving the immune response or if ERV expression is a consequence of this phenomenon. The induction of interferon-linked genes has been associated with ERV activation: ERVs are thought to play a part in boosting this response[47]. In addition, ERV-derived peptides may contribute to protein aggregation, which may further enhance an inflammatory response [48]. Future mechanistic *in-vitro* modeling in human oligodendrocyte cultures will be essential to resolve this question.

In summary, we here describe a novel role for oligodendrocytes in the initiation of the neuroinflammatory response following acute severe human TBI. We describe the activation of an interferon response in oligodendrocytes that is linked to the activation of regulatory genes of molecules present in immune cells. This transcriptional switch in oligodendrocytes correlates with the activation of ERVs. Our results could lead to entirely new treatment opportunities for acute TBI, as well as for chronic neurodegenerative disorders.

Methods

Ethical statement

All clinical and experimental research described herein was approved by the regional ethical review board (decision numbers 2005/103, 2008/303, 2009/89, and 2010/379). Written informed consent was obtained from the TBI patients' closest relatives and from the patients themselves if they had sufficiently recovered from their injury at >6 months post injury.

TBI and control cohorts

Detailed demographic and clinical characteristics are shown in Table 1. Twelve patients with severe TBI, defined as post-resuscitation Glasgow Coma Scale score ≤ 8, were included. The patients were >18 years old, and no patient had any other known neurological disorder or Down's syndrome. All patients had been mechanically ventilated and sedated and continuous measurements of intracranial pressure (ICP) and cerebral perfusion pressure (CPP) were performed at the neurocritical care unit at Uppsala University Hospital [49].

The age (mean ± SD) of the TBI patients (10 males, 2 females) was 49.5 ± 18.2 years. The tissue samples were obtained from patients suffering severe, life-threatening focal TBI. Due to space-occupying brain swelling causing midline shift and compression of the basal cisterns, or due to markedly increased intracranial pressure (ICP) refractory to medical neurointensive care measures, contused brain regions (typically the injured part of a temporal or frontal lobe) were surgically removed between 4 hours and 8 days after injury [21]. There were no complications associated with the surgical method. The patients needed prolonged neurocritical, neurosurgical, and then

neurorehabilitative care. At the time of follow-up, one patient was deceased. All available patients' outcomes are shown in Table 1.

Sampling and preparation of brain tissue

Surgically removed brain tissues were immediately placed in a sterile pre-labeled container and subsequently stored at −80°C until analyzed. Half of the tissue was put in a routinely-used fixative, 4% buffered formalin (Histo-Lab Products AB, Gothenburg, Sweden, catalogue no. 02176). The samples were fixed for 24-72 h and then paraffin-embedded and processed by hardware Tissue tek VIP (Sakura, CA, USA). Small samples were taken from the fresh-frozen contused brain tissue for snRNA-seq.

Immunohistochemistry

Immunohistochemistry was performed on microtome sections according to previously-published protocols. [17]

Single-nuclei RNA sequencing analysis

The FACS-based isolation of nuclei from frozen brain tissue was performed as previously described and loaded onto 10x Genomics Single Cell 3' Chip [17]. Sequencing library samples were multiplexed and sequenced on a Novaseq machine using a 150-cycle kit with the recommended read length from 10x Genomics. Raw base calls were demultiplexed to obtained sample-specific FastQ files and reads were aligned to GRCh38 genome assembly using the Cell Ranger pipeline (10x Genomics, Cellranger count v5) with default parameters (--include-introns was used for nuclei mapping). The resulting matrix files were used for further downstream analysis. Seurat

(version 3.1.1) and R (version 3.4) were used for bioinformatics analysis. Low-quality cells and genes were filtered out based on fraction of total number of genes detected in each cell (±3 nmads). All samples were merged and batch effects were removed using the Harmony algorithm [50]. Cells with at least 1,000 detected genes were retained and the data was normalized to transcript copies per 10,000, and log-normalized to reduce sequencing depth variability. For visualization and clustering, manifolds were calculated using UMAP methods (*RunUMAP,* Seurat) and 20 precomputed principal components and the shared nearest neighbor algorithm modularity optimization-based clustering algorithm (FindClusters, Seurat) with a resolution of 0.1. The clusters obtained were annotated using canonical marker gene expression. Differential gene expression between TBI and control samples across respective cell types was carried out using the Seurat function FindMarkers (Wilcox test, padj < 0.01). Cell-cycle scores were computed using the Seurat function CellCycleScoring. Gene ontology overrepresentation analysis was performed using the enrichGO function in the clusterProfiler R package using differentially-expressed genes in each cell type (padj < 0.01) [51]. Enrichment scores for genes related to selected immune-related GO terms were calculated using the AddModuleScore function from Seurat using five features as controls (ctrl = 5).

TE quantification per cell cluster

To ease the usage of the workflow we made use of, we created a Python package named trusTEr to process single-cell RNA-seq data to quantify TE expression per cluster of cells. The package consists of three classes (Experiment, Sample, and Cluster) for the abstraction of the projects, and a module (jobHandler.py) to handle SLURM jobs. For more detailed information about how the

classes are related, please refer to https://raquelgarza.github.io/truster/. The workflow has the following steps:

1) Extraction of reads from samples' BAM files (tsv_to_bam() in class Cluster). After the clustering formation, trusTEr expects one text file per cluster containing all barcodes of a sample in a cluster (done by default by get_clusters() and merge_clusters()). It will extract the given barcodes from the sample's BAM file (outs/possorted_genome_bam.bam output from Cell Ranger count) using the subset-bam software from 10x Genomics (version 1.0). Outputs a BAM file per cluster containing all alignments from the cells in the cluster.
2) Filter duplicates (filter_UMIs() in class Cluster). Given the number of PCR duplicates we could be carrying in these BAM files, we wrote a Python script (filterUMIs) to filter most of these out. The filter_UMIs function keeps reads with unique combinations of cell barcodes, UMI, and sequence, so that only unique molecules are kept.
3) Convert to FastQ (bam_to_fastq() in class Cluster). Using bamtofastq from 10x Genomics (version 1.2.0), the previous BAM files are output as fastQ files.
4) Concatenate lanes (concatenate_lanes() in class Cluster). Concatenates the different lanes from the same library as output from bamtofastq. This step outputs one fastQ file per cluster.
5) The quantification was performed in groups of samples (TBI and control groups) to increase statistical power. This step concatenates the fastQ files of the samples' clusters within a single group (see the groups parameters in the process_clusters function in the Experiment class, or in each of the steps' corresponding functions).
6) Map cluster (map_cluster() in class Cluster). Mapping the reads to a reference genome using STAR aligner [52]. For subfamily quantification (default in the map_cluster function)

–outFilterMultimapNmax 100, --winAnchorMultimapNmax 200 were used. To quantify individual elements (argument unique set to True), unique mapping was performed using –outFilterMultimapNmax 1 and –outFilterMismatchNoverLmax 0.03.

7) TE count (TE_count() in class Cluster). TE quantification of the BAM files produced for each cluster. For the purposes of this paper, we used TEcount from the TEToolkit with the curated TE GTF for hg38 provided by the authors (--TE), gencode version 36 as the gene GTF (--GTF). We ran it in multi mode (--mode multi) as forward stranded (--stranded yes) [53].

8) Normalization of TE counts. TE quantification is normalized by cluster size (number of cells in a cluster), and it's stored within the Seurat object. The matrices output and the Seurat assay only includes the normalized TE subfamily information output by TEcount.

Bulk RNA sequencing analysis

Total RNA was isolated from nuclei using the Rneasy Mini Kit (Qiagen). Libraries were generated using Illumina TruSeq Stranded mRNA library prep kit (poly-A selection) and were sequenced on an Illumina NextSeq500 machine (paired-end 2×150 bp).

Reads were mapped using STAR (2.6.0c) [52] with GRCh38.p13 as the genome index, and gencode version 38 to guide the mapping (--sjdbGTFfile). To avoid ambiguous reads due to the quantification of ERVs, the number of loci allowed for a read to map (--outFilterMultimapNmax) was set to 1, and the number of mismatches allowed (--outFilterMismatchNoverLmax) was set to 3%.

The RetroTector software [25], (https://github.com/PatricJernLab/RetroTector) was used to mine the human genome (version GRCh38/hg38 downloaded from https://hgdownload.soe.ucsc.edu/) for

ERVs as previously described.[54] Predicted positions and structures for ERVs scoring 300 and above are summarized in Suppl. Tables 1 and 2. Using the GTF version of the output file, read were quantified using featureCounts (Subread version 1.6.3) [55], forcing matching strandness of the reads to the features being quantified (-s 2).

Read count matrices were then input to DESeq2 (version 1.28.1) and fold changes shrunk using DESeq2::lfcShrink [56]. Volcano plots were visualized using the R package EnhancedVolcano (version 1.14.0) (https://github.com/kevinblighe/EnhancedVolcano). For further details, please refer to the corresponding Rmarkdown (TBI_bulk.Rmd) on github.

Data and code availability

Raw data of single-nuclei and bulk RNA-seq of the control post-mortem samples, and processed data of all samples can be found at the GEO repository GSE209552. Pipelines can be found at https://github.com/raquelgarza/truster and documentation at https://raquelgarza.github.io/truster/. Rmarkdowns for all the visualization and bulk RNA-seq analysis can be found at https://github.com/raquelgarza/tbi_Garza2022.

Acknowledgements

We would like to thank Roger A. Barker for critical input on the study and. Alafuzoff, J. Johansson, U. Jarl, M Vejgården and A. Hammarberg for technical assistance. We are grateful to all members of the Jakobsson and Marklund labs. The work was supported by grants from the Swedish Research Council (2018-02694 to J.J., 2018-02500 to N.M., 2021-01740 to P.J.), the Swedish Brain

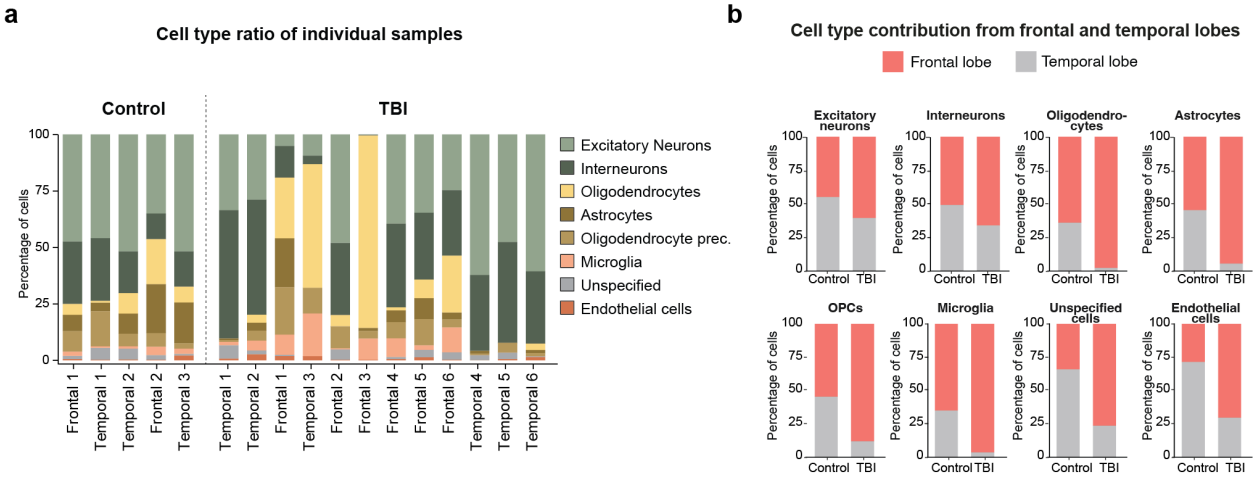
Foundation (FO2019-0098 to J.J., FO2020-0147to NM) and the Swedish Government Initiative for Strategic Research Areas (MultiPark & StemTherapy).

Author contributions

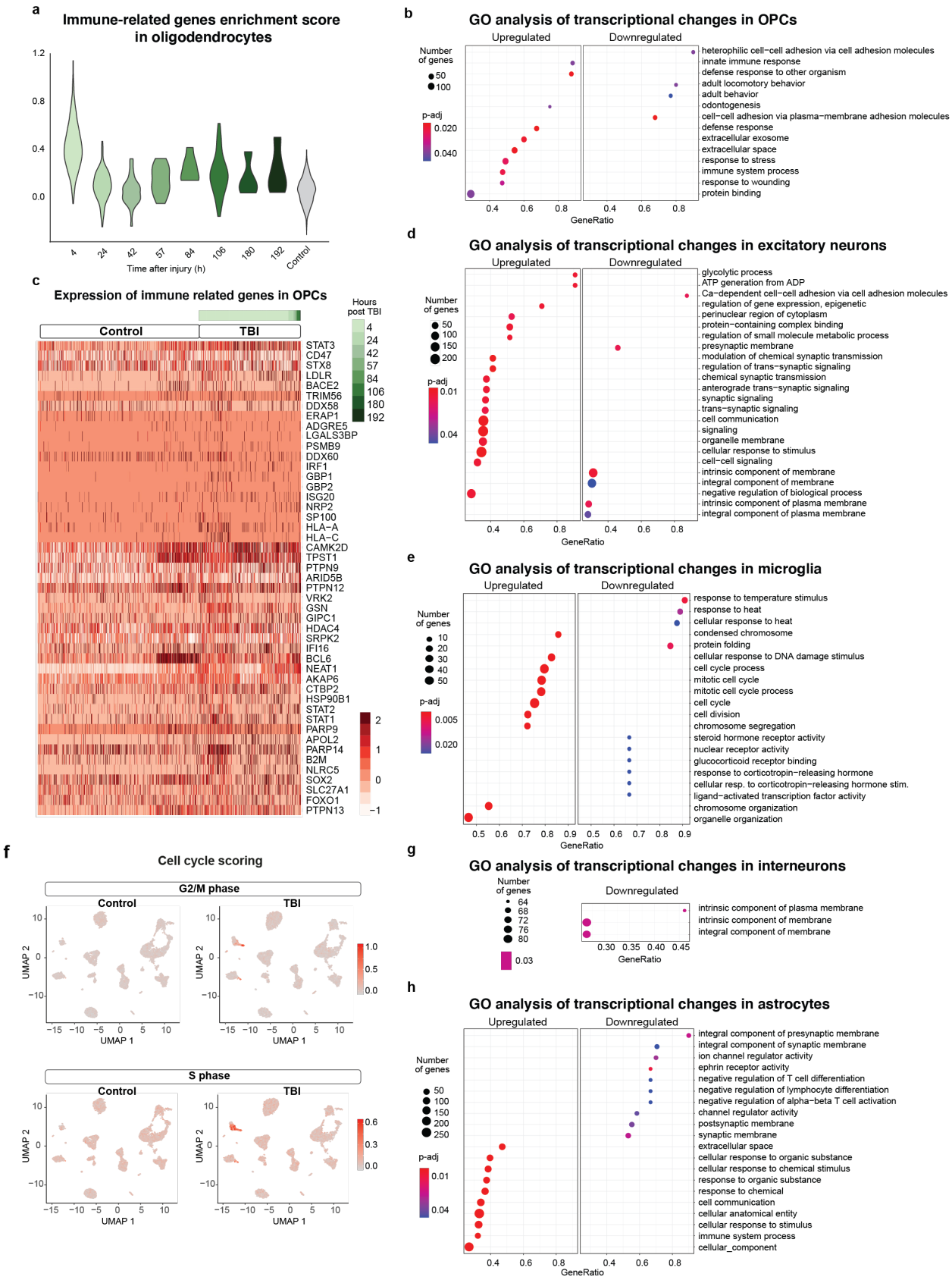
All of the authors took part in designing the study and interpreting the data. R.G., J.J., and N.M. conceived and designed the study. D.A., S.A.H, and M.J. performed the clinical and experimental research. R.G. and Y.S. performed the bioinformatic analyses. M.I., M.G.H., P.J., and E.E. contributed reagents and expertise. R.G., J.J., and N.M. wrote the manuscript, and all authors reviewed the final version.

Competing interests

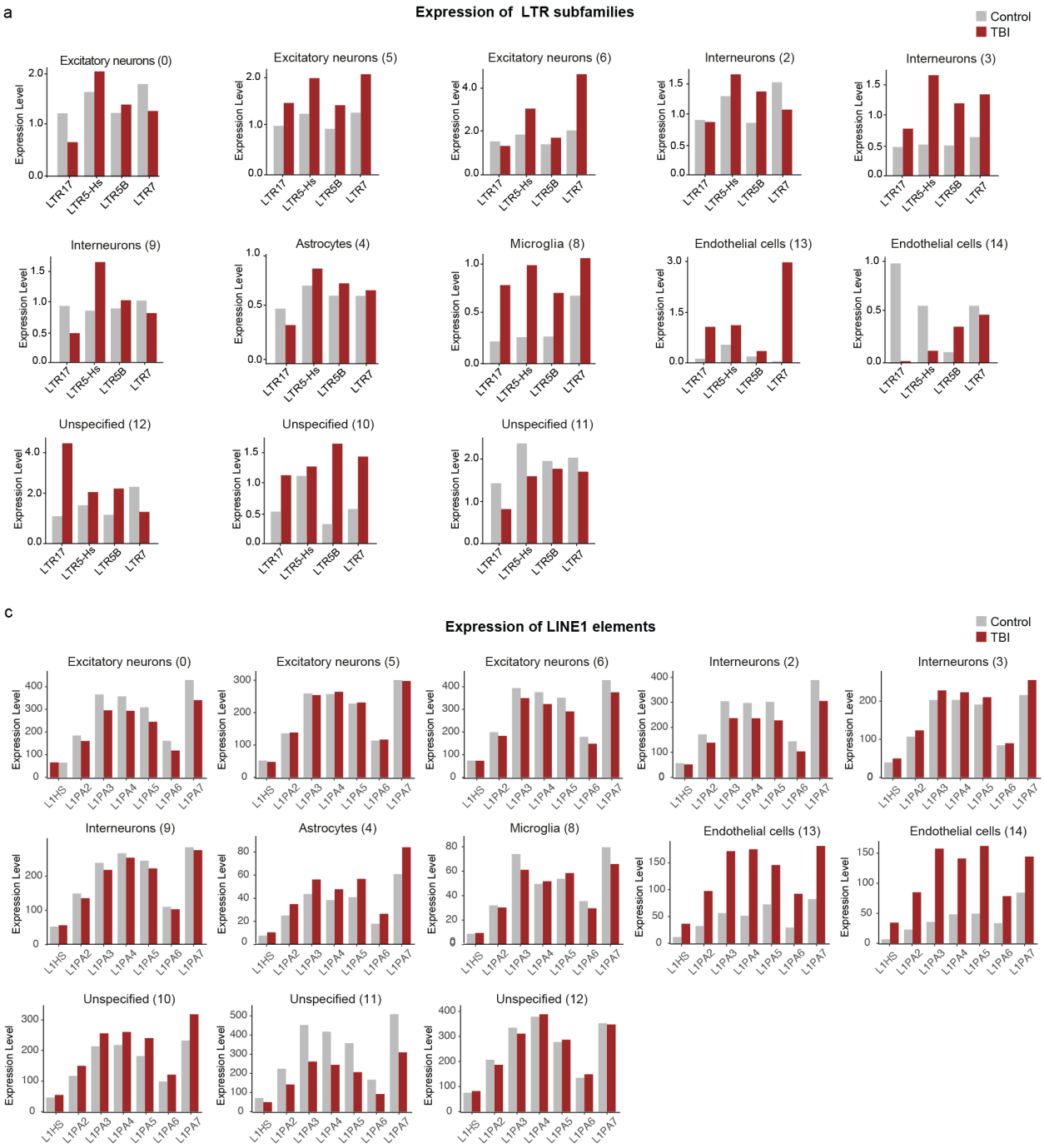
The authors declare no competing interests.



Suppl. Figure 1 a) Cell-type ratio found in each sample per condition. b) Percentage of cells belonging to samples collected from the frontal or temporal lobe in each cell type.


a
Immune-related genes enrichment score in oligodendrocytes
1.2
0.8
0.4
0.0
4
24
42
57
84
106
180
192
Control
Time after injury (h)
b
GO analysis of transcriptional changes in OPCs
Upregulated
Downregulated
Number of genes
50
100
p-adj
0.020
0.040
heterophilic cell−cell adhesion via cell adhesion molecules
innate immune response
defense response to other organism
adult locomotory behavior
adult behavior
odontogenesis
cell−cell adhesion via plasma−membrane adhesion molecules
defense response
extracellular exosome
extracellular space
response to stress
immune system process
response to wounding
protein binding
GeneRatio
c
Expression of immune related genes in OPCs
Control
TBI
Hours post TBI
4
24
42
57
84
106
180
192
STAT3
CD47
STX8
LDLR
BACE2
TRIM56
DDX58
ERAP1
ADGRE5
LGALS3BP
PSMB9
DDX60
IRF1
GBP1
GBP2
ISG20
NRP2
SP100
HLA−A
HLA−C
CAMK2D
TPST1
PTPN9
ARID5B
PTPN12
VRK2
GSN
GIPC1
HDAC4
SRPK2
IFI16
BCL6
NEAT1
AKAP6
CTBP2
HSP90B1
STAT2
STAT1
PARP9
APOL2
PARP14
B2M
NLRC5
SOX2
SLC27A1
FOXO1
PTPN13
2
1
0
−1
d
GO analysis of transcriptional changes in excitatory neurons
Upregulated
Downregulated
Number of genes
50
100
150
200
p-adj
0.01
0.04
glycolytic process
ATP generation from ADP
Ca−dependent cell−cell adhesion via cell adhesion molecules
regulation of gene expression, epigenetic
perinuclear region of cytoplasm
protein−containing complex binding
regulation of small molecule metabolic process
presynaptic membrane
modulation of chemical synaptic transmission
regulation of trans−synaptic signaling
chemical synaptic transmission
anterograde trans−synaptic signaling
synaptic signaling
trans−synaptic signaling
cell communication
signaling
organelle membrane
cellular response to stimulus
cell−cell signaling
intrinsic component of membrane
integral component of membrane
negative regulation of biological process
intrinsic component of plasma membrane
integral component of plasma membrane
GeneRatio
e
GO analysis of transcriptional changes in microglia
Upregulated
Downregulated
Number of genes
10
20
30
40
50
p-adj
0.005
0.020
response to temperature stimulus
response to heat
cellular response to heat
condensed chromosome
protein folding
cellular response to DNA damage stimulus
cell cycle process
mitotic cell cycle
mitotic cell cycle process
cell cycle
cell division
chromosome segregation
steroid hormone receptor activity
nuclear receptor activity
glucocorticoid receptor binding
response to corticotropin−releasing hormone
cellular resp. to corticotropin−releasing hormone stim.
ligand−activated transcription factor activity
chromosome organization
organelle organization
GeneRatio
f
Cell cycle scoring
G2/M phase
Control
TBI
UMAP 2
UMAP 1
1.0
0.5
0.0
S phase
Control
TBI
UMAP 2
UMAP 1
0.6
0.3
0.0
g
GO analysis of transcriptional changes in interneurons
Number of genes
64
68
72
76
80
0.03
Downregulated
intrinsic component of plasma membrane
intrinsic component of membrane
integral component of membrane
GeneRatio
h
GO analysis of transcriptional changes in astrocytes
Upregulated
Downregulated
Number of genes
50
100
150
200
250
p-adj
0.01
0.04
integral component of presynaptic membrane
integral component of synaptic membrane
ion channel regulator activity
ephrin receptor activity
negative regulation of T cell differentiation
negative regulation of lymphocyte differentiation
negative regulation of alpha−beta T cell activation
channel regulator activity
postsynaptic membrane
synaptic membrane
extracellular space
cellular response to organic substance
cellular response to chemical stimulus
response to organic substance
response to chemical
cell communication
cellular anatomical entity
cellular response to stimulus
immune system process
cellular_component
GeneRatio

Suppl. Figure 2 a) Violin plots showing enrichment scores of immune-related genes per time point of injury (GO terms: Interferon-gamma-mediated signaling pathway, Response to interferon-gamma, Cellular response to interferon-gamma, Innate immune response, Cellular-response to cytokine stimulus, Response to cytokine, and Defense response). b) Gene set enrichment analysis (GSEA) of differentially-expressed genes (padj < 0.01) found in OPCs. c) Expression of immune-related genes shown in Figure 2a in OPCs. d) Gene set enrichment analysis (GSEA) of differentially-expressed genes (padj < 0.01) found in excitatory neurons and e) microglia. f) Enrichment score of cell-cycle-related genes (top G2/M phase, bottom S phase). g) Gene set enrichment analysis (GSEA) of differentially-expressed genes (padj < 0.01) found in interneurons and h) astrocytes.



Suppl. Figure 3 a) Expression of LTR subfamilies in the cell types not shown in Figure 3d. b) Expression of LINE1 subfamilies in the cell types not shown in Figure 3e.

References

1. Cole, J.H.*, et al.* Spatial patterns of progressive brain volume loss after moderate-severe traumatic brain injury. *Brain* **141**, 822-836 (2018).
2. Johnson, V.E.*, et al.* Inflammation and white matter degeneration persist for years after a single traumatic brain injury. *Brain* **136**, 28-42 (2013).
3. Graham, N.S.N.*, et al.* Diffuse axonal injury predicts neurodegeneration after moderate-severe traumatic brain injury. *Brain* **143**, 3685-3698 (2020).
4. Raj, R., Kaprio, J., Jousilahti, P., Korja, M. & Siironen, J. Risk of Dementia After Hospitalization Due to Traumatic Brain Injury: A Longitudinal, Population-Based Study. *Neurology* (2022).
5. Loane, D.J., Kumar, A., Stoica, B.A., Cabatbat, R. & Faden, A.I. Progressive neurodegeneration after experimental brain trauma: association with chronic microglial activation. *J Neuropathol Exp Neurol* **73**, 14-29 (2014).
6. Ramlackhansingh, A.F.*, et al.* Inflammation after trauma: microglial activation and traumatic brain injury. *Ann Neurol* **70**, 374-383 (2011).
7. Ritzel, R.M.*, et al.* Chronic Alterations in Systemic Immune Function after Traumatic Brain Injury. *J Neurotrauma* **35**, 1419-1436 (2018).
8. Barrett, J.P.*, et al.* Interferon-beta Plays a Detrimental Role in Experimental Traumatic Brain Injury by Enhancing Neuroinflammation That Drives Chronic Neurodegeneration. *J Neurosci* **40**, 2357-2370 (2020).
9. Flygt, J.*, et al.* Neutralization of Interleukin-1beta following Diffuse Traumatic Brain Injury in the Mouse Attenuates the Loss of Mature Oligodendrocytes. *J Neurotrauma* **35**, 2837-2849 (2018).

10. Xing, J.*, et al.* Single-Cell RNA Sequencing Reveals Cellular and Transcriptional Changes Associated With Traumatic Brain Injury. *Front Genet* **13**, 861428 (2022).

11. Raabe, F.J.*, et al.* Expression of Lineage Transcription Factors Identifies Differences in Transition States of Induced Human Oligodendrocyte Differentiation. *Cells* **11**(2022).

12. Floriddia, E.M.*, et al.* Distinct oligodendrocyte populations have spatial preference and different responses to spinal cord injury. *Nat Commun* **11**, 5860 (2020).

13. Seeker, L.A. & Williams, A. Oligodendroglia heterogeneity in the human central nervous system. *Acta Neuropathol* **143**, 143-157 (2022).

14. Falcao, A.M.*, et al.* Disease-specific oligodendrocyte lineage cells arise in multiple sclerosis. *Nat Med* **24**, 1837-1844 (2018).

15. Meijer, M.*, et al.* Epigenomic priming of immune genes implicates oligodendroglia in multiple sclerosis susceptibility. *Neuron* **110**, 1193-1210 e1113 (2022).

16. Jonsson, M.E., Garza, R., Johansson, P.A. & Jakobsson, J. Transposable Elements: A Common Feature of Neurodevelopmental and Neurodegenerative Disorders. *Trends Genet* **36**, 610-623 (2020).

17. Jonsson, M.E.*, et al.* Activation of endogenous retroviruses during brain development causes an inflammatory response. *EMBO J* **40**, e106423 (2021).

18. Hurst, T.P. & Magiorkinis, G. Activation of the innate immune response by endogenous retroviruses. *J Gen Virol* **96**, 1207-1218 (2015).

19. Ishak, C.A., Classon, M. & De Carvalho, D.D. Deregulation of Retroelements as an Emerging Therapeutic Opportunity in Cancer. *Trends Cancer* **4**, 583-597 (2018).

20. Saleh, A., Macia, A. & Muotri, A.R. Transposable Elements, Inflammation, and Neurological Disease. *Front Neurol* **10**, 894 (2019).

21. Abu Hamdeh, S.*, et al.* Differential DNA Methylation of the Genes for Amyloid Precursor Protein, Tau, and Neurofilaments in Human Traumatic Brain Injury. *J Neurotrauma* **38**, 1679-1688 (2021).

22. Kobayashi, K.S. & Van Den Elsen, P.J. NLRC5: a key regulator of MHC class I-dependent immune responses. *Nature Reviews Immunology* **12**, 813-820 (2012).

23. Paul, P.*, et al.* A Genome-wide Multidimensional RNAi Screen Reveals Pathways Controlling MHC Class II Antigen Presentation. *Cell* **145**, 268-283 (2011).

24. Jin, Y., Tam, O.H., Paniagua, E. & Hammell, M. TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics* **31**, 3593-3599 (2015).

25. Sperber, G.O., Airola, T., Jern, P. & Blomberg, J. Automated recognition of retroviral sequences in genomic data—RetroTector©. *Nucleic Acids Research* **35**, 4964-4976 (2007).

26. Holmin, S., Soderlund, J., Biberfeld, P. & Mathiesen, T. Intracerebral inflammation after human brain contusion. *Neurosurgery* **42**, 291-298; discussion 298-299 (1998).

27. Marklund, N.*, et al.* Tau aggregation and increased neuroinflammation in athletes after sports-related concussions and in traumatic brain injury patients - A PET/MR study. *Neuroimage Clin* **30**, 102665 (2021).

28. Todd, B.P.*, et al.* Traumatic brain injury results in unique microglial and astrocyte transcriptomes enriched for type I interferon response. *J Neuroinflammation* **18**, 151 (2021).

29. Witcher, K.G.*, et al.* Traumatic Brain Injury Causes Chronic Cortical Inflammation and Neuronal Dysfunction Mediated by Microglia. *J Neurosci* **41**, 1597-1616 (2021).

30. Giacci, M.K., *et al.* Oligodendroglia Are Particularly Vulnerable to Oxidative Damage after Neurotrauma In Vivo. *J Neurosci* **38**, 6491-6504 (2018).

31. Xu, G.Y., Liu, S., Hughes, M.G. & McAdoo, D.J. Glutamate-induced losses of oligodendrocytes and neurons and activation of caspase-3 in the rat spinal cord. *Neuroscience* **153**, 1034-1047 (2008).

32. Dent, K.A., *et al.* Oligodendrocyte birth and death following traumatic brain injury in adult mice. *PLoS One* **10**, e0121541 (2015).

33. Flygt, J., Djupsjo, A., Lenne, F. & Marklund, N. Myelin loss and oligodendrocyte pathology in white matter tracts following traumatic brain injury in the rat. *Eur J Neurosci* **38**, 2153-2165 (2013).

34. Mierzwa, A.J., Marion, C.M., Sullivan, G.M., McDaniel, D.P. & Armstrong, R.C. Components of myelin damage and repair in the progression of white matter pathology after mild traumatic brain injury. *J Neuropathol Exp Neurol* **74**, 218-232 (2015).

35. Flygt, J., *et al.* Human Traumatic Brain Injury Results in Oligodendrocyte Death and Increases the Number of Oligodendrocyte Progenitor Cells. *J Neuropathol Exp Neurol* **75**, 503-515 (2016).

36. Jern, P. & Coffin, J.M. Effects of retroviruses on host genome function. *Annu Rev Genet* **42**, 709-732 (2008).

37. Andrews, W.D., *et al.* Detection of reverse transcriptase activity in the serum of patients with motor neurone disease. *J Med Virol* **61**, 527-532 (2000).

38. Douville, R., Liu, J., Rothstein, J. & Nath, A. Identification of active loci of a human endogenous retrovirus in neurons of patients with amyotrophic lateral sclerosis. *Ann Neurol* **69**, 141-151 (2011).

39. Garson, J.A., *et al.* Quantitative analysis of human endogenous retrovirus-K transcripts in postmortem premotor cortex fails to confirm elevated expression of HERV-K RNA in amyotrophic lateral sclerosis. *Acta Neuropathol Commun* **7**, 45 (2019).

40. Guo, C., *et al.* Tau Activates Transposable Elements in Alzheimer's Disease. *Cell Rep* **23**, 2874-2880 (2018).

41. Li, W., Jin, Y., Prazak, L., Hammell, M. & Dubnau, J. Transposable elements in TDP-43-mediated neurodegenerative disorders. *PLoS One* **7**, e44099 (2012).

42. Li, W., *et al.* Human endogenous retrovirus-K contributes to motor neuron disease. *Sci Transl Med* **7**, 307ra153 (2015).

43. MacGowan, D.J., *et al.* A controlled study of reverse transcriptase in serum and CSF of HIV-negative patients with ALS. *Neurology* **68**, 1944-1946 (2007).

44. Mayer, J., *et al.* Transcriptional profiling of HERV-K(HML-2) in amyotrophic lateral sclerosis and potential implications for expression of HML-2 proteins. *Mol Neurodegener* **13**, 39 (2018).

45. Steele, A.J., *et al.* Detection of serum reverse transcriptase activity in patients with ALS and unaffected blood relatives. *Neurology* **64**, 454-458 (2005).

46. Krug, L., *et al.* Retrotransposon activation contributes to neurodegeneration in a Drosophila TDP-43 model of ALS. *PLoS Genet* **13**, e1006635 (2017).

47. Chuong, E.B., Elde, N.C. & Feschotte, C. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science* **351**, 1083-1087 (2016).

48. Chang, H.Y. & Dubnau, J. Endogenous Retroviruses and TDP-43 Proteinopathy Form a Sustaining Feedback to Drive the Intercellular Spread of Neurodegeneration. *bioRxiv* (2022).

49. Svedung Wettervik, T.*, et al.* High Arterial Glucose is Associated with Poor Pressure Autoregulation, High Cerebral Lactate/Pyruvate Ratio and Poor Outcome Following Traumatic Brain Injury. *Neurocrit Care* **31**, 526-533 (2019).

50. Korsunsky, I.*, et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* **16**, 1289-1296 (2019).

51. Yu, G., Wang, L.G., Han, Y. & He, Q.Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287 (2012).

52. Dobin, A.*, et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).

53. Jin, Y., Tam, O.H., Paniagua, E. & Hammell, M. TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics* **31**, 3593-3599 (2015).

54. Hayward, A., Cornwallis, C.K. & Jern, P. Pan-vertebrate comparative genomics unmasks retrovirus macroevolution. *Proc Natl Acad Sci U S A* **112**, 464-469 (2015).

55. Liao, Y., Smyth, G.K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).

56. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**(2014).