Genetics and Genomics
Neuroscience

Steady-state neuron-predominant LINE-1 encoded ORF1p protein and LINE-1 RNA increase with aging in the mouse and human brain

Reviewed Preprint
v1 • September 6, 2024
Not revised

Tom Bonnifet, Sandra Sinnassamy, Olivia Massiani-Beaudoin, Philippe Mailly, Héloïse Monnet, Damarys Loew, Berangère Lombard, Nicolas Servant, Rajiv L Joshi, Julia Fuchs

CIRB, Collège de France, Université PSL, INSERM, CNRS, 75005, Paris, France • Orion Technological Core, CIRB, Collège de France, Université PSL, INSERM, CNRS, 75005, Paris, France • Institut Curie, Université PSL, Centre de Recherche, CurieCoreTech Spectrométrie de Masse Protéomique, 26 rue d'Ulm, Paris 75248 Cedex 05, France • Institut Curie, INSERM U900, Mines Paris Tech, Université PSL

Abstract

Recent studies have established a reciprocal causal link between aging and the activation of transposable elements, characterized in particular by a de-repression of LINE-1 retrotransposons. These LINE-1 elements represent 21% of the human genome, but only a minority of these sequences retain the coding potential essential for their mobility. LINE-1 encoded proteins can induce cell toxicity implicated in aging and neurodegenerative diseases. However, our knowledge of the expression and localization of LINE-1-encoded proteins in the central nervous system is limited. Using a novel approach combining atlas-based brain mapping with deep-learning algorithms on large-scale pyramidal brain images, we unveil a heterogeneous, neuron-predominant and widespread ORF1p expression throughout the murine brain at steady-state. In aged mice, ORF1p expression increases significantly which is corroborated in human post-mortem dopaminergic neurons by an increase in young LINE-1 elements including those with open reading frames. Mass spectrometry analysis of endogenous mouse ORF1p revealed novel, neuron-specific protein interactors. These findings contribute to a comprehensive description of the dynamics of LINE-1 and ORF1p expression in the brain at steady-state and in aging and provide insights on ORF1p protein interactions in the brain.

eLife assessment

Bonnifet et al. present data on the expression and interacting partners of the transposable element L1 in the mammalian brain. The work includes **important** findings addressing the potential role of L1 in aging and neurodegenerative disease. However, several aspects of experimental evidence presented are preliminary and the study remains **incomplete** in its current form.

https://doi.org/10.7554/eLife.100687.1.sa4

Introduction

Only about 2 % of the human genome are DNA sequences that will be translated into protein. The remaining 98 % are comprised of introns, regulatory elements, non-coding RNA, pseudogenes and repetitive elements including transposable elements. However, some sequences in what is generally considered “non-coding genome” do in fact contain sequences which encode proteins. This is true for specific lncRNAs which can encode peptides or functional proteins [1] but also for a few copies of two transposable element families, Long INterspersed Element-1 (LINE-1) and Human Endogenous RetroViruses (HERV). Non-functional copies of retrotransposons, to which LINE-1 and HERV belong, cover about 44% [2] of the human genome as remnants of an evolutionary ancient activity. Depending on the source, about 100 [3] to 146 [4] full-length LINE-1 elements with two open reading frames encoding ORF1 and ORF2 are present in the Human reference genome (GRCh38 Genome Assembly) and several incomplete HERV sequences encoding either or any combination of envelope (env), gag, pro or pol [5]. The LINE-1 encoded protein ORF1p, an RNA binding protein with “cis” preference [6,7], and ORF2p, an endonuclease and reverse transcriptase [8,9] are required for the mobility of LINE-1 elements. As many other transposable elements (TEs) including HERVs, LINE-1 elements are repressed by multiple cellular pathways. It was thus generally thought that TEs are repressed in somatic cells with no expression at steady-state [10–12]. However, the aging process reduces the reliability of these repressive mechanisms [13]. It is now, 31 years after the initial proposition of the “transposon theory of aging” by Driver and McKechnie [14], generally accepted that TE activation can be both, a cause and a consequence of aging [15,16].

Sparse data has shown that the LINE-1 encoded protein ORF1p is expressed at steady-state in the mouse ventral midbrain [17], the mouse hippocampus [18] and in some regions of the human post-mortem brain [19] and recent data informed about the presence of full-length transcripts in cancer cells, human epithelial cells and mouse hippocampal neurons [20]. Repression of LINE-1 might thus be incomplete and if so, it remains unclear how cells then prevent cell toxicity associated with LINE-1 encoded protein activity. Indeed, LINE-1 encoded proteins have been demonstrated to induce genomic instability (ORF2p endonuclease-mediated [17,21–26]) and inflammation (ORF2p reverse transcriptase-mediated [27–29]) and these cellular activities might be causally related to organismal aging, cancer, autoimmune and neurological diseases [30]. For instance, LINE-1 activation can drive neurodegeneration of mouse dopaminergic neurons [17], of drosophila neurons [31,32] and of mouse Purkinje neurons [33] which can be at least partially rescued with nucleoside analogue reverse transcriptase inhibitors (NRTIs) or other anti-LINE-1 strategies. NRTIs are currently being tested in several clinical trials designed to target either the RT of HERVs or the RT encoded by the LINE-1 ORF2 protein. It is not known today, however, to which extend LINE-1 encoded proteins are expressed at steady-state throughout the mouse and human brain, whether there is cell-type specificity and whether activation of LINE-1 encoded proteins is associated with brain aging or human neurodegeneration. Here, using a deep-learning assisted cellular detection methodology applied to pyramidal large-scale images of the mouse brain mapped to the Allen mouse brain atlas combined with post-mortem human brain imaging, co-IP mass spectrometry and transcriptomic analysis of LINE-1 expression, we describe a brain-wide map of ORF1p expression and interacting proteins at steady-state and in the context of aging. We find a heterogenous but widespread expression of ORF1p in the mouse brain with predominant expression in neurons. In aged mice, neuronal ORF1p expression increases brain-wide and in some brain regions to up to 27%. In human dopaminergic neurons, young LINE-1 transcripts and specific full-length and coding LINE-1 copies are increased in aged individuals. We further describe endogenous mouse ORF1p interacting proteins revealing known interactors and unexpected interacting proteins belonging to GO categories related to RNA metabolism, chromatin remodeling, cytoskeleton and the synapse.

Results

Widespread and heterogenous expression of the LINE-1 encoded ORF1p protein in the wildtype mouse brain

To investigate the expression pattern and intensities of endogenous LINE-1 encoded ORF1p protein throughout the entire mouse brain, we devised a deep-learning assisted cellular detection methodology applied to pyramidal large-scale images using a comprehensive workflow complemented by an approach based on confocal imaging as schematized in **Figure 1A**. Briefly, starting from sagittal slide scanner images of the mouse brain, we defined anatomical brain regions by mapping the Allen Brain Atlas onto the slide scanner images using Aligning Big Brains & Atlases (ABBA). We then employed a deep-learning detection method to identify all cell nuclei (Hoechst) and categorize all detected cells into neuronal cells (NeuN+) or non-neuronal cells (NeuN-) and ORF1p-expressing cells (ORF1p+) or cells that do not express ORF1p (ORF1p-). This workflow allowed us then to characterize the cell identity of ORF1p+ cells and ORF1p intensity throughout the whole brain but also in specific anatomical regions. In parallel, we completed the approach using confocal microscopy on selected anatomical regions allowing for comparison with higher resolution. Importantly, the specificity of the ORF1p antibody, a widely used, commercially available antibody [18,34–38], was confirmed by blocking the ORF1p antibody with purified mouse ORF1p protein resulting in the complete absence of immunofluorescence staining (**Suppl Fig. 1A**), by using an in-house antibody against mouse ORF1p[17] which colocalized with the anti-ORF1p antibody used (**Suppl Fig. 1B**, quantified in **Suppl Fig. 1C**), and by immunoprecipitation and mass spectrometry used in this study (see below). Unexpectedly, we found a generalized and widespread expression of ORF1p throughout the brain of wildtype mice (**Fig. 1B**; Swiss OF1 mice, three months-old; whole brain except regions with particularly high cellular density (cerebellum, hippocampus, olfactory bulb) which impedes nuclei detection by deep-learning. ORF1p is detectable in all regions and subregions analyzed with heterogenous expression patterns (density and intensity) per region/subregions. The ten regions shown in **Figure 1B** exemplify visible different densities of ORF1p+ cells with varying levels of expression. Notably, the expression pattern of ORF1p in the hippocampus is similar to what has recently been published [18] (**Fig. 1B**, panel 2). Throughout the entire brain, the mean density of ORF1p+ cells per mm² was ≈ 305 ±18 (mean ±SEM), representing up to 20% of all detected cells (**Fig. 1C**). ORF1p+ cells in each mouse brain analyzed showed up to eight-fold disparities in intensity between low- and high-expressed cells (**Fig. 1C**). We then quantified nine anatomical regions according to the Allen Brain Atlas on four brains of three-month old mice (**Fig. 1D**) using the automated workflow (**Fig. 1A**) with regard to cell density (**Fig. 1D**), cell proportions (**Fig. 1E**) and fluorescent intensity of ORF1p+ cells (**Fig. 1F**). This approach permitted the analysis of about 10 000 ORF1p+ cells per animal highlighting the power of our large-scale analysis. Densities of ORF1p+ cells ranged from the lowest density in the hindbrain with 154 ±19 cells per mm^2 (mean ±SEM) to the highest density of ORF1p+ cells in the isocortex with 451 ±44 cells per mm^2 (mean ±SEM) and the thalamus with 446 ±50 cells per mm^2 (mean ±SEM). The proportion of ORF1p+ cells per anatomical region fluctuated between 10% ±2.1 (ventral striatum, mean ±SEM) and 31% ±1.6 (thalamus, mean ±SEM). The dorsal striatum ("striatal dorsal" in the Allen Brain Atlas denomination) exhibited the lowest ORF1p expression intensity (658 ±3 mean ±SEM) of all regions tested, the hindbrain the highest mean intensity of ORF1p per cell (mean ±SD 1221 ±548) as illustrated in **Figure 1B** and quantified in **Figure 1D** and **1F**. Interestingly, cell density did not correlate with expression levels. Dorsal and ventral striatum for instance displayed similar ORF1p intensities per cell but exhibited significant differences in ORF1p cell density and proportion. The "midbrain motor" region as defined by the Allen Brain Atlas showed an intermediate cell density (mean ±SEM 265 ±16 cells per mm^2) and a rather high ORF1p expression intensity (mean ±SEM 1006 ±533). Statistical analysis comparing mean density of ORF1p+cells per mm^2 or mean intensity per ORF1p+ cells among regions confirmed the

heterogeneity concerning ORF1p expression throughout the mouse brain (**Fig. 1D**, 1F). Slide scanner and confocal images revealed an exceptionally high ORF1p expression intensity in the ventral region of the midbrain which we identified as the *Substantia nigra pars compacta (SNpc)*. This region displayed an important density of ORF1p+ cells and a comparatively high level of ORF1p expression as illustrated by confocal imaging (**Fig. 1B**, panel 8), but could not be quantified independently with our brain-wide approach due to the geometrically-complex anatomy of this region and its small size (subregion-level in the Allen Brain Atlas hierarchy). Another region which could not be included in our brain-wide analysis was the cerebellum due to its extremely high density of cell nuclei. However, slide scanner and confocal imaging (**Fig. 1B**, panel 10) revealed that ORF1p is expressed in Purkinje cells, while not detectable in the molecular or granular layers.

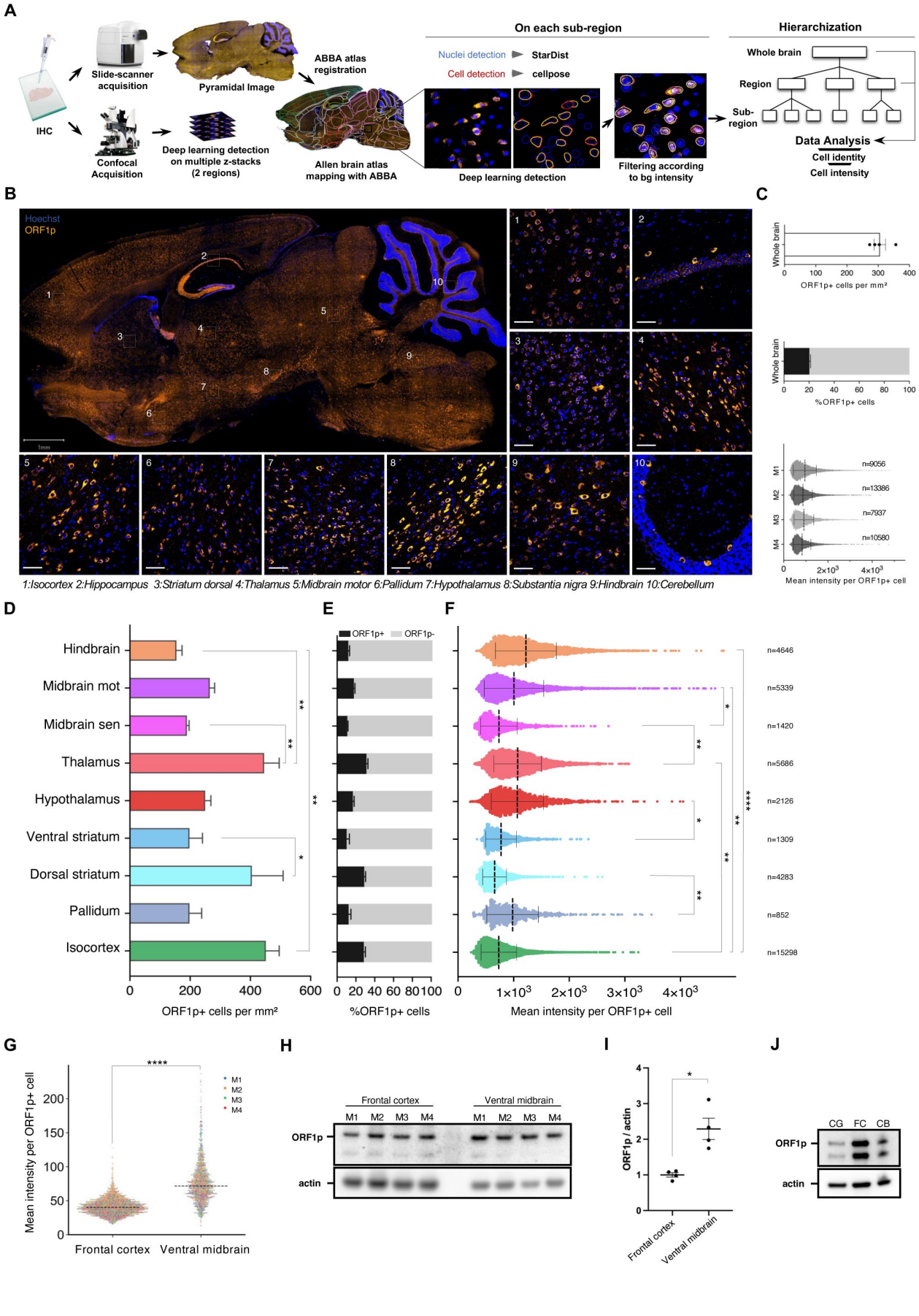

A
IHC
Slide-scanner acquisition
Pyramidal Image
ABBA atlas registration
Confocal Acquisition
Deep learning detection on multiple z-stacks (2 regions)
Allen brain atlas mapping with ABBA
On each sub-region
Nuclei detection
StarDist
Cell detection
cellpose
Deep learning detection
Filtering according to bg intensity
Hierarchization
Whole brain
Region
Sub-region
Data Analysis
Cell identity
Cell intensity
B
Hoechst
ORF1p
1mm
1:Isocortex 2:Hippocampus 3:Striatum dorsal 4:Thalamus 5:Midbrain motor 6:Pallidum 7:Hypothalamus 8:Substantia nigra 9:Hindbrain 10:Cerebellum
C
Whole brain
ORF1p+ cells per mm²
%ORF1p+ cells
Mean intensity per ORF1p+ cell
n=9056
n=13386
n=7937
n=10580
D
Hindbrain
Midbrain mot
Midbrain sen
Thalamus
Hypothalamus
Ventral striatum
Dorsal striatum
Pallidum
Isocortex
ORF1p+ cells per mm²
E
ORF1p+
ORF1p-
%ORF1p+ cells
F
Mean intensity per ORF1p+ cell
n=4646
n=5339
n=1420
n=5686
n=2126
n=1309
n=4283
n=852
n=15298
G
Mean intensity per ORF1p+ cell
Frontal cortex
Ventral midbrain
M1
M2
M3
M4
H
Frontal cortex
Ventral midbrain
ORF1p
actin
I
ORF1p / actin
Frontal cortex
Ventral midbrain
J
CG
FC
CB
ORF1p
actin

Fig. 1

Widespread and heterogenous expression of ORF1p protein in the mouse brain

(A) Schematic representation of the unbiased cell detection pipeline on large scale and confocal images. Immunofluorescent images on sagittal mouse brain slices were acquired on a digital pathology slide-scanner or on a confocal microscope (DNA stain: Hoechst, neuronal marker; NeuN, protein of interest: ORF1p). Pyramidal images aquired with the slide scanner were then aligned with the hierarchical anatomical annotation of the Allen Brain Atlas using ABBA. Once the regions defined, a deep-learning based detection of cell nuclei (Hoechst staining, Stardist) and cell cytoplasm (NeuN staining, Cellpose) was performed on each sub-region of the atlas. Objects were filtered according to the background intensity measured in each sub-region for each channel (NeuN and ORF1p). The identity and intensity measures were analyzed at the regional and whole brain level. In parallel, confocal images (multiple z-stacks) of two selected regions (frontal cortex and ventral midbrain) were also acquired and identity and intensity were quantified using Cellpose and Stardist.

(B) Widespread and heterogenous expression of the LINE-1 encoded protein ORF1p in the mouse brain. Representative image of ORF1p immunostaining (orange) of a sagittal section of the brain of a young (three months-old) mouse acquired on a slide scanner. Scale bar = 1mm. (1-10) Representative images of immunostainings showing ORF1p expression (orange) in 10 different regions of the mouse brain acquired on a confocal microscope. Nuclei are represented in blue (Hoechst), scale bar = 50µm. (1) Isocortex, (2) Hippocampus, (3) Striatum dorsal, (4) Thalamus, (5) Midbrain motor, (6) Pallidum, (7) Hypothalamus, (8) Substantia nigra pars compacta, (9) Hindbrain, (10) Cerebellum. ORF1p expression profile in the mouse brain. The entire mouse brain with the exception of the olfactory bulb and the cerebellum were analyzed according to the pipeline on large-scale images described in (A). Bar plot showing the total number of ORF1p+ cells per mm^2 in the mouse brain. Data is represented as mean ±SEM, n=4 mice (top). Bar plot indicating the proportion of ORF1p+ cells compared to all cells detected. Data is represented as mean ±SEM, n=4 mice, 202001 total cells analyzed (middle). Scatter plot showing the mean intensity of ORF1p per ORF1p+ cell. Data is represented as mean ±SD, n=4 mice, 40999 ORF1p+ cells analyzed (bottom).

(D-F) ORF1p expression profile (density, proportion and expression) in defined anatomical regions of the mouse brain. Nine anatomical regions as defined by the Allen Brain Atlas and mapped onto sagittal brain slices (four three-month-old Swiss/ OF1) with ABBA were analyzed using the pipeline on large scale images described in (A). (D) ORF1p+ cell density in 9 different regions. Bar plot showing the number of ORF1p+ cells per mm^2. Data is represented as mean ±SEM; *p<0.05; **p<0.01; adjusted p-value, one-way ANOVA followed by a Benjamin-Hochberg test (E) Proportion of ORF1p positive cells in 9 different regions. Bar plot showing the proportion of ORF1p+ cells among all cells detected per region. Data is represented as mean ±SEM. (F) Mean ORF1p expression per cell in 9 different regions. Dot plot showing the mean intensity of ORF1p signal per ORF1p+ cell in 9 different regions. Data is represented as mean ±SD. The number of analyzed cells per region is indicated in the figure. *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001; adjusted p-value, nested one-way ANOVA followed by Sidak' multiple comparison test.

(G) ORF1p expression in the frontal cortex and ventral midbrain. Confocal images with multiple z-stacks were analyzed using Cellpose and Stardist. Dot plot representing the mean expression of ORF1p per ORF1p+ cells. Individual (four three-month-old Swiss/ OF1) mice are represented each by a different color, the scattered line represents the median. ****p<0.0001, nested one-way ANOVA. Total cells analyzed = 4645.

(H-I) ORF1p expression in the frontal cortex and the ventral midbrain. (H) Western blots showing ORF1p (top) and actin expression (bottom) in four individual mice per region which were quantified in (I) using actin as a reference control. The signal intensity is plotted as the fold change of ORF1p expression in the ventral midbrain to ORF1p expression in the frontal cortex. *p<0.05; two-sided, unpaired student's-test.

(J) ORF1p expression in three regions of the human brain. Western blot showing ORF1p expression in the cingulate gyrus (CG), frontal cortex (FC) and cerebellum (CB) of post-mortem tissues from a healthy individual. ORF1p (Top), Actin (bottom).

In order to confirm ORF1p expression by an independent method, we performed Western blot analysis on six micro-dissected regions from the mouse brain (Swiss/OF1 mouse, three-month old). As shown in **Suppl Fig. 1D**, ORF1p is expressed in all six regions with varying expression levels confirming the overall presence of ORF1p throughout the brain. We then chose two regions with significantly divergent ORF1p expression intensities as identified and quantified on pyramidal large-scale images: the frontal cortex (low) and the ventral midbrain (intermediate to high). We confirmed a significant higher expression of ORF1p in the ventral midbrain compared to the frontal cortex using an approach based on the unbiased, automated quantification of multiple z-stacks using a confocal microscope (**Fig. 1G**) and by Western-blotting on micro-dissected regions (**Fig. 1H,I**). In concordance with the findings stemming from the large-scale image quantification pipeline (**Fig. 1F**), the ventral midbrain showed ≈ 2-times higher expression of ORF1p than the frontal cortex as quantified in **Figure 1G** (1.8-fold) and Figure I (2.3-fold) validating our cellular detection methodology for pyramidal large-scale imaging and underscoring the heterogeneity of ORF1p expression levels in the mouse brain. To investigate intra-individual expression patterns of ORF1p in the post-mortem human brain, we analyzed three brain regions of a neurologically-healthy individual (**Fig. 1J**) by Western blotting. ORF1p was expressed at different levels in the cingulate gyrus, the frontal cortex and the cerebellum underscoring a widespread expression of human ORF1p across the human brain. In summary, our findings reveal the consistent presence of ORF1p expression throughout the mouse brain in all anatomical regions analyzed with high regional variability in terms of density of ORF1p+ cells per mm2 per region and ORF1p+ cell intensity per region. This finding raises several questions concerning cell-type identity of ORF1p expressing cells and potential functions or consequences of ORF1p expression in the mouse and human brain at steady-state.

ORF1p is predominantly expressed in neurons

Following our observation of a wide-spread expression of endogenous ORF1p throughout the brain, we first addressed the question of the cellular identity of ORF1p+ cells. To this end, we used the neuron-specific marker NeuN, commonly used to identify post-mitotic neurons in the central nervous system [39]. This allowed us to determine the proportion of neuronal (NeuN+) or non-neuronal cells (NeuN-) expressing ORF1p (ORF1p+) or not (ORF1p-). Making use of our large-scale imaging approach (**Fig. 1A**), we observed drastic dissimilarities in detected cellular proportions between the white and grey brain matter. As expected, we observed only 1% of NeuN+ cells in the white matter (corpus callosum; **Fig. 2A**) validating both, the neuronal marker NeuN as such and the ABBA superposition of the Allen Brain Atlas onto the sagittal brain slices. In the grey matter, our approach detected 30.5% NeuN+ cells (dark red and yellow bars in **Fig. 2A**) which, according to the literature, should include all post-mitotic neurons with only minor exceptions [39–42] and corresponds to the reported proportion of neurons present in the mouse brain [43]. The nine identified grey matter regions in **Fig. 2A** display the proportions of the different cell types per region. The proportion of all cells in a given region which are positive for ORF1p (dark red bars) differed between regions (lowest proportion: hindbrain: 7%; highest proportion dorsal striatum: 26.6%). In the isocortex and the midbrain motor-related regions, the majority of neurons detected express ORF1p (54% and 59% by large-scale analysis, **Suppl Fig. 2A**; 68.7% and 68.8% by confocal imaging (**Fig. 2B**, quantified in C and **Suppl Fig. 2B**), respectively), while in the midbrain sensory related regions the proportion dropped to 25% whereas it reached 82% in the thalamus (**Suppl Fig. 2A**). Altogether, nearly half of all NeuN+ cells throughout the mouse brain expressed ORF1p (mean of all regions: 48.2%; **Suppl Fig. 2A**). Regarding the cell identity of ORF1p+ cells brain-wide, more than 70% were identified as neuronal by the large-scale approach (**Suppl Fig. 2C**). This contrasted somewhat with results obtained by the second approach using confocal imaging on multiple z-stacks which indicated that 91.3% (frontal cortex) and 88.5% (ventral midbrain) of ORF1p+ cells were neuronal (**Fig. 2D**). This difference in percentages of ORF1p+ expressing neurons among all neurons between the large-scale image cell detection methodology and the confocal workflow is most probably due to technical limitations inherent to our large-scale pipeline. Indeed, with the latter approach, region-dependent differences in cell density and signal intensity levels might be the cause for an underestimation of the proportion of

ORF1p+ cells being neuronal due to difficulties in cell detection by StarDist/Cellpose (high cell density) on a single focal plan, technical difficulties which are widely reduced by the multiple z-stack based approach when using a confocal microscope. Notably, frontal cortex and ventral midbrain present similar proportion of neurons expressing ORF1p and ORF1p although the percentage of NeuN+ cells between these two regions is significantly different (**Suppl Fig. 2D**). As we could not rule out that ORF1p might also be expressed in non-neuronal cells, we turned to non-neuronal markers specific for different glial cell populations using two different astrocytic markers (GFAP, Sox9), the astro- and oligodendrocytic marker S100β and the microglial marker Iba1[43,44] and performed co-staining with ORF1p followed by confocal imaging as illustrated in **Figure 2E**. We screened multiple images of frontal cortex, ventral midbrain, hippocampus and striatum and did not find a single ORF1p+ cell, which could unambiguously be defined as non-neuronal. This indicated that ORF1p is not or only very rarely expressed in non-neuronal cells. To further confirm the predominant presence of expression of ORF1p in neurons and the absence of ORF1p expression in non-neuronal cells, we used fluorescence-activated cell sorting (FACS) to isolate neurons (using a NeuN antibody) and non-neuronal cells (NeuN-) from the adult mouse brain followed by Western blotting with an antibody against ORF1p (**Fig. 2F**, 2G). We detected ORF1p exclusively in the neuronal population, confirming the results based on two different imaging approaches. Finally, to assess whether predominant, if not exclusive ORF1p expression in neurons is mouse brain specific or a pattern also applicable to the human brain, we investigated the identity of ORF1p expressing cells in the cingulate gyrus of a healthy human brain. Similar to what we found in the mouse brain, we observed sparse NeuN expression in the white (**Suppl Fig. 2E**) and extensive NeuN staining in the grey matter corresponding to the cortical layers (right, separated by a dashed white line) with ORF1p+ cells predominantly located in the grey matter (images shown in **Figure 2H** located in the grey matter). All cells stained by ORF1p were co-stained with NeuN indicating that ORF1p was expressed in neuronal cells in the human brain (**Fig. 2H**). However, due to the lower signal quality inherent to human post-mortem sections compared to mouse sections, the identity of ORF1p+ cells was estimated to be 80% neuronal by the automated image analysis pipeline, although no ORF1p+ / NeuN-cells could be clearly identified (**Fig. 2I**). Of all neurons identified, 37.2% were ORF1p+ (**Fig. 2J**), indicating that, similar to the mouse brain, only a fraction of neurons express ORF1p (**Fig. 2H**, right).

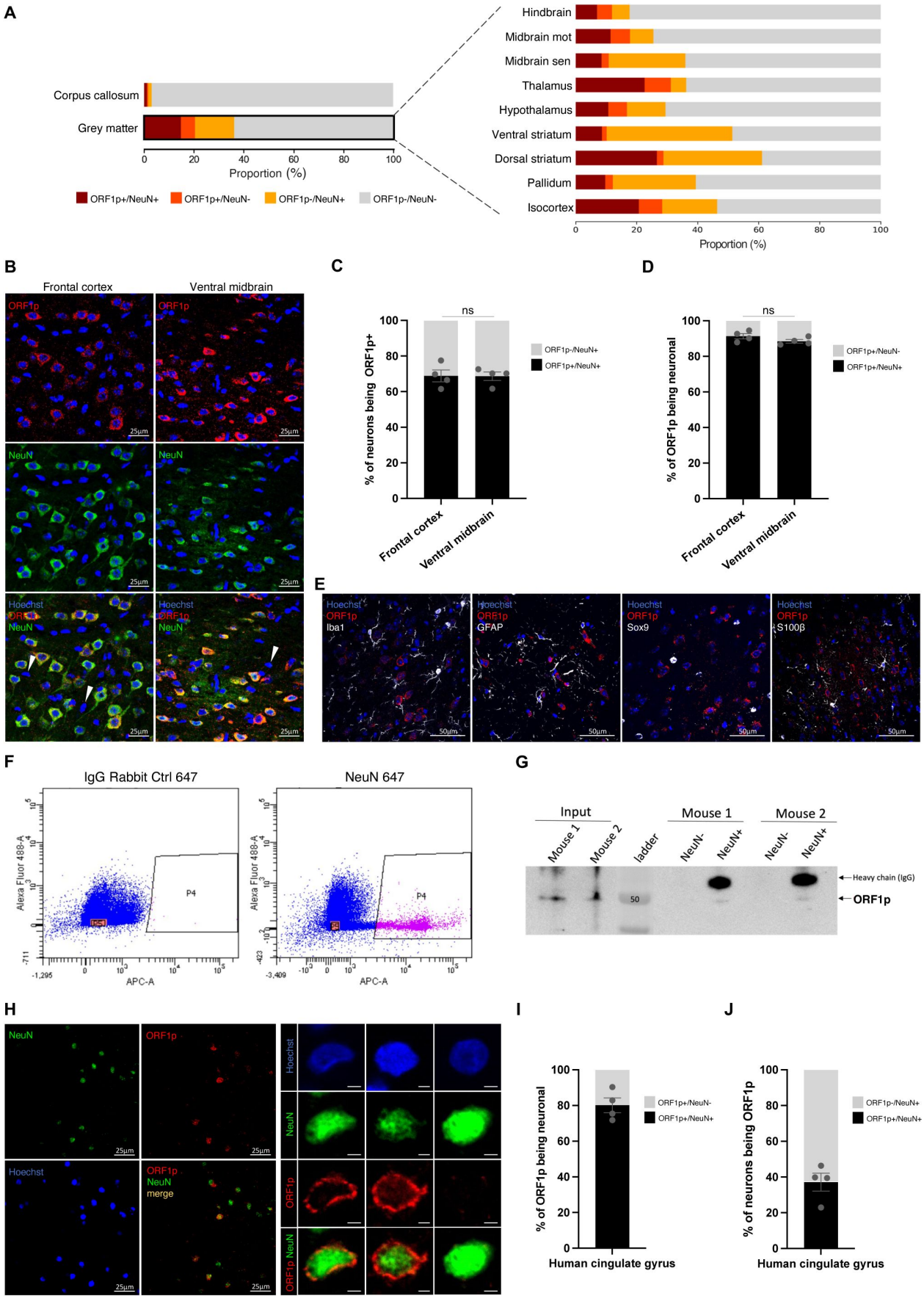

A
Corpus callosum
Grey matter
0 20 40 60 80 100
Proportion (%)
ORF1p+/NeuN+
ORF1p+/NeuN-
ORF1p-/NeuN+
ORF1p-/NeuN-
Hindbrain
Midbrain mot
Midbrain sen
Thalamus
Hypothalamus
Ventral striatum
Dorsal striatum
Pallidum
Isocortex
0 20 40 60 80 100
Proportion (%)
B
Frontal cortex
Ventral midbrain
ORF1p
NeuN
Hoechst
25µm
C
ns
% of neurons being ORF1p+
ORF1p-/NeuN+
ORF1p+/NeuN+
Frontal cortex
Ventral midbrain
D
ns
% of ORF1p being neuronal
ORF1p+/NeuN-
ORF1p+/NeuN+
Frontal cortex
Ventral midbrain
E
Hoechst
ORF1p
Iba1
GFAP
Sox9
S100β
50µm
F
IgG Rabbit Ctrl 647
NeuN 647
Alexa Fluor 488-A
APC-A
P4
G
Input
Mouse 1
Mouse 2
ladder
NeuN-
NeuN+
50
Heavy chain (IgG)
ORF1p
H
NeuN
ORF1p
Hoechst
merge
25µm
I
% of ORF1p being neuronal
ORF1p+/NeuN-
ORF1p+/NeuN+
Human cingulate gyrus
J
% of neurons being ORF1p
ORF1p-/NeuN+
ORF1p+/NeuN+
Human cingulate gyrus

Fig. 2

ORF1p is predominantly expressed in neurons in the mouse brain

(A) ORF1p expression is absent in the white matter (corpus callosum) and predominantly expressed in neurons. Proportion of ORF1p+/NeuN+, ORF1p+/NeuN-, ORF1p-/NeuN+ and ORF1p-/NeuN-cells in the white matter (corpus callosum) and the grey matter (left) and in nine different regions (right) analyzed by the cell detection pipeline on large scale images presented in Figure1A. Exact values can be found in Suppl_Table1.

(B) Representative confocal microscopy images showing ORF1p (red) and NeuN expression (green) in two different regions of the mouse brain. The bottom images show the merge of the two stainings, an overlap of both markers is represented in orange. z-projections; scalebar = 25µm.

(C) Proportion of neurons expressing ORF1p in the frontal cortex and ventral midbrain quantified on confocal images. ns: non-significant; chi-square test on the cell number of the different cell-types analyzed; n=4 mice, data is represented as mean ± SEM.

(D) Proportion of ORF1p+ cells identified as NeuN+ or NeuN- in two different regions, analyzed by confocal microscopy on multiple z-stacks. ns: non-significant; chi-square test, n=4 mice.

(E) ORF1p does not colocalize with glial or microglial cell markers. Representative confocal microscopy images showing ORF1p staining (red) and three different glial cell (GFAP, Sox9, S100β) or microglial (Iba1) markers (white). Note that Iba1 antibody (rabbit) was used with the ORF1p 09 antibody (guinea pig, in house) z-projections, scalebar = 50µm.

(F-G) Separation of neuronal and non–neuronal - cells by FACS confirms predominant neuronal expression of ORF1p. (F) Neuronal (NeuN+) and non-neuronal (NeuN-) cells isolated by fluorescent activated cell sorting (FACS). Dot plots showing autofluorescence versus an appropriate control antibody (IgG rabbit 647; left) and an antibody against NeuN (name of the AB 657, right). The P4 window represents isolated NeuN+ cells (pink) and the P5 fraction NeuN- cells (orange) containing the same number of cells as sorted in P4 for comparison, others NeuN- are represented in blue. (G) Western blot. ORF1p expression (top), in NeuN- and NeuN+ FACS-sorted cells stemming from Figure F.

(H) Representative confocal microscopy images showing ORF1p (red), NeuN (green) and Hoechst (blue) in the cingulate gyrus of the human brain. z-projection; scalebar = 25µm (left). Example of individuals neurons expressing ORF1p or not are shown on the right panel. z-projection; scalebar = 5 µm (right).

(I) Proportion of ORF1p+ cells identified as NeuN+ or NeuN- in the human cingulate gyrus, analyzed by confocal microscopy on multiple z-stacks.

(J) Proportion of neurons expressing ORF1p in the human cingulate gyrus, analyzed by confocal microscopy on multiple z-stacks.

In summary, ORF1p expression in the mouse and human brain is widely restricted to neurons of which a proportion express ORF1p. This raises the question of the function and consequences of ORF1p expression specifically in neurons but also on the dynamic regulation of this expression upon exogenous (exposome) or endogenous (aging) challenges.

ORF1p expression is increased in the aged mouse brain

ORF1p is expressed at steady-state throughout the brain, but if this expression is dynamically regulated is not known. Aging has unequivocally been linked to LINE-1 regulation [16,45] both as a trigger and as a consequence of LINE-1 activation but whether this is true for the brain has

not been thoroughly investigated with ORF1p as a read-out. We therefore addressed the question whether advanced age (16-month-old mice) was paralleled by an enhanced presence of ORF1p expression in the brain compared to young, three-month old mice. In the context of aging, we globally observed a reduction in the proportion of ORF1p+/NeuN+ cells using the cell detection workflow applied to large-scale images described in **Figure 1A**, phenomenon mainly driven by the midbrain motor, the dorsal striatum, the pallidum and the thalamus regions (**Fig. 3A**, dark red bars, Suppl_Table1). The confocal approach applied to two regions, the frontal cortex and the ventral midbrain (**Fig. 3B**) confirmed a loss of ORF1p+/NeuN+ cells in the ventral midbrain with no change in cell proportions in the frontal cortex in accordance with the large-scale imaging approach (**Fig. 3A**). The predominantly neuronal identity of ORF1p+ cells, however, was unchanged (**Suppl Fig. 3A**) just as the proportion of neurons expressing ORF1p (**Suppl Fig. 3B**). We observed a significant decrease of NeuN+ cells in the aged ventral midbrain (**Fig. 3B**, **Suppl Fig. 3C**). We next analyzed ORF1p expression levels in the context of brain aging. Interestingly, the mean intensity of ORF1p expression increased significantly with age throughout the brain (13% increase brain-wide; **Fig. 3C**). Frequency distribution analysis unveiled a shift in ORF1p mean expression per cell in aged mice (**Fig. 3D**). Importantly, the Hoechst mean intensity within nuclei of ORF1p+ cells, serving as an internal control, showed no significant change (**Fig. 3E**). Among nine analyzed regions, five demonstrated a general increase in ORF1p mean intensity per cell in aged mice ($p \leq 0.05$), a change independent from inter-individual variations in both young and aged mice (**Fig. 3F**). An increase of ORF1p expression was also observed in three others regions albeit not significant. The only exception was the isocortex which remained unchanged with aging. The general increase of ORF1p expression (fold change intensity) in the whole brain, reaching nearly a 30% increase in some regions, is represented on the heatmap in **Figure 1G**. These results were confirmed by the confocal imaging approach: ORF1p expression in the frontal cortex remained unchanged and the ventral midbrain region increased significantly in aged mice as quantified in **Figure 3H** and shown with a representative image in **Figure 3I**. Overall, these results highlight an age-dependent increase in ORF1p expression in neurons throughout the brain with some regions showing an increase of up to 27 % in ORF1p intensities.

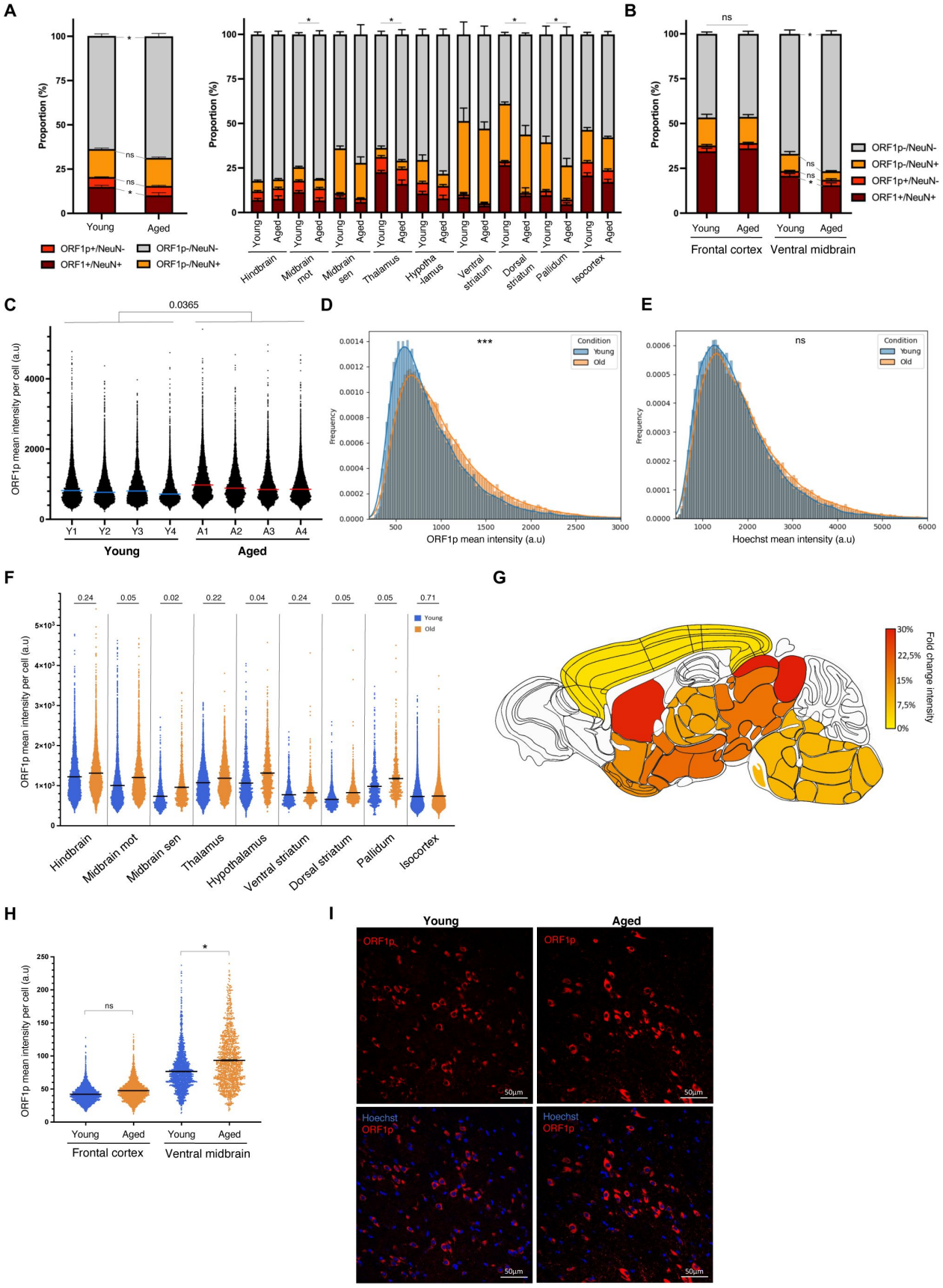

A
Proportion (%)
Young
Aged
ns
ORF1p+/NeuN-
ORF1+/NeuN+
ORF1p-/NeuN-
ORF1p-/NeuN+
Hindbrain
Midbrain mot
Midbrain sen
Thalamus
Hypotha -lamus
Ventral striatum
Dorsal striatum
Pallidum
Isocortex
B
Frontal cortex
Ventral midbrain
C
0.0365
ORF1p mean intensity per cell (a.u)
Y1
Y2
Y3
Y4
A1
A2
A3
A4
D
***
Condition
Frequency
ORF1p mean intensity (a.u)
E
Hoechst mean intensity (a.u)
F
0.24
0.05
0.02
0.22
0.04
0.24
0.05
0.05
0.71
G
Fold change intensity
30%
22,5%
15%
7,5%
0%
H
I
Hoechst
ORF1p
50µm

Fig. 3

ORF1p expression is increased throughout the whole mouse brain in the context of aging

(A) Proportion of ORF1p+/NeuN+, ORF1p+/NeuN+, ORF1p+/NeuN-, ORF1p-/NeuN+ and ORF1p-/NeuN-cell-type in the whole brain (left) and the different analyzed regions (right) of young (three-month aged, n=4) and aged (16-month-old, n=4) mice using the cell detection pipeline on large scale images presented in **Figure 1A**, data is represented as mean ± SEM. Exact values can be found in Suppl_Table1.

(B) Proportion of ORF1p+/NeuN+, ORF1p+/NeuN+, ORF1p+/NeuN-, ORF1p-/NeuN+ and ORF1p-/NeuN-cell-type in two different regions of young and aged mice, analyzed on multiple z-stack confocal images. ns: non-significant; *p<0.05 calculated using two-way ANOVA with sidak's multiple comparisons test on the cell number of the different cell-types analyzed; data is represented as mean ± SEM.

(C) ORF1p mean expression per ORF1p+ cell in the brain analyzed on large-scale images. Dot plot showing the ORF1p mean expression per ORF1p+ cell in young (n=4) and aged (n=4) mice in the whole brain (except cerebellum and olfactory bulb). 74985 total cells were analyzed; * p<0.05, two-way ANOVA with sidak's multiple comparisons; data is represented as mean ± SEM.

(D) Frequency distribution of ORF1p mean intensity in ORF1p+ cells. ***p<0.001, Kolmogorov-Smirnov test.

(E) Frequency distribution of Hoechst mean intensity in the nuclei of OrF1p+ cells. ns: non-significant, Kolmogorov-Smirnov test.

(F) Mean ORF1p expression per ORF1p+ cell in nine different anatomical regions. Dot plot showing the ORF1p mean expression per ORF1p positive cell (n=74985). Adjusted p-value are represented, two-tailed nested t-test followed by a Benjamin, Krieger and Yukutieli test; n=4 young and n=4 aged mice per region, data is represented as mean ± SEM.

(G) Color-coded representation of fold-changes of ORF1p expression with aging. Represented is the fold-change in percent (aged vs young) of the "mean of the mean" ORF1p expression per ORF1p+ cell quantified mapped onto the nine different regions analyzed as shown in (F).

(H) ORF1p expression is increased in the ventral midbrain of aged mice. Dot plot representing ORF1p expression in two different regions of young and aged mice analyzed on confocal images with multiple z-stacks; total cells analyzed = 8381 ns: non-significant *p<0.05, two-tailed one-way ANOVA; dashed lines represent the medians.

(I) Representative confocal microscopy acquisition showing increased ORF1p expression (red) in the ventral midbrain region of aged mice (one z plan is shown). Cell nuclei are shown in blue (Hoechst staining). Scalebar = 50µm.

Coding LINE-1 transcripts are increased in aged human dopaminergic neurons

Following the observation of increased ORF1p expression in the aged mouse brain, among which the ventral midbrain, and given the age-related susceptibility of dopaminergic neurons in the *SNpc* to cell death and to degeneration in PD [46], we turned to a RNA-seq dataset of laser-captured micro-dissected post-mortem human dopaminergic neurons of brain-healthy individuals [47], in order to interrogate full-length LINE-1 mRNA expression profiles as a function of age. To avoid read-length bias to which TE analysis is particularly sensitive, we analyzed only the data derived from 50bp paired-end reads of linearly amplified total RNA as this dataset represented all age categories (n=41; with ages ranging from 38 to 97; mean age: 79.88 (SD ±12.07); n=6 ≤65y; n=35

>65y; mean PMI: 7.07 (SD ±7.84), mean RIN: 7.09 (±0.94)). As age-related dysregulation of TEs might not be linear, we considered individuals with ages-at-death younger or equal to 65 years as "young" (n=6, 38-65 years, mean age 57.5 years (SD ±9.9)) and individuals older than 65 years as "aged" (n=35, 65-97 years, mean age 83 years (SD±7.8). The expression of the dopaminergic markers tyrosine hydroxylase (TH) and LMX1B were similar in both populations indicating no apparent change of dopaminergic identity of analyzed melanin-positive dopaminergic neurons (**Suppl Fig. 4A**). Next, we compared the expression of repeat elements at the class, family and name level based on the repeat masker annotation implemented in the UCSC genome browser using a commonly used mapping strategy for repeats [48]. No overt dysregulation of repeat elements at either level of repeat element hierarchy was observed (**Suppl Fig. 4C-F**), however there was a significant increase in several younger LINE-1 elements including L1HS and L1PA2 at the "name" level (**Fig. 4A,B**). This was not observed for HERVK-int, a human endogenous retrovirus family with some copies having retained coding potential (**Fig. 4B**) or other potentially active TEs like HERVH-int, HERV-Fc1 or SVA-F with the exception of a trend for an increase in AluYa5 transcripts in the >65y group, a young Alu family mobilized by the LINE-1 retrotransposition machinery (**Suppl Fig. 4G**). Interestingly, L1HS expression was highly correlated with L1PA2 expression and this correlation extended to almost all younger LINE-1 subfamilies weaning down with evolutionary distance (**Fig. 4C**). This was not true for other active TEs as L1HS was negatively correlated with HERVK-int expression (**Fig. 4C**). Several regulators of LINE-1 activity have been identified [17,49] and correlation of their expression with L1HS might allow to infer their relevance of interaction (activation or repression) with L1HS in human dopaminergic neurons. Spearman correlation analysis revealed three known repressors of LINE-1 activity; EN1 (Engrailed 1 [17], **Suppl Fig. 5A**) with important functions for dopaminergic neuron homeostasis [50], CBX5/HP1a, a heterochromatin binding protein binding to the histone mark H3K9me3, thereby mediating epigenetic repression [51] (**Suppl Fig. 5B**) and XRCC5/6, also known as Ku86/Ku70, which are essential for DNA double-stranded break repair through the nonhomologous end joining (NHEJ) pathway and limit LINE-1 full-length insertions [52] (**Suppl Fig. 5C**). The transcripts of these genes showed, although not statistically significant, a trend for decreased expression in the elderly (**Suppl Fig. 5D-G**). Based on the increase of young LINE-1 families L1HS and L1PA2 in aged human dopaminergic neurons and the finding that ORF1p was increased in the aged mouse brain, we focused our attention on LINE-1 elements with coding potential for ORF1 and ORF2 according to the L1Basev2 annotation which are specific elements comprised in the L1HS and L1PA2 annotation at the "name" level. Most of the 146 full-length and coding LINE-1 termed UIDs (= Unique Identifier) in the L1Base are L1HS elements (76.03%), whereas the remaining 35 UIDs belong to the evolutionary older L1PA2 family (**Suppl Fig. 6B**). The L1Base annotation is based on the human reference genome (GRCh38) and annotates 146 human full-length (>6kB), intact LINE-1 elements (ORF1 and ORF2 intact) with a unique identifier from 1 to 146 [4]. Attribution of sequencing reads to a specific, individual TE copy is problematic [53] and several approaches have been proposed to circumvent this problem including the mapping of unique reads [48].While several tools using expectation maximization algorithms in assigning multi-mapping reads have been developed and successfully tested in simulations [48,54], we used a different approach in mapping unique reads to the L1Base annotation of full-length LINE-1. Specific "hot" LINE-1 loci in a given cellular context have been identified [3], but usage of the L1Base annotation enabled an unbiased approach albeit ignoring polymorphic LINE-1 sequences. Unique read mapping strategies for repeat elements, especially young LINE-1 elements, will unavoidably underestimate LINE-1 locus-specific expression levels [48], but will be most accurate in assigning reads while allowing the comparison of two different conditions analyzed in parallel. Assuming that expression of UIDs was correlated with mappability, we plotted a mappability count of each UID (see methods) against its mean normalized read count expression of the six individuals ≤65y. Non-parametric Spearman correlation revealed no correlation between UID mappability and expression (**Suppl Fig. 6A**) indicating no apparent bias between the two parameters. However, individual UID dependency of mappability on expression cannot be excluded, especially for high expressing UIDs like UID-16 for

example (**Suppl Fig. 6A,E**). Expression of LINE-1 at the locus-level has been attributed to artefacts not representing autonomous transcription including differential high intronic read counts [55], pervasive transcription or reads attributable to passive co-transcription with genes when the LINE-1 element is intronic [56]. To evaluate the latter, we determined the number of intronic (46.58%) and intergenic UIDs (78/146; **Suppl Fig. 6C**) and identified the corresponding genes for intronic UIDs (**Suppl Fig. 6D**). Of the 146 UIDs, 140 passed the threshold of >3 reads in at least 6 individuals. Differential expression of UID between "young" and "aged" dopaminergic neurons revealed several significantly deregulated full-length LINE-1 loci (**Fig. 5A**). However, while no single locus stood out, paired analysis of the expression of all UIDs indicated a general increase (**Fig. 5B**, especially of low expressed UIDs. The comparative analysis of the sum expression of UIDs per individual comparing young (≤65y) with elderly human dopaminergic neurons, however, did not reach statistical significance (**Fig. 5C**). Several specific loci were dysregulated in particular, for instance UID-68 (**Fig. 5A**), a L1HS element located on chromosome 7 (chr7: 141920659-141926712) in between two genes, CLEC5A (C-type lectin domain containing 5A) and OR9A4 (olfactory receptor family 9 subfamily A member 4). To rule out any influence of "hosting" gene transcription interference on measurable UID-68 expression differences (**Fig. 5D**, left), we performed Spearman correlation which did not indicate any correlation between CLEC5A or OR9A4 expression with UID-68 (**Fig. 5D**). Further, UID-68 had a high mappability count of 16 (range of all UIDs: 1-30, mean 9.0 (SD ±6.05), **Suppl Fig. 6A**) indicating that UID-68 might be a candidate for an age-dependent gain of activity. L1HS UID-129, located on chromosome 15 (chr15: 54926081-54932099) is intergenic, at ≈ 2Mb distance upstream and ≈ 200Mb downstream from the next genes. However, the mappability count of UID-129 was only 1, indicating possible mapping biases inherent to the unique read-based mapping strategy employed (**Fig. 5E** and **Suppl Fig. 6A**). Another UID increased in individuals >65y, the L1HS UID-37 (chr10:98,782,942-98,788,971, minus strand), located in intron 3 of the HPSE2 gene (mappability count of 13, minus strand), showed no correlation with its "hosting" gene indicating potential autonomous transcription of this LINE-1 element and suggesting its contribution to the increase of full-length LINE-1 transcripts (**Suppl Fig. 7A**). We also inspected UID-127 (chr13:40,356,291-40,362,321, mappability count: 14), a L1PA2 element which slightly decreased in elder individuals (**Fig. 5A**). We found a positive association with LINC00598 (intron 6 of 6; ≤65y: r=0.2, p=0.71; >65y: r=0.37, p=0.03), a non-coding RNA which hosts UID-127 in its 6th intron, indicating transcriptional co-regulation potentially indicative of non-autonomous transcription of UID-127 (**Suppl Fig. 7B**). The decrease in the expression of UID-137 (**Fig. 5A**) was mostly driven by one young individual with high expression (**Suppl Fig. 7C**) and thus not reflecting an overall decrease (**Suppl Fig. 7C**, p=0.23). In conclusion, TE expression analysis of this human dataset covering an age-span of 59 years indicates an increase in the expression of young LINE-1 elements including those which have coding potential in elderly dopaminergic neurons. A slight net sum increase of UID transcripts/cell might be sufficient for the production of "above steady-state" levels of ORF1p and ORF2p. Other TEs with coding potential, namely members of the HERV family, were not increased. Further, correlation analyses suggest that L1HS expression might possibly be controlled by the homeoprotein EN1, a protein specifically expressed in dopaminergic neurons in the ventral midbrain [50], the heterochromatin binding protein HP1, two known regulators of LINE-1, and the DNA repair proteins XRCC5/6.

Endogenous ORF1p interactors in the mouse brain

To go further in our understanding of steady-state neuronal ORF1p expression across the mouse brain, we immunoprecipitated ORF1p and performed quantitative label-free LC-MS/MS to identify potential protein partners of ORF1p in the mouse brain. We successfully immunoprecipitated endogenous ORF1p from whole brain lysates (**Fig. 6A**) and identified a total of 424 potential protein interactors associated with ORF1p (Suppl_Table2) in 5 independent experiments (n=5 mice). Using Gene Ontology (GO) analysis, we identified several interacting proteins belonging to GO terms related to known functions of the ORF1p protein in RNA binding, preferentially [57] but not exclusively *in cis* [58], for instance RNA decapping and mRNA catabolic process, or related to

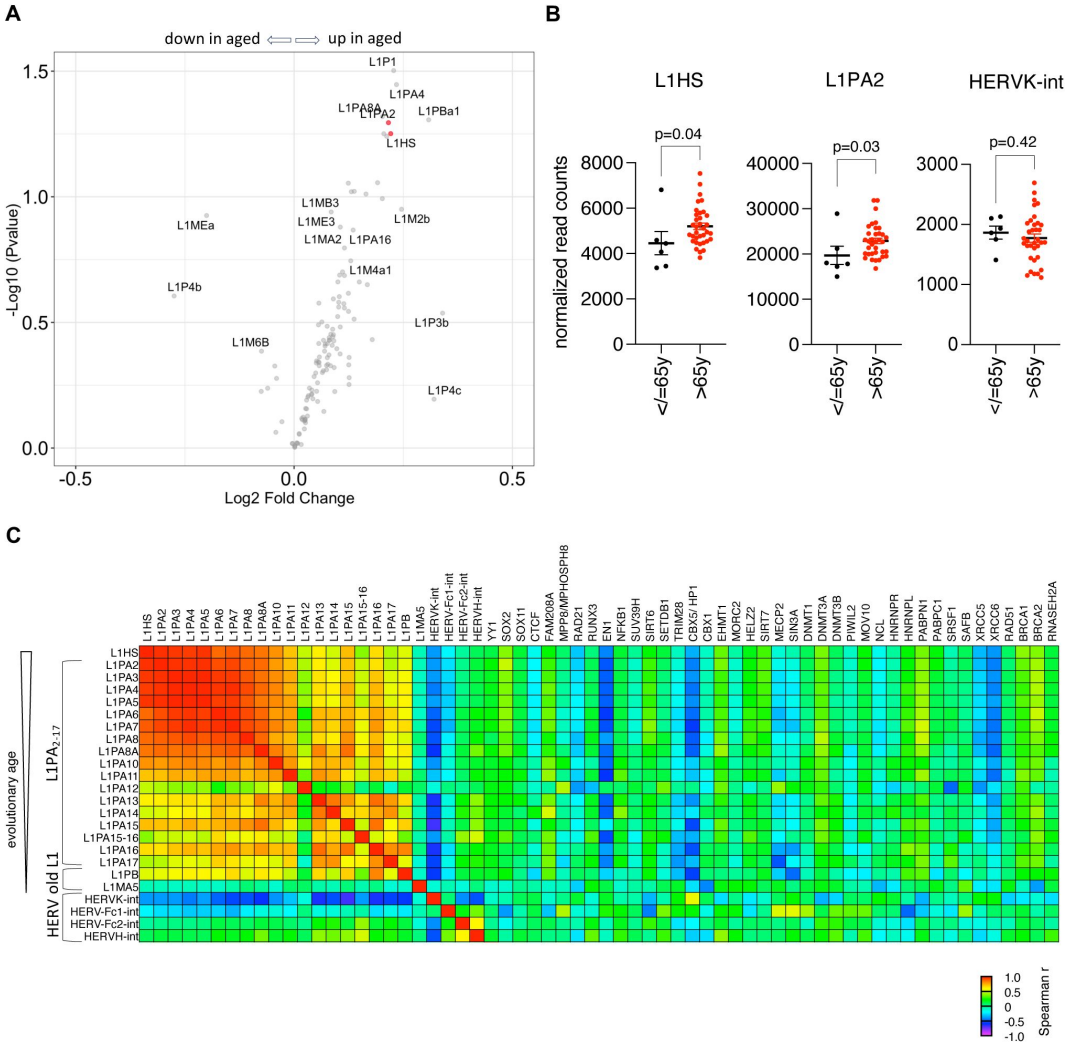


Fig. 4

Young LINE-1 elements are increased in aged human dopaminergic neurons

TE transcript expression in RNA-seq data of laser-captured micro-dissected post-mortem human dopaminergic neurons of brain-healthy individuals was analyzed using RepeatMasker (multimappers) or the L1Base (unique reads).
(A) Volcano plot of differential analysis of LINE-1 expression using DESeq2 comparing young (≤65y, n=6) or aged (>65y, n=35) human dopaminergic neurons at the “name” level of RepeatMasker. Young LINE-1 elements, including the two families L1HS and L1PA2 that have coding copies, are highlighted in red.
(B) Scatter plots of normalized read counts (“name” level) of the young L1HS and L1PA2 families as well as the human endogenous virus family HERVK-int, another TE family with coding potential comparing young (≤65y, n=6) or aged (>65y, n=35) human dopaminergic neurons. Mann-Whitney test, $p<0.05$.
(C) Correlation of the expression of LINE-1 elements with known regulators in human dopaminergic neurons. Spearman correlation of evolutionary close (L1HS, L1PA2-17) and distant LINE-1 (L1PB and L1MA5) as well as HERV elements with coding potential (HERV-Kint, HERV-Fc1, HERV-Fc2 and HERV-H-int) with known regulators of their expression. HERV-W and TREX1 did not pass the normalized read count threshold of >3 reads in >6 individuals.

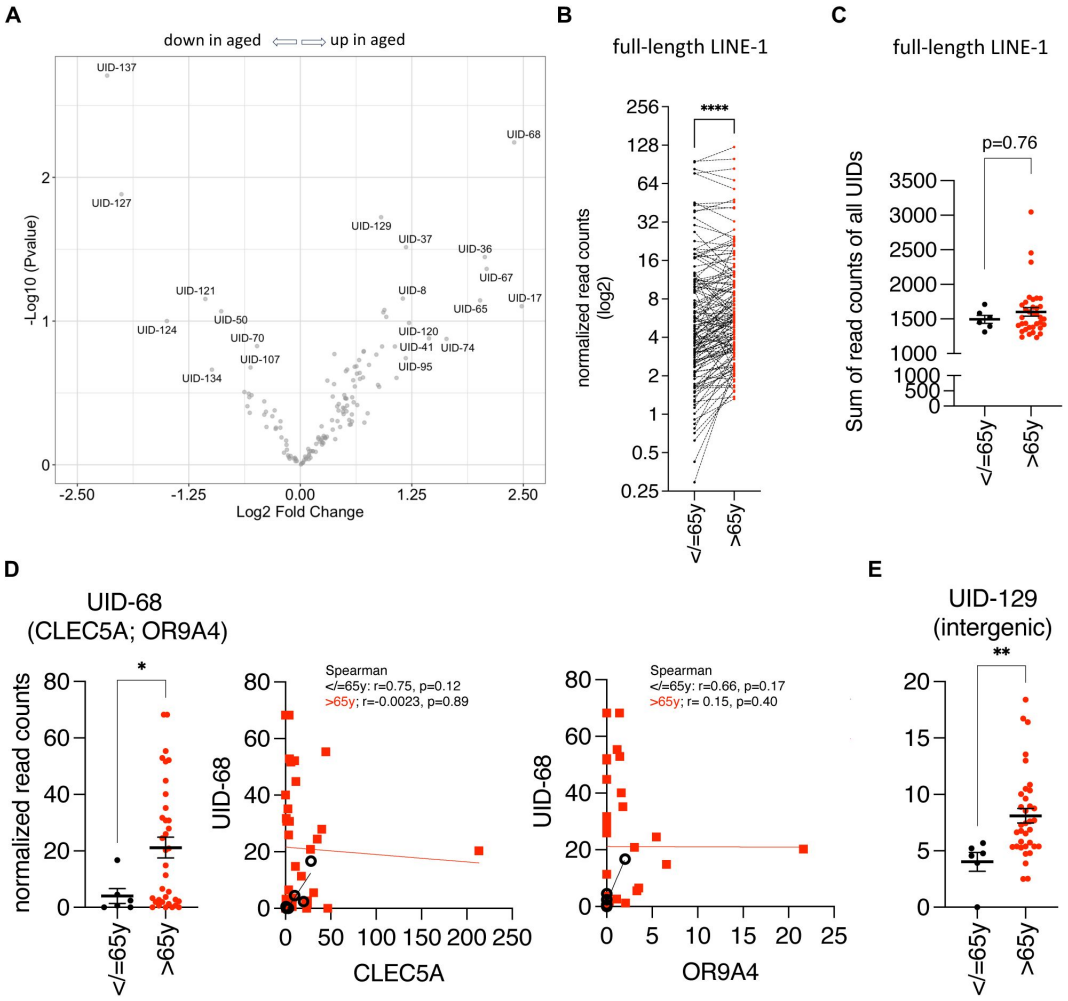


Fig. 5

Dysregulation of locus-specific full-length LINE-1 elements in aged human dopaminergic neurons

(A) Volcano plot of differential expression analysis of TE expression using DEseq2 comparing young (≤65y, n=6) and aged (>65y, n=35) human dopaminergic neurons at the locus-level of specific full-length LINE-1 elements (140 of 146 "UID's" as annotated in L1Base; threshold >3 reads in at least 6 individuals).
(B) Pairwise comparison of the expression of 140 out of 146 full-length LINE-1 elements comparing young (≤65y, n=6) and aged (>65y, n=35) human dopaminergic neurons. Wilcoxon matched signed rank test, p<0.0001left panel).
(C) The sum of read counts of all UIDs per individual were plotted comparing young (≤65y, n=6) and aged (>65y, n=35) human dopaminergic neurons.
(D-E) Dysregulated locus-specific full-length LINE-1 elements (UID-68 and UID-129) are plotted as scatter plots comparing young (≤65y, n=6) and aged (>65y, n=35) human dopaminergic neurons. (D) UID-68 is located adjacent to the genes CLEC5A and OR9A4 (left). Spearman correlation analysis of the expression of UID-68 and CLEC5A (middle) or OR9A4 (right) in young (≤65y, n=6, black dots) or aged (>65y, n=35, red squares) human dopaminergic neurons. (E) UID-129 is intergenic.

the known presence of ORF1p in ribonucleoprotein particles [59,60] (GO: cytoplasmic ribonucleotide granule) or the presence of ORF1p in p-bodies [58] as shown in **Figure 6B** and listed in Suppl_Table3. Other GO terms that emerged, to our knowledge not previously associated with ORF1p, were related to cGMP-mediated signaling (GO: cGMP-mediated signaling and 3'-5'phosphodiesterase activity: i.e. PDE4A, PDE4B, PDE4DIP) and the cytoskeleton (GO: microtubule depolymerization, cytoskeleton organization, microtubule and tubulin binding, cytoskeletal motor activity and protein binding). cGMP signaling is regulated by 3'-5' phosphodiesterases (PDEs) which degrade 3',5'-cyclic guanosine monophosphate (cGMP) and 3',5'-cyclic adenosine monophosphate (cAMP), an activity essential for cell physiology for the integration of extra- and intracellular signals including neuronal excitability, synaptic transmission and neuroplasticity [61,62]. Further, several ORF1p interacting proteins were constituents of the mating-type switching/sucrose nonfermenting complex (SWI/SNF complex), i.e. ARID1A, ARID1B, SMARCA2, SMARCB1, SMARCC2), an ATP-dependent chromatin remodeler complex disrupting nucleosome/DNA contacts to facilitate DNA/chromatin accessibility by shifting, removing or exchanging nucleosomes along DNA [63,64]. Finally, we also observed proteins belonging to the GO term "neuronal cell body", corroborating with the neuron-specific presence of ORF1p in the brain. A comparative analysis with previous mass spectrometry studies [60,65–70] aimed at identifying ORF1p interacting proteins unveiled significantly more common proteins than randomly expected (overrepresentation test; representation factor 2.6, p< 5.4e-08; **Fig. 6C**), including LARP1, STAU2, ATXN2, RALY, TARBP2 or DDX21 (for a full list see Suppl_Table4). The presence of a significant number of overlapping ORF1p interactors in different non-neuronal human cells (HEK[60,65,66], HeLa[67], human breast and ovarian tumors[70] and hESCs[68]) and mouse brain cells (our study), suggest conserved key interactors between both species and between cell types, with a subset of these proteins regulating RNA degradation and translation potentially relevant for the LINE-1 lifecycle itself. ORF1p interactors found in mouse spermatocytes [69] were also present in our analysis including CNOT10, CNOT11, PRKRA and FXR2 among others (Suppl_Table4). To unravel the physical interactions between the identified interactors of endogenous ORF1p within the mouse brain, we used the STRING database (Search Tool for Recurring Instances of Neighboring Genes, *https://string-db.org/*). This analysis generated a network representation, where physical interactions are represented by edges (**Fig. 6D**). In analogy with the GO term analysis, ORF1p displayed interactions with various clusters, including well-known RNA decapping complexes directed against LINE-1 RNA, which also encompassed DCP2 and DCP1A which had not previously been identified as interacting with ORF1p [71]. Furthermore, ORF1p exhibited interactions with the SWI/SNF complex (highlighted in red) as well as subunits of the RNA polymerase II complex suggesting a direct or indirect association with accessible chromatin, a hitherto unknown interaction of ORF1p with chromatin compartments within the nucleus. Notably, a multitude of novel interactors belonged to the "neuronal cell body" and "neuron projections" clusters, proposing potential neuron-specific partners of ORF1p such as Grm2/5, Bai1, Epha4, Kcnn2, Grik2 and Dmd among others. A last cluster, formed by Ncoa5 (Nuclear Receptor Coactivator 5), Nxf1 (Nuclear RNA Export Factor 1), Ranbp2 and Nup133 (both nucleoporins), might imply a role for these interactions in L1-RNA nuclear export and/or a mechanism for the LINE-1 RNP to gain access to the nucleus in post-mitotic neurons. Altogether, the identification of known and novel interactors of ORF1p in the mouse brain suggests roles of ORF1p in the LINE-1 life cycle (RNA binding and metabolism, RNP formation, nuclear access) but also suggests potential novel physiological roles of ORF1p in the brain related to cytoskeleton organization, cGMP signaling, neuron-specific functions (i.e. synaptic signaling, Suppl_Table3) and chromatin organization and/or transcription regulation.

Discussion

While LINE-1 derepression in aging has been extensively explored in peripheral tissues and various pathologies, including cancer, our understanding of LINE-1, particularly ORF1p expression, in the central nervous system remains limited [20,72,73]. A recent search of ORF1p

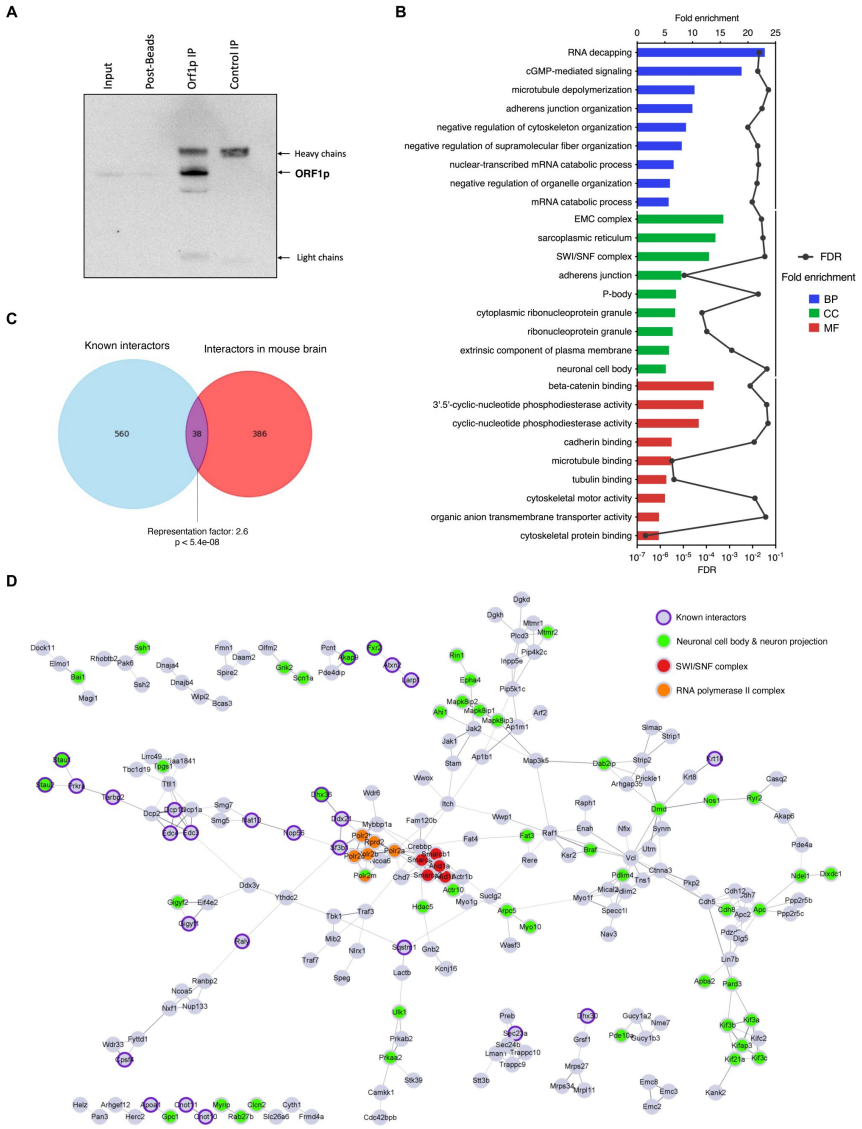


Fig. 6

Endogenous ORF1p interactors in the mouse brain

Immunoprecipitation (IP) of endogenous ORF1p from the mouse brain. WB against ORF1p showing ORF1p enrichment after IP but no signal in the IgG control. Five independent samples were then prepared for proteomic analysis by mass spectrometry (LC-MS/MS).

(B) GO slim enrichment analysis of proteins selected as endogenous ORF1p protein partners in the mouse brain after quantitative LC-MS/MS. ORF1p-immunoprecipitated proteins were categorized into GO slim terms. The nine GO slim term with the highest fold-change are plotted. Fold enrichment is depicted on the upper axis and displayed as bars, the FDR value appears on the lower axis and is represented by the black points. BP: Biological Process, CC: Cellular Component, MF: Molecular Function.

(C) Venn diagram showing common interactors (purple) between interactors of endogenous ORF1p in the mouse brain identified in this study (red) and known (published) interactors of ORF1p (blue). Statistical significance of the overlap between the two groups of proteins was tested by an overrepresentation test (*http://nemates.org/MA/progs/overlap_stats.html*).

(D) ORF1p associates with the SWI/SNF complex (red), RNA pol II complex (orange) and interactors belonging to GO terms related to neuronal cell body & neuron projection (green). Known interactors previously published [60,65–70] are indicated with a purple ring. STRING network of physical interactions where nodes represent proteins partners identified in (A) and edges thickness represents the strength of shared physical complexes. Only proteins sharing physical interactions were represented.

peptides in mass spectrometry data spanning 29 different healthy tissues did not reveal the presence of ORF1p in the brain, suggesting that its presence might lie below detection limits [20]. Only a few studies explored ORF1p encoded LINE-1 protein expression in the brain, most focusing on a specific region (in mice [17,33] in rats [74] and in human post-mortem brain [19]), but it remained unclear if ORF1p is expressed throughout the entire brain, exhibits cell-type specificity, and most intriguingly, if its expression is influenced by the aging process. Here, we demonstrate that ORF1p is expressed throughout the entire mouse brain and in at least three regions of the human post-mortem brain at steady-state. Leveraging a comprehensive workflow that incorporates brain atlas registration and machine learning algorithms (as described in the methods section), we quantified tens of thousands of brain cells, enabling a profound analysis of cell proportions, cell identities, densities and ORF1p expression levels across the entire brain. Surprisingly, more than one-fifth of detected cells expressed ORF1p. Strikingly, regional variations in ORF1p expression levels were observed, with each region exhibiting distinct proportions, cell density, and signal intensity of ORF1p+ cells. In a non-neurologically diseased human brain, ORF1p is expressed in all three regions examined, that is the cingulate gyrus, the frontal cortex and the cerebellum. This is in accordance to an earlier study using histological staining, which found ORF1p expression in the human frontal cortex, the hippocampus, in basal ganglia, thalamus, midbrain and the spinal cord [19]. This suggests, similarly to the mouse brain, a generalized expression across the human brain. On the transcriptomic level using long-read sequencing of GTEx tissues, brain and liver were highlighted as the organs displaying the highest expression of putatively active, full-length LINE-1 elements [75]. However, when the authors looked at sub-regions, they found transcript expression in cerebellar hemispheres and the putamen, but not in the caudate and the anterior cingulate gyrus and frontal cortex [75]. This is in contrast to our data and the data from Sur et al, where ORF1p was found to be expressed in the latter two regions using two different antibodies. We used the anti-human LINE-1 ORF1p antibody clone 4H1, a well characterized antibody [73,76]. While the sample size for the staining of human post-mortem tissues certainly needs to be increased in order to draw quantitative conclusions, the presence of the protein in two independent studies does point to a steady-state expression of ORF1p in the human brain.

In the mouse brain, we find ORF1p to be expressed predominantly if not exclusively in neurons using immunofluorescence and fluorescence-activated cell sorting (FACS). This result is consistent with previous studies, such as the identification of ORF1p in excitatory neurons within the mouse frontal cortex [77], in parvalbumin neurons in the hippocampus [18], its presence in neurons in the ventral midbrain [17] and the recognition of morphological similarities between stained neurons and ORF1p+ cells in a post-mortem hippocampus sample of a healthy individual [19]. We also detected ORF1p in Purkinje cells in the mouse and in bulk human cerebellum. Neuronal specificity or preference of LINE-1 expression was also shown on the transcriptomic level in recent studies investigating LINE-1 expression in the mouse hippocampus, where neuronal LINE-1 expression exceeded that of astrocytes and microglia by approximately twofold [20], is abundant in parvalbumin interneurons [18] and single-nuclei RNA-seq data from the mouse hippocampus and frontal cortex which confirmed globally that repetitive elements including LINE-1 are more active in neurons than in glial cells [77]. In the human brain, LINE-1 transcripts are found in greater quantities in neurons compared to non-neuronal cells by single-nucleus sequencing [78]. Furthermore, retrotransposition-competent LINE-1 elements (similar to UID) are found expressed exclusively in neurons [79]. While ORF1p expression is suggested to be expressed in microglia under experimental autoimmune encephalomyelitis conditions in the spinal cord [80], no evidence of such expression was observed in non-neuronal cells under non-pathological condition.

On average, throughout the mouse brain, the majority of neurons was positive for ORF1p and in some regions (i.e. the thalamus) around 80% of neurons expressed ORF1p. Comparing the results of both imaging approaches, the percentages of neurons expressing ORF1p in the ventral midbrain and frontal cortex were roughly similar (around 70% of neurons expressed ORF1p as quantified by confocal imaging and about 60% of neurons were identified as ORF1p+ using the slide scanner

approach). In the human cingulate gyrus, we found that 37.2% of neurons express ORF1p and that 80% of cells expressing ORF1p were neurons, which are proportions similar to some regions of the mouse brain. It is however possible that these percentages are underestimated due to technical issues inherent to the machine-learning based algorithm for cell detection as our observations often indicated a positive signal in neurons which were classified as negative due to a particular shape or our stringent intensity threshold. A question which arises based on these findings is whether specific features distinguish ORF1p+ and ORF1p-neurons. One hint comes from a recent study suggesting that in the mouse hippocampus, it is the parvalbumin positive neurons that predominantly express ORF1p [18]. We have made a similar observation in the mouse ventral midbrain, where TH-positive dopaminergic neurons express higher levels of ORF1p compared to surrounding, non-dopaminergic neurons [17] (**Fig. 1B**, panel 8). In the cerebellum, we observed ORF1p staining in Purkinje cells but not in the surrounding granular and molecular layer neurons (**Fig. 1B**, panel 10). Parvalbumin positive neurons are inhibitory neurons, so are Purkinje cells. However, dopaminergic neurons are modulatory neurons exerting excitatory and inhibitory effects depending on the brain region they act on. Specific neurons in the granular layer (i.e. Golgi and unipolar brush cells) of the cerebellum are inhibitory, but ORF1p negative, indicating that the decisive feature might not be the excitatory or inhibitory nature of a neuron. Another possibility is a cell-type specific chromatin organization permissive for the expression of LINE-1 and future single-cell studies in the mouse and human brain might reveal those differences.

Because transposable elements are known to become active in somatic tissues during aging [15,16,28,81,82], we aimed to investigate whether there was a corresponding increase at the protein level. In aged mice, ORF1p expression significantly increased throughout the mouse brain consistent with a previously documented increase in ORF1p outside the central nervous system in aged rats [74,83] and aged mice [81]. By quantifying the mean intensity of ORF1p in over 70 000 cells identified as ORF1p+, we were able to characterize the extent of this increase in each anatomical sub-region. Remarkably, apart from the isocortex which did not show any change, ORF1p expression increased in all other brain regions by 7% to 27%, indicating a generalized increase of ORF1p expression in neurons throughout the brain (13%). We did not detect any change in cell identity of ORF1p expressing cells, that is, ORF1p expression remained predominantly if not exclusively neuronal. Intriguingly, we observed a loss of NeuN+ cells, particularly those expressing ORF1p, throughout the brain which was more pronounced in specific regions (ventral midbrain) than in others (isocortex). The loss of NeuN+ cells could either be due to a loss of neuronal identity, as described recently in the context of neuronal aging [84] and in the context of Alzheimer disease and related tauopathies [85], to a neurodegenerative process *per se* or to gliosis related to the aging process [86] (as we observe a slight increase of ORF1p-/NeuN-cells per mm^2).

Interestingly, a region with a strong increase in ORF1p expression with aging (ventral midbrain) also had a significant loss of NeuN+ cells while a region with no change in ORF1p expression with aging (isocortex, frontal cortex) did not lose NeuN+ cells. However, further investigations are necessary to validate a correlation and to investigate an underlying mechanism. An increase of ORF1p might have several direct or indirect consequences on a cell or here, on a neuron. As ORF1p is translated from a polycistronic LINE-1 RNA together with ORF2p, albeit in much higher amounts (the estimated ratio ORF1p to ORF2p is 240:1) [87], it can be expected that a LINE-1 ribonucleotide particles are formed and ORF2-dependent cell toxicity in form of genomic instability [17,21] and single-stranded cytoplasmic DNA triggered inflammation [27,28,81] might result. This has been shown in mouse dopaminergic neurons where oxidative stress induced LINE-1 causally contributed to neurodegeneration [17].

Neurodegeneration was partially prevented by anti-LINE-1 strategies among which NRTIs [17] and similar LINE-1 protein-dependent neuronal toxicity has been shown in drosophila [31,32] and the mouse cerebellum [33].

In order to test whether an increase in LINE-1 is a feature of human brain aging, we turned to a unique RNA-seq dataset of human laser-captured dopaminergic neurons of 41 individuals ranging from 38 to 99 years [47]. In accordance with our focus on LINE-1 sequences which are full-length and coding, we developed a rationale to interrogate LINE-1 families with representatives that are coding (L1HS, L1PA2, multimappers; RepeatMasker) and to specifically investigate full-length LINE-1 elements that have intact open reading frames for ORF1p and ORF2p (unique reads; L1Basev2 [4]). Indeed, we find an increase in L1HS and L1PA2 elements in individuals ≥65y as well as an increase in specific full-length LINE-1 elements but only a trend for increase of all full-length LINE-1 in sum in the elderly. This analysis has technical limitations inherent to transcriptomic analysis of repeat elements especially as it is based on short-read sequences. Nevertheless, we tried to rule out several biases by demonstrating that mappability did not correlate with expression and that the expression of intronic full-length LINE-1 elements is not correlated with the expression of their "hosting" gene. Interestingly, dysregulated full-length LINE-1 elements in aged dopaminergic neurons did not correspond to those identified in bladder cancer [88] indicating the intricate nature of this expression across tissues and pathological conditions. Overall, a slight net sum increase of UID transcripts/cell might be sufficient for the production of "above steady-state" levels ORF1p and ORF2p. Further, a dissociation of LINE-1 transcript and protein levels in aging has been observed recently in excitatory neurons of the mouse cortex. In the absence of transcriptional changes of LINE-1, protein levels of ORF1p were increased [77].

We can only speculate about the reason for an increase in ORF1p in the aged brain. A recent single-cell epigenome analysis of the mouse brain suggested a specific decay of heterochromatin in excitatory neurons of the mouse brain with age which was paralleled by an increase in ORF1p, albeit equally in excitatory and inhibitory neurons, again not indicating any dependency of ORF1p regulation on the excitatory or inhibitory nature of neurons [77]. Chromatin and particularly heterochromatin disorganization are a primary hallmark of aging [82] but other repressive cellular pathways which control the LINE-1 life cycle might also fail with aging [13]. Another possibility is a loss of accessibility of repressive factors to the LINE-1 promoter or an age-dependent decrease in their expression. Matrix correlation analysis of several known LINE-1 regulators, both positive and negative, revealed possible regulators of young LINE-1 sequences in human dopaminergic neurons. Despite known and most probable cell-type unspecific repressive factors like the heterochromatin binding protein CBX5/HP1 [51] or the DNA repair proteins XRCC5 and XRCC6 [49], we identified the homeoprotein EN1 as negatively correlated with young LINE-1 elements including L1HS and L1PA2. EN1 is an essential protein for mouse dopaminergic neuronal survival [50] and binds, in its properties as a transcription factor, to the promoter of LINE-1 [17]. As EN1 is specifically expressed in dopaminergic neurons in the ventral midbrain, our findings suggests that EN1 controls LINE-1 expression in human dopaminergic neurons as well and serves as an example for a neuronal sub-type specific regulation of LINE-1.

The heterogenous, brain-wide presence of ORF1p expression at steady-state is intriguing. In cancer cell lines or mouse spermatocytes, ORF1p interacts with several "host" proteins, some if not most of which are related to the LINE-1 life cycle. However, a profile of endogenous ORF1p interactors in the mouse brain might inform on possible other and organ-specific functions besides its binding to the LINE-1 RNA in "cis" [69]. Among the total 424 potential interactors of endogenous ORF1p in the mouse brain, 38 partners had been previously identified by mass spectrometry in human cancers, cancerous cell lines and mouse spermatocytes [60,65–70] (Suppl_Table4). Further, GO term analysis contained expected categories like "P-body", mRNA metabolism related categories and "ribonucleoprotein granule". We also identified NXF1 as a protein partner of ORF1p, a protein found to interact with LINE-1 RNA related to its nuclear export [89]. This suggests the conservation of key interactors probably essential for completing or repressing the LINE-1 life cycle in both species, despite the divergence of mouse and human ORF1p protein sequences [90]. Along these lines, several ORF1p protein partners we identified might complete the list of post-transcriptional regulators implicated in LINE-1 silencing. Recent work conducted on human cancerous cell lines has demonstrated that MOV10 orchestrates the recruitment of DCP2 for LINE-1

RNA decapping [71]. In our analysis, we identified DCP2 along with DCP1A, known to enhance the decapping activity of DCP2 [91], and DCP1b, a pivotal component of the mRNA decapping complex [92]. Intriguingly, MOV10 was not detected in our mass spectrometry analysis, despite its established role in recruiting DCP2 and forming a complex with L1-RNP to mediate LINE-1 RNA decapping, as reported by Liu et al [71]. However, we found two enhancers of mRNA decapping, EDC3 and EDC4, both core components of P-bodies, a membrane-less organelle known to contain L1-RNP [58]. Multiple ubiquitin-ligase proteins were found although not appearing as a significantly enriched GO term. These results complete the picture of the post-transcriptional and translational control of ORF1p and suggest that these mechanisms, despite a steady-state expression, are operational in neurons. Further, several neuron-specific interactors were identified belonging to GO term categories “neuron projection” (75 proteins) and “neuronal cell body” (5 proteins), again pointing to the neuron-predominant expression of ORF1p in the mouse brain. Other interesting aspects were raised from this analysis. Among significantly enriched GO terms, several were related to the cytoskeleton, the functional consequences of which need to be determined in future studies. Our screen also identified PDE10A as an interactor of ORF1p in the mouse brain, a PDE almost exclusively expressed in medium spiny neurons of the striatum and a target for treatment of neurological diseases related to basal ganglia function like Huntington’s disease, schizophrenia and Tourette syndrome [93]. Interestingly, PDE10A inhibition is related to beta-catenin signaling, another GO term which emerged from our screen [94]. Finally, we found components of RNA polymerase II and the SWI/SNF complex as partners of ORF1p. This further indicates that ORF1p has access to the nucleus in mouse brain neurons as described for other cells [95,96], implying that ORF1p potentially has access to chromatin. These findings give rise to intriguing questions regarding the potential function of ORF1p in neuron in health and disease as (i) ORF1p is widely distributed throughout the brain under normal physiological conditions, (ii) ORF1p shows a wide range expression levels within and inbetween regions, (iii) ORF1p is expressed predominantly if not exclusively in neurons, (iv) but not in all neurons and (v) interacts with proteins that might not directly relate to the LINE-1 life cycle, some of which are neuron-specific. In addition, physicochemical properties of ORF1p to form compacted nucleic acid - bound complexes with sequestration potential were shown [90,97]. Future loss-of-function studies should help to shed light on the necessity of ORF1p for neuronal functions if they exist. Our data spurs the idea of a possible “physiological” function of ORF1p as an integrative protein with exapted function in neuronal homeostasis and a loss of restriction in the aged brain limiting LINE-1 expression to steady-state levels.

Materials & Methods

Animals

Swiss OF1 wild-type mice (Janvier) were housed on a 12h light/dark cycle with free access to water and food. Mice were sacrificed at 3-month or 16-month. Animal experiments were performed according to the EU directive 2010/63/EU.

Mouse tissue dissection and protein extraction

Tissues were extracted from 3-month-old and 16-month-old Swiss/OF1 mice. Briefly, the two hemispheres were separated in ice cold PBS -/-. For each mouse, one hemisphere was rinsed and fixed in 4% PFA for 1h followed by 24h of incubation in 30% sucrose. Hemispheres were kept at −20°C until being sliced on a freezing microtome (Epredia, HM 450) with a 20µm thickness. The other hemisphere was dissected in ice cold PBS -/- 1X and 6 brain regions were rinsed, cut in small pieces and dissociated separately using a large (21G) to small gauge (27G) needle in RIPA lysis buffer for 5 min. Lysates were kept on ice for 25 min, were sonicated for 15 min and supernatants were collected after a 30 min centrifugation at 4°C at 14 000 rpm.

Proteins were quantified and Laemmli buffer was added before boiling for 10min at 95°C to be used for Western Blot.

Human Samples

Cerebellum, frontal cortex and cingulate gyrus human samples were provided by Neuro-CEB biobank from a 78-year-old healthy male individual and conserved at −80°C.

Human Samples pulverization and protein extraction

We used the dry pulverizer Cryoprep (Covaris) for pulverization of tissue blocs. Each sample was disposed in a liquid-nitrogen precooled Tissue-tube bag and dry cryo-pulverized with one impact at the maximum level. The pulverized brain sample was then weighed and resuspended in lysis buffer (mg/v) (0.32M sucrose, 5mM $CaCl_2$, 3mM $Mg(CH3COOH)_2$, 0.1mM EDTA, 10mM Tris-HCL pH8, 1mM DTT, 0.1% TritonX-100 and Protease Inhibitors), kept on ice for 30 min with gentle up-and-down pipetting until homogenization. We added 2X RIPA buffer (v/v) to totals fractions for 30 min on ice. We then sonicated samples 2 times for 15 min. AtlasSupernatants were collected after a 30 min centrifugation at 14 000 rpm at 4°C, proteins were quantified and Laemmli buffer was added to be used for Western Blot. All samples were boiled 10 min at 95°C to be used for Western Blot.

Western-Blot

We used 1.5mm NuPAGE 4-12% Bis-Tris-Gel (Invitrogen™). Proteins samples were loaded and gel migration was performed with NuPAGE™ MES SDS Running Buffer (Invitrogen™) for 45 minutes at 200mV. Gels were transferred onto a methanol activated PVDF membrane (Immobilon) in a buffer containing: Tris 25 mM, pH=8,3 and Glycine 192 mM, during 1h30 at 400 mA. Membranes were blocked 30 min with 5% milk in TBST (0,2% Tween 20, 150 mM NaCl, 10 mM Tris pH:8). Primary antibodies were diluted in 5% milk in TBST, and membranes were incubated o/n at 4 C°. After 3 x 10 min washing in TBST, membranes were incubated for 1h30 with the respective secondary antibodies diluted at a concentration of 1/2000 in 5% milk TBST. Membranes were washed 3 x 10 min in TBST and were revealed by the LAS-4000 Fujifilm system using Clarity Western ECL Substrate (Bio Rad) or Maxi Clarity Western ECL Substrate (Bio Rad).

Immunostaining

Sagittal mouse brains slices were fixed for 10min in PFA 4% and rinsed 3 times for 10min in PBS -/-. Slices were then incubated 20 min in glycine 100mM, washed 3 times for 5min in PBS and immersed in 10mM citrate pH 6 at 62°C during 45min for antigen retrieval. Slices were then immersed 3 times in PBS with Triton X-100 0.2% and incubated in blocking buffer for 1,5 h (PBS with Triton X-100 0.2% and FBS (10%) previously inactivated 20min at 56°C (Gibeco, 16141061). Primary antibodies (ORF1p antibody: abcam ab216324; NeuN antibody: GeneTex GTX00837) were diluted (1/200 and 1/500 respectively) in blocking buffer and incubated with slices overnight at 4°C and then washed 3 times for 10 min with PBS. For validation, an in-house ORF1p antibody was used (09) (guinea pig, 1/200). For non-neuronal marker (GFAP antibody: Millipore AB5541; Iba1 antibody: GeneTex GTX101495; Sox9 antibody: RnDsystems AF3075; S100β antibody: Sigma S2532), antibodies were diluted at 1/500.

Suitable secondary antibodies (Invitrogen) and Hoechst (Invitrogen, 15586276) were incubated for 1,5h at 1/2000 in PBS with inactivated FBS (10%) and washed 3 times 10 min in PBS. To quench tissue autofluorescence, especially lipofuscin, TrueBlack Plus (Biotium) in PBS was used during 10min. Slices were rinsed 3 times in PBS and mounted with Fluoromount (Invitrogen™). For human cingulate gyrus stainings, the same protocol was performed, with the difference that a human ORF1p antibody (Abcam 245249) was used. Mouse and human brain slices were imaged by the Axioscan 7 Digital Slide Scanner (Zeiss) or a Spinning Disk W1 confocal microscope (Yogogawa).

Blocked peptide

The ORF1p antibody (abcam ab216324) was incubated 2h on a turning wheel with excess (4:1) of mouse ORF1p recombinant protein as in [17] before the blocked antibody was used in the above-described immunofluorescence protocol.

Quantification of confocal acquisitions

Analysis was performed with a custom-written plugin developed for the Fiji software, using Bio-Formats [98] and 3D Image Suite [99] libraries. Code is freely available online at *https://github.com/orion-cirb/DAPI_NEUN_ORF1P*. Nuclei were detected using Hoechst channel downscaled by a factor of 2 with the 2D-stitched version of Cellpose [100] (percentile normalization = [1-99], model = 'cyto', diameter = 30, flow threshold = 0.4, cell probability threshold = 0.0, stitching threshold = 0.75). The segmented image is rescaled to its original size, and the obtained 3D nuclei are filtered by volume to avoid false positive detections. NeuN and ORF1p+ cells were detected in their respective channel in the same manner as nuclei, but with different Cellpose settings (model = 'cyto2', diameter = 40, flow threshold = 0.4, cell probability threshold = 0.0, stitching threshold = 0.75). Then, each cell was associated with a nucleus having at least half of its volume in contact with it. Cells without any associated nucleus were filtered out. Each nucleus was thus labeled according to NeuN and/or ORF1p positivity.

ABBA Registration and Qupath analysis

We registered each sagittal brain section with the Allen Mouse Brain Atlas (CCFv3 [101]), using the Aligning Big Brains & Atlases plugin [102] in Fiji. Registration results were imported into QuPath software [103] for downstream processing.

In each brain subregion, cell analysis was done with custom Groovy scripts developed for QuPath. Code is freely available online at *https://github.com/orion-cirb/QuPath_ORF1P*. In brief, Hoechst nuclei were detected with StarDist 2D [104], applying the DSB 2018 pretrained model with the following parameters: percentile normalization = [1-99], probability threshold = 0.75, overlap threshold = 0.25. Cells in NeuN and ORF1p channels were detected with Cellpose 2D (percentile normalization = [1-99], model = 'cyto2', diameter = 30, flow threshold = 0.4, cell probability threshold = 0.0). Nuclei and cells were then filtered by area and intensity in their respective channel, once again to avoid false positive detections. Minimum intensity threshold was based on the channel background noise. This noise was estimated for each subregion as the mean intensity of pixels not belonging to any nucleus or cell detected in the corresponding channel. Finally, each cell was associated with a nucleus having its centroid located inside the cell mask. Cells without any associated nucleus were filtered out. Each cell containing a nucleus was thus identified as NeuN+ or NeuN- and ORF1p+ or ORF1p-. Intensity values were normalized by subtracting the background noise computed in the corresponding channel and subregion. As a last step, subregional results were merged into regional ones and data were analyzed using the Pandas Python library [105].

FACS

Mouse brains were dissociated with Adult Brain Dissociation kit (Miltenyi Biotec, 130-107-677) and incubated with the coupled antibody NeuN Alexa 647 (Abcam, ab190565) or the control isotype IgG Alexa 647 (Abcam, ab199093). Stained cells were filtered a last time with a 40µm filter before FACS sorting (FACS ARIA II). Neuronal and non-neuronal cells were separately collected in PBS -/- 2m EDTA and then centrifugate (5min at 700rpm). Pellets were resuspended in RIPA for protein extraction in an appropriate volume in order to achieve equal cell concentrations (10 000 cells/µl).

RNA-seq analysis

The RNA-seq dataset from Dong et al [47] was downloaded from dbGAP (phs001556.v1.p1) and contains unstranded paired-end 50bp and 75bp reads from pooled laser-capture micro-dissected dopaminergic neurons from human post-mortem brain (107 samples) from 93 individuals w/o brain disease. RNA-seq had been done on total and linearly amplified RNA. We focused our analysis on data obtained with 50bp reads, in order to avoid mappability bias, while still regrouping all age categories (n=41; with ages ranging from 38 to 97 (mean age: 79.88 (SD ±12.07); n=6 ≤65y; n=35 >65y; mean PMI: 7.07 (SD ±7.84), mean RIN: 7.09 (±0.94)). Sequencing reads were aligned on the Human reference genome (hg38) using the STAR mapper (v2.7.0a) 3 and two different sets of parameters. Genome-wide individual repeat quantification was performed using uniquely mapped reads and the following STAR parameters: -- outFilterMultimapNmax 1–- alignEndsType EndToEnd–outFilterMismatchNmax 999–- outFilterMismatchNoverLmax 0.06. Repeats class, family and name quantification was performed using a random mapping procedure and the following parameters : -- outFilterMultimapNmax 5000–outSAMmultNmax 1–- alignEndsType EndToEnd–- outFilterMismatchNmax 999–outFilterMismatchNoverLmax 0.06–- outSAMprimaryFlag OneBestScore–-outMultimapperOrder Random. Repeats annotations were downloaded from the UCSC Table Browser (repeatMasker database: *https://genome.ucsc.edu/cgi-bin/hgTables*) and coordinates of LINE-1 full length and coding elements were downloaded from the L1base database 2 (*http://l1base.charite.de/l1base.php*; [4]) selecting LINE-1 full length elements containing two predicted complete open reading frames for ORF1 and ORF2 (UID= Unique IDentifier) from the LINE-1 database (*http://l1base.charite.de/l1base.php*) and corrected genomic intervals with the repeat masker annotation of the corresponding genomic locus. Repeat quantification from the aligned data was done using a gtf file composed of all genes (Gencode v29) and all individual repeat elements. This strategy was used to avoid overestimation of repeat elements due to overlaps with expressed genes. For individual repeat quantification of the full length L1 elements (L1base), we therefore used a gtf of all genes and all L1base entries, and ran the FeatureCounts tool [106] with the following parameters: -g gene_id -s 0 -p. In the context of the family-based analysis, we used a gtf with all genes and all annotated repeats elements and ran FeatureCounts with -g gene_family -s 0 -p -M. Before DeSeq2 analysis, we remove all genes and repeat elements with less than 10 reads in a minimum of n individuals, n being the number of individuals in the condition containing the fewest individuals (“young” condition: n=6, 38-65y, mean 57.5 years). We use the same conditions with genes and UIDs with less than 3 reads in a minimum of n individuals. Finally, we calculated the scaling factors using DeSeq2 on all genes + all repeat elements or all gene + UID according to the quantification method and then applied these scaling factors to the corresponding counts tables.

In order to test for the mappability of each UID (= full-length and coding LINE-1), we extracted the bed track « main on human:umap50 (genome hg38) from the UCSC genome browser (≈ 7Mio regions) directly into Galaxy (*usegalaxy.org*) and joined genomic intervals with a minimum overlap of 45bp of this dataset with a dataset containing the annotation of UIDs extracted from L1Basev2 [4] corrected in length with repeat masker and completed with information on whether the UID is intra- or intergenic and, if intragenic, in which gene (NM_ID, chr, strand, start, end, gene length, number of exons, gene symbol) the fl-LINE-1 is located, which resulted in 1266 regions. We then used the “group on data and group by” function in Galaxy and counted the number of overlapping 50kmers with all 146 UIDs (=mappability score). Correlation analysis (non-parametric Spearman) was then done between the mappability score and the normalized read counts.

Immpunoprecipitation of ORF1p from the mouse brain

For immunoprecipitation, we used ORF1p (abcam, ab245122) and IgG rabbit (abcam, ab172730) antibodies. The antibodies were coupled to magnetic beads using the Dynabeads® Antibody Coupling Kit (Invitogen, 14311D) according to the manufacture”s recommendations. We used 5µg of antibody for 1 mg of beads and used 1.5mg of beads for IP. Individual mouse brain lysates (n=5),

homogenized using dounce and sonicated, were incubated with ORF1p or IgG-control coupled beads and a small fraction was kept as input. Each of these two tubes containing coupled beads and brain lysates were diluted in 5 ml buffer (10 mM Tris HCl, 150 mM NaCl, protease inhibitor). The samples were then incubated overnight on a wheel at 4°C. Samples were then washed 3 times with 1 ml buffer (10 mM Tris HCl pH 8, 200 mM NaCl) using a magnet and then resuspended in the same buffer. The samples were boiled in Laemmli buffer (95°C, 10 min) and 20 µl of each sample were loaded on a 4-12% Nupage gel (Invitrogen, NP0336) to be revealed by WB. For samples used in Mass Spectrometry study, beads were washed with buffer (10 mM Tris HCl pH 8, 200 mM NaCl) using a magnet. After 3 washes with 1ml buffer the beads were washed twice with 100 µL of 25 mM NH4HCO3 (ABC buffer). Finally, beads were resuspended in 100 µl of 25mM ABC buffer and digested by adding 0.20 µg of trypsine/LysC (Promega) for 1 hour at 37 °C. A second round of digestion was applied simultaneously on the beads by adding 100 µL of 25 mM ABC buffer and to the previous digest by adding 0.20 µg of trypsin/LysC for 1 hour at 37 °C. Samples were then loaded into homemade C18 StageTips packed by stacking three AttractSPE® disk (#SPE-Disks-Bio-C18-100.47.20 Affinisep) into a 200 µL micropipette tip for desalting. Peptides were eluted using a ratio of 40:60 CH3CN:H2O + 0.1% formic acid and vacuum concentrated to dryness with a SpeedVac device. Peptides were reconstituted in 10 µL of injection buffer in 0.3% trifluoroacetic acid (TFA) before liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis.

Mass Spectrometry

Online chromatography was performed with an RSLCnano system (Ultimate 3000, Thermo Scientific) coupled to a Q Exactive HF-X with a Nanospay Flex ion source (Thermo Scientific). Peptides were first trapped on a C18 column (75 µm inner diameter × 2 cm; nanoViper Acclaim PepMapTM 100, Thermo Scientific) with buffer A (2/98 MeCN/H2O in 0.1% formic acid) at a flow rate of 2.5 µL/min over 4 min. Separation was then performed on a 50 cm x 75 µm C18 column (nanoViper Acclaim PepMapTM RSLC, 2 µm, 100Å, Thermo Scientific) regulated to a temperature of 50°C with a linear gradient of 2% to 30% buffer B (100% MeCN in 0.1% formic acid) at a flow rate of 300 nL/min over 91 min. MS full scans were performed in the ultrahigh-field Orbitrap mass analyzer in ranges m/z 375–1500 with a resolution of 120 000 at m/z 200. The top 20 intense ions were subjected to Orbitrap for further fragmentation via high energy collision dissociation (HCD) activation and a resolution of 15 000 with the intensity threshold kept at 1.3 × 105. We selected ions with charge state from 2+ to 6+ for screening. Normalized collision energy (NCE) was set at 27 and the dynamic exclusion of 40s. For identification, the data were searched against the Mus Musculus (UP000000589_10090 012019) Uniprot database using Sequest HT through proteome discoverer (version 2.4). Enzyme specificity was set to trypsin and a maximum of two-missed cleavage sites were allowed. Oxidized methionine, Met-loss, Met-loss-Acetyl and N-terminal acetylation were set as variable modifications. Maximum allowed mass deviation was set to 10 ppm for monoisotopic precursor ions and 0.02 Da for MS/MS peaks. The resulting files were further processed using myProMS97 v3.10.0. FDR calculation used Percolator and was set to 1% at the peptide level for the whole study. The label free quantification was performed by peptide Extracted Ion Chromatograms (XICs), reextracted by conditions and computed with MassChroQ version 2.2.21 [107]. For protein quantification, XICs from proteotypic peptides shared between compared conditions (TopN matching) with missed cleavages were used. Median and scale normalization at peptide level was applied on the total signal to correct the XICs for each biological replicate (n=5). To estimate the significance of the change in protein abundance, a linear model (adjusted on peptides and biological replicates) was performed, and p-values were adjusted using the Benjamini–Hochberg FDR procedure. Proteins with at least three peptides, identified in each biological replicates of ORF1p condition, a 10-fold enrichment and an adjusted p-value ≤ 0.05 were considered significantly enriched in sample comparisons. Unique proteins were considered with at least three peptides in all replicates. Protein selected with these criteria were used for Gene Ontology enrichment analysis and string network analysis.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (*http://proteomecentral.proteomexchange.org*) via the PRIDE partner repository [108] with the dataset identifier PXD047160.

GO term and STRING network analysis

Gene Ontology analysis was performed using GO PANTHER (PANTHER: Making genome-scale phylogenetics accessible to allPaul D. Thomas, Dustin Albou and Huaiyu MiProtein Society. 2022;31(1):8-22. doi:10.1002/pro.4218) and String network physical interactions were retrieved using the STRING database v11.5 (*https://string-db.org/*) and then implemented in Cytoscape software [109].

Statistical analyses

In column comparisons, data in each column were tested for normality using two normality and lognormality tests (Shapiro-Wilk test and Kolmogorov-Smirnov test). Data which passed the normality tests were analyzed subsequently by a parametric test, data which did not pass the normality tests were analyzed by a non-parametric statistical test as indicated in the figure legends. The significance threshold was defined as $p<0.05$. Statistical analyses were done with PRISM software (v10).

Acknowledgements

This work was supported by the Fondation de France (00086320, to J.F.), the Fondation du Collège de France (to J.F. and T.B.), the Fondation NRJ/Institut de France (to J.F.) and the National French Agency for Research (ANR-20-CE16-0022 NEURAGE). We gratefully acknowledge the Orion Technological Core (IMACHEM-IBiSA) of CIRB, member of the France-BioImaging research infrastructure, especially Estelle Anceaume and Julien Dumont for assistance with slide scanner and spinning disk acquisition and Magalie Fradet for assistance with FACS analysis. We also thank the Fondation Bettencourt Schueller for their support.

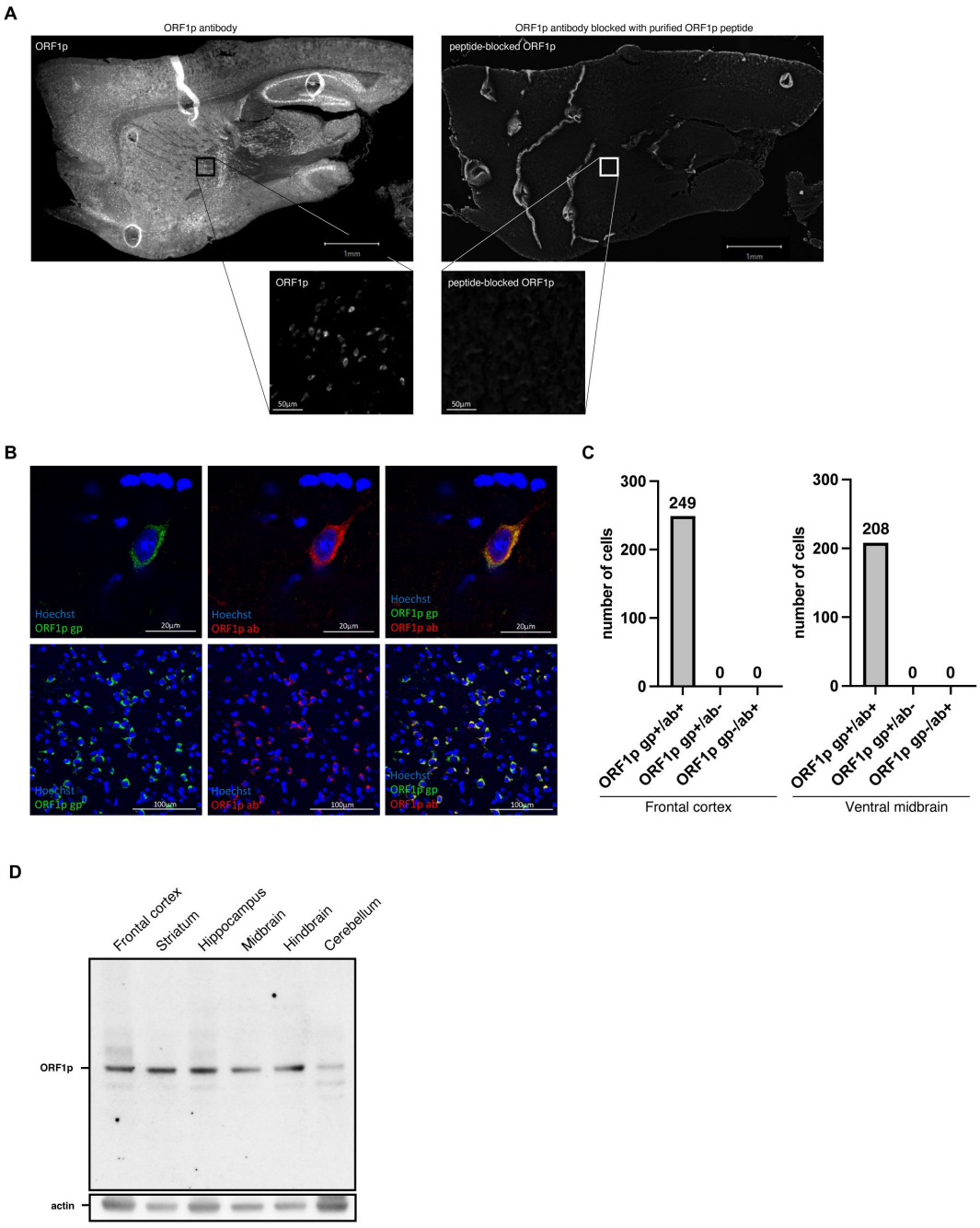
Contributions

T.B. carried out most of the experimental work and analyzed the data. T.B., S.S. and J.F. analyzed the transcriptomic data. O.M.B. and S.S contributed to the experimental work and performed the FACS analysis. T.B. and J.F. wrote the manuscript. T.B., P.M. and H.M. developed the image analysis pipeline. B.L. carried out the MS experimental work and D.L. supervised MS and performed data analysis with T.B.. Read alignment, quality control and mapping were done by N.S.. J.F. and R.L.J. conceived and supervised the project.

Competing interests

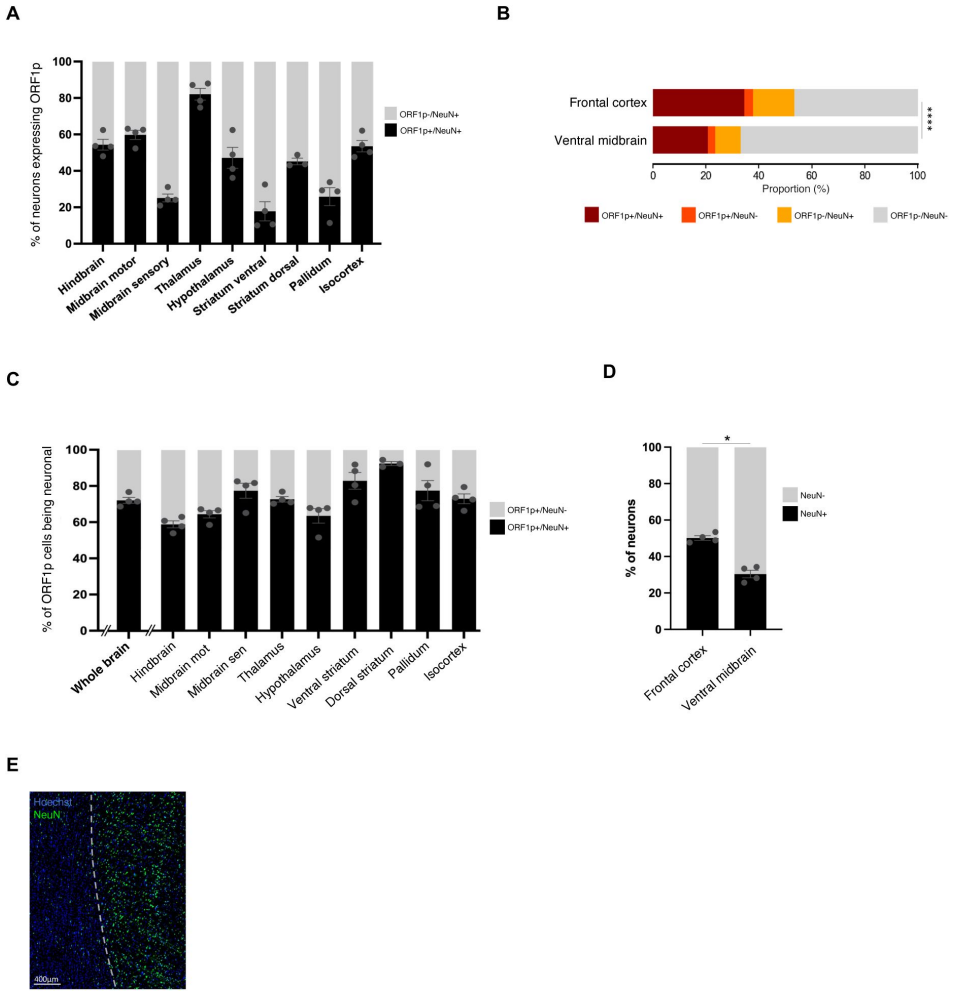
The authors declare no competing interests.

Supplementary Figures



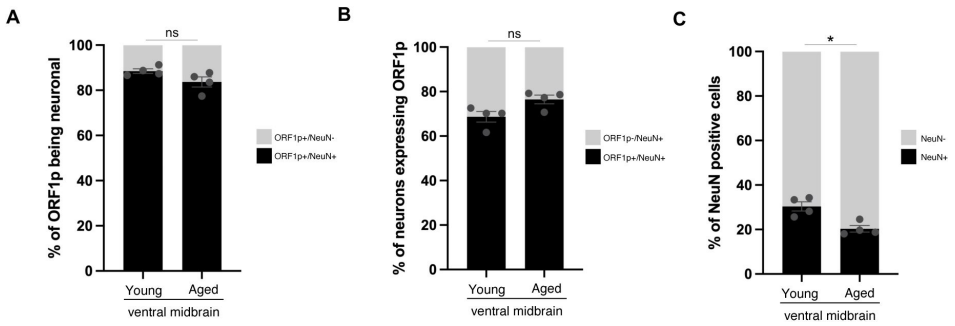
Suppl Fig. 1

(A) Selective recognition of ORF1p antibody. IHC showing ORF1p positives cells in sagittal mouse brain slice (left) and abolition of the signal when blocking the antibody with purified ORF1p (right).
(B) Representative acquisition showing ORF1p obtained with the widely used, commercially available ORF1p ab antibody (abcam ab216324) used in this study (red) and with an in-house ORF1p gp antibody (guinea pig, green) in mouse brain. Scalebar = 20µm (top) and 100µm (bottom).
(C) Quantification of double positives (gp+/ab+) cells using ORF1p ab antibody (abcam ab216324) and in-house ORF1p gp antibody (guinea pig) versus single-positive cells (gp+/ab- and gp-/ab+) in mouse frontal cortex (left) and ventral midbrain (right).
(D) ORF1p is expressed in six different brain regions in the mouse.
Brain regions were micro-dissected from a three-month old mouse brain. Western blot showing ORF1p expression in 6 brain regions. ORF1p (Top), Actin (bottom).



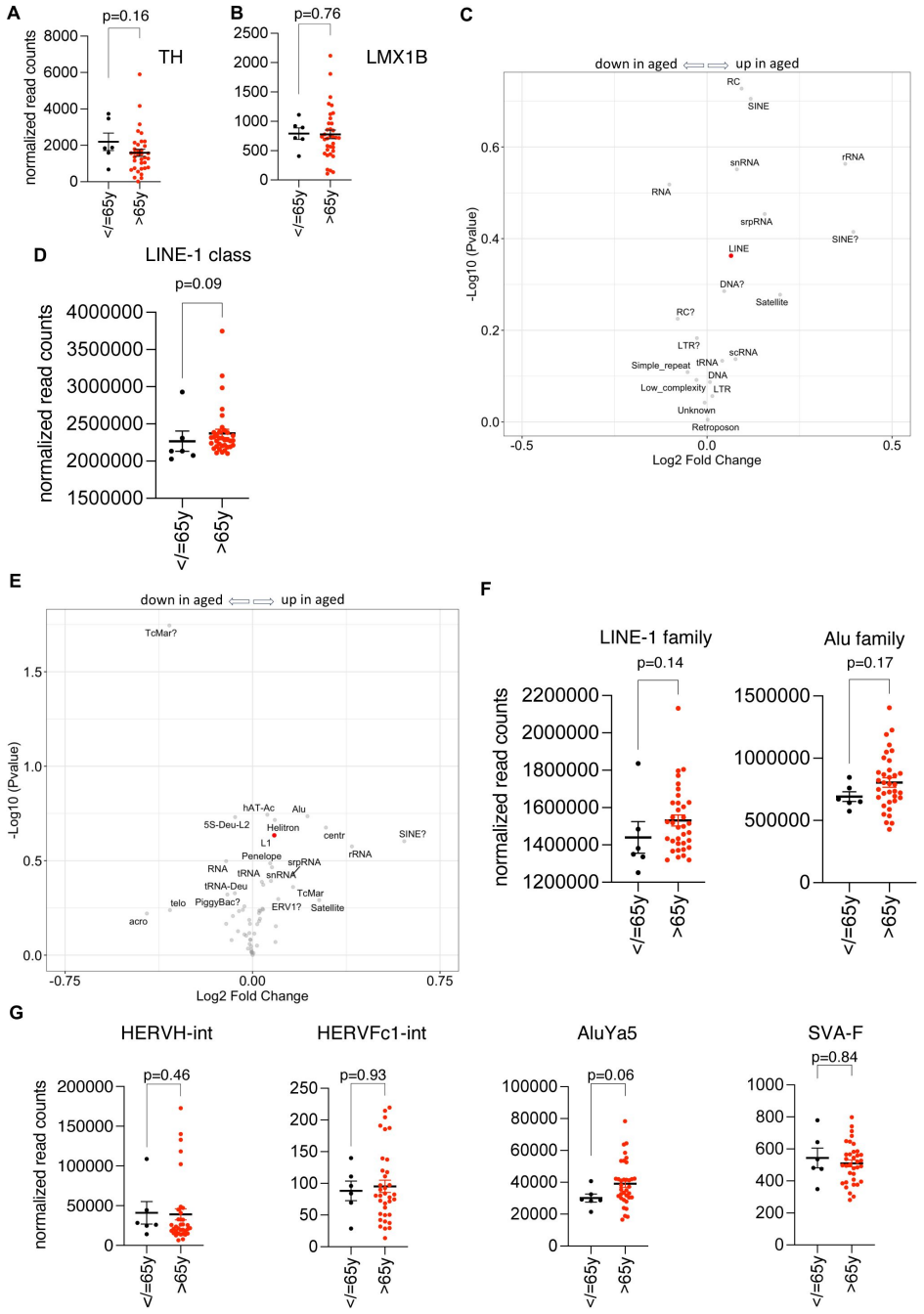
Suppl Fig. 2

(A) Proportion of neurons (NeuN+) expressing ORF1p in different regions of the mouse brain as quantified using the cell detection pipeline on large scale images.
(B) ORF1p is predominantly expressed in neurons. Proportion of ORF1p+/NeuN+, ORF1p+/NeuN+, ORF1p+/NeuN-, ORF1p-/NeuN+ and ORF1p-/NeuN-cells, ****p<0.0001; calculated using chi-square test on the cell number of the four different cell-types analyzed by confocal microscopy on multiple z-stacks.
(C) ORF1p cell identity. Proportion of ORF1p+ cells identified as NeuN+ (black) or NeuN- (grey), in the whole brain (left) and in 9 different regions analyzed (right) using the cell detection pipeline on large scale images presented in **Figure 1A**; data is represented as mean ± SEM, n=4 mice.
(D) Proportion of neurons in the frontal cortex and ventral midbrain quantified using confocal approach. *p<0.05, chi-square test on the cell number of the different cell-types analyzed; n=4 mice, data is represented as mean ± SEM.
(E) Representative slide-scanner acquisition of a human cingulate gyrus section showing NeuN positives cells (green) mostly located in the grey matter (right) compared to the white matter from a brain-healthy individual; scale bar = 400µm. Zoom into the grey matter region showing ORF1p is presented in **Figure 2H**.



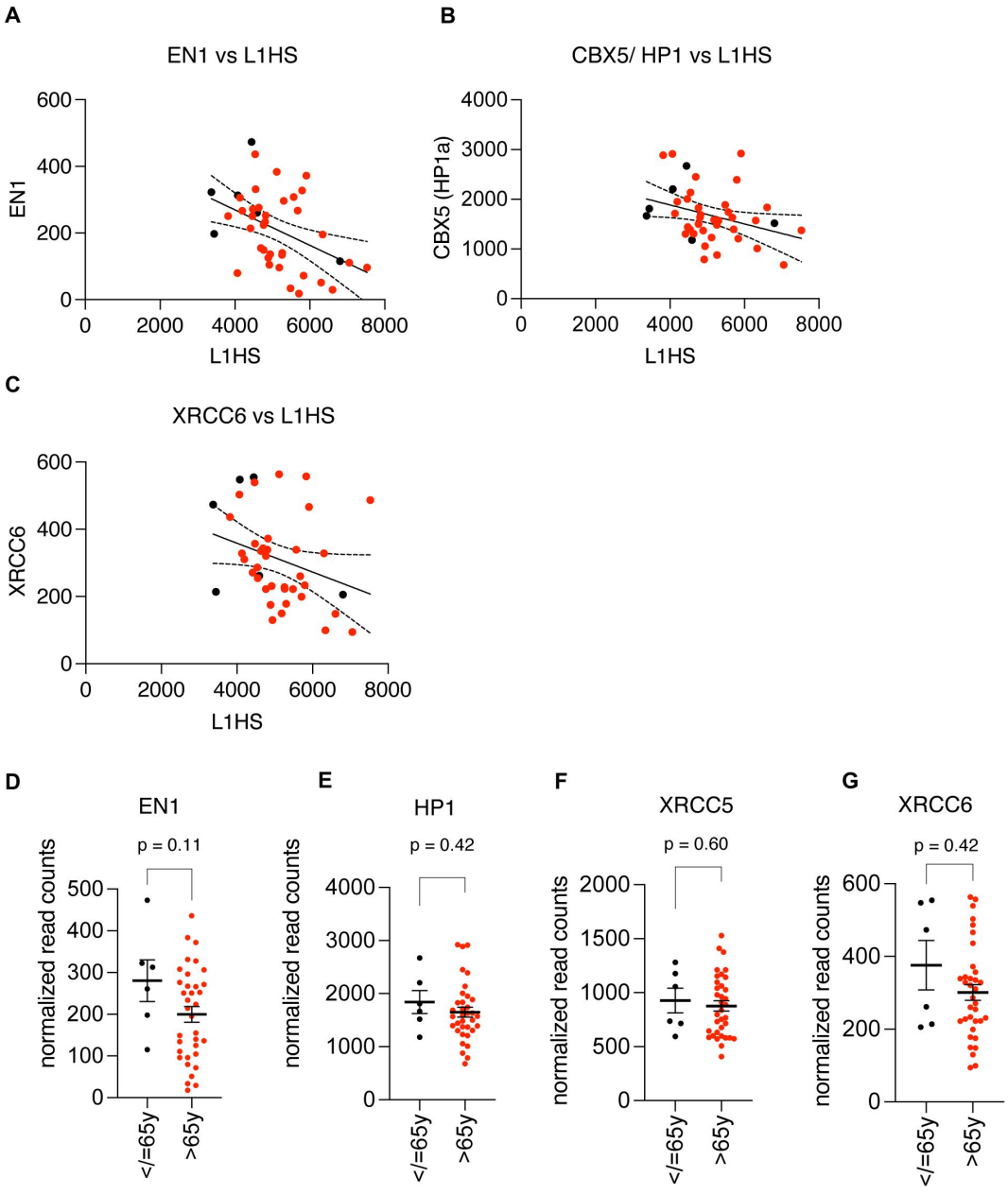
Suppl Fig. 3

(A) Proportion of ORF1p+ cells being neuronal in the ventral midbrain comparing young and aged mice as quantified using confocal approach. Kolmogorov-Smirnov test; data is represented as mean ± SEM.
(B) Proportion of neurons expressing ORF1p in the ventral midbrain comparing young and aged mice as quantified using confocal approach. Kolmogorov-Smirnov test; data is represented as mean ± SEM.
(C) Proportion of neurons in the ventral midbrain comparing young and aged mice as quantified using confocal approach. Kolmogorov-Smirnov test; data is represented as mean ± SEM.



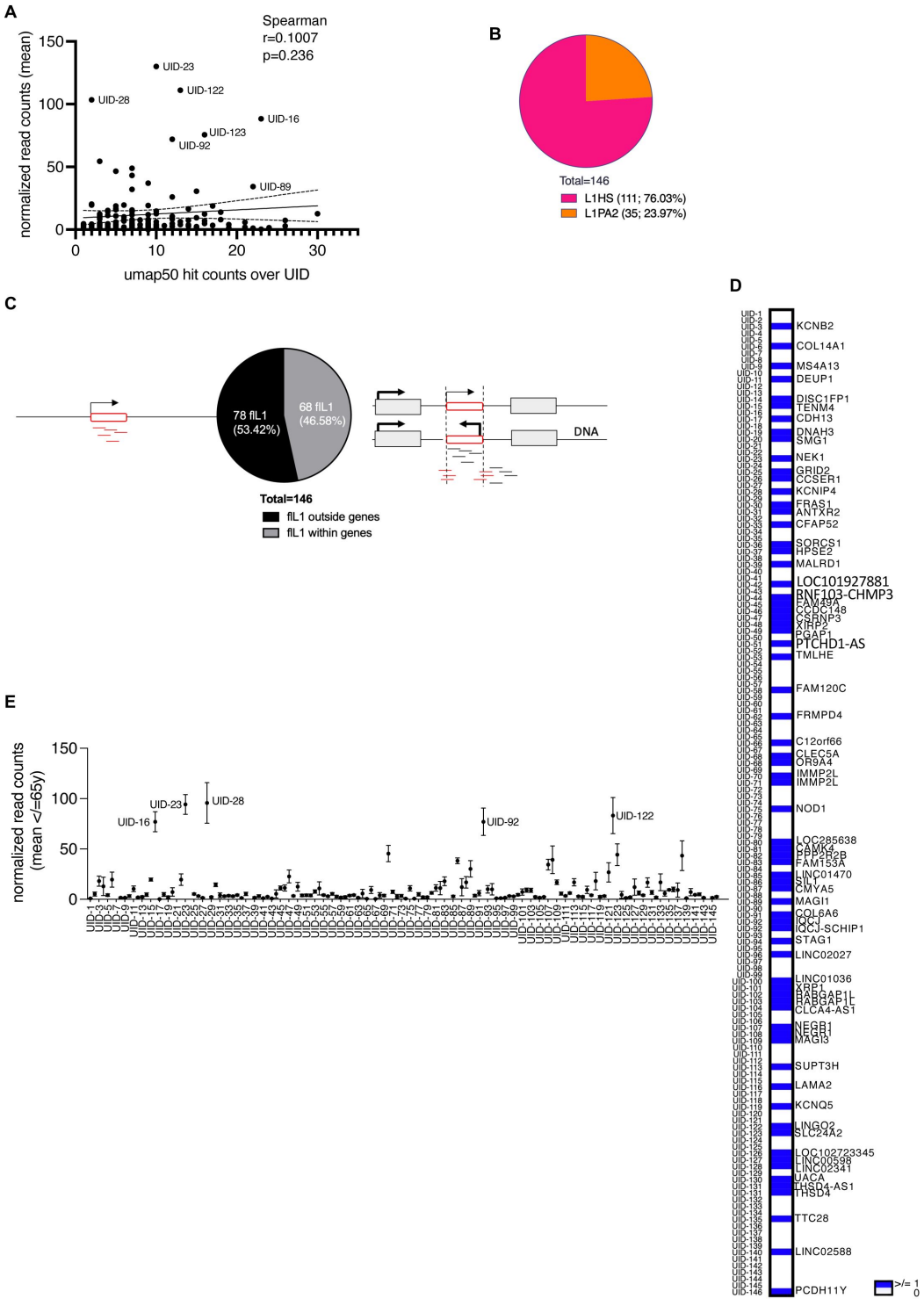
Suppl Fig. 4

(A-B) Comparison of the expression of dopaminergic markers tyrosine hydroxylase (TH, A) and LMX1B (B) between young (≤65y, n=6)) and aged (>65y, n=35) human dopaminergic neurons. Mann Whitney test.
(C, E) Volcano plot of differential expression analysis of TE expression using DEseq2 comparing young (≤65y, n=6) and aged (>65y, n=35) human dopaminergic neurons at the “class” (C) and “family” (E) level of RepeatMasker.
(D) Scatter plot comparing the expression of LINE at the “class” level between young (≤65y, n=6)) and aged (>65y, n=35) human dopaminergic neurons. Mann Whitney test.
(F) Scatter plot comparing the expression of LINE and Alu at the “family” level between young (≤65y, n=6)) and aged (>65y, n=35) human dopaminergic neurons. Mann Whitney test.
(G) Scatter plot comparing the expression of HERVH-int, HERV-Fc1 and two non-coding, non-autonomous but active TEs in the human genome, AluYa5 and SVA-F at the “name” level between young (≤65y, n=6)) and aged (>65y, n=35) human dopaminergic neurons. Mann Whitney test.



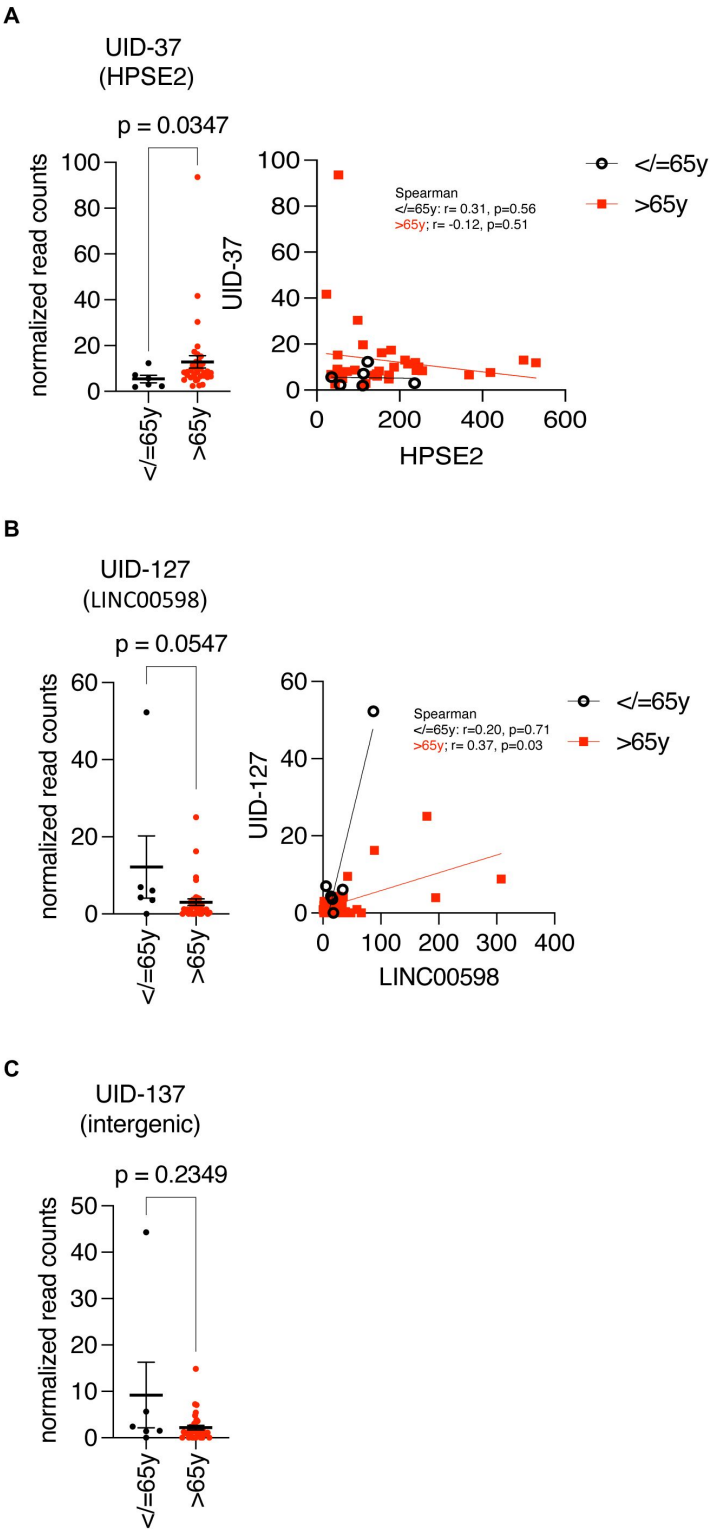
Suppl Fig. 5

(A-C) Correlation analyses of L1HS expression with Engrailed 1 (EN1, A, Spearman r=-0.43, p=0.002), CBX5/HP1 (B, Spearman r=-0.35, p=0.01) and XRCC6 expression (C, Spearman r= −0.394, p=0.005). Normalized read counts are plotted. Black dots correspond to young individuals (≤65y), red dots correspond to aged individuals (>65y).
(D-G) Scatter plots comparing the expression of EN1, CBX5/HP1, XRCC5 and XRCC6 between young (≤65y, n=6)) and aged (>65y, n=35) human dopaminergic neurons. Student's t-test (EN1) or Mann Whitney test (HP1, XRCC5/6).



Suppl Fig. 6

(A) Correlation of mappability (UMAP hit counts over UID, see methods) and UID expression (normalized read counts). Spearman correlation.
(B-D) *In silico* analysis of annotated full-length LINE-1 elements as in L1Basev2 (human reference genome hg38). (B) Percentage of L1HS and L1PA2 elements among the 146 full-length elements (UID1-146). (C) Percentage of full-length LINE-1 elements located inside or outside a gene. (D) Presence (blue, with gene symbol) or absence (white) of a “hosting” gene among the 146 annotated full-length LINE-1 in the human reference genome
(E) Mean expression of all 146 full-length LINE-1 elements in dopaminergic neurons of all individuals ≤65y.



Suppl Fig. 7

(A-C) Dysregulated locus-specific full-length LINE-1 elements are plotted as scatter plots comparing young (≤65y, n=6) and aged (>65y, n=35) human dopaminergic neurons. (A) UID-37 is located in an intron of the gene HPSE2 (left). Spearman correlation analysis of the expression of UID-37 and HPSE2 in young (≤65y, n=6, black dots) and aged (>65y, n=35, red squares) human dopaminergic neurons. (B) UID-127 is located within the 6th intron of the non-coding RNA LINC00598 (left). Spearman correlation analysis of the expression of UID-127 and LINC00598 in young (≤65y, n=6, black dots) and aged (>65y, n=35, red squares) human dopaminergic neurons. (C) UID-137 is intergenic.

References

1. Lu S, Zhang J, Lian X, et al. (2019) **A hidden human proteome encoded by 'non-coding' genes** *Nucleic Acids Research* **47**:8111–8125 https://doi.org/10.1093/nar/gkz646

2. Lander ES, Linton LM, Birren B, et al. (2001) **Initial sequencing and analysis of the human genome** *Nature* **409**:860–921 https://doi.org/10.1038/35057062

3. Brouha B, Schustak J, Badge RM, et al. (2003) **Hot L1s account for the bulk of retrotransposition in the human population** *Proceedings of the National Academy of Sciences of the United States of America* **100**:5280–5285 https://doi.org/10.1073/pnas.0831042100

4. Penzkofer T, Jäger M, Figlerowicz M, et al. (2017) **L1Base 2: more retrotransposition-active LINE-1s, more mammalian genomes** *Nucleic Acids Research* **45**:D68–D73 https://doi.org/10.1093/nar/gkw925

5. Mao J, Zhang Q, Cong YS (2021) **Human endogenous retroviruses in development and disease** *Computational and Structural Biotechnology Journal* **19**:5978–5986 https://doi.org/10.1016/j.csbj.2021.10.037

6. Esnault C, Maestre J, Heidmann T (2000) **Human LINE retrotransposons generate processed pseudogenes** *Nat Genet* **24**:363–367 https://doi.org/10.1038/74184

7. Wei W, Gilbert N, Ooi SL, et al. (2001) **Human L1 retrotransposition: cis preference versus trans complementation** *Molecular and Cellular Biology* **21**:1429–1439 https://doi.org/10.1128/MCB.21.4.1429-1439.2001

8. Mathias SL, Scott AF, Kazazian HH, Boeke JD, Gabriel A (1991) **Reverse transcriptase encoded by a human transposable element** *Science* **254**:1808–1810 https://doi.org/10.1126/science.1722352

9. Feng Q, Moran JV, Kazazian HH, Boeke JD (1996) **Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition** *Cell* **87**:905–916

10. Chen L, Dahlstrom JE, Lee SH, Rangasamy D (2012) **Naturally occurring endo-siRNA silences LINE-1 retrotransposons in human cells through DNA methylation** *Epigenetics : official journal of the DNA Methylation Society* **7**:758–771 https://doi.org/10.4161/epi.20706

11. Goodier JL, Cheung LE, Kazazian HH (2012) **MOV10 RNA helicase is a potent inhibitor of retrotransposition in cells** *PLoS Genetics* **8** https://doi.org/10.1371/journal.pgen.1002941

12. Philippe C, Vargas-Landin DB, Doucet AJ, et al. (2016) **Activation of individual L1 retrotransposon instances is restricted to cell-type dependent permissive loci** *eLife* **5** https://doi.org/10.7554/eLife.13926

13. Peze-Heidsieck E, Bonnifet T, Znaidi R, et al. (2021) **Retrotransposons as a Source of DNA Damage in Neurodegeneration** *Frontiers in aging neuroscience* **13** https://doi.org/10.3389/fnagi.2021.786897

14. Driver CJ, McKechnie SW (1992) **Transposable elements as a factor in the aging of Drosophila melanogaster** *Annals of the New York Academy of Sciences* **673**:83–91 https://doi.org/10.1111/j.1749-6632.1992.tb27439.x

15. Copley KE, Shorter J (2023) **Repetitive elements in aging and neurodegeneration** *Trends in Genetics* https://doi.org/10.1016/j.tig.2023.02.008

16. Gorbunova V, Seluanov A, Mita P, et al. (2021) **The role of retrotransposable elements in ageing and age-associated diseases** *Nature* :1–11 https://doi.org/10.1038/s41586-021-03542-y

17. Blaudin de Thé FX, Rekaik H, Peze-Heidsieck E, et al. (2018) **Engrailed homeoprotein blocks degeneration in adult dopaminergic neurons through LINE-1 repression** *The EMBO Journal* **37** https://doi.org/10.15252/embj.201797374

18. Bodea GO, Botto JM, Ferreiro ME, et al. (2024) **LINE-1 retrotransposons contribute to mouse PV interneuron development** *Nat Neurosci. Published online May* **21** https://doi.org/10.1038/s41593-024-01650-2

19. Sur D, Kustwar RK, Budania S, et al. (2017) **Detection of the LINE-1 retrotransposonRNA-binding protein ORF1p in differentanatomical regions of the human brain** *Mobile DNA* :1–12 https://doi.org/10.1186/s13100-017-0101-4

20. McKerrow W, Kagermazova L, Doudican N, et al. (2023) **LINE-1 retrotransposon expression in cancerous, epithelial and neuronal cells revealed by 5' single-cell RNA-Seq** *Nucleic Acids Research* **51**:2033–2045 https://doi.org/10.1093/nar/gkad049

21. Gasior SL, Wakeman TP, Xu B, Deininger PL (2006) **The human LINE-1 retrotransposon creates DNA double-strand breaks** *Journal of molecular biology* **357**:1383–1393 https://doi.org/10.1016/j.jmb.2006.01.089

22. Belgnaoui SM, Gosden RG, Semmes OJ, Haoudi A (2006) **Human LINE-1 retrotransposon induces DNA damage and apoptosis in cancer cells** *Cancer cell international* **6** https://doi.org/10.1186/1475-2867-6-13

23. De Cecco M, Criscione SW, Peckham EJ, et al. (2013) **Genomes of replicatively senescent cells undergo global epigenetic changes leading to gene silencing and activation of transposable elements** *Aging Cell* **12**:247–256 https://doi.org/10.1111/acel.12047

24. Van Meter M, Kashyap M, Rezazadeh S, et al. (2014) **SIRT6 represses LINE1 retrotransposons by ribosylating KAP1 but this repression fails with stress and age** *Nature communications* **5** https://doi.org/10.1038/ncomms6011

25. Sturm Á, Ivics Z, Vellai T (2015) **The mechanism of ageing: primary role of transposable elements in genome disintegration** *Cell Mol Life Sci* **72**:1839–1847 https://doi.org/10.1007/s00018-015-1896-0

26. Wallace N, Wagstaff BJ, Deininger PL, Roy-Engel AM (2008) **LINE-1 ORF1 protein enhances Alu SINE retrotransposition** *Gene* **419**:1–6 https://doi.org/10.1016/j.gene.2008.04.007

27. Thomas CA, Tejwani L, Trujillo CA, et al. (2017) **Modeling of TREX1-Dependent Autoimmune Disease using Human Stem Cells Highlights L1 Accumulation as a Source of Neuroinflammation** *Cell stem cell* **21**:319–331 https://doi.org/10.1016/j.stem.2017.07.009

28. De Cecco M, Ito T, Petrashen AP, et al. (2019) **L1 drives IFN in senescent cells and promotes age-associated inflammation** *Nature* **566**:73–78 https://doi.org/10.1038/s41586-018-0784-9

29. Luqman-Fatah A, Watanabe Y, Uno K, Ishikawa F, Moran JV, Miyoshi T (2023) **The interferon stimulated gene-encoded protein HELZ2 inhibits human LINE-1 retrotransposition and LINE-1 RNA-mediated type I interferon induction** *Nat Commun* **14** https://doi.org/10.1038/s41467-022-35757-6

30. Gazquez-Gutierrez A, Witteveldt J, Heras S R, Macias S (2021) **Sensing of transposable elements by the antiviral innate immune system** *RNA (New York, NY)* https://doi.org/10.1261/rna.078721.121

31. Krug L, Chatterjee N, Borges-Monroy R, et al. (2017) **Retrotransposon activation contributes to neurodegeneration in a Drosophila TDP-43 model of ALS** *PLoS Genetics* **13** https://doi.org/10.1371/journal.pgen.1006635

32. Casale AM, Liguori F, Ansaloni F, et al. (2022) **Transposable element activation promotes neurodegeneration in a Drosophila model of Huntington's disease** *ISCIENCE* **25** https://doi.org/10.1016/j.isci.2021.103702

33. Takahashi T, Stoiljkovic M, Song E, et al. (2022) **LINE-1 activation in the cerebellum drives ataxia** *Neuron* :1–19 https://doi.org/10.1016/j.neuron.2022.08.011

34. De Luca C, Gupta A, Bortvin A (2023) **Retrotransposon LINE-1 bodies in the cytoplasm of piRNA-deficient mouse spermatocytes: Ribonucleoproteins overcoming the integrated stress response. Feschotte C, ed** *PLoS Genet* **19** https://doi.org/10.1371/journal.pgen.1010797

35. Spencley AL, Bar S, Swigut T, et al. (2023) **Co-transcriptional genome surveillance by HUSH is coupled to termination machinery** *Mol Cell* **83**:1623–1639 https://doi.org/10.1016/j.molcel.2023.04.014

36. Garland W, Müller I, Wu M, et al. (2022) **Chromatin modifier HUSH co-operates with RNA decay factor NEXT to restrict transposable element expression** *Mol Cell* **82**:1691–1707 https://doi.org/10.1016/j.molcel.2022.03.004

37. Shirane K, Miura F, Ito T, Lorincz MC (2020) **NSD1-deposited H3K36me2 directs de novo methylation in the mouse male germline and counteracts Polycomb-associated silencing** *Nat Genet* **52**:1088–1098 https://doi.org/10.1038/s41588-020-0689-z

38. Guerrero A, Herranz N, Sun B, et al. (2019) **Cardiac glycosides are broad-spectrum senolytics** *Nat Metab* **1**:1074–1088 https://doi.org/10.1038/s42255-019-0122-z

39. Gusel'nikova VV, Korzhevskiy DE (2015) **NeuN As a Neuronal Nuclear Antigen and Neuron Differentiation Marker** *Acta Naturae* **7**:42–47 https://doi.org/10.32607/20758251-2015-7-2-42-47

40. Cannon JR, Greenamyre JT (2009) **NeuN is not a reliable marker of dopamine neurons in rat substantia nigra** *Neuroscience Letters* **464**:14–17 https://doi.org/10.1016/j.neulet.2009.08.023

41. Sukhorukova EG (2014) **Nuclear Protein NeuN in Neurons in the Human Substantia Nigra** *Neurosci Behav Physi* **44**:539–541 https://doi.org/10.1007/s11055-014-9947-9

42. Kumar SS, Buckmaster PS (2007) **Neuron-specific nuclear antigen NeuN is not detectable in gerbil subtantia nigra pars reticulata** *Brain Res* **1142**:54–60 https://doi.org/10.1016/j.brainres.2007.01.027

43. Keller D, Erö C, Markram H (2018) **Cell Densities in the Mouse Brain: A Systematic Review** *Front Neuroanat* **12** https://doi.org/10.3389/fnana.2018.00083

44. Sun W, Cornwell A, Li J, et al. (2017) **SOX9 Is an Astrocyte-Specific Nuclear Marker in the Adult Brain Outside the Neurogenic Regions** *J Neurosci* **37**:4493–4507 https://doi.org/10.1523/JNEUROSCI.3199-16.2017

45. Yushkova E, Moskalev A (2023) **Transposable elements and their role in aging** *Ageing Res Rev* **86** https://doi.org/10.1016/j.arr.2023.101881

46. Gibb WR, Lees AJ (1991) **Anatomy, pigmentation, ventral and dorsal subpopulations of the substantia nigra, and differential cell death in Parkinson's disease** *J Neurol Neurosurg Psychiatry* **54**:388–396 https://doi.org/10.1136/jnnp.54.5.388

47. Dong X, Liao Z, Gritsch D, et al. (2018) **Enhancers active in dopamine neurons are a primary link between genetic variation and neuropsychiatric disease** *Nature Neuroscience* :1–19 https://doi.org/10.1038/s41593-018-0223-0

48. Teissandier A, Servant N, Barillot E, Bourc'his D (2019) **Tools and best practices for retrotransposon analysis using high-throughput sequencing data** *Mobile DNA* **10** https://doi.org/10.1186/s13100-019-0192-1

49. Liu N, Lee CH, Swigut T, et al. (2018) **Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators** *Nature* **553**:228–232 https://doi.org/10.1038/nature25179

50. Rekaik H, Blaudin de Thé FX, Prochiantz A, Fuchs J, Joshi RL (2015) **Dissecting the role of Engrailed in adult dopaminergic neurons–Insights into Parkinson disease pathogenesis** *FEBS Letters* **589**:3786–3794 https://doi.org/10.1016/j.febslet.2015.10.002

51. Maeda R, Tachibana M (2022) **HP1 maintains protein stability of H3K9 methyltransferases and demethylases** *EMBO reports* **23** https://doi.org/10.15252/embr.202153581

52. Suzuki J, Yamaguchi K, Kajikawa M, et al. (2009) **Genetic Evidence That the Non-Homologous End-Joining Repair Pathway Is Involved in LINE Retrotransposition** *PLoS Genetics* **5** https://doi.org/10.1371/journal.pgen.1000461.g005

53. Goerner-Potvin P, Bourque G (2018) **Computational tools to unmask transposable elements** *Nature Reviews Genetics* :1–17 https://doi.org/10.1038/s41576-018-0050-x

54. Schwarz R, Koch P, Wilbrandt J, Hoffmann S (2022) **Locus-specific expression analysis of transposable elements** *Briefings in Bioinformatics* **23** https://doi.org/10.1093/bib/bbab417

55. Faulkner GJ (2023) **Elevated L1 expression in ataxia telangiectasia likely explained by an RNA-seq batch effect** *Neuron* **111**:610–611 https://doi.org/10.1016/j.neuron.2023.02.007

56. Lanciano S, Cristofari G (2020) **Measuring and interpreting transposable element expression** *Nature Reviews Genetics* **509** https://doi.org/10.1038/s41576-020-0251-y

57. Martin SL, Li J, Weisz JA (2000) **Deletion analysis defines distinct functional domains for protein-protein and nucleic acid interactions in the ORF1 protein of mouse LINE-1** *J Mol Biol* **304**:11–20 https://doi.org/10.1006/jmbi.2000.4182

58. Briggs EM, McKerrow W, Mita P, Boeke JD, Logan SK, Fenyo D (2021) **RIP-seq reveals LINE-1 ORF1p association with p-body enriched mRNAs** *Mobile DNA* :1–13 https://doi.org/10.1186/s13100-021-00233-3

59. Kulpa DA, Moran JV (2005) **Ribonucleoprotein particle formation is necessary but not sufficient for LINE-1 retrotransposition** *Human Molecular Genetics* **14**:3237–3248 https://doi.org/10.1093/hmg/ddi354

60. Taylor MS, Altukhov I, Molloy KR, et al. (2018) **Dissection of affinity captured LINE-1 macromolecular complexes** *eLife* **7**:210–241 https://doi.org/10.7554/eLife.30094

61. Kelly MP, Adamowicz W, Bove S, et al. (2014) **Select 3',5'-cyclic nucleotide phosphodiesterases exhibit altered expression in the aged rodent brain** *Cell Signal* **26**:383–397 https://doi.org/10.1016/j.cellsig.2013.10.007

62. Lakics V, Karran EH, Boess FG (2010) **Quantitative comparison of phosphodiesterase mRNA distribution in human brain and peripheral tissues** *Neuropharmacology* **59**:367–374 https://doi.org/10.1016/j.neuropharm.2010.05.004

63. Mandel S, Gozes I (2007) **Activity-dependent neuroprotective protein constitutes a novel element in the SWI/SNF chromatin remodeling complex** *J Biol Chem* **282**:34448–34456 https://doi.org/10.1074/jbc.M704756200

64. Singh A, Modak SB, Chaturvedi MM, Purohit JS (2023) **SWI/SNF Chromatin Remodelers: Structural, Functional and Mechanistic Implications** *Cell Biochem Biophys* **81**:167–187 https://doi.org/10.1007/s12013-023-01140-5

65. Goodier JL, Cheung LE, Kazazian HH (2013) **Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition** *Nucleic Acids Research* **41**:7401–7419 https://doi.org/10.1093/nar/gkt512

66. Taylor MS, LaCava J, Mita P, et al. (2013) **Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition** *Cell* **155**:1034–1048 https://doi.org/10.1016/j.cell.2013.10.021

67. Moldovan JB, Moran JV (2015) **The Zinc-Finger Antiviral Protein ZAP Inhibits LINE and Alu Retrotransposition** *PLoS Genetics* **11** https://doi.org/10.1371/journal.pgen.1005121

68. Vuong LM, Pan S, Donovan PJ (2019) **Proteome Profile of Endogenous Retrotransposon-Associated Complexes in Human Embryonic Stem Cells** *Proteomics* **19** https://doi.org/10.1002/pmic.201900169

69. De Luca C, Gupta A, Bortvin A (2023) **Ribonucleoprotein Condensation Driven by Retrotransposon LINE-1 Sustains RNA Integrity and Translation in Mouse Spermatocytes** *Molecular Biology* https://doi.org/10.1101/2023.01.09.523313

70. Ardeljan D, Wang X, Oghbaie M, et al. (2020) **LINE-1 ORF2p expression is nearly imperceptible in human cancers** *Mobile DNA* **11** https://doi.org/10.1186/s13100-019-0191-2

71. Liu Q, Yi D, Ding J, et al. (2023) **MOV10 recruits DCP2 to decap human LINE-1 RNA by forming large cytoplasmic granules with phase separation properties** *EMBO reports* **24** https://doi.org/10.15252/embr.202256512

72. McKerrow W, Wang X, Mendez-Dorantes C, et al. (2022) **LINE-1 expression in cancer correlates with p53 mutation, copy number alteration, and S phase checkpoint** *Proc Natl Acad Sci U S A* **119** https://doi.org/10.1073/pnas.2115999119

73. Rodić N, Sharma R, Sharma R, et al. (2014) **Long Interspersed Element-1 Protein Expression Is a Hallmark of Many Human Cancers** *Am J Pathol* **184**:1280–1286 https://doi.org/10.1016/j.ajpath.2014.01.007

74. Osburn SC, Mesquita PHC, Neal FK, et al. (2022) **Long-term voluntary wheel running effects on markers of long interspersed nuclear element-1 in skeletal muscle, liver, and brain tissue of female rats** *American Journal of Physiology-Cell Physiology* **323**:C907–C919 https://doi.org/10.1152/ajpcell.00234.2022

75. Rybacki K, Xia M, Ahsan MU, Xing J, Wang K (2023) **Assessing the Expression of Long INterspersed Elements (LINEs) via Long-Read Sequencing in Diverse Human Tissues and Cell Lines** *Genes* **14** https://doi.org/10.3390/genes14101893

76. Doucet-O'Hare TT, Rodić N, Sharma R, et al. (2015) **LINE-1 expression and retrotransposition in Barrett's esophagus and esophageal carcinoma** *Proceedings of the National Academy of Sciences* **112**:E4894–E4900 https://doi.org/10.1073/pnas.1502474112

77. Zhang Y, Amaral ML, Zhu C, et al. (2022) **Single-cell epigenome analysis reveals age-associated decay of heterochromatin domains in excitatory neurons in the mouse brain** *Cell Res* **32**:1008–1021 https://doi.org/10.1038/s41422-022-00719-6

78. Garza R, Atacho D, Adami A, et al. (2023) **L1 retrotransposons drive human neuronal transcriptome complexity and functional diversification** *Science Advances* **9** https://doi.org/10.1126/sciadv.adh9543

79. Watanabe R, Nakachi Y, Matsubara H, et al. (2023) **Identification of epigenetically active L1 promoters in the human brain and their relationship with psychiatric disorders** *Neuroscience Research* https://doi.org/10.1016/j.neures.2023.05.001

80. Takahashi F, Zhang C, Hohjoh H, et al. (2022) **Immune-mediated neurodegenerative trait provoked by multimodal derepression of long-interspersed nuclear element-1** *iScience* **25** https://doi.org/10.1016/j.isci.2022.104278

81. Simon M, Van Meter M, Ablaeva J, et al. (2019) **LINE1 Derepression in Aged Wild-Type and SIRT6-Deficient Mice Drives Inflammation** *Cell metabolism* **29**:871–885 https://doi.org/10.1016/j.cmet.2019.02.014

82. Lopez-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. (2023) **Hallmarks of aging: An expanding universe** *Cell* https://doi.org/10.1016/j.cell.2022.11.001

83. Mumford PW, Romero MA, Osburn SC, et al. (2019) **Skeletal muscle LINE-1 retrotransposon activity is upregulated in older versus younger rats** *Am J Physiol Regul Integr Comp Physiol* **317**:R397–R406 https://doi.org/10.1152/ajpregu.00110.2019

84. Traxler L, Lucciola R, Herdy JR, Jones JR, Mertens J, Gage FH (2023) **Neural cell state shifts and fate loss in ageing and age-related diseases** *Nat Rev Neurol* **19**:434–443 https://doi.org/10.1038/s41582-023-00815-0

85. Frost B (2023) **Alzheimer's disease and related tauopathies: disorders of disrupted neuronal identity** *Trends in Neurosciences* https://doi.org/10.1016/j.tins.2023.07.006

86. Unger JW (1998) **Glial reaction in aging and Alzheimer's disease** *Microsc Res Tech* **43**:24–28 https://doi.org/10.1002/(SICI)1097-0029(19981001)43:1<24::AID-JEMT4>3.0.CO;2-P

87. Dai L, LaCava J, Taylor MS, Boeke JD (2014) **Expression and detection of LINE-1 ORF-encoded proteins** *Mobile genetic elements* **4** https://doi.org/10.4161/mge.29319

88. Whongsiri P, Goering W, Lautwein T, et al. (2020) **Many Different LINE-1 Retroelements Are Activated in Bladder Cancer** *International Journal of Molecular Sciences* **21**:9433–14 https://doi.org/10.3390/ijms21249433

89. Lindtner S, Felber BK, Kjems J (2002) **An element in the 3' untranslated region of human LINE-1 retrotransposon mRNA binds NXF1(TAP) and can function as a nuclear export element** *RNA* **8**:345–356 https://doi.org/10.1017/s1355838202027759

90. Naufer MN, Furano AV, Williams MC (2019) **Protein-nucleic acid interactions of LINE-1 ORF1p** *Seminars in Cell & Developmental Biology* **86**:140–149 https://doi.org/10.1016/j.semcdb.2018.03.019

91. Garneau NL, Wilusz J, Wilusz CJ (2007) **The highways and byways of mRNA decay** *Nat Rev Mol Cell Biol* **8**:113–126 https://doi.org/10.1038/nrm2104

92. Lykke-Andersen J (2002) **Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay** *Mol Cell Biol* **22**:8114–8121 https://doi.org/10.1128/MCB.22.23.8114-8121.2002

93. Threlfell S, Sammut S, Menniti FS, Schmidt CJ, West AR (2009) **Inhibition of Phosphodiesterase 10A Increases the Responsiveness of Striatal Projection Neurons to Cortical Stimulation** *J Pharmacol Exp Ther* **328**:785–795 https://doi.org/10.1124/jpet.108.146332

94. Li N, Chen X, Zhu B, et al. (2015) **Suppression of β-catenin/TCF transcriptional activity and colon tumor cell growth by dual inhibition of PDE5 and 10** *Oncotarget* **6**:27403–27415 https://doi.org/10.18632/oncotarget.4741

95. Pereira GC, Sanchez L, Schaughency PM, et al. (2018) **Properties of LINE-1 proteins and repeat element expression in the context of amyotrophic lateral sclerosis** *Mobile DNA* **9** https://doi.org/10.1186/s13100-018-0138-z

96. Mita P, Wudzinska A, Sun X, et al. (2018) **LINE-1 protein localization and functional dynamics during the cell cycle** *eLife* **7** https://doi.org/10.7554/eLife.30058

97. Newton JC, Naik MT, Li GY, et al. (2021) **Phase separation of the LINE-1 ORF1 protein is mediated by the N-terminus and coiled-coil domain** *Biophysical journal* **120**:2181–2191 https://doi.org/10.1016/j.bpj.2021.03.028

98. Linkert M, Rueden CT, Allan C, et al. (2010) **Metadata matters: access to image data in the real world** *J Cell Biol* **189**:777–782 https://doi.org/10.1083/jcb.201004104

99. Ollion J, Cochennec J, Loll F, Escudé C, Boudier T (2013) **TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization** *Bioinformatics (Oxford, England)* **29**:1840–1841 https://doi.org/10.1093/bioinformatics/btt276

100. Stringer C, Wang T, Michaelos M, Pachitariu M (2021) **Cellpose: a generalist algorithm for cellular segmentation** *Nat Methods* **18**:100–106 https://doi.org/10.1038/s41592-020-01018-x

101. Wang Q, Ding SL, Li Y, et al. (2020) **The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas** *Cell* **181**:936–953 https://doi.org/10.1016/j.cell.2020.04.007

102. Chiaruttini N, Burri O, Haub P, Guiet R, Sordet-Dessimoz J, Seitz A (2022) **An Open-Source Whole Slide Image Registration Workflow at Cellular Precision Using Fiji** *QuPath and Elastix. Front Comput Sci* **3** https://doi.org/10.3389/fcomp.2021.780026

103. Bankhead P, Loughrey MB, Fernández JA, et al. (2017) **QuPath: Open source software for digital pathology image analysis** *Sci Rep* **7** https://doi.org/10.1038/s41598-017-17204-5

104. Schmidt U, Weigert M, Broaddus C, Myers G, Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G (2018) **Cell Detection with Star-Convex Polygons** *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018: Lecture Notes in Computer Science* :265–273 https://doi.org/10.1007/978-3-030-00934-2_30

105. McKinney W (2010) **Data Structures for Statistical Computing in Python** *SciPy Proceedings* :56–61 https://doi.org/10.25080/Majora-92bf1922-00a

106. Liao Y, Smyth GK, Shi W (2019) **The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads** *Nucleic Acids Res* **47** https://doi.org/10.1093/nar/gkz114

107. Valot B, Langella O, Nano E, Zivy M (2011) **MassChroQ: a versatile tool for mass spectrometry quantification** *Proteomics* **11**:3572–3577 https://doi.org/10.1002/pmic.201100120

108. Perez-Riverol Y, Bai J, Bandla C, et al. (2022) **The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences** *Nucleic Acids Res* **50**:D543–D552 https://doi.org/10.1093/nar/gkab1038

109. Shannon P, Markiel A, Ozier O, et al. (2003) **Cytoscape: a software environment for integrated models of biomolecular interaction networks** *Genome Res* **13**:2498–2504 https://doi.org/10.1101/gr.1239303

Editors

Reviewing Editor
Ashley Webb
Buck Institute for Research on Aging, Novato, United States of America

Senior Editor
Pankaj Kapahi
Buck Institute for Research on Aging, Novato, United States of America

Reviewer #1 (Public Review):

Summary:

In this study, Bonnifet et al. profile the presence of L1 ORF1p in the mouse and human brain. They claim that ORF1p is expressed in the human and mouse brain at a steady state and that there is an age-dependent increase in expression. This is a timely report as two recent papers have extensively documented the presence of full-length L1 transcripts in the mouse and human brain (PMID: 38773348 & PMID: 37910626). Thus, the finding that L1 ORF1p is consistently expressed in the brain is not surprising, but important to document.

Strengths:

Several parts of this manuscript appear to be well done and include the necessary controls. In particular, the evidence for steady-state expression of ORF1p in the mouse brain appears robust.

Weaknesses:

Several parts of the manuscript appear to be more preliminary and need further experiments to validate their claims. In particular, the data suggesting expression of L1 ORF1p in the human brain and the data suggesting increased expression in the aged brain need further validation. Detailed comments:

(1) The expression of ORF1p in the human brain shown in Figure 1j is not convincing. Why are there two strong bands in the WB? How can the authors be sure that this signal represents ORF1p expression and not non-specific labelling? Additional validations and controls are needed to verify the specificity of this signal.

(2) The data shown in Figure 2g are not convincing. How can the authors be sure that this signal represents ORF1p expression and not non-specific labelling? Extensive additional validations and controls are needed to verify the specificity of this signal.

(3) The data showing a reduction in ORF1p expression in the aged mouse brain is confusing and maybe even misleading. Although there is an increase in the intensity of the ORF1p signal in ORF1p+ cells, the data clearly shows that fewer cells express ORF1p in the aged brain. If these changes indicate an overall loss or gain of ORF1p, expression in the aged brain is not resolved. Thus, conclusions should be more carefully phrased in this section. It is important to show the quantification of NeuN+ and NeuN- cells in young vs aged (not only the proportions as shown in Figure 3b) to determine if the difference in the number of ORF1p+ cells is due to loss of neurons or perhaps a sampling issue. More so, it would be essential to perform WB and/or proteomics experiments to complement the IHC data for the aged mouse samples.

(4) The transcriptomic data presented in Figure 4 and Figure 5 are not convincing. Quantification of transposon expression on short read sequencing has important limitations. Longer reads and complementary approaches are needed to study the expression of evolutionarily young L1s (see PMID: 38773348 & PMID: 37910626 for examples of the current state of the art). Given the read length and the unstranded sequencing approach, I would at least ask the authors to add genome browser tracks of the upregulated loci so that we can properly assess the clarity of the results. I would also suggest adding the mappability profile of the elements in question. In addition, since this manuscript focuses on ORF1p, it would be essential to document changes in protein levels (and not just transcripts) in the ageing human brain.

(5) More information is needed on RNAseq of microdissections of dopaminergic neurons from 'healthy' post-mortem samples of different ages. No further information on these

samples is provided. I would suggest adding a table with the clinical information of these samples (especially age, sex, and cause of death). The authors should also discuss whether this experiment has sufficient power. The human ageing cohort seems very small to me.

(6) The findings in this manuscript apply to both human and mouse brains. However, the landscape of the evolutionarily young L1 subfamilies between these two species is very different and should be part of the discussion. For example, the regulatory sequences that drive L1 expression are quite different in human and mouse L1s. This should be discussed.

(7) On page 3 the authors write: "generally accepted that TE activation can be both, a cause and consequence of aging". This statement does not reflect the current state of the field. On the contrary, this is still an area of extensive investigation and many of the findings supporting this hypothesis need to be confirmed in independent studies. This statement should be revised to reflect this reality.

https://doi.org/10.7554/eLife.100687.1.sa3

Reviewer #2 (Public Review):

Summary:

Bonnifet et al. sought to characterize the expression pattern of L1 ORF1p expression across the entire mouse brain, in young and aged animals, and to corroborate their characterization with Western blotting for L1 ORF1p and L1 RNA expression data from human samples. They also queried L1 ORF1p interacting partners in the mouse brain by IP-MS.

Strengths:

A major strength of the study is the use of two approaches: a deep-learning detection method to distinguish neuronal vs. non-neuronal cells and ORF1p+ cells vs. ORF1p- cells across large-scale images encompassing multiple brain regions mapped by comparison to the Allen Brain Atlas, and confocal imaging to give higher resolution on specific brain regions. These results are also corroborated by Western blotting on six mouse brain regions. Extension of their analysis to post-mortem human samples, to the extent possible, is another strength of the paper. The identification of novel ORF1p interactors in the brain is also a strength in that it provides a novel dataset for future studies.

Weaknesses:

The main weakness of the study is that cell type specificity of ORF1p expression was not examined beyond neuron (NeuN+) vs non-neuron (NeuN-). Indeed, a recent study (Bodea et al. 2024, Nature Neuroscience) found that ORF1p expression is characteristic of parvalbumin-positive interneurons, and it would be very interesting to query whether other neuronal subtypes in different brain regions are distinguished by ORF1p expression. The data suggesting that ORF1p expression is increased in aged mouse brains is intriguing, although it seems to be based upon modestly (up to 27%, dependent on brain region) higher intensity of ORF1p staining rather than a higher proportion of ORF1+ neurons. Indeed, the proportion of NeuN+/Orf1p+ cells actually decreased in aged animals. It is difficult to interpret the significance and validity of the increase in intensity, as Hoechst staining of DNA, rather than immunostaining for a protein known to be stably expressed in young and aged neurons, was used as a control for staining intensity. The main weakness of the IP-MS portion of the study is that none of the interactors were individually validated or subjected to follow-up analyses. The list of interactors was compared to previously published datasets, but not to ORF1p interactors in any other mouse tissue.

The authors achieved the goals of broadly characterizing ORF1p expression across different regions of the mouse brain, and identifying putative ORF1p interactors in the mouse brain. However, findings from both parts of the study are somewhat superficial in depth.

This provides a useful dataset to the field, which likely will be used to justify and support numerous future studies into L1 activity in the aging mammalian brain and in neurodegenerative disease. Similarly, the list of ORF1p interacting proteins in the brain will likely be taken up and studied in greater depth.

https://doi.org/10.7554/eLife.100687.1.sa2

Reviewer #3 (Public Review):

The question about whether L1 exhibits normal/homeostatic expression in the brain (and in general) is interesting and important. L1 is thought to be repressed in most somatic cells (with the exception of some stem/progenitor compartments). However, to our knowledge, this has not been authoritatively / systematically examined and the literature is still developing with respect to this topic. The full gamut of biological and pathobiological roles of L1 remains to be shown and elucidated and this area has garnered rapidly increasing interest, year-by-year. With respect to the brain, L1 (and repeat sequences in general) have been linked with neurodegeneration, and this is thought to be an aging-related consequence or contributor (or both) of inflammation. This study provides an impressive and apparently comprehensive imaging analysis of differential L1 ORF1p expression in mouse brain (with some supporting analysis of the human brain), compatible with a narrative of non-pathological expression of retrotransposition-competent L1 sequences. We believe this will encourage and support further research into the functional roles of L1 in normal brain function and how this may give way to pathological consequences in concert with aging. However, we have concerns with conclusions drawn, in some cases regardless of the lack of statistical support from the data. We note a lack of clarity about how the 3rd party pre-trained machine learning models perform on the authors' imaging data (validation/monitoring tests are not reported), as well as issues (among others) with the particular implementation of co-immunoprecipitation (ORF1p is not among the highly enriched proteins and apparently does not reach statistical significance for the comparison) - neither of which may be sufficiently rigorous.

https://doi.org/10.7554/eLife.100687.1.sa1

Author Response:

> ***Public Reviews:***
>
> ***Reviewer #1 (Public Review):***
>
> *Summary:*
>
> *In this study, Bonnifet et al. profile the presence of L1 ORF1p in the mouse and human brain. They claim that ORF1p is expressed in the human and mouse brain at a steady state and that there is an age-dependent increase in expression. This is a timely report as two recent papers have extensively documented the presence of full-length L1 transcripts in the mouse and human brain (PMID: 38773348 & PMID: 37910626). Thus, the finding that L1 ORF1p is consistently expressed in the brain is not surprising, but important to document.*

Thank you for recognizing the importance of this study. The two cited papers have indeed reported the presence of full-length transcripts in the mouse and human brain. However, the first (PMID: 38773348) report has shown evidence of flL1 RNA and ORF1 protein expression in the mouse hippocampus (but not elsewhere) and the second (PMID: 37910626) shows full-length LINE-1 RNA expression and H3K4me3-ChIP data in the frontal and temporal lobe of the human brain, but not protein expression.

> *Strengths:*
>
> *Several parts of this manuscript appear to be well done and include the necessary controls. In particular, the evidence for steady-state expression of ORF1p in the mouse brain appears robust.*
>
> *Weaknesses:*
>
> *Several parts of the manuscript appear to be more preliminary and need further experiments to validate their claims. In particular, the data suggesting expression of L1 ORF1p in the human brain and the data suggesting increased expression in the aged brain need further validation. Detailed comments:*
>
> *(1) The expression of ORF1p in the human brain shown in Figure 1j is not convincing. Why are there two strong bands in the WB? How can the authors be sure that this signal represents ORF1p expression and not nonspecific labelling? Additional validations and controls are needed to verify the specificity of this signal.*

We have validated the antibody (Abcam 245249 - https://www.abcam.com/en-us/products/primary-antibodies/line-1-orf1p-antibody-epr22227-6-ab245249), which we use for Western blotting experiments like in Fig1j), by several means. We have done immunoprecipitations (IPs) and co-immunoprecipitations (co-IPs) followed by quantitative mass spectrometry (LC-MS/MS). We efficiently detect ORF1p in IPs (Western blot) and by quantitative mass spectrometry (5 independent samples per IP-ORF1p and IP-IgG: ORF1p/IgG ratio: 40.86; adj p-value 8.7e-07; human neurons in culture). We also did co-IPs followed by Western blot using two different antibodies, the Millipore or the Abcam antibody to immunoprecipitate and the Abcam antibody for Western blotting (the Millipore AB does not work well on WB in our hands) which consistently showed a double band indicating that both bands are ORF1p-derived. We can provide this data to the revised manuscript, although some of it (the MS data) is subject of another study in preparation. Abcam also reports a double band, and they suspect that the lower band is a truncated form (see the link to their website above). ORF1p Western blots done by other labs with different antibodies have detected a second band in human samples

(1) Sato, S. et al. LINE-1 ORF1p as a candidate biomarker in high grade serous ovarian carcinoma. Sci Rep 13, 1537 (2023) in Figure 1D

(2) McKerrow, W. et al. LINE-1 expression in cancer correlates with p53 mutation, copy number alteration, and S phase checkpoint. Proc. Natl. Acad. Sci. U.S.A. 119, e2115999119 (2022)) showing a Western blot of an inducible LINE-1 (ORFeus) detected by the MABC1152 ORF1p antibody from Millipore Sigma in Figure 7 3) in a publication in eLife (Walter et al. eLife 2016;5:e11418. DOI: 10.7554/eLife.11418) in mouse ES cells with an antibody made in-house from another lab (gift) – Figure 2B

The lower band might thus be a truncated form of ORF1p or a degradation product which appears to be shared by mouse and human ORF1p. We will mention this in the revised version of the paper. In addition, we have used the very well characterized antibody from Millipore (https://www.merckmillipore.com/CH/en/product/Anti-LINE-1-ORF1p-Antibody-clone-4H1,MM_NF-MABC1152?ReferrerURL=https%3A%2F%2Fwww.google.com%2F) for

immunostainings and detect ORF1p staining in human neurons in the very same brain regions (Fig 2H) including the cerebellum (selectively in Purkinje cells as in mice in Fig1B panel 10: human images not shown).

Altogether, based on our experimental validations and evidence from the literature, we are very confident that it is ORF1p that we detect on the blots.

> *(2) The data shown in Figure 2g are not convincing. How can the authors be sure that this signal represents ORF1p expression and not non-specific labelling? Extensive additional validations and controls are needed to verify the specificity of this signal.*

Figure 2g shows a Western blot using an extensively used and well characterized ORF1p antibody from abcam (mouse ORF1p - (https://www.abcam.com/en-us/products/primary-antibodies/line-1-orf1p-antibody-epr21844108-ab216324; cited in at least 11 publications) after FACS-sorting of neurons (NeuN+) of the mouse brain. We have validated this ORF1p antibody ourselves in IPs (see Fig 6A) and co-IP followed by mass spectrometry (LC/MS-MS; see Fig 6, where we detect ORF1p exclusively in the 5 independent ORF1p-IP samples and not at all in 5 independent IgG-IP control samples, see Suppl Table 2). This together makes us very confident that we are looking at a specific ORF1p signal. Please note that in the IP of ORF1p shown in Fig6A, there is a double band as well, strongly suggesting that the lower band might be a truncated or processed form of ORF1p. As stated above, this double band has been detected in other studies (Walter et al. eLife 2016;5:e11418. DOI: 10.7554/eLife.11418) in mouse ES cells using an in-house generated antibody against mouse ORF1p. Thus, with either commercial or in-house generated antibodies in some mouse and human samples, there is a double band corresponding to full-length ORF1p and a truncated or processed version of it.

We noticed that we have not added the references of the primary antibodies used in Western blot experiments in the manuscript, which will be corrected in the revised version.

> *(3) The data showing a reduction in ORF1p expression in the aged mouse brain is confusing and maybe even misleading. Although there is an increase in the intensity of the ORF1p signal in ORF1p+ cells, the data clearly shows that fewer cells express ORF1p in the aged brain. If these changes indicate an overall loss or gain of ORF1p, expression in the aged brain is not resolved. Thus, conclusions should be more carefully phrased in this section. It is important to show the quantification of NeuN+ and NeuN- cells in young vs aged (not only the proportions as shown in Figure 3b) to determine if the difference in the number of ORF1p+ cells is due to loss of neurons or perhaps a sampling issue. More so, it would be essential to perform WB and/or proteomics experiments to complement the IHC data for the aged mouse samples.*

The data presented in Fig3 C-I show a modest but widespread and reproducible increase in expression of ORF1p per cell. What decreases is the proportion of ORF1p+/NeuN+ cells (Fig3A, B), indicating that fewer cells might express ORF1p in the brain. However, the proportion or number/mm2 of ORF1p+ cells overall does not decrease significantly, neither does the proportion or number/mm2 of NeuN+ cells (data will be added to the revision). We show data of the % of NeuN+ and NeuN- cells in the ventral midbrain (Suppl Fig3C, quantified on confocal images)) which indeed indicates that in this region, there are less neurons in the aged mouse brain compared to the young. There might thus be a very regional decrease in neurons with age in the midbrain motor region. We will, however, as suggested, plot the number of NeuN+ and NeuN- cells per mm2 for the whole brain as well as the different regions in young vs aged to compare actual cell numbers per volume. While it is true that we cannot say that there is an overall loss or gain of ORF1p expression in the aged mouse brain, we believe that this is not of the highest importance as what most likely matters biologically in the context of aging is the quantity of ORF1p per cell (and possibly full-length LINE-1 RNA and ORF2p) and not “per brain”.

We also plan on doing Western blots on mouse brain tissues from young and aged individuals, however, we might run into limits regarding tissue availability of aged mice.

> *(4) The transcriptomic data presented in Figure 4 and Figure 5 are not convincing. Quantification of transposon expression on short read sequencing has important limitations. Longer reads and complementary approaches are needed to study the expression of evolutionarily young L1s (see PMID: 38773348 & PMID: 37910626 for examples of the current state of the art). Given the read length and the unstranded sequencing approach, I would at least ask the authors to add genome browser tracks of the upregulated loci so that we can properly assess the clarity of the results. I would also suggest adding the mappability profile of the elements in question. In addition, since this manuscript focuses on ORF1p, it would be essential to document changes in protein levels (and not just transcripts) in the ageing human brain.*

We agree that there are limitations to the analysis of TEs with short read sequencing and we will add more text on this aspect in a revised version. The approaches shown in PMID: 38773348 & PMID: 37910626 or even a combination of them, would be ideal of course. However, here we reanalyzed a unique existing dataset (Dong et al, Nature Neuroscience, 2018; http://dx.doi.org/10.1038/s41593-018-0223-0), which contains RNA-seq data of human post-mortem dopaminergic neurons in a relatively high number of brain-healthy individuals of a wide age range including some "young" individuals which is rare in post-mortem studies. Such data is unfortunately not available with long read sequencing or any other more appropriate approach yet. Limitations are evident, but all limitations will apply equally to both groups of individuals that we compare. We will add genome browser tracks of the differentially expressed elements. The general mappability profile of the full-length LINE-1 "UIDs" is shown in Suppl Fig 6A. We will color-highlight the specific elements in this graph and will add genome browser data for these elements in a revised version.

We will not be able to document changes in protein levels in aged human dopaminergic neurons as we do not have access to this material. We have tried to obtain human substantia nigra tissues but were not able to get sufficient amounts to do laser-capture microdissection or FACS analyses, especially of young individuals. There are still important limitations to tissue availability, especially of regions of interest like the substantia nigra pars compacta affected in Parkinson's disease.

> *(5) More information is needed on RNAseq of microdissections of dopaminergic neurons from 'healthy' postmortem samples of different ages. No further information on these samples is provided. I would suggest adding a table with the clinical information of these samples (especially age, sex, and cause of death). The authors should also discuss whether this experiment has sufficient power. The human ageing cohort seems very small to me.*

This is a re-analysis of a published dataset (Dong et al, Nat Neurosci, 2018; doi:10.1038/s41593-018-0223-0), available through dbgap (phs001556.v1.p1). In this original article, the criteria for inclusion as a brain-healthy control were as follows:

"…Subjects… were without clinicopathological diagnosis of a neurodegenerative disease meeting the following stringent inclusion and exclusion criteria. Inclusion criteria: (i) absence of clinical or neuropathological diagnosis of a neurodegenerative disease, for example, PD according to the UKPDBB criteria47, Alzheimer's disease according to NIA-Reagan criteria48, or dementia with Lewy bodies by revised consensus criteria49; for the purpose of this analysis incidental Lewy body cases (not meeting clinicopathological diagnostic criteria for PD or other neurodegenerative disease) were accepted for inclusion; (ii) PMI ≤ 48 h; (iii) RIN50 ≥ 6.0 by Agilent Bioanalyzer (good RNA integrity); and (iv) visible

ribosomal peaks on the electropherogram. Exclusion criteria were: (i) a primary intracerebral event as the cause of death; (2) brain tumor (except incidental meningiomas); (3) systemic disorders likely to cause chronic brain damage."

We do not have access to the cause of death, but we will add available metadata to the manuscript.

We will perform a post-hoc power analysis and add the result to the revision.

> *(6) The findings in this manuscript apply to both human and mouse brains. However, the landscape of the evolutionarily young L1 subfamilies between these two species is very different and should be part of the discussion. For example, the regulatory sequences that drive L1 expression are quite different in human and mouse L1s. This should be discussed.*

Indeed, they are very different. We will add this to the discussion.

> *(7) On page 3 the authors write: "generally accepted that TE activation can be both, a cause and consequence of aging". This statement does not reflect the current state of the field. On the contrary, this is still an area of extensive investigation and many of the findings supporting this hypothesis need to be confirmed in independent studies. This statement should be revised to reflect this reality.*

We agree, this is overstated, we will change this sentence accordingly.

> ***Reviewer #2 (Public Review):***
>
> *Summary:*
>
> *Bonnifet et al. sought to characterize the expression pattern of L1 ORF1p expression across the entire mouse brain, in young and aged animals, and to corroborate their characterization with Western blotting for L1 ORF1p and L1 RNA expression data from human samples. They also queried L1 ORF1p interacting partners in the mouse brain by IP-MS.*
>
> *Strengths:*
>
> *A major strength of the study is the use of two approaches: a deep-learning detection method to distinguish neuronal vs. non-neuronal cells and ORF1p+ cells vs. ORF1p- cells across large-scale images encompassing multiple brain regions mapped by comparison to the Allen Brain Atlas, and confocal imaging to give higher resolution on specific brain regions. These results are also corroborated by Western blotting on six mouse brain regions. Extension of their analysis to post-mortem human samples, to the extent possible, is another strength of the paper. The identification of novel ORF1p interactors in the brain is also a strength in that it provides a novel dataset for future studies.*

Thank you for highlighting the strength of our study.

> *Weaknesses:*
>
> *The main weakness of the study is that cell type specificity of ORF1p expression was not examined beyond neuron (NeuN+) vs non-neuron (NeuN-). Indeed, a recent study (Bodea et al. 2024, Nature Neuroscience) found that ORF1p expression is characteristic of parvalbumin-positive interneurons, and it would be very interesting to query whether other neuronal subtypes in different brain regions are distinguished by ORF1p expression.*

We agree that this point is important to address. We do provide indications for this in the manuscript. For instance, we detect staining in mouse and human Purkinje cells of the cerebellum in accordance with data from Takahashi et al, Neuron, 2022; DOI: 10.1016/j.neuron.2022.08.011. We also know from previous work, that in the mouse ventral midbrain, dopaminergic neurons (TH+/NeuN+) express ORF1p and that these neurons express higher levels of ORF1p than adjacent non-dopaminergic neurons (TH-/NeuN+; Blaudin de Thé et al, EMBO J, 2018). Others have shown evidence of full-length L1 RNA expression in both excitatory and inhibitory neurons but much less expression in non-neuronal cells (Garza et al, SciAdv, 2023). In sum, although this has not been investigated systematically brain-wide, it does not seem as if ORF1p expression is restricted to PV cells overall. We will deepen the discussion of this aspect in the revised manuscript. To address this question experimentally, we will try to perform ORF1p stainings on different brain regions together with PV stainings and add this data to a revised version, if possible.

> *The data suggesting that ORF1p expression is increased in aged mouse brains is intriguing, although it seems to be based upon modestly (up to 27%, dependent on brain region) higher intensity of ORF1p staining rather than a higher proportion of ORF1+ neurons. Indeed, the proportion of NeuN+/Orf1p+ cells actually decreased in aged animals. It is difficult to interpret the significance and validity of the increase in intensity, as Hoechst staining of DNA, rather than immunostaining for a protein known to be stably expressed in young and aged neurons, was used as a control for staining intensity.*

It would have been indeed interesting to have another marker than DNA as a control. However, this requires a protein that is indeed stably expressed throughout the brain and throughout age. We are not aware of a protein for which this has been established. DNA staining with Hoechst does control for technical artefacts. We have whole-brain imaging data for the protein Rbfox3 (NeuN) which we used as a marker for cell identity. If this protein turns out to be stable, we could add this data to a revised version.

> *The main weakness of the IP-MS portion of the study is that none of the interactors were individually validated or subjected to follow-up analyses. The list of interactors was compared to previously published datasets, but not to ORF1p interactors in any other mouse tissue.*

As stated in the manuscript, the list of previously published datasets does include a mouse dataset with ORF1p interacting proteins in mouse spermatocytes (please see line 434-435: “ORF1p interactors found in mouse spermatocytes were also present in our analysis including CNOT10, CNOT11, PRKRA and FXR2 among others (Suppl_Table4).”) -> De Luca, C., Gupta, A. & Bortvin, A. Retrotransposon LINE-1 bodies in the cytoplasm of piRNA-deficient mouse spermatocytes: Ribonucleoproteins overcoming the integrated stress response. *PLoS Genet* 19, e1010797 (2023)). We indeed did not validate any interactors for several reasons (economic reasons and time constraints (post-doc leaving)). However, we feel that the significant overlap with previously published interactors highlights the validity of our data and we anticipate that this list of ORF1p protein interactors in the mouse brain will be of further use for the community.

> *The authors achieved the goals of broadly characterizing ORF1p expression across different regions of the mouse brain, and identifying putative ORF1p interactors in the mouse brain. However, findings from both parts of the study are somewhat superficial in depth.*
>
> *This provides a useful dataset to the field, which likely will be used to justify and support numerous future studies into L1 activity in the aging mammalian brain and in*

neurodegenerative disease. Similarly, the list of ORF1p interacting proteins in the brain will likely be taken up and studied in greater depth.

Reviewer #3 (Public Review):

The question about whether L1 exhibits normal/homeostatic expression in the brain (and in general) is interesting and important. L1 is thought to be repressed in most somatic cells (with the exception of some stem/progenitor compartments). However, to our knowledge, this has not been authoritatively / systematically examined and the literature is still developing with respect to this topic. The full gamut of biological and pathobiological roles of L1 remains to be shown and elucidated and this area has garnered rapidly increasing interest, year-by-year. With respect to the brain, L1 (and repeat sequences in general) have been linked with neurodegeneration, and this is thought to be an aging-related consequence or contributor (or both) of inflammation. This study provides an impressive and apparently comprehensive imaging analysis of differential L1 ORF1p expression in mouse brain (with some supporting analysis of the human brain), compatible with a narrative of non-pathological expression of retrotransposition-competent L1 sequences. We believe this will encourage and support further research into the functional roles of L1 in normal brain function and how this may give way to pathological consequences in concert with aging. However, we have concerns with conclusions drawn, in some cases regardless of the lack of statistical support from the data. We note a lack of clarity about how the 3rd party pre-trained machine learning models perform on the authors' imaging data (validation/monitoring tests are not reported), as well as issues (among others) with the particular implementation of co-immunoprecipitation (ORF1p is not among the highly enriched proteins and apparently does not reach statistical significance for the comparison) - neither of which may be sufficiently rigorous.

Thank you for your comments on our manuscript.

In a revised version and a more in-depth response, we will address the concerns about the machine learning paradigm. Concerning the co-IP-MS, we can confirm that ORF1p is among the highly enriched proteins as it was not found in the IgG control (in 5 independent samples), only in the ORF1p-IP (in 5 out of 5 independent samples). This is what the infinite sign in Suppl Table 2 indicates and this is why there is no p-value assigned as infinite/0 doesn't allow to calculate a p-value. We will make this clearer in a revised version of the manuscript.

https://doi.org/10.7554/eLife.100687.1.sa0