Title:

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

Authors:

Tom Bonnifet[1], Sandra Sinnassamy[1], Olivia Massiani Beaudoin[1], Philippe Mailly[2], Héloïse Monnet[2], Damarys Loew[3], Berangère Lombard[3], Nicolas Servant[4], Rajiv Joshi[1]* & Julia Fuchs[1]*

Affiliations:

[1] CIRB, Collège de France, PSL University, INSERM, CNRS, 75005, Paris, France.

[2] Orion Imaging Facility, CIRB, Collège de France, PSL University, INSERM, CNRS, Labex Memolife, 75005, Paris, France.

[3] Institut Curie, PSL University, Centre de Recherche, CurieCoreTech Spectrométrie de Masse Protéomique, 26 rue d'Ulm, Paris 75248 Cedex 05, France.

[4] Institut Curie, INSERM U900, Mines Paris Tech, PSL University.

***For correspondance:**

J.F. (Julia.fuchs@college-de-france.fr); R.L.J. (rajiv.joshi@college-de-france.fr)

Key words: Transposable elements, retrotransposons, LINE-1, ORF1p, aging, brain mapping, mass spectrometry, ORF1p protein interactors, human post-mortem brain, dopaminergic neurons, deep learning.

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.

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% of the human genome as remnants of an evolutionary ancient activity. Depending on the source, about 100 [2] to 146 [3] 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 [4] . The LINE-1 encoded protein ORF1p, an RNA binding protein with "cis" preference [5,6], and ORF2p, an endonuclease and reverse transcriptase [7,8] are required for the mobility of LINE-1 elements. LINE-1 elements, as many other transposable elements including HERVs, are repressed by multiple cellular pathways. Due to this, it was generally thought that TEs are repressed in somatic cells with no expression at steady-state [9–11]. However, aging might reduce the reliability of these repressive mechanisms [12]. It is now, 31 years after the initial proposition of the "transposon theory of aging" by Driver and McKechnie [13], generally accepted that TE activation can be both, a cause and a consequence of aging [14,15]. Sparse data has shown that the LINE-1 encoded protein ORF1p is expressed at steady-state in the mouse ventral midbrain [16] and in some regions of the human post-mortem brain [17] and recent data informed about the presence of full-length transcripts in various tissues [18]. Repression of LINE-1 might

74 thus be incomplete and if so, it remains unclear how cells then prevent cell toxicity associated
75 with LINE-1 encoded protein activity. Indeed, LINE-1 encoded proteins have been
76 demonstrated to induce genomic instability (ORF2p endonuclease-mediated [16,19–24];
77 inflammation (ORF2p reverse transcriptase-mediated [25–27] and/or nuclear envelope fragility
78 (ORF1p-mediated [28]) and these cellular activities might be causally related to organismal aging
79 and/or cancer , autoimmune and neurological diseases [29]. For instance, LINE-1 activation can
80 drive neurodegeneration of mouse dopaminergic neurons [16], of drosophila neurons [30,31] and of
81 mouse Purkinje neurons [32] which can be at least partially rescued with nucleoside analogue
82 reverse transcriptase inhibitors (NRTIs) or other anti-LINE-1 strategies. NRTIs are currently
83 being tested in several clinical trials designed to target either the RT of HERVs or the RT
84 encoded by the LINE-1 ORF2 protein. It is not known today, however, to which extend LINE-
85 1 encoded proteins are expressed at steady-state throughout the mouse and human brain,
86 whether there is cell-type specificity and whether activation of LINE-1 encoded proteins is
87 associated with brain aging or human neurodegeneration. Here, using a deep-learning assisted
88 cellular detection methodology applied to pyramidal large-scale images of the mouse brain
89 mapped to the Allen mouse brain atlas combined with post-mortem human brain imaging, co-
90 IP mass spectrometry and transcriptomic analysis of LINE-1 expression, we describe a brain-
91 wide map of ORF1p expression and interacting proteins at steady-state and in the context of
92 aging. We find a heterogenous but widespread expression of ORF1p in the mouse brain with
93 predominant expression in neurons. In aged mice, neuronal ORF1p expression increases brain-
94 wide and in some brain regions to up to 27%. In human dopaminergic neurons, young LINE-1
95 transcripts and specific full-length and coding LINE-1 copies are increased in aged individuals.
96 We further describe endogenous mouse ORF1p interacting proteins revealing known
97 interactors and unexpected interacting proteins belonging to GO categories related to RNA
98 metabolism, chromatin remodeling, cytoskeleton and the synapse.

99 **Results**

100

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

103 To investigate the expression pattern and intensities of endogenous LINE-1 encoded ORF1p
104 protein throughout the entire mouse brain, we devised a deep-learning assisted cellular
105 detection methodology applied to pyramidal large-scale images using a comprehensive
106 workflow complemented by an approach based on confocal imaging as schematized in Figure
107 1A. Briefly, starting from sagittal slide scanner images of the mouse brain, we defined
108 anatomical brain regions by mapping the Allen Brain Atlas onto the slide scanner images using
109 Aligning Big Brains & Atlases (ABBA). We then employed a deep-learning detection method
110 to identify all cell nuclei (Hoechst) and categorize all detected cells into neuronal cells (NeuN+)
111 or non-neuronal cells (NeuN-) and ORF1p-expressing cells (ORF1p+) and cells that do not
112 express ORF1p (ORF1p-). This workflow allowed us then to characterize cell identity and
113 ORF1p intensity throughout the whole brain but also in specific anatomical regions. In parallel,
114 we completed the approach using confocal microscopy on selected anatomical regions allowing
115 for comparison with higher resolution. Importantly, ORF1p antibody specificity was validated
116 by blocking the ORF1p antibody with purified mouse ORF1p resulting in the complete absence
117 of immunofluorescence staining (SupplFigure1A) as well as by immunoprecipitation and mass
118 spectrometry (Fig 6). Unexpectedly, we found a generalized and widespread expression of
119 ORF1p throughout the brain of wildtype mice (Swiss OF1 mice, three months-old; Figure 1B).
120 ORF1p is detectable in all regions and subregions analyzed. The expression pattern (density
121 and intensity) per region/subregions was however heterogeneous. The ten regions shown in
122 Figure 1B exemplify visible different densities of ORF1p+ cells with varying levels of
123 expression. Throughout the entire brain, we found a relatively high density of ORF1p+ cells

124 with a mean density of ≈ 305 ±18 ORF1p+ cells per mm^2 (mean ±SEM), representing up to
125 20% of all detected cells (Figure 1C). Among ORF1p+ cells within each analyzed mouse brain,
126 ORF1p intensities between low and high expressing cells were up to eight-fold different (Figure
127 1C). We then quantified nine anatomical regions according to the Allen Brain Atlas on four
128 brains of three-month old mice (Fig1D) with our unbiased workflow (Fig1A) regarding cell
129 density (Figure 1D), cell proportions (Figure 1E) and fluorescent intensity of ORF1p+ cells
130 (Figure 1F). This approach permitted the analysis of about 10 000 ORF1p+ cells per animal
131 highlighting the power of our large-scale analysis. Densities of ORF1p+ cells ranged from the
132 lowest density in the hindbrain with 154 ±19 cells per mm^2 (mean ±SEM) to the highest density
133 of ORF1p+ cells in the isocortex with 451 ±44 cells per mm^2 (mean ±SEM) and the thalamus
134 with 446 ±50 cells per mm^2 (mean ±SEM). This correlated with the percentage of ORF1p+ cells
135 among all cells identified per region (Figure 1E, r^2=0.96). The percentage of ORF1p+ cells per
136 anatomical region fluctuated between 10% ±2.1 (ventral striatum, mean ±SEM) and 31% ±1.6
137 (thalamus, mean ±SEM). The dorsal striatum ("striatal dorsal" in the Allen Brain Atlas
138 denomination) exhibited the lowest ORF1p expression intensity (658 ±3 mean ±SEM) of all
139 regions tested, the hindbrain the highest mean intensity of ORF1p per cell (mean 1221 ±8
140 (SEM)) as illustrated in Figure 1B and quantified in Figure 1D and 1F. Interestingly, cell density
141 did not correlate with expression levels. Dorsal and ventral striatum for instance display similar
142 ORF1p intensities per cell but exhibit significant differences in ORF1p cell density and
143 proportion. The "midbrain motor" region as defined by the Allen Brain Atlas showed an
144 intermediate cell density (mean 265 ±16 (SEM)) and a rather high ORF1p expression intensity
145 (mean 1006 ±7.3 (SEM)). Statistical analysis comparing mean density of ORF1p+cells per mm^2
146 or mean intensity per ORF1p+ cells among regions confirmed the heterogeneity concerning
147 ORF1p expression throughout the mouse brain (Fig1D, 1F). Slide scanner and confocal images
148 revealed an exceptionally high ORF1p expression intensity in the ventral region of the midbrain

149 which we identified as the *Substantia nigra pars compacta (SNpc),* an anatomical region
150 containing dopaminergic neurons which are vulnerable in Parkinson disease (PD). This region
151 displayed an important density of ORF1p+ cells and a comparatively high level of ORF1p
152 expression as illustrated by confocal imaging (Figure 1B, panel 8), but could not be quantified
153 independently with our brain-wide approach due to the geometrically-complex anatomy of this
154 region and its small size (subregion-level in the Allen Brain Atlas hierarchy). Another region
155 which could not be included in our brain-wide analysis was the cerebellum due to its extremely
156 high density of cell nuclei. However, slide scanner and confocal imaging (Figure 1B, panel 10)
157 revealed that ORF1p is highly expressed in Purkinje cells, while not detectable in the molecular
158 or granular layers.
159 In order to confirm ORF1p expression by an independent method, we performed Western blot
160 analysis on six micro-dissected regions from the mouse brain (Swiss/OF1 mouse, three-month
161 old). As shown in SupplFig1C, ORF1p is expressed in all six regions with varying expression
162 levels confirming the overall presence of ORF1p throughout the brain. We then chose two
163 regions with significantly divergent ORF1p expression intensities as identified and quantified
164 on pyramidal large-scale images: the frontal cortex (low) and the ventral midbrain (intermediate
165 to high). We confirmed a significant higher expression of ORF1p in the ventral midbrain
166 compared to the frontal cortex using an approach based on the unbiased, automated
167 quantification of multiple z-stacks using a confocal microscope (Figure 1G) and by Western-
168 blotting on micro-dissected regions (Figure 1H, 1I). In concordance with the findings stemming
169 from the large-scale image quantification pipeline (Figure 1F), the ventral midbrain showed ≈
170 2-times higher expression of ORF1p than the frontal cortex as quantified in Figure 1G (1.8-
171 fold) and Figure I (2.3-fold) validating our cellular detection methodology for pyramidal large-
172 scale imaging and underscoring the heterogeneity of ORF1p expression levels in the mouse
173 brain. To investigate intra-individual expression patterns of ORF1p in the post-mortem human

174 brain, we analyzed three brain regions of a neurologically-healthy individual (Figure 1J) by
175 Western blotting. We detected two bands, the upper one corresponding to ORF1p (≈ 40kD) and
176 a lower band (≈ 36kD) of unknown nature, both of which are also detectable in HeLa or HEK
177 cells (data not shown). ORF1p was expressed at different levels in the cingulate gyrus, the
178 frontal cortex and the cerebellum underscoring a heterogenous widespread expression of human
179 ORF1p across the human brain. In summary, our findings reveal the consistent presence of
180 ORF1p expression throughout the mouse brain in all anatomical regions analyzed with high
181 regional variability in terms of density of ORF1p+ cells per mm2 per region and ORF1p+
182 intensity per cell per region. This indicates a consistent presence of a LINE-1 encoded protein
183 in brain cells and raises several questions concerning cell-type identity of ORF1p expressing
184 cells and potential functions or consequences of ORF1p expression in the mouse and human
185 brain.
186
187 **ORF1p is predominantly expressed in neurons**
188 Following up on our observation of wide-spread expression patterns of endogenous ORF1p
189 throughout the brain, we first addressed the question of the cellular identity of ORF1p+ cells.
190 To this end, we used the neuron-specific marker NeuN, which has been used for many decades
191 to identify post-mitotic neurons in the central nervous system [33] allowing us to determine the
192 proportion of neuronal (NeuN+) or non-neuronal cells (NeuN-) expressing ORF1p (ORF1p+)
193 or not (ORF1p-). Making use of our large-scale imaging approach (Figure 1A), we observe
194 drastic dissimilarities in detected cellular proportions between the white and grey brain matter.
195 As expected, we observed only 1% of NeuN+ cells in the white matter (corpus callosum; Figure
196 2A) validating both, the neuronal marker NeuN as such and the ABBA superposition of the
197 Allen Brain Atlas onto the sagittal brain slices. Interestingly, among the non-neuronal cells
198 predominant in this region, only 5 cells (mean value of 4 animals) cells were detected as

199 ORF1p+ indicating a near absence of ORF1p+ cells in non-neuronal cells of the white matter.
200 In the grey matter, our approach detected 30.5% NeuN+ cells (dark red and yellow bars in
201 Figure 2A) which, according to the literature, should include all post-mitotic neurons with only
202 minor exceptions [33–36] and corresponds to the reported proportion of neurons present in the
203 mouse brain [37]. The nine identified grey matter regions in Figure 2A display the proportions of
204 the different cell types per region. The proportion of all cells in a given region which are positive
205 for ORF1p (dark red bars) differs between regions (lowest proportion: hindbrain: 7%; highest
206 proportion dorsal striatum: 26.6%). In the isocortex and the midbrain motor-related regions, the
207 majority of neurons detected express ORF1p (54% and 59% by large-scale analysis, Figure 2A;
208 68.8% and 68.7% by confocal imaging (Figure 2B, quantified in C), respectively), while in the
209 midbrain sensory related regions the proportion dropped to 25% whereas it reached 82% in the
210 thalamus (Suppl Figure 2A). Altogether, nearly half of all NeuN+ cells throughout the mouse
211 brain expressed ORF1p (mean of all regions: 48.2%; Figure 2E). Regarding the cell identity of
212 ORF1p+ cells brain-wide (except the cerebellum), more than 70% were identified as neuronal
213 by the large-scale approach (Suppl Figure 2A). The regional analysis indicated that 73%
214 (isocortex) and 64.4% (midbrain motor) of ORF1p+ cells were also NeuN+ (Suppl Figure 2A).
215 This contrasted with results obtained by the second approach using confocal imaging on
216 multiple z-stacks which indicated that 91.3% (frontal cortex) and 88.5% (ventral midbrain) of
217 ORF1p+ cells were neuronal (Figure 2D). This difference in numbers observed between our
218 large-scale image cell detection methodology and the confocal workflow are most probably due
219 to technical limitations inherent to our large-scale pipeline. Indeed, with the latter approach,
220 region-dependent differences in cell density and signal intensity levels might be the cause for
221 an underestimation of the proportion of ORF1p+ cells being neuronal due to difficulties in cell
222 detection by StarDist/Cellpose (high cell density) on a single focal plan, technical difficulties
223 which are widely reduced by the multiple z-stack based approach when using a confocal

224 microscope. Notably, we also observed a difference in the % of NeuN+ cells between these two
225 regions with significantly less NeuN+ cells in the ventral midbrain compared to the frontal
226 cortex (SupplFig2B). As we could not rule out that ORF1p might also be expressed in non-
227 neuronal cells, we turned to non-neuronal markers specific for different glial cell populations
228 using two different astrocytic markers (GFAP, Sox9), the astro- and oligodendrocytic marker
229 S100b and the microglial marker Iba1[37,38] and performed co-stainings with ORF1p followed by
230 confocal imaging as illustrated in Figure 2G. We screened multiple images of frontal cortex,
231 ventral midbrain, hippocampus and striatum and did not find a single cell, which could
232 unambiguously be defined as non-neuronal, to express ORF1p. This indicated that ORF1p is
233 not or only very rarely expressed in non-neuronal cells. To assess whether predominant, if not
234 exclusive ORF1p expression in neurons is mouse brain specific or a pattern also applicable to
235 the human brain, we investigated the identity of ORF1p expressing cells in the cingulate gyrus
236 of a healthy human brain. Similar to what we found in the mouse brain, we observed sparse
237 NeuN expression in the white (Figure 2H, left) and extensive NeuN staining in the grey matter
238 corresponding to the cortical layers (right, separated by a dashed white line) with ORF1p+ cells
239 predominantly located in the grey matter. All cells stained by ORF1p were co-stained with
240 NeuN indicating that ORF1p was expressed in neuronal cells also in the human brain
241 (Figure2H, right). However, as in the mouse brain, not all neurons expressed ORF1p (Figure
242 2H, bottom right). Finally, to further confirm the absence of ORF1p expression in non-neuronal
243 cells and the presence of expression of ORF1p in neurons, we used fluorescence-activated cell
244 sorting (FACS) to isolate neurons (using a NeuN antibody) from non-neuronal cells (NeuN-)
245 from the adult mouse brain followed by Western blotting with an antibody against ORF1p
246 (Figure 2I, 2J). We detected ORF1p exclusively in the neuronal fraction, again confirming our
247 prior results based on two different imaging approaches. In summary, ORF1p expression in the
248 mouse brain is widely restricted to neurons of which the majority express ORF1p. This raises

249 the question of the function and consequences of ORF1p expression specifically in neurons but
250 also on the dynamic regulation of this expression upon exogenous (exposome) or endogenous
251 (aging) challenges.
252
253 **ORF1p expression is increased in the aged mouse brain**
254 ORF1p is expressed at steady-state throughout the brain, but whether there is a dynamic
255 regulation of this expression is not known. Aging has unequivocally been linked to LINE-1
256 regulation [15,39] both as a trigger and as a consequence of LINE-1 activation but whether this is
257 true for the brain has not been thoroughly investigated using ORF1p as a read-out. We therefore
258 addressed the question whether advanced age (16-month-old mice) was paralleled by a dynamic
259 regulation of ORF1p expression in the brain compared to young, three-month old mice. In the
260 context of aging, we globally observed a reduction in the proportion of ORF1p+/NeuN+ cells
261 using the cell detection workflow applied to large-scale images described in Figure 1A,
262 phenomenon mainly driven by the midbrain motor, the dorsal striatum, the pallidum and the
263 thalamus regions (Figure 3A, dark red bars, SupplTable1). The confocal approach applied to
264 two regions, the frontal cortex and the ventral midbrain (Figure 3B) confirmed a loss of
265 ORF1p+/NeuN+ cells in the ventral midbrain with no change in cell proportions in the frontal
266 cortex in accordance with the large-scale imaging approach (Figure 3A). The predominantly
267 neuronal identity of ORF1p+ cells, however, was unchanged (SupplFig3A) just as the
268 proportion of neurons expressing ORF1p (SupplFig3B). We observed a significant decrease of
269 NeuN+ cells in the ventral midbrain (Figure 3B, SupplFig3C). We next analyzed ORF1p
270 expression levels in the context of brain aging. Interestingly, the mean intensity of ORF1p
271 expression increased significantly with age throughout the brain (13% increase brain-wide;
272 Figure 3C). Frequency distribution analysis unveiled a shift in ORF1p mean expression per
273 cell in aged mice (Figure 3D). Importantly, the Hoechst mean intensity within nuclei of

274 ORF1p+ cells, serving as an internal control, showed no significant change (Figure 3E). Among
275 nine analyzed regions, five demonstrated a general increase in ORF1p mean intensity per cell
276 in aged mice (p</=0.05), a change independent from inter-individual variations in both young
277 and aged mice (Figure 3F). An increase of ORF1p expression was also observed in three others
278 regions but albeit not significant. The only exception was the isocortex which remained
279 unchanged with aging. The general increase of ORF1p expression (fold change intensity) in the
280 whole brain, reaching nearly a 30% increase in some regions, is represented on the heatmap in
281 Figure 1G. These results were confirmed by the confocal imaging approach: ORF1p expression
282 in the frontal cortex remained unchanged and the ventral midbrain region increased
283 significantly in aged mice as quantified in Figure 3H and shown with a representative image in
284 Figure 3I. Overall, these results highlight an age-dependent increase in ORF1p expression in
285 neurons throughout the brain with some regions showing an increase of up to 30% in ORF1p
286 intensities.
287
288 **Coding LINE-1 transcripts are increased in aged human dopaminergic neurons**
289 Following the observation of increased ORF1p expression in the aged mouse brain, among
290 which the ventral midbrain, and given the age-related susceptibility of dopaminergic neurons
291 in the *SNpc* to degenerate in PD, we turned to a RNA-seq dataset of laser-captured micro-
292 dissected post-mortem human dopaminergic neurons of brain-healthy individuals [40], in order to
293 interrogate full-length LINE-1 expression profiles as a function of age. To avoid read-length
294 bias to which TE analysis is particularly sensitive, we analyzed only the data derived from 50bp
295 paired-end reads of linearly amplified total RNA as this dataset regrouped all age categories
296 (n=41; with ages ranging from 38 to 97 (mean age: 79.88 (SD ±12.07); n=6 </=65y; n=35 >65y;
297 mean PMI: 7.07 (SD ±7.84), mean RIN: 7.09 (±0.94). As age-related dysregulation of TEs
298 might not be linear, we considered individuals with ages-at-death younger or equal to 65 years

299 as “young” (n=6, 38-65 years, mean age 57.5 years (SD ±9.9)) and individuals older than 65
300 years as “aged” (n=35, 65-97 years, mean age 83 years (SD±7.8). The expression of the
301 dopaminergic markers tyrosine hydroxylase (TH) and LMX1B were not different in both
302 populations indicating no apparent change of dopaminergic identity of analyzed melanin-
303 positive dopaminergic neurons (Suppl Figure 4A). Next, we compared the expression of repeat
304 elements at the class, family and name level based on the repeat masker annotation implemented
305 in the UCSC genome browser using a commonly used mapping strategy for repeats [41]. No overt
306 dysregulation of repeat elements at either level of repeat element hierarchy was observed (Suppl
307 Figures 4C,D,E,F), however there was a significant increase in several younger LINE-1
308 elements including L1HS and L1PA2 at the “name” level (Figure 4A,B). This was not observed
309 for HERVK-int, a human endogenous retrovirus family with some copies having retained
310 coding potential (Figure 4B) or other potentially active TEs like HERVH-int, HERV-Fc1 or
311 SVA-F with the exception of a trend for an increase in AluYa5 transcripts in the >65y group, a
312 young Alu family mobilized by the LINE-1 retrotransposition machinery (Suppl Figure 4G).
313 Interestingly, L1HS expression was highly correlated with L1PA2 expression and this
314 correlation extended to almost all younger LINE-1 subfamilies weaning down with
315 evolutionary distance (Figure 5A). This was not true for other active TEs as L1HS was
316 negatively correlated with HERVK-int expression (Figure 5A). Several regulators of LINE-1
317 activity have been identified [16,42] and correlation of their expression with L1HS might allow to
318 infer their relevance of interaction (activation or repression) with L1HS in human dopaminergic
319 neurons. Besides the co-regulation of young LINE-1 elements (L1HS, L1PA2-11, Figure 5A),
320 Spearman correlation analysis revealed three known repressors of LINE-1 activity; EN1
321 (Engrailed 1 [16], Suppl Figure 5A) with important functions for dopaminergic neuron
322 homeostasis [43], CBX5/HP1a, a heterochromatin binding protein binding to the histone mark
323 H3K9me3, thereby mediating epigenetic repression [44] (Suppl Figure 5B) and XRCC5/6, also

324 known as Ku86/Ku70, which are essential for DNA double-stranded break repair through the
325 nonhomologous end joining (NHEJ) pathway and limit LINE-1 full-length insertions [45] (Suppl
326 Figure 5C). The transcripts of these genes showed, although not statistically significant, a trend
327 for decreased expression in the elderly (Suppl Figures 5D-G). Based on the increase of young
328 LINE-1 families L1HS and L1PA2 in aged human dopaminergic neurons and the finding that
329 ORF1p was increased in the aged mouse brain, we focused our attention on LINE-1 elements
330 with coding potential for ORF1 and ORF2 according to the L1Basev2 annotation which are
331 specific elements comprised in the L1HS and L1PA2 annotation at the "name" level. Most of
332 the 146 full-length and coding LINE-1 termed UIDs (= Unique Identifier) in the L1Base are
333 L1HS elements (76.03%), whereas the remaining 35 UIDs belong to the evolutionary older
334 L1PA2 family (Suppl Figure 6B). The L1Base annotation is based on the human reference
335 genome (GRCh38) and annotates 146 human full-length (>6kB), intact LINE-1 elements
336 (ORF1 and ORF2 intact) with a unique identifier from 1 to 146 [3]. Attribution of sequencing
337 reads to a specific, individual TE copy is problematic [46] and several approaches have been
338 proposed to circumvent this problem including the mapping of unique reads [41].While several
339 tools using expectation maximization algorithms in assigning multi-mapping reads have been
340 developed and successfully tested in simulations [41,47], we used a different approach in mapping
341 unique reads to the L1Base annotation of full-length LINE-1. Specific "hot" LINE-1 loci in a
342 given cellular context have been identified [2], but usage of the L1Base annotation enabled an
343 unbiased approach albeit ignoring polymorphic LINE-1 sequences. Unique read mapping
344 strategies for repeat elements, especially young LINE-1 elements, will unavoidably
345 underestimate LINE-1 locus-specific expression levels [41], but will be most accurate in assigning
346 reads while allowing the comparison of two different conditions analyzed in parallel. Assuming
347 that expression of UIDs was correlated with mappability, we plotted a mappability count of
348 each UID (see methods) against its mean normalized read count expression of the six

349 individuals </=65y. Non-parametric Spearman correlation revealed no correlation between UID
350 mappability and expression (SupplFig6A) indicating no apparent bias between the two
351 parameters. However, individual UID dependency of mappability on expression cannot be
352 excluded, especially for high expressing UIDs like UID-16 for example (Suppl Figures 6A,E).
353 Expression of LINE-1 at the locus-level has been attributed to artefacts not representing
354 autonomous transcription including differential high intronic read counts [48] pervasive
355 transcription or reads attributable to passive co-transcription with genes when the LINE-1
356 element is intronic [49]. To evaluate the latter, we determined the number of intronic (46.58%)
357 and intergenic UIDs (78/146; Suppl Figure 6D) and identified the corresponding genes for
358 intronic UIDs (Suppl Figure 6C). Of the 146 UIDs, 140 passed the threshold of >3 reads in at
359 least 6 individuals. Differential expression of UID between "young" and "aged" dopaminergic
360 neurons revealed several significantly deregulated full-length LINE-1 loci. However, while no
361 single locus stood out, paired analysis of the expression of all UIDs indicated a general increase
362 (Figure 4D), especially of low expressed UIDs. The comparative analysis of the sum expression
363 of UIDs per individual comparing young (</=65y) with elderly human dopaminergic neurons,
364 however, did not reach statistical significance (Figure 4E). Several specific loci were
365 dysregulated in particular, for instance UID-68 (Figure 4C), a L1HS element located on
366 chromosome 7 (chr7: 141920659-141926712) in between two genes, CLEC5A (C-type lectin
367 domain containing 5A) and OR9A4 (olfactory receptor family 9 subfamily A member 4). To
368 rule out any influence of "hosting" gene transcription interference on measurable UID-68
369 expression differences (Fig4F, left), we performed Spearman correlation which did not indicate
370 any correlation between CLEC5A or OR9A4 expression with UID-68 (Fig4F). Further, UID-
371 68 had a high mappability count of 16 (range of all UIDs: 1-30, mean 9.0 (SD ±6.05),
372 SupplFig6A) indicating that UID-68 might be a candidate for an age-dependent gain of activity.
373 L1HS UID-129, located on chromosome 15 (chr15: 54926081-54932099) is intergenic, at ≈

374 2Mb distance upstream and ≈ 200Mb downstream from the next genes. However, the
375 mappability count of UID-129 was only 1, indicating possible mapping biases inherent to the
376 unique read-based mapping strategy employed (Figure 4G and SupplFig6A). Another UID
377 increased in individuals >65y, the L1HS UID-37 (chr10:98,782,942-98,788,971, minus strand),
378 located in intron 3 of the HPSE2 gene (mappability count of 13, minus strand), showed no
379 correlation with its "hosting" gene indicating potential autonomous transcription of this LINE-
380 1 element and suggesting its contribution to the increase of full-length LINE-1 transcripts
381 (SupplFig7A). We also inspected UID-127 (chr13:40,356,291-40,362,321, mappability count:
382 14), a L1PA2 element which slightly decreased in elder individuals (Fig4C). We found a
383 positive association with LINC00598 (intron 6 of 6; </=65y: r=0.2, p=0.71; >65y: r=0.37,
384 p=0.03), a non-coding RNA which hosts UID-127 in its 6th intron, indicating transcriptional co-
385 regulation potentially indicative of non-autonomous transcription of UID-127 (SupplFig7B).
386 The decrease in the expression of UID-137 (Fig4C) was mostly driven by one young individual
387 with high expression (SupplFig7C) and thus not reflecting an overall decrease (SupplFig7C,
388 p=0.23). In conclusion, TE expression analysis of this human dataset covering an age-span of
389 59 years indicates an increase in the expression of young LINE-1 elements including those
390 which have coding potential in elderly dopaminergic neurons. A slight net sum increase of UID
391 transcripts/cell might be sufficient for the production of "above steady-state" levels of ORF1p
392 and ORF2p. Other TEs with coding potential, namely members of the HERV family, were not
393 increased. Further, correlation analyses suggest that L1HS expression might possibly be
394 controlled by the homeoprotein Engrailed 1, a protein specifically expressed in dopaminergic
395 neurons in the ventral midbrain [43], the heterochromatin binding protein HP1, two known
396 regulators of LINE-1, and the DNA repair proteins XRCC5/6.
397
398 **Endogenous ORF1p interactors in the mouse brain**

399 To go further in our understanding of steady-state neuronal ORF1p expression across the mouse
400 brain, we immunoprecipitated ORF1p and performed label-free and quantitative LC/MS-MS to
401 identify potential protein partners of ORF1p in the mouse brain. We successfully
402 immunoprecipitated endogenous ORF1p from whole brain lysates (Figure 6A) and identified a
403 total of 424 potential protein interactors associated with ORF1p in 5 independent experiments
404 (n=5 mice, TableS1). Using Gene Ontology (GO) analysis, we identified several interacting
405 proteins belonging to GO terms related to known functions of the ORF1p protein in RNA
406 binding, preferentially [50] but not exclusively *in cis* [51], for instance RNA decapping and mRNA
407 catabolic process, or related to the known presence of ORF1p in ribonucleoprotein particles
408 [52,53] (GO: cytoplasmic ribonucleotide granule) or the presence of ORF1p in p-bodies [51] as shown
409 in Figure 6B and listed in TableS2. Other GO terms that emerged, to our knowledge not
410 previously associated with ORF1p, were related to cGMP-mediated signaling (GO: cGMP-
411 mediated signaling and 3'-5'phosphodiesterase activity: i.e. PDE4A, PDE4B, PDE4DIP) and
412 the cytoskeleton (GO: microtubule depolymerization, cytoskeleton organization, microtubule
413 and tubulin binding, cytoskeletal motor activity and protein binding). cGMP signaling is
414 regulated by 3'-5' phosphodiesterases (PDEs) which degrade 3',5'-cyclic guanosine
415 monophosphate (cGMP) and 3',5'-cyclic adenosine monophosphate (cAMP), an activity
416 essential for cell physiology for the integration of extra- and intracellular signals including
417 neuronal excitability, synaptic transmission and neuroplasticity [54,55]. Further, several ORF1p
418 interacting proteins were constituents of the mating-type switching/sucrose nonfermenting
419 complex (SWI/SNF complex), i.e. ARID1A, ARID1B, SMARCA2, SMARCB1, SMARCC2),
420 an ATP-dependent chromatin remodeler complex disrupting nucleosome/DNA contacts to
421 facilitate DNA/chromatin accessibility by shifting, removing or exchanging nucleosomes along
422 DNA [56,57]. Finally, we also observed proteins belonging to the GO term "neuronal cell body",
423 corroborating with the neuron-specific presence of ORF1p in the brain. A comparative analysis

424 with previous mass spectrometry studies [53,58–63] aimed at identifying ORF1p interacting proteins
425 unveiled significantly more common proteins than expected by chance (overrepresentation test;
426 representation factor 2.6, p< 5.4e-08; Figure 6C), including LARP1, STAU2, ATXN2, RALY,
427 TARBP2 or DDX21 (for a full list see TableS3). The presence of a significant number of
428 overlapping ORF1p interactors in different non-neuronal human cells (HEK[53,58,59], HeLa[60],
429 human breast and ovarian tumors[63] and hESCs[61]) and mouse brain cells (our study), suggest
430 conserved key interactors between both species and between cell types, with a subset of these
431 proteins regulating RNA degradation and translation potentially relevant for the LINE-1
432 lifecycle itself. ORF1p interactors found in mouse spermatocytes [62] were also present in our
433 analysis including CNOT10, CNOT11, PRKRA and FXR2 among others (TableS3). To unravel
434 the physical interactions between the identified interactors of endogenous ORF1p within the
435 mouse brain, we used the STRING database (Search Tool for Recurring Instances of
436 Neighboring Genes, https://string-db.org/). This analysis generated a network representation,
437 where physical interactions are represented by edges (Figure 6D). In analogy with the GO term
438 analysis, ORF1p displayed interactions with various clusters, including well-known RNA
439 decapping complexes directed against LINE-1 RNA, which also encompassed DCP2 and
440 DCP1A which had not previously been identified as interacting with ORF1p [64]. Furthermore,
441 ORF1p exhibited interactions with the SWI/SNF complex (highlighted in red) as well as
442 subunits of the RNA polymerase II complex suggesting a direct or indirect association with
443 accessible chromatin, a hitherto unknown interaction of ORF1p with chromatin compartments
444 within the nucleus. Notably, a multitude of novel interactors belonged to the "neuronal cell
445 body" and "neuron projections" clusters, proposing potential neuron-specific partners of
446 ORF1p such as Grm2/5, Bai1, Epha4, Kcnn2, Grik2 and Dmd among others. A last cluster,
447 formed by Ncoa5 (Nuclear Receptor Coactivator 5), Nxf1 (Nuclear RNA Export Factor 1),
448 Ranbp2 and Nup133 (both nucleoporins), might imply a role for these interactions in L1-RNA

449 nuclear export and/or a mechanism for the LINE-1 RNP to gain access to the nucleus in post-
450 mitotic neurons, a process we described recently in human neurons [28]. Altogether, the
451 identification of known and novel interactors of ORF1p in the mouse brain suggests roles of
452 ORF1p in the LINE-1 life cycle (RNA binding and metabolism, RNP formation, nuclear access)
453 but also suggests potential novel physiological roles of ORF1p in the brain related to
454 cytoskeleton organization, cGMP signaling, neuron-specific functions (i.e. synaptic signaling,
455 TableS2) and chromatin organization and/or transcription regulation.
456
457
458 **Discussion**
459 While LINE-1 derepression in aging has been extensively explored in peripheral tissues and
460 various pathologies, including cancer, our understanding of LINE-1, particularly ORF1p
461 expression, in the central nervous system within this context remains limited [18,65,66]. A recent
462 search of ORF1p peptides in mass spectrometry data spanning 29 different healthy tissues did
463 not reveal the presence of ORF1p in the brain, suggesting that its presence might lie below
464 detection limits [18]. Only a few studies explored ORF1p encoded LINE-1 protein expression in
465 the brain, most focusing on a specific region (in mice [16,32] in rats [68] and in human post-mortem
466 brain [17]), but it remained unclear if ORF1p is expressed throughout the entire brain, exhibits
467 cell-type specificity, and most intriguingly, if its expression is influenced by the aging process.
468 Here, we demonstrate that ORF1p is expressed throughout the entire mouse brain and in at least
469 three regions of the human post-mortem brain at steady-state. Leveraging a comprehensive
470 workflow that incorporates brain atlas registration and machine learning algorithms (as
471 described in the methods section), we quantified tens of thousands of brain cells, enabling a
472 profound analysis of cell proportions, cell identities, densities and ORF1p expression levels
473 across the entire brain. Surprisingly, more than one-fifth of detected cells expressed ORF1p.

474 Strikingly, regional variations in ORF1p expression levels were observed, with each region
475 exhibiting distinct proportions, cell density, and signal intensity of ORF1p+ cells. In a non-
476 neurologically diseased human brain, ORF1p is expressed in all three regions examined, that is
477 the cingulate gyrus, the frontal cortex and the cerebellum. This is in accordance to an earlier
478 study using histological stainings, which found ORF1p expression in the human frontal cortex,
479 the hippocampus, in basal ganglia, thalamus, midbrain and the spinal cord [17]. This suggests,
480 similarly to the mouse brain, a generalized expression across the human brain. On the
481 transcriptomic level using long-read sequencing of GTEx tissues, brain and liver were
482 highlighted as the organs displaying the highest expression of putatively active, full-length
483 LINE-1 elements [67]. However, when the authors looked at sub-regions, they found transcript
484 expression in cerebellar hemispheres and the putamen, but not in the caudate and the anterior
485 cingulate gyrus and frontal cortex [67]. This is in contrast to our data and the data from Sur et al,
486 where ORF1p was found to be expressed in the latter two regions using two different
487 antibodies. We used the anti-human LINE-1 ORF1p antibody clone 4H1, a well characterized
488 antibody [66,69]. While the sample size of human post-mortem stainings certainly needs to be
489 increased in order to draw quantitative conclusions, the presence of the protein in two
490 independent studies does point to a steady-state expression of ORF1p in the human brain.
491 In the mouse brain, we find ORF1p to be expressed predominantly if not exclusively in neurons
492 using immunofluorescence and fluorescence-activated cell sorting (FACS). This result is
493 consistent with previous studies, such as the identification of ORF1p in excitatory neurons
494 within the mouse frontal cortex [70], its presence in neurons in the ventral midbrain [16] and the
495 recognition of morphological similarities between stained neurons and ORF1p+ cells in a post-
496 mortem hippocampus sample of a healthy individual [17]. We also detected ORF1p in Purkinje
497 cells in the mouse and in bulk human cerebellum. Neuronal specificity or preference of LINE-
498 1 expression was also shown on the transcriptomic level in a recent study investigating LINE-

499 1 expression in the mouse hippocampus, where neuronal LINE-1 expression exceeded that of
500 astrocytes and microglia by approximately twofold [18] and single-nuclei RNA-seq data from the
501 mouse hippocampus and frontal cortex which confirmed globally that repetitive elements
502 including LINE-1 are more active in neurons than in glial cells [70]. While ORF1p expression is
503 suggested to be expressed in microglia under experimental autoimmune encephalomyelitis
504 conditions in the spinal cord [71], no evidence of such expression was observed in non-neuronal
505 cells under non-pathological condition.

506 On average, throughout the mouse brain, the majority of neurons were positive for
507 ORF1p and in some regions (i.e. the thalamus) around 80% of neurons expressed ORF1p.
508 Comparing the results of both imaging approaches, the percentages of neurons expressing
509 ORF1p in the ventral midbrain and frontal cortex were roughly similar (around 70% of neurons
510 expressed ORF1p as quantified by confocal imaging and about 60% of neurons were identified
511 as ORF1p+ using the slide scanner approach). It is possible that these percentages are
512 underestimated due to technical issues inherent to the machine-learning based algorithm for cell
513 detection as our observations often indicated a positive signal in neurons which were classified
514 as negative due to a particular shape or our stringent intensity threshold. A question which arises
515 based on these findings is whether specific features distinguish ORF1p+ and ORF1p- neurons.
516 While we cannot conclude on this question, our results permit to exclude certain possibilities.
517 There is no specific and clustered ORF1p expression in anatomically defined brain nuclei which
518 are known to share a specific neurotransmitter (i.e. dopaminergic neurons in the *SNpc)* which
519 suggests that ORF1p expression does not seem to be strictly neuronal subtype-dependent.
520 However, there is some degree of cell-type specificity as mouse dopaminergic neurons in the
521 *SNpc* express higher levels of ORF1p compared to surrounding, non-dopaminergic neurons [16]
522 (Figure 1B, panel 8). Further, we observe ORF1p staining in Purkinje cells but not in the
523 surrounding granular and molecular layer neurons of the cerebellum. Purkinje cells are

524 inhibitory neurons just as some neurons in the granular layer (i.e. golgi and unipolar brush cells)
525 which are ORF1p negative, so the decisive feature does not appear to be the excitatory or
526 inhibitory nature of a neuron either. Another possibility is a cell-type specific chromatin
527 organization permissive for the expression of LINE-1 and future single-cell studies in the mouse
528 and human brain might reveal those differences.

529 Because transposable elements are known to become active in somatic tissues during
530 aging [14,15,26,72,73], we aimed to investigate whether there was a corresponding increase at the
531 protein level. In aged mice, ORF1p expression significantly increased throughout the mouse
532 brain consistent with a previously documented increase in ORF1p outside the central nervous
533 system in aged rats [68,74] and aged mice [72]. By quantifying the mean intensity of ORF1p in over
534 70 000 cells identified as ORF1p+, we were able to characterize the extent of this increase in
535 each anatomical sub-region. Remarkably, apart from the isocortex which did not show any
536 change, ORF1p expression increased in all other brain regions by 7% to 27%, indicating a
537 generalized increase of ORF1p expression in neurons throughout the brain (13%). We did not
538 detect any change in cell identity of ORF1p expressing cells, that is, ORF1p expression
539 remained predominantly if not exclusively neuronal. Intriguingly, we observed a loss of NeuN+
540 cells, particularly those expressing ORF1p, throughout the brain which was more pronounced
541 in specific regions (ventral midbrain) than in others (isocortex). The loss of NeuN+ cells could
542 either be due to a loss of neuronal identity, as described recently in the context of neuronal
543 aging [75] and in the context of Alzheimer disease and related tauopathies [76], to a
544 neurodegenerative process *per se* or to gliosis related to the aging process [77] (as we observe a
545 slight increase of ORF1p-/NeuN- cells per mm^2).
546 Interestingly, a region with a strong increase in ORF1p expression with aging (ventral midbrain)
547 also had a significant loss of NeuN+ cells while a region with no change in ORF1p expression
548 with aging (isocortex, frontal cortex) did not lose NeuN+ cells. However, further investigations

549 are necessary to validate a correlation and to investigate an underlying mechanism. An increase
550 of ORF1p might have several direct or indirect consequences on a cell or here, on a neuron. As
551 ORF1p is translated from a polycistronic LINE-1 RNA together with ORF2p, albeit in much
552 higher amounts (the estimated ratio ORF1p to ORF2p is 240:1) [78], it can be expected that a
553 LINE-1 ribonucleotide particles are formed and ORF2-dependent cell toxicity in form of
554 genomic instability [16,19] and single-stranded cytoplasmic DNA triggered inflammation [25,26,72]
555 might result. This has been shown in mouse dopaminergic neurons where oxidative stress
556 induced LINE-1 causally contributed to neurodegeneration [16]. Neurodegeneration was partially
557 prevented by anti-LINE-1 strategies among which NRTIs [16] and similar LINE-1 protein-
558 dependent neuronal toxicity has been shown in drosophila [30,31] and the mouse cerebellum [32].
559 Recently, ORF1p has been evoked as a destabilizer of the nuclear envelope with negative
560 consequences on nucleocytoplasmic transport and chromatin organization, independently on
561 the known endonuclease or reverse transcriptase-dependent toxic properties of ORF2p [28].
562 In order to test whether an increase in LINE-1 is a feature of human brain aging, we
563 turned to a unique RNA-seq dataset of human laser-captured dopaminergic neurons of 41
564 individuals ranging from 38 to 99 years [40]. In accordance with our focus on LINE-1 sequences
565 which are full-length and coding, we developed a rationale to interrogate LINE-1 families with
566 representatives that are coding (L1HS, L1PA2, multimappers; RepeatMasker) and to
567 specifically investigate full-length LINE-1 elements that have intact open reading frames for
568 ORF1p and ORF2p (unique reads; L1Basev2 [3]). Indeed, we find an increase in L1HS and
569 L1PA2 elements in individuals >/=65y as well as an increase in specific full-length LINE-1
570 elements but only a trend for increase of all full-length LINE-1 in sum in the elderly. This
571 analysis has technical limitations inherent to transcriptomic analysis of repeat elements
572 especially as it is based on short-read sequences. Nevertheless, we tried to rule out several biases
573 by demonstrating that mappability did not correlate with expression and that the expression of

574 intronic full-length LINE-1 elements is not correlated with the expression of their "hosting"
575 gene. Interestingly, dysregulated full-length LINE-1 elements in aged dopaminergic neurons
576 did not correspond to those identified in bladder cancer [79] indicating the intricate nature of this
577 expression across tissues and pathological conditions. Overall, a slight net sum increase of UID
578 transcripts/cell might be sufficient for the production of "above steady-state" levels ORF1p and
579 ORF2p. Further, a dissociation of LINE-1 transcript and protein levels in aging has been
580 observed recently in excitatory neurons of the mouse cortex. In the absence of transcriptional
581 changes of LINE-1, protein levels of ORF1p were increased [70].

582 We can only speculate about the reason for an increase in ORF1p in the aged brain. A
583 recent single-cell epigenome analysis of the mouse brain suggested a specific decay of
584 heterochromatin in excitatory neurons of the mouse brain with age which was paralleled by an
585 increase in ORF1p, albeit equally in excitatory and inhibitory neurons, again not indicating any
586 dependency of ORF1p regulation on the excitatory or inhibitory nature of neurons [70]. Chromatin
587 and particularly heterochromatin disorganization are a primary hallmark of aging [73] but other
588 repressive cellular pathways which control the LINE-1 life cycle might also fail with aging [12].
589 Another possibility is a loss of accessibility of repressive factors to the LINE-1 promoter or an
590 age-dependent decrease in their expression. Matrix correlation analysis of several known LINE-
591 1 regulators, both positive and negative, revealed possible regulators of young LINE-1
592 sequences in human dopaminergic neurons. Despite known and most probable cell-type
593 unspecific repressive factors like the heterochromatin binding protein CBX5/HP1 [44] or the DNA
594 repair proteins XRCC5 and XRCC6 [42], we identified the homeoprotein Engrailed as negatively
595 correlated with young LINE-1 elements including L1HS and L1PA2. Engrailed is an essential
596 protein for mouse dopaminergic neuronal survival [43] and binds, in its properties as a
597 transcription factor, to the promoter of LINE-1 [16]. As Engrailed is specifically expressed in
598 dopaminergic neurons in the ventral midbrain, our findings suggests that Engrailed controls

599 LINE-1 expression in human dopaminergic neurons as well and serves as an example for a
600 neuronal sub-type specific regulation of LINE-1.

601 The heterogenous, brain-wide presence of ORF1p expression at steady-state is
602 intriguing. In cancer cell lines or mouse spermatocytes, ORF1p interacts with several "host"
603 proteins, some if not most of which are related to the LINE-1 life cycle. However, a profile of
604 endogenous ORF1p interactors in the mouse brain might inform on possible other and organ-
605 specific functions besides its binding to the LINE-1 RNA in "cis" [62]. Among the total 424
606 potential interactors of endogenous ORF1p in the mouse brain, 38 partners had been previously
607 identified by mass spectrometry in human cancers, cancerous cell lines and mouse
608 spermatocytes [53,58–63] (TableS2). Further, GO term analysis contained expected categories like
609 "P-body", mRNA metabolism related categories and "ribonucleoprotein granule". We also
610 identified NXF1 as a protein partner of ORF1p, a protein found to interact with LINE-1 RNA
611 related to its nuclear export [80]. This suggests the conservation of key interactors probably
612 essential for completing or repressing the LINE-1 life cycle in both species, despite the
613 divergence of mouse and human ORF1p protein sequences [81]. Along these lines, several ORF1p
614 protein partners we identified might complete the list of post-transcriptional regulators
615 implicated in LINE-1 silencing. Recent work conducted on human cancerous cell lines has
616 demonstrated that MOV10 orchestrates the recruitment of DCP2 for LINE-1 RNA decapping
617 [64]. In our analysis, we identified DCP2 along with DCP1A, known to enhance the decapping
618 activity of DCP2 [82], and DCP1b, a pivotal component of the mRNA decapping complex [83].
619 Intriguingly, MOV10 was not detected in our mass spectrometry analysis, despite its established
620 role in recruiting DCP2 and forming a complex with L1-RNP to mediate LINE-1 RNA
621 decapping, as reported by Liu et al. However, we found two enhancers of mRNA decapping,
622 EDC3 and EDC4, both core components of P-bodies, a membrane-less organelle known to
623 contain L1-RNP [51]. Multiple ubiquitin-ligase proteins were found although not appearing as a

624 significantly enriched GO term. These results complete the picture of the post-transcriptional
625 and translational control of ORF1p and suggest that these mechanisms, despite a steady-state
626 expression, are operational in neurons. However, several neuron-specific interactors were
627 identified belonging to GO term categories "neuron projection" (75 proteins) and "neuronal cell
628 body" (5 proteins), again pointing to the neuron-predominant expression of ORF1p in the
629 mouse brain. Other interesting aspects were raised from this analysis. Among significantly
630 enriched GO terms, several were related to the cytoskeleton, the functional consequences of
631 which need to be determined in future studies. Our screen also identified PDE10A as an
632 interactor of ORF1p in the mouse brain, a PDE almost exclusively expressed in medium spiny
633 neurons of the striatum and a target for treatment of neurological diseases related to basal
634 ganglia function like Huntington's disease, schizophrenia and Tourette syndrome [84].
635 Interestingly, PDE10A inhibition is related to beta-catenin signaling, another GO term which
636 emerged from our screen [85]. Finally, we found components of RNA polymerase II and the
637 SWI/SNF complex as partners of ORF1p. This further indicates that ORF1p has access to the
638 nucleus in mouse brain neurons as described for other cells [86,87] and recently in human
639 dopaminergic neurons *in vitro* [28], implying that ORF1p potentially has access to chromatin.
640 These findings give rise to intriguing questions concerning the potential function of ORF1p as
641 ORF1p is widely distributed throughout the brain under normal physiological conditions and
642 interacts with a variety of proteins which might not directly relate to the LINE-1 life cycle.
643 Moreover, ORF1p exhibits a prominent neuronal presence, with some neurons expressing up
644 to eight times more ORF1p than others and neurons that do not express ORF1p at all. Future
645 loss-of-function studies should help to shed light on the necessity of ORF1p for neuronal
646 functions if they exist. Our data spurs the idea of a possible "physiological" function of ORF1p
647 as an integrative protein with exapted function in neuronal homeostasis and a loss of restriction
648 in the aged brain limiting LINE-1 expression to steady-state levels.

649

650 ## Materials & Methods

651 **Animals**

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

655

656 **Mouse tissue dissection and protein extraction**

657 Tissues were extracted from 3-month-old and 16-month-old Swiss/OF1 mice. Briefly, the two
658 hemispheres were separated in ice cold PBS -/-. For each mouse, one hemisphere was rinsed
659 and fixed in 4% PFA for 1h followed by 24h of incubation in 30% sucrose and then washed for
660 10 min with PBS 1X 3 times. Hemispheres were kept at -20°C until being sliced on a freezing
661 microtome (Epredia, HM 450) with a 20µm thickness. The other hemisphere was dissected in
662 ice cold PBS -/- 1X and 6 brain regions were rinsed, cut in small pieces and dissociated
663 separately using a large (21G) to small gauge (27G) needle in RIPA lysis buffer for 5 min.
664 Lysates were kept on ice for 25 min, were sonicated for 15 min and supernatants were collected
665 after a 30 min centrifugation at 4°C at 14 000 rpm. Proteins were quantified and Laemmli buffer
666 was added before boiling for 10min at 95°C to be used for Western Blot.

667

668 **Human Samples:**

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

671

672 **Human Samples pulverization and protein extraction**

673 We used the dry pulverizer Cryoprep (Covaris) for pulverization of tissue blocs. Each sample
674 was disposed in a liquid-nitrogen precooled Tissue-tube bag and dry cryo-pulverized with one
675 impact at the maximum level. The pulverized brain sample was then weighed and resuspended
676 in lysis buffer (mg/v) (0.32M sucrose, 5mM $CaCl_2$, 3mM $Mg(CH3COOH)_2$, 0.1mM EDTA,
677 10mM Tris-HCL pH8, 1mM DTT, 0.1% TritonX-100 and Protease Inhibitors), kept on ice for
678 30 min with gentle up-and-down pipetting until homogenization. We added 2X RIPA buffer
679 (v/v) to totals fractions for 30 min on ice. We then sonicated samples 2 times for 15 min.
680 AtlasSupernatants were collected after a 30 min centrifugation at 14 000 rpm at 4°C, proteins
681 were quantified and Laemmli buffer was added to be used for Western Blot. All samples were
682 boiled 10 min at 95°C to be used for Western Blot.
683
684 **Western-Blot:**
685 We used 1.5mm NuPAGE 4-12% Bis-Tris-Gel (Invitrogen™). Proteins samples were loaded
686 and gel migration was performed with NuPAGE™ MES SDS Running Buffer (Invitrogen™)
687 for 45 minutes at 200mV. Gels were transferred onto a methanol activated PVDF membrane
688 (Immobilon) in a buffer containing: Tris 25 mM, pH=8,3 and Glycine 192 mM, during 1h30 at
689 400 mA. Membranes were blocked 30 min with 5% milk in TBST (0,2% Tween 20, 150 mM
690 NaCl, 10 mM Tris pH:8). Primary antibodies were diluted in 5% milk in TBST, and membranes
691 were incubated o/n at 4 C°. After 3 x 10 min washing in TBST, membranes were incubated for
692 1h30 with the respective secondary antibodies diluted at a concentration of 1/2000 in 5% milk
693 TBST. Membranes were washed 3 x 10 min in TBST and were revealed by the LAS-4000
694 Fujifilm system using Clarity Western ECL Substrate (Bio Rad) or Maxi Clarity Western ECL
695 Substrate (Bio Rad).
696
697 **Immunostaining:**

698 Sagittal mouse brains slices were fixed for 10min in PFA 4% and rinsed 3 times for 10min in
699 PBS -/-. Slices were then incubated 20 min in glycine 100mM, washed 3 times for 5min in PBS
700 and immersed in 10mM citrate pH 6 at 62°C during 45min for antigen retrieval. Slices were
701 then immersed 3 times in PBS with Triton X-100 0.2% and incubated in blocking buffer for 1,5
702 h (PBS with Triton X-100 0.2% and FBS (10%) previously inactivated 20min at 56°C (Gibeco,
703 16141061). Primary antibodies (ORF1p antibody: abcam ab216324; NeuN antibody: GeneTex
704 GTX00837) were diluted (1/200 and 1/500 respectively) in blocking buffer and incubated with
705 slices overnight at 4°C and then washed 3 times for 10 min with PBS. Suitable secondary
706 antibodies and Hoechst (Invitrogen, 15586276) were incubated for 1,5h at 1/2000 in PBS with
707 inactivated FBS (10%) and washed 3 times 10 min in PBS. To quench tissue autofluorescence,
708 especially lipofuscin, TrueBlack Plus (Biotium) in PBS was used during 10min. Slices were
709 rinsed 3 times in PBS and mounted with Fluoromount (Invitrogen™). For human cingulate
710 gyrus stainings, the same protocol was performed, with the difference that a human ORF1p
711 antibody (Abcam 245249) was used. Mouse and human brain slices were imaged by the
712 Axioscan 7 Digital Slide Scanner (Zeiss) or a Spinning Disk W1 (Yogogawa).
713
714 **Blocked peptide**
715 The ORF1p antibody (abcam ab216324) was incubated 2h on a turning wheel with excess (4:1)
716 of mouse ORF1p recombinant protein as in [16] before the blocked antibody was used in the
717 above-described immunofluorescence protocol.
718
719 **Quantification of confocal acquisitions**
720 Analysis was performed with a custom-written plugin [88] developed for the Fiji software [89], using
721 Bio-Formats [90] and 3D Image Suite [91] libraries. Nuclei were detected using Hoechst channel
722 with the 2D-stitched version of Cellpose [92] (model = 'cyto', diameter = 30, flow threshold =

723 0.4, cell probability threshold = 0.0, stitching threshold = 0.75). Z-stack images were rescaled
724 to obtain 3D nuclei that are filtered by volume to avoid false positive detections. NeuN and
725 ORF1p+ cells were detected in their respective channel in the same manner as nuclei, but with
726 different Cellpose settings (model = 'cyto2', diameter = 40, flow threshold = 0.4, cell
727 probability threshold = 0.0, stitching threshold = 0.75). Then, each cell was associated with a
728 nucleus having at least half of its volume in contact with it. Cells without any associated nucleus
729 were filtered out. Each nucleus was thus labeled according to NeuN and/or ORF1p positivity.
730

731 **ABBA Registration and Qupath analysis**

732 We registered each sagittal brain section with the Allen Mouse Brain Atlas (CCFv3 [93]), using
733 the Aligning Big Brains & Atlases plugin [94] in Fiji. Registration brain regions and subregions
734 were imported into QuPath software [95] for downstream processing.
735 In each brain subregion, cell analysis was done with custom Groovy scripts [88] developed for
736 QuPath. In brief, Hoechst nuclei were detected with StarDist 2D [10], applying the DSB 2018
737 pretrained model with the following parameters: percentile normalization = [1 - 99], probability
738 threshold = 0.75, overlap threshold = 0.25. Cells in NeuN and ORF1p channels were detected
739 with Cellpose 2D (percentile normalization = [1 - 99], model = 'cyto2', diameter = 30, flow
740 threshold = 0.4, cell probability threshold = 0.0). Nuclei and cells were then filtered by area and
741 intensity mean in their respective channel, once again to avoid false positive detections.
742 Minimum intensity threshold was based on the channel background noise. This noise was
743 estimated for each subregion as the mean intensity of pixels not belonging to any nucleus or
744 cell detected in the corresponding channel. Finally, each cell was associated with a nucleus
745 having its centroid located inside the cell mask. Cells without any associated nucleus were
746 filtered out. Each cell containing a nucleus was thus identified as NeuN+ or NeuN- and ORF1p+
747 or ORF1p-. Intensity means values were normalized by subtracting the background noise

748 computed in the corresponding channel and subregion. As a last step, subregional results were
749 merged into regional ones and data were analyzed using the Pandas Python library [96].
750
751 **FACS**
752 Mouse brains were dissociated with Adult Brain Dissociation kit (Miltenyi Biotec, 130-107-
753 677) and incubated with the coupled antibody NeuN Alexa 647 (Abcam, ab190565) or the
754 control isotype IgG Alexa 647 (Abcam, ab199093). Stained cells were filtered a last time with
755 a 40µm filter before FACS sorting (FACS ARIA II). Neuronal and non-neuronal cells were
756 collected in RIPA for protein extraction in an appropriate volume in order to achieve equal
757 cell concentrations (10 000 cells/µl).
758
759 **RNA-seq analysis**
760 The RNA-seq dataset from Dong et al [40] was downloaded from dbGAP (phs001556.v1.p1) and
761 contains unstranded paired-end 50bp and 75bp reads from pooled laser-capture micro-dissected
762 dopaminergic neurons from human post-mortem brain (107 samples) from 93 individuals w/o
763 brain disease. RNA-seq had been done on total and linearly amplified RNA. We focused our
764 analysis on data obtained with 50bp reads, in order to avoid mappability bias, while still
765 regrouping all age categories (n=41; with ages ranging from 38 to 97 (mean age: 79.88 (SD
766 ±12.07); n=6 </=65y; n=35 >65y; mean PMI: 7.07 (SD ±7.84), mean RIN: 7.09 (±0.94)).
767 Sequencing reads were aligned on the Human reference genome (hg38) using the STAR mapper
768 (v2.7.0a) 3 and two different sets of parameters. Genome-wide individual repeat quantification
769 was performed using uniquely mapped reads and the following STAR parameters: --
770 outFilterMultimapNmax 1 --alignEndsType EndToEnd --outFilterMismatchNmax 999 --
771 outFilterMismatchNoverLmax 0.06. Repeats class, family and name quantification was
772 performed using a random mapping procedure and the following parameters : --

773 outFilterMultimapNmax 5000 --outSAMmultNmax 1 --alignEndsType EndToEnd --
774 outFilterMismatchNmax 999 --outFilterMismatchNoverLmax 0.06 --outSAMprimaryFlag
775 OneBestScore --outMultimapperOrder Random.
776 Repeats annotations were downloaded from the UCSC Table Browser (repeatMasker database:
777 https://genome.ucsc.edu/cgi-bin/hgTables) and coordinates of LINE-1 full length and coding
778 elements were downloaded from the L1base database 2 (http://l1base.charite.de/l1base.php; [3])
779 selecting LINE-1 full length elements containing two predicted complete open reading frames
780 for ORF1 and ORF2 (UID= Unique IDentifier) from the LINE-1 database
781 (http://l1base.charite.de/l1base.php) and corrected genomic intervals with the repeat masker
782 annotation of the corresponding genomic locus. Repeat quantification from the aligned data
783 was done using a gtf file composed of all genes (Gencode v29) and all individual repeat
784 elements. This strategy was used to avoid overestimation of repeat elements due to overlaps
785 with expressed genes. For individual repeat quantification of the full length L1 elements
786 (L1base), we therefore used a gtf of all genes and all L1base entries, and ran the FeatureCounts
787 tool [97] with the following parameters: -g gene_id -s 0 -p. In the context of the family-based
788 analysis, we used a gtf with all genes and all annotated repeats elements and ran FeatureCounts
789 with -g gene_family -s 0 -p -M. Before DeSeq2 analysis, we remove all genes and repeat
790 elements with less than 10 reads in a minimum of n individuals, n being the number of
791 individuals in the condition containing the fewest individuals ("young" condition: n=6, 38-65y,
792 mean 57.5 years). We use the same conditions with genes and UIDs with less than 3 reads in a
793 minimum of n individuals. Finally, we calculated the scaling factors using DeSeq2 on all genes
794 + all repeat elements or all gene + UID according to the quantification method and then applied
795 these scaling factors to the corresponding counts tables.
796 In order to test for the mappability of each UID (= full-length and coding LINE-1), we extracted
797 the bed track « main on human:umap50 (genome hg38) from the UCSC genome browser (≈

798 7Mio regions) directly into Galaxy (usegalaxy.org) and joined genomic intervals with a
799 minimum overlap of 45bp of this dataset with a dataset containing the annotation of UIDs
800 extracted from L1Basev2 [3] corrected in length with repeat masker and completed with
801 information on whether the UID is intra- or intergenic and, if intragenic, in which gene
802 (NM_ID, chr, strand, start, end, gene length, number of exons, gene symbol) the fl-LINE-1 is
803 located, which resulted in 1266 regions. We then used the "group on data and group by"
804 function in Galaxy and counted the number of overlapping 50kmers with all 146 UIDs
805 (=mappability score). Correlation analysis (non-parametric Spearman) was then done between
806 the mappability score and the normalized read counts.
807
808 **Immpunoprecipitation of ORF1p from the mouse brain**
809 For immunoprecipitation, we used ORF1p (abcam, ab245122) and IgG rabbit (abcam,
810 ab172730) antibodies. The antibodies were coupled to magnetic beads using the Dynabeads®
811 Antibody Coupling Kit (Invitogen, 14311D) according to the manufacturer's recommendations.
812 We used 5µg of antibody for 1 mg of beads and used 1.5mg of beads for IP. Individual mouse
813 brain lysates (n=5), homogenized using dounce and sonicated, were incubated with ORF1p or
814 IgG-control coupled beads and a small fraction was kept as input. Each of these two tubes
815 containing coupled beads and brain lysates were diluted in 5 ml buffer (10 mM Tris HCl, 150
816 mM NaCl, protease inhibitor). The samples were then incubated overnight on a wheel at 4°C.
817 Samples were then washed 3 times with 1 ml buffer (10 mM Tris HCl pH 8, 150 mM NaCl)
818 using a magnet and then resuspended in the same buffer. The samples were boiled in Laemmli
819 buffer (95°C, 10 min) and 20 µl of each sample were deposited on a 4-12% Nupage gel
820 (Invitrogen, NP0336) to be revealed by WB. For samples used in Mass Spectrometry study,
821 beads were washed with buffer (10 mM Tris HCl pH 8, 150 mM NaCl) using a magnet. After
822 3 washes with 1ml buffer the beads were washed twice with 100 µL of 25 mM NH4HCO3

823 (ABC buffer). Finally, beads were resuspended in 100 µl of 25mM ABC buffer and digested
824 by adding 0.20 µg of trypsine/LysC (Promega) for 1 hour at 37 °C. A second round of digestion
825 was applied simultaneously on the beads by adding 100 µL of 25 mM ABC buffer and to the
826 previous digest by adding 0.20 µg of trypsin/LysC for 1 hour at 37 °C. Samples were then
827 loaded into homemade C18 StageTips packed by stacking three AttractSPE® disk (#SPE-
828 Disks-Bio-C18-100.47.20 Affinisep) into a 200 µL micropipette tip for desalting. Peptides were
829 eluted using a ratio of 40:60 CH3CN:H2O + 0.1% formic acid and vacuum concentrated to
830 dryness with a SpeedVac device. Peptides were reconstituted in 10 µL of injection buffer in
831 0.3% trifluoroacetic acid (TFA) before liquid chromatography-tandem mass spectrometry (LC-
832 MS/MS) analysis.
833
834 **Mass Spectrometry**
835 Online chromatography was performed with an RSLCnano system (Ultimate 3000, Thermo
836 Scientific) coupled to a Q Exactive HF-X with a Nanospay Flex ion source (Thermo Scientific).
837 Peptides were first trapped on a C18 column (75 µm inner diameter × 2 cm; nanoViper Acclaim
838 PepMapTM 100, Thermo Scientific) with buffer A (2/98 MeCN/H2O in 0.1% formic acid) at
839 a flow rate of 2.5 µL/min over 4 min. Separation was then performed on a 50 cm x 75 µm C18
840 column (nanoViper Acclaim PepMapTM RSLC, 2 µm, 100Å, Thermo Scientific) regulated to
841 a temperature of 50°C with a linear gradient of 2% to 30% buffer B (100% MeCN in 0.1%
842 formic acid) at a flow rate of 300 nL/min over 91 min. MS full scans were performed in the
843 ultrahigh-field Orbitrap mass analyzer in ranges m/z 375–1500 with a resolution of 120 000 at
844 m/z 200. The top 20 intense ions were subjected to Orbitrap for further fragmentation via high
845 energy collision dissociation (HCD) activation and a resolution of 15 000 with the intensity
846 threshold kept at 1.3 × 105. We selected ions with charge state from 2+ to 6+ for screening.
847 Normalized collision energy (NCE) was set at 27 and the dynamic exclusion of 40s. For

848 identification, the data were searched against the Mus Musculus (UP000000589_10090
849 012019) Uniprot database using Sequest HT through proteome discoverer (version 2.4).
850 Enzyme specificity was set to trypsin and a maximum of two-missed cleavage sites were
851 allowed. Oxidized methionine, Met-loss, Met-loss-Acetyl and N-terminal acetylation were set
852 as variable modifications. Maximum allowed mass deviation was set to 10 ppm for
853 monoisotopic precursor ions and 0.02 Da for MS/MS peaks. The resulting files were further
854 processed using myProMS97 v3.10.0. FDR calculation used Percolator and was set to 1% at
855 the peptide level for the whole study. The label free quantification was performed by peptide
856 Extracted Ion Chromatograms (XICs), reextracted by conditions and computed with
857 MassChroQ version 2.2.21 [98]. For protein quantification, XICs from proteotypic peptides
858 shared between compared conditions (TopN matching) with missed cleavages were used.
859 Median and scale normalization at peptide level was applied on the total signal to correct the
860 XICs for each biological replicate (n=5). To estimate the significance of the change in protein
861 abundance, a linear model (adjusted on peptides and biological replicates) was performed, and
862 p-values were adjusted using the Benjamini–Hochberg FDR procedure. Proteins with at least
863 three peptides, identified in each biological replicates of ORF1p condition, a 10-fold enrichment
864 and an adjusted p-value ≤ 0.05 were considered significantly enriched in sample comparisons.
865 Unique proteins were considered with at least three peptides in all replicates. Protein selected
866 with these criteria were used for Gene Ontology enrichment analysis and string network
867 analysis.
868 The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consorti-
869 um (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository [99] with the
870 dataset identifier PXD047160.
871
872 **GO term and STRING network analysis**

873 Gene Ontology analysis was performed using GO PANTHER (PANTHER: Making genome-
874 scale phylogenetics accessible to allPaul D. Thomas, Dustin Albou and Huaiyu MiProtein
875 Society. 2022;31(1):8-22. doi:10.1002/pro.4218) and String network physical interactions were
876 retrieved using the STRING database v11.5 (https://string-db.org/) and then implemented in
877 Cytoscape software [100].
878
879 **Statistical analyses**
880 In column comparisons, data in each column were tested for normality using two normality and
881 lognormality tests (Shapiro-Wilk test and Kolmogorov-Smirnov test). Data which passed the
882 normality tests were analyzed subsequently by a parametric test, data which did not pass the
883 normality tests were analyzed by a non-parametric statistical test as indicated in the figure
884 legends. The significance threshold was defined as $p<0.05$. Statistical analyses were done with
885 PRISM software (v10).
886
887
888 **Acknowledgments**
889 This work was supported by the Fondation de France (00086320, to J.F.), the Fondation du
890 Collège de France (to J.F. and T.B.), the Fondation NRJ/Institut de France (to J.F.) and the
891 National French Agency for Research (ANR-20-CE16-0022 NEURAGE). We gratefully
892 acknowledge the Collège de France Orion imaging facility (IMACHEM-IBiSA), member of
893 the French National Research Infrastructure France-BioImaging (ANR-10-INBS-04), which
894 received support from the program «Investissements d'Avenir» ANR-10-LABX-54
895 MEMOLIFE. We also thank the Fondation Bettencourt Schueller for their support.
896
897

Contributions

J.F. and R.L.J. conceived the project. 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..

Competing interests

The authors declare no competing interests.

923 **Figure Legends**

924

925 **Figure 1: Widespread and heterogenous expression of ORF1p protein in the mouse brain**

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

938 (B) Widespread and heterogenous expression of the LINE-1 encoded protein ORF1p in the
939 mouse brain. Representative image of ORF1p immunostaining (orange) of a sagittal section of
940 the brain of a young (three months-old) mouse acquired on a slide scanner. Scale bar = 1mm.
941 (1-10) Representative images of immunostainings showing ORF1p expression (orange) in 10
942 different regions of the mouse brain acquired on a confocal microscope. Nuclei are represented
943 in blue (Hoechst), scale bar = 50µm. (1) Isocortex, (2) Hippocampus, (3) Striatum dorsal, (4)
944 Thalamus, (5) Midbrain motor, (6) Pallidum, (7) Hypothalamus, (8) Substantia nigra pars
945 compacta, (9) Hindbrain, (10) Cerebellum.

946 (C) ORF1p expression profile in the mouse brain. The entire mouse brain with the exception of
947 the olfactory bulb and the cerebellum were analyzed according to the pipeline on large-scale

948 images described in (A). Bar plot showing the total number of ORF1p+ cells per mm² in the
949 mouse brain. Data is represented as mean ±SEM, n=4 mice (top). Bar plot indicating the
950 proportion of ORF1p+ cells compared to all cells detected. Data is represented as mean ±SEM,
951 n=4 mice, 202001 total cells analyzed (middle). Scatter plot showing the mean intensity of
952 ORF1p per ORF1p+ cell. Data is represented as mean ±SD, n=4 mice, 40999 ORF1p+ cells
953 analyzed (bottom).
954 (D-F) ORF1p expression profile (density, proportion and expression) in defined anatomical
955 regions of the mouse brain. Nine anatomical regions as defined by the Allen Brain Atlas and
956 mapped onto sagittal brain slices with ABBA were analyzed using the pipeline on large scale
957 images described in (A). (D) ORF1p+ cell density in 9 different regions. Bar plot showing the
958 number of ORF1p+ cells per mm². Data is represented as mean ±SEM; *p<0.05; **p<0.01;
959 adjusted p-value, one-way ANOVA followed by a Benjamin-Hochberg test. (E) Proportion of
960 ORF1p positive cells in 9 different regions. Bar plot showing the proportion of ORF1p+ cells
961 among all cells detected per region. Data is represented as mean ±SEM. (F) Mean ORF1p
962 expression per cell in 9 different regions. plot showing the mean intensity of ORF1p signal per
963 ORF1p+ cell in 9 different regions. Data is represented as mean ±SD. The number of analyzed
964 cells per region is indicated in the figure. *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001;
965 adjusted p-value, nested one-way ANOVA followed by Sidak' multiple comparison test.
966 (G) ORF1p expression in the frontal cortex and ventral midbrain. Confocal images with
967 multiple z-stacks were analyzed using Cellpose and Stardist. Dot plot representing the mean
968 expression of ORF1p per ORF1p+ cells. Individual (four three-month-old Swiss/ OF1) mice
969 are represented each by a different color, the scattered line represents the median.
970 ****p<0.0001, nested one-way ANOVA. Total cells analyzed = 4645.
971 (H-I) ORF1p expression in the frontal cortex and the ventral midbrain. (H) Western blots
972 showing ORF1p (top) and actin expression (bottom) in four individual mice per region which

973 were quantified in (I) using actin as a reference control. The signal intensity is plotted as the
974 fold change of ORF1p expression in the ventral midbrain to ORF1p expression in the frontal
975 cortex. *p<0.05 ob; two-sided, unpaired student's-test.
976
977 **Figure 2: ORF1p is predominantly expressed in neurons in the mouse brain**
978 (A) ORF1p expression is absent in the white matter (corpus callosum) and predominantly
979 expressed in neurons. Proportion of ORF1p+/NeuN+, ORF1p+/NeuN-, ORF1p-/NeuN+ and
980 ORF1p-/NeuN- cells in the white matter (corpus callosum) and the grey matter (left) and in
981 nine different regions (right) analyzed by the cell detection pipeline on large scale images
982 presented in *Figure1A*. Exact values can be found in *Suppl. Table 1*.
983 (B) Representative confocal microscopy images showing ORF1p (red) and NeuN expression
984 (green) in two different regions of the mouse brain. The bottom images show the merge of the
985 two stainings, an overlap of both markers is represented in orange. z-projections; scalebar =
986 25µm
987 (C) ORF1p is predominantly expressed in neurons. Proportion of ORF1p+/NeuN+,
988 ORF1p+/NeuN+, ORF1p+/NeuN-, ORF1p-/NeuN+ and ORF1p-/NeuN- cells, ****p<0.0001;
989 calculated using chi-square test on the cell number of the four different cell-types analyzed by
990 confocal microscopy on multiple z-stacks (p2).
991 (D) Proportion of ORF1p+ cells identified as NeuN+ or NeuN- in two different regions,
992 analyzed by confocal microscopy on multiple z-stacks (p2). ns: non-significant; chi-square test,
993 n=4 mice.
994 (E) Proportion of neurons (NeuN+) expressing ORF1p in different regions of the mouse brain
995 as quantified using the cell detection pipeline on large scale images (p1).
996 (F) Proportion of neurons expressing ORF1p in the frontal cortex and ventral midbrain
997 quantified on confocal images using confocal approach. ns: non-significant; chi-square test on

998 the cell number of the different cell-types analyzed; n=4 mice, data is represented as mean ±
999 SEM.
1000 (G) ORF1p does not colocalize with glial or microglial cell markers. Representative confocal
1001 microscopy images showing ORF1p staining (red) and three different glial cell (GFAP, Sox9,
1002 S100b) or microglial (Iba1) markers (white). z-projections, scalebar = 50µm.
1003 (H) Representative acquisition showing NeuN positives cells (green) In grey (right) but not in
1004 white matter (left) in a cingular gyrus in healthy human brain acquire on slide scanner; scale
1005 bar = 400µm (left) and zoom on grey matter with a confocale showing ORF1p in red;
1006 representative image scale bar = 5µm (right).
1007 (I-J) Isolation of neuronal cells by FACS confirms predominant neuronal expression of ORF1p.
1008 (I) Neuronal (NeuN+) cells isolated by fluorescent activated cell sorting (FACS). Dot plots
1009 showing autofluorescence versus an appropriate control antibody (IgG rabbit 647; left) and an
1010 antibody against NeuN (name of the AB 657, right). The P4 window represents isolated NeuN+
1011 cells (pink) and the P5 fraction NeuN- cells (orange) containing the same number of cells as
1012 sorted in P4 for comparison, others NeuN- are represented in blue. (J) Western blot. ORF1p
1013 expression (top), in NeuN- and NeuN+ FACS-sorted cells stemming from I.
1014
1015 **Figure 3: ORF1p expression is increased throughout the whole mouse brain in the context**
1016 **of aging**
1017 (A) Proportion of ORF1p+/NeuN+, ORF1p+/NeuN+, ORF1p+/NeuN-, ORF1p-/NeuN+ and
1018 ORF1p-/NeuN- cell-type in the whole brain (left) and the different analyzed regions (right) of
1019 young (three-month aged, n=4) and aged (16-month-old, n=4) mice using the cell detection
1020 pipeline on large scale images presented in Figure 1A (p1), data is represented as mean ± SEM.
1021 (B) Proportion of ORF1p+/NeuN+, ORF1p+/NeuN+, ORF1p+/NeuN-, ORF1p-/NeuN+ and
1022 ORF1p-/NeuN- cell-type in two different regions of young and aged mice, analyzed on multiple

1023 z-stack confocal images (p2). ns: non-significant; *p<0.05 calculated using two-way ANOVA
1024 with sidak's multiple comparisons test on the cell number of the different cell-types analyzed;
1025 data is represented as mean ± SEM.
1026 (C) ORF1p mean expression per ORF1p+ cell in the brain analyzed on large-scale images (p1).
1027 Plot showing the ORF1p mean expression per ORF1p+ cell in young (n=4) and aged (n=4)
1028 mice in the whole brain (except cerebellum and olfactory bulb). 74985 total cells were analyzed;
1029 * p<0.05, two-way ANOVA with sidak's multiple comparisons; data is represented as mean ±
1030 SEM.
1031 (D) Frequency distribution of ORF1p mean intensity in ORF1p+ cells. ***p<0.001,
1032 Kolmogorov-Smirnov test.
1033 (E) Frequency distribution of Hoechst mean intensity in the nuclei of ORF1p+ cells. ns: non-
1034 significant, Kolmogorov-Smirnov test.
1035 (F) Mean ORF1p expression per ORF1p+ cell in nine different anatomical regions. Plot
1036 showing the ORF1p mean expression per ORF1p positive cell (n=74985). Adjusted p-value are
1037 represented, two-tailed nested t-test followed by a Benjamin, Krieger and Yukutieli test; n=4
1038 young and n=4 aged mice per region, data is represented as mean ± SEM.
1039 (G) Color-coded representation of fold-changes of ORF1p expression changes with aging.
1040 Represented is the fold-change in percent (aged vs young) of the "mean of the mean" ORF1p
1041 expression per ORF1p+ cell quantified mapped onto the nine different regions analyzed as
1042 shown in (F).
1043 (H) ORF1p expression is increased in the ventral midbrain of aged mice. Dot plot representing
1044 ORF1p expression in two different regions of young and aged mice analyzed on confocal
1045 images with multiple z-stacks; total cells analyzed = 8381 ns: non-significant *p<0.05, two-
1046 tailed one-way ANOVA; dashed lines represent the medians.

1047 (I) Representative confocal microscopy acquisition showing increased ORF1p expression (red)
1048 in the ventral midbrain region of aged mice (one z plan is shown). Cell nuclei are shown in blue
1049 (Hoechst staining). Scalebar = 50µm.
1050
1051 **Figure 4: Young LINE-1 elements are increased in aged human dopaminergic neurons**
1052 TE transcript expression in RNA-seq data of laser-captured micro-dissected post-mortem
1053 human dopaminergic neurons of brain-healthy individuals was analyzed using RepeatMasker
1054 (multimappers) or the L1Base (unique reads).
1055 (A) Volcano plot of differential analysis of LINE-1 expression using DESeq2 comparing young
1056 (</=65y, n=6) or aged (>65y, n=35) human dopaminergic neurons at the "name" level of
1057 RepeatMasker. Young LINE-1 elements, including the two families L1HS and L1PA2 that have
1058 coding copies, are highlighted in red.
1059 (B) Scatter plots of normalized read counts ("name" level) of the young L1HS and L1PA2
1060 families as well as the human endogenous virus family HERVK-int, another TE family with
1061 coding potential comparing young (</=65y, n=6) or aged (>65y, n=35) human dopaminergic
1062 neurons. Mann-Whitney test, $p<0.05$.
1063 (C) Volcano plot of differential expression analysis of TE expression using DEseq2 comparing
1064 young (</=65y, n=6) and aged (>65y, n=35) human dopaminergic neurons at the locus-level of
1065 specific full-length LINE-1 elements (140 of 146 "UID's" as annotated in L1Base; threshold
1066 >3 reads in at least 6 individuals).
1067 (D) Pairwise comparison of the expression of 140 out of 146 full-length LINE-1 elements
1068 comparing young (</=65y, n=6) and aged (>65y, n=35) human dopaminergic neurons.
1069 Wilcoxon matched signed rank test, $p<0.0001$ (left panel).
1070 (E) The sum of read counts of all UIDs per individual were plotted comparing young (</=65y,
1071 n=6) and aged (>65y, n=35) human dopaminergic neurons.

1072 (F-G) Dysregulated locus-specific full-length LINE-1 elements (UID-68 and UID-129) are
1073 plotted as scatter plots comparing young (</=65y, n=6) and aged (>65y, n=35) human
1074 dopaminergic neurons. (F) UID-68 is located adjacent to the genes CLEC5A and OR9A4 (left).
1075 Spearman correlation analysis of the expression of UID-68 and CLEC5A (middle) or OR9A4
1076 (right) in young (</=65y, n=6, black dots) or aged (>65y, n=35, red squares) human
1077 dopaminergic neurons. (G) UID-129 is intergenic.
1078
1079 **Figure 5: Potential regulators of LINE-1 expression in human dopaminergic neurons**
1080 Correlation of the expression of evolutionary related and distant LINE-1 elements with known
1081 regulators in human dopaminergic neurons
1082 (A) Spearman correlation of evolutionary close (L1HS, L1PA2-17) and distant LINE-1 (L1PB
1083 and L1MA5) as well as HERV elements with coding potential (HERV-Kint, HERV-Fc1,
1084 HERV-Fc2 and HERV-H-int) with known regulators of their expression. HERV-W and
1085 TREX1 did not pass the normalized read count threshold of >3 reads in >6 individuals.
1086
1087 **Figure 6: Endogenous ORF1p interactors in the mouse brain**
1088 Immunoprecipitation (IP) of endogenous ORF1p from the mouse brain. WB against ORF1p
1089 showing ORF1p enrichment after IP but no signal in the IgG control. Five independent samples
1090 were then prepared for proteomic analysis by mass spectrometry (LC-MS/MS).
1091 (B) GO slim enrichment analysis of proteins identified as endogenous ORF1p protein partners
1092 in the mouse brain by LC/MS-MS. ORF1p-immunoprecipitated proteins were categorized into
1093 GO slim terms. The nine GO slim term with the highest fold-change are plotted. Fold
1094 enrichment is depicted on the upper axis and displayed as bars, the FDR value appears on the
1095 lower axis and is represented by the black points. BP: Biological Process, CC: Cellular
1096 Component, MF: Molecular Function.

1097 (C) Venn diagram showing common interactors (purple) between interactors of endogenous
1098 ORF1p in the mouse brain identified in this study (red) and known (published) interactors of
1099 ORF1p (blue). Statistical significance of the overlap between the two groups of proteins was
1100 tested by an overrepresentation test (http://nemates.org/MA/progs/overlap_stats.html).
1101 (D) ORF1p associates with the SWI/SNF complex (red), RNA pol II complex (orange) and
1102 interactors belonging to GO terms related to neuronal cell body & neuron projection (green).
1103 Known interactors previously published [53,58–63] are indicated with a purple ring. STRING
1104 network of physical interactions where nodes represent proteins partners identified in (A) and
1105 edges thickness represents the strength of shared physical complexes. Only proteins sharing
1106 physical interactions were represented.
1107
1108
1109 **Supplementary Figures**
1110
1111 Suppl Figure 1
1112 (A) Selective recognition of ORF1p antibody. IHC showing ORF1p positives cells in sagittal
1113 mouse brain slice (left) and abolition of the signal when blocking the antibody with purified
1114 ORF1p (right).
1115 (B) ORF1p is expressed in six different brain regions in the mouse.
1116 Brain regions were micro-dissected from a three-month old mouse brain. Western blot showing
1117 ORF1p expression in 6 brain regions. ORF1p (Top), Actin (bottom).
1118
1119 Suppl Figure 2
1120 (A) ORF1p cell identity. Proportion of ORF1p+ cells identified as NeuN+ (black) or NeuN-
1121 (grey), in the whole brain (left) and in 9 different regions analyzed (right) using the cell

1122 detection pipeline on large scale images presented in Figure 1A; data is represented as mean ±
1123 SEM, n=4 mice.
1124 (B) Proportion of neurons in the frontal cortex and ventral midbrain quantified using confocal
1125 approach. *p<0.05, chi-square test on the cell number of the different cell-types analyzed; n=4
1126 mice, data is represented as mean ± SEM.
1127
1128 Suppl Figure 3
1129 (A) Proportion of ORF1p+ cells being neuronal in the ventral midbrain comparing young and
1130 aged mice as quantified using confocal approach. Kolmogorov-Smirnov test; data is
1131 represented as mean ± SEM.
1132 (B) Proportion of neurons expressing ORF1p in the ventral midbrain comparing young and
1133 aged mice as quantified using confocal approach. Kolmogorov-Smirnov test; data is
1134 represented as mean ± SEM.
1135 (C) Proportion of neurons in the ventral midbrain comparing young and aged mice as quantified
1136 using confocal approach**.** Kolmogorov-Smirnov test; data is represented as mean ± SEM.
1137
1138 Suppl Fig4
1139 (A-B) Comparison of the expression of dopaminergic markers tyrosine hydroxylase (TH, A)
1140 and LMX1B (B) between young (</=65y, n=6)) and aged (>65y, n=35) human dopaminergic
1141 neurons. Mann Whitney test.
1142 (C, E) Volcano plot of differential expression analysis of TE expression using DEseq2
1143 comparing young (</=65y, n=6) and aged (>65y, n=35) human dopaminergic neurons at the
1144 "class" (C) and "family" (E) level of RepeatMasker.
1145 (D) Scatter plot comparing the expression of LINE at the "class" level between young (</=65y,
1146 n=6)) and aged (>65y, n=35) human dopaminergic neurons. Mann Whitney test.

1147 (F) Scatter plot comparing the expression of LINE and Alu at the "family" level between young
1148 (</=65y, n=6)) and aged (>65y, n=35) human dopaminergic neurons. Mann Whitney test.
1149 (G) Scatter plot comparing the expression of HERVH-int, HERV-Fc1 and two non-coding,
1150 non-autonomous but active TEs in the human genome, AluYa5 and SVA-F at the "name" level
1151 between young (</=65y, n=6)) and aged (>65y, n=35) human dopaminergic neurons. Mann
1152 Whitney test.
1153
1154 Suppl Fig5
1155 (A-C) Correlation analyses of L1HS expression with Engrailed 1 (EN1, A, Spearman r=-0.43,
1156 p=0.002), CBX5/HP1 (B, Spearman r=-0.35, p=0.01) and XRCC6 expression (C, Spearman r=
1157 -0.394, p=0.005). Normalized read counts are plotted. Black dots correspond to young
1158 individuals (</=65y), red dots correspond to aged individuals (>65y).
1159 (D-G) Scatter plots comparing the expression of EN1, CBX5/HP1, XRCC5 and XRCC6
1160 between young (</=65y, n=6)) and aged (>65y, n=35) human dopaminergic neurons. Student's
1161 t-test (EN1) or Mann Whitney test (HP1, XRCC5/6).
1162
1163 Suppl Fig6
1164 (A) Correlation of mappability (UMAP hit counts over UID, see methods) and UID expression
1165 (normalized read counts). Spearman correlation.
1166 (B-D) *In silico* analysis of annotated full-length LINE-1 elements as in L1Basev2 (human
1167 reference genome hg38). (B) Percentage of L1HS and L1PA2 elements among the 146 full-
1168 length elements (UID1-146). (C) Presence (blue, with gene symbol) or absence (white) of a
1169 "hosting" gene among the 146b annotated full-length LINE-1 in the human reference genome
1170 (D) Percentage of full-length LINE-1 elements located inside or outside a gene.

(E) Mean expression of all 146 full-length LINE-1 elements in dopaminergic neurons of all individuals </=65y.

Suppl Fig7

(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. A hidden human proteome encoded by “non-coding” genes - PubMed. Accessed October 19, 2023. https://pubmed.ncbi.nlm.nih.gov/31340039/
2. Brouha B, Schustak J, Badge RM, et al. 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*. 2003;100(9):5280-5285. doi:10.1073/pnas.0831042100
3. Penzkofer T, Jäger M, Figlerowicz M, et al. L1Base 2: more retrotransposition-active LINE-1s, more mammalian genomes. *Nucleic Acids Research*. 2017;45(D1):D68-D73. doi:10.1093/nar/gkw925
4. Mao J, Zhang Q, Cong YS. Human endogenous retroviruses in development and disease. *Computational and Structural Biotechnology Journal*. 2021;19:5978-5986. doi:10.1016/j.csbj.2021.10.037
5. Esnault C, Maestre J, Heidmann T. Human LINE retrotransposons generate processed pseudogenes. *Nat Genet*. 2000;24(4):363-367. doi:10.1038/74184
6. Wei W, Gilbert N, Ooi SL, et al. Human L1 retrotransposition: cis preference versus trans complementation. *Molecular and Cellular Biology*. 2001;21(4):1429-1439. doi:10.1128/MCB.21.4.1429-1439.2001
7. Mathias SL, Scott AF, Kazazian HH, Boeke JD, Gabriel A. Reverse transcriptase encoded by a human transposable element. *Science*. 1991;254(5039):1808-1810.

1205 doi:10.1126/science.1722352
1206 8. Feng Q, Moran JV, Kazazian HH, Boeke JD. Human L1 retrotransposon encodes a
1207 conserved endonuclease required for retrotransposition. *Cell*. 1996;87(5):905-916.
1208 9. Chen L, Dahlstrom JE, Lee SH, Rangasamy D. Naturally occurring endo-siRNA
1209 silences LINE-1 retrotransposons in human cells through DNA methylation. *Epigenetics :*
1210 *official journal of the DNA Methylation Society*. 2012;7(7):758-771. doi:10.4161/epi.20706
1211 10. Goodier JL, Cheung LE, Kazazian HH. MOV10 RNA helicase is a potent inhibitor of
1212 retrotransposition in cells. *PLoS Genetics*. 2012;8(10):e1002941.
1213 doi:10.1371/journal.pgen.1002941
1214 11. Philippe C, Vargas-Landin DB, Doucet AJ, et al. Activation of individual L1
1215 retrotransposon instances is restricted to cell-type dependent permissive loci. *eLife*. 2016;5.
1216 doi:10.7554/eLife.13926
1217 12. Peze-Heidsieck E, Bonnifet T, Znaidi R, et al. Retrotransposons as a Source of DNA
1218 Damage in Neurodegeneration. *Frontiers in aging neuroscience*. 2021;13:786897.
1219 doi:10.3389/fnagi.2021.786897
1220 13. Driver CJ, McKechnie SW. Transposable elements as a factor in the aging of
1221 Drosophila melanogaster. *Annals of the New York Academy of Sciences*. 1992;673(1):83-91.
1222 doi:10.1111/j.1749-6632.1992.tb27439.x
1223 14. Copley KE, Shorter J. Repetitive elements in aging and neurodegeneration. *Trends in*
1224 *Genetics*. Published online March 2023:S0168952523000343. doi:10.1016/j.tig.2023.02.008
1225 15. Gorbunova V, Seluanov A, Mita P, et al. The role of retrotransposable elements in
1226 ageing and age-associated diseases. *Nature*. Published online July 2021:1-11.
1227 doi:10.1038/s41586-021-03542-y
1228 16. Blaudin de Thé FX, Rekaik H, Peze-Heidsieck E, et al. Engrailed homeoprotein blocks
1229 degeneration in adult dopaminergic neurons through LINE-1 repression. *The EMBO Journal*.
1230 2018;37(15). doi:10.15252/embj.201797374
1231 17. Sur D, Kustwar RK, Budania S, et al. Detection of the LINE-1 retrotransposonRNA-
1232 binding protein ORF1p in differentanatomical regions of the human brain. *Mobile DNA*.
1233 Published online November 2017:1-12. doi:10.1186/s13100-017-0101-4
1234 18. McKerrow W, Kagermazova L, Doudican N, et al. LINE-1 retrotransposon expression
1235 in cancerous, epithelial and neuronal cells revealed by 5′ single-cell RNA-Seq. *Nucleic Acids*
1236 *Research*. 2023;51(5):2033-2045. doi:10.1093/nar/gkad049
1237 19. Gasior SL, Wakeman TP, Xu B, Deininger PL. The human LINE-1 retrotransposon
1238 creates DNA double-strand breaks. *Journal of molecular biology*. 2006;357(5):1383-1393.
1239 doi:10.1016/j.jmb.2006.01.089
1240 20. Belgnaoui SM, Gosden RG, Semmes OJ, Haoudi A. Human LINE-1 retrotransposon
1241 induces DNA damage and apoptosis in cancer cells. *Cancer cell international*. 2006;6:13.
1242 doi:10.1186/1475-2867-6-13
1243 21. De Cecco M, Criscione SW, Peckham EJ, et al. Genomes of replicatively senescent
1244 cells undergo global epigenetic changes leading to gene silencing and activation of
1245 transposable elements. *Aging Cell*. 2013;12(2):247-256. doi:10.1111/acel.12047
1246 22. Van Meter M, Kashyap M, Rezazadeh S, et al. SIRT6 represses LINE1
1247 retrotransposons by ribosylating KAP1 but this repression fails with stress and age. *Nature*
1248 *communications*. 2014;5:5011. doi:10.1038/ncomms6011
1249 23. Sturm Á, Ivics Z, Vellai T. The mechanism of ageing: primary role of transposable
1250 elements in genome disintegration. *Cell Mol Life Sci*. 2015;72(10):1839-1847.
1251 doi:10.1007/s00018-015-1896-0
1252 24. Wallace N, Wagstaff BJ, Deininger PL, Roy-Engel AM. LINE-1 ORF1 protein
1253 enhances Alu SINE retrotransposition. *Gene*. 2008;419(1-2):1-6.
1254 doi:10.1016/j.gene.2008.04.007

1255 25. Thomas CA, Tejwani L, Trujillo CA, et al. Modeling of TREX1-Dependent
1256 Autoimmune Disease using Human Stem Cells Highlights L1 Accumulation as a Source of
1257 Neuroinflammation. *Cell stem cell*. 2017;21(3):319-331.e8. doi:10.1016/j.stem.2017.07.009
1258 26. De Cecco M, Ito T, Petrashen AP, et al. L1 drives IFN in senescent cells and promotes
1259 age-associated inflammation. *Nature*. 2019;566(7742):73-78. doi:10.1038/s41586-018-0784-9
1260 27. Luqman-Fatah A, Watanabe Y, Uno K, Ishikawa F, Moran JV, Miyoshi T. The
1261 interferon stimulated gene-encoded protein HELZ2 inhibits human LINE-1 retrotransposition
1262 and LINE-1 RNA-mediated type I interferon induction. *Nat Commun*. 2023;14(1):203.
1263 doi:10.1038/s41467-022-35757-6
1264 28. Znaidi R, Massiani-Beaudoin O, Mailly P, Monnet H, Joshi RL, Fuchs J. Nuclear
1265 translocation of LINE-1 encoded ORF1p alters nuclear envelope integrity and disrupts
1266 nucleocytoplasmic transport in neurons. bioRxiv, 2023. doi:10.1101/2023.08.10.552479
1267 29. Gazquez-Gutierrez A, Witteveldt J, R Heras S, Macias S. Sensing of transposable
1268 elements by the antiviral innate immune system. *RNA (New York, NY)*. Published online April
1269 2021. doi:10.1261/rna.078721.121
1270 30. Krug L, Chatterjee N, Borges-Monroy R, et al. Retrotransposon activation contributes
1271 to neurodegeneration in a Drosophila TDP-43 model of ALS. *PLoS Genetics*.
1272 2017;13(3):e1006635. doi:10.1371/journal.pgen.1006635
1273 31. Casale AM, Liguori F, Ansaloni F, et al. Transposable element activation promotes
1274 neurodegeneration in a Drosophila model of Huntington's disease. *ISCIENCE*.
1275 2022;25(1):103702. doi:10.1016/j.isci.2021.103702
1276 32. Takahashi T, Stoiljkovic M, Song E, et al. LINE-1 activation in the cerebellum drives
1277 ataxia. *Neuron*. Published online August 2022:1-19. doi:10.1016/j.neuron.2022.08.011
1278 33. Gusel'nikova VV, Korzhevskiy DE. NeuN As a Neuronal Nuclear Antigen and
1279 Neuron Differentiation Marker. *Acta Naturae*. 2015;7(2):42-47. doi:10.32607/20758251-
1280 2015-7-2-42-47
1281 34. Cannon JR, Greenamyre JT. NeuN is not a reliable marker of dopamine neurons in rat
1282 substantia nigra. *Neuroscience Letters*. 2009;464(1):14-17. doi:10.1016/j.neulet.2009.08.023
1283 35. Sukhorukova EG. Nuclear Protein NeuN in Neurons in the Human Substantia Nigra.
1284 *Neurosci Behav Physi*. 2014;44(5):539-541. doi:10.1007/s11055-014-9947-9
1285 36. Kumar SS, Buckmaster PS. Neuron-specific nuclear antigen NeuN is not detectable in
1286 gerbil subtantia nigra pars reticulata. *Brain Res*. 2007;1142:54-60.
1287 doi:10.1016/j.brainres.2007.01.027
1288 37. Keller D, Erö C, Markram H. Cell Densities in the Mouse Brain: A Systematic
1289 Review. *Front Neuroanat*. 2018;12:83. doi:10.3389/fnana.2018.00083
1290 38. Sun W, Cornwell A, Li J, et al. SOX9 Is an Astrocyte-Specific Nuclear Marker in the
1291 Adult Brain Outside the Neurogenic Regions. *J Neurosci*. 2017;37(17):4493-4507.
1292 doi:10.1523/JNEUROSCI.3199-16.2017
1293 39. Yushkova E, Moskalev A. Transposable elements and their role in aging. *Ageing Res*
1294 *Rev*. 2023;86:101881. doi:10.1016/j.arr.2023.101881
1295 40. Dong X, Liao Z, Gritsch D, et al. Enhancers active in dopamine neurons are a primary
1296 link between genetic variation and neuropsychiatric disease. *Nature Neuroscience*. Published
1297 online September 2018:1-19. doi:10.1038/s41593-018-0223-0
1298 41. Teissandier A, Servant N, Barillot E, Bourc'his D. Tools and best practices for
1299 retrotransposon analysis using high-throughput sequencing data. *Mobile DNA*. 2019;10(1):52.
1300 doi:10.1186/s13100-019-0192-1
1301 42. Liu N, Lee CH, Swigut T, et al. Selective silencing of euchromatic L1s revealed by
1302 genome-wide screens for L1 regulators. *Nature*. 2018;553(7687):228-232.
1303 doi:10.1038/nature25179
1304 43. Rekaik H, Blaudin de Thé FX, Prochiantz A, Fuchs J, Joshi RL. Dissecting the role of

1305 Engrailed in adult dopaminergic neurons–Insights into Parkinson disease pathogenesis. *FEBS*
1306 *Letters*. 2015;589(24 Pt A):3786-3794. doi:10.1016/j.febslet.2015.10.002
1307 44. Maeda R, Tachibana M. HP1 maintains protein stability of H3K9 methyltransferases
1308 and demethylases. *EMBO reports*. 2022;23(4):e53581. doi:10.15252/embr.202153581
1309 45. Suzuki J, Yamaguchi K, Kajikawa M, et al. Genetic Evidence That the Non-
1310 Homologous End-Joining Repair Pathway Is Involved in LINE Retrotransposition. *PLoS*
1311 *Genetics*. 2009;5(4):e1000461. doi:10.1371/journal.pgen.1000461.g005
1312 46. Goerner-Potvin P, Bourque G. Computational tools to unmask transposable elements.
1313 *Nature Reviews Genetics*. Published online September 2018:1-17. doi:10.1038/s41576-018-
1314 0050-x
1315 47. Schwarz R, Koch P, Wilbrandt J, Hoffmann S. Locus-specific expression analysis of
1316 transposable elements. *Briefings in Bioinformatics*. 2022;23(1):bbab417.
1317 doi:10.1093/bib/bbab417
1318 48. Faulkner GJ. Elevated L1 expression in ataxia telangiectasia likely explained by an
1319 RNA-seq batch effect. *Neuron*. 2023;111(5):610-611. doi:10.1016/j.neuron.2023.02.007
1320 49. Lanciano S, Cristofari G. Measuring and interpreting transposable element expression.
1321 *Nature Reviews Genetics*. 2020;509:7. doi:10.1038/s41576-020-0251-y
1322 50. Martin SL, Li J, Weisz JA. Deletion analysis defines distinct functional domains for
1323 protein-protein and nucleic acid interactions in the ORF1 protein of mouse LINE-1. *J Mol*
1324 *Biol*. 2000;304(1):11-20. doi:10.1006/jmbi.2000.4182
1325 51. Briggs EM, McKerrow W, Mita P, Boeke JD, Logan SK, Fenyo D. RIP-seq reveals
1326 LINE-1 ORF1p association with p-body enriched mRNAs. Published online February 2021:1-
1327 13. doi:10.1186/s13100-021-00233-3
1328 52. Kulpa DA, Moran JV. Ribonucleoprotein particle formation is necessary but not
1329 sufficient for LINE-1 retrotransposition. *Human Molecular Genetics*. 2005;14(21):3237-3248.
1330 doi:10.1093/hmg/ddi354
1331 53. Taylor MS, Altukhov I, Molloy KR, et al. Dissection of affinity captured LINE-1
1332 macromolecular complexes. *eLife*. 2018;7:210-241. doi:10.7554/eLife.30094
1333 54. Kelly MP, Adamowicz W, Bove S, et al. Select 3',5'-cyclic nucleotide
1334 phosphodiesterases exhibit altered expression in the aged rodent brain. *Cell Signal*.
1335 2014;26(2):383-397. doi:10.1016/j.cellsig.2013.10.007
1336 55. Lakics V, Karran EH, Boess FG. Quantitative comparison of phosphodiesterase
1337 mRNA distribution in human brain and peripheral tissues. *Neuropharmacology*.
1338 2010;59(6):367-374. doi:10.1016/j.neuropharm.2010.05.004
1339 56. Mandel S, Gozes I. Activity-dependent neuroprotective protein constitutes a novel
1340 element in the SWI/SNF chromatin remodeling complex. *J Biol Chem*. 2007;282(47):34448-
1341 34456. doi:10.1074/jbc.M704756200
1342 57. Singh A, Modak SB, Chaturvedi MM, Purohit JS. SWI/SNF Chromatin Remodelers:
1343 Structural, Functional and Mechanistic Implications. *Cell Biochem Biophys*. 2023;81(2):167-
1344 187. doi:10.1007/s12013-023-01140-5
1345 58. Goodier JL, Cheung LE, Kazazian HH. Mapping the LINE1 ORF1 protein
1346 interactome reveals associated inhibitors of human retrotransposition. *Nucleic Acids*
1347 *Research*. 2013;41(15):7401-7419. doi:10.1093/nar/gkt512
1348 59. Taylor MS, LaCava J, Mita P, et al. Affinity proteomics reveals human host factors
1349 implicated in discrete stages of LINE-1 retrotransposition. *Cell*. 2013;155(5):1034-1048.
1350 doi:10.1016/j.cell.2013.10.021
1351 60. Moldovan JB, Moran JV. The Zinc-Finger Antiviral Protein ZAP Inhibits LINE and
1352 Alu Retrotransposition. *PLoS Genetics*. 2015;11(5):e1005121.
1353 doi:10.1371/journal.pgen.1005121
1354 61. Vuong LM, Pan S, Donovan PJ. Proteome Profile of Endogenous Retrotransposon-

Associated Complexes in Human Embryonic Stem Cells. *Proteomics*. 2019;19(15):e1900169. doi:10.1002/pmic.201900169

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

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

64. Liu Q, Yi D, Ding J, et al. MOV10 recruits DCP2 to decap human LINE-1 RNA by forming large cytoplasmic granules with phase separation properties. *EMBO reports*. 2023;24(9):e56512. doi:10.15252/embr.202256512

65. McKerrow W, Wang X, Mendez-Dorantes C, 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*. 2022;119(8):e2115999119. doi:10.1073/pnas.2115999119

66. Rodić N, Sharma R, Sharma R, et al. Long Interspersed Element-1 Protein Expression Is a Hallmark of Many Human Cancers. *Am J Pathol*. 2014;184(5):1280-1286. doi:10.1016/j.ajpath.2014.01.007

67. Rybacki K, Xia M, Ahsan MU, Xing J, Wang K. Assessing the Expression of Long INterspersed Elements (LINEs) via Long-Read Sequencing in Diverse Human Tissues and Cell Lines. *Genes*. 2023;14(10):1893. doi:10.3390/genes14101893

68. Osburn SC, Mesquita PHC, Neal FK, et al. 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*. 2022;323(3):C907-C919. doi:10.1152/ajpcell.00234.2022

69. Doucet-O'Hare TT, Rodić N, Sharma R, et al. LINE-1 expression and retrotransposition in Barrett's esophagus and esophageal carcinoma. *Proceedings of the National Academy of Sciences*. 2015;112(35):E4894-E4900. doi:10.1073/pnas.1502474112

70. Zhang Y, Amaral ML, Zhu C, et al. Single-cell epigenome analysis reveals age-associated decay of heterochromatin domains in excitatory neurons in the mouse brain. *Cell Res*. 2022;32(11):1008-1021. doi:10.1038/s41422-022-00719-6

71. Takahashi F, Zhang C, Hohjoh H, et al. Immune-mediated neurodegenerative trait provoked by multimodal derepression of long-interspersed nuclear element-1. *iScience*. 2022;25(5):104278. doi:10.1016/j.isci.2022.104278

72. Simon M, Van Meter M, Ablaeva J, et al. LINE1 Derepression in Aged Wild-Type and SIRT6-Deficient Mice Drives Inflammation. *Cell metabolism*. 2019;29(4):871-885.e5. doi:10.1016/j.cmet.2019.02.014

73. Lopez-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. Hallmarks of aging: An expanding universe. *Cell*. Published online January 2023:S0092867422013770. doi:10.1016/j.cell.2022.11.001

74. Mumford PW, Romero MA, Osburn SC, et al. Skeletal muscle LINE-1 retrotransposon activity is upregulated in older versus younger rats. *Am J Physiol Regul Integr Comp Physiol*. 2019;317(3):R397-R406. doi:10.1152/ajpregu.00110.2019

75. Traxler L, Lucciola R, Herdy JR, Jones JR, Mertens J, Gage FH. Neural cell state shifts and fate loss in ageing and age-related diseases. *Nat Rev Neurol*. 2023;19(7):434-443. doi:10.1038/s41582-023-00815-0

76. Frost B. Alzheimer's disease and related tauopathies: disorders of disrupted neuronal identity. *Trends in Neurosciences*. Published online August 2023:S0166223623001662. doi:10.1016/j.tins.2023.07.006

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

1405 78. Dai L, LaCava J, Taylor MS, Boeke JD. Expression and detection of LINE-1 ORF-
1406 encoded proteins. *Mobile genetic elements*. 2014;4:e29319. doi:10.4161/mge.29319
1407 79. Whongsiri P, Goering W, Lautwein T, et al. Many Different LINE-1 Retroelements
1408 Are Activated in Bladder Cancer. *International Journal of Molecular Sciences*.
1409 2020;21(24):9433-14. doi:10.3390/ijms21249433
1410 80. Lindtner S, Felber BK, Kjems J. An element in the 3' untranslated region of human
1411 LINE-1 retrotransposon mRNA binds NXF1(TAP) and can function as a nuclear export
1412 element. *RNA*. 2002;8(3):345-356. doi:10.1017/s1355838202027759
1413 81. Naufer MN, Furano AV, Williams MC. Protein-nucleic acid interactions of LINE-1
1414 ORF1p. *Seminars in Cell & Developmental Biology*. 2019;86:140-149.
1415 doi:10.1016/j.semcdb.2018.03.019
1416 82. Garneau NL, Wilusz J, Wilusz CJ. The highways and byways of mRNA decay. *Nat*
1417 *Rev Mol Cell Biol*. 2007;8(2):113-126. doi:10.1038/nrm2104
1418 83. Lykke-Andersen J. Identification of a human decapping complex associated with hUpf
1419 proteins in nonsense-mediated decay. *Mol Cell Biol*. 2002;22(23):8114-8121.
1420 doi:10.1128/MCB.22.23.8114-8121.2002
1421 84. Threlfell S, Sammut S, Menniti FS, Schmidt CJ, West AR. Inhibition of
1422 Phosphodiesterase 10A Increases the Responsiveness of Striatal Projection Neurons to
1423 Cortical Stimulation. *J Pharmacol Exp Ther*. 2009;328(3):785-795.
1424 doi:10.1124/jpet.108.146332
1425 85. Li N, Chen X, Zhu B, et al. Suppression of β-catenin/TCF transcriptional activity and
1426 colon tumor cell growth by dual inhibition of PDE5 and 10. *Oncotarget*. 2015;6(29):27403-
1427 27415. doi:10.18632/oncotarget.4741
1428 86. Pereira GC, Sanchez L, Schaughency PM, et al. Properties of LINE-1 proteins and
1429 repeat element expression in the context of amyotrophic lateral sclerosis. *Mobile DNA*.
1430 2018;9(1):35. doi:10.1186/s13100-018-0138-z
1431 87. Mita P, Wudzinska A, Sun X, et al. LINE-1 protein localization and functional
1432 dynamics during the cell cycle. *eLife*. 2018;7. doi:10.7554/eLife.30058
1433 88. facility I analysis. DAPI_NEUN_ORF1P. Published online June 21, 2023. Accessed
1434 October 20, 2023. https://github.com/orion-cirb/DAPI_NEUN_ORF1P
1435 89. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: an open-source platform for
1436 biological-image analysis. *Nature methods*. 2012;9(7):676-682. doi:10.1038/nmeth.2019
1437 90. Linkert M, Rueden CT, Allan C, et al. Metadata matters: access to image data in the
1438 real world. *J Cell Biol*. 2010;189(5):777-782. doi:10.1083/jcb.201004104
1439 91. Ollion J, Cochennec J, Loll F, Escudé C, Boudier T. TANGO: a generic tool for high-
1440 throughput 3D image analysis for studying nuclear organization. *Bioinformatics (Oxford,*
1441 *England)*. 2013;29(14):1840-1841. doi:10.1093/bioinformatics/btt276
1442 92. Stringer C, Wang T, Michaelos M, Pachitariu M. Cellpose: a generalist algorithm for
1443 cellular segmentation. *Nat Methods*. 2021;18(1):100-106. doi:10.1038/s41592-020-01018-x
1444 93. Wang Q, Ding SL, Li Y, et al. The Allen Mouse Brain Common Coordinate
1445 Framework: A 3D Reference Atlas. *Cell*. 2020;181(4):936-953.e20.
1446 doi:10.1016/j.cell.2020.04.007
1447 94. Chiaruttini N, Burri O, Haub P, Guiet R, Sordet-Dessimoz J, Seitz A. An Open-Source
1448 Whole Slide Image Registration Workflow at Cellular Precision Using Fiji, QuPath and
1449 Elastix. *Front Comput Sci*. 2022;3:780026. doi:10.3389/fcomp.2021.780026
1450 95. Bankhead P, Loughrey MB, Fernández JA, et al. QuPath: Open source software for
1451 digital pathology image analysis. *Sci Rep*. 2017;7(1):16878. doi:10.1038/s41598-017-17204-5
1452 96. McKinney W. Data Structures for Statistical Computing in Python. In: ; 2010:56-61.
1453 doi:10.25080/Majora-92bf1922-00a
1454 97. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and

1455 better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res*.
1456 2019;47(8):e47. doi:10.1093/nar/gkz114
1457 98. Valot B, Langella O, Nano E, Zivy M. MassChroQ: a versatile tool for mass
1458 spectrometry quantification. *Proteomics*. 2011;11(17):3572-3577.
1459 doi:10.1002/pmic.201100120
1460 99. Perez-Riverol Y, Bai J, Bandla C, et al. The PRIDE database resources in 2022: a hub
1461 for mass spectrometry-based proteomics evidences. *Nucleic Acids Res*. 2022;50(D1):D543-
1462 D552. doi:10.1093/nar/gkab1038
1463 100. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for
1464 integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498-
1465 2504. doi:10.1101/gr.1239303
1466

Figure 1:

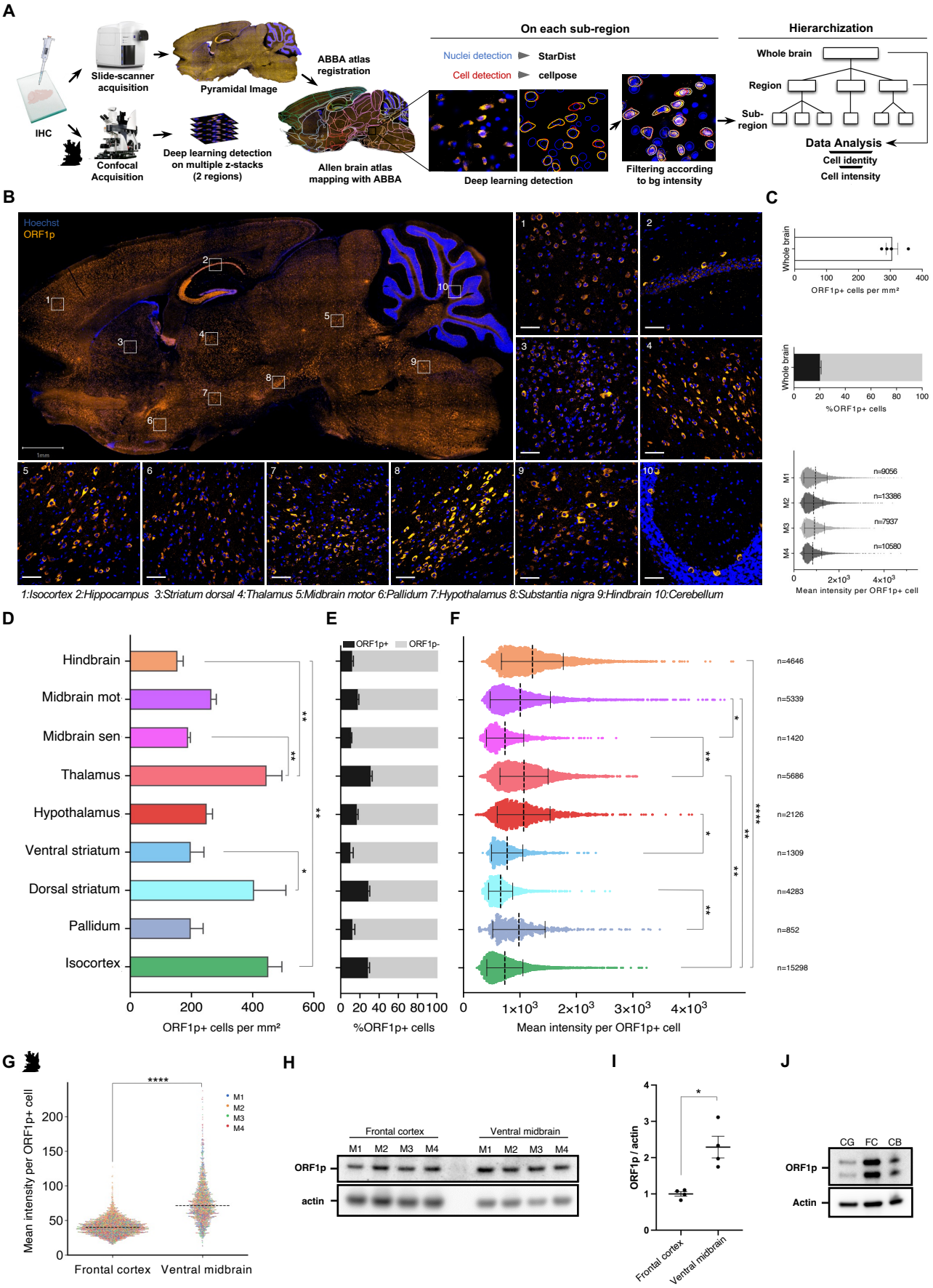

A
IHC
Slide-scanner acquisition
Pyramidal Image
ABBA atlas registration
Confocal Acquisition
Deep learning detection on multiple z-stacks (2 regions)
Allen brain atlas mapping with ABBA
On each sub-region
Nuclei detection
StarDist
Cell detection
cellpose
Deep learning detection
Filtering according to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

Figure 2:

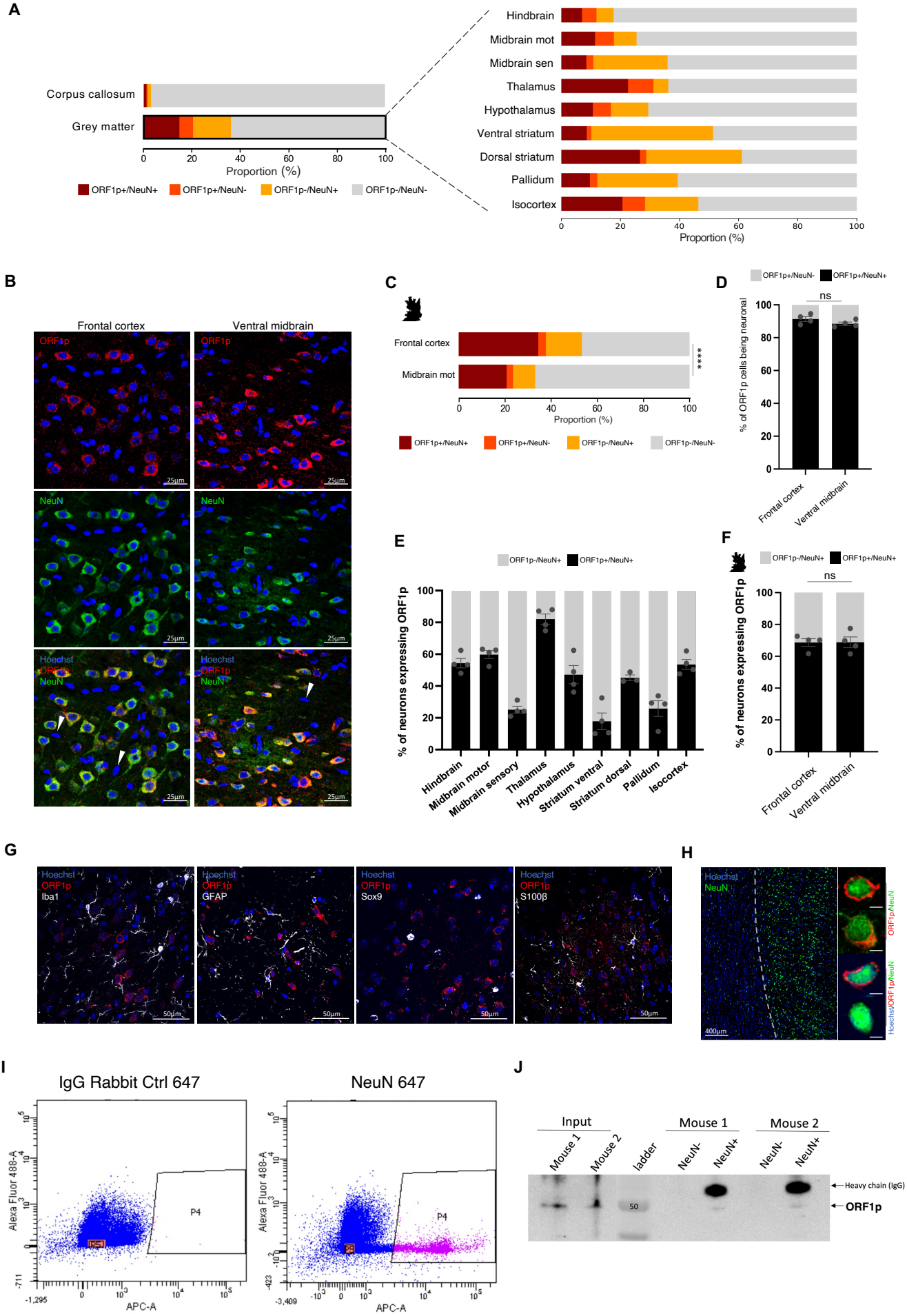

A
Corpus callosum
Grey matter
Proportion (%)
ORF1p+/NeuN+
ORF1p+/NeuN-
ORF1p-/NeuN+
ORF1p-/NeuN-
Hindbrain
Midbrain mot
Midbrain sen
Thalamus
Hypothalamus
Ventral striatum
Dorsal striatum
Pallidum
Isocortex
B
Frontal cortex
Ventral midbrain
ORF1p
NeuN
Hoechst
25µm
C
Frontal cortex
Midbrain mot
***
D
ns
% of ORF1p cells being neuronal
E
% of neurons expressing ORF1p
Hindbrain
Midbrain motor
Midbrain sensory
Thalamus
Hypothalamus
Striatum ventral
Striatum dorsal
Pallidum
Isocortex
F
G
Iba1
GFAP
Sox9
S100β
50µm
H
400µm
ORF1p/NeuN
Hoechst/ORF1p/NeuN
I
IgG Rabbit Ctrl 647
NeuN 647
Alexa Fluor 488-A
APC-A
P4
J
Input
Mouse 1
Mouse 2
ladder
NeuN-
NeuN+
Heavy chain (IgG)
ORF1p
50

Figure 3:

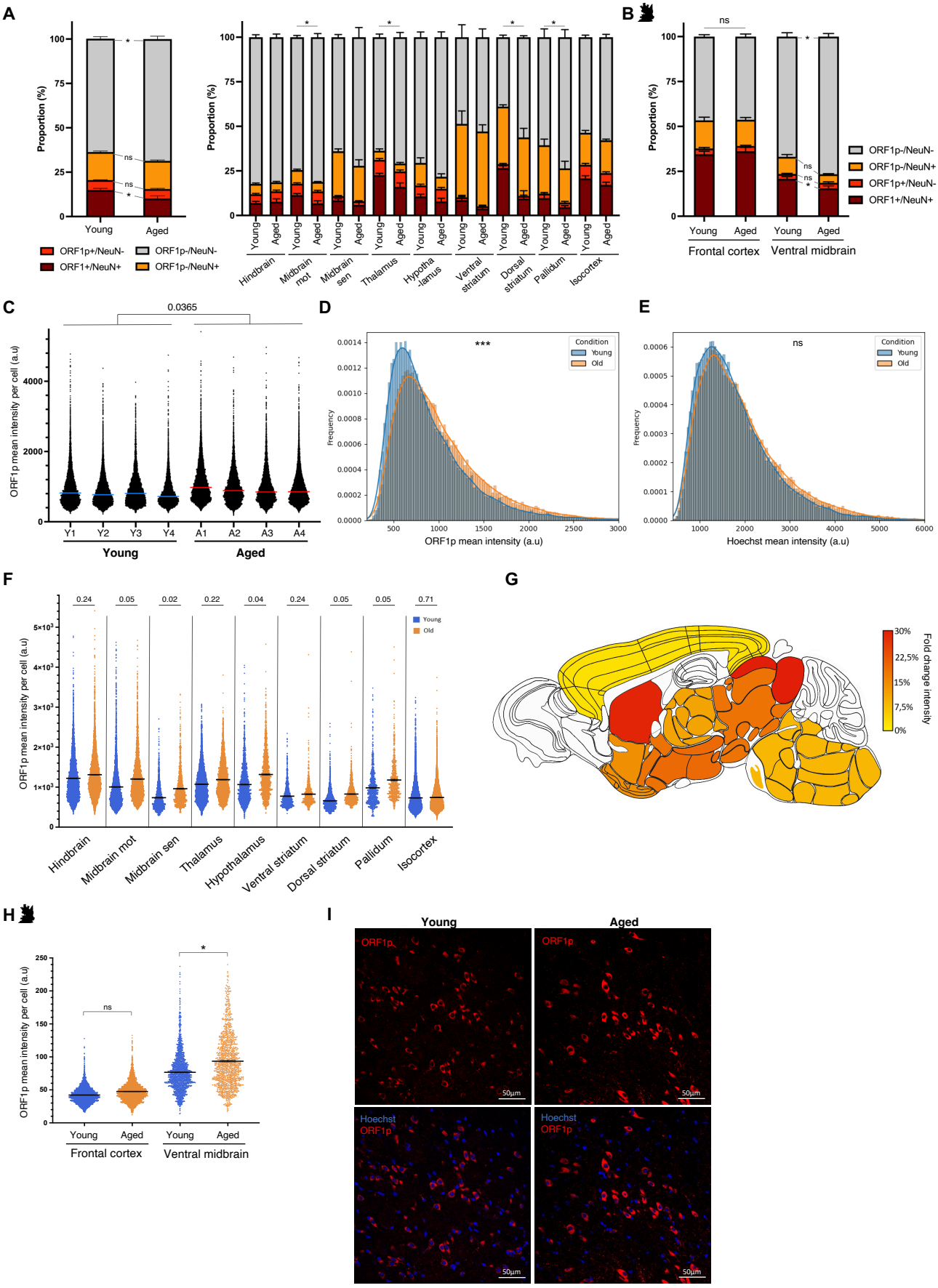

A
Proportion (%)
Young
Aged
ORF1p+/NeuN-
ORF1+/NeuN+
ORF1p-/NeuN-
ORF1p-/NeuN+
Hindbrain
Midbrain mot
Midbrain sen
Thalamus
Hypothalamus
Ventral striatum
Dorsal striatum
Pallidum
Isocortex
B
ns
Frontal cortex
Ventral midbrain
C
0.0365
ORF1p mean intensity per cell (a.u)
Y1
Y2
Y3
Y4
A1
A2
A3
A4
D
***
Condition
Young
Old
Frequency
ORF1p mean intensity (a.u)
E
ns
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
ns
I
Young
Aged
ORF1p
Hoechst
50µm

Figure 4:

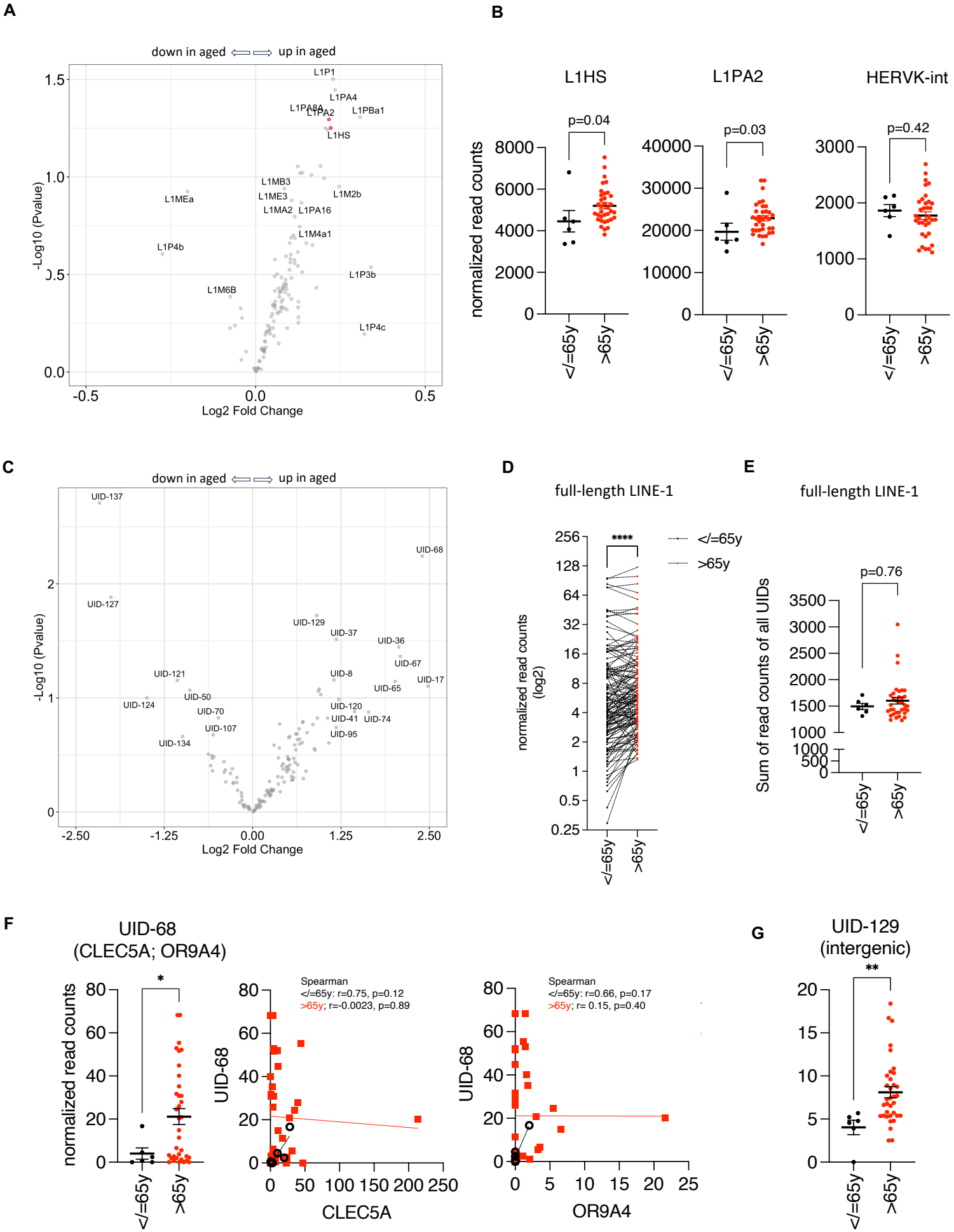

A
down in aged ⟸⟹ up in aged
-Log10 (Pvalue)
Log2 Fold Change
L1P1
L1PA4
L1PA8A
L1PA2
L1PBa1
L1HS
L1MB3
L1MEa
L1ME3
L1M2b
L1MA2
L1PA16
L1M4a1
L1P4b
L1P3b
L1M6B
L1P4c
B
L1HS
L1PA2
HERVK-int
p=0.04
p=0.03
p=0.42
normalized read counts
</=65y
>65y
C
down in aged ⟸⟹ up in aged
-Log10 (Pvalue)
Log2 Fold Change
UID-137
UID-68
UID-127
UID-129
UID-37
UID-36
UID-67
UID-121
UID-8
UID-17
UID-65
UID-50
UID-124
UID-70
UID-120
UID-41
UID-74
UID-107
UID-95
UID-134
D
full-length LINE-1
****
</=65y
>65y
normalized read counts (log2)
E
full-length LINE-1
p=0.76
Sum of read counts of all UIDs
F
UID-68
(CLEC5A; OR9A4)
*
normalized read counts
Spearman
</=65y: r=0.75, p=0.12
>65y; r=-0.0023, p=0.89
UID-68
CLEC5A
Spearman
</=65y: r=0.66, p=0.17
>65y; r= 0.15, p=0.40
OR9A4
G
UID-129
(intergenic)
**

Figure 5:

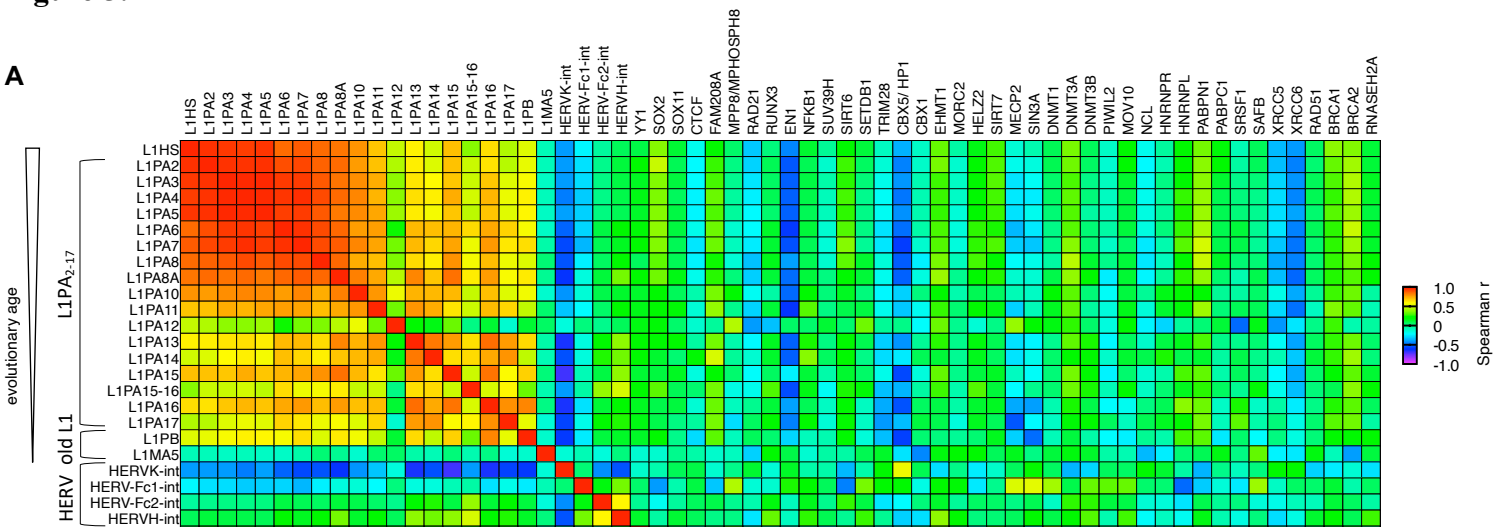
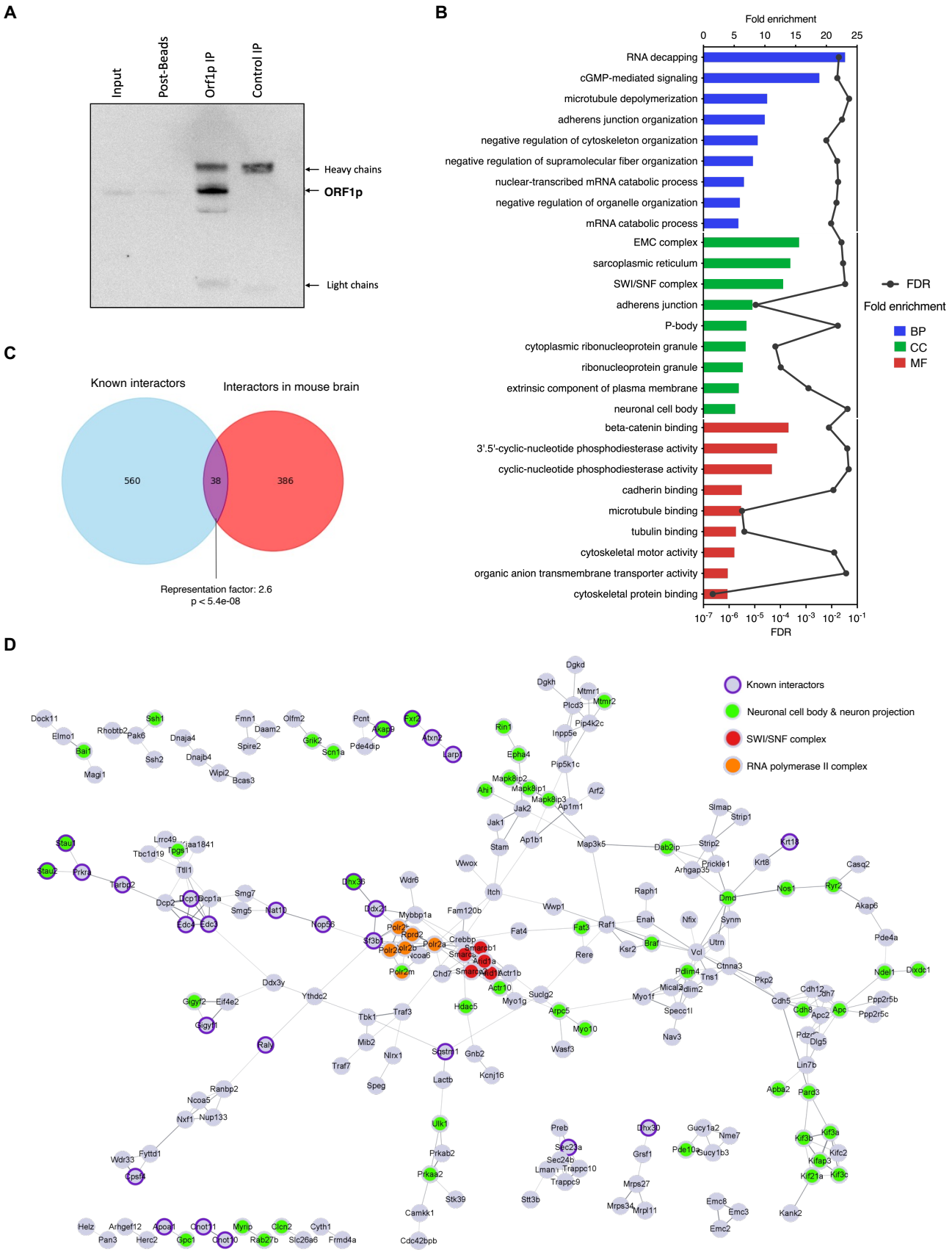
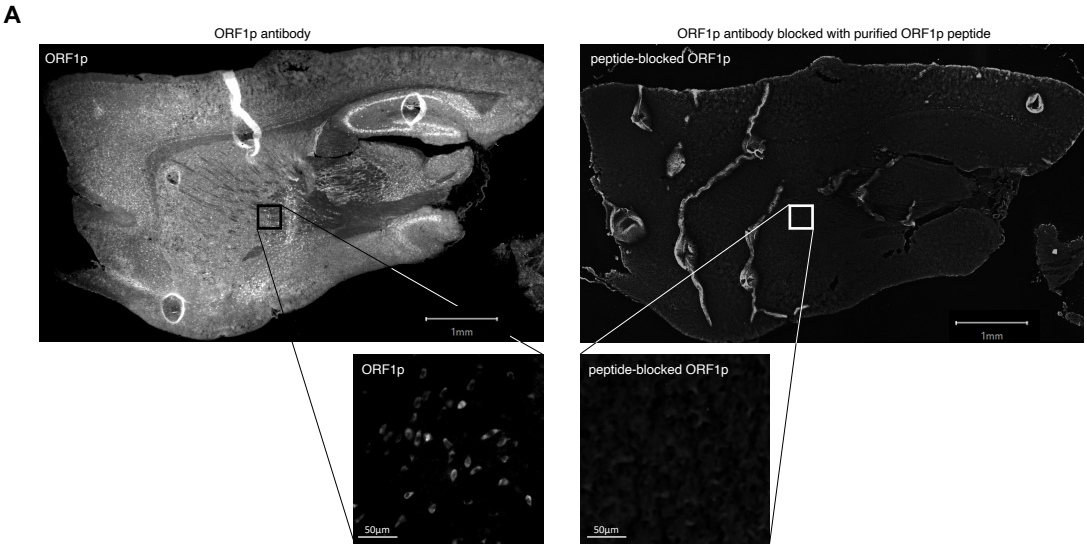
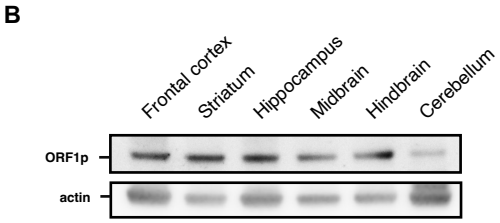

A
L1HS
L1PA2
L1PA3
L1PA4
L1PA5
L1PA6
L1PA7
L1PA8
L1PA8A
L1PA10
L1PA11
L1PA12
L1PA13
L1PA14
L1PA15
L1PA15-16
L1PA16
L1PA17
L1PB
L1MA5
HERVK-int
HERV-Fc1-int
HERV-Fc2-int
HERVH-int
YY1
SOX2
SOX11
CTCF
FAM208A
MPP8/MPHOSPH8
RAD21
RUNX3
EN1
NFKB1
SUV39H
SIRT6
SETDB1
TRIM28
CBX5/HP1
CBX1
EHMT1
MORC2
HELZ2
SIRT7
MECP2
SIN3A
DNMT1
DNMT3A
DNMT3B
PIWIL2
MOV10
NCL
HNRNPR
HNRNPL
PABPN1
PABPC1
SRSF1
SAFB
XRCC5
XRCC6
RAD51
BRCA1
BRCA2
RNASEH2A
evolutionary age
L1PA2-17
old L1
HERV
1.0
0.5
0
-0.5
-1.0
Spearman r

Figure 6:

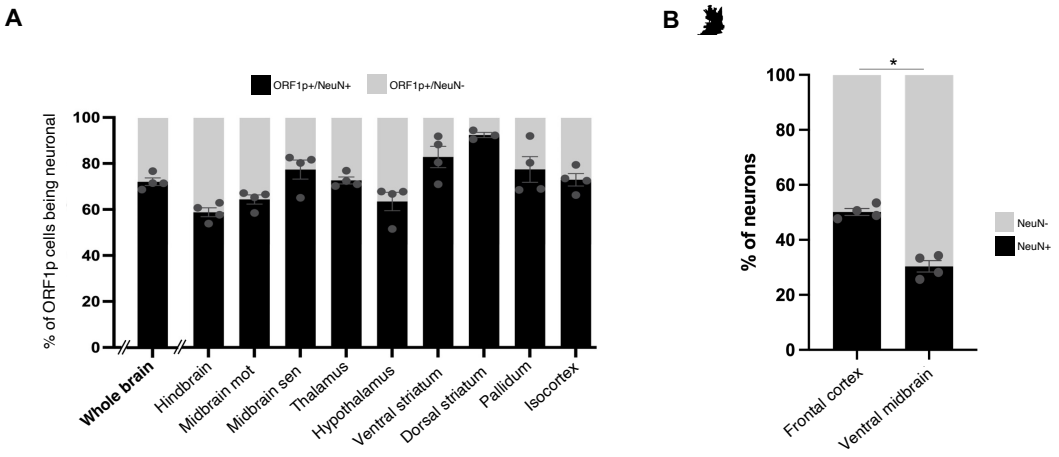

A
Input
Post-Beads
Orf1p IP
Control IP
Heavy chains
ORF1p
Light chains
B
Fold enrichment
0
5
10
15
20
25
RNA decapping
cGMP-mediated signaling
microtubule depolymerization
adherens junction organization
negative regulation of cytoskeleton organization
negative regulation of supramolecular fiber organization
nuclear-transcribed mRNA catabolic process
negative regulation of organelle organization
mRNA catabolic process
EMC complex
sarcoplasmic reticulum
SWI/SNF complex
adherens junction
P-body
cytoplasmic ribonucleoprotein granule
ribonucleoprotein granule
extrinsic component of plasma membrane
neuronal cell body
beta-catenin binding
3'.5'-cyclic-nucleotide phosphodiesterase activity
cyclic-nucleotide phosphodiesterase activity
cadherin binding
microtubule binding
tubulin binding
cytoskeletal motor activity
organic anion transmembrane transporter activity
cytoskeletal protein binding
FDR
Fold enrichment
BP
CC
MF
10^{-7} 10^{-6} 10^{-5} 10^{-4} 10^{-3} 10^{-2} 10^{-1}
FDR
C
Known interactors
Interactors in mouse brain
560
38
386
Representation factor: 2.6
p < 5.4e-08
D
Known interactors
Neuronal cell body & neuron projection
SWI/SNF complex
RNA polymerase II complex

Suppl Figure 1:

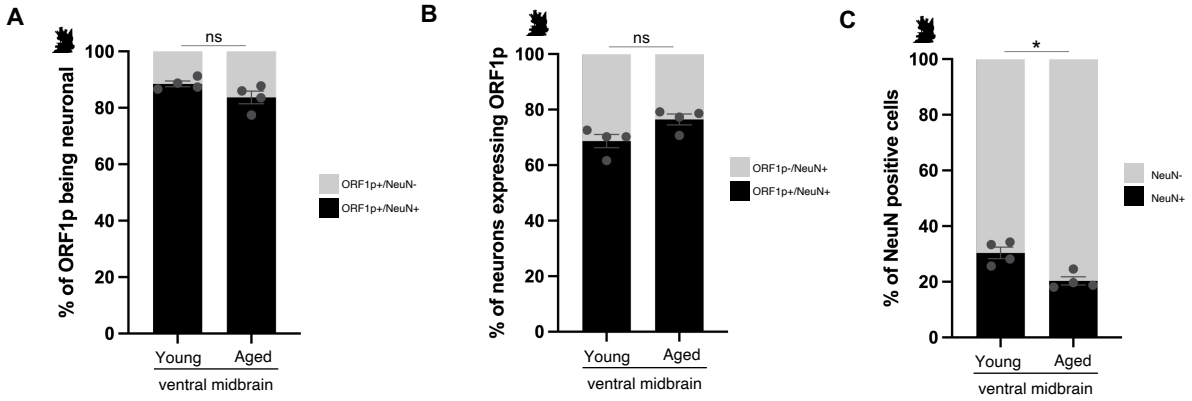




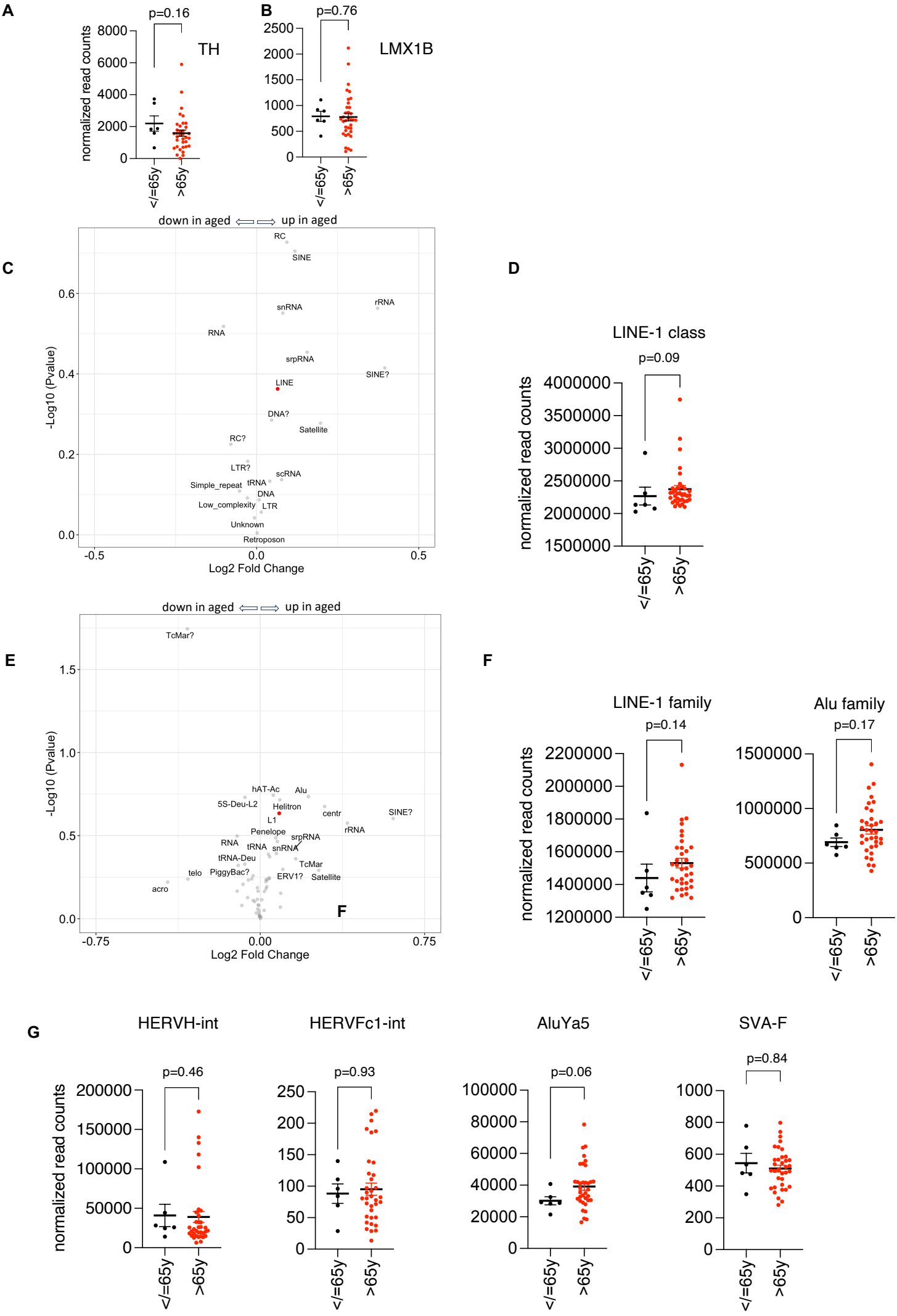
Suppl Figure 2:



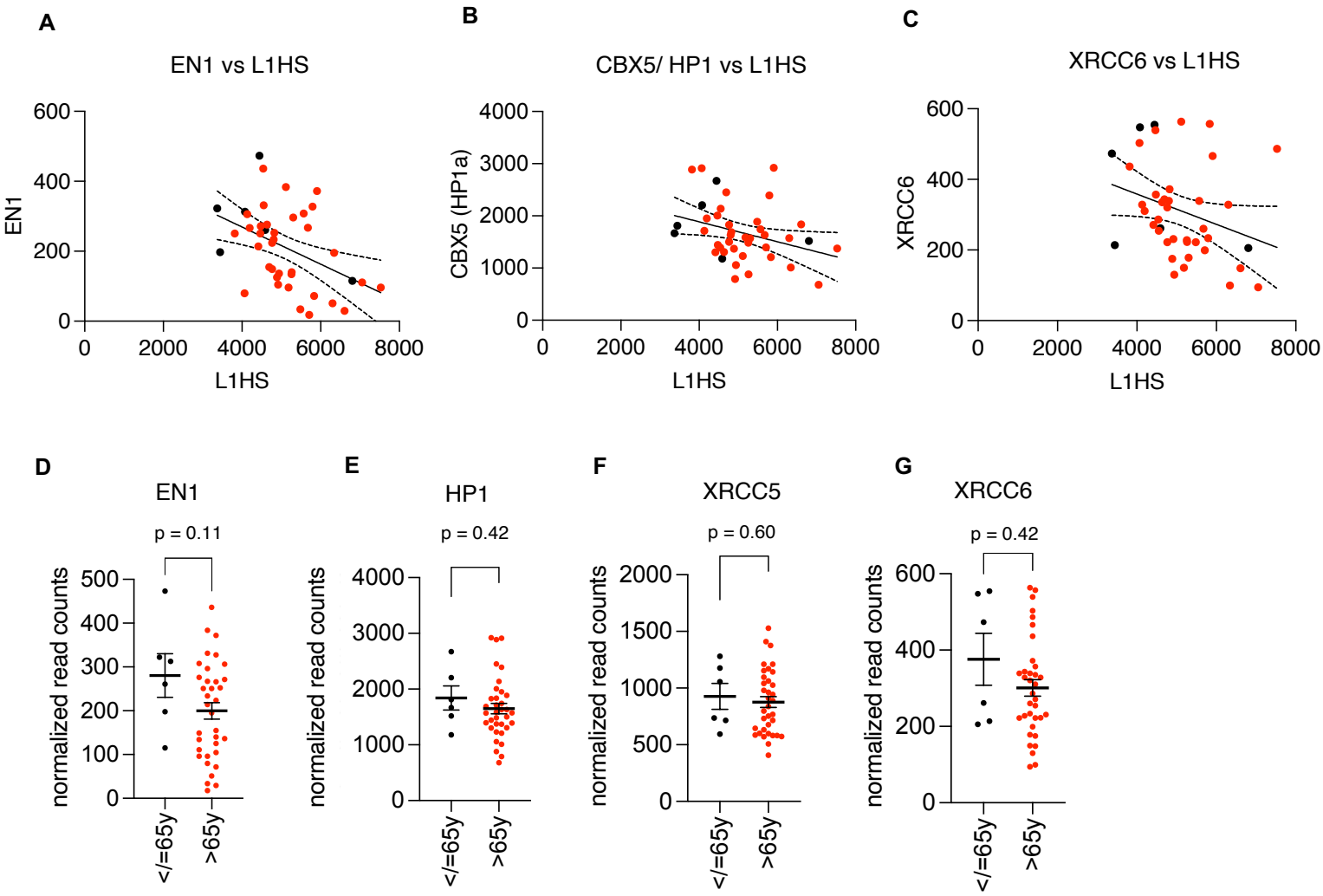
Suppl Figure 3:



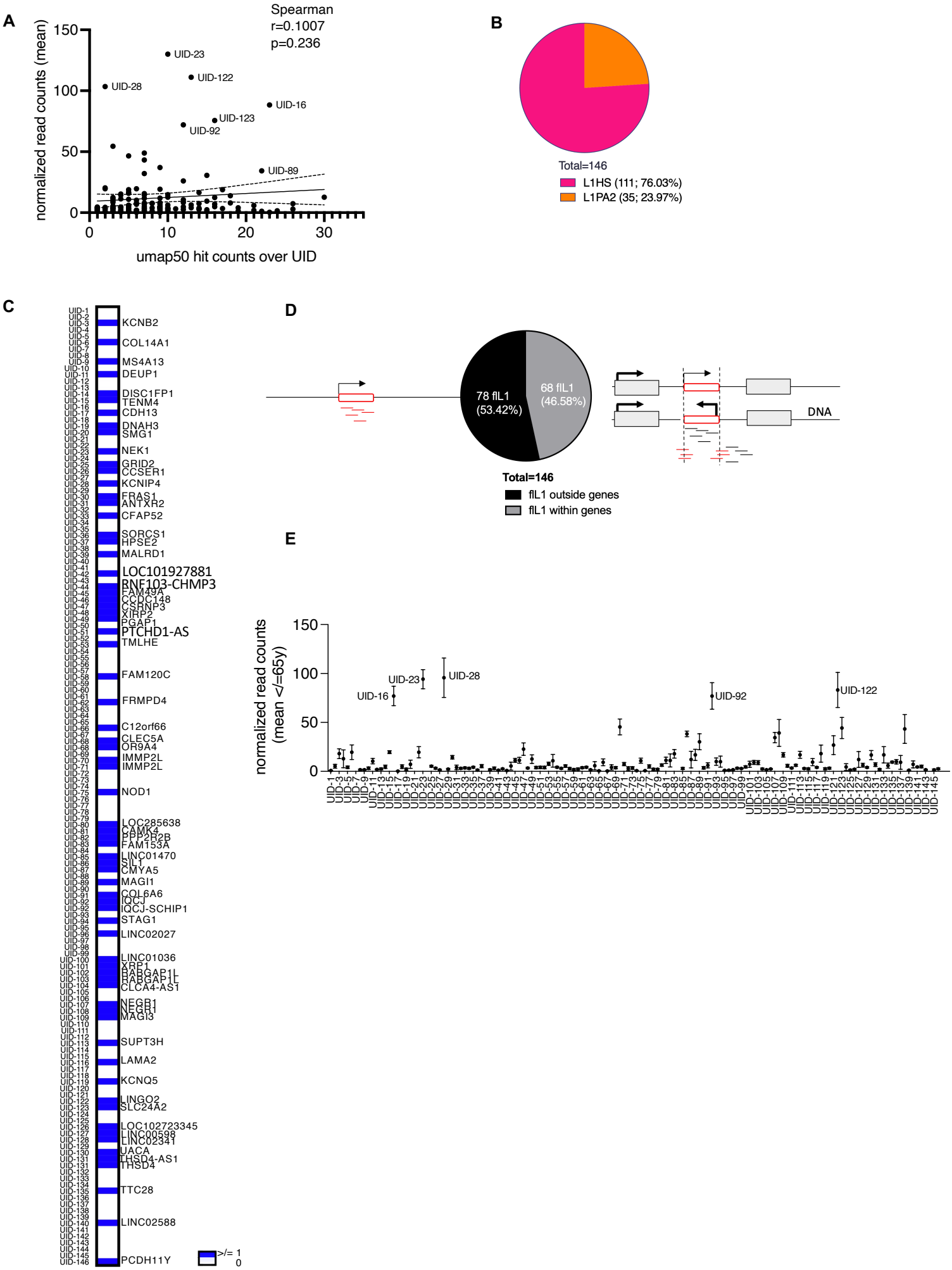
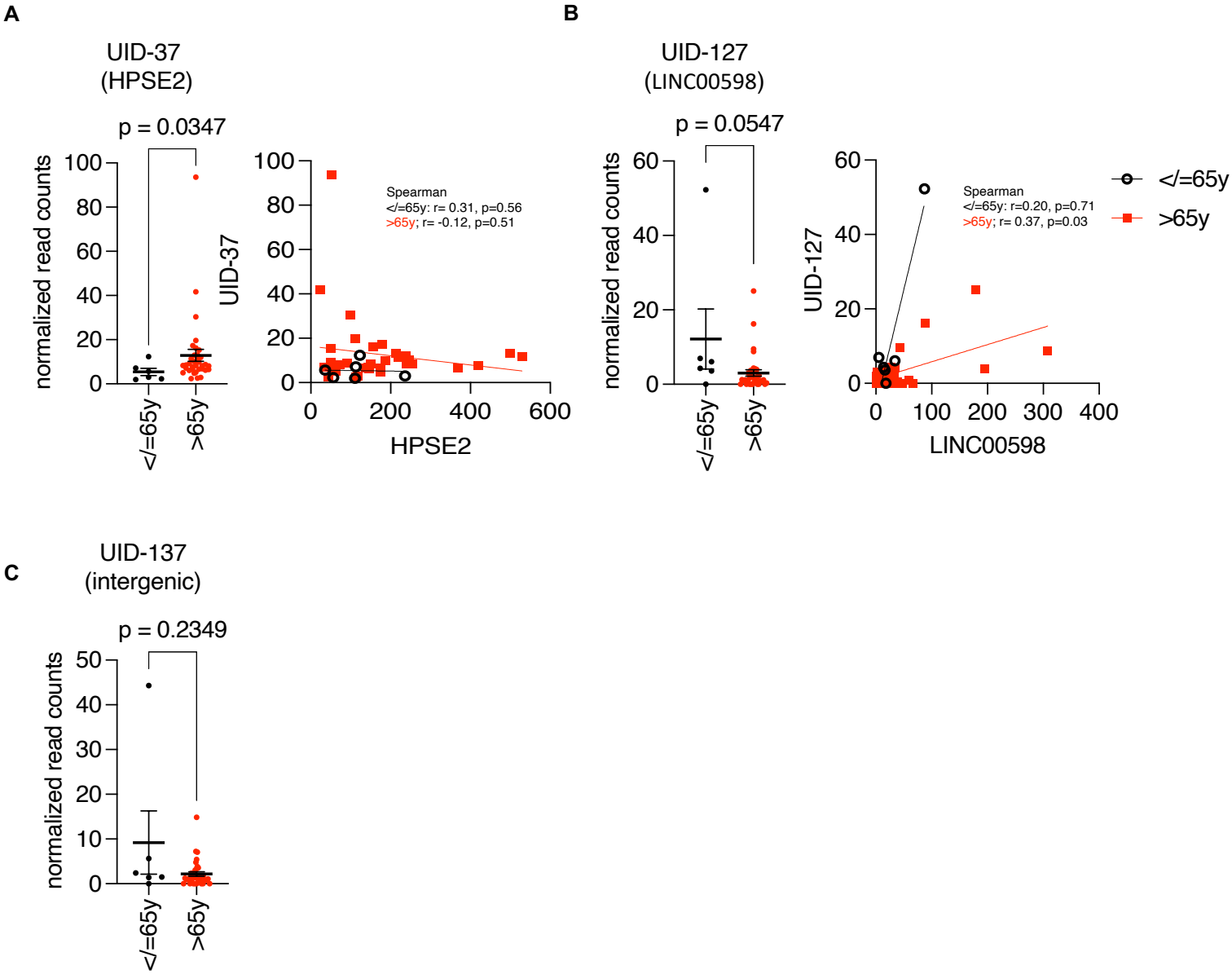
Suppl Figure 4:



Suppl Figure 5:



Suppl Figure 6:



Suppl Figure 7: