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 L. Joshi[1]* & Julia Fuchs[1]*

Affiliations:

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

[2] Orion Technological Core, CIRB, Collège de France, Université PSL, INSERM, CNRS, 75005, Paris, France.

[3] Institut Curie, Université PSL, 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, Université PSL

***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% [2] of the human genome as remnants of an evolutionary ancient activity. Depending on the source, about 100 [3] to 146 [4] full-length LINE-1 elements with two open reading frames encoding ORF1 and ORF2 are present in the Human reference genome (GRCh38 Genome Assembly) and several incomplete HERV sequences encoding either or any combination of envelope (env), gag, pro or pol [5] . The LINE-1 encoded protein ORF1p, an RNA binding protein with "cis" preference [6,7], and ORF2p, an endonuclease and reverse transcriptase [8,9] are required for the mobility of LINE-1 elements. As many other transposable elements (TEs) including HERVs, LINE-1 elements are repressed by multiple cellular pathways. It was thus generally thought that TEs are repressed in somatic cells with no expression at steady-state [10–12]. However, the aging process reduces the reliability of these repressive mechanisms [13]. It is now, 31 years after the initial proposition of the "transposon theory of aging" by Driver and McKechnie [14], generally accepted that TE activation can be both, a cause and a consequence of aging [15,16].

Sparse data has shown that the LINE-1 encoded protein ORF1p is expressed at steady-state in the mouse ventral midbrain [17] , the mouse hippocampus [18] and in some regions of the human post-mortem brain [19] and recent data informed about the presence of full-length

74 transcripts in cancer cells, human epithelial cells and mouse hippocampal neurons [20].
75 Repression of LINE-1 might thus be incomplete and if so, it remains unclear how cells then
76 prevent cell toxicity associated with LINE-1 encoded protein activity. Indeed, LINE-1 encoded
77 proteins have been demonstrated to induce genomic instability (ORF2p endonuclease-mediated
78 [17,21–26]) and inflammation (ORF2p reverse transcriptase-mediated [27–29]) and these cellular
79 activities might be causally related to organismal aging, cancer, autoimmune and neurological
80 diseases [30]. For instance, LINE-1 activation can drive neurodegeneration of mouse
81 dopaminergic neurons [17], of drosophila neurons [31,32] and of mouse Purkinje neurons [33] which
82 can be at least partially rescued with nucleoside analogue reverse transcriptase inhibitors
83 (NRTIs) or other anti-LINE-1 strategies. NRTIs are currently being tested in several clinical
84 trials designed to target either the RT of HERVs or the RT encoded by the LINE-1 ORF2
85 protein. It is not known today, however, to which extend LINE-1 encoded proteins are
86 expressed at steady-state throughout the mouse and human brain, whether there is cell-type
87 specificity and whether activation of LINE-1 encoded proteins is associated with brain aging
88 or human neurodegeneration. Here, using a deep-learning assisted cellular detection
89 methodology applied to pyramidal large-scale images of the mouse brain mapped to the Allen
90 mouse brain atlas combined with post-mortem human brain imaging, co-IP mass spectrometry
91 and transcriptomic analysis of LINE-1 expression, we describe a brain-wide map of ORF1p
92 expression and interacting proteins at steady-state and in the context of aging. We find a
93 heterogenous but widespread expression of ORF1p in the mouse brain with predominant
94 expression in neurons. In aged mice, neuronal ORF1p expression increases brain-wide and in
95 some brain regions to up to 27%. In human dopaminergic neurons, young LINE-1 transcripts
96 and specific full-length and coding LINE-1 copies are increased in aged individuals. We further
97 describe endogenous mouse ORF1p interacting proteins revealing known interactors and

98 unexpected interacting proteins belonging to GO categories related to RNA metabolism,
99 chromatin remodeling, cytoskeleton and the synapse.
100
101 **Results**
102
103 **Widespread and heterogenous expression of the LINE-1 encoded ORF1p protein in the**
104 **wildtype mouse brain**
105 To investigate the expression pattern and intensities of endogenous LINE-1 encoded ORF1p
106 protein throughout the entire mouse brain, we devised a deep-learning assisted cellular
107 detection methodology applied to pyramidal large-scale images using a comprehensive
108 workflow complemented by an approach based on confocal imaging as schematized in Figure
109 1A. Briefly, starting from sagittal slide scanner images of the mouse brain, we defined
110 anatomical brain regions by mapping the Allen Brain Atlas onto the slide scanner images using
111 Aligning Big Brains & Atlases (ABBA). We then employed a deep-learning detection method
112 to identify all cell nuclei (Hoechst) and categorize all detected cells into neuronal cells (NeuN+)
113 or non-neuronal cells (NeuN-) and ORF1p-expressing cells (ORF1p+) or cells that do not
114 express ORF1p (ORF1p-). This workflow allowed us then to characterize the cell identity of
115 ORF1p+ cells and ORF1p intensity throughout the whole brain but also in specific anatomical
116 regions. In parallel, we completed the approach using confocal microscopy on selected
117 anatomical regions allowing for comparison with higher resolution. Importantly, the specificity
118 of the ORF1p antibody, a widely used, commercially available antibody [18,34–38], was confirmed
119 by blocking the ORF1p antibody with purified mouse ORF1p protein resulting in the complete
120 absence of immunofluorescence staining (Suppl Fig. 1A), by using an in-house antibody against
121 mouse ORF1p[17] which colocalized with the anti-ORF1p antibody used (Suppl Fig. 1B,
122 quantified in Suppl Fig. 1C), and by immunoprecipitation and mass spectrometry used in this

123 study (see below). Unexpectedly, we found a generalized and widespread expression of ORF1p
124 throughout the brain of wildtype mice (Fig. 1B; Swiss OF1 mice, three months-old; whole brain
125 except regions with particularly high cellular density (cerebellum, hippocampus, olfactory
126 bulb) which impedes nuclei detection by deep-learning. ORF1p is detectable in all regions and
127 subregions analyzed with heterogenous expression patterns (density and intensity) per
128 region/subregions. The ten regions shown in Figure 1B exemplify visible different densities of
129 ORF1p+ cells with varying levels of expression. Notably, the expression pattern of ORF1p in
130 the hippocampus is similar to what has recently been published [18] (Fig. 1B, panel 2).
131 Throughout the entire brain, the mean density of ORF1p+ cells per mm^2 was ≈ 305 ±18 (mean
132 ±SEM), representing up to 20% of all detected cells (Fig. 1C). ORF1p+ cells in each mouse
133 brain analyzed showed up to eight-fold disparities in intensity between low- and high-expressed
134 cells (Fig. 1C). We then quantified nine anatomical regions according to the Allen Brain Atlas
135 on four brains of three-month old mice (Fig. 1D) using the automated workflow (Fig. 1A) with
136 regard to cell density (Fig. 1D), cell proportions (Fig. 1E) and fluorescent intensity of ORF1p+
137 cells (Fig. 1F). This approach permitted the analysis of about 10 000 ORF1p+ cells per animal
138 highlighting the power of our large-scale analysis. Densities of ORF1p+ cells ranged from the
139 lowest density in the hindbrain with 154 ±19 cells per mm^2 (mean ±SEM) to the highest density
140 of ORF1p+ cells in the isocortex with 451 ±44 cells per mm^2 (mean ±SEM) and the thalamus
141 with 446 ±50 cells per mm^2 (mean ±SEM). The proportion of ORF1p+ cells per anatomical
142 region fluctuated between 10% ±2.1 (ventral striatum, mean ±SEM) and 31% ±1.6 (thalamus,
143 mean ±SEM). The dorsal striatum ("striatal dorsal" in the Allen Brain Atlas denomination)
144 exhibited the lowest ORF1p expression intensity (658 ±3 mean ±SEM) of all regions tested, the
145 hindbrain the highest mean intensity of ORF1p per cell (mean ±SD 1221 ±548) as illustrated in
146 Figure 1B and quantified in Figure 1D and 1F. Interestingly, cell density did not correlate with
147 expression levels. Dorsal and ventral striatum for instance displayed similar ORF1p intensities

148 per cell but exhibited significant differences in ORF1p cell density and proportion. The
149 "midbrain motor" region as defined by the Allen Brain Atlas showed an intermediate cell
150 density (mean ±SEM 265 ±16 cells per mm^2) and a rather high ORF1p expression intensity
151 (mean ±SEM 1006 ±533). Statistical analysis comparing mean density of ORF1p+cells per
152 mm^2 or mean intensity per ORF1p+ cells among regions confirmed the heterogeneity
153 concerning ORF1p expression throughout the mouse brain (Fig. 1D, 1F). Slide scanner and
154 confocal images revealed an exceptionally high ORF1p expression intensity in the ventral
155 region of the midbrain which we identified as the *Substantia nigra pars compacta (SNpc)*. This
156 region displayed an important density of ORF1p+ cells and a comparatively high level of
157 ORF1p expression as illustrated by confocal imaging (Fig. 1B, panel 8), but could not be
158 quantified independently with our brain-wide approach due to the geometrically-complex
159 anatomy of this region and its small size (subregion-level in the Allen Brain Atlas hierarchy).
160 Another region which could not be included in our brain-wide analysis was the cerebellum due
161 to its extremely high density of cell nuclei. However, slide scanner and confocal imaging (Fig.
162 1B, panel 10) revealed that ORF1p is expressed in Purkinje cells, while not detectable in the
163 molecular or granular layers.

164 In order to confirm ORF1p expression by an independent method, we performed
165 Western blot analysis on six micro-dissected regions from the mouse brain (Swiss/OF1 mouse,
166 three-month old). As shown in Suppl Fig. 1D, ORF1p is expressed in all six regions with
167 varying expression levels confirming the overall presence of ORF1p throughout the brain. We
168 then chose two regions with significantly divergent ORF1p expression intensities as identified
169 and quantified on pyramidal large-scale images: the frontal cortex (low) and the ventral
170 midbrain (intermediate to high). We confirmed a significant higher expression of ORF1p in the
171 ventral midbrain compared to the frontal cortex using an approach based on the unbiased,
172 automated quantification of multiple z-stacks using a confocal microscope (Fig. 1G) and by

173 Western-blotting on micro-dissected regions (Fig. 1H,I). In concordance with the findings
174 stemming from the large-scale image quantification pipeline (Fig. 1F), the ventral midbrain
175 showed ≈ 2-times higher expression of ORF1p than the frontal cortex as quantified in Figure
176 1G (1.8-fold) and Figure I (2.3-fold) validating our cellular detection methodology for
177 pyramidal large-scale imaging and underscoring the heterogeneity of ORF1p expression levels
178 in the mouse brain. To investigate intra-individual expression patterns of ORF1p in the post-
179 mortem human brain, we analyzed three brain regions of a neurologically-healthy individual
180 (Fig. 1J) by Western blotting. ORF1p was expressed at different levels in the cingulate gyrus,
181 the frontal cortex and the cerebellum underscoring a widespread expression of human ORF1p
182 across the human brain. In summary, our findings reveal the consistent presence of ORF1p
183 expression throughout the mouse brain in all anatomical regions analyzed with high regional
184 variability in terms of density of ORF1p+ cells per mm2 per region and ORF1p+ cell intensity
185 per region. This finding raises several questions concerning cell-type identity of ORF1p
186 expressing cells and potential functions or consequences of ORF1p expression in the mouse
187 and human brain at steady-state.
188
189 **ORF1p is predominantly expressed in neurons**
190 Following our observation of a wide-spread expression of endogenous ORF1p throughout the
191 brain, we first addressed the question of the cellular identity of ORF1p+ cells. To this end, we
192 used the neuron-specific marker NeuN, commonly used to identify post-mitotic neurons in the
193 central nervous system [39]. This allowed us to determine the proportion of neuronal (NeuN+) or
194 non-neuronal cells (NeuN-) expressing ORF1p (ORF1p+) or not (ORF1p-). Making use of our
195 large-scale imaging approach (Fig. 1A), we observed drastic dissimilarities in detected cellular
196 proportions between the white and grey brain matter. As expected, we observed only 1% of
197 NeuN+ cells in the white matter (corpus callosum; Fig. 2A) validating both, the neuronal marker

198 NeuN as such and the ABBA superposition of the Allen Brain Atlas onto the sagittal brain
199 slices. In the grey matter, our approach detected 30.5% NeuN+ cells (dark red and yellow bars
200 in Fig. 2A) which, according to the literature, should include all post-mitotic neurons with only
201 minor exceptions [39–42] and corresponds to the reported proportion of neurons present in the
202 mouse brain [43]. The nine identified grey matter regions in Fig. 2A display the proportions of
203 the different cell types per region. The proportion of all cells in a given region which are positive
204 for ORF1p (dark red bars) differed between regions (lowest proportion: hindbrain: 7%; highest
205 proportion dorsal striatum: 26.6%). In the isocortex and the midbrain motor-related regions, the
206 majority of neurons detected express ORF1p (54% and 59% by large-scale analysis, Suppl Fig.
207 2A; 68.7% and 68.8% by confocal imaging (Fig. 2B, quantified in C and Suppl Fig. 2B),
208 respectively), while in the midbrain sensory related regions the proportion dropped to 25%
209 whereas it reached 82% in the thalamus (Suppl Fig. 2A). Altogether, nearly half of all NeuN+
210 cells throughout the mouse brain expressed ORF1p (mean of all regions: 48.2%; Suppl Fig.
211 2A). Regarding the cell identity of ORF1p+ cells brain-wide, more than 70% were identified as
212 neuronal by the large-scale approach (Suppl Fig. 2C). This contrasted somewhat with results
213 obtained by the second approach using confocal imaging on multiple z-stacks which indicated
214 that 91.3% (frontal cortex) and 88.5% (ventral midbrain) of ORF1p+ cells were neuronal (Fig.
215 2D). This difference in percentages of ORF1p+ expressing neurons among all neurons between
216 the large-scale image cell detection methodology and the confocal workflow is most probably
217 due to technical limitations inherent to our large-scale pipeline. Indeed, with the latter approach,
218 region-dependent differences in cell density and signal intensity levels might be the cause for
219 an underestimation of the proportion of ORF1p+ cells being neuronal due to difficulties in cell
220 detection by StarDist/Cellpose (high cell density) on a single focal plan, technical difficulties
221 which are widely reduced by the multiple z-stack based approach when using a confocal
222 microscope. Notably, frontal cortex and ventral midbrain present similar proportion of neurons

223 expressing ORF1p and ORF1p although the percentage of NeuN+ cells between these two
224 regions is significantly different (Suppl Fig. 2D). As we could not rule out that ORF1p might
225 also be expressed in non-neuronal cells, we turned to non-neuronal markers specific for
226 different glial cell populations using two different astrocytic markers (GFAP, Sox9), the astro-
227 and oligodendrocytic marker S100β and the microglial marker Iba1[43,44] and performed co-
228 staining with ORF1p followed by confocal imaging as illustrated in Figure 2E. We screened
229 multiple images of frontal cortex, ventral midbrain, hippocampus and striatum and did not find
230 a single ORF1p+ cell, which could unambiguously be defined as non-neuronal. This indicated
231 that ORF1p is not or only very rarely expressed in non-neuronal cells. To further confirm the
232 predominant presence of expression of ORF1p in neurons and the absence of ORF1p expression
233 in non-neuronal cells, we used fluorescence-activated cell sorting (FACS) to isolate neurons
234 (using a NeuN antibody) and non-neuronal cells (NeuN-) from the adult mouse brain followed
235 by Western blotting with an antibody against ORF1p (Fig. 2F, 2G). We detected ORF1p
236 exclusively in the neuronal population, confirming the results based on two different imaging
237 approaches. Finally, to assess whether predominant, if not exclusive ORF1p expression in
238 neurons is mouse brain specific or a pattern also applicable to the human brain, we investigated
239 the identity of ORF1p expressing cells in the cingulate gyrus of a healthy human brain. Similar
240 to what we found in the mouse brain, we observed sparse NeuN expression in the white (Suppl
241 Fig. 2E) and extensive NeuN staining in the grey matter corresponding to the cortical layers
242 (right, separated by a dashed white line) with ORF1p+ cells predominantly located in the grey
243 matter (images shown in Figure 2H located in the grey matter). All cells stained by ORF1p
244 were co-stained with NeuN indicating that ORF1p was expressed in neuronal cells in the human
245 brain (Fig. 2H). However, due to the lower signal quality inherent to human post-mortem
246 sections compared to mouse sections, the identity of ORF1p+ cells was estimated to be 80%
247 neuronal by the automated image analysis pipeline, although no ORF1p+ / NeuN- cells could

248 be clearly identified (Fig. 2I). Of all neurons identified, 37.2% were ORF1p+ (Fig. 2J),
249 indicating that, similar to the mouse brain, only a fraction of neurons express ORF1p (Fig. 2H,
250 right).
251 In summary, ORF1p expression in the mouse and human brain is widely restricted to neurons
252 of which a proportion express ORF1p. This raises the question of the function and consequences
253 of ORF1p expression specifically in neurons but also on the dynamic regulation of this
254 expression upon exogenous (exposome) or endogenous (aging) challenges.
255
256 **ORF1p expression is increased in the aged mouse brain**
257 ORF1p is expressed at steady-state throughout the brain, but if this expression is dynamically
258 regulated is not known. Aging has unequivocally been linked to LINE-1 regulation [16,45] both as
259 a trigger and as a consequence of LINE-1 activation but whether this is true for the brain has
260 not been thoroughly investigated with ORF1p as a read-out. We therefore addressed the
261 question whether advanced age (16-month-old mice) was paralleled by an enhanced presence
262 of ORF1p expression in the brain compared to young, three-month old mice. In the context of
263 aging, we globally observed a reduction in the proportion of ORF1p+/NeuN+ cells using the
264 cell detection workflow applied to large-scale images described in Figure 1A, phenomenon
265 mainly driven by the midbrain motor, the dorsal striatum, the pallidum and the thalamus regions
266 (Fig. 3A, dark red bars, Suppl_Table1). The confocal approach applied to two regions, the
267 frontal cortex and the ventral midbrain (Fig. 3B) confirmed a loss of ORF1p+/NeuN+ cells in
268 the ventral midbrain with no change in cell proportions in the frontal cortex in accordance with
269 the large-scale imaging approach (Fig. 3A). The predominantly neuronal identity of ORF1p+
270 cells, however, was unchanged (Suppl Fig. 3A) just as the proportion of neurons expressing
271 ORF1p (Suppl Fig. 3B). We observed a significant decrease of NeuN+ cells in the aged ventral
272 midbrain (Fig. 3B, Suppl Fig. 3C). We next analyzed ORF1p expression levels in the context

273 of brain aging. Interestingly, the mean intensity of ORF1p expression increased significantly
274 with age throughout the brain (13% increase brain-wide; Fig. 3C). Frequency distribution
275 analysis unveiled a shift in ORF1p mean expression per cell in aged mice (Fig. 3D).
276 Importantly, the Hoechst mean intensity within nuclei of ORF1p+ cells, serving as an internal
277 control, showed no significant change (Fig. 3E). Among nine analyzed regions, five
278 demonstrated a general increase in ORF1p mean intensity per cell in aged mice ($p \leq 0.05$), a
279 change independent from inter-individual variations in both young and aged mice (Fig. 3F). An
280 increase of ORF1p expression was also observed in three others regions albeit not significant.
281 The only exception was the isocortex which remained unchanged with aging. The general
282 increase of ORF1p expression (fold change intensity) in the whole brain, reaching nearly a 30%
283 increase in some regions, is represented on the heatmap in Figure 1G. These results were
284 confirmed by the confocal imaging approach: ORF1p expression in the frontal cortex remained
285 unchanged and the ventral midbrain region increased significantly in aged mice as quantified
286 in Figure 3H and shown with a representative image in Figure 3I. Overall, these results
287 highlight an age-dependent increase in ORF1p expression in neurons throughout the brain with
288 some regions showing an increase of up to 27 % in ORF1p intensities.
289

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

291 Following the observation of increased ORF1p expression in the aged mouse brain, among
292 which the ventral midbrain, and given the age-related susceptibility of dopaminergic neurons
293 in the *SNpc* to cell death and to degeneration in PD [46], we turned to a RNA-seq dataset of laser-
294 captured micro-dissected post-mortem human dopaminergic neurons of brain-healthy
295 individuals [47], in order to interrogate full-length LINE-1 mRNA expression profiles as a
296 function of age. To avoid read-length bias to which TE analysis is particularly sensitive, we
297 analyzed only the data derived from 50bp paired-end reads of linearly amplified total RNA as

298 this dataset represented all age categories (n=41; with ages ranging from 38 to 97; mean age:
299 79.88 (SD ±12.07); n=6 ≤65y; n=35 >65y; mean PMI: 7.07 (SD ±7.84), mean RIN: 7.09
300 (±0.94)). As age-related dysregulation of TEs might not be linear, we considered individuals
301 with ages-at-death younger or equal to 65 years as "young" (n=6, 38-65 years, mean age 57.5
302 years (SD ±9.9)) and individuals older than 65 years as "aged" (n=35, 65-97 years, mean age
303 83 years (SD±7.8). The expression of the dopaminergic markers tyrosine hydroxylase (TH) and
304 LMX1B were similar in both populations indicating no apparent change of dopaminergic
305 identity of analyzed melanin-positive dopaminergic neurons (Suppl Fig. 4A). Next, we
306 compared the expression of repeat elements at the class, family and name level based on the
307 repeat masker annotation implemented in the UCSC genome browser using a commonly used
308 mapping strategy for repeats [48]. No overt dysregulation of repeat elements at either level of
309 repeat element hierarchy was observed (Suppl Fig. 4C-F), however there was a significant
310 increase in several younger LINE-1 elements including L1HS and L1PA2 at the "name" level
311 (Fig. 4A,B). This was not observed for HERVK-int, a human endogenous retrovirus family
312 with some copies having retained coding potential (Fig. 4B) or other potentially active TEs like
313 HERVH-int, HERV-Fc1 or SVA-F with the exception of a trend for an increase in AluYa5
314 transcripts in the >65y group, a young Alu family mobilized by the LINE-1 retrotransposition
315 machinery (Suppl Fig. 4G). Interestingly, L1HS expression was highly correlated with L1PA2
316 expression and this correlation extended to almost all younger LINE-1 subfamilies weaning
317 down with evolutionary distance (Fig. 4C). This was not true for other active TEs as L1HS was
318 negatively correlated with HERVK-int expression (Fig. 4C). Several regulators of LINE-1
319 activity have been identified [17,49] and correlation of their expression with L1HS might allow to
320 infer their relevance of interaction (activation or repression) with L1HS in human dopaminergic
321 neurons. Spearman correlation analysis revealed three known repressors of LINE-1 activity;
322 EN1 (Engrailed 1 [17], Suppl Fig. 5A) with important functions for dopaminergic neuron

323 homeostasis [50], CBX5/HP1a, a heterochromatin binding protein binding to the histone mark
324 H3K9me3, thereby mediating epigenetic repression [51] (Suppl Fig. 5B) and XRCC5/6, also
325 known as Ku86/Ku70, which are essential for DNA double-stranded break repair through the
326 nonhomologous end joining (NHEJ) pathway and limit LINE-1 full-length insertions [52] (Suppl
327 Fig. 5C). The transcripts of these genes showed, although not statistically significant, a trend
328 for decreased expression in the elderly (Suppl Fig. 5D-G). Based on the increase of young
329 LINE-1 families L1HS and L1PA2 in aged human dopaminergic neurons and the finding that
330 ORF1p was increased in the aged mouse brain, we focused our attention on LINE-1 elements
331 with coding potential for ORF1 and ORF2 according to the L1Basev2 annotation which are
332 specific elements comprised in the L1HS and L1PA2 annotation at the "name" level. Most of
333 the 146 full-length and coding LINE-1 termed UIDs (= Unique Identifier) in the L1Base are
334 L1HS elements (76.03%), whereas the remaining 35 UIDs belong to the evolutionary older
335 L1PA2 family (Suppl Fig. 6B). The L1Base annotation is based on the human reference genome
336 (GRCh38) and annotates 146 human full-length (>6kB), intact LINE-1 elements (ORF1 and
337 ORF2 intact) with a unique identifier from 1 to 146 [4]. Attribution of sequencing reads to a
338 specific, individual TE copy is problematic [53] and several approaches have been proposed to
339 circumvent this problem including the mapping of unique reads [48].While several tools using
340 expectation maximization algorithms in assigning multi-mapping reads have been developed
341 and successfully tested in simulations [48,54], we used a different approach in mapping unique
342 reads to the L1Base annotation of full-length LINE-1. Specific "hot" LINE-1 loci in a given
343 cellular context have been identified [3], but usage of the L1Base annotation enabled an unbiased
344 approach albeit ignoring polymorphic LINE-1 sequences. Unique read mapping strategies for
345 repeat elements, especially young LINE-1 elements, will unavoidably underestimate LINE-1
346 locus-specific expression levels [48], but will be most accurate in assigning reads while allowing
347 the comparison of two different conditions analyzed in parallel. Assuming that expression of

348 UIDs was correlated with mappability, we plotted a mappability count of each UID (see
349 methods) against its mean normalized read count expression of the six individuals ≤65y. Non-
350 parametric Spearman correlation revealed no correlation between UID mappability and
351 expression (Suppl Fig. 6A) indicating no apparent bias between the two parameters. However,
352 individual UID dependency of mappability on expression cannot be excluded, especially for
353 high expressing UIDs like UID-16 for example (Suppl Fig. 6A,E). Expression of LINE-1 at the
354 locus-level has been attributed to artefacts not representing autonomous transcription including
355 differential high intronic read counts [55], pervasive transcription or reads attributable to passive
356 co-transcription with genes when the LINE-1 element is intronic [56]. To evaluate the latter, we
357 determined the number of intronic (46.58%) and intergenic UIDs (78/146; Suppl Fig. 6C) and
358 identified the corresponding genes for intronic UIDs (Suppl Fig. 6D). Of the 146 UIDs, 140
359 passed the threshold of >3 reads in at least 6 individuals. Differential expression of UID
360 between "young" and "aged" dopaminergic neurons revealed several significantly deregulated
361 full-length LINE-1 loci (Fig. 5A). However, while no single locus stood out, paired analysis of
362 the expression of all UIDs indicated a general increase (Fig. 5B, especially of low expressed
363 UIDs. The comparative analysis of the sum expression of UIDs per individual comparing young
364 (≤65y) with elderly human dopaminergic neurons, however, did not reach statistical
365 significance (Fig. 5C). Several specific loci were dysregulated in particular, for instance UID-
366 68 (Fig. 5A), a L1HS element located on chromosome 7 (chr7: 141920659-141926712) in
367 between two genes, CLEC5A (C-type lectin domain containing 5A) and OR9A4 (olfactory
368 receptor family 9 subfamily A member 4). To rule out any influence of "hosting" gene
369 transcription interference on measurable UID-68 expression differences (Fig. 5D, left), we
370 performed Spearman correlation which did not indicate any correlation between CLEC5A or
371 OR9A4 expression with UID-68 (Fig. 5D). Further, UID-68 had a high mappability count of
372 16 (range of all UIDs: 1-30, mean 9.0 (SD ±6.05), Suppl Fig. 6A) indicating that UID-68 might

373 be a candidate for an age-dependent gain of activity. L1HS UID-129, located on chromosome
374 15 (chr15: 54926081-54932099) is intergenic, at ≈ 2Mb distance upstream and ≈ 200Mb
375 downstream from the next genes. However, the mappability count of UID-129 was only 1,
376 indicating possible mapping biases inherent to the unique read-based mapping strategy
377 employed (Fig. 5E and Suppl Fig. 6A). Another UID increased in individuals >65y, the L1HS
378 UID-37 (chr10:98,782,942-98,788,971, minus strand), located in intron 3 of the HPSE2 gene
379 (mappability count of 13, minus strand), showed no correlation with its "hosting" gene
380 indicating potential autonomous transcription of this LINE-1 element and suggesting its
381 contribution to the increase of full-length LINE-1 transcripts (Suppl Fig. 7A). We also inspected
382 UID-127 (chr13:40,356,291-40,362,321, mappability count: 14), a L1PA2 element which
383 slightly decreased in elder individuals (Fig. 5A). We found a positive association with
384 LINC00598 (intron 6 of 6; ≤65y: r=0.2, p=0.71; >65y: r=0.37, p=0.03), a non-coding RNA
385 which hosts UID-127 in its 6th intron, indicating transcriptional co-regulation potentially
386 indicative of non-autonomous transcription of UID-127 (Suppl Fig. 7B). The decrease in the
387 expression of UID-137 (Fig. 5A) was mostly driven by one young individual with high
388 expression (Suppl Fig. 7C) and thus not reflecting an overall decrease (Suppl Fig. 7C, p=0.23).
389 In conclusion, TE expression analysis of this human dataset covering an age-span of 59 years
390 indicates an increase in the expression of young LINE-1 elements including those which have
391 coding potential in elderly dopaminergic neurons. A slight net sum increase of UID
392 transcripts/cell might be sufficient for the production of "above steady-state" levels of ORF1p
393 and ORF2p. Other TEs with coding potential, namely members of the HERV family, were not
394 increased. Further, correlation analyses suggest that L1HS expression might possibly be
395 controlled by the homeoprotein EN1, a protein specifically expressed in dopaminergic neurons
396 in the ventral midbrain [50], the heterochromatin binding protein HP1, two known regulators of
397 LINE-1, and the DNA repair proteins XRCC5/6.

398

399 **Endogenous ORF1p interactors in the mouse brain**

400 To go further in our understanding of steady-state neuronal ORF1p expression across the mouse
401 brain, we immunoprecipitated ORF1p and performed quantitative label-free LC-MS/MS to
402 identify potential protein partners of ORF1p in the mouse brain. We successfully
403 immunoprecipitated endogenous ORF1p from whole brain lysates (Fig. 6A) and identified a
404 total of 424 potential protein interactors associated with ORF1p (Suppl_Table2) in 5
405 independent experiments (n=5 mice). Using Gene Ontology (GO) analysis, we identified
406 several interacting proteins belonging to GO terms related to known functions of the ORF1p
407 protein in RNA binding, preferentially [57] but not exclusively *in cis* [58], for instance RNA
408 decapping and mRNA catabolic process, or related to the known presence of ORF1p in
409 ribonucleoprotein particles [59,60] (GO: cytoplasmic ribonucleotide granule) or the presence of
410 ORF1p in p-bodies [58] as shown in Figure 6B and listed in Suppl_Table3. Other GO terms that
411 emerged, to our knowledge not previously associated with ORF1p, were related to cGMP-
412 mediated signaling (GO: cGMP-mediated signaling and 3'-5'phosphodiesterase activity: i.e.
413 PDE4A, PDE4B, PDE4DIP) and the cytoskeleton (GO: microtubule depolymerization,
414 cytoskeleton organization, microtubule and tubulin binding, cytoskeletal motor activity and
415 protein binding). cGMP signaling is regulated by 3'-5' phosphodiesterases (PDEs) which
416 degrade 3',5'-cyclic guanosine monophosphate (cGMP) and 3',5'-cyclic adenosine
417 monophosphate (cAMP), an activity essential for cell physiology for the integration of extra-
418 and intracellular signals including neuronal excitability, synaptic transmission and
419 neuroplasticity [61,62]. Further, several ORF1p interacting proteins were constituents of the
420 mating-type switching/sucrose nonfermenting complex (SWI/SNF complex), i.e. ARID1A,
421 ARID1B, SMARCA2, SMARCB1, SMARCC2), an ATP-dependent chromatin remodeler
422 complex disrupting nucleosome/DNA contacts to facilitate DNA/chromatin accessibility by

423 shifting, removing or exchanging nucleosomes along DNA [63,64]. Finally, we also observed
424 proteins belonging to the GO term "neuronal cell body", corroborating with the neuron-specific
425 presence of ORF1p in the brain. A comparative analysis with previous mass spectrometry
426 studies [60,65–70] aimed at identifying ORF1p interacting proteins unveiled significantly more
427 common proteins than randomly expected (overrepresentation test; representation factor 2.6,
428 p< 5.4e-08; Fig. 6C), including LARP1, STAU2, ATXN2, RALY, TARBP2 or DDX21 (for a
429 full list see Suppl_Table4). The presence of a significant number of overlapping ORF1p
430 interactors in different non-neuronal human cells (HEK[60,65,66], HeLa[67], human breast and
431 ovarian tumors[70] and hESCs[68]) and mouse brain cells (our study), suggest conserved key
432 interactors between both species and between cell types, with a subset of these proteins
433 regulating RNA degradation and translation potentially relevant for the LINE-1 lifecycle itself.
434 ORF1p interactors found in mouse spermatocytes [69] were also present in our analysis including
435 CNOT10, CNOT11, PRKRA and FXR2 among others (Suppl_Table4). To unravel the physical
436 interactions between the identified interactors of endogenous ORF1p within the mouse brain,
437 we used the STRING database (Search Tool for Recurring Instances of Neighboring Genes,
438 https://string-db.org/). This analysis generated a network representation, where physical
439 interactions are represented by edges (Fig. 6D). In analogy with the GO term analysis, ORF1p
440 displayed interactions with various clusters, including well-known RNA decapping complexes
441 directed against LINE-1 RNA, which also encompassed DCP2 and DCP1A which had not
442 previously been identified as interacting with ORF1p [71]. Furthermore, ORF1p exhibited
443 interactions with the SWI/SNF complex (highlighted in red) as well as subunits of the RNA
444 polymerase II complex suggesting a direct or indirect association with accessible chromatin, a
445 hitherto unknown interaction of ORF1p with chromatin compartments within the nucleus.
446 Notably, a multitude of novel interactors belonged to the "neuronal cell body" and "neuron
447 projections" clusters, proposing potential neuron-specific partners of ORF1p such as Grm2/5,

448 Bai1, Epha4, Kcnn2, Grik2 and Dmd among others. A last cluster, formed by Ncoa5 (Nuclear
449 Receptor Coactivator 5), Nxf1 (Nuclear RNA Export Factor 1), Ranbp2 and Nup133 (both
450 nucleoporins), might imply a role for these interactions in L1-RNA nuclear export and/or a
451 mechanism for the LINE-1 RNP to gain access to the nucleus in post-mitotic neurons.
452 Altogether, the identification of known and novel interactors of ORF1p in the mouse brain
453 suggests roles of ORF1p in the LINE-1 life cycle (RNA binding and metabolism, RNP
454 formation, nuclear access) but also suggests potential novel physiological roles of ORF1p in
455 the brain related to cytoskeleton organization, cGMP signaling, neuron-specific functions (i.e.
456 synaptic signaling, Suppl_Table3) and chromatin organization and/or transcription regulation.
457
458
459 **Discussion**
460 While LINE-1 derepression in aging has been extensively explored in peripheral tissues and
461 various pathologies, including cancer, our understanding of LINE-1, particularly ORF1p
462 expression, in the central nervous system remains limited [20,72,73]. A recent search of ORF1p
463 peptides in mass spectrometry data spanning 29 different healthy tissues did not reveal the
464 presence of ORF1p in the brain, suggesting that its presence might lie below detection limits [20].
465 Only a few studies explored ORF1p encoded LINE-1 protein expression in the brain, most
466 focusing on a specific region (in mice [17,33] in rats [74] and in human post-mortem brain [19]), but it
467 remained unclear if ORF1p is expressed throughout the entire brain, exhibits cell-type
468 specificity, and most intriguingly, if its expression is influenced by the aging process. Here, we
469 demonstrate that ORF1p is expressed throughout the entire mouse brain and in at least three
470 regions of the human post-mortem brain at steady-state. Leveraging a comprehensive workflow
471 that incorporates brain atlas registration and machine learning algorithms (as described in the
472 methods section), we quantified tens of thousands of brain cells, enabling a profound analysis

473 of cell proportions, cell identities, densities and ORF1p expression levels across the entire brain.
474 Surprisingly, more than one-fifth of detected cells expressed ORF1p. Strikingly, regional
475 variations in ORF1p expression levels were observed, with each region exhibiting distinct
476 proportions, cell density, and signal intensity of ORF1p+ cells. In a non-neurologically diseased
477 human brain, ORF1p is expressed in all three regions examined, that is the cingulate gyrus, the
478 frontal cortex and the cerebellum. This is in accordance to an earlier study using histological
479 staining, which found ORF1p expression in the human frontal cortex, the hippocampus, in basal
480 ganglia, thalamus, midbrain and the spinal cord [19]. This suggests, similarly to the mouse brain,
481 a generalized expression across the human brain. On the transcriptomic level using long-read
482 sequencing of GTEx tissues, brain and liver were highlighted as the organs displaying the
483 highest expression of putatively active, full-length LINE-1 elements [75]. However, when the
484 authors looked at sub-regions, they found transcript expression in cerebellar hemispheres and
485 the putamen, but not in the caudate and the anterior cingulate gyrus and frontal cortex [75]. This
486 is in contrast to our data and the data from Sur et al, where ORF1p was found to be expressed
487 in the latter two regions using two different antibodies. We used the anti-human LINE-1 ORF1p
488 antibody clone 4H1, a well characterized antibody [73,76]. While the sample size for the staining
489 of human post-mortem tissues certainly needs to be increased in order to draw quantitative
490 conclusions, the presence of the protein in two independent studies does point to a steady-state
491 expression of ORF1p in the human brain.

492 In the mouse brain, we find ORF1p to be expressed predominantly if not exclusively in
493 neurons using immunofluorescence and fluorescence-activated cell sorting (FACS). This result
494 is consistent with previous studies, such as the identification of ORF1p in excitatory neurons
495 within the mouse frontal cortex [77], in parvalbumin neurons in the hippocampus [18], its presence
496 in neurons in the ventral midbrain [17] and the recognition of morphological similarities between
497 stained neurons and ORF1p+ cells in a post-mortem hippocampus sample of a healthy

498 individual [19]. We also detected ORF1p in Purkinje cells in the mouse and in bulk human
499 cerebellum. Neuronal specificity or preference of LINE-1 expression was also shown on the
500 transcriptomic level in recent studies investigating LINE-1 expression in the mouse
501 hippocampus, where neuronal LINE-1 expression exceeded that of astrocytes and microglia by
502 approximately twofold [20], is abundant in parvalbumin interneurons [18] and single-nuclei RNA-
503 seq data from the mouse hippocampus and frontal cortex which confirmed globally that
504 repetitive elements including LINE-1 are more active in neurons than in glial cells [77]. In the
505 human brain, LINE-1 transcripts are found in greater quantities in neurons compared to non-
506 neuronal cells by single-nucleus sequencing [78]. Furthermore, retrotransposition-competent
507 LINE-1 elements (similar to UID) are found expressed exclusively in neurons [79]. While ORF1p
508 expression is suggested to be expressed in microglia under experimental autoimmune
509 encephalomyelitis conditions in the spinal cord [80], no evidence of such expression was observed
510 in non-neuronal cells under non-pathological condition.

511 On average, throughout the mouse brain, the majority of neurons was positive for
512 ORF1p and in some regions (i.e. the thalamus) around 80% of neurons expressed ORF1p.
513 Comparing the results of both imaging approaches, the percentages of neurons expressing
514 ORF1p in the ventral midbrain and frontal cortex were roughly similar (around 70% of neurons
515 expressed ORF1p as quantified by confocal imaging and about 60% of neurons were identified
516 as ORF1p+ using the slide scanner approach). In the human cingulate gyrus, we found that
517 37.2% of neurons express ORF1p and that 80% of cells expressing ORF1p were neurons, which
518 are proportions similar to some regions of the mouse brain. It is however possible that these
519 percentages are underestimated due to technical issues inherent to the machine-learning based
520 algorithm for cell detection as our observations often indicated a positive signal in neurons
521 which were classified as negative due to a particular shape or our stringent intensity threshold.
522 A question which arises based on these findings is whether specific features distinguish

523 ORF1p+ and ORF1p- neurons. One hint comes from a recent study suggesting that in the mouse
524 hippocampus, it is the parvalbumin positive neurons that predominantly express ORF1p [18]. We
525 have made a similar observation in the mouse ventral midbrain, where TH-positive
526 dopaminergic neurons express higher levels of ORF1p compared to surrounding, non-
527 dopaminergic neurons [17] (Fig. 1B, panel 8). In the cerebellum, we observed ORF1p staining in
528 Purkinje cells but not in the surrounding granular and molecular layer neurons (Fig. 1B, panel
529 10). Parvalbumin positive neurons are inhibitory neurons, so are Purkinje cells. However,
530 dopaminergic neurons are modulatory neurons exerting excitatory and inhibitory effects
531 depending on the brain region they act on. Specific neurons in the granular layer (i.e. Golgi and
532 unipolar brush cells) of the cerebellum are inhibitory, but ORF1p negative, indicating that the
533 decisive feature might not be the excitatory or inhibitory nature of a neuron. Another possibility
534 is a cell-type specific chromatin organization permissive for the expression of LINE-1 and
535 future single-cell studies in the mouse and human brain might reveal those differences.

536 Because transposable elements are known to become active in somatic tissues during
537 aging [15,16,28,81,82], we aimed to investigate whether there was a corresponding increase at the
538 protein level. In aged mice, ORF1p expression significantly increased throughout the mouse
539 brain consistent with a previously documented increase in ORF1p outside the central nervous
540 system in aged rats [74,83] and aged mice [81]. By quantifying the mean intensity of ORF1p in over
541 70 000 cells identified as ORF1p+, we were able to characterize the extent of this increase in
542 each anatomical sub-region. Remarkably, apart from the isocortex which did not show any
543 change, ORF1p expression increased in all other brain regions by 7% to 27%, indicating a
544 generalized increase of ORF1p expression in neurons throughout the brain (13%). We did not
545 detect any change in cell identity of ORF1p expressing cells, that is, ORF1p expression
546 remained predominantly if not exclusively neuronal. Intriguingly, we observed a loss of NeuN+
547 cells, particularly those expressing ORF1p, throughout the brain which was more pronounced

548 in specific regions (ventral midbrain) than in others (isocortex). The loss of NeuN+ cells could
549 either be due to a loss of neuronal identity, as described recently in the context of neuronal
550 aging [84] and in the context of Alzheimer disease and related tauopathies [85], to a
551 neurodegenerative process *per se* or to gliosis related to the aging process [86] (as we observe a
552 slight increase of ORF1p-/NeuN- cells per mm^2).

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

564 Neurodegeneration was partially prevented by anti-LINE-1 strategies among which
565 NRTIs [17] and similar LINE-1 protein-dependent neuronal toxicity has been shown in drosophila
566 [31,32] and the mouse cerebellum [33].

567 In order to test whether an increase in LINE-1 is a feature of human brain aging, we
568 turned to a unique RNA-seq dataset of human laser-captured dopaminergic neurons of 41
569 individuals ranging from 38 to 99 years [47]. In accordance with our focus on LINE-1 sequences
570 which are full-length and coding, we developed a rationale to interrogate LINE-1 families with
571 representatives that are coding (L1HS, L1PA2, multimappers; RepeatMasker) and to
572 specifically investigate full-length LINE-1 elements that have intact open reading frames for

573 ORF1p and ORF2p (unique reads; L1Basev2 [4]). Indeed, we find an increase in L1HS and
574 L1PA2 elements in individuals ≥65y as well as an increase in specific full-length LINE-1
575 elements but only a trend for increase of all full-length LINE-1 in sum in the elderly. This
576 analysis has technical limitations inherent to transcriptomic analysis of repeat elements
577 especially as it is based on short-read sequences. Nevertheless, we tried to rule out several biases
578 by demonstrating that mappability did not correlate with expression and that the expression of
579 intronic full-length LINE-1 elements is not correlated with the expression of their "hosting"
580 gene. Interestingly, dysregulated full-length LINE-1 elements in aged dopaminergic neurons
581 did not correspond to those identified in bladder cancer [88] indicating the intricate nature of this
582 expression across tissues and pathological conditions. Overall, a slight net sum increase of UID
583 transcripts/cell might be sufficient for the production of "above steady-state" levels ORF1p
584 and ORF2p. Further, a dissociation of LINE-1 transcript and protein levels in aging has been
585 observed recently in excitatory neurons of the mouse cortex. In the absence of transcriptional
586 changes of LINE-1, protein levels of ORF1p were increased [77].

587 We can only speculate about the reason for an increase in ORF1p in the aged brain. A
588 recent single-cell epigenome analysis of the mouse brain suggested a specific decay of
589 heterochromatin in excitatory neurons of the mouse brain with age which was paralleled by an
590 increase in ORF1p, albeit equally in excitatory and inhibitory neurons, again not indicating any
591 dependency of ORF1p regulation on the excitatory or inhibitory nature of neurons [77]. Chromatin
592 and particularly heterochromatin disorganization are a primary hallmark of aging [82] but other
593 repressive cellular pathways which control the LINE-1 life cycle might also fail with aging [13].
594 Another possibility is a loss of accessibility of repressive factors to the LINE-1 promoter or an
595 age-dependent decrease in their expression. Matrix correlation analysis of several known LINE-
596 1 regulators, both positive and negative, revealed possible regulators of young LINE-1
597 sequences in human dopaminergic neurons. Despite known and most probable cell-type

598 unspecific repressive factors like the heterochromatin binding protein CBX5/HP1 [51] or the DNA
599 repair proteins XRCC5 and XRCC6 [49], we identified the homeoprotein EN1 as negatively
600 correlated with young LINE-1 elements including L1HS and L1PA2. EN1 is an essential
601 protein for mouse dopaminergic neuronal survival [50] and binds, in its properties as a
602 transcription factor, to the promoter of LINE-1 [17]. As EN1 is specifically expressed in
603 dopaminergic neurons in the ventral midbrain, our findings suggests that EN1 controls LINE-1
604 expression in human dopaminergic neurons as well and serves as an example for a neuronal
605 sub-type specific regulation of LINE-1.

606 The heterogenous, brain-wide presence of ORF1p expression at steady-state is
607 intriguing. In cancer cell lines or mouse spermatocytes, ORF1p interacts with several "host"
608 proteins, some if not most of which are related to the LINE-1 life cycle. However, a profile of
609 endogenous ORF1p interactors in the mouse brain might inform on possible other and organ-
610 specific functions besides its binding to the LINE-1 RNA in "cis" [69]. Among the total 424
611 potential interactors of endogenous ORF1p in the mouse brain, 38 partners had been previously
612 identified by mass spectrometry in human cancers, cancerous cell lines and mouse
613 spermatocytes [60,65–70] (Suppl_Table4). Further, GO term analysis contained expected categories
614 like "P-body", mRNA metabolism related categories and "ribonucleoprotein granule". We also
615 identified NXF1 as a protein partner of ORF1p, a protein found to interact with LINE-1 RNA
616 related to its nuclear export [89]. This suggests the conservation of key interactors probably
617 essential for completing or repressing the LINE-1 life cycle in both species, despite the
618 divergence of mouse and human ORF1p protein sequences [90]. Along these lines, several ORF1p
619 protein partners we identified might complete the list of post-transcriptional regulators
620 implicated in LINE-1 silencing. Recent work conducted on human cancerous cell lines has
621 demonstrated that MOV10 orchestrates the recruitment of DCP2 for LINE-1 RNA decapping
622 [71]. In our analysis, we identified DCP2 along with DCP1A, known to enhance the decapping

623 activity of DCP2 [91], and DCP1b, a pivotal component of the mRNA decapping complex [92].
624 Intriguingly, MOV10 was not detected in our mass spectrometry analysis, despite its established
625 role in recruiting DCP2 and forming a complex with L1-RNP to mediate LINE-1 RNA
626 decapping, as reported by Liu et al [71]. However, we found two enhancers of mRNA decapping,
627 EDC3 and EDC4, both core components of P-bodies, a membrane-less organelle known to
628 contain L1-RNP [58]. Multiple ubiquitin-ligase proteins were found although not appearing as a
629 significantly enriched GO term. These results complete the picture of the post-transcriptional
630 and translational control of ORF1p and suggest that these mechanisms, despite a steady-state
631 expression, are operational in neurons. Further, several neuron-specific interactors were
632 identified belonging to GO term categories "neuron projection" (75 proteins) and "neuronal cell
633 body" (5 proteins), again pointing to the neuron-predominant expression of ORF1p in the
634 mouse brain. Other interesting aspects were raised from this analysis. Among significantly
635 enriched GO terms, several were related to the cytoskeleton, the functional consequences of
636 which need to be determined in future studies. Our screen also identified PDE10A as an
637 interactor of ORF1p in the mouse brain, a PDE almost exclusively expressed in medium spiny
638 neurons of the striatum and a target for treatment of neurological diseases related to basal
639 ganglia function like Huntington's disease, schizophrenia and Tourette syndrome [93].
640 Interestingly, PDE10A inhibition is related to beta-catenin signaling, another GO term which
641 emerged from our screen [94]. Finally, we found components of RNA polymerase II and the
642 SWI/SNF complex as partners of ORF1p. This further indicates that ORF1p has access to the
643 nucleus in mouse brain neurons as described for other cells [95,96], implying that ORF1p
644 potentially has access to chromatin. These findings give rise to intriguing questions regarding
645 the potential function of ORF1p in neuron in health and disease as (i) ORF1p is widely
646 distributed throughout the brain under normal physiological conditions, (ii) ORF1p shows a
647 wide range expression levels within and inbetween regions, (iii) ORF1p is expressed

648 predominantly if not exclusively in neurons, (iv) but not in all neurons and (v) interacts with
649 proteins that might not directly relate to the LINE-1 life cycle, some of which are neuron-
650 specific. In addition, physicochemical properties of ORF1p to form compacted nucleic acid -
651 bound complexes with sequestration potential were shown [90,97]. Future loss-of-function studies
652 should help to shed light on the necessity of ORF1p for neuronal functions if they exist. Our
653 data spurs the idea of a possible "physiological" function of ORF1p as an integrative protein
654 with exapted function in neuronal homeostasis and a loss of restriction in the aged brain limiting
655 LINE-1 expression to steady-state levels.
656
657
658 **Materials & Methods**
659 **Animals**
660 Swiss OF1 wild-type mice (Janvier) were housed on a 12h light/dark cycle with free access to
661 water and food. Mice were sacrificed at 3-month or 16-month. Animal experiments were
662 performed according to the EU directive 2010/63/EU.
663
664 **Mouse tissue dissection and protein extraction**
665 Tissues were extracted from 3-month-old and 16-month-old Swiss/OF1 mice. Briefly, the two
666 hemispheres were separated in ice cold PBS -/-. For each mouse, one hemisphere was rinsed
667 and fixed in 4% PFA for 1h followed by 24h of incubation in 30% sucrose. Hemispheres were
668 kept at -20°C until being sliced on a freezing microtome (Epredia, HM 450) with a 20µm
669 thickness. The other hemisphere was dissected in ice cold PBS -/- 1X and 6 brain regions were
670 rinsed, cut in small pieces and dissociated separately using a large (21G) to small gauge (27G)
671 needle in RIPA lysis buffer for 5 min. Lysates were kept on ice for 25 min, were sonicated for
672 15 min and supernatants were collected after a 30 min centrifugation at 4°C at 14 000 rpm.

673 Proteins were quantified and Laemmli buffer was added before boiling for 10min at 95°C to be
674 used for Western Blot.
675
676 **Human Samples:**
677 Cerebellum, frontal cortex and cingulate gyrus human samples were provided by Neuro-CEB
678 biobank from a 78-year-old healthy male individual and conserved at -80°C.
679
680 **Human Samples pulverization and protein extraction**
681 We used the dry pulverizer Cryoprep (Covaris) for pulverization of tissue blocs. Each sample
682 was disposed in a liquid-nitrogen precooled Tissue-tube bag and dry cryo-pulverized with one
683 impact at the maximum level. The pulverized brain sample was then weighed and resuspended
684 in lysis buffer (mg/v) (0.32M sucrose, 5mM $CaCl_2$, 3mM $Mg(CH3COOH)_2$, 0.1mM EDTA,
685 10mM Tris-HCL pH8, 1mM DTT, 0.1% TritonX-100 and Protease Inhibitors), kept on ice for
686 30 min with gentle up-and-down pipetting until homogenization. We added 2X RIPA buffer
687 (v/v) to totals fractions for 30 min on ice. We then sonicated samples 2 times for 15 min.
688 AtlasSupernatants were collected after a 30 min centrifugation at 14 000 rpm at 4°C, proteins
689 were quantified and Laemmli buffer was added to be used for Western Blot. All samples were
690 boiled 10 min at 95°C to be used for Western Blot.
691
692 **Western-Blot:**
693 We used 1.5mm NuPAGE 4-12% Bis-Tris-Gel (Invitrogen™). Proteins samples were loaded
694 and gel migration was performed with NuPAGE™ MES SDS Running Buffer (Invitrogen™)
695 for 45 minutes at 200mV. Gels were transferred onto a methanol activated PVDF membrane
696 (Immobilon) in a buffer containing: Tris 25 mM, pH=8,3 and Glycine 192 mM, during 1h30 at
697 400 mA. Membranes were blocked 30 min with 5% milk in TBST (0,2% Tween 20, 150 mM

698 NaCl, 10 mM Tris pH:8). Primary antibodies were diluted in 5% milk in TBST, and membranes
699 were incubated o/n at 4 C°. After 3 x 10 min washing in TBST, membranes were incubated for
700 1h30 with the respective secondary antibodies diluted at a concentration of 1/2000 in 5% milk
701 TBST. Membranes were washed 3 x 10 min in TBST and were revealed by the LAS-4000
702 Fujifilm system using Clarity Western ECL Substrate (Bio Rad) or Maxi Clarity Western ECL
703 Substrate (Bio Rad).
704
705 **Immunostaining:**
706 Sagittal mouse brains slices were fixed for 10min in PFA 4% and rinsed 3 times for 10min in
707 PBS -/-. Slices were then incubated 20 min in glycine 100mM, washed 3 times for 5min in PBS
708 and immersed in 10mM citrate pH 6 at 62°C during 45min for antigen retrieval. Slices were
709 then immersed 3 times in PBS with Triton X-100 0.2% and incubated in blocking buffer for 1,5
710 h (PBS with Triton X-100 0.2% and FBS (10%) previously inactivated 20min at 56°C (Gibeco,
711 16141061). Primary antibodies (ORF1p antibody: abcam ab216324; NeuN antibody: GeneTex
712 GTX00837) were diluted (1/200 and 1/500 respectively) in blocking buffer and incubated with
713 slices overnight at 4°C and then washed 3 times for 10 min with PBS. For validation, an in-
714 house ORF1p antibody was used (09) (guinea pig, 1/200). For non-neuronal marker (GFAP
715 antibody: Millipore AB5541; Iba1 antibody: GeneTex GTX101495; Sox9 antibody:
716 RnDsystems AF3075; S100β antibody: Sigma S2532), antibodies were diluted at 1/500.
717 Suitable secondary antibodies (Invitrogen) and Hoechst (Invitrogen, 15586276) were incubated
718 for 1,5h at 1/2000 in PBS with inactivated FBS (10%) and washed 3 times 10 min in PBS. To
719 quench tissue autofluorescence, especially lipofuscin, TrueBlack Plus (Biotium) in PBS was
720 used during 10min. Slices were rinsed 3 times in PBS and mounted with Fluoromount
721 (Invitrogen™). For human cingulate gyrus stainings, the same protocol was performed, with
722 the difference that a human ORF1p antibody (Abcam 245249) was used. Mouse and human

723 brain slices were imaged by the Axioscan 7 Digital Slide Scanner (Zeiss) or a Spinning Disk
724 W1 confocal microscope (Yogogawa).
725
726 **Blocked peptide**
727 The ORF1p antibody (abcam ab216324) was incubated 2h on a turning wheel with excess (4:1)
728 of mouse ORF1p recombinant protein as in [17] before the blocked antibody was used in the
729 above-described immunofluorescence protocol.
730
731 **Quantification of confocal acquisitions**
732 Analysis was performed with a custom-written plugin developed for the Fiji software, using
733 Bio-Formats [98] and 3D Image Suite [99] libraries. Code is freely available online at
734 https://github.com/orion-cirb/DAPI_NEUN_ORF1P. Nuclei were detected using Hoechst
735 channel downscaled by a factor of 2 with the 2D-stitched version of Cellpose [100] (percentile
736 normalization = [1-99], model = 'cyto', diameter = 30, flow threshold = 0.4, cell probability
737 threshold = 0.0, stitching threshold = 0.75). The segmented image is rescaled to its original size,
738 and the obtained 3D nuclei are filtered by volume to avoid false positive detections. NeuN and
739 ORF1p+ cells were detected in their respective channel in the same manner as nuclei, but with
740 different Cellpose settings (model = 'cyto2', diameter = 40, flow threshold = 0.4, cell
741 probability threshold = 0.0, stitching threshold = 0.75). Then, each cell was associated with a
742 nucleus having at least half of its volume in contact with it. Cells without any associated nucleus
743 were filtered out. Each nucleus was thus labeled according to NeuN and/or ORF1p positivity.
744
745 **ABBA Registration and Qupath analysis**

746 We registered each sagittal brain section with the Allen Mouse Brain Atlas (CCFv3 [101]), using
747 the Aligning Big Brains & Atlases plugin [102] in Fiji. Registration results were imported into
748 QuPath software [103] for downstream processing.
749 In each brain subregion, cell analysis was done with custom Groovy scripts developed for
750 QuPath. Code is freely available online at https://github.com/orion-cirb/QuPath_ORF1P. In
751 brief, Hoechst nuclei were detected with StarDist 2D [104], applying the DSB 2018 pretrained
752 model with the following parameters: percentile normalization = [1-99], probability threshold
753 = 0.75, overlap threshold = 0.25. Cells in NeuN and ORF1p channels were detected with
754 Cellpose 2D (percentile normalization = [1-99], model = 'cyto2', diameter = 30, flow threshold
755 = 0.4, cell probability threshold = 0.0). Nuclei and cells were then filtered by area and intensity
756 in their respective channel, once again to avoid false positive detections. Minimum intensity
757 threshold was based on the channel background noise. This noise was estimated for each
758 subregion as the mean intensity of pixels not belonging to any nucleus or cell detected in the
759 corresponding channel. Finally, each cell was associated with a nucleus having its centroid
760 located inside the cell mask. Cells without any associated nucleus were filtered out. Each cell
761 containing a nucleus was thus identified as NeuN+ or NeuN- and ORF1p+ or ORF1p-. Intensity
762 values were normalized by subtracting the background noise computed in the corresponding
763 channel and subregion. As a last step, subregional results were merged into regional ones and
764 data were analyzed using the Pandas Python library [105].
765
766 **FACS**
767 Mouse brains were dissociated with Adult Brain Dissociation kit (Miltenyi Biotec, 130-107-
768 677) and incubated with the coupled antibody NeuN Alexa 647 (Abcam, ab190565) or the
769 control isotype IgG Alexa 647 (Abcam, ab199093). Stained cells were filtered a last time with
770 a 40µm filter before FACS sorting (FACS ARIA II). Neuronal and non-neuronal cells were

771 separately collected in PBS -/- 2m EDTA and then centrifugate (5min at 700rpm). Pellets
772 were resuspended in RIPA for protein extraction in an appropriate volume in order to achieve
773 equal cell concentrations (10 000 cells/µl).
774
775 **RNA-seq analysis**
776 The RNA-seq dataset from Dong et al [47] was downloaded from dbGAP (phs001556.v1.p1) and
777 contains unstranded paired-end 50bp and 75bp reads from pooled laser-capture micro-dissected
778 dopaminergic neurons from human post-mortem brain (107 samples) from 93 individuals w/o
779 brain disease. RNA-seq had been done on total and linearly amplified RNA. We focused our
780 analysis on data obtained with 50bp reads, in order to avoid mappability bias, while still
781 regrouping all age categories (n=41; with ages ranging from 38 to 97 (mean age: 79.88 (SD
782 ±12.07); n=6 ≤65y; n=35 >65y; mean PMI: 7.07 (SD ±7.84), mean RIN: 7.09 (±0.94)).
783 Sequencing reads were aligned on the Human reference genome (hg38) using the STAR mapper
784 (v2.7.0a) 3 and two different sets of parameters. Genome-wide individual repeat quantification
785 was performed using uniquely mapped reads and the following STAR parameters: --
786 outFilterMultimapNmax 1–-alignEndsType EndToEnd–-outFilterMismatchNmax 999–-
787 outFilterMismatchNoverLmax 0.06. Repeats class, family and name quantification was
788 performed using a random mapping procedure and the following parameters : --
789 outFilterMultimapNmax 5000–-outSAMmultNmax 1–-alignEndsType EndToEnd–-
790 outFilterMismatchNmax 999–-outFilterMismatchNoverLmax 0.06–-outSAMprimaryFlag
791 OneBestScore–-outMultimapperOrder Random.
792 Repeats annotations were downloaded from the UCSC Table Browser (repeatMasker database:
793 https://genome.ucsc.edu/cgi-bin/hgTables) and coordinates of LINE-1 full length and coding
794 elements were downloaded from the L1base database 2 (http://l1base.charite.de/l1base.php; [4])
795 selecting LINE-1 full length elements containing two predicted complete open reading frames

796 for ORF1 and ORF2 (UID= Unique IDentifier) from the LINE-1 database
797 (http://l1base.charite.de/l1base.php) and corrected genomic intervals with the repeat masker
798 annotation of the corresponding genomic locus. Repeat quantification from the aligned data
799 was done using a gtf file composed of all genes (Gencode v29) and all individual repeat
800 elements. This strategy was used to avoid overestimation of repeat elements due to overlaps
801 with expressed genes. For individual repeat quantification of the full length L1 elements
802 (L1base), we therefore used a gtf of all genes and all L1base entries, and ran the FeatureCounts
803 tool [106] with the following parameters: -g gene_id -s 0 -p. In the context of the family-based
804 analysis, we used a gtf with all genes and all annotated repeats elements and ran FeatureCounts
805 with -g gene_family -s 0 -p -M. Before DeSeq2 analysis, we remove all genes and repeat
806 elements with less than 10 reads in a minimum of n individuals, n being the number of
807 individuals in the condition containing the fewest individuals ("young" condition: n=6, 38-65y,
808 mean 57.5 years). We use the same conditions with genes and UIDs with less than 3 reads in a
809 minimum of n individuals. Finally, we calculated the scaling factors using DeSeq2 on all genes
810 + all repeat elements or all gene + UID according to the quantification method and then applied
811 these scaling factors to the corresponding counts tables.
812 In order to test for the mappability of each UID (= full-length and coding LINE-1), we extracted
813 the bed track « main on human:umap50 (genome hg38) from the UCSC genome browser (≈
814 7Mio regions) directly into Galaxy (usegalaxy.org) and joined genomic intervals with a
815 minimum overlap of 45bp of this dataset with a dataset containing the annotation of UIDs
816 extracted from L1Basev2 [4] corrected in length with repeat masker and completed with
817 information on whether the UID is intra- or intergenic and, if intragenic, in which gene
818 (NM_ID, chr, strand, start, end, gene length, number of exons, gene symbol) the fl-LINE-1 is
819 located, which resulted in 1266 regions. We then used the "group on data and group by"
820 function in Galaxy and counted the number of overlapping 50kmers with all 146 UIDs

821 (=mappability score). Correlation analysis (non-parametric Spearman) was then done between
822 the mappability score and the normalized read counts.
823
824 **Immpunoprecipitation of ORF1p from the mouse brain**
825 For immunoprecipitation, we used ORF1p (abcam, ab245122) and IgG rabbit (abcam,
826 ab172730) antibodies. The antibodies were coupled to magnetic beads using the Dynabeads®
827 Antibody Coupling Kit (Invitogen, 14311D) according to the manufacture''s recommendations.
828 We used 5µg of antibody for 1 mg of beads and used 1.5mg of beads for IP. Individual mouse
829 brain lysates (n=5), homogenized using dounce and sonicated, were incubated with ORF1p or
830 IgG-control coupled beads and a small fraction was kept as input. Each of these two tubes
831 containing coupled beads and brain lysates were diluted in 5 ml buffer (10 mM Tris HCl, 150
832 mM NaCl, protease inhibitor). The samples were then incubated overnight on a wheel at 4°C.
833 Samples were then washed 3 times with 1 ml buffer (10 mM Tris HCl pH 8, 200 mM NaCl)
834 using a magnet and then resuspended in the same buffer. The samples were boiled in Laemmli
835 buffer (95°C, 10 min) and 20 µl of each sample were loaded on a 4-12% Nupage gel (Invitrogen,
836 NP0336) to be revealed by WB. For samples used in Mass Spectrometry study, beads were
837 washed with buffer (10 mM Tris HCl pH 8, 200 mM NaCl) using a magnet. After 3 washes
838 with 1ml buffer the beads were washed twice with 100 µL of 25 mM NH4HCO3 (ABC buffer).
839 Finally, beads were resuspended in 100 µl of 25mM ABC buffer and digested by adding 0.20
840 µg of trypsine/LysC (Promega) for 1 hour at 37 °C. A second round of digestion was applied
841 simultaneously on the beads by adding 100 µL of 25 mM ABC buffer and to the previous digest
842 by adding 0.20 µg of trypsin/LysC for 1 hour at 37 °C. Samples were then loaded into
843 homemade C18 StageTips packed by stacking three AttractSPE® disk (#SPE-Disks-Bio-C18-
844 100.47.20 Affinisep) into a 200 µL micropipette tip for desalting. Peptides were eluted using a
845 ratio of 40:60 CH3CN:H2O + 0.1% formic acid and vacuum concentrated to dryness with a

846 SpeedVac device. Peptides were reconstituted in 10 µL of injection buffer in 0.3%
847 trifluoroacetic acid (TFA) before liquid chromatography-tandem mass spectrometry (LC-
848 MS/MS) analysis.
849
850 **Mass Spectrometry**
851 Online chromatography was performed with an RSLCnano system (Ultimate 3000, Thermo
852 Scientific) coupled to a Q Exactive HF-X with a Nanospay Flex ion source (Thermo Scientific).
853 Peptides were first trapped on a C18 column (75 µm inner diameter × 2 cm; nanoViper Acclaim
854 PepMapTM 100, Thermo Scientific) with buffer A (2/98 MeCN/H2O in 0.1% formic acid) at
855 a flow rate of 2.5 µL/min over 4 min. Separation was then performed on a 50 cm x 75 µm C18
856 column (nanoViper Acclaim PepMapTM RSLC, 2 µm, 100Å, Thermo Scientific) regulated to
857 a temperature of 50°C with a linear gradient of 2% to 30% buffer B (100% MeCN in 0.1%
858 formic acid) at a flow rate of 300 nL/min over 91 min. MS full scans were performed in the
859 ultrahigh-field Orbitrap mass analyzer in ranges m/z 375–1500 with a resolution of 120 000 at
860 m/z 200. The top 20 intense ions were subjected to Orbitrap for further fragmentation via high
861 energy collision dissociation (HCD) activation and a resolution of 15 000 with the intensity
862 threshold kept at 1.3 × 105. We selected ions with charge state from 2+ to 6+ for screening.
863 Normalized collision energy (NCE) was set at 27 and the dynamic exclusion of 40s. For
864 identification, the data were searched against the Mus Musculus (UP000000589_10090
865 012019) Uniprot database using Sequest HT through proteome discoverer (version 2.4).
866 Enzyme specificity was set to trypsin and a maximum of two-missed cleavage sites were
867 allowed. Oxidized methionine, Met-loss, Met-loss-Acetyl and N-terminal acetylation were set
868 as variable modifications. Maximum allowed mass deviation was set to 10 ppm for
869 monoisotopic precursor ions and 0.02 Da for MS/MS peaks. The resulting files were further
870 processed using myProMS97 v3.10.0. FDR calculation used Percolator and was set to 1% at

871 the peptide level for the whole study. The label free quantification was performed by peptide
872 Extracted Ion Chromatograms (XICs), reextracted by conditions and computed with
873 MassChroQ version 2.2.21 [107]. For protein quantification, XICs from proteotypic peptides
874 shared between compared conditions (TopN matching) with missed cleavages were used.
875 Median and scale normalization at peptide level was applied on the total signal to correct the
876 XICs for each biological replicate (n=5). To estimate the significance of the change in protein
877 abundance, a linear model (adjusted on peptides and biological replicates) was performed, and
878 p-values were adjusted using the Benjamini–Hochberg FDR procedure. Proteins with at least
879 three peptides, identified in each biological replicates of ORF1p condition, a 10-fold enrichment
880 and an adjusted p-value ≤ 0.05 were considered significantly enriched in sample comparisons.
881 Unique proteins were considered with at least three peptides in all replicates. Protein selected
882 with these criteria were used for Gene Ontology enrichment analysis and string network
883 analysis.
884 The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consorti-
885 um (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository [108] with the
886 dataset identifier PXD047160.
887
888 **GO term and STRING network analysis**
889 Gene Ontology analysis was performed using GO PANTHER (PANTHER: Making genome-
890 scale phylogenetics accessible to allPaul D. Thomas, Dustin Albou and Huaiyu MiProtein
891 Society. 2022;31(1):8-22. doi:10.1002/pro.4218) and String network physical interactions were
892 retrieved using the STRING database v11.5 (https://string-db.org/) and then implemented in
893 Cytoscape software [109].
894
895

896 **Statistical analyses**

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

903

904

905 **Acknowledgments**

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

913

914 **Contributions**

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

Competing interests

The authors declare no competing interests.

946 **Figure Legends**

947

948 **Fig. 1: Widespread and heterogenous expression of ORF1p protein in the mouse brain**

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

961 (B) Widespread and heterogenous expression of the LINE-1 encoded protein ORF1p in the
962 mouse brain. Representative image of ORF1p immunostaining (orange) of a sagittal section of
963 the brain of a young (three months-old) mouse acquired on a slide scanner. Scale bar = 1mm.
964 (1-10) Representative images of immunostainings showing ORF1p expression (orange) in 10
965 different regions of the mouse brain acquired on a confocal microscope. Nuclei are represented
966 in blue (Hoechst), scale bar = 50µm. (1) Isocortex, (2) Hippocampus, (3) Striatum dorsal, (4)
967 Thalamus, (5) Midbrain motor, (6) Pallidum, (7) Hypothalamus, (8) Substantia nigra pars
968 compacta, (9) Hindbrain, (10) Cerebellum. ORF1p expression profile in the mouse brain. The
969 entire mouse brain with the exception of the olfactory bulb and the cerebellum were analyzed
970 according to the pipeline on large-scale images described in (A). Bar plot showing the total

971 number of ORF1p+ cells per mm² in the mouse brain. Data is represented as mean ±SEM, n=4
972 mice (top). Bar plot indicating the proportion of ORF1p+ cells compared to all cells detected.
973 Data is represented as mean ±SEM, n=4 mice, 202001 total cells analyzed (middle). Scatter
974 plot showing the mean intensity of ORF1p per ORF1p+ cell. Data is represented as mean ±SD,
975 n=4 mice, 40999 ORF1p+ cells analyzed (bottom).
976 (D-F) ORF1p expression profile (density, proportion and expression) in defined anatomical
977 regions of the mouse brain. Nine anatomical regions as defined by the Allen Brain Atlas and
978 mapped onto sagittal brain slices (four three-month-old Swiss/ OF1) with ABBA were analyzed
979 using the pipeline on large scale images described in (A). (D) ORF1p+ cell density in 9 different
980 regions. Bar plot showing the number of ORF1p+ cells per mm². Data is represented as mean
981 ±SEM; *p<0.05; **p<0.01; adjusted p-value, one-way ANOVA followed by a Benjamin-
982 Hochberg test (E) Proportion of ORF1p positive cells in 9 different regions. Bar plot showing
983 the proportion of ORF1p+ cells among all cells detected per region. Data is represented as mean
984 ±SEM. (F) Mean ORF1p expression per cell in 9 different regions. Dot plot showing the mean
985 intensity of ORF1p signal per ORF1p+ cell in 9 different regions. Data is represented as mean
986 ±SD. The number of analyzed cells per region is indicated in the figure. *p<0.05; **p<0.01;
987 ***p<0.001; ****p<0.0001; adjusted p-value, nested one-way ANOVA followed by Sidak'
988 multiple comparison test.
989 (G) ORF1p expression in the frontal cortex and ventral midbrain. Confocal images with
990 multiple z-stacks were analyzed using Cellpose and Stardist. Dot plot representing the mean
991 expression of ORF1p per ORF1p+ cells. Individual (four three-month-old Swiss/ OF1) mice
992 are represented each by a different color, the scattered line represents the median.
993 ****p<0.0001, nested one-way ANOVA. Total cells analyzed = 4645.
994 (H-I) ORF1p expression in the frontal cortex and the ventral midbrain. (H) Western blots
995 showing ORF1p (top) and actin expression (bottom) in four individual mice per region which

996 were quantified in (I) using actin as a reference control. The signal intensity is plotted as the
997 fold change of ORF1p expression in the ventral midbrain to ORF1p expression in the frontal
998 cortex. *p<0.05; two-sided, unpaired student's-test.
999 (J) ORF1p expression in three regions of the human brain. Western blot showing ORF1p
1000 expression in the cingulate gyrus (CG), frontal cortex (FC) and cerebellum (CB) of post-mortem
1001 tissues from a healthy individual. ORF1p (Top), Actin (bottom).
1002
1003 **Fig. 2: ORF1p is predominantly expressed in neurons in the mouse brain**
1004 (A) ORF1p expression is absent in the white matter (corpus callosum) and predominantly
1005 expressed in neurons. Proportion of ORF1p+/NeuN+, ORF1p+/NeuN-, ORF1p-/NeuN+ and
1006 ORF1p-/NeuN- cells in the white matter (corpus callosum) and the grey matter (left) and in
1007 nine different regions (right) analyzed by the cell detection pipeline on large scale images
1008 presented in Figure1A. Exact values can be found in Suppl_Table1.
1009 (B) Representative confocal microscopy images showing ORF1p (red) and NeuN expression
1010 (green) in two different regions of the mouse brain. The bottom images show the merge of the
1011 two stainings, an overlap of both markers is represented in orange. z-projections; scalebar =
1012 25µm.
1013 (C) Proportion of neurons expressing ORF1p in the frontal cortex and ventral midbrain
1014 quantified on confocal images. ns: non-significant; chi-square test on the cell number of the
1015 different cell-types analyzed; n=4 mice, data is represented as mean ± SEM.
1016 (D) Proportion of ORF1p+ cells identified as NeuN+ or NeuN- in two different regions,
1017 analyzed by confocal microscopy on multiple z-stacks. ns: non-significant; chi-square test, n=4
1018 mice.
1019 (E) ORF1p does not colocalize with glial or microglial cell markers. Representative confocal
1020 microscopy images showing ORF1p staining (red) and three different glial cell (GFAP, Sox9,

1021 S100β) or microglial (Iba1) markers (white). Note that Iba1 antibody (rabbit) was used with
1022 the ORF1p 09 antibody (guinea pig, in house) z-projections, scalebar = 50µm.
1023 (F-G) Separation of neuronal and non–neuronal - cells by FACS confirms predominant
1024 neuronal expression of ORF1p. (F) Neuronal (NeuN+) and non-neuronal (NeuN-) cells isolated
1025 by fluorescent activated cell sorting (FACS). Dot plots showing autofluorescence versus an
1026 appropriate control antibody (IgG rabbit 647; left) and an antibody against NeuN (name of the
1027 AB 657, right). The P4 window represents isolated NeuN+ cells (pink) and the P5 fraction
1028 NeuN- cells (orange) containing the same number of cells as sorted in P4 for comparison, others
1029 NeuN- are represented in blue. (G) Western blot. ORF1p expression (top), in NeuN- and NeuN+
1030 FACS-sorted cells stemming from Figure F.
1031 (H) Representative confocal microscopy images showing ORF1p (red), NeuN (green) and
1032 Hoechst (blue) in the cingulate gyrus of the human brain. z-projection; scalebar = 25µm (left).
1033 Example of individuals neurons expressing ORF1p or not are shown on the right panel. z-
1034 projection; scalebar = 5 µm (right).
1035 (I) Proportion of ORF1p+ cells identified as NeuN+ or NeuN- in the human cingulate gyrus,
1036 analyzed by confocal microscopy on multiple z-stacks.
1037 (J) Proportion of neurons expressing ORF1p in the human cingulate gyrus, analyzed by
1038 confocal microscopy on multiple z-stacks.
1039
1040 **Fig. 3: ORF1p expression is increased throughout the whole mouse brain in the context**
1041 **of aging**
1042 (A) Proportion of ORF1p+/NeuN+, ORF1p+/NeuN+, ORF1p+/NeuN-, ORF1p-/NeuN+ and
1043 ORF1p-/NeuN- cell-type in the whole brain (left) and the different analyzed regions (right) of
1044 young (three-month aged, n=4) and aged (16-month-old, n=4) mice using the cell detection

1045 pipeline on large scale images presented in Figure 1A, data is represented as mean ± SEM.
1046 Exact values can be found in Suppl_Table1.
1047 (B) Proportion of ORF1p+/NeuN+, ORF1p+/NeuN+, ORF1p+/NeuN-, ORF1p-/NeuN+ and
1048 ORF1p-/NeuN- cell-type in two different regions of young and aged mice, analyzed on multiple
1049 z-stack confocal images. ns: non-significant; *p<0.05 calculated using two-way ANOVA with
1050 sidak's multiple comparisons test on the cell number of the different cell-types analyzed; data
1051 is represented as mean ± SEM.
1052 (C) ORF1p mean expression per ORF1p+ cell in the brain analyzed on large-scale images. Dot
1053 plot showing the ORF1p mean expression per ORF1p+ cell in young (n=4) and aged (n=4) mice
1054 in the whole brain (except cerebellum and olfactory bulb). 74985 total cells were analyzed; *
1055 p<0.05, two-way ANOVA with sidak's multiple comparisons; data is represented as mean ±
1056 SEM.
1057 (D) Frequency distribution of ORF1p mean intensity in ORF1p+ cells. ***p<0.001,
1058 Kolmogorov-Smirnov test.
1059 (E) Frequency distribution of Hoechst mean intensity in the nuclei of OrF1p+ cells. ns: non-
1060 significant, Kolmogorov-Smirnov test.
1061 (F) Mean ORF1p expression per ORF1p+ cell in nine different anatomical regions. Dot plot
1062 showing the ORF1p mean expression per ORF1p positive cell (n=74985). Adjusted p-value are
1063 represented, two-tailed nested t-test followed by a Benjamin, Krieger and Yukutieli test; n=4
1064 young and n=4 aged mice per region, data is represented as mean ± SEM.
1065 (G) Color-coded representation of fold-changes of ORF1p expression with aging. Represented
1066 is the fold-change in percent (aged vs young) of the "mean of the mean" ORF1p expression per
1067 ORF1p+ cell quantified mapped onto the nine different regions analyzed as shown in (F).
1068 (H) ORF1p expression is increased in the ventral midbrain of aged mice. Dot plot representing
1069 ORF1p expression in two different regions of young and aged mice analyzed on confocal

1070 images with multiple z-stacks; total cells analyzed = 8381 ns: non-significant *p<0.05, two-
1071 tailed one-way ANOVA; dashed lines represent the medians.
1072 (I) Representative confocal microscopy acquisition showing increased ORF1p expression (red)
1073 in the ventral midbrain region of aged mice (one z plan is shown). Cell nuclei are shown in blue
1074 (Hoechst staining). Scalebar = 50µm.
1075
1076 **Fig. 4: Young LINE-1 elements are increased in aged human dopaminergic neurons**
1077 TE transcript expression in RNA-seq data of laser-captured micro-dissected post-mortem
1078 human dopaminergic neurons of brain-healthy individuals was analyzed using RepeatMasker
1079 (multimappers) or the L1Base (unique reads).
1080 (A) Volcano plot of differential analysis of LINE-1 expression using DESeq2 comparing young
1081 (≤65y, n=6) or aged (>65y, n=35) human dopaminergic neurons at the "name" level of
1082 RepeatMasker. Young LINE-1 elements, including the two families L1HS and L1PA2 that have
1083 coding copies, are highlighted in red.
1084 (B) Scatter plots of normalized read counts ("name" level) of the young L1HS and L1PA2
1085 families as well as the human endogenous virus family HERVK-int, another TE family with
1086 coding potential comparing young (≤65y, n=6) or aged (>65y, n=35) human dopaminergic
1087 neurons. Mann-Whitney test, p<0.05.
1088 (C) Correlation of the expression of LINE-1 elements with known regulators in human
1089 dopaminergic neurons. Spearman correlation of evolutionary close (L1HS, L1PA2-17) and
1090 distant LINE-1 (L1PB and L1MA5) as well as HERV elements with coding potential (HERV-
1091 Kint, HERV-Fc1, HERV-Fc2 and HERV-H-int) with known regulators of their expression.
1092 HERV-W and TREX1 did not pass the normalized read count threshold of >3 reads in >6
1093 individuals.
1094

1095 **Fig. 5: Dysregulation of locus-specific full-length LINE-1 elements in aged human**
1096 **dopaminergic neurons**
1097 (A) Volcano plot of differential expression analysis of TE expression using DEseq2 comparing
1098 young (≤65y, n=6) and aged (>65y, n=35) human dopaminergic neurons at the locus-level of
1099 specific full-length LINE-1 elements (140 of 146 "UID's" as annotated in L1Base; threshold
1100 >3 reads in at least 6 individuals).
1101 (B) Pairwise comparison of the expression of 140 out of 146 full-length LINE-1 elements
1102 comparing young (≤65y, n=6) and aged (>65y, n=35) human dopaminergic neurons. Wilcoxon
1103 matched signed rank test, p<0.0001left panel).
1104 (C) The sum of read counts of all UIDs per individual were plotted comparing young (≤65y,
1105 n=6) and aged (>65y, n=35) human dopaminergic neurons.
1106 (D-E) Dysregulated locus-specific full-length LINE-1 elements (UID-68 and UID-129) are
1107 plotted as scatter plots comparing young (≤65y, n=6) and aged (>65y, n=35) human
1108 dopaminergic neurons. (D) UID-68 is located adjacent to the genes CLEC5A and OR9A4 (left).
1109 Spearman correlation analysis of the expression of UID-68 and CLEC5A (middle) or OR9A4
1110 (right) in young (≤65y, n=6, black dots) or aged (>65y, n=35, red squares) human dopaminergic
1111 neurons. (E) UID-129 is intergenic.
1112
1113 **Fig. 6: Endogenous ORF1p interactors in the mouse brain**
1114 Immunoprecipitation (IP) of endogenous ORF1p from the mouse brain. WB against ORF1p
1115 showing ORF1p enrichment after IP but no signal in the IgG control. Five independent samples
1116 were then prepared for proteomic analysis by mass spectrometry (LC-MS/MS).
1117 (B) GO slim enrichment analysis of proteins selected as endogenous ORF1p protein partners in
1118 the mouse brain after quantitative LC-MS/MS. ORF1p-immunoprecipitated proteins were
1119 categorized into GO slim terms. The nine GO slim term with the highest fold-change are

1120 plotted. Fold enrichment is depicted on the upper axis and displayed as bars, the FDR value
1121 appears on the lower axis and is represented by the black points. BP: Biological Process, CC:
1122 Cellular Component, MF: Molecular Function.
1123 (C) Venn diagram showing common interactors (purple) between interactors of endogenous
1124 ORF1p in the mouse brain identified in this study (red) and known (published) interactors of
1125 ORF1p (blue). Statistical significance of the overlap between the two groups of proteins was
1126 tested by an overrepresentation test (http://nemates.org/MA/progs/overlap_stats.html).
1127 (D) ORF1p associates with the SWI/SNF complex (red), RNA pol II complex (orange) and
1128 interactors belonging to GO terms related to neuronal cell body & neuron projection (green).
1129 Known interactors previously published [60,65–70] are indicated with a purple ring. STRING
1130 network of physical interactions where nodes represent proteins partners identified in (A) and
1131 edges thickness represents the strength of shared physical complexes. Only proteins sharing
1132 physical interactions were represented.
1133
1134
1135 **Supplementary Figures**
1136
1137 Suppl Fig. 1
1138 (A) Selective recognition of ORF1p antibody. IHC showing ORF1p positives cells in sagittal
1139 mouse brain slice (left) and abolition of the signal when blocking the antibody with purified
1140 ORF1p (right).
1141 (B) Representative acquisition showing ORF1p obtained with the widely used, commercially
1142 available ORF1p ab antibody (abcam ab216324) used in this study (red) and with an in-house
1143 ORF1p gp antibody (guinea pig, green) in mouse brain. Scalebar = 20µm (top) and 100µm
1144 (bottom).

1145 (C) Quantification of double positives (gp+/ab+) cells using ORF1p ab antibody (abcam
1146 ab216324) and in-house ORF1p gp antibody (guinea pig) versus single-positive cells (gp+/ab-
1147 and gp-/ab+) in mouse frontal cortex (left) and ventral midbrain (right).
1148 (D) ORF1p is expressed in six different brain regions in the mouse.
1149 Brain regions were micro-dissected from a three-month old mouse brain. Western blot showing
1150 ORF1p expression in 6 brain regions. ORF1p (Top), Actin (bottom).
1151
1152 Suppl Fig. 2
1153 (A) Proportion of neurons (NeuN+) expressing ORF1p in different regions of the mouse brain
1154 as quantified using the cell detection pipeline on large scale images.
1155 (B) ORF1p is predominantly expressed in neurons. Proportion of ORF1p+/NeuN+,
1156 ORF1p+/NeuN+, ORF1p+/NeuN-, ORF1p-/NeuN+ and ORF1p-/NeuN- cells, ****$p<0.0001$;
1157 calculated using chi-square test on the cell number of the four different cell-types analyzed by
1158 confocal microscopy on multiple z-stacks.
1159 (C) ORF1p cell identity. Proportion of ORF1p+ cells identified as NeuN+ (black) or NeuN-
1160 (grey), in the whole brain (left) and in 9 different regions analyzed (right) using the cell
1161 detection pipeline on large scale images presented in Figure 1A; data is represented as mean ±
1162 SEM, n=4 mice.
1163 (D) Proportion of neurons in the frontal cortex and ventral midbrain quantified using confocal
1164 approach. *$p<0.05$, chi-square test on the cell number of the different cell-types analyzed; n=4
1165 mice, data is represented as mean ± SEM.
1166 (E) Representative slide-scanner acquisition of a human cingulate gyrus section showing NeuN
1167 positives cells (green) mostly located in the grey matter (right) compared to the white matter
1168 from a brain-healthy individual; scale bar = 400µm. Zoom into the grey matter region showing
1169 ORF1p is presented in Figure 2H.

1170 Suppl Fig. 3

1171 (A) Proportion of ORF1p+ cells being neuronal in the ventral midbrain comparing young and
1172 aged mice as quantified using confocal approach. Kolmogorov-Smirnov test; data is
1173 represented as mean ± SEM.

1174 (B) Proportion of neurons expressing ORF1p in the ventral midbrain comparing young and
1175 aged mice as quantified using confocal approach. Kolmogorov-Smirnov test; data is
1176 represented as mean ± SEM.

1177 (C) Proportion of neurons in the ventral midbrain comparing young and aged mice as quantified
1178 using confocal approach**.** Kolmogorov-Smirnov test; data is represented as mean ± SEM.

1179

1180 Suppl Fig. 4

1181 (A-B) Comparison of the expression of dopaminergic markers tyrosine hydroxylase (TH, A)
1182 and LMX1B (B) between young (≤65y, n=6)) and aged (>65y, n=35) human dopaminergic
1183 neurons. Mann Whitney test.

1184 (C, E) Volcano plot of differential expression analysis of TE expression using DEseq2
1185 comparing young (≤65y, n=6) and aged (>65y, n=35) human dopaminergic neurons at the
1186 "class" (C) and "family" (E) level of RepeatMasker.

1187 (D) Scatter plot comparing the expression of LINE at the "class" level between young (≤65y,
1188 n=6)) and aged (>65y, n=35) human dopaminergic neurons. Mann Whitney test.

1189 (F) Scatter plot comparing the expression of LINE and Alu at the "family" level between young
1190 (≤65y, n=6)) and aged (>65y, n=35) human dopaminergic neurons. Mann Whitney test.

1191 (G) Scatter plot comparing the expression of HERVH-int, HERV-Fc1 and two non-coding,
1192 non-autonomous but active TEs in the human genome, AluYa5 and SVA-F at the "name" level
1193 between young (≤65y, n=6)) and aged (>65y, n=35) human dopaminergic neurons. Mann
1194 Whitney test.

1195 Suppl Fig. 5

1196 (A-C) Correlation analyses of L1HS expression with Engrailed 1 (EN1, A, Spearman r=-0.43,
1197 p=0.002), CBX5/HP1 (B, Spearman r=-0.35, p=0.01) and XRCC6 expression (C, Spearman r=
1198 -0.394, p=0.005). Normalized read counts are plotted. Black dots correspond to young
1199 individuals (≤65y), red dots correspond to aged individuals (>65y).

1200 (D-G) Scatter plots comparing the expression of EN1, CBX5/HP1, XRCC5 and XRCC6
1201 between young (≤65y, n=6)) and aged (>65y, n=35) human dopaminergic neurons. Student's t-
1202 test (EN1) or Mann Whitney test (HP1, XRCC5/6).

1203

1204 Suppl Fig. 6

1205 (A) Correlation of mappability (UMAP hit counts over UID, see methods) and UID expression
1206 (normalized read counts). Spearman correlation.

1207 (B-D) *In silico* analysis of annotated full-length LINE-1 elements as in L1Basev2 (human
1208 reference genome hg38). (B) Percentage of L1HS and L1PA2 elements among the 146 full-
1209 length elements (UID1-146). (C) Percentage of full-length LINE-1 elements located inside or
1210 outside a gene. (D) Presence (blue, with gene symbol) or absence (white) of a "hosting" gene
1211 among the 146 annotated full-length LINE-1 in the human reference genome

1212 (E) Mean expression of all 146 full-length LINE-1 elements in dopaminergic neurons of all
1213 individuals ≤65y.

1214

1215 Suppl Fig. 7

1216 (A-C) Dysregulated locus-specific full-length LINE-1 elements are plotted as scatter plots
1217 comparing young (≤65y, n=6) and aged (>65y, n=35) human dopaminergic neurons. (A) UID-
1218 37 is located in an intron of the gene HPSE2 (left). Spearman correlation analysis of the
1219 expression of UID-37 and HPSE2 in young (≤65y, n=6, black dots) and aged (>65y, n=35, red

1220 squares) human dopaminergic neurons. (B) UID-127 is located within the 6th intron of the non-
1221 coding RNA LINC00598 (left). Spearman correlation analysis of the expression of UID-127
1222 and LINC00598 in young (≤65y, n=6, black dots) and aged (>65y, n=35, red squares) human
1223 dopaminergic neurons. (C) UID-137 is intergenic.
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244

References

1. A hidden human proteome encoded by “non-coding” genes - PubMed. Accessed October 19, 2023. https://pubmed.ncbi.nlm.nih.gov/31340039/
2. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921. doi:10.1038/35057062
3. 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
4. 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
5. 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
6. Esnault C, Maestre J, Heidmann T. Human LINE retrotransposons generate processed pseudogenes. *Nat Genet*. 2000;24(4):363-367. doi:10.1038/74184
7. 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
8. Mathias SL, Scott AF, Kazazian HH, Boeke JD, Gabriel A. Reverse transcriptase encoded by a human transposable element. *Science*. 1991;254(5039):1808-1810. doi:10.1126/science.1722352
9. Feng Q, Moran JV, Kazazian HH, Boeke JD. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell*. 1996;87(5):905-916.
10. Chen L, Dahlstrom JE, Lee SH, Rangasamy D. Naturally occurring endo-siRNA silences LINE-1 retrotransposons in human cells through DNA methylation. *Epigenetics : official journal of the DNA Methylation Society*. 2012;7(7):758-771. doi:10.4161/epi.20706
11. Goodier JL, Cheung LE, Kazazian HH. MOV10 RNA helicase is a potent inhibitor of retrotransposition in cells. *PLoS Genetics*. 2012;8(10):e1002941. doi:10.1371/journal.pgen.1002941
12. Philippe C, Vargas-Landin DB, Doucet AJ, et al. Activation of individual L1 retrotransposon instances is restricted to cell-type dependent permissive loci. *eLife*. 2016;5. doi:10.7554/eLife.13926
13. Peze-Heidsieck E, Bonnifet T, Znaidi R, et al. Retrotransposons as a Source of DNA Damage in Neurodegeneration. *Frontiers in aging neuroscience*. 2021;13:786897. doi:10.3389/fnagi.2021.786897
14. Driver CJ, McKechnie SW. Transposable elements as a factor in the aging of Drosophila melanogaster. *Annals of the New York Academy of Sciences*. 1992;673(1):83-91. doi:10.1111/j.1749-6632.1992.tb27439.x
15. Copley KE, Shorter J. Repetitive elements in aging and neurodegeneration. *Trends in Genetics*. Published online March 2023:S0168952523000343. doi:10.1016/j.tig.2023.02.008
16. Gorbunova V, Seluanov A, Mita P, et al. The role of retrotransposable elements in ageing and age-associated diseases. *Nature*. Published online July 2021:1-11. doi:10.1038/s41586-021-03542-y
17. Blaudin de Thé FX, Rekaik H, Peze-Heidsieck E, et al. Engrailed homeoprotein blocks degeneration in adult dopaminergic neurons through LINE-1 repression. *The EMBO Journal*. 2018;37(15). doi:10.15252/embj.201797374
18. Bodea GO, Botto JM, Ferreiro ME, et al. LINE-1 retrotransposons contribute to mouse PV interneuron development. *Nat Neurosci*. Published online May 21, 2024.

1294 doi:10.1038/s41593-024-01650-2
1295 19. Sur D, Kustwar RK, Budania S, et al. Detection of the LINE-1 retrotransposonRNA-
1296 binding protein ORF1p in differentanatomical regions of the human brain. *Mobile DNA*.
1297 Published online November 2017:1-12. doi:10.1186/s13100-017-0101-4
1298 20. McKerrow W, Kagermazova L, Doudican N, et al. LINE-1 retrotransposon expression
1299 in cancerous, epithelial and neuronal cells revealed by 5′ single-cell RNA-Seq. *Nucleic Acids*
1300 *Research*. 2023;51(5):2033-2045. doi:10.1093/nar/gkad049
1301 21. Gasior SL, Wakeman TP, Xu B, Deininger PL. The human LINE-1 retrotransposon
1302 creates DNA double-strand breaks. *Journal of molecular biology*. 2006;357(5):1383-1393.
1303 doi:10.1016/j.jmb.2006.01.089
1304 22. Belgnaoui SM, Gosden RG, Semmes OJ, Haoudi A. Human LINE-1 retrotransposon
1305 induces DNA damage and apoptosis in cancer cells. *Cancer cell international*. 2006;6:13.
1306 doi:10.1186/1475-2867-6-13
1307 23. De Cecco M, Criscione SW, Peckham EJ, et al. Genomes of replicatively senescent
1308 cells undergo global epigenetic changes leading to gene silencing and activation of
1309 transposable elements. *Aging Cell*. 2013;12(2):247-256. doi:10.1111/acel.12047
1310 24. Van Meter M, Kashyap M, Rezazadeh S, et al. SIRT6 represses LINE1
1311 retrotransposons by ribosylating KAP1 but this repression fails with stress and age. *Nature*
1312 *communications*. 2014;5:5011. doi:10.1038/ncomms6011
1313 25. Sturm Á, Ivics Z, Vellai T. The mechanism of ageing: primary role of transposable
1314 elements in genome disintegration. *Cell Mol Life Sci*. 2015;72(10):1839-1847.
1315 doi:10.1007/s00018-015-1896-0
1316 26. Wallace N, Wagstaff BJ, Deininger PL, Roy-Engel AM. LINE-1 ORF1 protein
1317 enhances Alu SINE retrotransposition. *Gene*. 2008;419(1-2):1-6.
1318 doi:10.1016/j.gene.2008.04.007
1319 27. Thomas CA, Tejwani L, Trujillo CA, et al. Modeling of TREX1-Dependent
1320 Autoimmune Disease using Human Stem Cells Highlights L1 Accumulation as a Source of
1321 Neuroinflammation. *Cell stem cell*. 2017;21(3):319-331.e8. doi:10.1016/j.stem.2017.07.009
1322 28. De Cecco M, Ito T, Petrashen AP, et al. L1 drives IFN in senescent cells and promotes
1323 age-associated inflammation. *Nature*. 2019;566(7742):73-78. doi:10.1038/s41586-018-0784-9
1324 29. Luqman-Fatah A, Watanabe Y, Uno K, Ishikawa F, Moran JV, Miyoshi T. The
1325 interferon stimulated gene-encoded protein HELZ2 inhibits human LINE-1 retrotransposition
1326 and LINE-1 RNA-mediated type I interferon induction. *Nat Commun*. 2023;14(1):203.
1327 doi:10.1038/s41467-022-35757-6
1328 30. Gazquez-Gutierrez A, Witteveldt J, R Heras S, Macias S. Sensing of transposable
1329 elements by the antiviral innate immune system. *RNA (New York, NY)*. Published online April
1330 2021. doi:10.1261/rna.078721.121
1331 31. Krug L, Chatterjee N, Borges-Monroy R, et al. Retrotransposon activation contributes
1332 to neurodegeneration in a Drosophila TDP-43 model of ALS. *PLoS Genetics*.
1333 2017;13(3):e1006635. doi:10.1371/journal.pgen.1006635
1334 32. Casale AM, Liguori F, Ansaloni F, et al. Transposable element activation promotes
1335 neurodegeneration in a Drosophila model of Huntington's disease. *ISCIENCE*.
1336 2022;25(1):103702. doi:10.1016/j.isci.2021.103702
1337 33. Takahashi T, Stoiljkovic M, Song E, et al. LINE-1 activation in the cerebellum drives
1338 ataxia. *Neuron*. Published online August 2022:1-19. doi:10.1016/j.neuron.2022.08.011
1339 34. De Luca C, Gupta A, Bortvin A. Retrotransposon LINE-1 bodies in the cytoplasm of
1340 piRNA-deficient mouse spermatocytes: Ribonucleoproteins overcoming the integrated stress
1341 response. Feschotte C, ed. *PLoS Genet*. 2023;19(6):e1010797.
1342 doi:10.1371/journal.pgen.1010797
1343 35. Spencley AL, Bar S, Swigut T, et al. Co-transcriptional genome surveillance by

1344 HUSH is coupled to termination machinery. *Mol Cell*. 2023;83(10):1623-1639.e8.
1345 doi:10.1016/j.molcel.2023.04.014
1346 36. Garland W, Müller I, Wu M, et al. Chromatin modifier HUSH co-operates with RNA
1347 decay factor NEXT to restrict transposable element expression. *Mol Cell*. 2022;82(9):1691-
1348 1707.e8. doi:10.1016/j.molcel.2022.03.004
1349 37. Shirane K, Miura F, Ito T, Lorincz MC. NSD1-deposited H3K36me2 directs de novo
1350 methylation in the mouse male germline and counteracts Polycomb-associated silencing. *Nat*
1351 *Genet*. 2020;52(10):1088-1098. doi:10.1038/s41588-020-0689-z
1352 38. Guerrero A, Herranz N, Sun B, et al. Cardiac glycosides are broad-spectrum
1353 senolytics. *Nat Metab*. 2019;1(11):1074-1088. doi:10.1038/s42255-019-0122-z
1354 39. Gusel'nikova VV, Korzhevskiy DE. NeuN As a Neuronal Nuclear Antigen and
1355 Neuron Differentiation Marker. *Acta Naturae*. 2015;7(2):42-47. doi:10.32607/20758251-
1356 2015-7-2-42-47
1357 40. Cannon JR, Greenamyre JT. NeuN is not a reliable marker of dopamine neurons in rat
1358 substantia nigra. *Neuroscience Letters*. 2009;464(1):14-17. doi:10.1016/j.neulet.2009.08.023
1359 41. Sukhorukova EG. Nuclear Protein NeuN in Neurons in the Human Substantia Nigra.
1360 *Neurosci Behav Physi*. 2014;44(5):539-541. doi:10.1007/s11055-014-9947-9
1361 42. Kumar SS, Buckmaster PS. Neuron-specific nuclear antigen NeuN is not detectable in
1362 gerbil subtantia nigra pars reticulata. *Brain Res*. 2007;1142:54-60.
1363 doi:10.1016/j.brainres.2007.01.027
1364 43. Keller D, Erö C, Markram H. Cell Densities in the Mouse Brain: A Systematic
1365 Review. *Front Neuroanat*. 2018;12:83. doi:10.3389/fnana.2018.00083
1366 44. Sun W, Cornwell A, Li J, et al. SOX9 Is an Astrocyte-Specific Nuclear Marker in the
1367 Adult Brain Outside the Neurogenic Regions. *J Neurosci*. 2017;37(17):4493-4507.
1368 doi:10.1523/JNEUROSCI.3199-16.2017
1369 45. Yushkova E, Moskalev A. Transposable elements and their role in aging. *Ageing Res*
1370 *Rev*. 2023;86:101881. doi:10.1016/j.arr.2023.101881
1371 46. Gibb WR, Lees AJ. Anatomy, pigmentation, ventral and dorsal subpopulations of the
1372 substantia nigra, and differential cell death in Parkinson's disease. *J Neurol Neurosurg*
1373 *Psychiatry*. 1991;54(5):388-396. doi:10.1136/jnnp.54.5.388
1374 47. Dong X, Liao Z, Gritsch D, et al. Enhancers active in dopamine neurons are a primary
1375 link between genetic variation and neuropsychiatric disease. *Nature Neuroscience*. Published
1376 online September 2018:1-19. doi:10.1038/s41593-018-0223-0
1377 48. Teissandier A, Servant N, Barillot E, Bourc'his D. Tools and best practices for
1378 retrotransposon analysis using high-throughput sequencing data. *Mobile DNA*. 2019;10(1):52.
1379 doi:10.1186/s13100-019-0192-1
1380 49. Liu N, Lee CH, Swigut T, et al. Selective silencing of euchromatic L1s revealed by
1381 genome-wide screens for L1 regulators. *Nature*. 2018;553(7687):228-232.
1382 doi:10.1038/nature25179
1383 50. Rekaik H, Blaudin de Thé FX, Prochiantz A, Fuchs J, Joshi RL. Dissecting the role of
1384 Engrailed in adult dopaminergic neurons–Insights into Parkinson disease pathogenesis. *FEBS*
1385 *Letters*. 2015;589(24 Pt A):3786-3794. doi:10.1016/j.febslet.2015.10.002
1386 51. Maeda R, Tachibana M. HP1 maintains protein stability of H3K9 methyltransferases
1387 and demethylases. *EMBO reports*. 2022;23(4):e53581. doi:10.15252/embr.202153581
1388 52. Suzuki J, Yamaguchi K, Kajikawa M, et al. Genetic Evidence That the Non-
1389 Homologous End-Joining Repair Pathway Is Involved in LINE Retrotransposition. *PLoS*
1390 *Genetics*. 2009;5(4):e1000461. doi:10.1371/journal.pgen.1000461.g005
1391 53. Goerner-Potvin P, Bourque G. Computational tools to unmask transposable elements.
1392 *Nature Reviews Genetics*. Published online September 2018:1-17. doi:10.1038/s41576-018-
1393 0050-x

1394 54. Schwarz R, Koch P, Wilbrandt J, Hoffmann S. Locus-specific expression analysis of
1395 transposable elements. *Briefings in Bioinformatics*. 2022;23(1):bbab417.
1396 doi:10.1093/bib/bbab417
1397 55. Faulkner GJ. Elevated L1 expression in ataxia telangiectasia likely explained by an
1398 RNA-seq batch effect. *Neuron*. 2023;111(5):610-611. doi:10.1016/j.neuron.2023.02.007
1399 56. Lanciano S, Cristofari G. Measuring and interpreting transposable element expression.
1400 *Nature Reviews Genetics*. 2020;509:7. doi:10.1038/s41576-020-0251-y
1401 57. Martin SL, Li J, Weisz JA. Deletion analysis defines distinct functional domains for
1402 protein-protein and nucleic acid interactions in the ORF1 protein of mouse LINE-1. *J Mol*
1403 *Biol*. 2000;304(1):11-20. doi:10.1006/jmbi.2000.4182
1404 58. Briggs EM, McKerrow W, Mita P, Boeke JD, Logan SK, Fenyo D. RIP-seq reveals
1405 LINE-1 ORF1p association with p-body enriched mRNAs. Published online February 2021:1-
1406 13. doi:10.1186/s13100-021-00233-3
1407 59. Kulpa DA, Moran JV. Ribonucleoprotein particle formation is necessary but not
1408 sufficient for LINE-1 retrotransposition. *Human Molecular Genetics*. 2005;14(21):3237-3248.
1409 doi:10.1093/hmg/ddi354
1410 60. Taylor MS, Altukhov I, Molloy KR, et al. Dissection of affinity captured LINE-1
1411 macromolecular complexes. *eLife*. 2018;7:210-241. doi:10.7554/eLife.30094
1412 61. Kelly MP, Adamowicz W, Bove S, et al. Select 3',5'-cyclic nucleotide
1413 phosphodiesterases exhibit altered expression in the aged rodent brain. *Cell Signal*.
1414 2014;26(2):383-397. doi:10.1016/j.cellsig.2013.10.007
1415 62. Lakics V, Karran EH, Boess FG. Quantitative comparison of phosphodiesterase
1416 mRNA distribution in human brain and peripheral tissues. *Neuropharmacology*.
1417 2010;59(6):367-374. doi:10.1016/j.neuropharm.2010.05.004
1418 63. Mandel S, Gozes I. Activity-dependent neuroprotective protein constitutes a novel
1419 element in the SWI/SNF chromatin remodeling complex. *J Biol Chem*. 2007;282(47):34448-
1420 34456. doi:10.1074/jbc.M704756200
1421 64. Singh A, Modak SB, Chaturvedi MM, Purohit JS. SWI/SNF Chromatin Remodelers:
1422 Structural, Functional and Mechanistic Implications. *Cell Biochem Biophys*. 2023;81(2):167-
1423 187. doi:10.1007/s12013-023-01140-5
1424 65. Goodier JL, Cheung LE, Kazazian HH. Mapping the LINE1 ORF1 protein
1425 interactome reveals associated inhibitors of human retrotransposition. *Nucleic Acids*
1426 *Research*. 2013;41(15):7401-7419. doi:10.1093/nar/gkt512
1427 66. Taylor MS, LaCava J, Mita P, et al. Affinity proteomics reveals human host factors
1428 implicated in discrete stages of LINE-1 retrotransposition. *Cell*. 2013;155(5):1034-1048.
1429 doi:10.1016/j.cell.2013.10.021
1430 67. Moldovan JB, Moran JV. The Zinc-Finger Antiviral Protein ZAP Inhibits LINE and
1431 Alu Retrotransposition. *PLoS Genetics*. 2015;11(5):e1005121.
1432 doi:10.1371/journal.pgen.1005121
1433 68. Vuong LM, Pan S, Donovan PJ. Proteome Profile of Endogenous Retrotransposon-
1434 Associated Complexes in Human Embryonic Stem Cells. *Proteomics*. 2019;19(15):e1900169.
1435 doi:10.1002/pmic.201900169
1436 69. De Luca C, Gupta A, Bortvin A. *Ribonucleoprotein Condensation Driven by*
1437 *Retrotransposon LINE-1 Sustains RNA Integrity and Translation in Mouse Spermatocytes*.
1438 Molecular Biology; 2023. doi:10.1101/2023.01.09.523313
1439 70. Ardeljan D, Wang X, Oghbaie M, et al. LINE-1 ORF2p expression is nearly
1440 imperceptible in human cancers. *Mobile DNA*. 2020;11(1):1. doi:10.1186/s13100-019-0191-2
1441 71. Liu Q, Yi D, Ding J, et al. MOV10 recruits DCP2 to decap human LINE-1 RNA by
1442 forming large cytoplasmic granules with phase separation properties. *EMBO reports*.
1443 2023;24(9):e56512. doi:10.15252/embr.202256512

72. 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
73. 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
74. 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
75. 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
76. 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
77. 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
78. Garza R, Atacho D, Adami A, et al. L1 retrotransposons drive human neuronal transcriptome complexity and functional diversification.
79. Watanabe R, Nakachi Y, Matsubara H, et al. Identification of epigenetically active L1 promoters in the human brain and their relationship with psychiatric disorders. *Neuroscience Research*. Published online May 2023:S0168010223000913. doi:10.1016/j.neures.2023.05.001
80. 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
81. 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
82. 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
83. 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
84. 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
85. 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
86. 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
87. Dai L, LaCava J, Taylor MS, Boeke JD. Expression and detection of LINE-1 ORF-encoded proteins. *Mobile genetic elements*. 2014;4:e29319. doi:10.4161/mge.29319
88. Whongsiri P, Goering W, Lautwein T, et al. Many Different LINE-1 Retroelements Are Activated in Bladder Cancer. *International Journal of Molecular Sciences*.

2020;21(24):9433-14. doi:10.3390/ijms21249433
89. Lindtner S, Felber BK, Kjems J. An element in the 3' untranslated region of human LINE-1 retrotransposon mRNA binds NXF1(TAP) and can function as a nuclear export element. *RNA*. 2002;8(3):345-356. doi:10.1017/s1355838202027759
90. Naufer MN, Furano AV, Williams MC. Protein-nucleic acid interactions of LINE-1 ORF1p. *Seminars in Cell & Developmental Biology*. 2019;86:140-149. doi:10.1016/j.semcdb.2018.03.019
91. Garneau NL, Wilusz J, Wilusz CJ. The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol*. 2007;8(2):113-126. doi:10.1038/nrm2104
92. Lykke-Andersen J. Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay. *Mol Cell Biol*. 2002;22(23):8114-8121. doi:10.1128/MCB.22.23.8114-8121.2002
93. Threlfell S, Sammut S, Menniti FS, Schmidt CJ, West AR. Inhibition of Phosphodiesterase 10A Increases the Responsiveness of Striatal Projection Neurons to Cortical Stimulation. *J Pharmacol Exp Ther*. 2009;328(3):785-795. doi:10.1124/jpet.108.146332
94. Li N, Chen X, Zhu B, et al. Suppression of β-catenin/TCF transcriptional activity and colon tumor cell growth by dual inhibition of PDE5 and 10. *Oncotarget*. 2015;6(29):27403-27415. doi:10.18632/oncotarget.4741
95. Pereira GC, Sanchez L, Schaughency PM, et al. Properties of LINE-1 proteins and repeat element expression in the context of amyotrophic lateral sclerosis. *Mobile DNA*. 2018;9(1):35. doi:10.1186/s13100-018-0138-z
96. Mita P, Wudzinska A, Sun X, et al. LINE-1 protein localization and functional dynamics during the cell cycle. *eLife*. 2018;7. doi:10.7554/eLife.30058
97. Newton JC, Naik MT, Li GY, et al. Phase separation of the LINE-1 ORF1 protein is mediated by the N-terminus and coiled-coil domain. *Biophysical journal*. 2021;120(11):2181-2191. doi:10.1016/j.bpj.2021.03.028
98. Linkert M, Rueden CT, Allan C, et al. Metadata matters: access to image data in the real world. *J Cell Biol*. 2010;189(5):777-782. doi:10.1083/jcb.201004104
99. Ollion J, Cochennec J, Loll F, Escudé C, Boudier T. TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization. *Bioinformatics (Oxford, England)*. 2013;29(14):1840-1841. doi:10.1093/bioinformatics/btt276
100. Stringer C, Wang T, Michaelos M, Pachitariu M. Cellpose: a generalist algorithm for cellular segmentation. *Nat Methods*. 2021;18(1):100-106. doi:10.1038/s41592-020-01018-x
101. Wang Q, Ding SL, Li Y, et al. The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. *Cell*. 2020;181(4):936-953.e20. doi:10.1016/j.cell.2020.04.007
102. Chiaruttini N, Burri O, Haub P, Guiet R, Sordet-Dessimoz J, Seitz A. An Open-Source Whole Slide Image Registration Workflow at Cellular Precision Using Fiji, QuPath and Elastix. *Front Comput Sci*. 2022;3:780026. doi:10.3389/fcomp.2021.780026
103. Bankhead P, Loughrey MB, Fernández JA, et al. QuPath: Open source software for digital pathology image analysis. *Sci Rep*. 2017;7(1):16878. doi:10.1038/s41598-017-17204-5
104. Schmidt U, Weigert M, Broaddus C, Myers G. Cell Detection with Star-Convex Polygons. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, eds. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*. Lecture Notes in Computer Science. Springer International Publishing; 2018:265-273. doi:10.1007/978-3-030-00934-2_30
105. McKinney W. Data Structures for Statistical Computing in Python. In: ; 2010:56-61. doi:10.25080/Majora-92bf1922-00a
106. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and

1544 better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res*.
1545 2019;47(8):e47. doi:10.1093/nar/gkz114
1546 107. Valot B, Langella O, Nano E, Zivy M. MassChroQ: a versatile tool for mass
1547 spectrometry quantification. *Proteomics*. 2011;11(17):3572-3577.
1548 doi:10.1002/pmic.201100120
1549 108. Perez-Riverol Y, Bai J, Bandla C, et al. The PRIDE database resources in 2022: a hub
1550 for mass spectrometry-based proteomics evidences. *Nucleic Acids Res*. 2022;50(D1):D543-
1551 D552. doi:10.1093/nar/gkab1038
1552 109. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for
1553 integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498-
1554 2504. doi:10.1101/gr.1239303
1555

1556

1557

Fig. 1

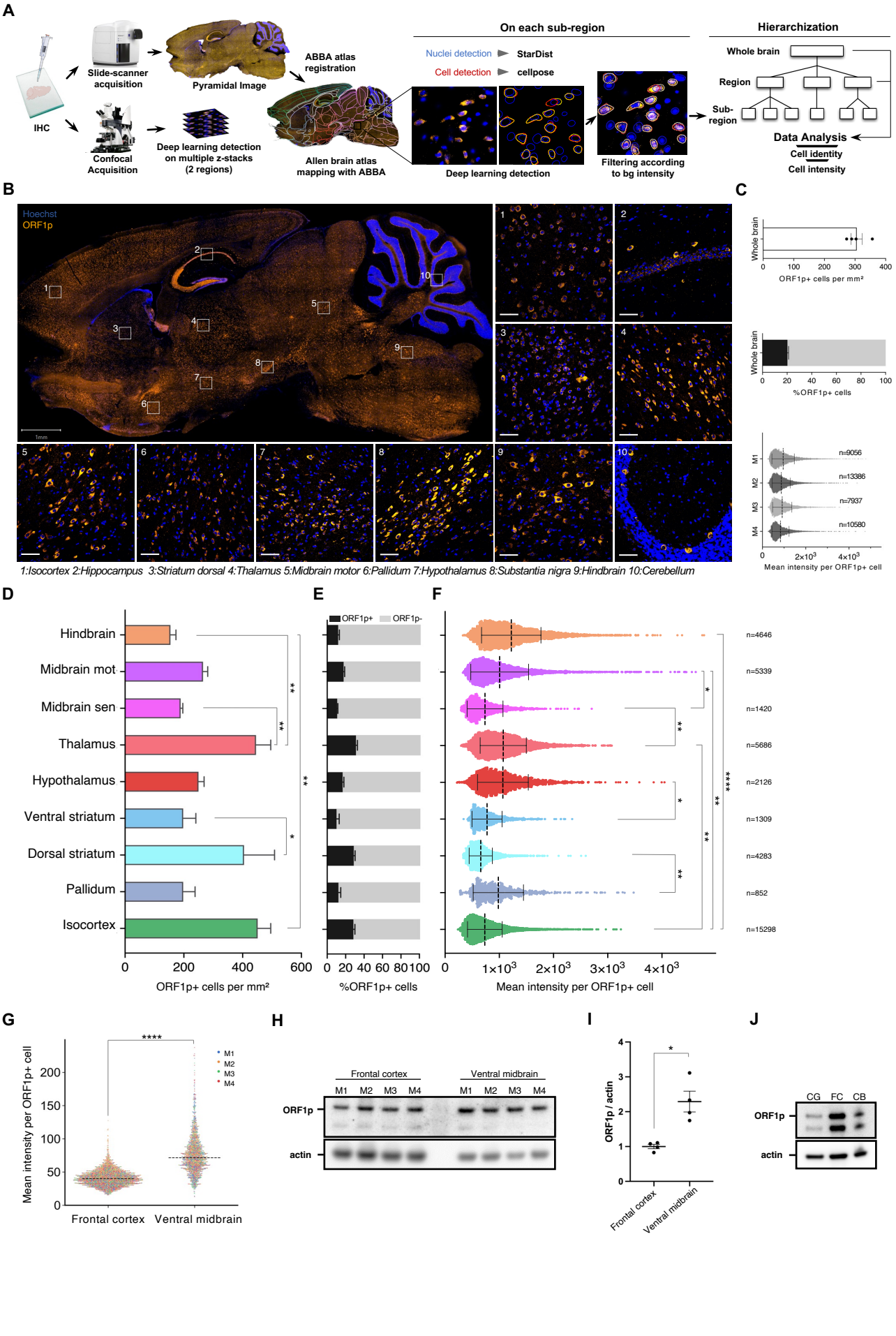

A
IHC
Slide-scanner acquisition
Pyramidal Image
ABBA atlas registration
Confocal Acquisition
Deep learning detection on multiple z-stacks (2 regions)
Allen brain atlas mapping with ABBA
On each sub-region
Nuclei detection
StarDist
Cell detection
cellpose
Deep learning detection
Filtering according to 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

Fig. 2

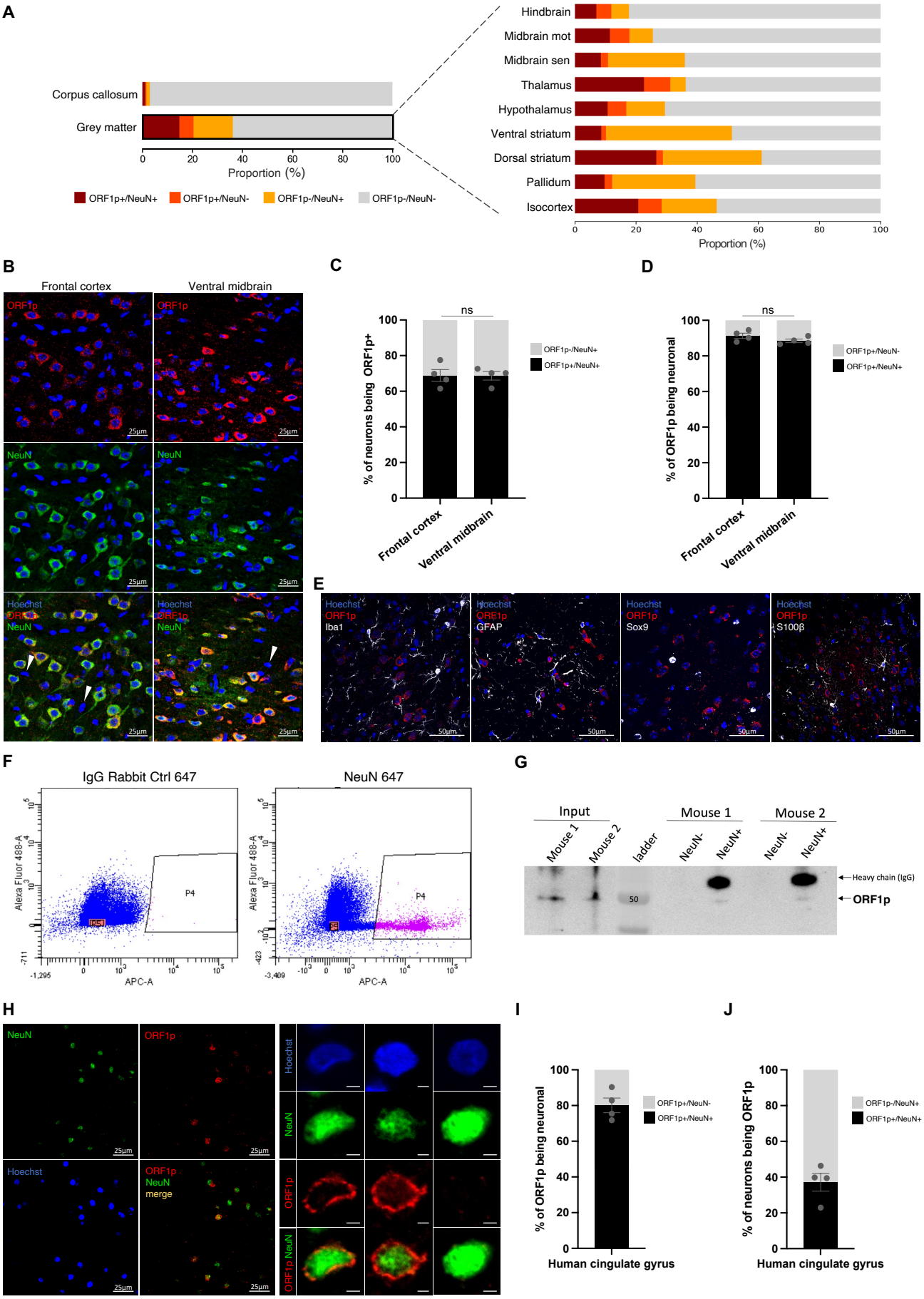

A
Corpus callosum
Grey matter
0
20
40
60
80
100
Proportion (%)
ORF1p+/NeuN+
ORF1p+/NeuN-
ORF1p-/NeuN+
ORF1p-/NeuN-
Hindbrain
Midbrain mot
Midbrain sen
Thalamus
Hypothalamus
Ventral striatum
Dorsal striatum
Pallidum
Isocortex
Proportion (%)
B
Frontal cortex
Ventral midbrain
ORF1p
NeuN
Hoechst
25µm
C
ns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

Fig. 3

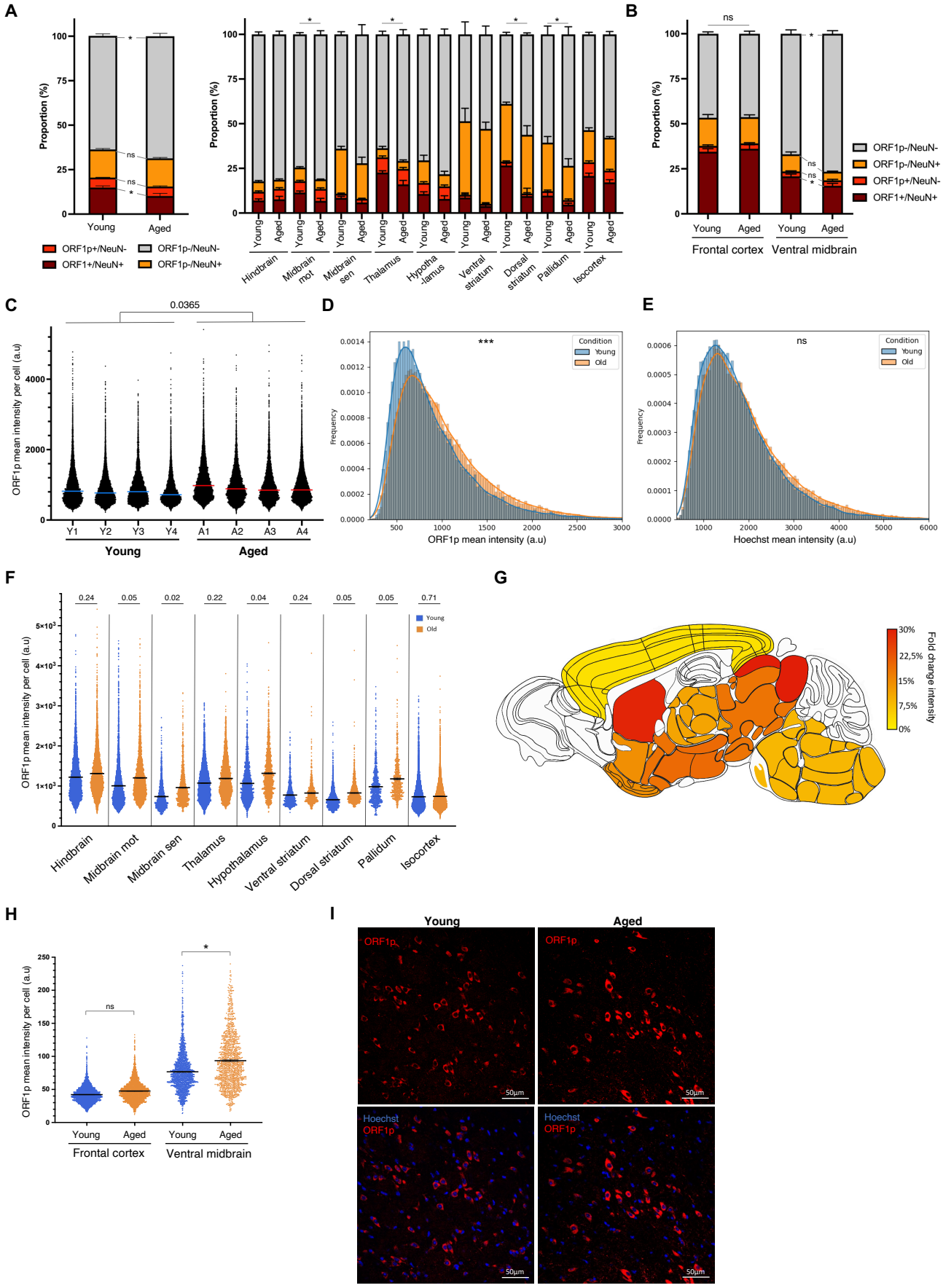

A
Proportion (%)
Young
Aged
ns
*
ORF1p+/NeuN-
ORF1+/NeuN+
ORF1p-/NeuN-
ORF1p-/NeuN+
Hindbrain
Midbrain mot
Midbrain sen
Thalamus
Hypothalamus
Ventral striatum
Dorsal striatum
Pallidum
Isocortex
B
Frontal cortex
Ventral midbrain
C
0.0365
ORF1p mean intensity per cell (a.u)
Y1
Y2
Y3
Y4
A1
A2
A3
A4
D
***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

Fig. 4

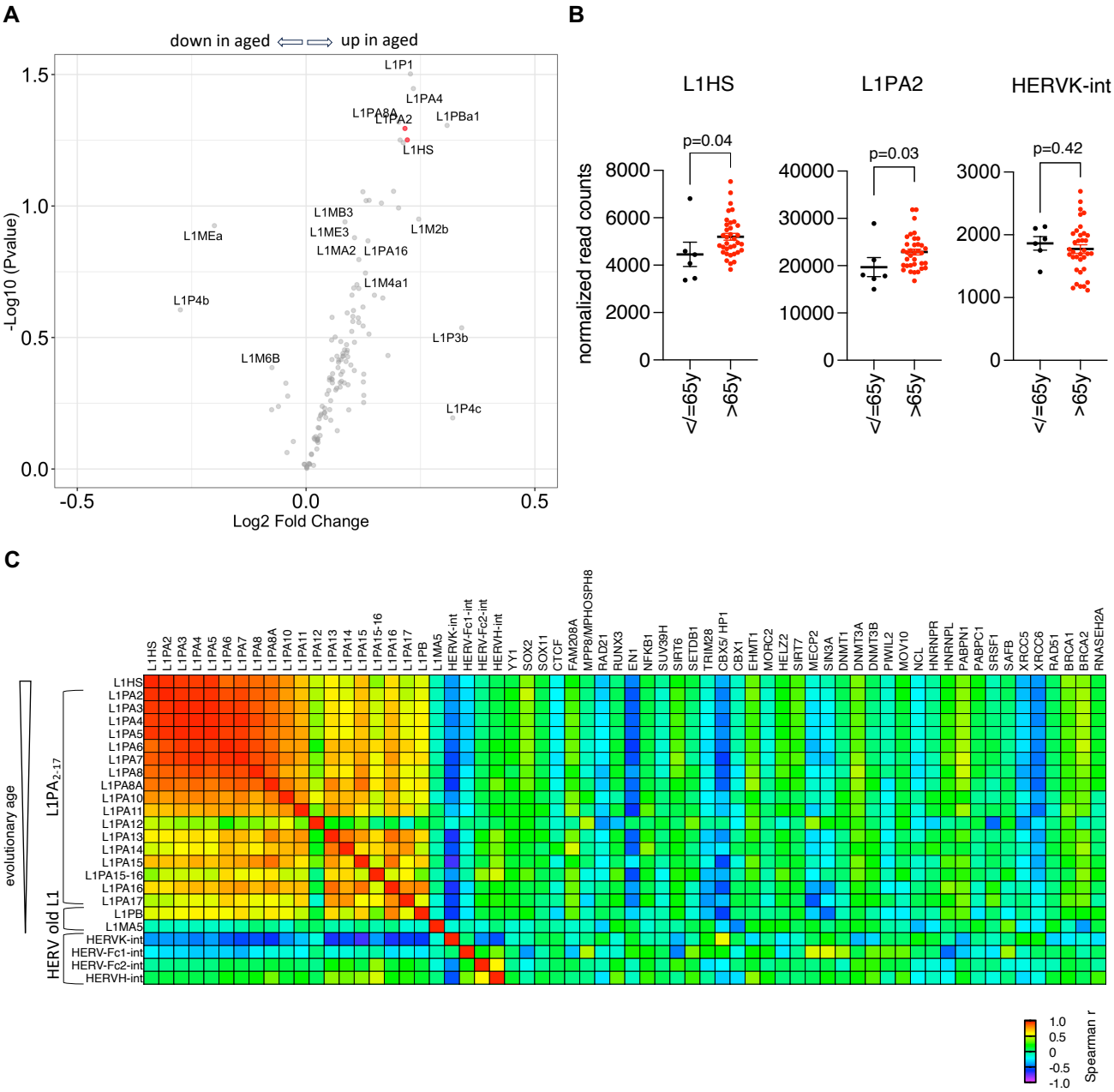

A
down in aged
up in aged
-Log10 (Pvalue)
Log2 Fold Change
L1P1
L1PA4
L1PA8A
L1PA2
L1PBa1
L1HS
L1MB3
L1MEa
L1ME3
L1M2b
L1MA2
L1PA16
L1M4a1
L1P4b
L1P3b
L1M6B
L1P4c
B
L1HS
L1PA2
HERVK-int
p=0.04
p=0.03
p=0.42
normalized read counts
</=65y
>65y
C
evolutionary age
L1PA2-17
L1
HERV old L1
Spearman r

Fig. 5

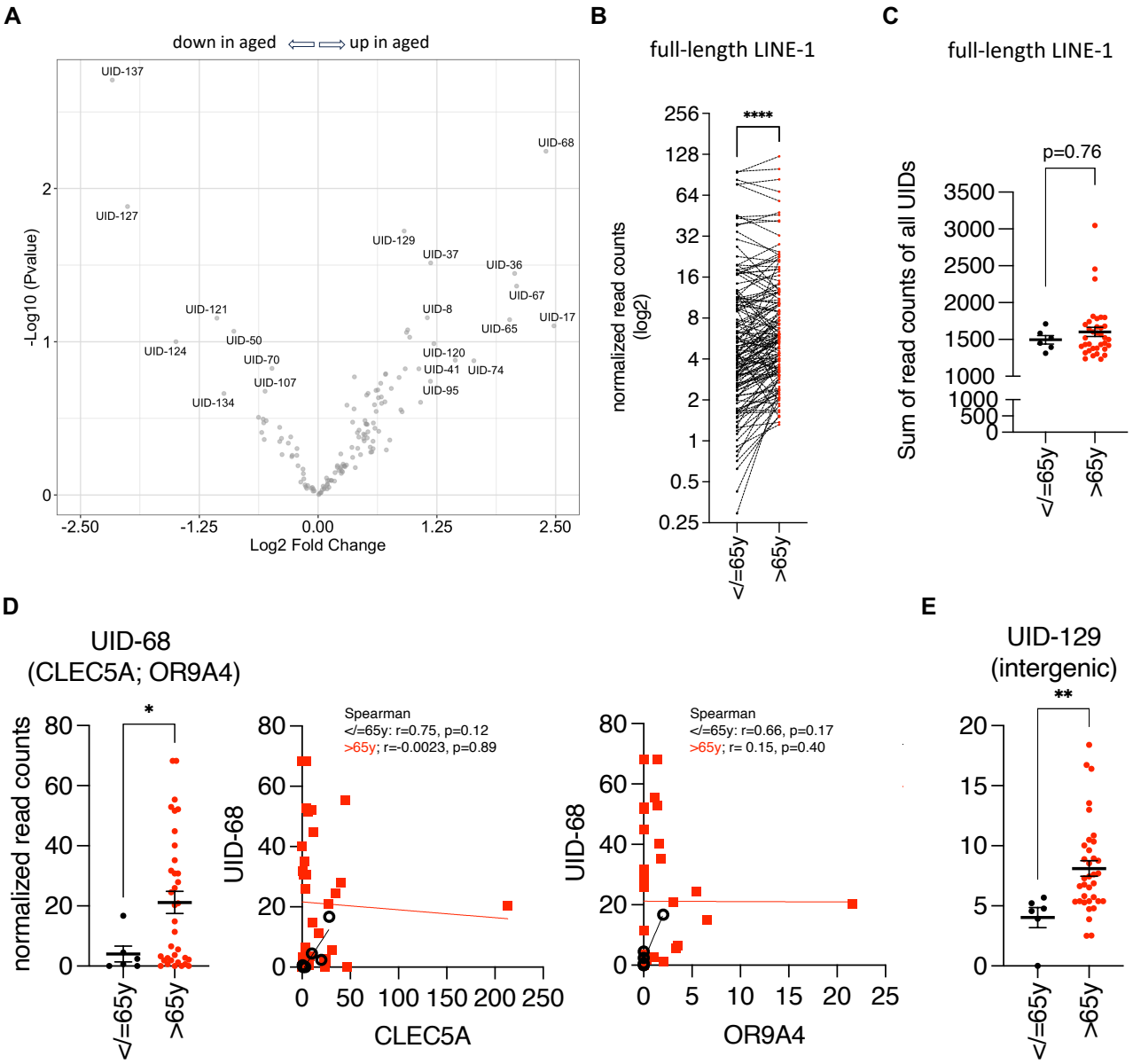
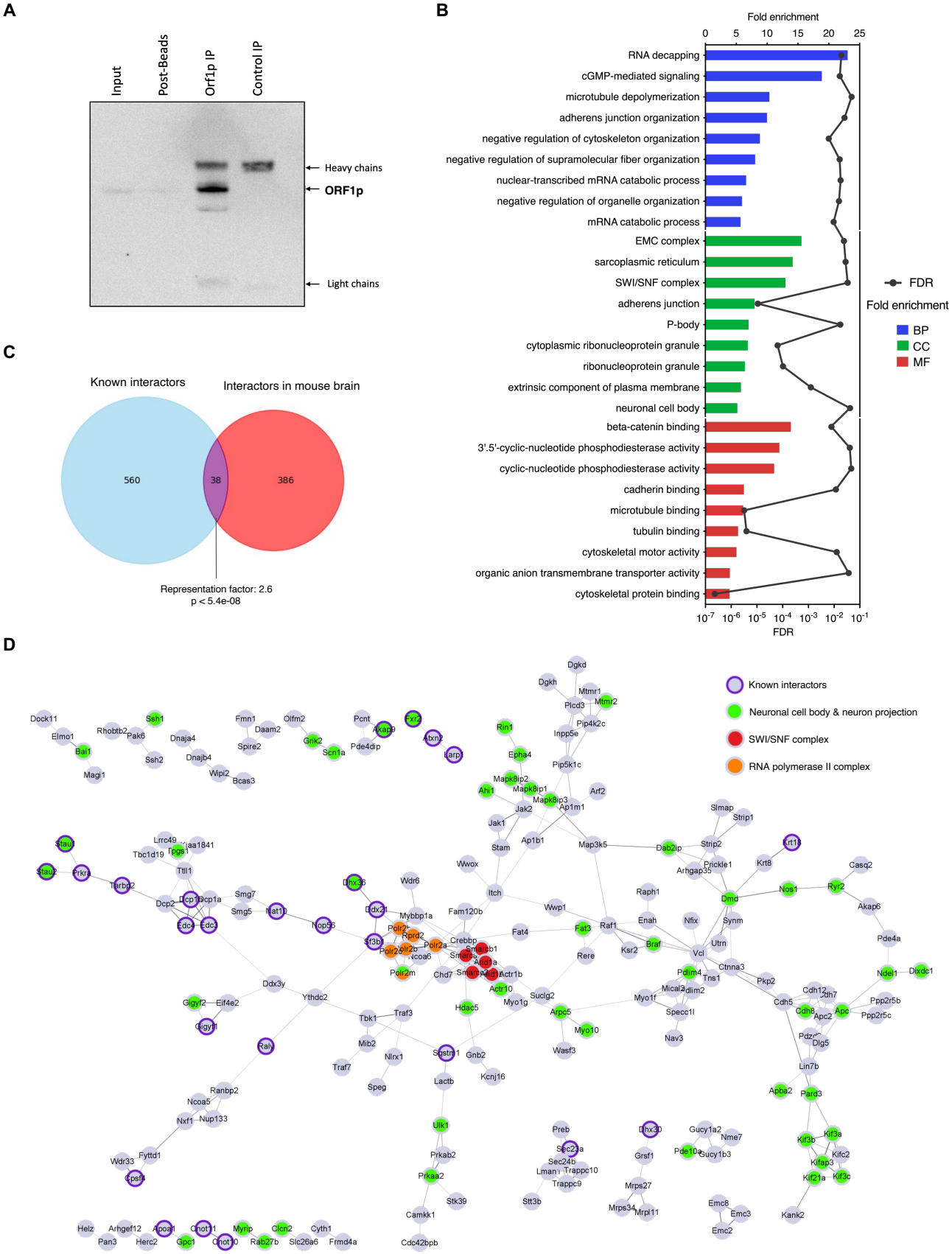
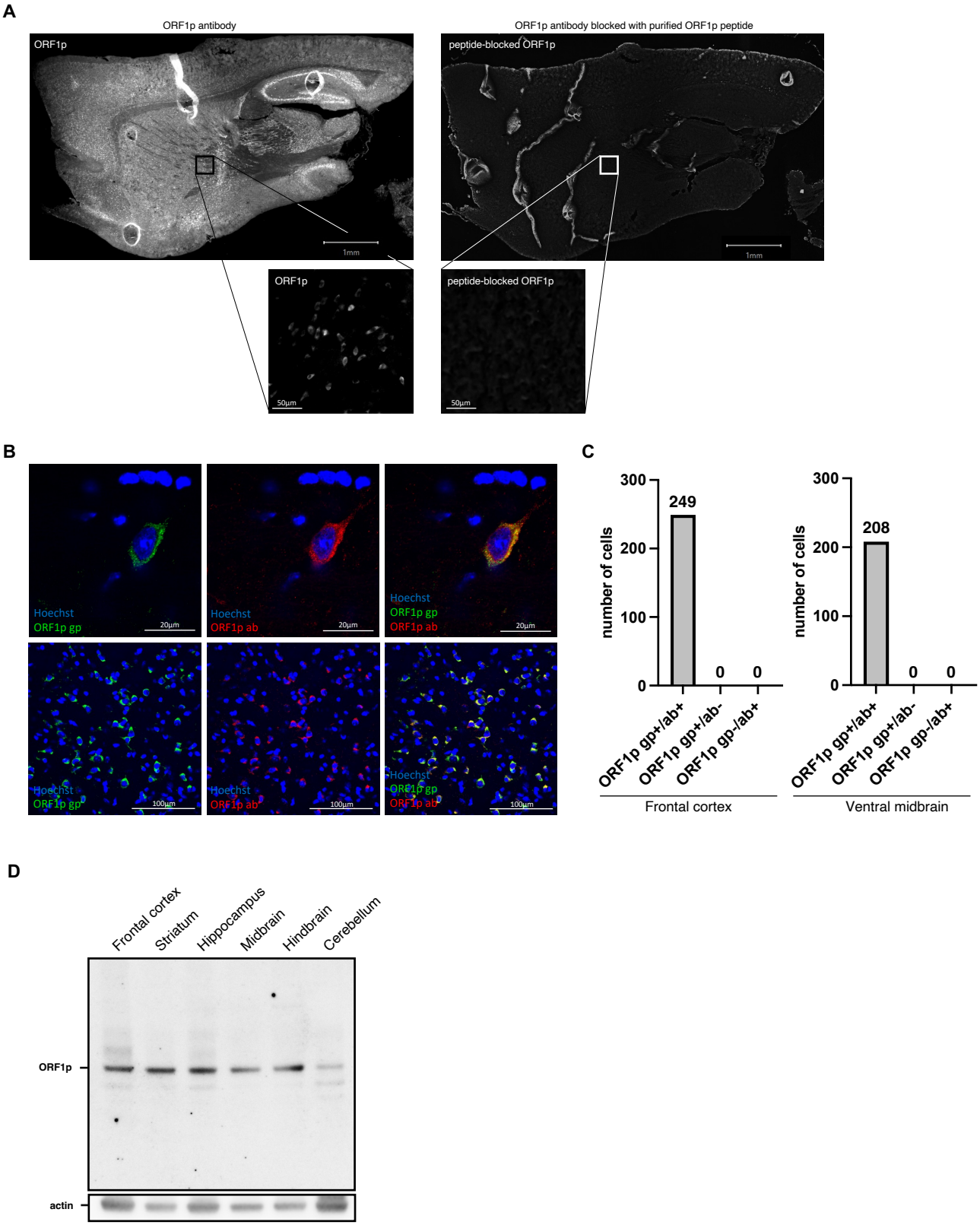

A
down in aged ⟸⟹ up in aged
-Log10 (Pvalue)
Log2 Fold Change
UID-137
UID-68
UID-127
UID-129
UID-37
UID-36
UID-67
UID-121
UID-8
UID-17
UID-65
UID-50
UID-124
UID-70
UID-120
UID-41
UID-74
UID-107
UID-95
UID-134
B
full-length LINE-1
normalized read counts (log2)
****
</=65y
>65y
C
full-length LINE-1
Sum of read counts of all UIDs
p=0.76
D
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
E
UID-129
(intergenic)
**

Fig. 6

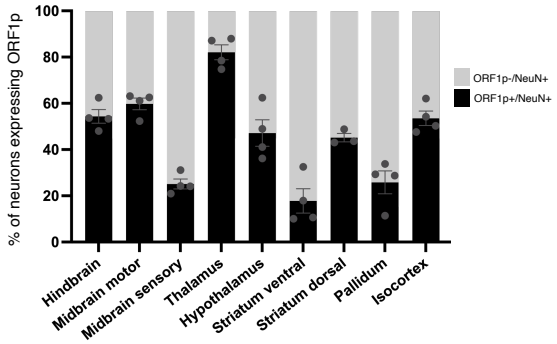

A
Input
Post-Beads
Orf1p IP
Control IP
Heavy chains
ORF1p
Light chains
B
Fold enrichment
0
5
10
15
20
25
RNA decapping
cGMP-mediated signaling
microtubule depolymerization
adherens junction organization
negative regulation of cytoskeleton organization
negative regulation of supramolecular fiber organization
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
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 Fig. 1

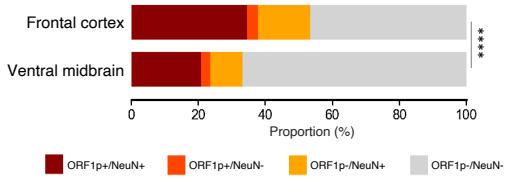

A
ORF1p antibody
ORF1p antibody blocked with purified ORF1p peptide
ORF1p
peptide-blocked ORF1p
1mm
1mm
ORF1p
peptide-blocked ORF1p
50µm
50µm
B
Hoechst
ORF1p gp
20µm
Hoechst
ORF1p ab
20µm
Hoechst
ORF1p gp
ORF1p ab
20µm
Hoechst
ORF1p gp
100µm
Hoechst
ORF1p ab
100µm
Hoechst
ORF1p gp
ORF1p ab
100µm
C
number of cells
300
200
100
0
249
0
0
ORF1p gp+/ab+
ORF1p gp+/ab-
ORF1p gp-/ab+
Frontal cortex
number of cells
300
200
100
0
208
0
0
ORF1p gp+/ab+
ORF1p gp+/ab-
ORF1p gp-/ab+
Ventral midbrain
D
Frontal cortex
Striatum
Hippocampus
Midbrain
Hindbrain
Cerebellum
ORF1p
actin

Suppl Fig. 2

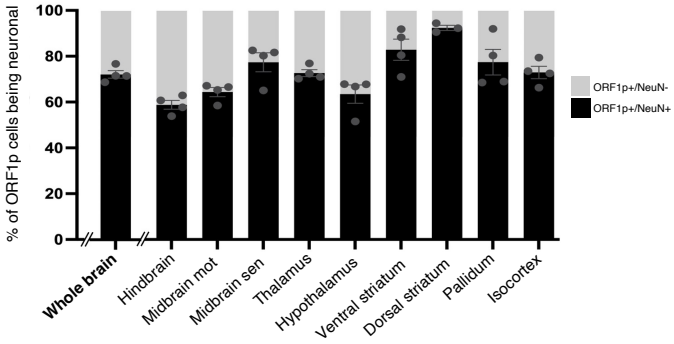
A



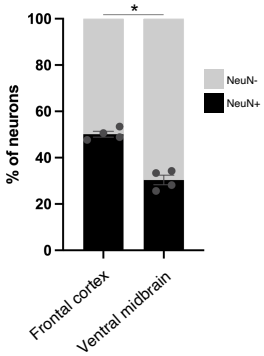
B



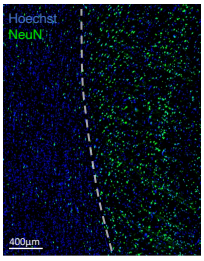
C



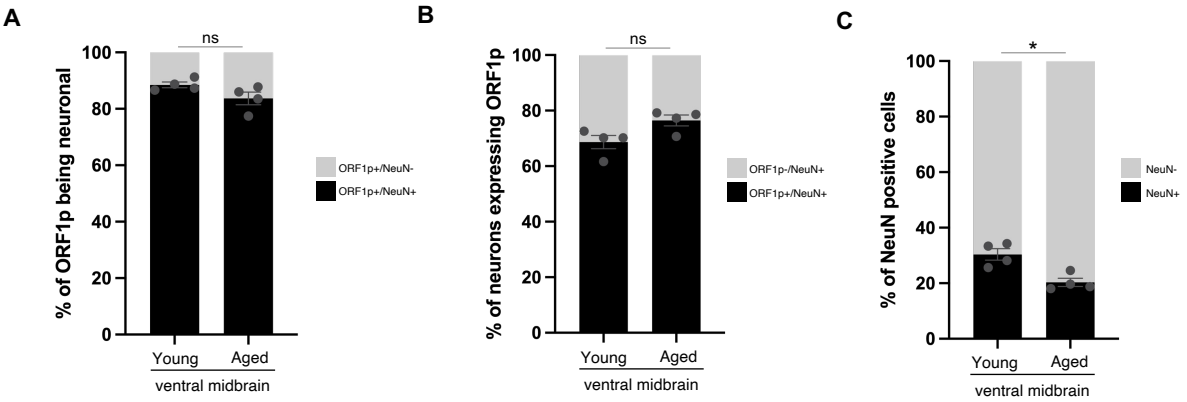
D



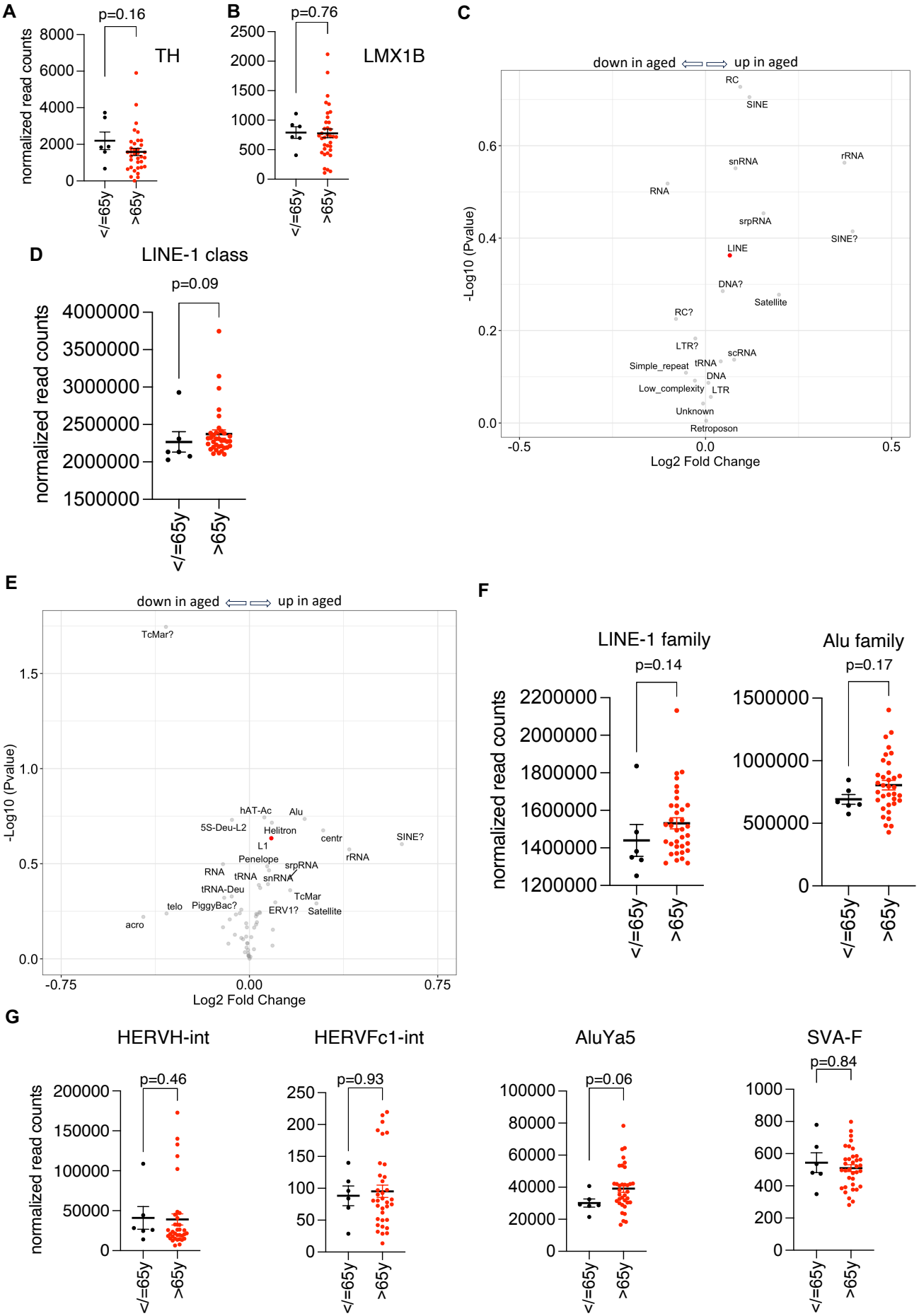
E



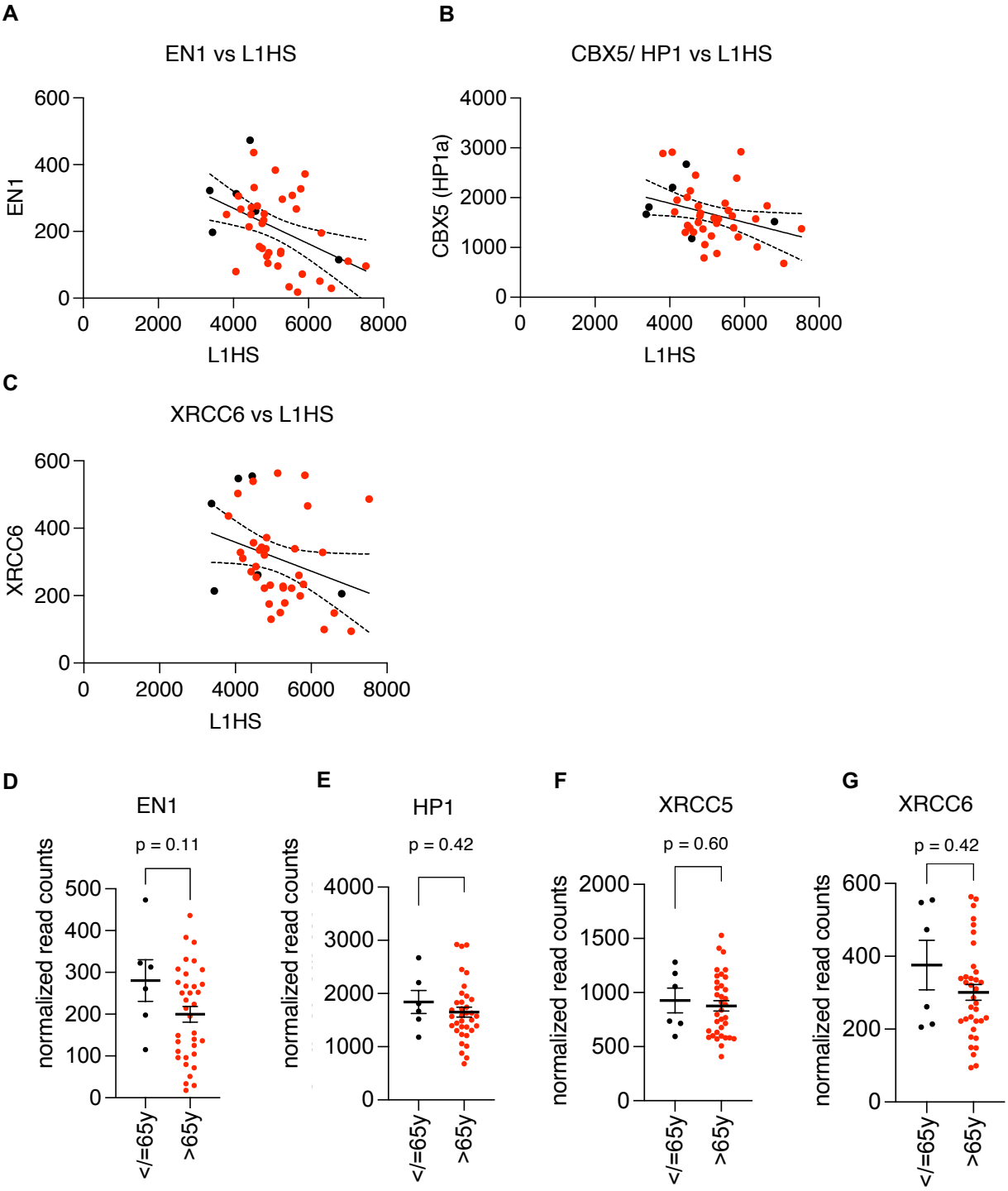
Suppl Fig. 3


A
ns
% of ORF1p being neuronal
100
80
60
40
20
0
Young
Aged
ventral midbrain
ORF1p+/NeuN-
ORF1p+/NeuN+
B
ns
% of neurons expressing ORF1p
100
80
60
40
20
0
Young
Aged
ventral midbrain
ORF1p-/NeuN+
ORF1p+/NeuN+
C
*
% of NeuN positive cells
100
80
60
40
20
0
Young
Aged
ventral midbrain
NeuN-
NeuN+

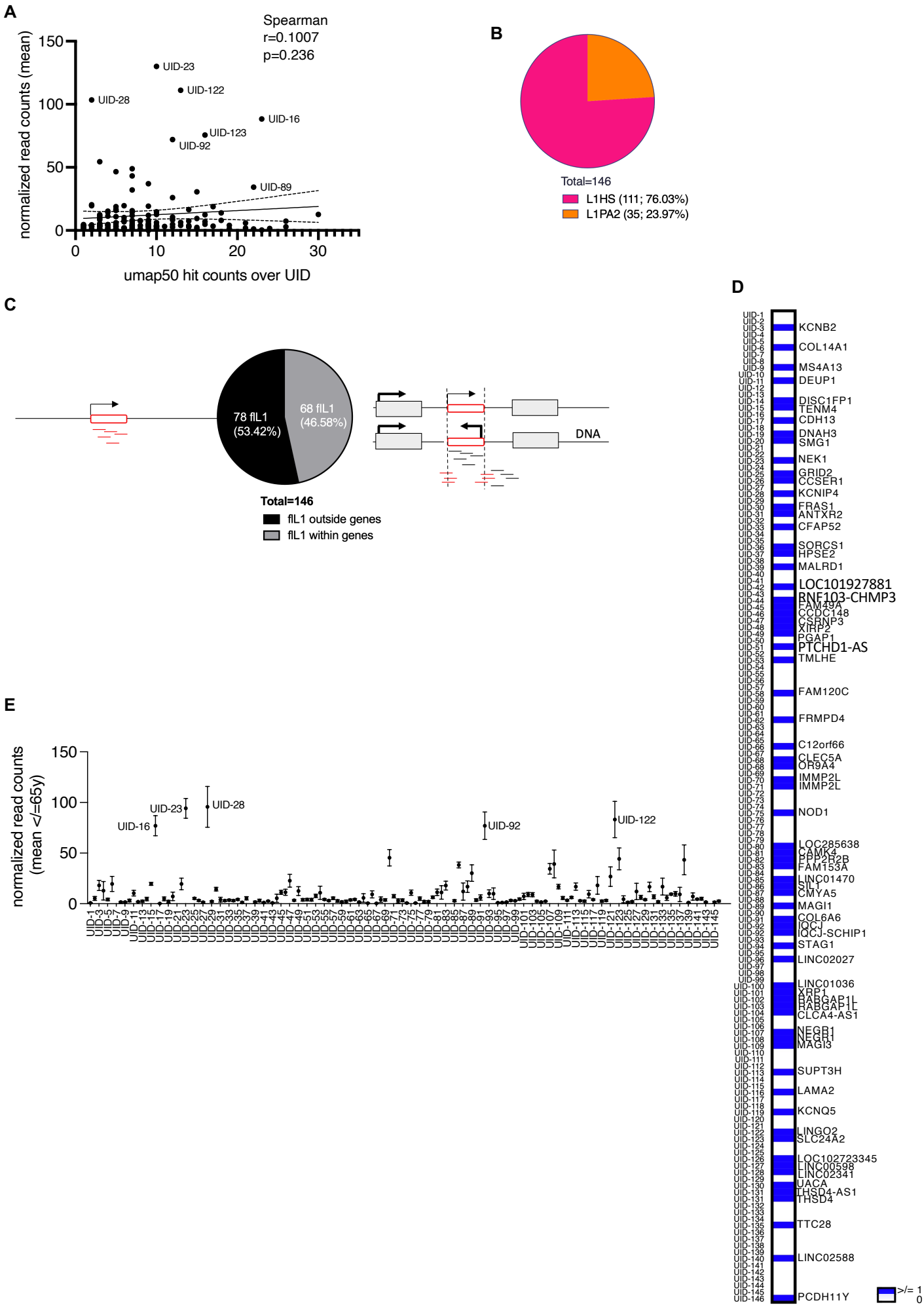
Suppl Fig. 4


A
p=0.16
TH
normalized read counts
8000
6000
4000
2000
0
</=65y
>65y
B
p=0.76
LMX1B
2500
2000
1500
1000
500
0
C
down in aged ⟸⟹ up in aged
-Log10 (Pvalue)
Log2 Fold Change
RC
SINE
snRNA
rRNA
RNA
srpRNA
SINE?
LINE
DNA?
Satellite
RC?
LTR?
scRNA
Simple_repeat
tRNA
DNA
Low_complexity
LTR
Unknown
Retroposon
D
LINE-1 class
p=0.09
4000000
3500000
3000000
2500000
2000000
1500000
E
down in aged ⟸⟹ up in aged
TcMar?
hAT-Ac
Alu
5S-Deu-L2
Helitron
centr
SINE?
L1
Penelope
rRNA
srpRNA
RNA
tRNA
snRNA
tRNA-Deu
TcMar
telo
PiggyBac?
ERV1?
Satellite
acro
F
LINE-1 family
p=0.14
Alu family
p=0.17
2200000
2000000
1800000
1600000
1400000
1200000
1500000
1000000
500000
0
G
HERVH-int
p=0.46
HERVFc1-int
p=0.93
AluYa5
p=0.06
SVA-F
p=0.84
200000
150000
100000
50000
250
200
150
100
50
100000
80000
60000
40000
20000
1000
800
600
400
200

Suppl Fig. 5

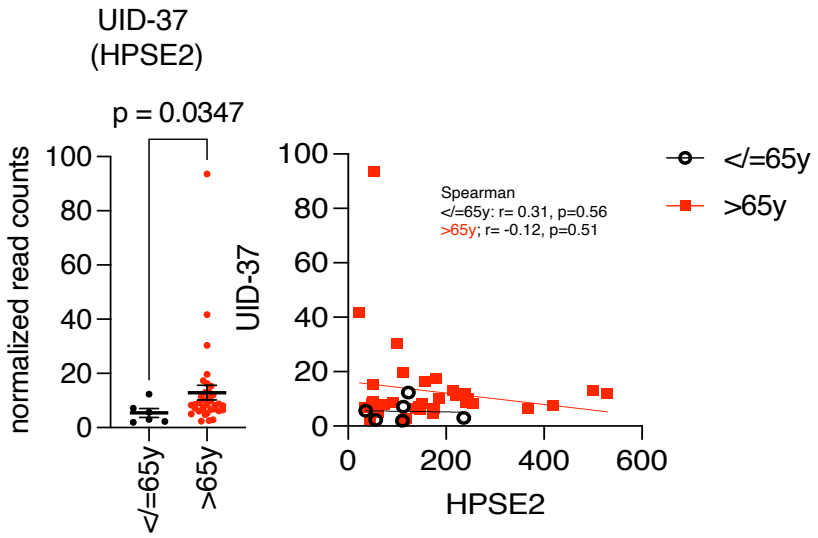

A
EN1 vs L1HS
EN1
L1HS
B
CBX5/ HP1 vs L1HS
CBX5 (HP1a)
L1HS
C
XRCC6 vs L1HS
XRCC6
L1HS
D
EN1
p = 0.11
normalized read counts
</=65y
>65y
E
HP1
p = 0.42
normalized read counts
</=65y
>65y
F
XRCC5
p = 0.60
normalized read counts
</=65y
>65y
G
XRCC6
p = 0.42
normalized read counts
</=65y
>65y

Suppl Fig. 6

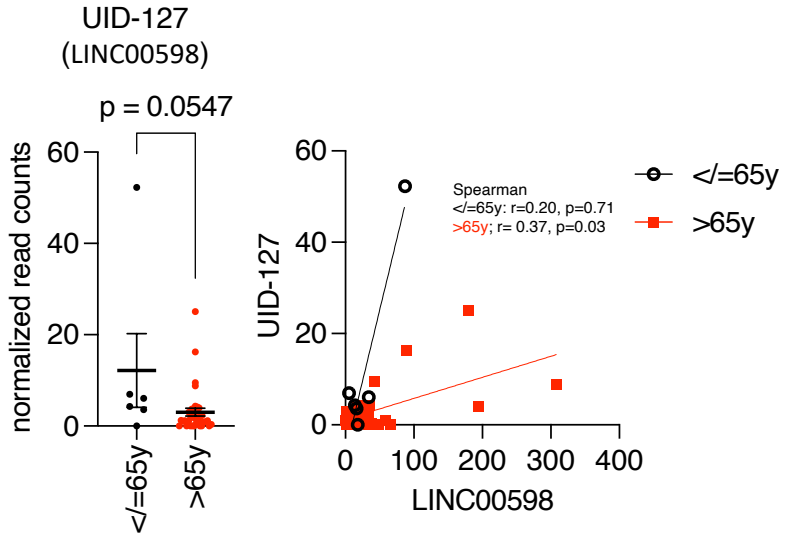
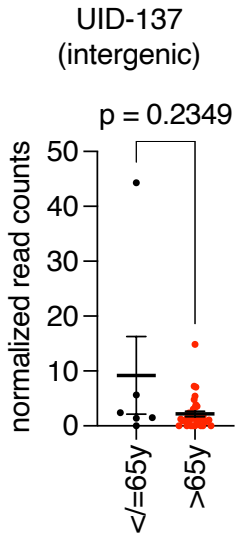

A
Spearman
r=0.1007
p=0.236
normalized read counts (mean)
umap50 hit counts over UID
UID-23
UID-122
UID-28
UID-16
UID-123
UID-92
UID-89
B
Total=146
L1HS (111; 76.03%)
L1PA2 (35; 23.97%)
C
78 flL1
(53.42%)
68 flL1
(46.58%)
DNA
Total=146
flL1 outside genes
flL1 within genes
D
KCNB2
COL14A1
MS4A13
DEUP1
DISC1FP1
TENM4
CDH13
DNAH3
SMG1
NEK1
GRID2
CCSER1
KCNIP4
FRAS1
ANTXR2
CFAP52
SORCS1
HPSE2
MALRD1
LOC101927881
RNF103-CHMP3
FAM49A
CCDC148
CSRNP3
XIRP2
PGAP1
PTCHD1-AS
TMLHE
FAM120C
FRMPD4
C12orf66
CLEC5A
OR9A4
IMMP2L
IMMP2L
NOD1
LOC285638
CAMK4
PPP2R2B
FAM153A
LINC01470
SIL1
CMYA5
MAGI1
COL6A6
IQCJ
IQCJ-SCHIP1
STAG1
LINC02027
LINC01036
XRP1
RABGAP1L
RABGAP1L
CLCA4-AS1
NEGR1
NEGR1
MAGI3
SUPT3H
LAMA2
KCNQ5
LINGO2
SLC24A2
LOC102723345
LINC00598
LINC02341
UACA
THSD4-AS1
THSD4
TTC28
LINC02588
PCDH11Y
>/= 1
0
E
normalized read counts
(mean </=65y)
UID-16
UID-23
UID-28
UID-92
UID-122

Suppl Fig. 7

A



B



C