

1 **Transposable elements contribute to the spatiotemporal microRNA landscape in human**
2 **brain development**

3 Christopher J. Playfoot¹, Shaoline Sheppard¹, Evarist Planet¹ and Didier Trono^{1,2}

4 ¹School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland

5 ²Corresponding author: Didier.Trono@epfl.ch

6 Short title: TE-embedded miRNAs in brain development

7 **Abstract**

8 Transposable elements (TEs) contribute to the evolution of gene regulatory networks and are
9 dynamically expressed throughout human brain development and disease. One gene regulatory
10 mechanism influenced by TEs is the miRNA system of post-transcriptional control. miRNA sequences
11 frequently overlap TE loci and this miRNA expression landscape is crucial for control of gene expression
12 in adult brain and different cellular contexts. Despite this, a thorough investigation of the
13 spatiotemporal expression of TE-embedded miRNAs in human brain development is lacking. Here, we
14 identify a spatiotemporally dynamic TE-embedded miRNA expression landscape between childhood
15 and adolescent stages of human brain development. These miRNAs sometimes arise from two apposed
16 TEs of the same subfamily, such as for L2 or MIR elements, but in the majority of cases stem from solo
17 TEs. They give rise to *in silico* predicted high-confidence pre-miRNA hairpin structures, likely represent
18 functional miRNAs and have predicted genic targets associated with neurogenesis. TE-embedded
19 miRNA expression is distinct in the cerebellum when compared to other brain regions, as has
20 previously been described for gene and TE expression. Furthermore, we detect expression of
21 previously non-annotated TE-embedded miRNAs throughout human brain development, suggestive of
22 a previously undetected miRNA control network. Together, as with non-TE-embedded miRNAs, TE-
23 embedded sequences give rise to spatiotemporally dynamic miRNA expression networks, the
24 implications of which for human brain development constitute extensive avenues of future

25 experimental research. To facilitate interactive exploration of these spatiotemporal miRNA expression
26 dynamics, we provide the “Brain miRTEplorer” web application freely accessible for the community.

27 **Introduction**

28 Transposable elements (TEs) account for around half of the human genome and have contributed to
29 the evolution of gene regulatory networks (Chuong et al. 2013, 2016, 2017; Garcia-Perez et al. 2016;
30 Pontis et al. 2019; Turelli et al. 2020; Playfoot et al. 2021). The majority of TEs have lost their capacity
31 to ‘copy and paste’ to new locations around the genome, instead being co-opted by the host organism
32 to perform a plethora of regulatory homeostatic functions during normal development (Elbarbary et
33 al. 2016; Chuong et al. 2017). One post-transcriptional regulatory mechanism in which TE-embedded
34 sequences have been co-opted is the microRNA (miRNA) system (Smalheiser and Torvik 2005;
35 Piriyaongsa et al. 2007; Roberts et al. 2014). Computational and experimental studies have shown
36 different classes of TEs (LINE, SINE and LTR) can act as functional sources of miRNA in different cellular
37 models. However, limited information exists for primary tissues, especially for tightly regulated
38 spatiotemporal developmental processes such as human brain development (Piriyaongsa et al. 2007;
39 Piriyaongsa and Jordan 2007; Ding et al. 2010; Frankel et al. 2014; Roberts et al. 2014; Spengler et al.
40 2014; Petri et al. 2019).

41 Recent studies in a small number of adult brains have highlighted the roles of TE-embedded miRNAs
42 from the L2 family. These are functional in neurotypical adult brains and are differentially expressed
43 in glioblastoma (Skalsky and Cullen 2011; Petri et al. 2019). Furthermore, miRNAs have critical roles in
44 mammalian neuronal homeostasis, highlighting the fundamental nature of miRNAs in neurogenesis,
45 alongside diverse roles in neurological disease and human evolution (Cao et al. 2007; Somel et al. 2011;
46 Qureshi and Mehler 2012; Petri et al. 2014; Topol et al. 2016; Sambandan et al. 2017; Jużwik et al.
47 2019; Woods and Van Vactor 2021). miRNAs are spatially and temporally expressed in the developing
48 human brain from birth to adolescence, however the contribution of TE-embedded sequences to this
49 process has never been investigated (Ziats and Rennert 2014). Indeed, the years proceeding birth and

50 throughout childhood represent a crucial window in human brain development, characterized by
51 extensive changes in size, cellular composition and functional processes such as synaptogenesis,
52 myelination and synaptic pruning (Silbereis et al. 2016; Dyck and Morrow 2017).

53 We therefore aimed to determine the prevalence of spatiotemporally expressed, annotated TE-
54 embedded miRNAs in the developing human brain by re-analysis of small RNA-seq data available from
55 the BrainSpan Atlas of the Developing Human Brain from one year old to 19-year-old brains (Miller et
56 al. 2014; Li et al. 2018). We computationally uncover dynamic spatiotemporal expression of numerous
57 annotated TE-embedded miRNAs and a small number of previously undetected novel putative TE-
58 embedded miRNAs, suggesting TE-sequence co-option as miRNAs may play a role in this important
59 neurodevelopmental window. We provide the “Brain miRTEplorer” web application to facilitate
60 interactive exploration of both annotated TE-embedded and non-TE-embedded miRNA
61 spatiotemporal expression data, freely accessible for the community at
62 <https://tronoapps.epfl.ch/BrainmiRTEplorer/>.

63 **Results**

64 **TEs contribute to the annotated miRNA transcriptional landscape in the human brain**

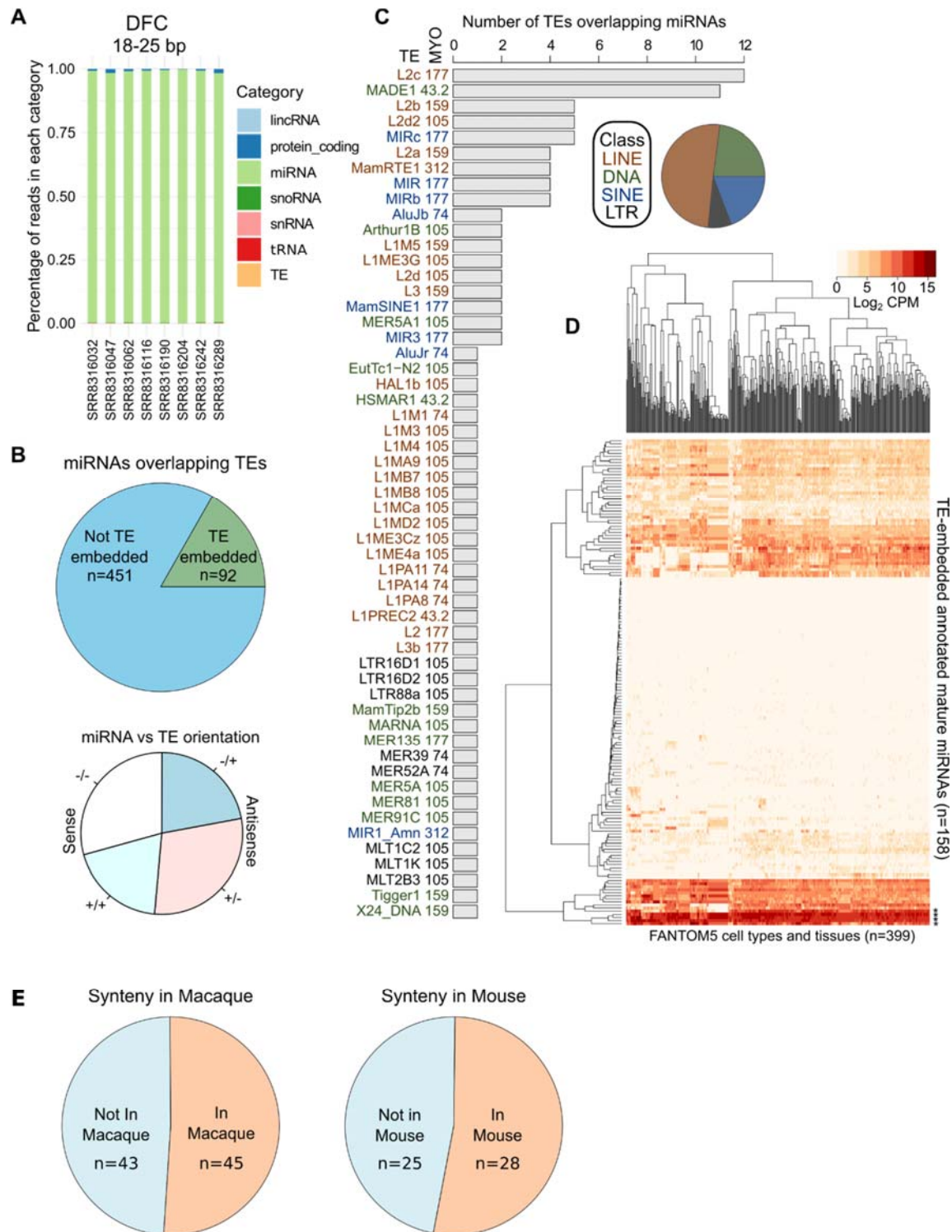
65 To determine spatiotemporal, small RNA expression in postnatal human brain development, we
66 analyzed small RNA-seq data from 174 samples from one year to 20 years of age, encompassing 16
67 different brain regions, from 16 donors (9 male and 7 female) available through the BrainSpan Atlas of
68 the Developing Human Brain (Supplemental Fig. S1) (Miller et al. 2014; Li et al. 2018). To enrich for
69 different small RNA moieties, we separated sequencing reads into lengths of 18 - 25bp, 26 – 37bp and
70 38 – 50bp and intersected with Ensembl annotations, miRBase, the GtRNADB database and our
71 modified merged TE RepeatMasker data set (Kozomara and Griffiths-Jones 2014; Chan and Lowe 2016;
72 Pontis et al. 2019; Turelli et al. 2020; Yates et al. 2019; Playfoot et al. 2021). As expected, the different
73 read lengths enriched for annotated miRNAs, tRNAs and snoRNAs respectively (Fig. 1A; Supplemental

74 Fig. S2). By retaining the miRNA derived 18 – 25bp reads, we detected the expression of 543/1871
75 annotated miRNAs (Fig. 1B; Supplemental Table S1 & S2).

76 To determine the overlap of annotated miRNAs with TEs, we intersected their genomic coordinates
77 with those from our curated RepeatMasker data set (Turelli et al. 2020; Playfoot et al. 2021). 17% of
78 annotated miRNAs were derived from TEs, in either sense and antisense orientation to the miRNA and
79 belonged to all known classes of elements, with representatives from various subfamilies and
80 evolutionary ages (Fig. 1B & C; Supplemental Table S1). Only 36/543 expressed miRNAs were
81 annotated as mirtrons (Da Fonseca et al. 2019), none of which were TE-embedded. This indicates that
82 TEs do not contribute to mirtrons in this context.

83 L2 family members of 105 - 177 million-year-old (MYO) contributed the most to annotated detectably
84 expressed miRNAs in the child and adolescent brain (Fig. 1C), with detection of all L2-embedded,
85 annotated miRNAs previously noted in adult brain and glioblastoma (Piriyapongsa et al. 2007; Petri et
86 al. 2019), pointing to their likely roles in earlier stages of brain development (Supplemental Table S1).
87 The previously described 43.2 MYO MADE1 elements and the 177 MYO MIR family elements also
88 heavily contributed to expressed miRNAs (Fig. 1C) (Piriyapongsa and Jordan 2007; Shao et al. 2010;
89 Borchert et al. 2011; Spengler et al. 2014).

90 To determine the potential importance of TE-embedded miRNAs versus non-TE-embedded miRNAs we
91 plotted the mean expression of all miRNAs in descending order (Supplemental Fig. S3A). One quarter
92 (24/94) of TE-embedded miRNAs were in the top 200 most expressed miRNAs (Supplemental Fig. S3A).



93
94 **Figure 1. TEs contribute to annotated miRNAs in the child and adolescent human brain.** (A) Stacked
95 bar chart indicating the percentage of 18-25bp reads overlapping different annotated genomic
96 features for samples from the dorsolateral frontal cortex (DFC). If a TE overlaps an annotated feature
97 (miRNA, tRNA etc.) the feature takes preference. (B) Pie charts indicating the number of miRBase
98 annotated miRNAs overlapping at least one TE (*top*), and their relative orientations (*bottom*). (C) Bar
99 chart indicating the number of TEs overlapping miRBase annotated miRNAs and their class and age in
100 million years old (MYO). (D) Expression in log₂ counts per million+1 (CPM+1) of mature TE-embedded

101 miRNA in 399 cell types and tissues from FANTOM5 (De Rie et al. 2017). * denotes miRNAs highlighted
102 in the text. Samples comprise largely primary cells such as epithelial, fibroblast, endothelial, connective
103 tissue, smooth muscle, immune, neural stem, dendritic and pluripotent stem cells, among others (full
104 list available in De Rie et al., 2017). (E) Pie charts indicating the number of human TE-embedded
105 miRNAs detected in syntenic locations in macaque and mouse miRBase annotations (see also
106 Supplemental Table S3 and methods for detail).

107

108 Four of these were in the top 50 most expressed and have described roles in neurogenesis and glioma
109 (Zhang et al. 2012; Huang et al. 2015; Ruan et al. 2015). Despite this, TE-embedded miRNAs were
110 significantly less expressed than non-TE-embedded miRNAs (Supplemental Fig. S3B), however even TE-
111 embedded miRNAs with low expression levels (e.g., hsa-mir-326_Arthur1B - 529th in the expression
112 list) have roles in neurogenic diseases like glioma (Kefas et al., 2009). This indicates that TE-embedded
113 miRNAs are expressed at generally lower levels than non-TE-embedded miRNAs but still play functional
114 roles in the brain.

115 We next aimed to determine if TE-embedded miRNAs were produced in other cell types and tissues by
116 analyzing miRNA expression data from 399 human samples comprising largely primary cells such as
117 epithelial, fibroblast, endothelial, connective tissue, smooth muscle, immune, neural stem, dendritic
118 and pluripotent stem cells, among others (De Rie et al. 2017). Mature TE-embedded miRNAs were
119 broadly expressed in the majority of cell types, with relatively ubiquitous, high levels for MIRc-
120 embedded hsa-miR-378a-3p, L2d2-embedded hsa-miR-28-3p, L2c-embedded hsa-miR-151a-3p/5p
121 and MamRTE1-embedded hsa-miR-130a-3p and lower expression for other TE-embedded miRNAs (Fig.
122 1D). A similar ubiquitous expression was detected for mature non-TE-embedded miRNAs
123 (Supplemental Fig. S4). Together, these data indicate that a multitude of TE-embedded miRNAs are
124 broadly expressed in the child and adolescent human brain, with appreciable expression in other cell
125 types.

126 To investigate if TE-embedded miRNAs are conserved and in syntenic locations among mammals, we
127 performed a series of lift-overs from the human (hg19) to the macaque (rheMac8) and mouse (mm10)
128 genomes. We next downloaded annotated miRNA coordinates from miRBase for macaque and mouse

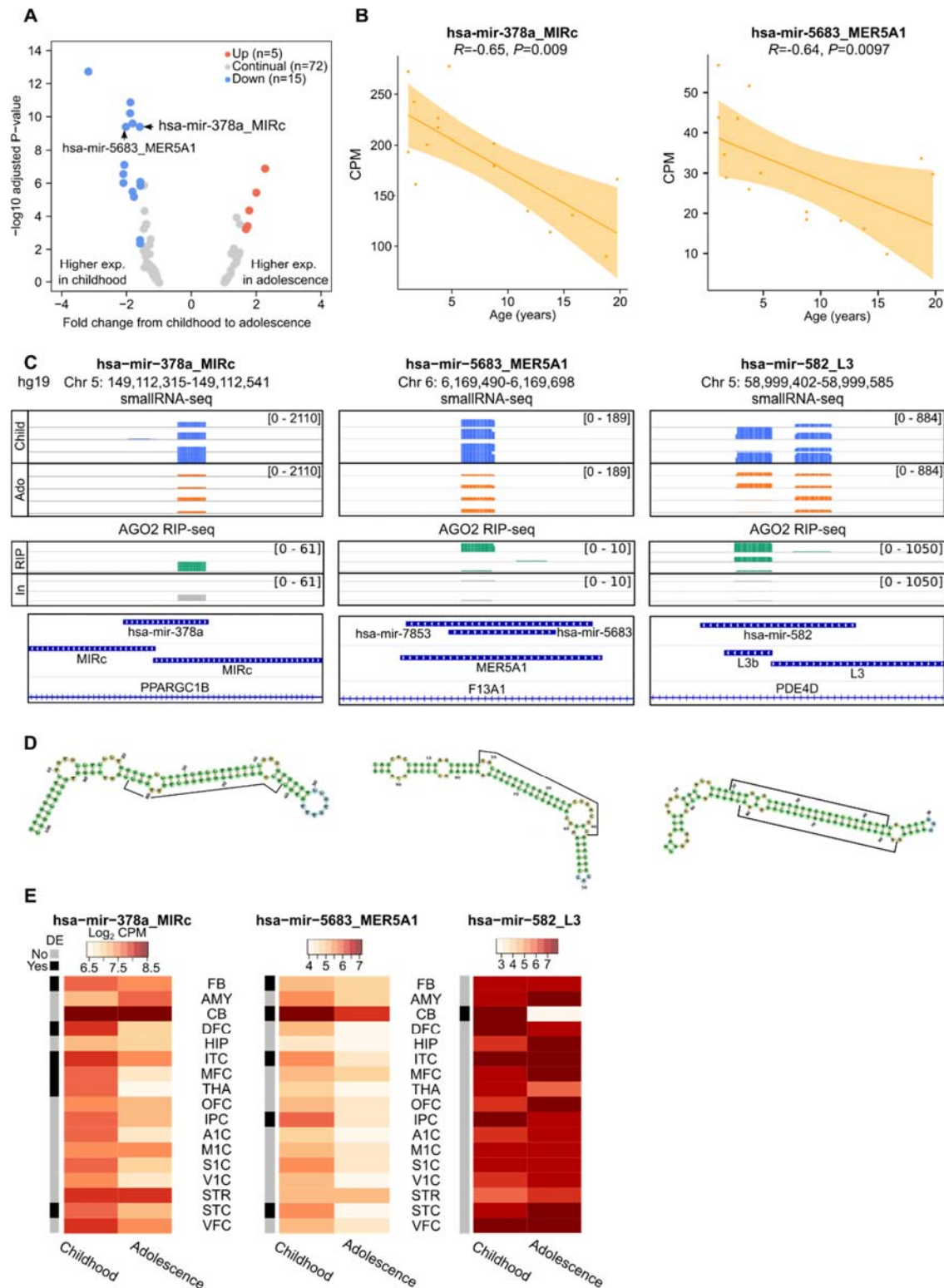
129 (Kozomara and Griffiths-Jones 2014). Of the 92 TE-embedded miRNAs detected in human (Fig. 1B), 88
130 remained after lift-over to macaque, whereas the last four had no syntenic homologue. An intersect
131 of these syntenic coordinates in macaque with the annotated miRNA coordinates from miRBase
132 detected 45/88 TE-embedded miRNA loci which were largely annotated with the same miRNA name
133 (Fig. 1E; Supplemental Table S3). Intersecting these with all TE coordinates from rheMac8 also showed
134 that 40/45 overlapped the same or closely related TE subfamilies in macaque (Supplemental Table S3).
135 In mouse, 53 loci remained when TE-embedded miRNA coordinates in hg19 were lifted over to mm10.
136 Of these, 28/53 overlapped coordinates from miRBase for mm10, again largely with the same name
137 (Fig. 1E; Supplemental Table S3). The subsequent intersect with all TE loci for mouse resulted in 19/28
138 TE-embedded miRNA loci in human also overlapping the same or closely related TE subfamily in mouse
139 (Supplemental Table S3). These were the older, more conserved TE subfamilies such as L2 (105 – 177
140 MYO), MIR (177 MYO) and some DNA elements (105 – 177 MYO) (Supplemental Table S3). Together,
141 these analyses indicate that a large subset of TE-embedded miRNAs expressed in the human brain are
142 conserved in mammalian lineages.

143 **TE-embedded miRNAs exhibit spatiotemporal expression patterns**

144 To investigate the temporal dynamics of TE-embedded miRNAs in brain development, we compared
145 their expression from childhood (1 to 5 years) to adolescence (9 to 20 years) (Supplemental Fig. S1).
146 We initially combined samples of forebrain (FB) origin, representing 124 samples from 16 donors, with
147 66 and 58 samples representing childhood and adolescence respectively (Supplemental Fig. S1). Data
148 was normalized using the trimmed mean of the log expression ratios (TMM) method which excludes
149 the top and bottom expressed miRNAs prior to computing library sizes to ensure very high or low
150 expressed miRNAs did not dominate the library size normalization (Robinson and Oshlack 2010).
151 Counts per million (CPM) values were generated after correcting by the TMM library size. There was
152 no difference in the number of reads, percentage of reads aligned to the genome or assigned to a
153 feature between childhood and adolescence categories (Supplemental Fig. S5A-C).

154 16% and 5.5% of TE-embedded miRNAs were significantly more highly expressed in childhood or
155 adolescence respectively, whilst 78% were continually expressed (Fig. 2A). Differentially expressed
156 miRNAs, again represented a suite of TE subfamilies and evolutionary ages (Supplemental Table S1).
157 There was no difference in the ages of differentially expressed or continual TE-embedded miRNAs
158 (Supplemental Fig. S6A).

159 To determine the relevance of TE-embedded miRNAs versus non-TE-embedded miRNAs, we compared
160 the fold changes of each. Indeed, non-TE-embedded miRNAs were differentially expressed to a
161 comparable extent and in similar proportions as the TE-embedded miRNAs (Supplemental Fig. S6B-D).
162 Of note, the TE-embedded miRNA, hsa-mir-548ba_MADE1, was the most upregulated miRNA in
163 childhood compared to adolescence, indicating the relevance of TE-embedded miRNAs in the temporal
164 neurodevelopmental context (Supplemental Fig. S6B & C). This indicates that temporal expression is
165 not restricted to TE-embedded miRNAs but is a broad feature of this class of post-transcriptional
166 regulators (Ziats and Rennert 2014).



167

168 **Figure 2. TE-embedded miRNAs are temporally expressed between child and adolescent human**
 169 **brains. (A) Volcano plot highlighting TE-embedded miRNAs significantly differentially expressed in FB**
 170 **(adjusted P -value ≤ 0.05 , 1.5-fold change). (B) Dot plots showing the correlation of expression and age**
 171 **for specific TE-embedded miRNAs. Shaded area represents the variance. (C) (Top) Integrated genome**

172 viewer (IGV) visualization of four childhood (Child; blue) and four adolescent (Ado; orange) BAM files
173 from DFC small RNA-seq data. (*Middle*) IGV visualization of three AGO2 RIP-seq (RIP; green) and three
174 input (In; grey) BAM files from adult brain from Petri et al. 2019 (GSE106810). Read count is shown
175 within square brackets. (*Bottom*) miRBase annotation, TE annotation and gene annotations for hg19.
176 (*D*) miRNA hairpin schematics from miRNAfold (Tempel and Tahi 2012; Tav et al. 2016) for the DNA
177 sequences in *C*. Each hairpin structure exhibits 90% of verified miRNA hairpin features as previously
178 defined (Tempel and Tahi 2012; Tav et al. 2016). 22bp peak sequences are highlighted by the black
179 bars on arms of the hairpin. (*E*) Heatmaps showing regional expression in log₂ counts per million (CPM),
180 alongside differential expression results (black and grey bar). For differential expression, a linear model
181 was generated for every expressed miRNA, with miRNA expression as response variable and stage as
182 explanatory variable. TMM normalized expression estimates were used as input for the modelling.
183 Region abbreviations are defined in Supplemental Fig. S1.

184

185 In order to confirm our differential expression results, we next matched the expression of TE-
186 embedded miRNAs in the FB with donor age. Of the 20 differentially expressed TE-embedded miRNAs,
187 12 also exhibited significant correlations or anti-correlations with this parameter (Supplemental Fig.
188 S7; Supplemental Table S4).

189 One of the most significantly differentially expressed TE-embedded miRNAs in the FB was the cancer-
190 and cell proliferation-associated hsa-mir-378a, which displayed higher expression in childhood and a
191 significant anti-correlation with donor age in the small RNA-seq data (Li et al. 2015; Velazquez-Torres
192 et al. 2018; Guo et al. 2019)(Fig. 2A, B *left* & C *left top*). To confirm our detection of this TE-embedded
193 miRNA, we reanalysed a publicly available Argonaute2 RNA-immunoprecipitation sequencing (AGO2
194 RIP-seq) data set from three adult human brains (Petri et al. 2019; GSE106810). AGO2 directly binds to
195 mature processed miRNAs for incorporation into the RISC complex for targeting of mRNA (Kobayashi
196 and Tomari 2016; Michlewski and Cáceres 2019), therefore AGO2-bound elements are likely to
197 represent *bona fide* miRNAs rather than mere degradation products. Enrichment of reads in one out
198 of three AGO2 RIP-seq samples was observed compared to the input sample, with a peak residing over
199 the same sequence as the small RNA-seq data (Fig. 2C *left middle*). This hsa-mir-378a miRNA is
200 embedded in two intronic, MIRc elements arranged in opposite orientations, facilitating high
201 confidence pre-miRNA hairpin precursor formation as determined by *in silico* miRNA folding analyses
202 to detect hairpins with 90% of verified miRNA hairpin features (Tempel and Tahi 2012; Tav et al. 2016)
203 (Fig. 2C *bottom left* & D *left*). The glycolysis-, cancer- and cell proliferation-associated hsa-mir-5683

204 was also significantly more expressed in childhood, with a significant anti-correlation with donor age
205 and was detectable in at least one AGO2 RIP-seq sample, however was embedded in a solo 105 MYO
206 MER5A1 element which also facilitated pre-miRNA hairpin formation (Miao et al. 2020; Rong et al.
207 2020) (Fig. 2A, B *right*, C *middle* & D *middle*). One TE-embedded miRNA which was continually
208 expressed in childhood and adolescent brains from small RNA-seq data and detectable in all samples
209 of the AGO2 RIP-seq data was the cancer- and neuron- associated hsa-mir-582 (Fig. 2C *right*) (Fang et
210 al. 2015; Zhang et al. 2015; Ding et al. 2019). This miRNA is embedded in two apposed L3 elements,
211 again leading to an *in silico* predicted high-confidence precursor hairpin structure (Fig. 2D *right*).
212 Indeed, 18/92 miRNAs overlapped at least two TEs, with varying genomic orientations (Supplemental
213 Table S5), although the majority of expressed TE-embedded miRNAs overlapped only one TE.

214 **Different regions exhibit diverse miRNA temporal expression patterns**

215 The temporal TE and gene expression profile of the human brain varies by region, notably with the
216 cerebellum (CB) displaying a different transposcriptional and transcriptional landscape when
217 compared to FB (Playfoot et al. 2021). We therefore next determined the temporal expression profile
218 of miRNAs in childhood and adolescence in different individual brain regions (Fig. 2E, Supplemental
219 Table S1, see methods). hsa-mir-378a MIRc exhibited significantly higher expression in childhood, not
220 only in combined FB samples, but also in individual FB regions such as the dorsolateral prefrontal cortex
221 (DFC), the inferior temporal cortex (ITC), the medial prefrontal cortex (MFC) and the superior temporal
222 cortex (STC), along with non-FB regions such as the mediodorsal nucleus of the thalamus (THA) (Fig.
223 2E *left*). Similarly, hsa-mir-5683 MER5A1 had significantly higher expression in childhood versus
224 adolescence in the FB combined, along with other individual FB regions, but also in the CB (Fig. 2E
225 *middle*). In both instances, expression in the CB was higher than for any other individual region. In
226 contrast, the L3-embedded hsa-mir-582 exhibited continual high expression across childhood and
227 adolescence for all regions, except the CB where hsa-mir-582 L3 expression was largely absent in
228 adolescence and restricted to childhood (Fig. 2E *right*). This provides a striking example of

229 spatiotemporal control of the miRNA transcriptional landscape. Overall, these data demonstrate that
230 the TE-embedded miRNA transcriptional landscape exhibits diverse spatiotemporal dynamics, with
231 sometimes overt differences between childhood and adolescence for FB and non-FB regions.

232 **TE-embedded miRNAs are spatially expressed**

233 Due to the temporal nature of miRNA expression in multiple brain regions, we next aimed to determine
234 spatial differences in TE-embedded miRNA expression, regardless of age. We performed 120
235 differential expression analyses, comparing each region to each other independent region. We fitted
236 a linear model for every expressed miRNA, with miRNA expression as response variable and region as
237 explanatory variable. Stage and patient were used as covariates to the statistical model. TMM
238 normalized expression estimates were used as input for the modelling. miRNAs with fold change larger
239 than 1.5 and P -value ≤ 0.05 were considered as differentially expressed.

240 Of these comparisons, the region with the largest number of differentially expressed TE-embedded
241 miRNAs was consistently the CB (Fig. 3A). The CB was responsible for half of the top 30 comparisons
242 with the highest number of differentially expressed TE-embedded miRNAs (Fig. 3A). The CB versus the
243 hippocampus (HIP) had the highest number of differentially expressed TE-embedded miRNAs, followed
244 by the CB versus striatum (STR), amygdala (AMY) and many regions of the FB (Fig. 3A). These data
245 suggest that the CB exhibits, not only different TE and gene expression compared to other brain regions
246 as previously described (Playfoot et al. 2021; Li et al. 2018), but also differences in TE-embedded
247 miRNA expression.

248 To determine the relevance of TE-embedded miRNAs compared to non-TE-embedded miRNAs in the
249 spatial context, we ordered all differentially expressed miRNAs on the basis of their fold change in the
250 cerebellum versus hippocampus (Supplemental Fig. S8A & B). The differential expression of TE-
251 embedded miRNAs was spread between small and large fold changes, similar to non-TE-embedded
252 miRNAs. For example, mir-1298_X24_DNA was the second most highly upregulated miRNA in the
253 hippocampus when compared to cerebellum (Supplemental Fig. S8A). Together, these results indicate

254 that TE-embedded miRNAs exhibit similar expression differences to non-TE-embedded miRNAs in
255 regional comparisons.

256

257

258

259

260

261

262

263

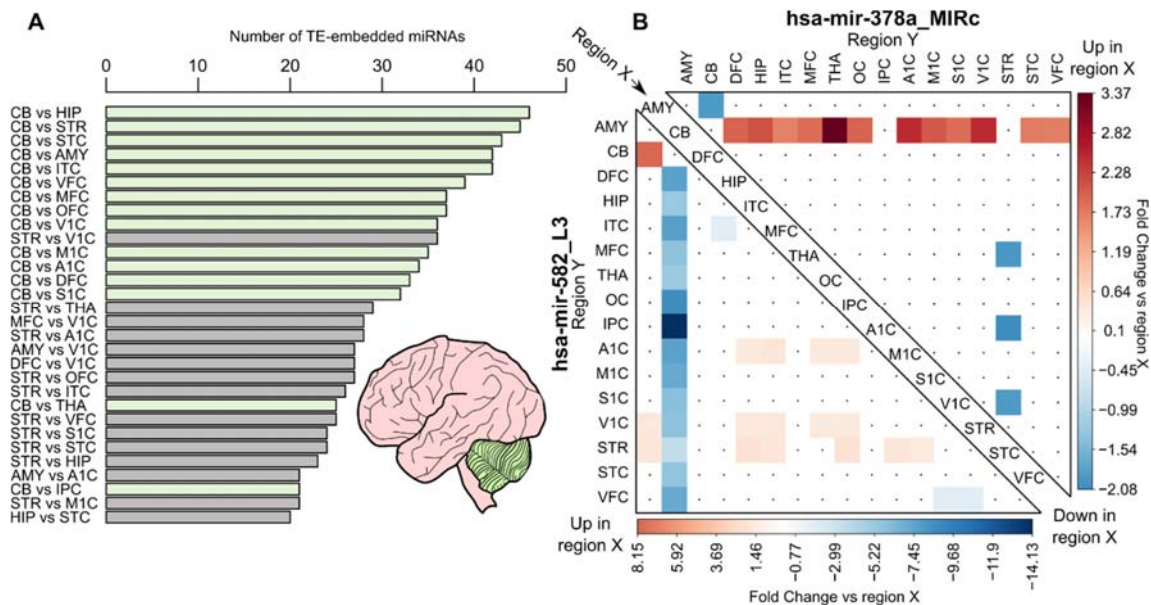
264

265

266

267

268



269

270 **Figure 3. TE-embedded miRNAs exhibit spatial expression with major differences in the cerebellum.**
 271 (A) Bar chart showing the number of differentially expressed TE-embedded miRNAs per regional
 272 comparisons (P -value ≤ 0.05 , 1.5-fold change up or down). Only the top 30 comparisons are shown.
 273 For differential expression, a linear model was generated for every expressed miRNA, with miRNA
 274 expression as response variable and region as explanatory variable. Stage and patient were used as
 275 covariates to the statistical model. TMM normalized expression estimates were used as input for the
 276 modelling. miRNAs with fold change larger than 1.5 and P -value ≤ 0.05 were considered as
 277 differentially expressed. (B) Heatmap comparing the fold change of region X (*center diagonal*) to region
 278 Y (*left and top*) for two TE-embedded miRNA loci described in Fig. 2. Only regions with significant fold
 279 changes are colored (P -value ≤ 0.05 , 1.5-fold change).

280

281

282

283

284

285

286

287

288

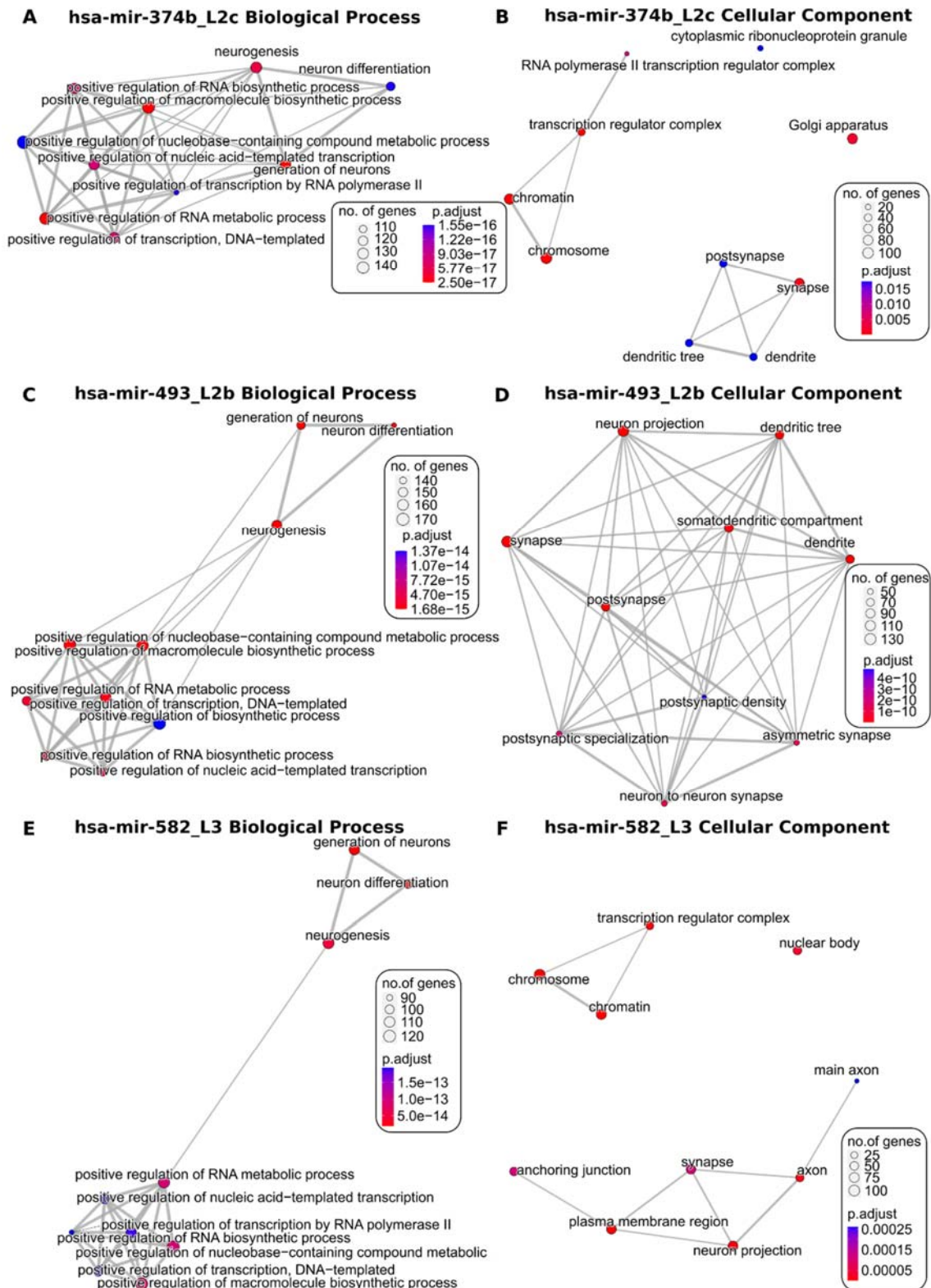
289 As hsa-mir-378a MIRc exhibited distinct temporal expression (Fig. 2) we next assessed its potential
290 spatial expression. Indeed, the CB exhibited significantly higher expression of hsa-mir-378a MIRc when
291 compared to most other regions (Fig. 3B). Conversely, hsa-mir-582 L3 exhibited significantly lower
292 expression in the CB compared to all other regions, suggestive of diverse regulatory control of different
293 miRNAs (Fig. 3B). A multitude of other examples of spatial miRNA expression suggests widespread
294 spatial regulation of not only TE-embedded miRNAs, but also non-TE-embedded miRNAs. These
295 dynamics can be interactively explored for all miRNAs with our Brain miRTEplorer application.

296 **TE-embedded miRNAs target neurogenesis-associated genes**

297 In order to determine possible functional relevance, we extracted predicted genic targets of TE-
298 embedded miRNAs from the TargetScan database (Supplemental Table S6) (Agarwal et al. 2015;
299 McGeary et al. 2019). We specifically focused on conserved predicted target sites (defined by
300 conserved branch lengths such as an 8mer or 7mer) of conserved miRNA families (defined by multiple-
301 sequence miRNA alignments), as annotated in TargetScan (Agarwal et al. 2015; McGeary et al. 2019).
302 Using this stringent list, we used two different gene ontology (GO) enrichment analysis tools
303 (clusterProfiler, Yu et al. 2012; and PantherDB, Mi et al. 2021). Biological process analyses indicated
304 that many target genes of TE-embedded miRNAs are enriched in neurogenesis-associated functions,
305 alongside other enriched pathways such as regulation of transcription and metabolic processes,
306 indicating the diversity of miRNA targets (Fig. 4A, C & E). For example, the L2c-embedded hsa-mir-374b
307 targets all four genes involved in striatal medium spiny neuron differentiation (GO:0021773) and three
308 out of four genes associated with glial cell fate specification (GO:0021780) and oligodendrocyte cell
309 fate specification (GO:0021778) (Supplemental Table S7). Similarly, the L2b-embedded hsa-mir-493
310 was enriched in positive regulation of synaptic vesicle exocytosis (GO:2000302) and neurotransmitter
311 receptor transport to plasma membrane (GO:0098877), among others such as actin polymerization-
312 dependent cell motility (GO:0070358) (Supplemental Table S7).

313 These two analyses also revealed significant enrichments in GO cell component terms for hsa-mir-374b
314 such as synapse (GO:0045202), golgi apparatus (GO:0005794) and transcription regulator complex
315 (GO:0005667) (Fig. 4B; Supplemental Table S7). hsa-mir-493 also exhibited diverse cellular component
316 GO enrichment terms such as NMDA selective glutamate receptor complex (GO:0017146), glial cell
317 projection (GO:0097386) and integral component of postsynaptic specialization membrane
318 (GO:0099060), among other neurogenesis and non-neurogenesis terms such as chromatin silencing
319 complex (GO:0005677) (Fig. 4D; Supplemental Table S7). The aforementioned continually expressed,
320 L3-embedded hsa-mir-582, was also broadly enriched in neurogenesis and non-neurogenesis-
321 associated biological process and cellular component terms such as astrocyte end-foot (GO:0097450),
322 main axon (GO:0044304) and glycoprotein complex (GO:0090665) (Fig. 4E; Supplemental Table S7).

323 Similar results were obtained for non-TE-embedded miRNAs such as the highly expressed hsa-mir-191,
324 enriched in neurogenesis processes such as synapse assembly (GO:0007416) and non-neurogenesis
325 such as pancreas development (GO:0031016) and cellular component terms such as spindle midzone
326 (GO:0051233) and chromatin (GO:0000785) (Supplemental Fig. S9A & B). The non-TE-embedded
327 miRNA with the largest fold change between childhood and adolescence, hsa-mir-211, also exhibited
328 similar GO enrichments for neurogenesis-associated biological processes and cellular components as
329 TE-embedded miRNAs (Supplemental Fig. S9C & D). Together, the GO enrichments detected for
330 different TE-embedded and non-TE-embedded miRNA predicted targets highlight the specialized
331 miRNA target networks in human neurogenesis and other diverse processes.



332

333 **Figure 4. TE-embedded miRNA predicted targets are enriched in neurogenesis and other diverse**
 334 **gene ontology terms.** ClusterProfiler emap network plots showing the top ten enriched biological
 335 process (A, C & E) and cellular component (B, D & F) GO terms for specific TE-embedded miRNAs. The

336 number of genes associated with each term is shown by point size and the adjusted *P-value* as the
337 indicated color. Edges connect overlapping gene sets which cluster together, indicating the relatedness
338 of terms. See also Supplemental Table S7 for GO enrichment from Panther DB.

339

340 **TEs contribute to novel putative miRNAs**

341 Most studies rely on mapping small RNA-seq reads directly to miRNA annotations provided in miRBase.
342 As a large proportion of annotated miRNAs are embedded in TEs, we reasoned that other TE loci could
343 be contributing to previously undetected, novel miRNAs expressed in the brain. We therefore further
344 investigated our unbiased, unique mapping to the whole genome used for detection of annotated TE-
345 embedded miRNAs. To ensure robustness and to limit false positives, we used our custom
346 RepeatMasker annotation (Turelli et al. 2020; Playfoot et al. 2021), alongside manual curation by
347 inspecting BAM files from childhood and adolescent samples of the DFC to detect a characteristic
348 ~22bp peak. These candidates were further refined by intersecting with genomic coordinates of TEs
349 with at least one read from the AGO2 RIP-seq data. As TEs are inherently repetitive and have many
350 thousands of copies, we next aimed to ensure that the sequences residing below putative miRNA peaks
351 were indeed novel by searching for the sequences in miRBase. This resulted in a stringent list of eight
352 novel non-annotated TE-embedded miRNA candidates (Supplemental Table S8). We next focused on
353 two of these which met our strict criteria. The first was embedded in two apposed head-to-head,
354 intronic MER3 elements and was confirmed with peaks detectable in the AGO2 RIP-seq data,
355 suggestive of processed miRNA (Fig. 5A & B *left*). Indeed, the 200bp sequence covering the miRNA
356 locus facilitated *in silico* hairpin structure formation with 90% of verified features and the 22bp 3p and
357 5p peak sequences contributing to each arm of the hairpin (Fig. 5C *left*). The same was observed for a
358 novel putative miRNA embedded in a single MER5A element, however AGO2 RIP-seq peaks overlapped
359 the probable miRNA star sequence (the peak with fewer reads in the Brainspan samples) (Fig. 5A, B &
360 C *right*). AGO2 alternative strand loading can lead to shifts in the target profile of the miRNA-Induced
361 Silencing Complex and may account for this (Medley et al. 2020).

362 To determine the evolutionary history of these two loci, we assessed the 22bp sequence using MULTIZ
363 alignments (Blanchette et al. 2004). Indeed, the MER3-embedded miRNA is present in rhesus macaque
364 but absent from mouse, whereas the MER5A element is present in rhesus macaque but with a deletion
365 in the seed region in mouse (Fig. 5D). To determine their novelty, the 22bp sequence of these
366 candidates were searched in miRBase and did not match any sequences. These two TE loci therefore
367 represent robust, novel TE-embedded miRNAs, the function of which remains to be elucidated.
368 Together, these data highlight the dynamic spatiotemporal nature of annotated and novel TE-
369 embedded miRNAs in the developing human brain and provides scope to investigate the disease and
370 functional relevance of TE sequence co-option as miRNAs throughout evolution.

371

372

373

374

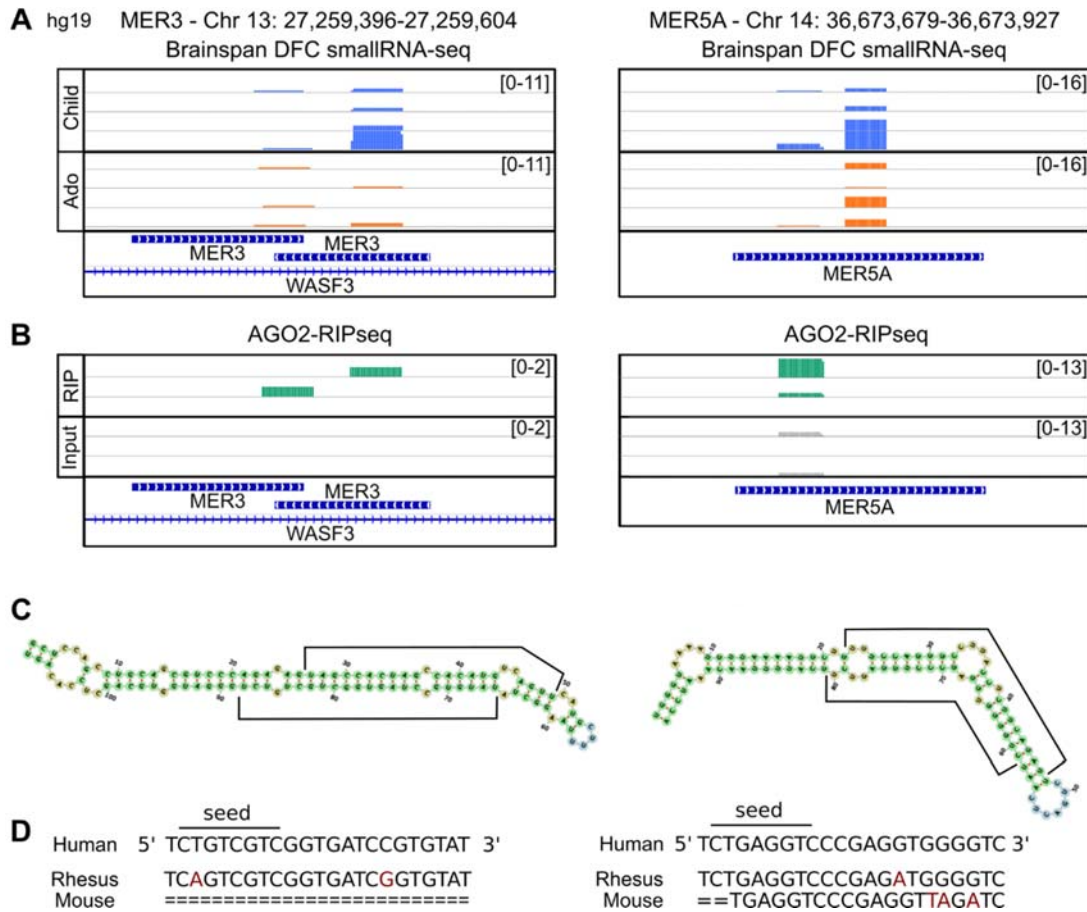
375

376

377

378

379



380

381 **Figure 5. Novel, non-annotated TE-embedded miRNAs are present in child and adolescent brains.** (A)
 382 IGV visualization of non-annotated TE-embedded miRNAs with classical 22bp peaks in four childhood
 383 (blue) DFC BAM files and four adolescent (orange) DFC BAM files, alongside TE and gene annotations
 384 for hg19. (B) IGV visualization of three AGO2 RIP-seq (RIP; green) and three input (In; grey) BAM files
 385 from adult brain (Petri et al. 2019; GSE106810) for the corresponding region in A. (C) miRNA hairpin
 386 schematics from miRNAfold (Tav et al. 2016) for the DNA sequences in A. 22bp peaks are highlighted
 387 by the black bars on both arms of the hairpin. (D) MULTIZ alignment from the UCSC genome browser
 388 of the 22bp miRNA sequence beneath the largest peak in A (Blanchette et al. 2004; Navarro Gonzalez
 389 et al. 2021). The putative seed region and orientation are indicated.

390

391

392

393

394

395 Discussion

396 Human brain development is a dynamic and highly regulated spatiotemporal process, however the
397 contribution of TEs to the miRNA mechanism of regulatory control has never been formally
398 investigated in this context. We show that the postnatal TE-embedded miRNA landscape is indeed
399 spatially and temporally dynamic, with alterations in TE-embedded miRNA expression from childhood
400 to adolescence, similar to non-TE-embedded miRNAs. Our previous work highlighted a distinct TE
401 expression switch during late prenatal and early postnatal developmental timepoints, accompanied by
402 coordinated reduction in expression of their controlling transcription factors, the KRAB-zinc finger
403 proteins (KZFPs) (Playfoot et al. 2021). Furthermore, we determined spatiotemporal TE-mediated
404 alternative promoter usage leading to novel mRNA transcript isoforms, indicative of direct TE-
405 dependent transcriptional innovation (Playfoot et al. 2021). Here, we expand the role of TEs in human
406 brain development to that of miRNAs; a more indirect, but no less important method of transcriptional
407 innovation.

408 One critical limitation of our study is the restriction to postnatal timepoints. As major gene and TE
409 expression changes occur during prenatal to postnatal transitional stages, future work should aim to
410 generate small RNA-seq data covering the whole timeframe of human brain development. miRNAs
411 were previously demonstrated to play critical roles in mouse prenatal brain development (Petri et al.
412 2014), and we found here that many human TE-embedded miRNAs were more highly expressed in
413 childhood when compared to adolescence. As many neurological disorders appear to have origins in
414 early development (Short and Baram 2019), it would be imperative to investigate both TE-embedded
415 and non-TE-embedded miRNA expression at prenatal stages. To date, the limited number of human
416 studies aiming to address this point were restricted by sample number, developmental stages and
417 regions (Nowakowski et al. 2018). Future work should also focus on the spatiotemporal control of TE-
418 embedded miRNA expression in other human developmental tissues, however current availability of
419 relevant data sets preclude this analysis.

420 Another limitation of our study is that we could not assess the effect of miRNA expression on their
421 mRNA targets. First, the effect size of miRNAs on target mRNAs is usually very small and the frequent
422 redundancy of miRNAs acting on the same mRNA target can mask the influence of a single effector
423 (Friedman et al. 2009; McGeary et al. 2019). In addition, brain samples comprising multiple different
424 cell types were analyzed here by bulk small RNA- or mRNA-sequencing, precluding a straightforward
425 interpretation of the relationship between miRNA and target mRNA expression. Advances in human
426 embryonic stem cell differentiation protocols have enabled *in vitro* study of different neurological cell
427 types and cerebral organoids, hence will facilitate this type of exploration, although these approaches
428 still fail to recapitulate the wide cellular diversity or maturity found in tissue samples.

429 The detection of novel, non-annotated TE-embedded miRNAs is suggestive of a previously undetected
430 TE-originating miRNA landscape. The volume of data assessed may have allowed the detection of
431 these, however computational limitations of using only uniquely mapping reads is especially acute for
432 young, more homogenous TE subfamilies which have accumulated less mutations. This can be further
433 compounded by the absence of unique molecular identifiers (UMI) in sequencing reads, leading to
434 misinterpretation of PCR duplications. Future work should experimentally assess putative, young
435 repetitive TE-embedded miRNAs, as they have the potential to expand significantly the RNA-based
436 regulome. Their repetitive nature likely facilitates post-transcriptional control of mRNA targets
437 containing the same TE subfamilies in their 3' UTRs, as has been shown for the annotated L2-embedded
438 miRNAs (Petri et al. 2019). These results also suggest a multifactorial role for TEs, whereby some TEs
439 give rise to mature miRNAs but are also targets of the miRNA microprocessor machinery themselves,
440 thus acting to restrict their movement when the TEs remain retrotransposition competent (Heras et
441 al. 2013, 2014). Indeed, brain-specific data sets determining direct targets of the microprocessor would
442 be useful to determine if these TE-embedded miRNAs are *bona fide* targets of the miRNA processing
443 machinery. Furthermore, it may be possible that the expression of TEs themselves may contribute to
444 the expression of miRNAs, again creating a regulatory feedback loop.

445 In summary, the spatiotemporal expression of TE-embedded miRNAs from childhood to adolescence
446 suggests a role for TEs in the fine tuning of transcriptional networks at the post-transcriptional level
447 throughout human brain development. Although these dynamics are not restricted to TE-embedded
448 miRNAs, these analyses provide a novel insight into a crucial understudied developmental window, as
449 the role of TE-embedded miRNAs has only been previously investigated in adult or disease contexts.

450 **Materials and Methods**

451 **Dataset download and preprocessing**

452 Raw small RNA-seq FASTQ files from the BrainSpan Atlas of the Developing Human Brain
453 (phs000755.v2.p1 provided by Dr. Nenad Sestan), were downloaded from the dbGaP-authorized
454 access platform (Miller et al. 2014; Li et al. 2018) (Supplemental Acknowledgements). The reads were
455 first trimmed to remove Illumina small RNA 3' sequencing adapters (TGGAATTCTCGGGTGCCAAGG)
456 using FLEXBAR (version 3.5.0) with parameters --adapter-trim-end RIGHT --min-read-length 18 (Dodt
457 et al. 2012). Trimmed reads were then divided by read length ranges of 18-25 nucleotides, 26-37
458 nucleotides and 38-50 nucleotides. Reads were then mapped to the human hg19 genome (GRCh37.p5)
459 using Bowtie (version 2.3.4.1) with parameter --very-sensitive-local (Langmead et al. 2009). Read
460 counts on different genomic features were quantified using featureCounts (version 1.6.2 of the
461 subread package) (Liao et al. 2014). Uniquely mapped reads were quantified with parameters -t exon
462 -g gene_id -Q 10 and multimapped reads with parameters -M --fraction -t exon -g gene_id -Q 0. We
463 used the parameters -s 1 and -s 2, to quantify sense and antisense reads respectively which were
464 subsequently merged, keeping only the strand with the most reads. To confirm specific read lengths
465 were enriching for specific RNA moieties, the annotation of snoRNA, snRNA, miscRNA, scRNA and
466 genes from Ensembl (GRCh37.p5, release 100) were used. For miRNAs and tRNAs, miRBase version 20
467 (Kozomara and Griffiths-Jones 2014) and tRNA annotations from GtRNAdb (release 19) were used
468 respectively (Chan and Lowe 2016). For repetitive sequences, a previously described in-house curated
469 version of the RepeatMasker database was used (where fragmented LTR and internal segments

470 belonging to a single integrant were merged) (Turelli et al. 2020; Playfoot et al. 2021). Exons of genes
471 and TEs overlapping small RNAs in the same orientation were removed using BEDTools intersect
472 (version 2.27.1) with default parameters, to prioritize reads falling on small RNAs (Quinlan and Hall
473 2010). To determine which expressed annotated miRNAs overlapped TEs, we used BEDTools to
474 intersect the miRBase and our custom RepeatMasker merged TE annotations with a minimum of one
475 base pair overlap. TE subfamily age estimates were obtained from DFAM (Hubley et al. 2016). BAM
476 files were visualized using the Integrative Genome Viewer (Robinson et al. 2011).

477 **Filtering and normalization**

478 Samples were sequenced with a read length of 51bp and samples with less than 1 million reads mapped
479 were removed. Features where the sum of the counts over all the samples was lower than the total
480 number of samples were removed. TEs overlapping gene exons were also removed using BEDTools
481 closest (Quinlan and Hall 2010). Normalization for the sequencing depth was performed for all features
482 on the sense and antisense with the trimmed mean of the log expression ratios (TMM) method as
483 implemented in the R package limma (version 3.46.0) (Ritchie et al. 2015). The TMM method excludes
484 the top and bottom expressed miRNAs prior to computing library sizes to ensure very high or low
485 expressed miRNAs did not dominate the library size normalization (Robinson and Oshlack 2010).
486 Counts per million (CPM) values were generated after correcting by the TMM library size. The
487 subsequent total number of mapped reads was used as library size.

488 **Differential expression analysis**

489 Samples from one year to five years were considered as childhood and nine years to 20 years as
490 adolescence (Supplemental Fig. S1). To perform the aggregated temporal FB differential expression,
491 the following brain regions were considered as FB: dorsolateral prefrontal cortex, inferior temporal
492 cortex, medial prefrontal cortex, orbital prefrontal cortex, posterior inferior parietal cortex, primary
493 auditory (A1) cortex, primary somatosensory (S1) cortex, primary visual (V1) cortex, superior temporal
494 cortex, ventrolateral prefrontal cortex and primary motor (M1) cortex (Supplemental Fig. S1).

495 Independent temporal comparisons were performed without aggregations of multiple regions. For
496 differential expression between regions, all samples regardless of age were used.

497 Differential gene expression analysis was performed using voom (Law et al. 2014) as it has been
498 implemented in the R package limma (version 3.46.0), with TMM normalized counts as input and using
499 staging and patient information as covariates when fitting the linear models. *P-values* were corrected
500 for multiple testing using the Benjamini-Hochberg's method (Benjamini and Hochberg 1995). A feature
501 was considered to be differentially expressed when the fold change between the groups compared
502 was higher than 1.5 and the adjusted or non-adjusted *P-value* ≤ 0.05 as stated in figure legends.

503 **Correlation analysis**

504 Correlation between age and miRNA expression was assessed using spearman correlation and *P-values*
505 were adjusted using the Bonferroni correction.

506 **Expression of TE-embedded miRNAs in other tissues**

507 Processed CPM expression data of mature miRNAs in 399 human samples (De Rie et al. 2017; file:
508 human.srna.cpm.txt) was downloaded and \log_2 CPMs of all annotated mature miRNAs non-overlapping
509 and overlapping-TE annotations were plotted with addition of a pseudo-count of 1.

510 **miRNA precursor secondary structure analyses**

511 To predict *in silico* miRNA precursor hairpin structures, the DNA sequence of a 200 to 300 bp window
512 around consistent 22bp peaks observed in BAM files was inputted to miRNAfold (Tempel and Tah
513 2012; Tav et al. 2016). A stringent threshold of 90% of verified features was used, to ensure only robust
514 hairpins with a very low false positive rate were returned (Tempel and Tah 2012; Tav et al. 2016).

515 **AGO2 RIP-seq data**

516 Three publicly available neurotypical adult brain AGO2 RIP-seq data sets (Petri et al. 2019; GSE106810)
517 were processed using the same pipeline as for the small RNA-seq analysis.

518 **Evolutionary conservation**

519 MULTIZ tracks from the UCSC genome browser were used to determine the presence of non-annotated
520 TE-embedded miRNA sequences in different species (Blanchette et al. 2004; Navarro Gonzalez et al.
521 2021). TE-embedded miRNA coordinates were lifted over from the human genome (hg19) to that of
522 the macaque (rheMac8) and the mouse genome (mm10) using the UCSC lift over tool (Hinrichs et al.
523 2006). BEDTools intersect (version 2.27.1) (Quinlan and Hall 2010) was used with default parameters
524 to intersect lift-over coordinates with miRNA coordinates from miRBase for macaque (rheMac8) and
525 mouse (mm10) (Kozomara and Griffiths-Jones 2014).

526 **miRNA target prediction**

527 miRNA target predictions were downloaded from TargetScan Human (Release 8.0) (Agarwal et al.
528 2015; McGeary et al. 2019; File: Predicted_Targets_Context_Scores.default_predictions.txt). We
529 utilized only the conserved target predictions for conserved miRNAs as defined in TargetScan (Agarwal
530 et al. 2015; McGeary et al. 2019). GO analysis and visualization was performed using clusterProfiler (Yu
531 et al., 2012) with default options and PantherDB (Release: 20210224) (Mi et al., 2021), with GO
532 ontology database (Release: 2021-08-18) and enrichment was assessed using Fisher's exact test
533 followed by false discovery rate adjustment using all human genes as background (Mi et al. 2013).

534 **Brain miRTEplorer application**

535 The Brain miRTEplorer application was implemented in R using the Shiny app package (Chang et al.,
536 2017). A description and example of usage is provided as Supplemental Fig. S10.

537 **Availability of data and materials**

538 No new data was generated during the course of this study. Processed data can be interactively
539 visualized using our "Brain miRTEplorer" application at
540 <https://tronoapps.epfl.ch/BrainmiRTEplorer/>.

541 **Competing interests**

542 The authors declare that they have no competing interests

543 **Funding**

544 This study was supported by grants from the Personalized Health and Related Technologies (PHRT-
545 508), the European Research Council (KRABnKAP, #268721; Transpos-X, #694658), and the Swiss
546 National Science Foundation (310030_152879 and 310030B_173337) to D.T.

547 **Author contributions**

548 C.P. conceived the study, performed bioinformatic analyses, interpreted the data and wrote the
549 manuscript. S.S. and E.P. performed bioinformatics analyses. C.P. and D.T. edited the manuscript. All
550 authors reviewed the manuscript.

551 **Acknowledgements**

552 We thank all members of the Trono Lab, Johan Jakobsson and Retha Ritter for helpful and insightful
553 discussions.

554 **References**

- 555 Agarwal V, Bell GW, Nam JW, Bartel DP. 2015. Predicting effective microRNA target sites in
556 mammalian mRNAs. *Elife* **4**: 1–38.
- 557 Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful
558 Approach to Multiple Controlling the False Discovery Rate: a Practical and Powerful Approach to
559 Multiple Testing. *J R Stat Soc* **57**: 289–300.
- 560 Blanchette M, Kent WJ, Riemer C, Elnitski L, Smith AFA, Roskin KM, Baertsch R, Rosenbloom K,
561 Clawson H, Green ED, et al. 2004. Aligning multiple genomic sequences with the threaded
562 blockset aligner. *Genome Res* **14**: 708–715.
- 563 Borchert GM, Holton NW, Williams JD, Hernan WL, Bishop IP, Dembosky JA, Elste JE, Gregoire NS,
564 Kim J-A, Koehler WW, et al. 2011. Comprehensive analysis of microRNA genomic loci identifies
565 pervasive repetitive-element origins. *Mob Genet Elements* **1**: 8–17.
- 566 Cao X, Pfaff SL, Gage FH. 2007. A functional study of miR-124 in the developing neural tube. *Genes*
567 *Dev* **21**: 531–536.
- 568 Chan PP, Lowe TM. 2016. GtRNADB 2.0: An expanded database of transfer RNA genes identified in
569 complete and draft genomes. *Nucleic Acids Res* **44**: D184–D189.
- 570 Chang, W., Cheng, J., Allaire, J., Xie, Y., & McPherson, J. 2017. Shiny: web application framework for
571 R. R package version, 1(5).

- 572 Chuong EB, Rumi M a K, Soares MJ, Baker JC. 2013. Endogenous retroviruses function as species-
573 specific enhancer elements in the placenta. *Nat Genet* 45: 325–329.
- 574 Chuong EB, Elde NC, Feschotte C. 2016. Regulatory evolution of innate immunity through co-option
575 of endogenous retroviruses. *Science* 351: 1083–1087.
- 576 Chuong EB, Elde NC, Feschotte C. 2017. Regulatory activities of transposable elements: from conflicts
577 to benefits. *Nat Rev Genet* 18: 71–86.
- 578 Da Fonseca BHR, Domingues D S, Paschoal AR. 2019. mirtronDB: a mirtron knowledge base.
579 *Bioinformatics*, 35: 3873-3874.
- 580 De Rie D, Abugessaisa I, Alam T, Arner E, Arner P, Ashoor H, Åström G, Babina M, Bertin N, Burroughs
581 AM, et al. 2017. An integrated expression atlas of miRNAs and their promoters in human and
582 mouse. *Nat Biotechnol* 35: 872–878.
- 583 Ding H, Gao S, Wang L, Wei Y, Zhang M. 2019. Overexpression of miR-582-5p Inhibits the Apoptosis
584 of Neuronal Cells after Cerebral Ischemic Stroke Through Regulating PAR-1/Rho/Rho Axis. *J*
585 *Stroke Cerebrovasc Dis* 28: 149–155.
- 586 Ding J, Huang S, Wu S, Zhao Y, Liang L, Yan M, Ge C, Yao J, Chen T, Wan D, et al. 2010. Gain of miR-
587 151 on chromosome 8q24.3 facilitates tumour cell migration and spreading through
588 downregulating RhoGDI. *Nat Cell Biol* 12: 390–399.
- 589 Dodt M, Roehr JT, Ahmed R, Dieterich C. 2012. FLEXBAR-flexible barcode and adapter processing for
590 next-generation sequencing platforms. *Biology (Basel)* 1: 895–905.
- 591 Dyck LIV, Morrow EM. 2017. Genetic control of postnatal human brain growth. *Curr Opin Neurol* 30:
592 114–124.
- 593 Elbarbary RA, Lucas BA, Maquat LE. 2016. Retrotransposons as regulators of gene expression. *Science*
594 351: aac7247.
- 595 Fang L, Cai J, Chen B, Wu S, Li R, Xu X, Yang Y, Guan H, Zhu X, Zhang L, et al. 2015. Aberrantly
596 expressed miR-582-3p maintains lung cancer stem cell-like traits by activating Wnt/ β -catenin
597 signalling. *Nat Commun* 6: 8640.
- 598 Frankel LB, Di Malta C, Wen J, Eskelinen EL, Ballabio A, Lund AH. 2014. A non-conserved miRNA
599 regulates lysosomal function and impacts on a human lysosomal storage disorder. *Nat Commun*
600 5: 5840.
- 601 Friedman RC, Farh KK, Burge CB, Bartel DP. 2009. Most mammalian mRNAs are conserved targets of
602 microRNAs. *Genome research*. 19: 92-105.
- 603 Garcia-Perez JL, Widmann TJ, Adams IR. 2016. The impact of transposable elements on mammalian
604 development. *Development* 143: 4101–4114.
- 605 Guo XB, Zhang XC, Chen P, Ma LM, Shen ZQ. 2019. MiR-378a-3p inhibits cellular proliferation and
606 migration in glioblastoma multiforme by targeting tetraspanin 17. *Oncol Rep* 42: 1957–1971.
- 607 Heras SR, Macias S, Cáceres JF, Garcia-Perez JL. 2014. Control of mammalian retrotransposons by
608 cellular RNA processing activities. *Mob Genet Elements* 4: e28439.
- 609 Heras SR, Macias S, Plass M, Fernandez N, Cano D, Eyra E, Garcia-Perez JL, Cáceres JF. 2013. The
610 Microprocessor controls the activity of mammalian retrotransposons. *Nat Struct Mol Biol* 20:
611 1173–81.

- 612 Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H, Diekhans M, Furey TS, Harte
613 RA, Hsu F, Hillman-Jackson J. 2006. The UCSC genome browser database: update 2006. *Nucleic acids*
614 *research*. **34**: D590-8.
- 615 Huang D, Qiu S, Ge R, He L, Li M, Li Y, Peng Y. 2015. miR-340 suppresses glioblastoma multiforme.
616 *Oncotarget*. **6**: 9257
- 617 Hubley R, Finn RD, Clements J, Eddy SR, Jones TA, Bao W, Smit AFA, Wheeler TJ. 2016. The Dfam
618 database of repetitive DNA families. *Nucleic Acids Res* **44**: D81–D89.
- 619 Jużwik CA, S. Drake S, Zhang Y, Paradis-Isler N, Sylvester A, Amar-Zifkin A, Douglas C, Morquette B,
620 Moore CS, Fournier AE. 2019. microRNA dysregulation in neurodegenerative diseases: A
621 systematic review. *Prog Neurobiol* **182**: 101664.
- 622 Kefas B, Comeau L, Floyd DH, Seleverstov O, Godlewski J, Schmittgen T, Jiang J, Li Y, Chiocca EA, Lee J,
623 Fine H. 2009. The neuronal microRNA miR-326 acts in a feedback loop with notch and has
624 therapeutic potential against brain tumors. *Journal of Neuroscience*. **29**: 15161-8.
- 625 Kobayashi H, Tomari Y. 2016. RISC assembly: Coordination between small RNAs and Argonaute
626 proteins. *Biochim Biophys Acta - Gene Regul Mech* **1859**: 71–81.
- 627 Kozomara A, Griffiths-Jones S. 2014. MiRBase: Annotating high confidence microRNAs using deep
628 sequencing data. *Nucleic Acids Res* **42**: 68–73.
- 629 Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short
630 DNA sequences to the human genome. *Genome Biol* **10**: R25.
- 631 Law CW, Chen Y, Shi W, Smyth GK. 2014. voom: precision weights unlock linear model analysis tools
632 for RNA-seq read counts. *Genome Biol* **15**: R29.
- 633 Li B, Wang Y, Li S, He H, Sun F, Wang C, Lu Y, Wang X, Tao B. 2015. Decreased expression of miR-378
634 correlates with tumor invasiveness and poor prognosis of patients with glioma. *Int J Clin Exp*
635 *Pathol* **8**: 7016–7021.
- 636 Li M, Santpere G, Imamura Kawasawa Y, Evgrafov O V., Gulden FO, Pochareddy S, Sunkin SM, Li Z,
637 Shin Y, Zhu Y, et al. 2018. Integrative functional genomic analysis of human brain development
638 and neuropsychiatric risks. *Science* **362**: eaat7615.
- 639 Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning
640 sequence reads to genomic features. *Bioinformatics* **30**: 923–930.
- 641 McGeary SE, Lin KS, Shi CY, Pham TM, Bisaria N, Kelley GM, Bartel DP. 2019. The biochemical basis of
642 microRNA targeting efficacy. *Science* **366**: eaav1741.
- 643 Medley JC, Panzade G, Zinovyeva AY. 2021. microRNA strand selection: Unwinding the rules. *Wiley*
644 *Interdisciplinary Reviews: RNA*. **12**: e1627.
- 645 Mi H, Muruganujan A, Casagrande JT, Thomas PD. 2013. Large-scale gene function analysis with the
646 panther classification system. *Nat Protoc* **8**: 1551–1566.
- 647 Mi H, Ebert D, Muruganujan A, Mills C, Albou LP, Mushayamaha T, Thomas PD. 2021. PANTHER
648 version 16: a revised family classification, tree-based classification tool, enhancer regions and
649 extensive API. *Nucleic acids research*. **49**: D394-403.
- 650 Miao Y, Li Q, Sun G, Wang L, Zhang D, Xu H, Xu Z. 2020. MiR-5683 suppresses glycolysis and
651 proliferation through targeting pyruvate dehydrogenase kinase 4 in gastric cancer. *Cancer Med*
652 **9**: 7231–7243.

- 653 Michlewski G, Cáceres JF. 2019. Post-transcriptional control of miRNA biogenesis. *Rna* **25**: 1–16.
- 654 Miller JA, Ding S-L, Sunkin SM, Smith KA, Ng L, Szafer A, Ebbert A, Riley ZL, Royall JJ, Aiona K, et al.
655 2014. Transcriptional landscape of the prenatal human brain. *Nature* **508**: 199–206.
- 656 Navarro Gonzalez J, Zweig AS, Speir ML, Schmelter D, Rosenbloom KR, Raney BJ, Powell CC, Nassar
657 LR, Maulding ND, Lee CM, et al. 2021. The UCSC genome browser database: 2021 update.
658 *Nucleic Acids Res* **49**: D1046–D1057.
- 659 Nowakowski TJ, Rani N, Golkaram M, Zhou HR, Alvarado B, Huch K, West JA, Leyrat A, Pollen AA,
660 Kriegstein AR, et al. 2018. Regulation of cell-type-specific transcriptomes by microRNA networks
661 during human brain development. *Nat Neurosci* **21**: 1784–1792.
- 662 Petri R, Brattås PL, Sharma Y, Jonsson ME, Piracs K, Bengzon J, Jakobsson J. 2019. LINE-2 transposable
663 elements are a source of functional human microRNAs and target sites. *PLoS Genet* **15**: 1–18.
- 664 Petri R, Malmevik J, Fasching L, Åkerblom M, Jakobsson J. 2014. MiRNAs in brain development. *Exp*
665 *Cell Res* **321**: 84–89.
- 666 Piriyaopongsa J, Jordan IK. 2007. A Family of Human MicroRNA Genes from Miniature Inverted-Repeat
667 Transposable Elements ed. A. Christoffels. *PLoS One* **2**: e203.
- 668 Piriyaopongsa J, Mariño-Ramírez L, Jordan IK. 2007. Origin and evolution of human microRNAs from
669 transposable elements. *Genetics* **176**: 1323–1337.
- 670 Playfoot CJ, Duc J, Sheppard S, Dind S, Coudray A, Planet E, Trono D. 2021. Transposable elements
671 and their KZFP controllers are drivers of transcriptional innovation in the developing human
672 brain. *Genome Res* **31**: 1531–1545.
- 673 Pontis J, Planet E, Offner S, Turelli P, Duc J, Coudray A, Theunissen TW, Jaenisch R, Trono D. 2019.
674 Hominoid-Specific Transposable Elements and KZFPs Facilitate Human Embryonic Genome
675 Activation and Control Transcription in Naive Human ESCs. *Cell Stem Cell* **24**: 724-735.e5.
- 676 Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features.
677 *Bioinformatics* **26**: 841–842.
- 678 Qureshi IA, Mehler MF. 2012. Emerging roles of non-coding RNAs in brain evolution, development,
679 plasticity and disease. *Nat Rev Neurosci* **13**: 528–541.
- 680 Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. limma powers differential
681 expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**: e47–e47.
- 682 Roberts JT, Cardin SE, Borchert GM. 2014. Burgeoning evidence indicates that microRNAs were
683 initially formed from transposable element sequences. *Mob Genet Elements* **4**: e29255.
- 684 Robinson MD, Oshlack A. 2010. A scaling normalization method for differential expression analysis of
685 RNA-seq data. *Genome biology*. **11**: 1-9.
- 686 Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011.
687 Integrative genomics viewer. *Nat Biotechnol* **29**: 24–26.
- 688 Rong Z, Rong Y, Li Y, Zhang L, Peng J, Zou B, Zhou N, Pan Z. 2020. Development of a Novel Six-miRNA-
689 Based Model to Predict Overall Survival Among Colon Adenocarcinoma Patients. *Front Oncol*
690 **10**: 1–13.
- 691 Ruan J, Lou S, Dai Q, Mao D, Ji J, Sun X. 2015. Tumor suppressor miR-181c attenuates proliferation,
692 invasion, and self-renewal abilities in glioblastoma. *Neuroreport* **26**: 66-73.
- 693 Sambandan S, Akbalik G, Kochen L, Rinne J, Kahlstatt J, Glock C, Tushev G, Alvarez-Castelao B, Heckel

- 694 A, Schuman EM. 2017. Activity-dependent spatially localized miRNA maturation in neuronal
695 dendrites. *Science* **355**: 634–637.
- 696 Shao N-Y, Hu H, Yan Z, Xu Y, Hu H, Menzel C, Li N, Chen W, Khaitovich P. 2010. Comprehensive survey
697 of human brain microRNA by deep sequencing. *BMC Genomics* **11**: 409.
- 698 Short AK, Baram TZ. 2019. Early-life adversity and neurological disease: age-old questions and novel
699 answers. *Nat Rev Neurol* **15**: 657–669.
- 700 Silbereis JC, Pochareddy S, Zhu Y, Li M, Sestan N. 2016. The Cellular and Molecular Landscapes of the
701 Developing Human Central Nervous System. *Neuron* **89**: 248–268.
- 702 Skalsky RL, Cullen BR. 2011. Reduced expression of brain-enriched microRNAs in glioblastomas
703 permits targeted regulation of a cell death gene. *PLoS One* **6**: 20–23.
- 704 Smalheiser N, Torvik V. 2005. Mammalian microRNAs derived from genomic repeats. *Trends Genet*
705 **21**: 322–326.
- 706 Somel M, Liu X, Tang L, Yan Z, Hu H, Guo S, Jiang X, Zhang X, Xu G, Xie G, et al. 2011. MicroRNA-
707 Driven Developmental Remodeling in the Brain Distinguishes Humans from Other Primates.
708 *PLoS Biol* **9**: e1001214.
- 709 Spengler RM, Oakley CK, Davidson BL. 2014. Functional microRNAs and target sites are created by
710 lineage-specific transposition. *Hum Mol Genet* **23**: 1783–1793.
- 711 Tav C, Tempel S, Poligny L, Tahi F. 2016. miRNAfold: a web server for fast miRNA precursor
712 prediction in genomes. *Nucleic Acids Res* **44**: W181–W184.
- 713 Tempel S, Tahi F. 2012. A fast ab-initio method for predicting miRNA precursors in genomes. *Nucleic*
714 *Acids Res* **40**: e80–e80.
- 715 Topol A, Zhu S, Hartley BJ, English J, Hauberg ME, Tran N, Rittenhouse CA, Simone A, Ruderfer DM,
716 Johnson J, et al. 2016. Dysregulation of miRNA-9 in a Subset of Schizophrenia Patient-Derived
717 Neural Progenitor Cells. *Cell Rep* **15**: 1024–1036.
- 718 Turelli P, Playfoot C, Grun D, Raclot C, Pontis J, Coudray A, Thorball C, Duc J, Pankevich E V.,
719 Deplancke B, et al. 2020. Primate-restricted KRAB zinc finger proteins and target
720 retrotransposons control gene expression in human neurons. *Sci Adv* **6**: eaba3200.
- 721 Velazquez-Torres G, Shoshan E, Ivan C, Huang L, Fuentes-Mattei E, Paret H, Kim SJ, Rodriguez-Aguayo
722 C, Xie V, Brooks D, et al. 2018. A-to-I miR-378a-3p editing can prevent melanoma progression
723 via regulation of PARVA expression. *Nat Commun* **9**: 461.
- 724 Woods BJ, Van Vactor D. 2021. miRNA: local guardians of presynaptic function in plasticity and
725 disease. *RNA Biol* **18**: 1014–1024.
- 726 Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Azov AG,
727 Bennett R, et al. 2019. Ensembl 2020. *Nucleic Acids Res* **48**: D682–D688.
- 728 Yu G, Wang LG, Han Y, He QY. 2012. clusterProfiler: an R package for comparing biological themes
729 among gene clusters. *Omics: a journal of integrative biology*. **16**: 284–7.
- 730 Zhang L, Dong LY, Li YJ, Hong Z, Wei WS. 2012. The microRNA miR-181c controls microglia-mediated
731 neuronal apoptosis by suppressing tumor necrosis factor. *Journal of neuroinflammation*. **9**: 1–2.
- 732 Zhang X, Zhang Y, Yang J, Li S, Chen J. 2015. Upregulation of miR-582-5p inhibits cell proliferation, cell
733 cycle progression and invasion by targeting Rab27a in human colorectal carcinoma. *Cancer*
734 *Gene Ther* **22**: 475–480.

735 Ziats MN, Rennert OM. 2014. Identification of differentially expressed microRNAs across the
736 developing human brain. *Mol Psychiatry* **19**: 848–852.

737

738



RNA
A PUBLICATION OF THE RNA SOCIETY

Transposable elements contribute to the spatiotemporal microRNA landscape in human brain development

Christopher J Playfoot, Shaoline Sheppard, Evarist Planet, et al.

RNA published online June 22, 2022

Supplemental Material <http://rnajournal.cshlp.org/content/suppl/2022/06/22/rna.079100.122.DC1>

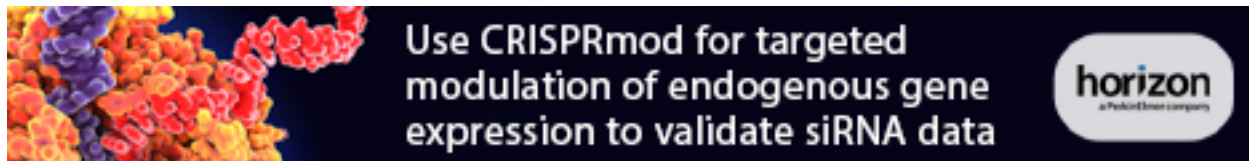
P<P Published online June 22, 2022 in advance of the print journal.

Accepted Manuscript Peer-reviewed and accepted for publication but not copyedited or typeset; accepted manuscript is likely to differ from the final, published version.

Open Access Freely available online through the *RNA* Open Access option.

Creative Commons License This article, published in *RNA*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

An advertisement banner with a dark background. On the left, there is a colorful 3D molecular model of a protein complex. The text in the center reads "Use CRISPRmod for targeted modulation of endogenous gene expression to validate siRNA data". On the right, there is a logo for "horizon" with the tagline "a PerkinElmer company" below it.

Use CRISPRmod for targeted modulation of endogenous gene expression to validate siRNA data

horizon
a PerkinElmer company

To subscribe to *RNA* go to:
<http://rnajournal.cshlp.org/subscriptions>