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

CHRISTOPHER J. PLAYFOOT, SHAOLINE SHEPPARD, EVARIST PLANET, and DIDIER TRONO

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

ABSTRACT

Transposable elements (TEs) contribute to the evolution of gene regulatory networks and are dynamically expressed throughout human brain development and disease. One gene regulatory mechanism influenced by TEs is the miRNA system of post-transcriptional control. miRNA sequences frequently overlap TE loci and this miRNA expression landscape is crucial for control of gene expression in adult brain and different cellular contexts. Despite this, a thorough investigation of the spatiotemporal expression of TE-embedded miRNAs in human brain development is lacking. Here, we identify a spatiotemporally dynamic TE-embedded miRNA expression landscape between childhood and adolescent stages of human brain development. These miRNAs sometimes arise from two apposed TEs of the same subfamily, such as for L2 or MIR elements, but in the majority of cases stem from solo TEs. They give rise to in silico predicted high-confidence pre-miRNA hairpin structures, likely represent functional miRNAs, and have predicted genic targets associated with neurogenesis. TE-embedded miRNA expression is distinct in the cerebellum when compared to other brain regions, as has previously been described for gene and TE expression. Furthermore, we detect expression of previously nonannotated TE-embedded miRNAs throughout human brain development, suggestive of a previously undetected miRNA control network. Together, as with non-TE-embedded miRNAs, TE-embedded sequences give rise to spatiotemporally dynamic miRNA expression networks, the implications of which for human brain development constitute extensive avenues of future experimental research. To facilitate interactive exploration of these spatiotemporal miRNA expression dynamics, we provide the "Brain miRTExplorer" web application freely accessible for the community.

Keywords: human brain; neurogenesis; transposable elements; microRNA

INTRODUCTION

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

Recent studies in a small number of adult brains have highlighted the roles of TE-embedded miRNAs from the L2 family. These are functional in neurotypical adult brains and are differentially expressed in glioblastoma (Skalsky and Cullen 2011; Petri et al. 2019). Furthermore, miRNAs have critical roles in mammalian neuronal homeostasis, highlighting the fundamental nature of miRNAs in neurogenesis, alongside diverse roles in neurological disease and human evolution (Cao et al. 2007; Somel et al. 2011; Qureshi and Mehler 2012; Petri et al. 2014; Topol et al. 2016; Sambandan et al. 2017; Juźwik et al. 2019; Woods and Van Vactor 2021). miRNAs are spatially and temporally expressed in the developing human brain from birth to adolescence; however, the contribution of TE-embedded

Corresponding author: Didier.Trono@epfl.ch

Article is online at http://www.rnajournal.org/cgi/doi/10.1261/rna.079100.122. Freely available online through the *RNA* Open Access option.

sequences to this process has never been investigated (Ziats and Rennert 2014). Indeed, the years proceeding birth and throughout childhood represent a crucial window in human brain development, characterized by extensive changes in size, cellular composition and functional processes such as synaptogenesis, myelination, and synaptic pruning (Silbereis et al. 2016; Dyck and Morrow 2017).

We therefore aimed to determine the prevalence of spatiotemporally expressed, annotated TE-embedded miRNAs in the developing human brain by reanalysis of small RNA-seq data available from the BrainSpan Atlas of the Developing Human Brain from 1- to 19-yr-old brains (Miller et al. 2014; Li et al. 2018). We computationally uncover dynamic spatiotemporal expression of numerous annotated TE-embedded miRNAs and a small number of previously undetected novel putative TE-embedded miRNAs, suggesting TE-sequence cooption as miRNAs may play a role in this important neurodevelopmental window. We provide the "Brain miRTExplorer" web application to facilitate interactive exploration of both annotated TE-embedded and non-TE-embedded miRNA spatiotemporal expression data, freely accessible for the community at https://tronoapps.epfl.ch/BrainmiRTExplorer/.

RESULTS

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

To determine spatiotemporal, small RNA expression in postnatal human brain development, we analyzed small RNA-seq data from 174 samples from 1 yr to 20 yr of age, encompassing 16 different brain regions, from 16 donors (nine male and seven female) available through the BrainSpan Atlas of the Developing Human Brain (Supplemental Fig. S1; Miller et al. 2014; Li et al. 2018). To enrich for different small RNA moieties, we separated sequencing reads into lengths of 18–25, 26–37, and 38–50 bp and intersected with Ensembl annotations, miRBase, the GtRNAdb database and our modified merged TE RepeatMasker data set (Kozomara and Griffiths-Jones 2014; Chan and Lowe 2016; Pontis et al. 2019; Turelli et al. 2020; Yates et al. 2019; Playfoot et al. 2021). As expected, the different read lengths enriched for annotated miRNAs, tRNAs and snoRNAs, respectively (Fig. 1A; Supplemental Fig. S2). By retaining the miRNA derived 18–25 bp reads, we detected the expression of 543/1871 annotated miRNAs (Fig. 1B; Supplemental Tables S1, S2).

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

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

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

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

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

To investigate if TE-embedded miRNAs are conserved and in syntenic locations among mammals, we performed

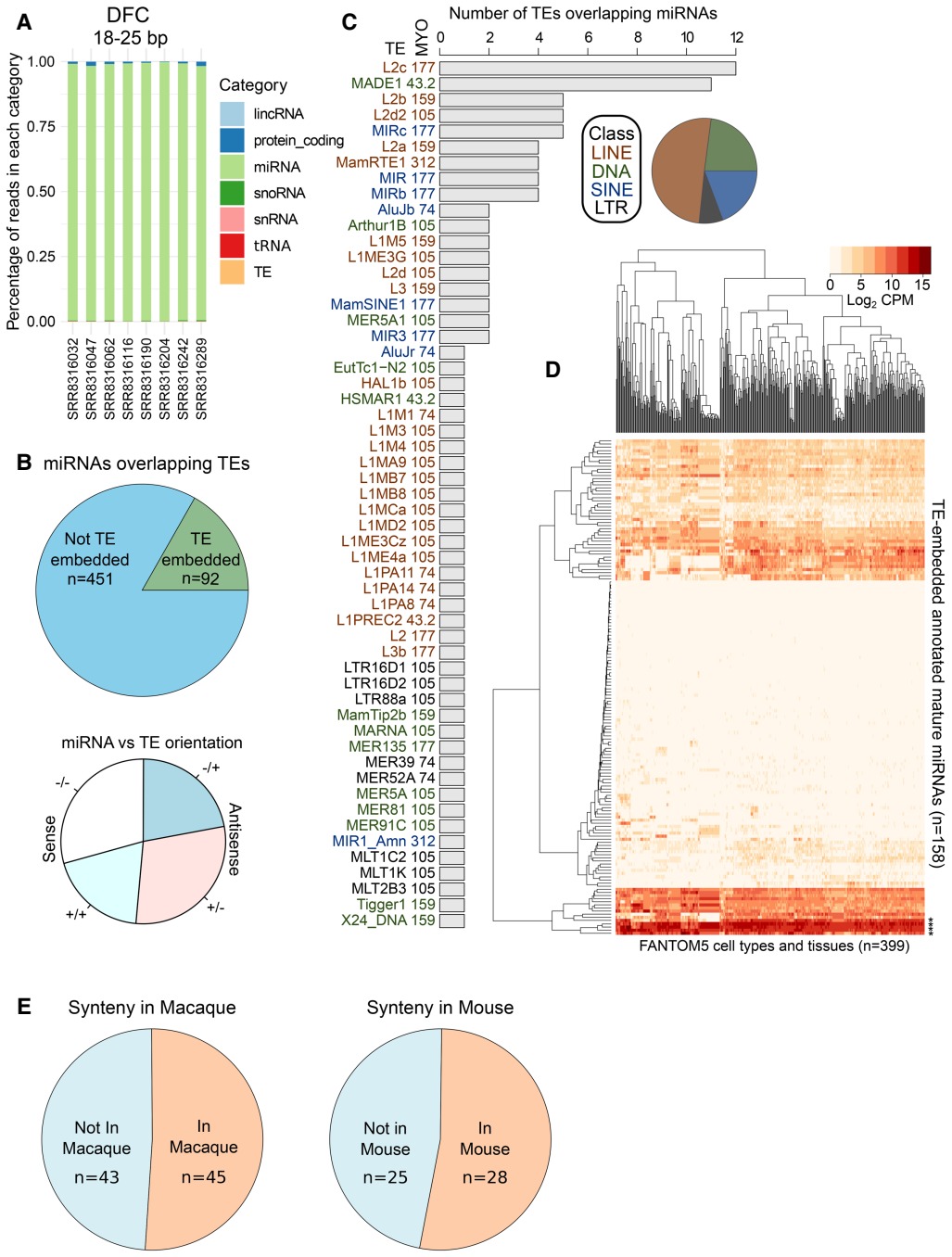

FIGURE 1. TEs contribute to annotated miRNAs in the child and adolescent human brain. (*A*) Stacked bar chart indicating the percentage of 18–25 bp reads overlapping different annotated genomic features for samples from the dorsolateral frontal cortex (DFC). If a TE overlaps an annotated feature (miRNA, tRNA, etc.) the feature takes preference. (*B*) Pie charts indicating the number of miRBase annotated miRNAs overlapping at least one TE (*top*), and their relative orientations (*bottom*). (*C*) Bar chart indicating the number of TEs overlapping miRBase annotated miRNAs and their class and age in million years old (MYO). (*D*) Expression in $\log_2$ counts per million+1 (CPM + 1) of mature TE-embedded miRNA in 399 cell types and tissues from FANTOM5 (De Rie et al. 2017). (*) Denotes miRNAs highlighted in the text. Samples comprise largely primary cells such as epithelial, fibroblast, endothelial, connective tissue, smooth muscle, immune, neural stem, dendritic, and pluripotent stem cells, among others (full list available in De Rie et al. 2017). (*E*) Pie charts indicating the number of human TE-embedded miRNAs detected in syntenic locations in macaque and mouse miRBase annotations (see also Supplemental Table S3 and Materials and Methods for detail).

a series of lift-overs from the human (hg19) to the macaque (rheMac8) and mouse (mm10) genomes. We next downloaded annotated miRNA coordinates from miRBase for macaque and mouse (Kozomara and Griffiths-Jones 2014). Of the 92 TE-embedded miRNAs detected in human (Fig. 1B), 88 remained after lift-over to macaque,

whereas the last four had no syntenic homolog. An intersect of these syntenic coordinates in macaque with the annotated miRNA coordinates from miRBase detected 45/88 TE-embedded miRNA loci which were largely annotated with the same miRNA name (Fig. 1E; Supplemental Table S3). Intersecting these with all TE coordinates from rheMac8 also showed that 40/45 overlapped the same or closely related TE subfamilies in macaque (Supplemental Table S3).

In mouse, 53 loci remained when TE-embedded miRNA coordinates in hg19 were lifted over to mm10. Of these, 28/53 overlapped coordinates from miRBase for mm10, again largely with the same name (Fig. 1E; Supplemental Table S3). The subsequent intersect with all TE loci for mouse resulted in 19/28 TE-embedded miRNA loci in human also overlapping the same or closely related TE subfamily in mouse (Supplemental Table S3). These were the older, more conserved TE subfamilies such as L2 (105–177 MYO), MIR (177 MYO), and some DNA elements (105–177 MYO) (Supplemental Table S3). Together, these analyses indicate that a large subset of TE-embedded miRNAs expressed in the human brain are conserved in mammalian lineages.

TE-embedded miRNAs exhibit spatiotemporal expression patterns

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

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

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

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

One of the most significantly differentially expressed TE-embedded miRNAs in the FB was the cancer- and cell proliferation-associated hsa-mir-378a, which displayed higher expression in childhood and a significant anticorrelation with donor age in the small RNA-seq data (Fig. 2A,B, left and C, left top; Li et al. 2015; Velazquez-Torres et al. 2018; Guo et al. 2019). To confirm our detection of this TE-embedded miRNA, we reanalyzed a publicly available Argonaute2 RNA-immunoprecipitation sequencing (AGO2 RIP-seq) data set from three adult human brains (Petri et al. 2019; GSE106810). AGO2 directly binds to mature processed miRNAs for incorporation into the RISC complex for targeting of mRNA (Kobayashi and Tomari 2016; Michlewski and Cáceres 2019), therefore AGO2-bound elements are likely to represent bona fide miRNAs rather than mere degradation products. Enrichment of reads in one out of three AGO2 RIP-seq samples was observed compared to the input sample, with a peak residing over the same sequence as the small RNA-seq data (Fig. 2C, left middle). This hsa-mir-378a miRNA is embedded in two intronic, MIRc elements arranged in opposite orientations, facilitating high confidence pre-miRNA hairpin precursor formation as determined by in silico miRNA folding analyses to detect hairpins with 90% of verified miRNA hairpin features (Fig. 2C, bottom left, and D, left; Tempel and Tahi 2012; Tav et al. 2016). The glycolysis-, cancer- and cell proliferation-associated hsa-mir-5683 was also significantly more expressed in childhood, with a significant anticorrelation with donor age and was detectable in at least one AGO2 RIP-seq sample, however was embedded in a solo 105 MYO MER5A1 element, which also facilitated pre-miRNA hairpin formation (Fig. 2A,B right; 2C, middle; and 2D, middle; Miao et al. 2020; Rong et al. 2020). One TE-embedded miRNA which was continually expressed in childhood and adolescent brains from small RNA-seq data and detectable in all samples of the AGO2 RIP-seq data was the cancer- and neuron-associated hsa-mir-582

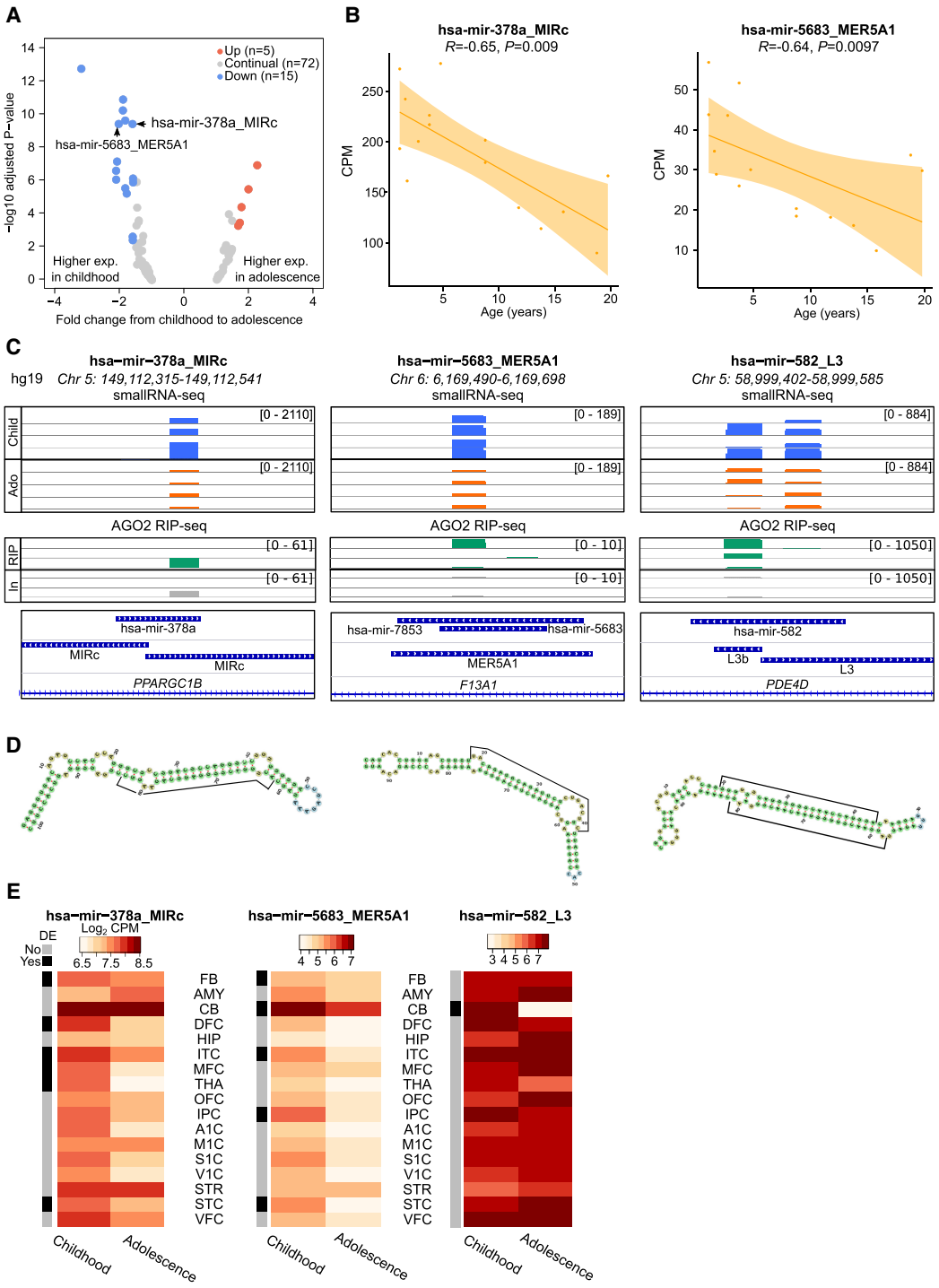

FIGURE 2. TE-embedded miRNAs are temporally expressed between child and adolescent human brains. (*A*) Volcano plot highlighting TE-embedded miRNAs significantly differentially expressed in FB (adjusted *P*-value ≤ 0.05, 1.5-fold change). (*B*) Dot plots showing the correlation of expression and age for specific TE-embedded miRNAs. Shaded area represents the variance. (*C*, *top*) Integrated genome viewer (IGV) visualization of four childhood (Child; blue) and four adolescent (Ado; orange) BAM files from DFC small RNA-seq data. (*Middle*) IGV visualization of three AGO2 RIP-seq (RIP; green) and three input (In; gray) BAM files from adult brain from Petri et al. 2019 (GSE106810). Read count is shown within square brackets. (*Bottom*) miRBase annotation, TE annotation, and gene annotations for hg19. (*D*) miRNA hairpin schematics from miRNAfold (Tempel and Tahi 2012; Tav et al. 2016) for the DNA sequences in *C*. Each hairpin structure exhibits 90% of verified miRNA hairpin features as previously defined (Tempel and Tahi 2012; Tav et al. 2016). Twenty-two bp peak sequences are highlighted by the black bars on arms of the hairpin. (*E*) Heatmaps showing regional expression in $\log_2$ counts per million (CPM), alongside differential expression results (black and gray bars). For differential expression, a linear model was generated for every expressed miRNA, with miRNA expression as response variable and stage as explanatory variable. TMM normalized expression estimates were used as input for the modeling. Region abbreviations are defined in Supplemental Figure S1.

(Fig. 2C, right; Fang et al. 2015; Zhang et al. 2015; Ding et al. 2019). This miRNA is embedded in two apposed L3 elements, again leading to an in silico predicted high-confidence precursor hairpin structure (Fig. 2D, right). Indeed, 18/92 miRNAs overlapped at least two TEs, with varying genomic orientations (Supplemental Table S5), although the majority of expressed TE-embedded miRNAs overlapped only one TE.

Different regions exhibit diverse miRNA temporal expression patterns

The temporal TE and gene expression profile of the human brain varies by region, notably with the cerebellum (CB) displaying a different transposcriptional and transcriptional landscape when compared to FB (Playfoot et al. 2021). We therefore next determined the temporal expression profile of miRNAs in childhood and adolescence in different individual brain regions (Fig. 2E; Supplemental Table S1, see Materials and Methods). hsa-mir-378a MIRc exhibited significantly higher expression in childhood, not only in combined FB samples, but also in individual FB regions such as the dorsolateral prefrontal cortex (DFC), the inferior temporal cortex (ITC), the medial prefrontal cortex (MFC) and the superior temporal cortex (STC), along with non-FB regions such as the mediodorsal nucleus of the thalamus (THA) (Fig. 2E, left). Similarly, hsa-mir-5683 MER5A1 had significantly higher expression in childhood versus adolescence in the FB combined, along with other individual FB regions, but also in the CB (Fig. 2E, middle). In both instances, expression in the CB was higher than for any other individual region. In contrast, the L3-embedded hsa-mir-582 exhibited continual high expression across childhood and adolescence for all regions, except the CB where hsa-mir-582 L3 expression was largely absent in adolescence and restricted to childhood (Fig. 2E, right). This provides a striking example of spatiotemporal control of the miRNA transcriptional landscape. Overall, these data demonstrate that the TE-embedded miRNA transcriptional landscape exhibits diverse spatiotemporal dynamics, with sometimes overt differences between childhood and adolescence for FB and non-FB regions.

TE-embedded miRNAs are spatially expressed

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

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

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

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

TE-embedded miRNAs target neurogenesis-associated genes

In order to determine possible functional relevance, we extracted predicted genic targets of TE-embedded miRNAs from the TargetScan database (Supplemental Table S6; Agarwal et al. 2015; McGeary et al. 2019). We specifically focused on conserved predicted target sites (defined by conserved branch lengths such as an 8mer or 7mer) of conserved miRNA families (defined by multiple-sequence miRNA alignments), as annotated in TargetScan (Agarwal et al. 2015; McGeary et al. 2019). Using this stringent list, we used two different gene ontology (GO) enrichment

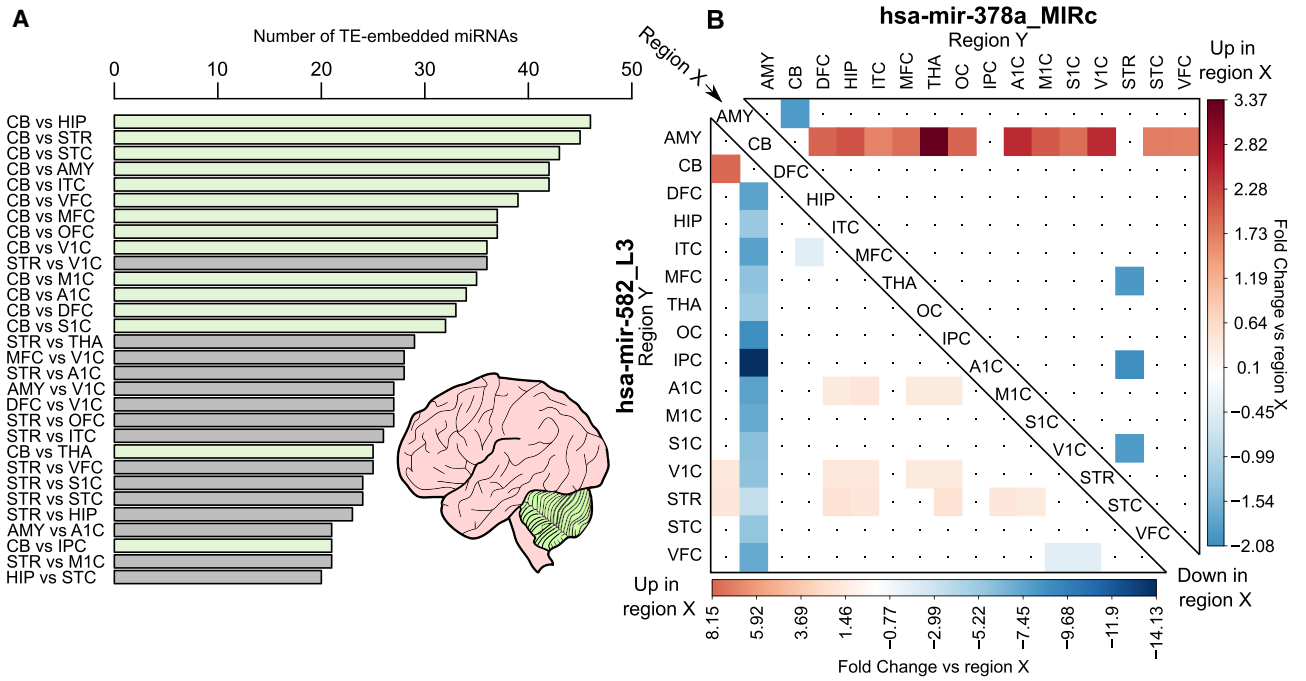

FIGURE 3. TE-embedded miRNAs exhibit spatial expression with major differences in the cerebellum. (*A*) Bar chart showing the number of differentially expressed TE-embedded miRNAs per regional comparisons (*P*-value ≤ 0.05, 1.5-fold change up or down). Only the top 30 comparisons are shown. For differential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

analysis tools (clusterProfiler, Yu et al. 2012; and Panther DB, Mi et al. 2021). Biological process analyses indicated that many target genes of TE-embedded miRNAs are enriched in neurogenesis-associated functions, alongside other enriched pathways such as regulation of transcription and metabolic processes, indicating the diversity of miRNA targets (Fig. 4A,C,E). For example, the L2c-embedded hsa-mir-374b targets all four genes involved in striatal medium spiny neuron differentiation (GO:0021773) and three out of four genes associated with glial cell fate specification (GO:0021780) and oligodendrocyte cell fate specification (GO:0021778) (Supplemental Table S7). Similarly, the L2b-embedded hsa-mir-493 was enriched in positive regulation of synaptic vesicle exocytosis (GO:2000302) and neurotransmitter receptor transport to plasma membrane (GO:00 98877), among others such as actin polymerization-dependent cell motility (GO:0070358) (Supplemental Table S7).

These two analyses also revealed significant enrichments in GO cell component terms for hsa-mir-374b such as synapse (GO:0045202), golgi apparatus (GO:0005794) and transcription regulator complex (GO:0005667) (Fig. 4B; Supplemental Table S7). hsa-mir-493 also exhibited diverse cellular component GO enrichment terms such as NMDA selective glutamate receptor complex (GO:001 7146), glial cell projection (GO:0097386) and integral component of postsynaptic specialization membrane (GO:00 99060), among other neurogenesis and non-neurogenesis terms such as chromatin silencing complex (GO:0005677) (Fig. 4D; Supplemental Table S7). The aforementioned continually expressed, L3-embedded hsa-mir-582, was also broadly enriched in neurogenesis and non-neurogenesis-associated biological process and cellular component terms such as astrocyte end-foot (GO:0097450), main axon (GO:0044304) and glycoprotein complex (GO:0090665) (Fig. 4E; Supplemental Table S7).

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

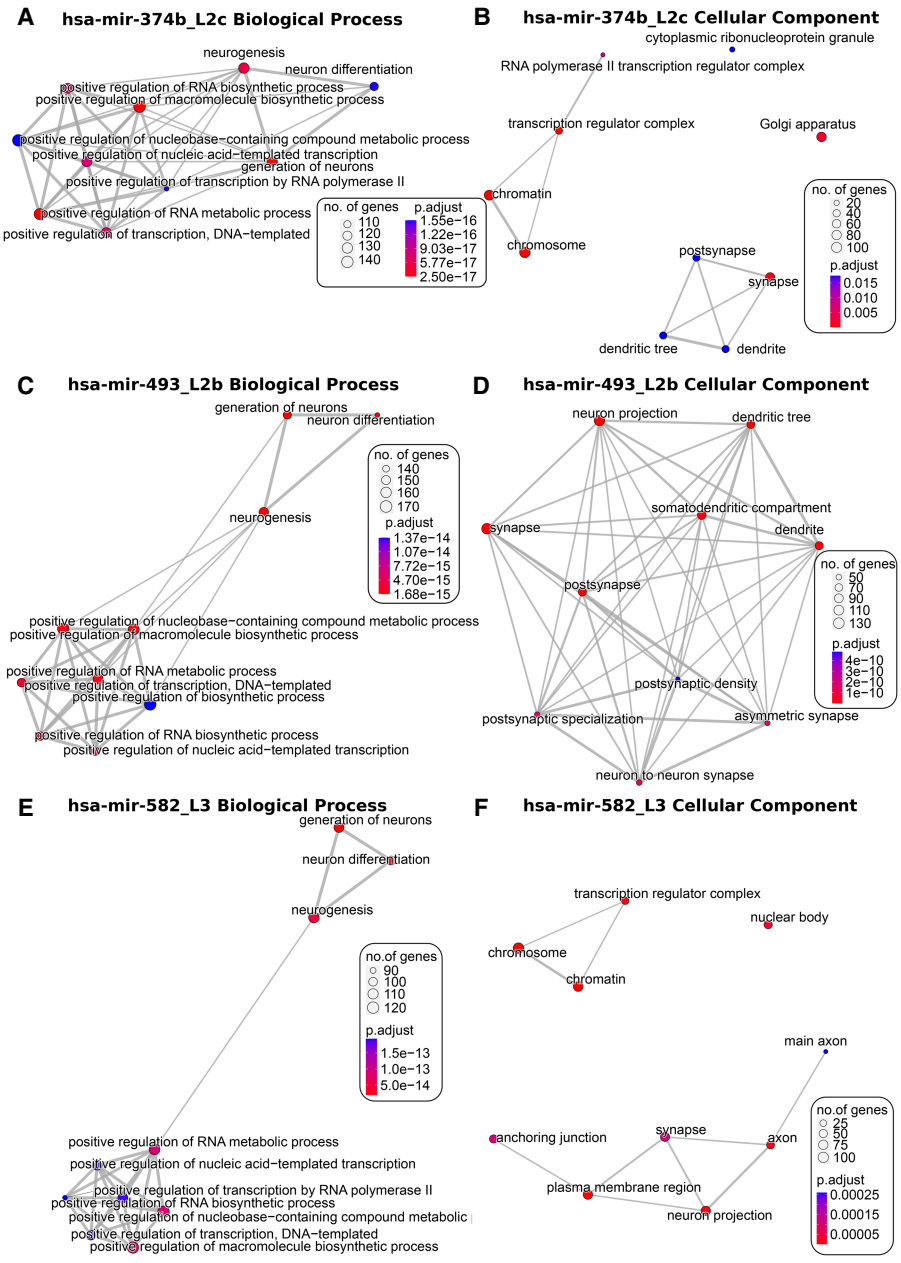

FIGURE 4. TE-embedded miRNA predicted targets are enriched in neurogenesis and other diverse gene ontology terms. ClusterProfiler emap network plots showing the top 10 enriched biological process (*A*,*C*,*E*) and cellular component (*B*,*D*,*F*) GO terms for specific TE-embedded miRNAs. The number of genes associated with each term is shown by point size and the adjusted *P*-value as the indicated color. Edges connect overlapping gene sets which cluster together, indicating the relatedness of terms. See also Supplemental Table S7 for GO enrichment from Panther DB.

highlight the specialized miRNA target networks in human neurogenesis and other diverse processes.

TEs contribute to novel putative miRNAs

Most studies rely on mapping small RNA-seq reads directly to miRNA annotations provided in miRBase. As a large proportion of annotated miRNAs are embedded in TEs, we reasoned that other TE loci could be contributing to previously undetected, novel miRNAs expressed in the brain. We therefore further investigated our unbiased, unique mapping to the whole genome used for detection of annotated TE-embedded miRNAs. To ensure robustness and to limit false positives, we used our custom RepeatMasker annotation (Turelli et al. 2020; Playfoot et al. 2021), alongside manual curation by inspecting BAM files from childhood and adolescent samples of the

DFC to detect a characteristic ~22 bp peak. These candidates were further refined by intersecting with genomic coordinates of TEs with at least one read from the AGO2 RIP-seq data. As TEs are inherently repetitive and have many thousands of copies, we next aimed to ensure that the sequences residing below putative miRNA peaks were indeed novel by searching for the sequences in miRBase. This resulted in a stringent list of eight novel nonannotated TE-embedded miRNA candidates (Supplemental Table S8). We next focused on two of these which met our strict criteria. The first was embedded in two apposed head-to-head, intronic MER3 elements and was confirmed with peaks detectable in the AGO2 RIP-seq data, suggestive of processed miRNA (Fig. 5A,B, left). Indeed, the 200 bp sequence covering the miRNA locus facilitated in silico hairpin structure formation with 90% of verified features and the 22 bp, 3p, and 5p peak sequences contributing to each arm of the hairpin (Fig. 5C, left). The same was observed for a novel putative miRNA embedded in a single MER5A element; however, AGO2 RIP-seq peaks overlapped the probable miRNA star sequence (the peak with fewer reads in the Brainspan samples) (Fig. 5A–C, right). AGO2 alternative strand loading can lead to shifts in the target profile of the miRNA-Induced Silencing Complex and may account for this (Medley et al. 2021).

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

DISCUSSION

Human brain development is a dynamic and highly regulated spatiotemporal process; however, the contribution of

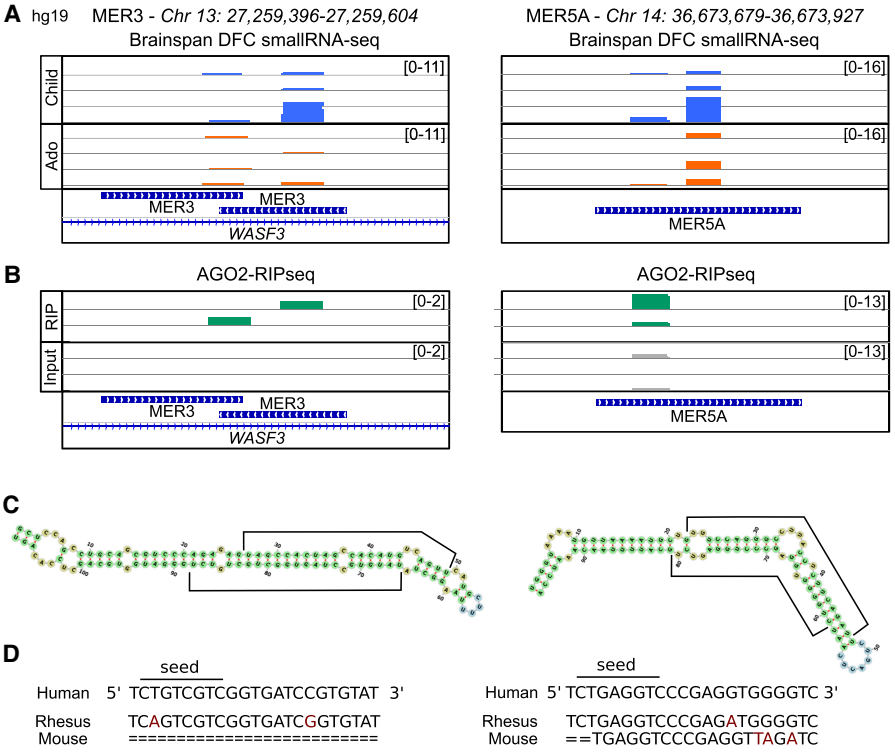

FIGURE 5. Novel, nonannotated TE-embedded miRNAs are present in child and adolescent brains. (*A*) IGV visualization of nonannotated TE-embedded miRNAs with classical 22 bp peaks in four childhood (blue) DFC BAM files and four adolescent (orange) DFC BAM files, alongside TE and gene annotations for hg19.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

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

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

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

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

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

MATERIALS AND METHODS

Data set download and preprocessing

Raw small RNA-seq FASTQ files from the BrainSpan Atlas of the Developing Human Brain (phs000755.v2.p1 provided by Dr. Nenad Sestan) were downloaded from the dbGaP-authorized access platform (Supplemental Acknowledgments; Miller et al. 2014; Li et al. 2018). The reads were first trimmed to remove Illumina small RNA 3′ sequencing adapters (TGGAATTCTCGG GTGCCAAGG) using FLEXBAR (version 3.5.0) with parameters –adapter-trim-end RIGHT –min-read-length 18 (Dodt et al. 2012). Trimmed reads were then divided by read length ranges of 18–25, 26–37, and 38–50 nt. Reads were then mapped to the human hg19 genome (GRCh37.p5) using Bowtie (version 2.3.4.1) with parameter –very-sensitive-local (Langmead et al. 2009). Read counts on different genomic features were quantified using featureCounts (version 1.6.2 of the Subread package) (Liao et al.

2014). Uniquely mapped reads were quantified with parameters -t exon -g gene_id -Q 10 and multimapped reads with parameters -M –fraction -t exon -g gene_id -Q 0. We used the parameters -s 1 and -s 2, to quantify sense and antisense reads, respectively, which were subsequently merged, keeping only the strand with the most reads. To confirm that specific read lengths were enriching for specific RNA moieties, the annotation of snoRNA, snRNA, miscRNA, scRNA and genes from Ensembl (GRCh37.p5, release 100) were used. For miRNAs and tRNAs, miRBase version 20 (Kozomara and Griffiths-Jones 2014) and tRNA annotations from GtRNAdb (release 19) were used, respectively (Chan and Lowe 2016). For repetitive sequences, a previously described in-house curated version of the RepeatMasker database was used (where fragmented LTR and internal segments belonging to a single integrant were merged) (Turelli et al. 2020; Playfoot et al. 2021). Exons of genes and TEs overlapping small RNAs in the same orientation were removed using BEDTools intersect (version 2.27.1) with default parameters, to prioritize reads falling on small RNAs (Quinlan and Hall 2010). To determine which expressed annotated miRNAs overlapped TEs, we used BEDTools to intersect the miRBase and our custom RepeatMasker merged TE annotations with a minimum of one base pair overlap. TE subfamily age estimates were obtained from DFAM (Hubley et al. 2016). BAM files were visualized using the Integrative Genome Viewer (Robinson et al. 2011).

Filtering and normalization

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

Differential expression analysis

Samples from 1 yr to 5 yr were considered as childhood and 9 yr to 20 yr as adolescence (Supplemental Fig. S1). To perform the aggregated temporal FB differential expression, the following brain regions were considered as FB: dorsolateral prefrontal cortex, inferior temporal cortex, medial prefrontal cortex, orbital prefrontal cortex, posterior inferior parietal cortex, primary auditory (A1) cortex, primary somatosensory (S1) cortex, primary visual (V1) cortex, superior temporal cortex, ventrolateral prefrontal cortex, and primary motor (M1) cortex (Supplemental Fig. S1). Independent temporal comparisons were performed without aggregations of multiple regions. For differential expression between regions, all samples regardless of age were used.

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

Correlation analysis

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

Expression of TE-embedded miRNAs in other tissues

Processed CPM expression data of mature miRNAs in 399 human samples (De Rie et al. 2017; file: human.srna.cpm.txt) were downloaded, and $\log_2$ CPMs of all annotated mature miRNAs nonoverlapping and overlapping-TE annotations were plotted with addition of a pseudocount of one.

miRNA precursor secondary structure analyses

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

AGO2 RIP-seq data

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

Evolutionary conservation

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

miRNA target prediction

miRNA target predictions were downloaded from TargetScan Human (Release 8.0) (Agarwal et al. 2015; McGeary et al. 2019;

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

Brain miRTExplorer application

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

DATA DEPOSITION

No new data were generated during the course of this study. Processed data can be interactively visualized using our "Brain miRTExplorer" application at https://tronoapps.epfl.ch/BrainmiRTExplorer/.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We thank all members of the Trono laboratory and Johan Jakobsson and Retha Ritter for helpful and insightful discussions. This study was supported by grants from the Personalized Health and Related Technologies (PHRT-508), the European Research Council (KRABnKAP, #268721; Transpos-X, #694658), and the Swiss National Science Foundation (310030_152879 and 310030B_173337) to D.T.

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

Received January 11, 2022; accepted June 15, 2022.

REFERENCES

Agarwal V, Bell GW, Nam JW, Bartel DP. 2015. Predicting effective microRNA target sites in mammalian mRNAs. *Elife* **4:** e05005. doi:10.7554/eLife.05005

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc* **57:** 289–300.

Blanchette M, Kent WJ, Riemer C, Elnitski L, Smith AFA, Roskin KM, Baertsch R, Rosenbloom K, Clawson H, Green ED, et al. 2004. Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res* **14:** 708–715. doi:10.1101/gr.1933104

Borchert GM, Holton NW, Williams JD, Hernan WL, Bishop IP, Dembosky JA, Elste JE, Gregoire NS, Kim J-A, Koehler WW, et al. 2011. Comprehensive analysis of microRNA genomic loci identifies pervasive repetitive-element origins. *Mob Genet Elements* **1:** 8–17. doi:10.4161/mge.1.1.15766

Cao X, Pfaff SL, Gage FH. 2007. A functional study of miR-124 in the developing neural tube. *Genes Dev* **21:** 531–536. doi:10.1101/gad.1519207

Chan PP, Lowe TM. 2016. GtRNAdb 2.0: An expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res* **44:** D184–D189. doi:10.1093/nar/gkv1309

Chang W, Cheng J, Allaire J, Xie Y, McPherson J. 2017. Shiny: web application framework for R. R package version, 1(5).

Chuong EB, Rumi MAK, Soares MJ, Baker JC. 2013. Endogenous retroviruses function as species-specific enhancer elements in the placenta. *Nat Genet* **45:** 325–329. doi:10.1038/ng.2553

Chuong EB, Elde NC, Feschotte C. 2016. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science* **351:** 1083–1087. doi:10.1126/science.aad5497

Chuong EB, Elde NC, Feschotte C. 2017. Regulatory activities of transposable elements: from conflicts to benefits. *Nat Rev Genet* **18:** 71–86. doi:10.1038/nrg.2016.139

Da Fonseca BHR, Domingues D S, Paschoal AR. 2019. mirtronDB: a mirtron knowledge base. *Bioinformatics* **35:** 3873–3874. doi:10.1093/bioinformatics/btz153

De Rie D, Abugessaisa I, Alam T, Arner E, Arner P, Ashoor H, Åström G, Babina M, Bertin N, Burroughs AM, et al. 2017. An integrated expression atlas of miRNAs and their promoters in human and mouse. *Nat Biotechnol* **35:** 872–878. doi:10.1038/nbt.3947

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-151 on chromosome 8q24.3 facilitates tumour cell migration and spreading through downregulating *RhoGDIA*. *Nat Cell Biol* **12:** 390–399. doi:10.1038/ncb2039

Ding H, Gao S, Wang L, Wei Y, Zhang M. 2019. Overexpression of miR-582-5p inhibits the apoptosis of neuronal cells after cerebral ischemic stroke through regulating PAR-1/Rho/Rho axis. *J Stroke Cerebrovasc Dis* **28:** 149–155. doi:10.1016/j.jstrokecerebrovasdis.2018.09.023

Dodt M, Roehr JT, Ahmed R, Dieterich C. 2012. FLEXBAR-flexible barcode and adapter processing for next-generation sequencing platforms. *Biology (Basel)* **1:** 895–905. doi:10.3390/biology1030895

Dyck LIV, Morrow EM. 2017. Genetic control of postnatal human brain growth. *Curr Opin Neurol* **30:** 114–124. doi:10.1097/WCO.0000000000000405

Elbarbary RA, Lucas BA, Maquat LE. 2016. Retrotransposons as regulators of gene expression. *Science* **351:** aac7247. doi:10.1126/science.aac7247

Fang L, Cai J, Chen B, Wu S, Li R, Xu X, Yang Y, Guan H, Zhu X, Zhang L, et al. 2015. Aberrantly expressed miR-582-3p maintains lung cancer stem cell-like traits by activating Wnt/β-catenin signalling. *Nat Commun* **6:** 8640. doi:10.1038/ncomms9640

Frankel LB, Di Malta C, Wen J, Eskelinen EL, Ballabio A, Lund AH. 2014. A non-conserved miRNA regulates lysosomal function and impacts on a human lysosomal storage disorder. *Nat Commun* **5:** 5840. doi:10.1038/ncomms6840

Friedman RC, Farh KK, Burge CB, Bartel DP. 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* **19:** 92–105. doi:10.1101/gr.082701.108

Garcia-Perez JL, Widmann TJ, Adams IR. 2016. The impact of transposable elements on mammalian development. *Development* **143:** 4101–4114. doi:10.1242/dev.132639

Guo XB, Zhang XC, Chen P, Ma LM, Shen ZQ. 2019. miR-378a-3p inhibits cellular proliferation and migration in glioblastoma multiforme by targeting tetraspanin 17. *Oncol Rep* **42:** 1957–1971.

Heras SR, Macias S, Plass M, Fernandez N, Cano D, Eyras E, Garcia-Perez JL, Cáceres JF. 2013. The Microprocessor controls

the activity of mammalian retrotransposons. *Nat Struct Mol Biol* **20:** 1173–1181. doi:10.1038/nsmb.2658

Heras SR, Macias S, Cáceres JF, Garcia-Perez JL. 2014. Control of mammalian retrotransposons by cellular RNA processing activities. *Mob Genet Elements* **4:** e28439. doi:10.4161/mge.28439

Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H, Diekhans M, Furey TS, Harte RA, Hsu F, et al. 2006. The UCSC Genome Browser database: update 2006. *Nucleic Acids Res* **34:** D590–D598. doi:10.1093/nar/gkj144

Huang D, Qiu S, Ge R, He L, Li M, Li Y, Peng Y. 2015. miR-340 suppresses glioblastoma multiforme. *Oncotarget* **6:** 9257. doi:10.18632/oncotarget.3288

Hubley R, Finn RD, Clements J, Eddy SR, Jones TA, Bao W, Smit AFA, Wheeler TJ. 2016. The Dfam database of repetitive DNA families. *Nucleic Acids Res* **44:** D81–D89. doi:10.1093/nar/gkv1272

Juźwik CA, Drake S S, Zhang Y, Paradis-Isler N, Sylvester A, Amar-Zifkin A, Douglas C, Morquette B, Moore CS, Fournier AE. 2019. microRNA dysregulation in neurodegenerative diseases: a systematic review. *Prog Neurobiol* **182:** 101664. doi:10.1016/j.pneurobio.2019.101664

Kefas B, Comeau L, Floyd DH, Seleverstov O, Godlewski J, Schmittgen T, Jiang J, Li Y, Chiocca EA, Lee J, et al. 2009. The neuronal microRNA miR-326 acts in a feedback loop with notch and has therapeutic potential against brain tumors. *J Neurosci* **29:** 15161–15168. doi:10.1523/JNEUROSCI.4966-09.2009

Kobayashi H, Tomari Y. 2016. RISC assembly: coordination between small RNAs and Argonaute proteins. *Biochim Biophys Acta Gene Regul Mech* **1859:** 71–81. doi:10.1016/j.bbagrm.2015.08.007

Kozomara A, Griffiths-Jones S. 2014. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res* **42:** 68–73. doi:10.1093/nar/gkt1181

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10:** R25. doi:10.1186/gb-2009-10-3-r25

Law CW, Chen Y, Shi W, Smyth GK. 2014. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* **15:** R29. doi:10.1186/gb-2014-15-2-r29

Li B, Wang Y, Li S, He H, Sun F, Wang C, Lu Y, Wang X, Tao B. 2015. Decreased expression of miR-378 correlates with tumor invasiveness and poor prognosis of patients with glioma. *Int J Clin Exp Pathol* **8:** 7016–7021.

Li M, Santpere G, Imamura Kawasawa Y, Evgrafov O V, Gulden FO, Pochareddy S, Sunkin SM, Li Z, Shin Y, Zhu Y, et al. 2018. Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* **362:** eaat7615. doi:10.1126/science.aat7615

Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30:** 923–930. doi:10.1093/bioinformatics/btt656

McGeary SE, Lin KS, Shi CY, Pham TM, Bisaria N, Kelley GM, Bartel DP. 2019. The biochemical basis of microRNA targeting efficacy. *Science* **366:** eaav1741. doi:10.1126/science.aav1741

Medley JC, Panzade G, Zinovyeva AY. 2021. microRNA strand selection: unwinding the rules. *Wiley Interdiscip Rev RNA* **12:** e1627. doi:10.1002/wrna.1627

Mi H, Muruganujan A, Casagrande JT, Thomas PD. 2013. Large-scale gene function analysis with the panther classification system. *Nat Protoc* **8:** 1551–1566. doi:10.1038/nprot.2013.092

Mi H, Ebert D, Muruganujan A, Mills C, Albou LP, Mushayamaha T, Thomas PD. 2021. PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res* **49:** D394–D403. doi:10.1093/nar/gkaa1106

Miao Y, Li Q, Sun G, Wang L, Zhang D, Xu H, Xu Z. 2020. miR-5683 suppresses glycolysis and proliferation through targeting pyruvate dehydrogenase kinase 4 in gastric cancer. *Cancer Med* **9:** 7231–7243. doi:10.1002/cam4.3344

Michlewski G, Cáceres JF. 2019. Post-transcriptional control of miRNA biogenesis. *RNA* **25:** 1–16. doi:10.1261/rna.068692.118

Miller JA, Ding S-L, Sunkin SM, Smith KA, Ng L, Szafer A, Ebbert A, Riley ZL, Royall JJ, Aiona K, et al. 2014. Transcriptional landscape of the prenatal human brain. *Nature* **508:** 199–206. doi:10.1038/nature13185

Navarro Gonzalez J, Zweig AS, Speir ML, Schmelter D, Rosenbloom KR, Raney BJ, Powell CC, Nassar LR, Maulding ND, Lee CM, et al. 2021. The UCSC Genome Browser database: 2021 update. *Nucleic Acids Res* **49:** D1046–D1057. doi:10.1093/nar/gkaa1070

Nowakowski TJ, Rani N, Golkaram M, Zhou HR, Alvarado B, Huch K, West JA, Leyrat A, Pollen AA, Kriegstein AR, et al. 2018. Regulation of cell-type-specific transcriptomes by microRNA networks during human brain development. *Nat Neurosci* **21:** 1784–1792. doi:10.1038/s41593-018-0265-3

Petri R, Malmevik J, Fasching L, Åkerblom M, Jakobsson J. 2014. miRNAs in brain development. *Exp Cell Res* **321:** 84–89. doi:10.1016/j.yexcr.2013.09.022

Petri R, Brattås PL, Sharma Y, Jonsson ME, Pircs K, Bengzon J, Jakobsson J. 2019. LINE-2 transposable elements are a source of functional human microRNAs and target sites. *PLoS Genet* **15:** e1008036. doi:10.1371/journal.pgen.1008036

Piriyapongsa J, Jordan IK. 2007. A family of human microRNA genes from miniature inverted-repeat transposable elements. *PLoS One* **2:** e203. doi:10.1371/journal.pone.0000203

Piriyapongsa J, Mariño-Ramírez L, Jordan IK. 2007. Origin and evolution of human microRNAs from transposable elements. *Genetics* **176:** 1323–1337. doi:10.1534/genetics.107.072553

Playfoot CJ, Duc J, Sheppard S, Dind S, Coudray A, Planet E, Trono D. 2021. Transposable elements and their KZFP controllers are drivers of transcriptional innovation in the developing human brain. *Genome Res* **31:** 1531–1545. doi:10.1101/gr.275133.120

Pontis J, Planet E, Offner S, Turelli P, Duc J, Coudray A, Theunissen TW, Jaenisch R, Trono D. 2019. Hominoid-specific transposable elements and KZFPs facilitate human embryonic genome activation and control transcription in naive human ESCs. *Cell Stem Cell* **24:** 724–735.e5. doi:10.1016/j.stem.2019.03.012

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26:** 841–842. doi:10.1093/bioinformatics/btq033

Qureshi IA, Mehler MF. 2012. Emerging roles of non-coding RNAs in brain evolution, development, plasticity and disease. *Nat Rev Neurosci* **13:** 528–541. doi:10.1038/nrn3234

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. *limma* powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43:** e47. doi:10.1093/nar/gkv007

Roberts JT, Cardin SE, Borchert GM. 2014. Burgeoning evidence indicates that microRNAs were initially formed from transposable element sequences. *Mob Genet Elements* **4:** e29255. doi:10.4161/mge.29255

Robinson MD, Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11:** R25. doi:10.1186/gb-2010-11-3-r25

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* **29:** 24–26. doi:10.1038/nbt.1754

Rong Z, Rong Y, Li Y, Zhang L, Peng J, Zou B, Zhou N, Pan Z. 2020. Development of a novel six-miRNA-based model to predict overall survival among colon adenocarcinoma patients. *Front Oncol* **10:** 26. doi:10.3389/fonc.2020.00026

Ruan J, Lou S, Dai Q, Mao D, Ji J, Sun X. 2015. Tumor suppressor miR-181c attenuates proliferation, invasion, and self-renewal abilities in glioblastoma. *Neuroreport* **26:** 66–73. doi:10.1097/WNR.0000000000000302

Sambandan S, Akbalik G, Kochen L, Rinne J, Kahlstatt J, Glock C, Tushev G, Alvarez-Castelao B, Heckel A, Schuman EM. 2017. Activity-dependent spatially localized miRNA maturation in neuronal dendrites. *Science* **355:** 634–637. doi:10.1126/science.aaf8995

Shao N-Y, Hu H, Yan Z, Xu Y, Hu H, Menzel C, Li N, Chen W, Khaitovich P. 2010. Comprehensive survey of human brain microRNA by deep sequencing. *BMC Genomics* **11:** 409. doi:10.1186/1471-2164-11-409

Short AK, Baram TZ. 2019. Early-life adversity and neurological disease: age-old questions and novel answers. *Nat Rev Neurol* **15:** 657–669. doi:10.1038/s41582-019-0246-5

Silbereis JC, Pochareddy S, Zhu Y, Li M, Sestan N. 2016. The cellular and molecular landscapes of the developing human central nervous system. *Neuron* **89:** 248–268. doi:10.1016/j.neuron.2015.12.008

Skalsky RL, Cullen BR. 2011. Reduced expression of brain-enriched microRNAs in glioblastomas permits targeted regulation of a cell death gene. *PLoS One* **6:** e24248. doi:10.1371/journal.pone.0024248

Smalheiser N, Torvik V. 2005. Mammalian microRNAs derived from genomic repeats. *Trends Genet* **21:** 322–326. doi:10.1016/j.tig.2005.04.008

Somel M, Liu X, Tang L, Yan Z, Hu H, Guo S, Jiang X, Zhang X, Xu G, Xie G, et al. 2011. MicroRNA-driven developmental remodeling in the brain distinguishes humans from other primates. *PLoS Biol* **9:** e1001214. doi:10.1371/journal.pbio.1001214

Spengler RM, Oakley CK, Davidson BL. 2014. Functional microRNAs and target sites are created by lineage-specific transposition. *Hum Mol Genet* **23:** 1783–1793. doi:10.1093/hmg/ddt569

Tav C, Tempel S, Poligny L, Tahi F. 2016. miRNAFold: a web server for fast miRNA precursor prediction in genomes. *Nucleic Acids Res* **44:** W181–W184. doi:10.1093/nar/gkw459

Tempel S, Tahi F. 2012. A fast ab-initio method for predicting miRNA precursors in genomes. *Nucleic Acids Res* **40:** e80. doi:10.1093/nar/gks146

Topol A, Zhu S, Hartley BJ, English J, Hauberg ME, Tran N, Rittenhouse CA, Simone A, Ruderfer DM, Johnson J, et al. 2016. Dysregulation of miRNA-9 in a subset of schizophrenia patient-derived neural progenitor cells. *Cell Rep* **15:** 1024–1036. doi:10.1016/j.celrep.2016.03.090

Turelli P, Playfoot C, Grun D, Raclot C, Pontis J, Coudray A, Thorball C, Duc J, Pankevich E V, Deplancke B, et al. 2020. Primate-restricted KRAB zinc finger proteins and target retrotransposons control gene expression in human neurons. *Sci Adv* **6:** eaba3200. doi:10.1126/sciadv.aba3200

Velazquez-Torres G, Shoshan E, Ivan C, Huang L, Fuentes-Mattei E, Paret H, Kim SJ, Rodriguez-Aguayo C, Xie V, Brooks D, et al. 2018. A-to-I miR-378a-3p editing can prevent melanoma progression via regulation of *PARVA* expression. *Nat Commun* **9:** 461. doi:10.1038/s41467-018-02851-7

Woods BJ, Van Vactor D. 2021. miRNA: local guardians of presynaptic function in plasticity and disease. *RNA Biol* **18:** 1014–1024. doi:10.1080/15476286.2020.1871214

Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Azov AG, Bennett R, et al. 2019. Ensembl 2020. *Nucleic Acids Res* **48:** D682–D688. doi:10.1093/nar/gkz966

Yu G, Wang LG, Han Y, He QY. 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16:** 284–287. doi:10.1089/omi.2011.0118

Zhang L, Dong LY, Li YJ, Hong Z, Wei WS. 2012. The microRNA miR-181c controls microglia-mediated neuronal apoptosis by suppressing tumor necrosis factor. *J Neuroinflammation* **9:** 211. doi:10.1186/1742-2094-9-211

Zhang X, Zhang Y, Yang J, Li S, Chen J. 2015. Upregulation of miR-582-5p inhibits cell proliferation, cell cycle progression and invasion by targeting Rab27a in human colorectal carcinoma. *Cancer Gene Ther* **22:** 475–480. doi:10.1038/cgt.2015.44

Ziats MN, Rennert OM. 2014. Identification of differentially expressed microRNAs across the developing human brain. *Mol Psychiatry* **19:** 848–852. doi:10.1038/mp.2013.93

MEET THE FIRST AUTHOR

Christopher Playfoot

***Meet the First Author(s)* is a new editorial feature within *RNA*, in which the first author(s) of research-based papers in each issue have the opportunity to introduce themselves and their work to readers of *RNA* and the RNA research community. Christopher Playfoot is the first author of this paper, "Transposable elements contribute to the spatiotemporal microRNA landscape in human brain development." Christopher did this work as a postdoctoral researcher in the Laboratory of Virology and Genetics of Didier Trono at the Swiss Federal Institute of Technology Lausanne in Switzerland.**

What are the major results described in your paper and how do they impact this branch of the field?

By reanalyzing 174 small RNA-seq samples from different regions of child and adolescent brains, we show that transposable elements contribute considerably to microRNA expression in the brain at different ages and in different regions. Predicted targets

Continued

of these microRNAs are involved in neurogenesis-associated processes, indicating that transposable elements contribute to gene expression networks fundamental to human brain development. This impacts the field as it indicates the spatiotemporal dynamic nature of transposable element-embedded microRNAs, which could be fundamental to neurodevelopment or disease.

What led you to study RNA or this aspect of RNA science?

MicroRNAs are an excellent example of transcriptional innovation which, in many cases, appears to have been driven by transposable elements throughout evolution. As both TEs and microRNAs are dynamically expressed in the brain, the contribution of transposable elements to microRNA expression in this tissue intrigued me. I have always found it fascinating how a 21 bp stretch of RNA can control gene expression.

Are there specific individuals or groups who have influenced your philosophy or approach to science?

My Ph.D. supervisor, Professor Ian Adams at The University of Edinburgh, made me the scientist I am today. He gave me the confidence and knowledge to work independently and ask the correct questions. Professor Didier Trono also influenced my philosophy to science immensely.

What are your subsequent near- or long-term career plans?

I am now a freelance scientific writer and editor, living and working on the road. It allows a flexible work-life balance, while maintaining a connection to the science I love. There is life beyond the lab!