Supplemental Material

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

Christopher J. Playfoot[1], Shaoline Sheppard[1], Evarist Planet[1] and Didier Trono[1, 2]

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

[2]Corresponding author: Didier.Trono@epfl.ch

Short title: TE-embedded miRNAs in brain development

Contents

Supplemental Figures

		Age (Years)													
		Childhood (C)					Adolescence (A)						C	A	Total
Brainspan Region	Abv.	1	2	3	4	5	9	12	14	16	19	20	Sum	Sum	Sum
Amygdala	AMY	2	1	0	2	0	2	1	0	1	0	1	5	5	10
Cerebellar cortex	CB	3	2	0	2	0	2	1	0	1	0	1	7	5	12
Dorsolateral prefrontal cortex	DFC	1	2	0	0	1	1	1	0	1	0	1	4	4	8
Hippocampus	HIP	1	1	1	2	0	2	1	1	0	0	1	5	5	10
Inferior temporal cortex	ITC	2	2	1	2	1	2	1	1	1	1	1	8	7	15
Medial prefrontal cortex	MFC	2	2	1	0	1	1	0	0	1	1	1	6	4	10
Mediodorsal nucleus of the thalamus	THA	1	1	1	1	1	0	0	1	1	0	1	5	3	8
Orbital prefrontal cortex	OFC	1	2	1	1	0	1	1	0	1	1	1	5	5	10
Posterior inferior parietal cortex	IPC	1	0	1	2	1	1	0	0	1	1	1	5	4	9
Primary auditory (A1) cortex	A1C	1	1	0	1	0	1	1	1	1	1	1	3	6	9
Primary motor (M1) cortex	M1C	2	1	1	2	1	1	1	0	1	1	1	7	5	12
Primary somatosensory (S1) cortex	S1C	2	1	1	2	1	1	1	0	1	1	1	7	5	12
Primary visual (V1) cortex	V1C	1	1	1	2	1	1	1	1	1	1	1	6	6	12
Striatum	STR	2	1	1	1	1	1	0	1	1	0	1	6	4	10
Superior temporal cortex	STC	2	2	1	2	1	2	1	0	1	1	1	8	6	14
Ventrolateral prefrontal cortex	VFC	2	2	1	1	1	2	1	0	1	1	1	7	6	13
												Sum	94	80	174
											Forebrain Sum		66	58	124

Figure S1. Summary of samples analysed from the Brainspan Atlas of the Developing Human Brain. Table highlighting the number of samples per brain region (forebrain regions highlighted in blue) and age. Red indicates no samples for that region and age, orange indicates between one and two samples for that region and age and green indicates three or more samples.

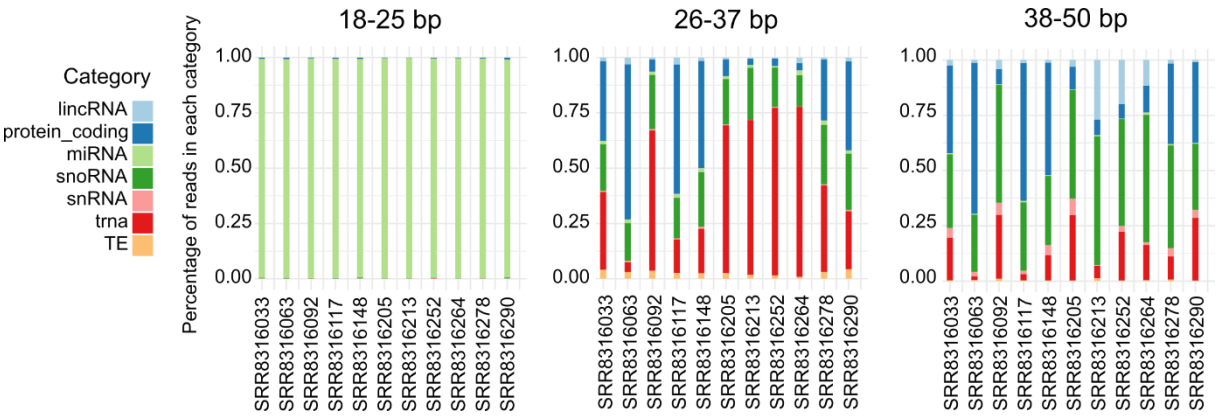


Figure S2. Different read lengths enrich for different small RNA moieties. Stacked bar charts indicating the percentage of 18-25bp, 26-37bp or 38-50bp reads overlapping different annotated genomic features for samples from the hippocampus (HIP).

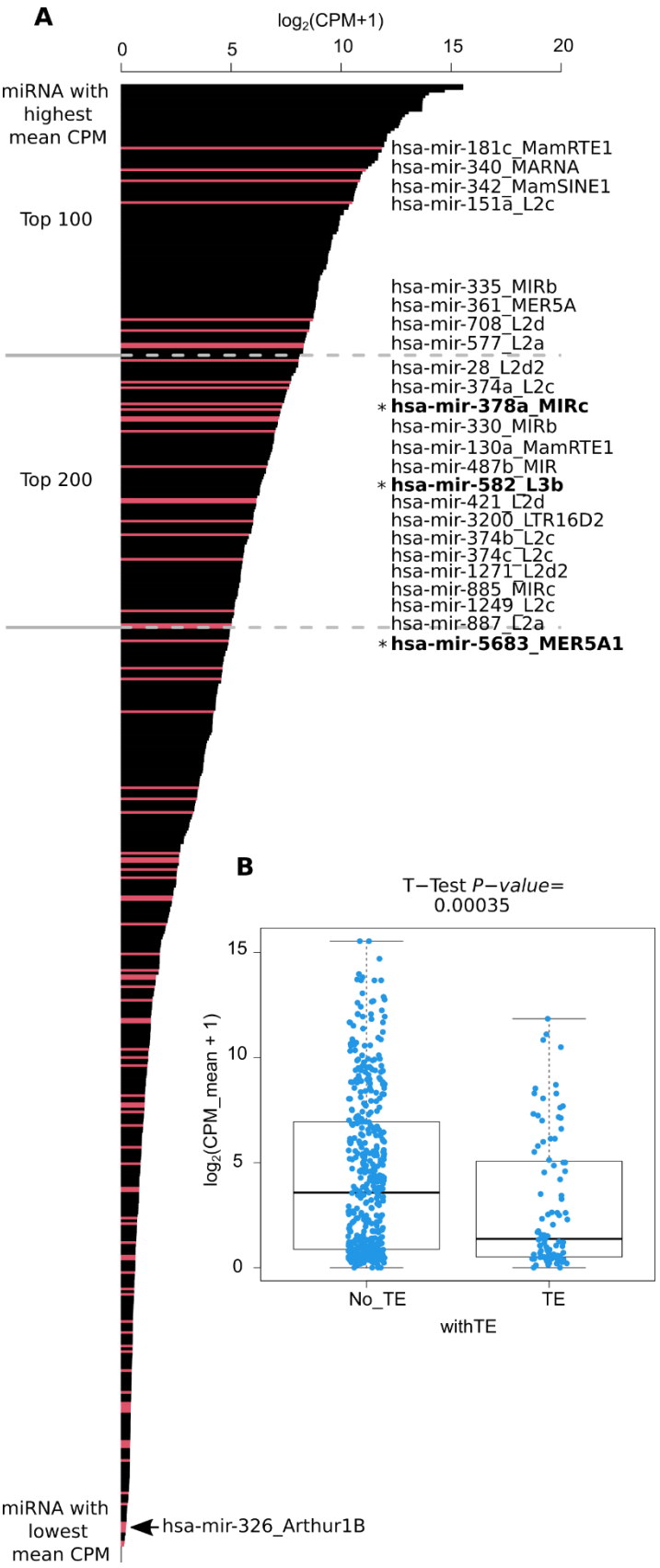

A
log2(CPM+1)
0
5
10
15
20
miRNA with highest mean CPM
Top 100
Top 200
hsa-mir-181c_MamRTE1
hsa-mir-340_MARNA
hsa-mir-342_MamSINE1
hsa-mir-151a_L2c
hsa-mir-335_MIRb
hsa-mir-361_MER5A
hsa-mir-708_L2d
hsa-mir-577_L2a
hsa-mir-28_L2d2
hsa-mir-374a_L2c
*hsa-mir-378a_MIRc
hsa-mir-330_MIRb
hsa-mir-130a_MamRTE1
hsa-mir-487b_MIR
*hsa-mir-582_L3b
hsa-mir-421_L2d
hsa-mir-3200_LTR16D2
hsa-mir-374b_L2c
hsa-mir-374c_L2c
hsa-mir-1271_L2d2
hsa-mir-885_MIRc
hsa-mir-1249_L2c
hsa-mir-887_L2a
*hsa-mir-5683_MER5A1
B
T−Test P−value= 0.00035
log2(CPM_mean + 1)
15
10
5
0
No_TE
TE
withTE
miRNA with lowest mean CPM
hsa-mir-326_Arthur1B

Figure S3. Comparison of expression levels of TE-embedded miRNAs versus non-TE-embedded miRNAs. (*A*) Barchart showing miRNA expression in log_2CPM+1 from highest to lowest expression. Red bars denote a TE-embedded miRNA. Names are shown in order of expression for TE-embedded miRNAs in the top 200 most expressed. Bold and * denote TE-embedded miRNAs discussed in Fig. 2. (*B*) Boxplot showing the log_2CPM+1 mean expression values for all non-TE-embedded miRNAs and TE-embedded miRNAs. *P-value* from Students T-test is shown above.

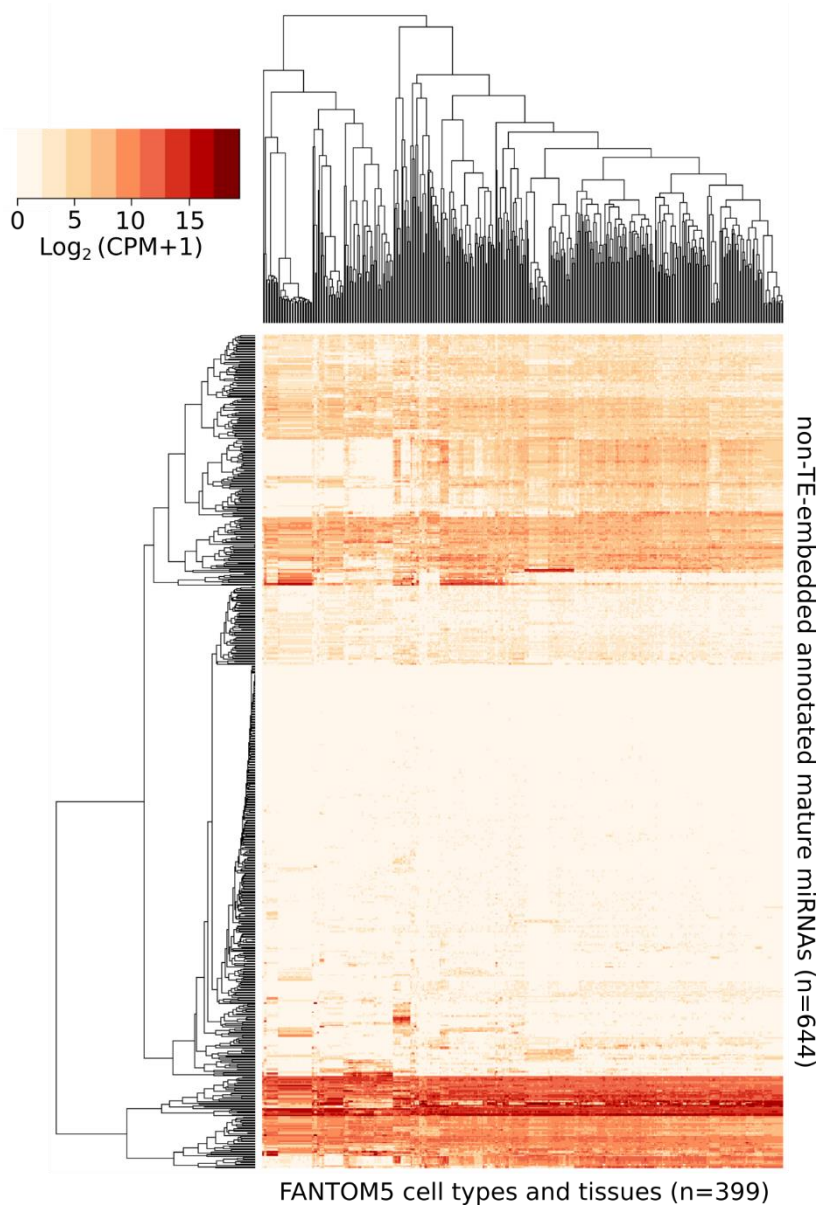


Figure S4. Expression of non-TE-embedded miRNAs in FANTOM5 dataset. Expression in $\log_2$ counts per million+1 (CPM+1) of mature non-TE-embedded miRNA in 399 cell types and tissues from FANTOM5 comprising largely primary cells such as epithelial, fibroblast, endothelial, connective tissue, smooth muscle, immune, neural stem, dendritic and pluripotent stem cells, among others (De Rie et al. 2017). Data for both 5p and 3p mature miRNAs are included, hence the greater number of miRNAs compared to Fig. 1B.

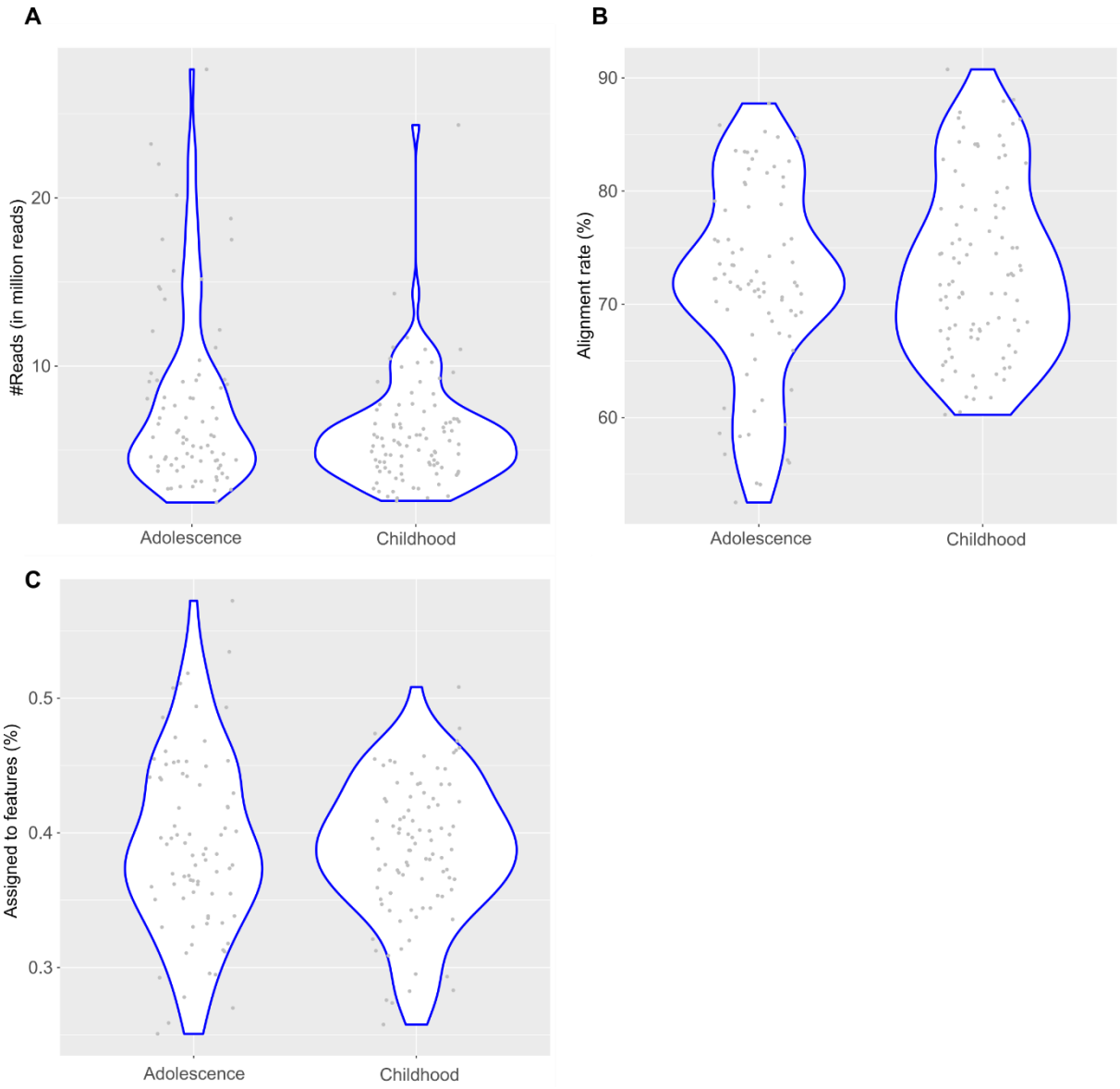
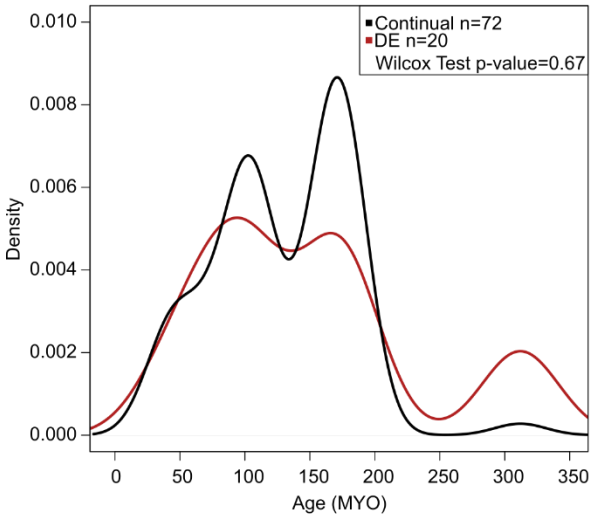
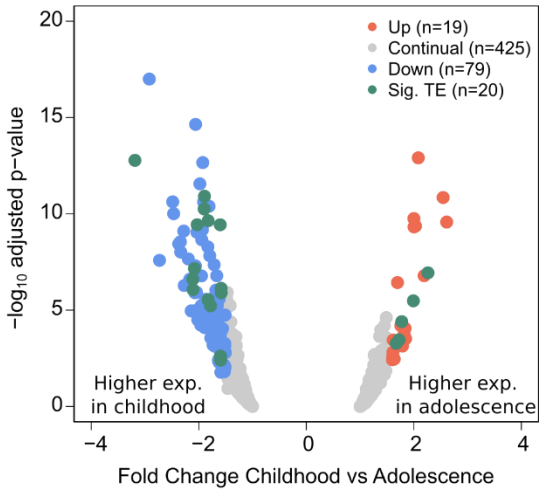


Figure. S5. Sample quality control. Violin plots showing (*A*) number of reads per million (*P-value*=0.08, Wilcoxon-test), (*B*) percentage read alignment rate (*P-value*=0.66, Wilcoxon-test) and (*C*) percentage of reads assigned to features (*P-value*=0.99, Wilcoxon-test) for small RNA-seq samples from adolescence and childhood.

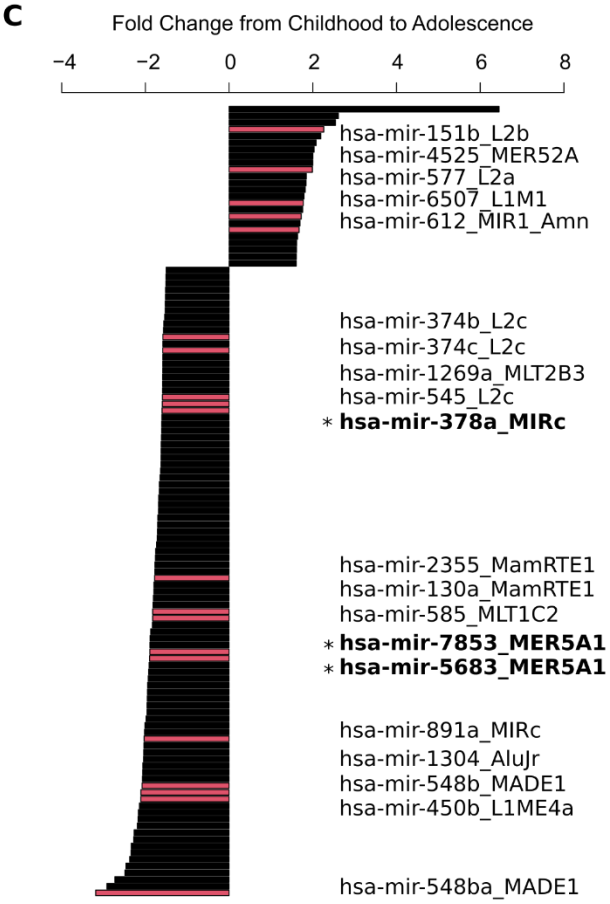
A



B



C



D

Comparison	Non-TE-embedded miRNA	TE-embedded miRNA
Number of miRNAs detected	451	92
Number up in adult	19 (4.2%)	5 (5.4%)
Number up in childhood	79 (17.5%)	15 (16.3%)
Overall expression (mean CPM) (*P-value=0.0004*)	821.8018	136.9767
Average fold change (childhood vs adolescence)	-0.595999791	-0.625604252

Figure S6. Differentially expressed TE-embedded miRNAs are a similar age to those continually expressed and their behaviour is similar to non-TE-embedded miRNAs (*A*) Density plot of TE age in million years old (MYO) of continually and differentially expressed TE-embedded miRNAs. (*B*) Volcano plot highlighting non-TE-embedded miRNAs significantly differentially expressed (adjusted *P-value* ≤ 0.05, 1.5-fold change). (*B*) Volcano plot highlighting non-TE-embedded and TE-embedded miRNAs significantly differentially expressed in FB (adjusted *P-value* ≤ 0.05, 1.5-fold change). (*C*) Barchart showing miRNA fold change from childhood to adolescence for all significantly differentially expressed miRNAs (adjusted *P-value* ≤ 0.05 and fold change ≥ 1.5 or ≤ -1.5). Red bars denote a TE-embedded miRNA. Names are shown in order of fold change. Bold and * denote TE-embedded miRNAs discussed in Fig. 2. (*D*) Table summarizing non-TE and TE-embedded miRNA comparisons. *P-value* represents students T-test on mean CPM values.

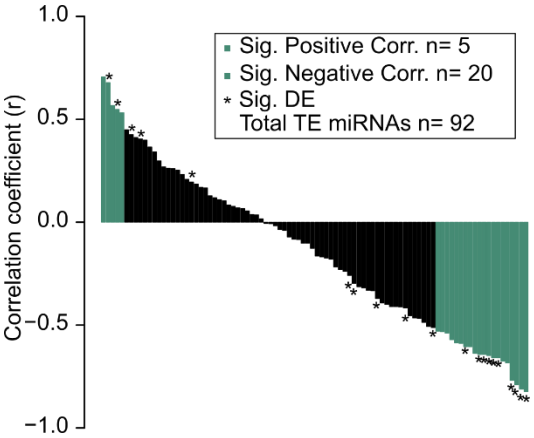


Figure S7. TE-embedded miRNA expression correlates with donor age. Bar chart indicating the spearman correlation coefficient of TE-embedded miRNA expression with donor age. Significance was determined as P-value ≤ 0.05. * denotes TE-embedded miRNA which are also differentially expressed in Fig. 2A. See also Supplemental Table S2.

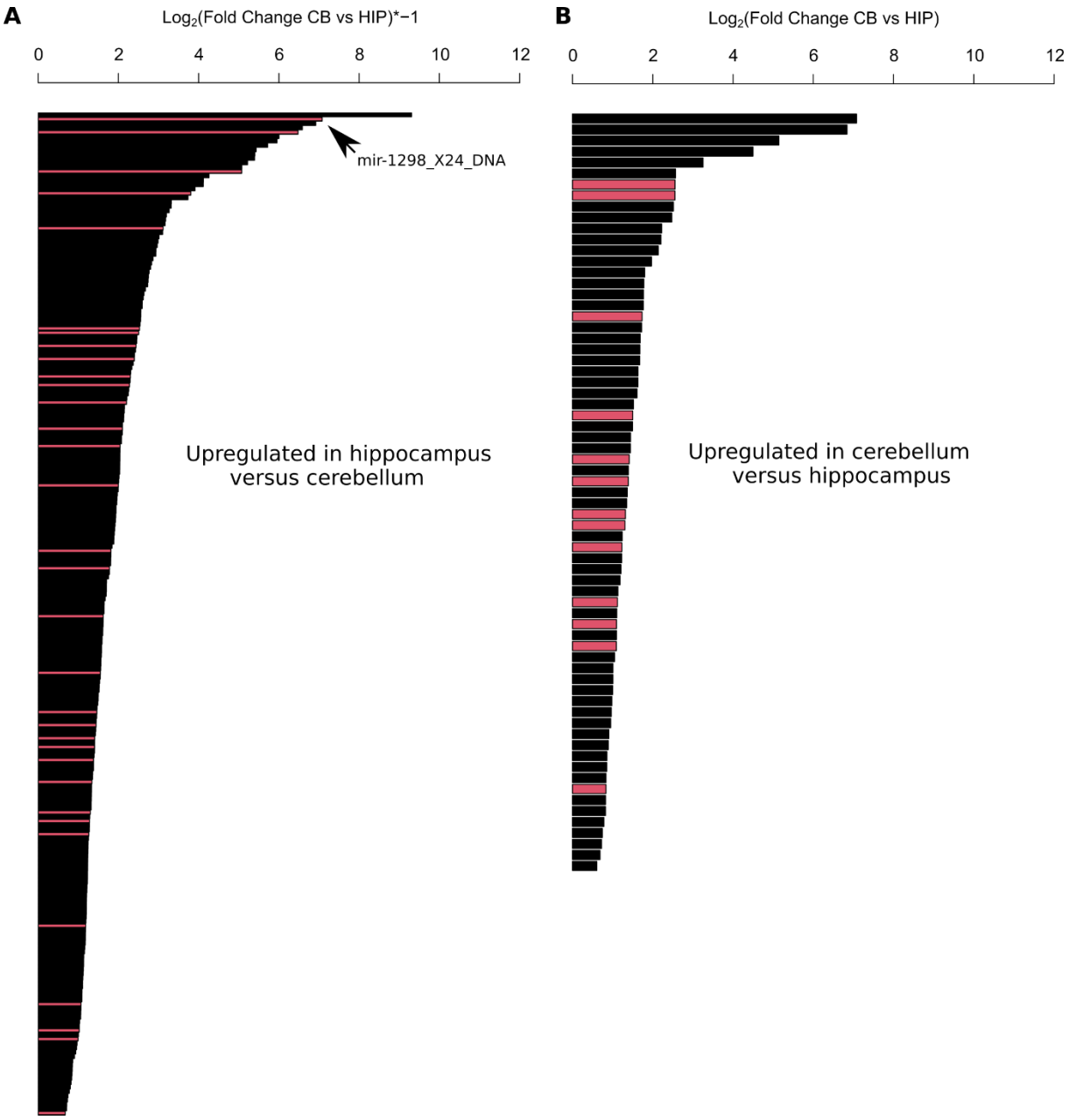


Figure S8. Comparison of non-TE-embedded miRNAs versus TE-embedded miRNAs in cerebellum versus hippocampus. (*A*) Barchart showing Log_2 Fold Change CB vs HIP *-1 for 230 miRNAs significantly upregulated in hippocampus compared to cerebellum. Red bars denote a TE-embedded miRNA. (*B*) Barchart showing Log_2 Fold Change CB vs HIP for 69 miRNAs significantly upregulated in cerebellum compared to hippocampus. Red bars denote a TE-embedded miRNA.

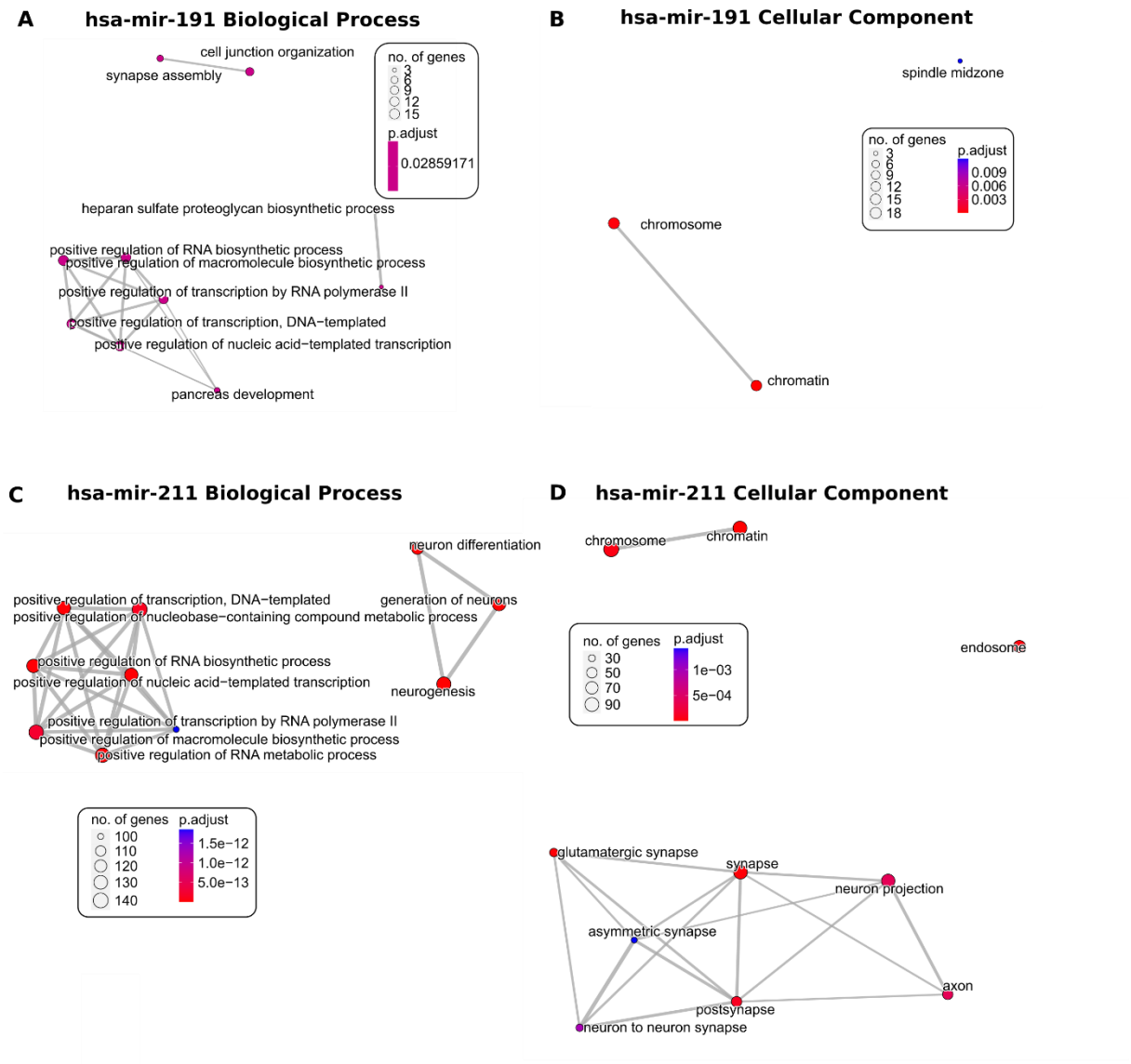


Figure S9. Non-TE-embedded miRNA predicted targets are enriched in neurogenesis and other diverse gene ontology terms. ClusterProfiler emap network plots showing the top ten enriched biological process (*A & C*) and cellular component (*B & D*) GO terms for specific non-TE-embedded miRNAs. The number 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.

Description:
Brain mirTExplorer is a quick and easy tool to investigate the behaviour of your favourite miRNA in childhood and adolescent brains. It uses small RNA-seq data from the Brainspan Atlas of the Developing Human Brain (Miller et al. 2014; Li et al. 2018) which we have reanalysed with an unbiased TE-centric approach.
The app provides heatplots of differential expression between childhood and adolescence for 16 different brain regions and differential expression between regions, regardless of age. Gene ontology analyses can be easily performed to highlight the top 10 enriched biological process, cellular component or molecular function GO terms.

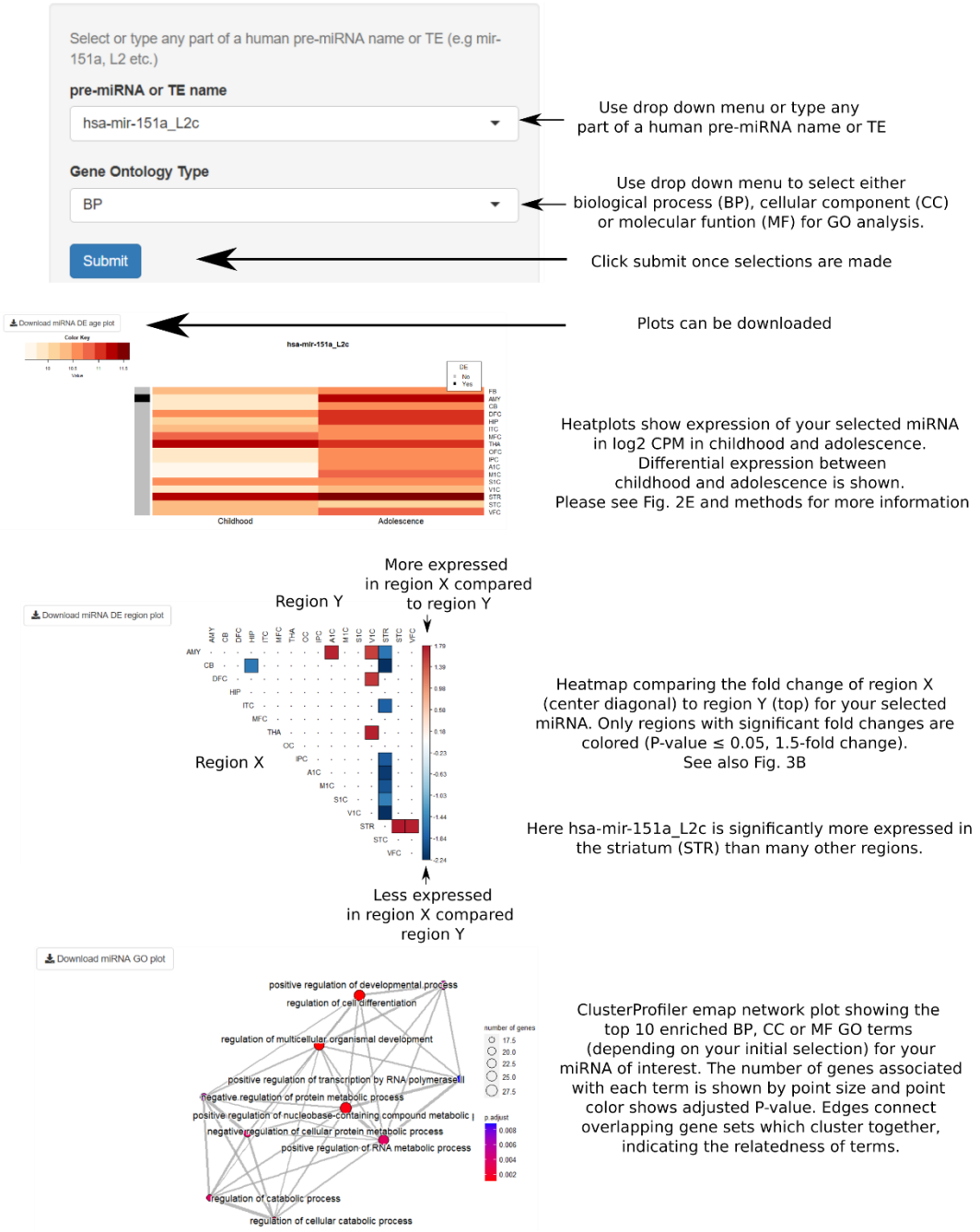


Figure S10. Example usage of the Brain mirTExplorer application.
https://tronoapps.epfl.ch/BrainmiRTExplorer/

Supplemental Table Descriptions

Supplemental Table S1 – Expression of all TE-embedded miRNA and non-TE-embedded miRNA loci in counts per million (CPM) in childhood and adolescence, alongside their differential expression between childhood and adolescence. TE age as defined in DFAM is also shown. The first tab represents the collated analysis on all forebrain samples and the adjusted *P-value* is shown and subsequent tabs provide differential expression analyses showing the non-adjusted *P-value* due to low sample numbers (see Fig. S1 for sample information).

Supplemental Table S2 - Summary table of miRNA sequences from miRBase for all detectably expressed miRNAs in small RNA-seq data from human brains.

Supplemental Table S3 – Summary table of miRNA conservation in macaque and mouse.

Supplemental Table S4 - Spearman correlation analysis of TE-embedded miRNA expression to donor age in collated forebrain samples.

Supplemental Table S5 – Information about miRNAs embedded in two individual TEs and their respective orientations.

Supplemental Table S6 – Predicted genic TE-embedded miRNA targets from TargetScan. Only conserved miRNAs and conserved targets are included and columns are as defined in Agarwal et al. 2015 and McGeary et al. 2019. File used: Predicted_Targets_Context_Scores.default_predictions.txt.

Supplemental Table S7 - PantherDB gene ontology biological process and cellular component analyses for hsa-mir374b_L2c, hsa-mir-493_L2b and hsa-mir-582_L3 targets from TargetScan (Supplemental Table S5).

Supplemental Table S8 - Expression of putative non-annotated TE-embedded miRNAs in counts per million (CPM) in childhood and adolescence in forebrain, alongside their differential expression between childhood and adolescence from Brainspan data set. TE age as defined in DFAM is also shown. Only those TE loci with detectable reads in the Brainspan data set, at least one read present in AGO2 RIP-seq samples and whose sequence is not detectable in miRBase are shown.

Supplemental Acknowledgements

Submission of BrainSpan Atlas of the Developing Human Brain data (phs000406.v2.p1) to dbGaP was provided by Dr. Nenad Sestan. Collection of the data and analysis was supported by grants from the National Institutes of Health (MH089929, MH081896, and MH090047). Additional support was provided by the Kavli Foundation, a James S. McDonnell Foundation Scholar Award, NARSAD, and the Foster-Davis Foundation.