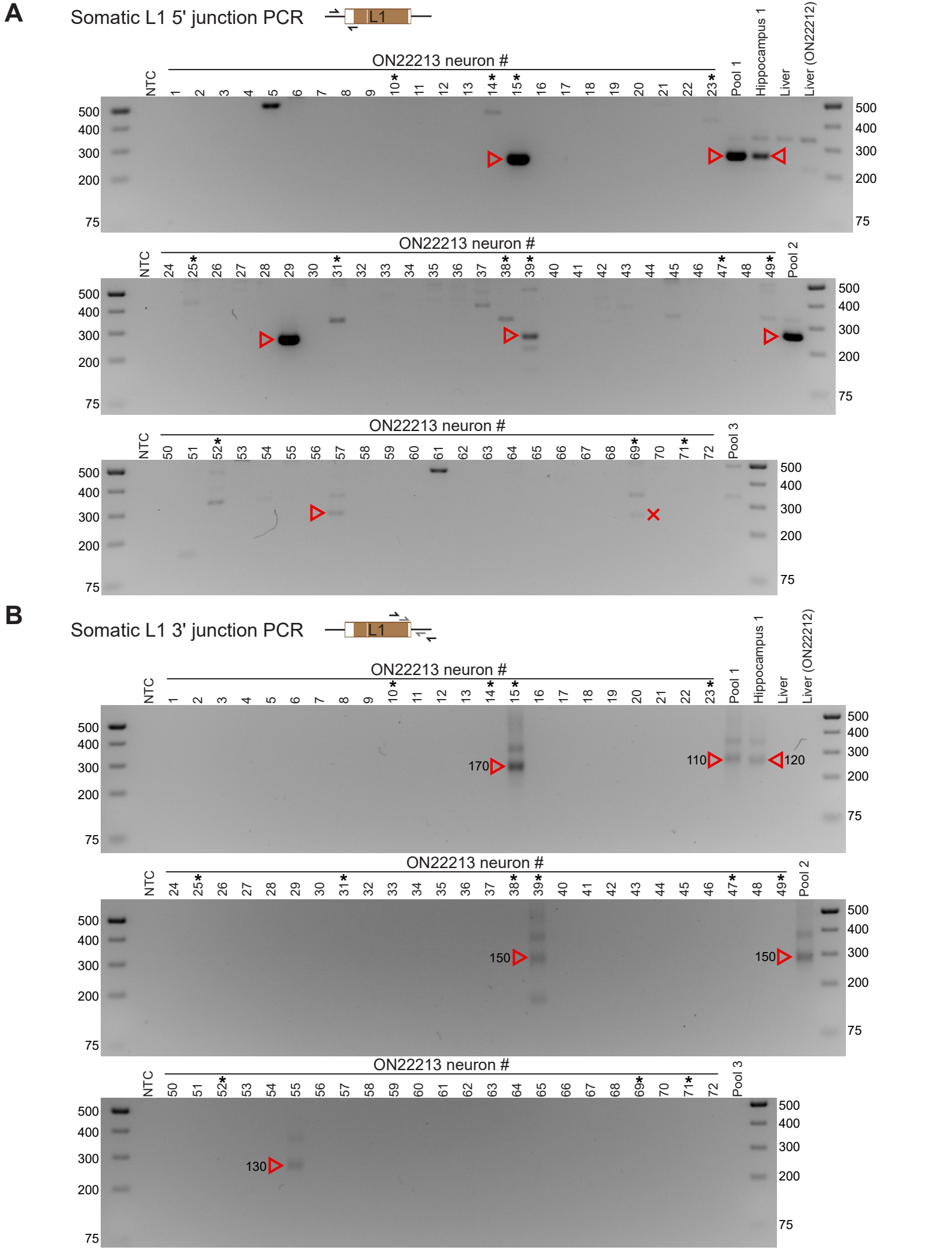
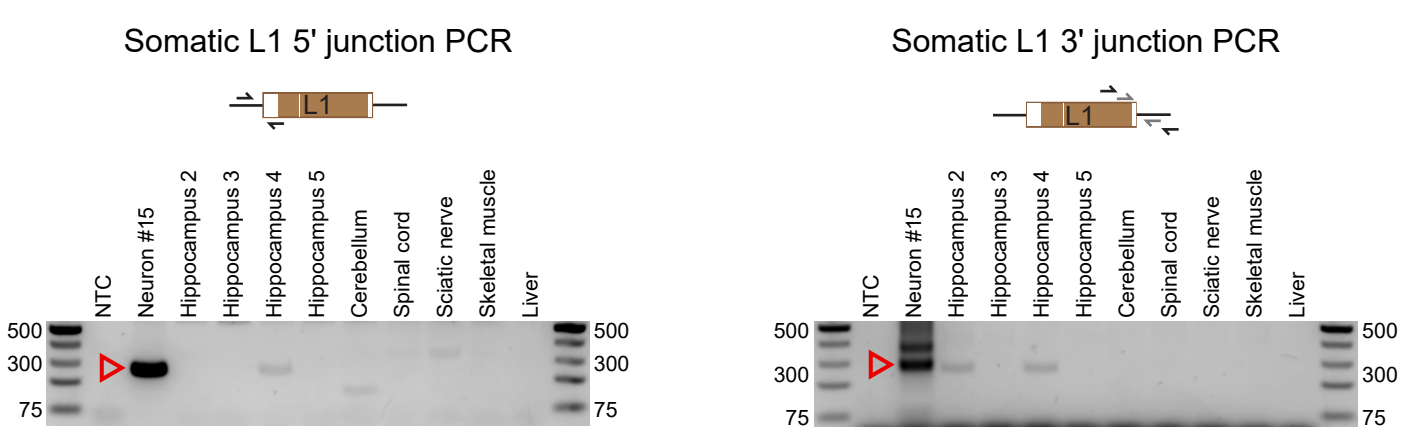
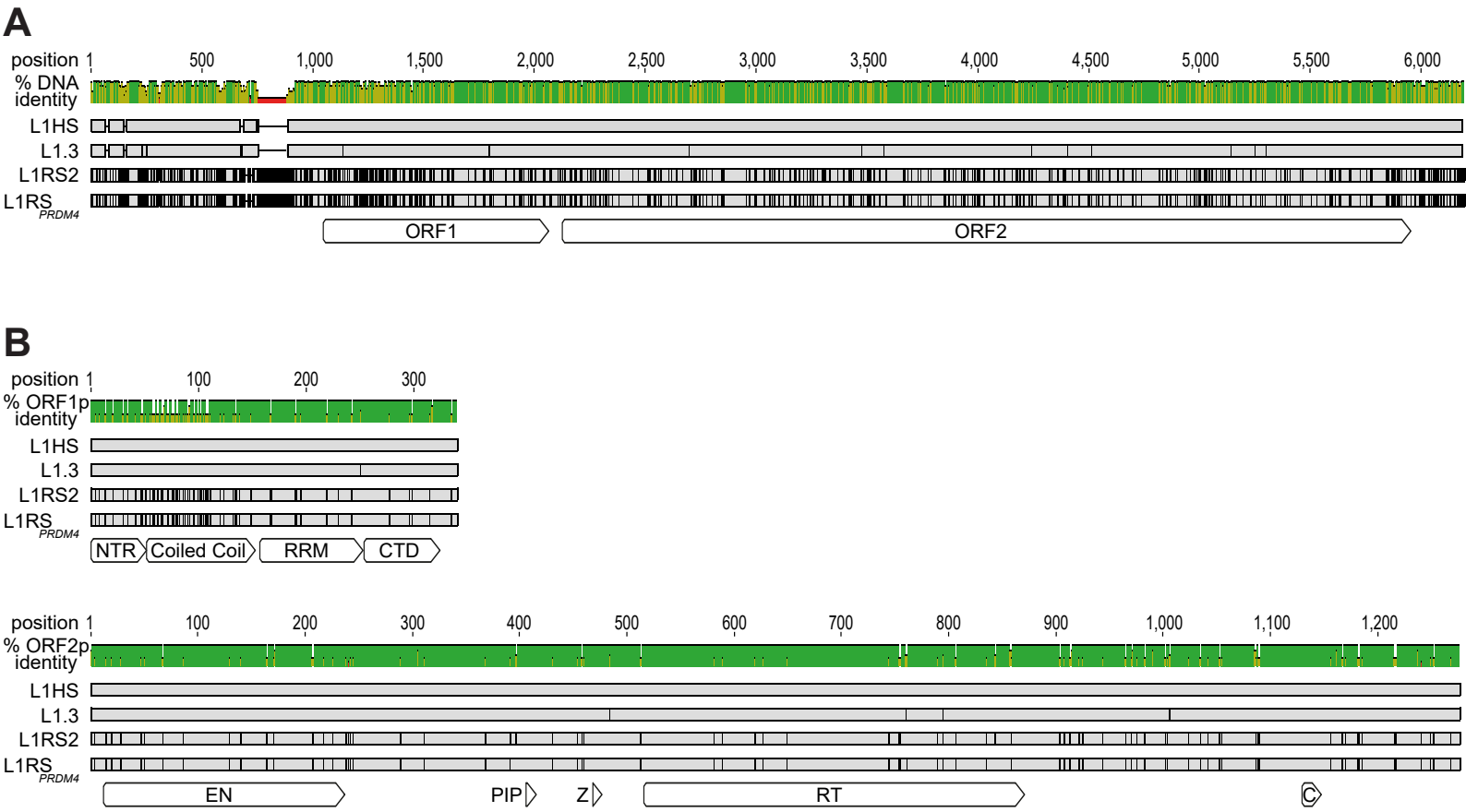

A
Somatic L1 5' junction PCR
L1
ON22213 neuron #
NTC
Pool 1
Hippocampus 1
Liver
Liver (ON22212)
Pool 2
Pool 3
500
400
300
200
75
B
Somatic L1 3' junction PCR
170
110
120
150
130
C

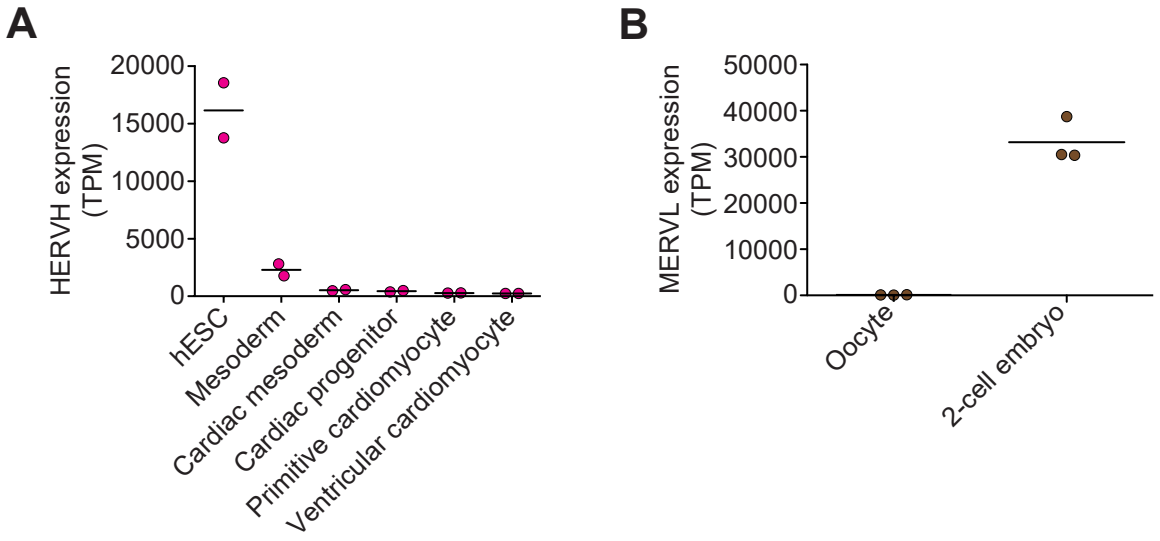
Supplemental Figure S1. $L1RS_{somatic}$ PCR validation results in additional single neurons. (*A*) PCR amplification results for the somatic L1 insertion 5′ junction (primer positions indicated in schematic at top). Reaction input consisted of non-template control (NTC), 72 MDA-amplified ON22213 hippocampal neurons, 3 pools of ~20 MDA-amplified ON22213 hippocampal neurons, bulk ON22213 hippocampus and liver DNA, and bulk ON22212 liver. scWGS and RC-seq were applied to neurons marked with an asterisk. Red arrowheads indicate amplicons confirmed as on-target by capillary sequencing. A red cross indicates an amplicon of approximately the expected size that was subsequently determined by capillary sequencing to be off-target. **(*B*)** As for (*A*), except for the 3′ junction and using nested PCR. Reactions involved different primer pairs for the first (black) and second (grey) rounds of amplification. Numbers next to confirmed on-target bands indicate the estimated L1 poly(A) tail length in that amplicon. **(*C*)** Bidirectional capillary sequencing electropherogram of the somatic L1 insertion 3′ junction amplicon for neuron #15. Note: in (*A*) and (*B*), neurons #29, #55 and #57 PCR amplified at one junction and not the other. We ascribed this result to MDA dropout rather than truncation of $L1RS_{somatic}$ *in vivo*.



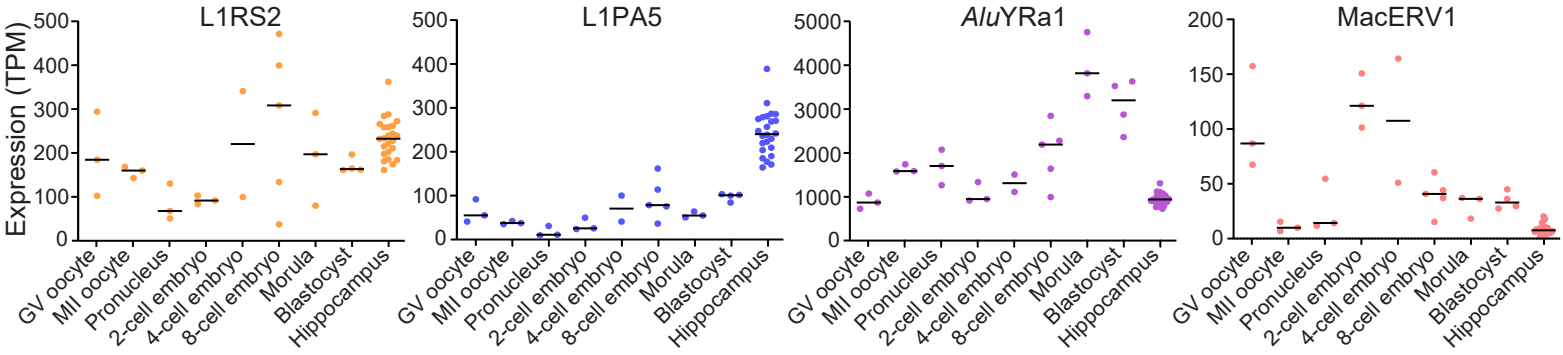
Supplemental Figure S2. Additional L1RS$_{somatic}$ junction PCR data using an extended set of animal ON22213 tissues. PCR amplification results are shown for the somatic L1 insertion 5′ junction (left) and 3′ junction (right), with primer positions indicated by schematics. Reaction inputs for each experiment consisted of non-template control (NTC), as well as DNA from an MDA-amplified hippocampal neuron (#15) where L1RS$_{somatic}$ was first detected, four additional bulk hippocampus samples (labeled 2-5), two representative central nervous system samples (cerebellum and spinal cord), peripheral (sciatic) nerve, and two non-ectoderm samples comprising skeletal muscle (mesoderm) and liver (endoderm). Red arrowheads indicate amplicons confirmed as on-target by capillary sequencing.



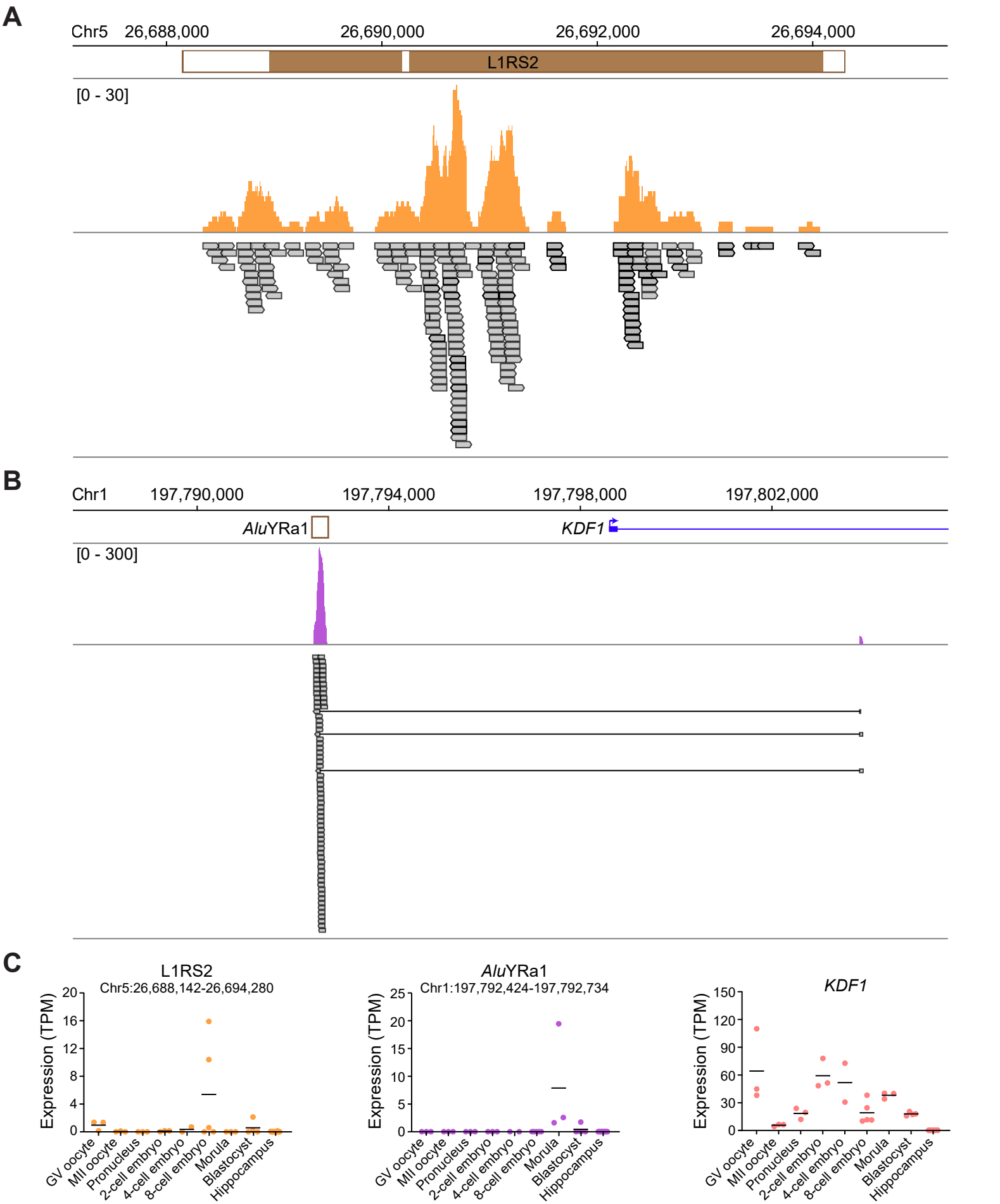
Supplemental Figure S3. Human and macaque mobile L1 sequence conservation. (*A*) Geneious Prime (www.geneious.com) ClustalW alignment of the L1HS consensus (Khan et al. 2006), L1.3 (Dombroski et al. 1993), L1RS2 consensus and ON22213 $L1RS_{PRDM4}$ nucleotide sequences. Grey boxes represent fragments conserved against the L1HS consensus, and vertical black strokes/boxes represent variations to the L1HS consensus sequence. **(*B*)** As for (*A*), except showing amino acid conservation for ORF1 (top) and ORF2 (bottom). ORF1p and ORF2p functional domains (Khazina and Weichenrieder, 2018; Moran et al. 1996; Taylor et al. 2013) are shown underneath.



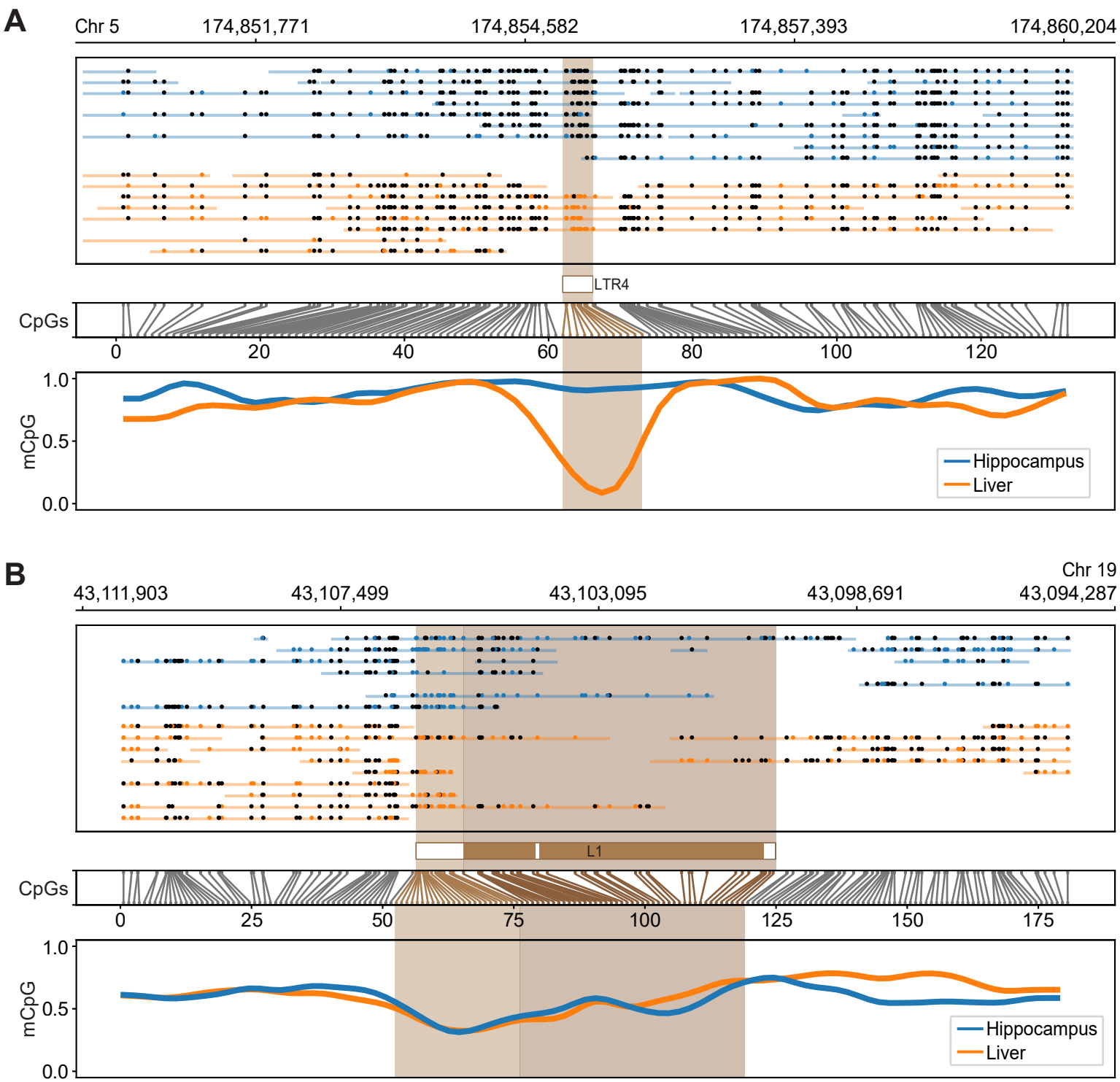
Supplemental Figure S4. Human and mouse pluripotent cell ERV transcription measured via RNA-seq. (*A*) Human endogenous retrovirus-H (HERVH) is a highly expressed TE subfamily in human embryonic stem cells (hESCs) and is down-regulated upon differentiation (Grow et al. 2015; Lu et al. 2014; Zhang et al. 2019). To test the computational approach we applied to macaque RNA-seq data, we analyzed a published human RNA-seq duplicate time course of cardiomyocyte differentiation from hESCs (Zhang et al. 2019). We identified the same pattern of high HERVH expression in hESCs attenuated as differentiation progresses. **(*B*)** Similarly, murine endogenous retrovirus-L (MERVL) is differentially expressed in mouse pre-implantation (2C) embryos when compared to oocytes (Macfarlan et al. 2012; Peaston et al. 2004; Svoboda et al. 2004). A re-analysis of the triplicate RNA-seq data from Macfarlan et al. identified strong and consistent MERVL expression in 2C-embryo samples. Note: horizontal bars in (*A*) and (*B*) represent the mean of duplicate and triplicate values, respectively.



Supplemental Figure S5. Young TE subfamily transcription analyzed with a second computational pipeline. RNA-seq counts were generated and normalized as tags-per-million (TPM) for the same samples shown in Fig. 4 and analyzed using TEtranscripts (Jin et al. 2015).



Supplemental Figure S6. Independently transcribed TE loci. (*A*) Integrative Genomics Viewer (Robinson et al. 2011) coverage (*top*) and alignment (*bottom*) tracks for an intergenic L1RS2. Plots display uniquely aligned RNA-seq reads from an 8-cell embryo replicate (SRR4242911) chosen due to maximum L1RS2 subfamily expression being observed at this time point. This particular element was the most highly expressed full-length L1RS2 copy in the SRR4242911 dataset. **(*B*)** As per (*A*), except showing an intergenic *Alu*YRa1 adjacent to *KDF1* and displaying data from a morula replicate (SRR4242914). **(*C*)** Expression profiles of the L1RS2 and *Alu*YRa1 insertions from (*A*) and (*B*), respectively, as well as that of the *KDF1* gene adjacent to the *Alu*YRa1, measured in RNA-seq tags-per-million (TPM). Data were obtained from prior analyses of germinal vesicle (GV) and metaphase II (MII) oocytes, pre-implantation embryo development stages (Wang et al. 2017) and adult hippocampus (Yin et al. 2020). Horizontal bars represent the mean of biological replicates.



Supplemental Figure S7. Hypomethylated macaque TE loci. (*A*) A MacERV1 solo LTR4 located on chromosome 5 and sharply hypomethylated in animal ON22213 liver. The first panel displays aligned ONT reads, with unmethylated CpGs colored in blue (hippocampus) and orange (liver), and methylated CpGs colored black. The second panel indicates the relationship between CpG positions in genome space and CpG space, including those corresponding to the LTR4 (shaded brown). The third panel indicates the fraction of methylated CpGs for each tissue across CpG space. **(*B*)** As for (*A*), except showing an L1RS2 located on chromosome 19 and less than 50% methylated in both hippocampus and liver.