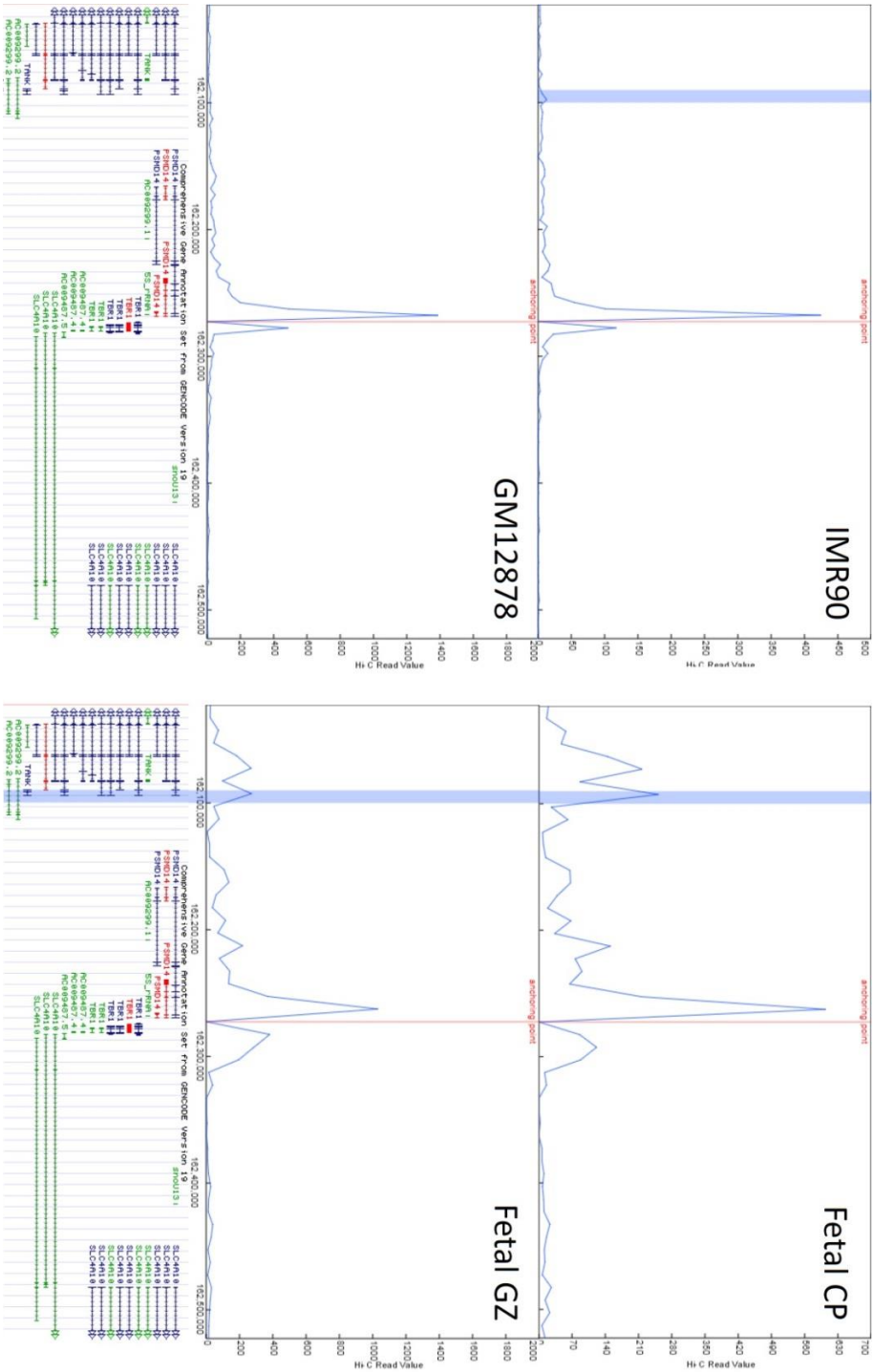
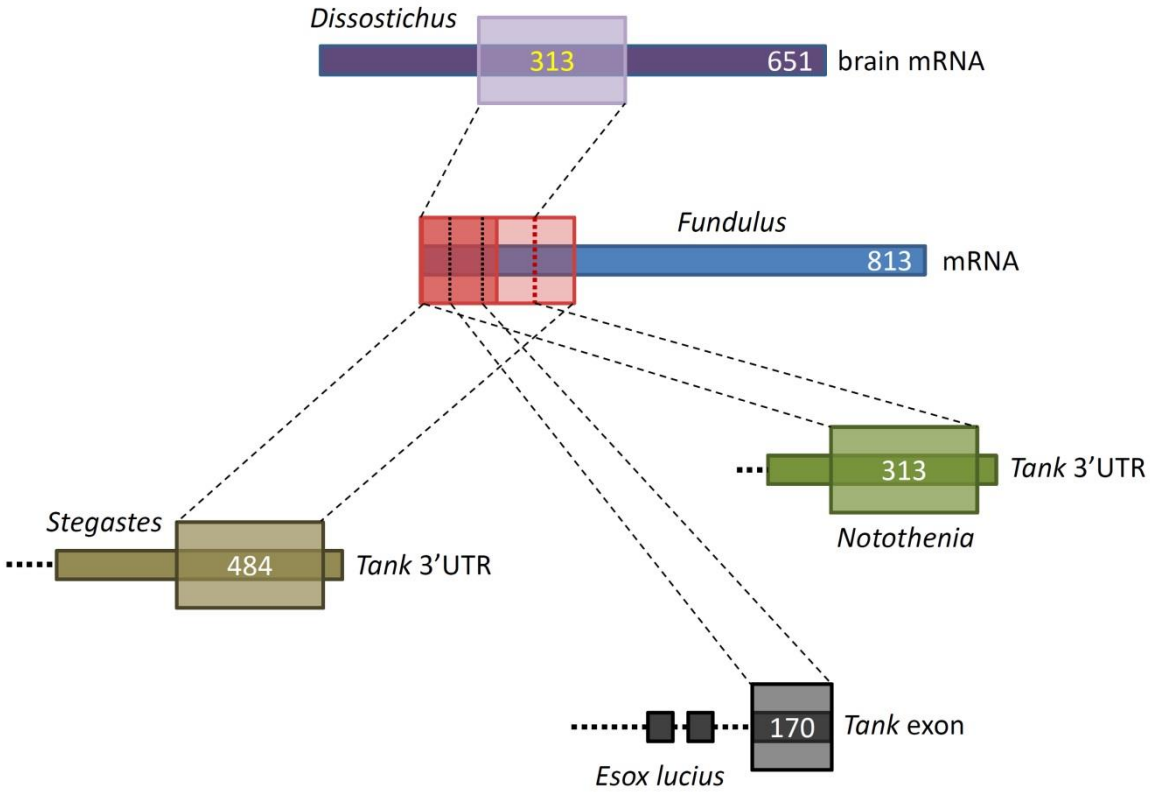
ADDITIONAL INFORMATION

Supplementary Figure S1. Virtual 4C plots using Hi-C data in non-brain cell liness (GM12878 and IMR90), fetal brain cortical plate (Fetal CP) and germinal zone (Fetal GZ). In each plot, the anchoring point (red line) marks the promoter of *TBR1*. There is a peak of contacts with BRE1 (region highlighted in blue at the 3'-end of *TANK* transcripts) only in brain samples. The bottom panel shows the GENCODE gene annotation comprehensive set, version 19. Contacts (Y-axis) are shown as RPKMs.

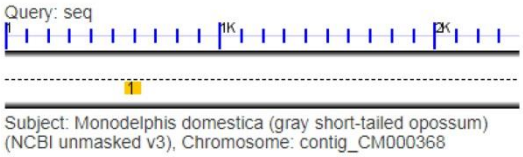


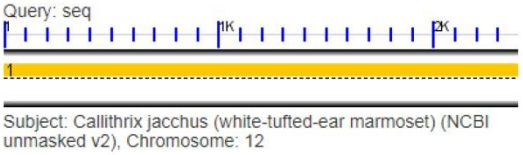
Supplementary Figure S2. COGE blast results using the human sequence of BRE1 as query. For each hit, a rule at the top represents the size of the query, and a yellow rectangle indicates the size and position of the hit with respect to the query. The oldest hit is found in elephant shark and it expands in later vertebrates (top to bottom, left column first). From chicken to primates, the hit covers the entire query.

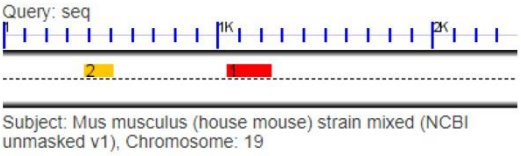
Supplementary Figure S3. Schematic representation of blast hits for the BRE1 sequence in expressed sequences from EST and RefSeq RNA databases. The whole query is present in an mRNA from *Fundulus grandis* (red rectangle); parts of the query also find significant hits in expressed sequences from other Acanthopterygii, frequently as part of the 3'-end of *Tank* mRNAs (see text for details).

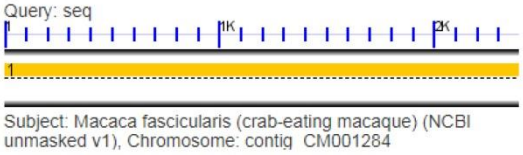


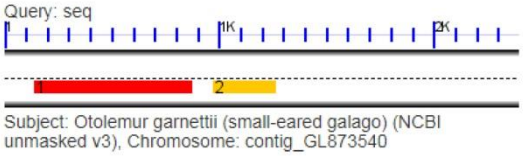
Supplementary Figure S4. COGE blast results using the human sequence of BRE2 as query (top to bottom, left column first). For each hit, a rule at the top represents the size of the query, and a yellow rectangle indicates the size and position of the hit with respect to the query. The oldest significant hit is found in opossum (top left) and a second hit appears in mouse; both hits are progressively extended in primates.

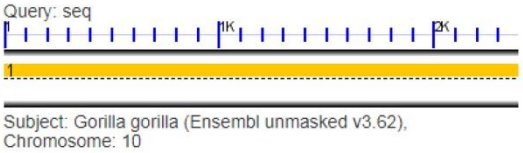


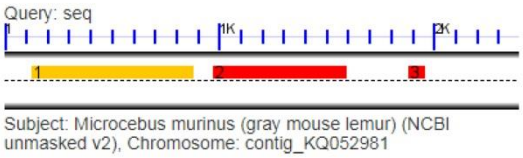


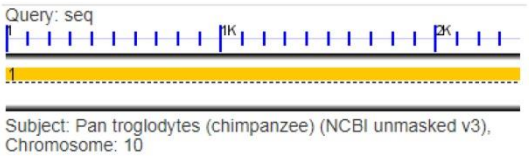




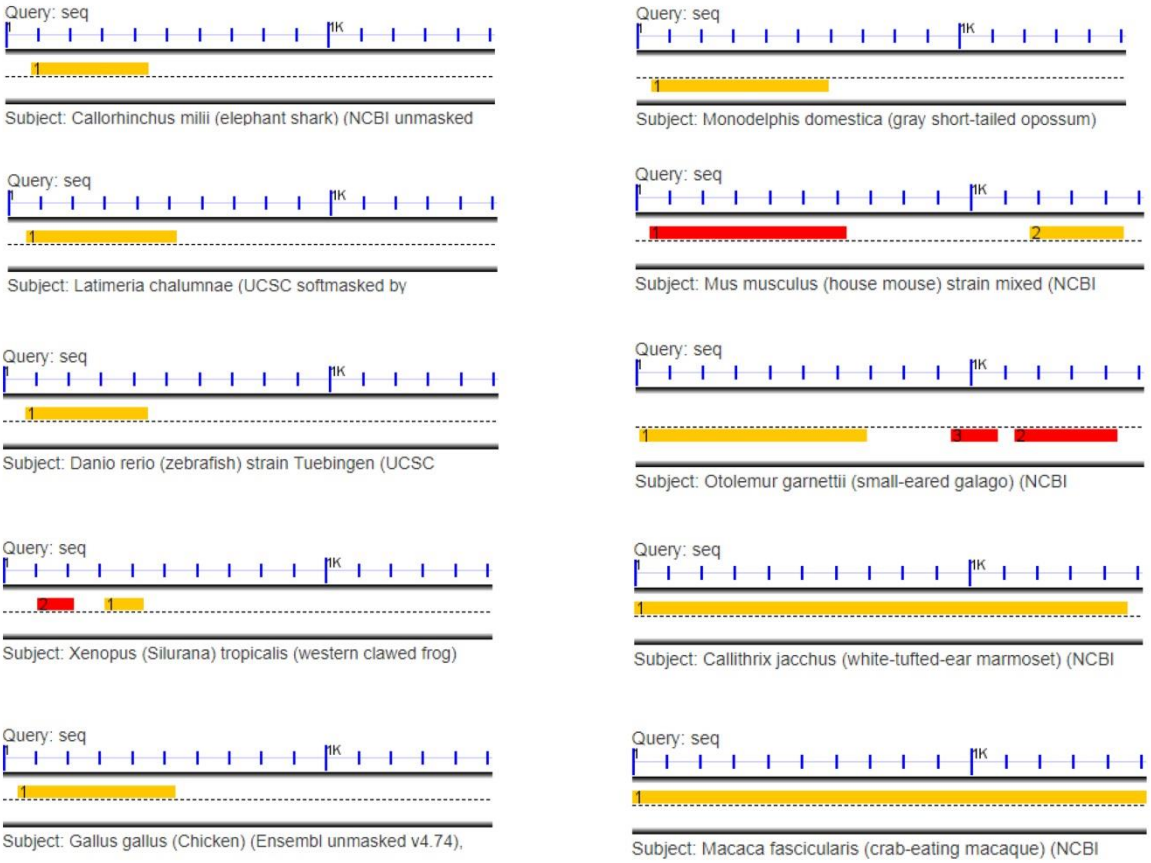
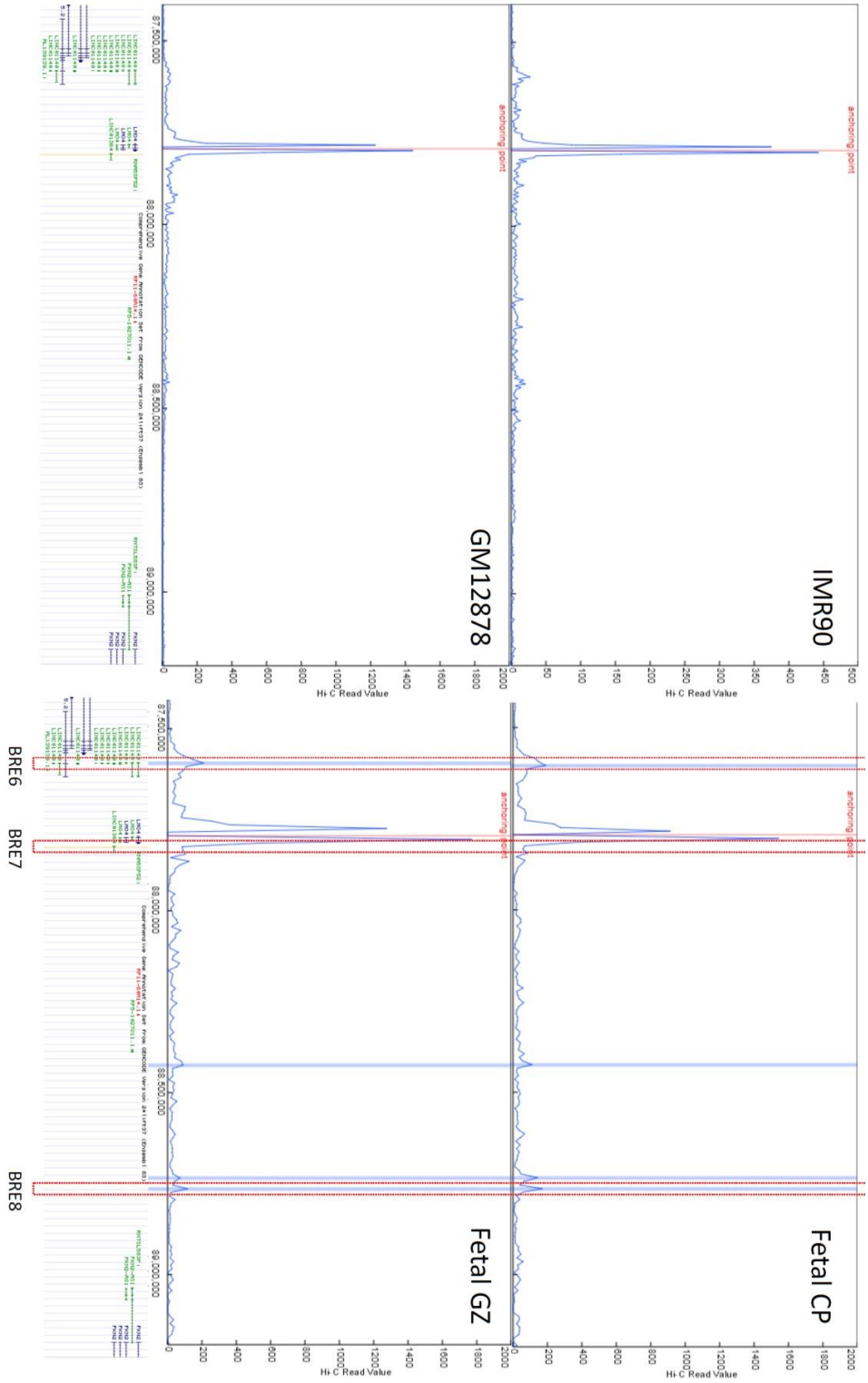








Supplementary Figure S5. COGE blast results using the human sequence of BRE3 as query (top to bottom, left column first). For each hit, a rule at the top represents the size of the query, and a yellow rectangle indicates the size and position of the hit with respect to the query. The oldest hit is found in elephant shark and it is conserved with the same size until chicken (with the exception of xenopus, where it is broken into two fragments). A second hit appears in mouse and both hits are progressively extended in primates.



Supplementary Figure S6. Virtual 4C plots using Hi-C data from non-brain cell lines (GM12878 and IMR90), fetal brain cortical plate (Fetal CP) and germinal zone (Fetal GZ). In each plot, the anchoring point (red line) marks the promoter of *LMO4*. Several peaks of contacts indicate the location of BRE6, BRE7 and BRE8 (highlighted by red rectangles), active in brain samples but not in GM12878 or IMR90 cells (note the different scale of the Y-axis in IMR90). The bottom panel shows the comprehensive gene annotation from GENCODE version 24lift37. Contacts (Y-axis) are shown as RPKMs.