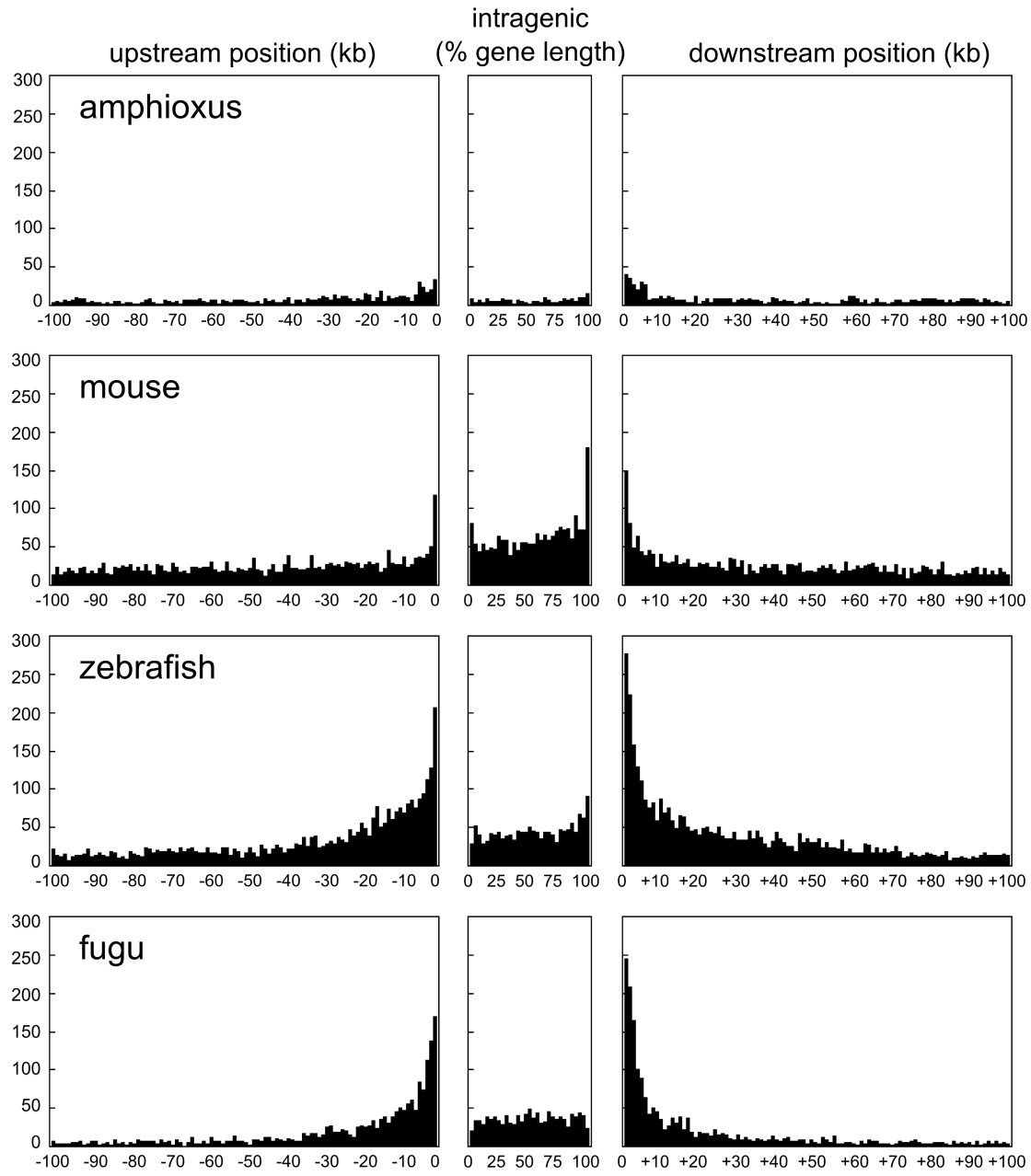
Deeply conserved chordate non-coding sequences preserve genome synteny but do not drive gene duplicate retention

Supplementary Figures 1-3

Andrew L. Hufton, Susanne Mathia, Helene Braun, Udo Georgi, Hans Lehrach, Martin Vingron, Albert J. Poustka, and Georgia Panopoulou

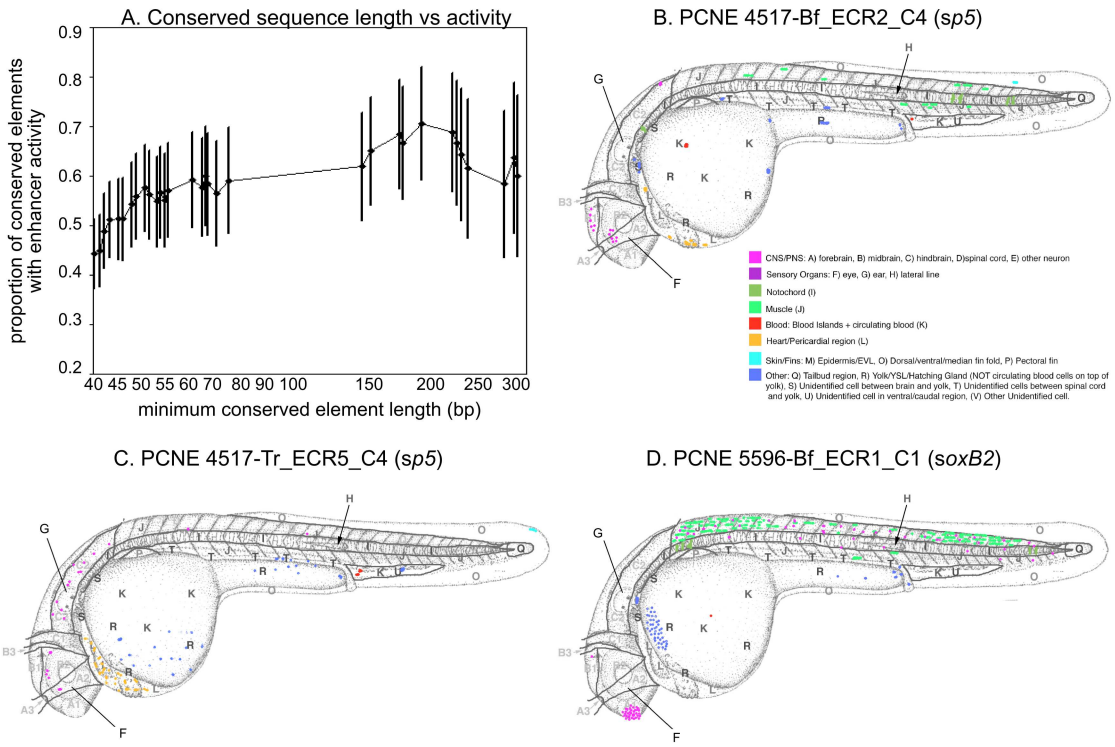
Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany

Supplementary Figure 1: Species specific PCNE locations



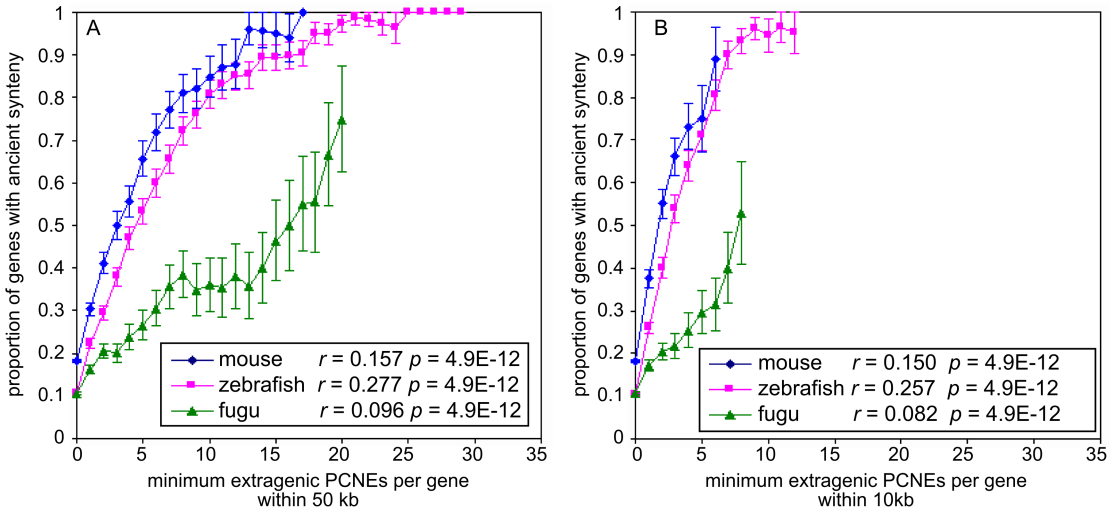
Histograms of PCNE locations for each of the four organisms studied. The right and left-most panels show the distributions of PCNEs upstream and downstream of their parent genes (bin width is 1 kb), while the middle panel shows the distribution of elements found within the introns and UTRs of their parent genes, using relative bin sizes.

Supplementary Figure 2: Many PCNEs are functional enhancers that drive expression in specific tissue patterns



(A) As the minimum length threshold is increased a greater proportion of tested conserved non-coding elements have *in vivo* enhancer activity. Vertical lines show standard error of the proportion. (B-D) show the tissue specific expression patterns driven in zebrafish by three example PCNEs. (B, C) show tissue patterns for two PCNEs linked to amphioxus and fugu *sp5*. Both drive expression primarily in neural and pericardial tissues of the zebrafish embryo, indicating conservation of regulatory function for possibly more than 800 Myr of divergence. (D) shows an the expression pattern for a PCNE associated with amphioxus *soxB2*, an ortholog of vertebrate *sox21*. *Sox21* plays a key role in vertebrate neuronal differentiation (Sandberg et al. 2005), and amphioxus *soxB2* is expressed primarily in the developing neural tube (Meulemans et al. 2003), indicating that these cis-regulatory elements may be part of an ancient genetic program regulating neural development.

Supplementary Figure 3: Conserved synteny association for extragenic PCNEs within 50 kb and 10 kb of their predicted target gene



The association between PCNEs and synteny conservation in three vertebrate organisms for extragenic PCNEs within (A) 50 kb, or (B) 10 kb of their gene. Statistical analysis was performed as described for Fig. 4. Compared to Fig. 4A, the trendline becomes compressed to the right as smaller datasets are used, but the point biserial correlation (r) remains similar.

References

Meulemans, D., McCauley, D., and Bronner-Fraser, M., 2003. Id expression in amphioxus and lamprey highlights the role of gene cooption during neural crest evolution. *Dev Biol*, **264**(2):430–442.

Sandberg, M., Källström, M., and Muhr, J., 2005. Sox21 promotes the progression of vertebrate neurogenesis. *Nat Neurosci*, **8**(8):995–1001.