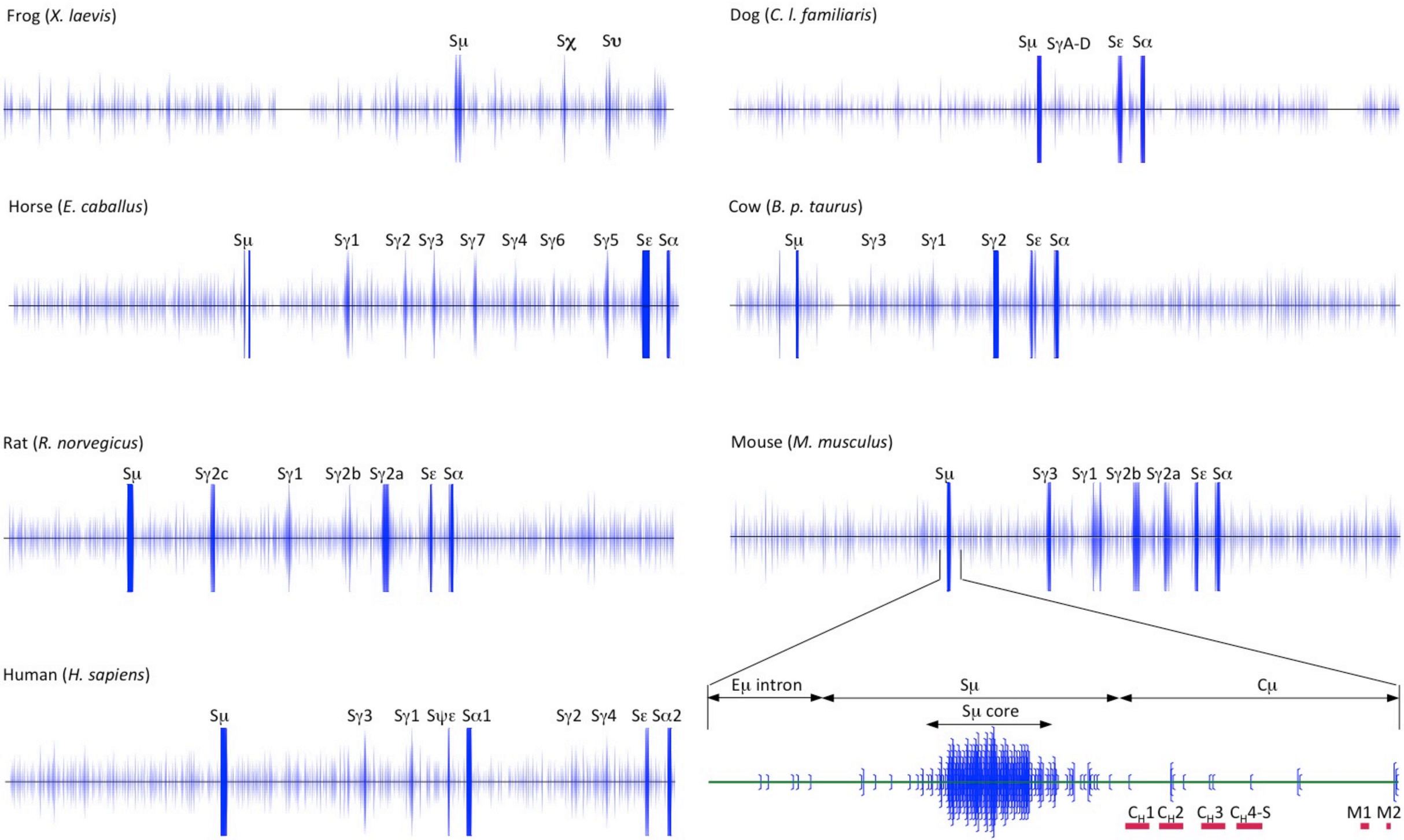

Supplementary Figure 1 | **5'-AGCT-3' repeats in S region DNA.** 5'-AGCT-3' is an evolutionarily conserved motif recurring at a high frequency in S regions of the IgH locus in species that support CSR, from *X. laevis* to human (each vertical line represents a 5'-AGCT-3' motif in 400-kb DNA that includes all D_H and J_H gene segments, S and C_H regions – lines are stacked when multiple 5'-AGCT-3' motifs are closely spaced). 5'-AGCT-3' accounts for 45% of the Sμ core DNA (shown in the magnified map of the mouse *Igh* DNA from the end of the J_H4 gene segment to Cμ region, including the Eμ intron, Sμ and the C_H1, C_H2, C_H3, C_H4-S, M1 and M2 exons of Cμ), while less than 2% of DNA in C_H regions and the genome at large. 5'-AGCT-3' is an iteration of the CSR and SHM hotspot 5'-RGYW-3', within which DSBs, S-S junctions and point-mutations preferentially segregate. 5'-AGCT-3' is also an iteration of 5'-WGCW-3', which includes 5'-WGC-3', the preferred substrate for AID dC DNA deamination activity, in both DNA strands. Due to their palindromic nature, 5'-AGCT-3' repeats can form complex secondary DNA structures, including cruciform structures.