SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1. Poly(A) sites of mouse genes. (A) Histogram of the genomic distance between adjacent poly(A) sites. (B) Histogram of the distance between adjacent poly(A) sites, both located in the 3'-most exon (median = 345 nt). (C) Histogram of the distance between the stop codon and its closest downstream poly(A) site (median = 385 nt). The x-axes in all graphs are in base-2 logarithmic scale. For each histogram, a gaussian smoothing kernel method is used to generate a density line. The genomic position of the first cleavage site of a poly(A) site is used to represent its location in all graphs.

Supplementary Figure 2. Multiple cleavage sites in poly(A) sites of mouse genes. (A) Histogram of the genomic distance between adjacent cleavage sites in mouse genes. (B) Histogram of the distance between the first cleavage site and other downstream cleavage sites when multiple cleavage sites are present in a poly(A) site (mean = 6.5 nt, median = 4 nt). (C) Correlation between the number of cleavage sites and the number of supporting cDNA/EST sequences for poly(A) sites (Pearson correlation coefficient R = 0.75). (D) Histogram of the distance between the first cleavage site and the PAS when only one PAS is present.

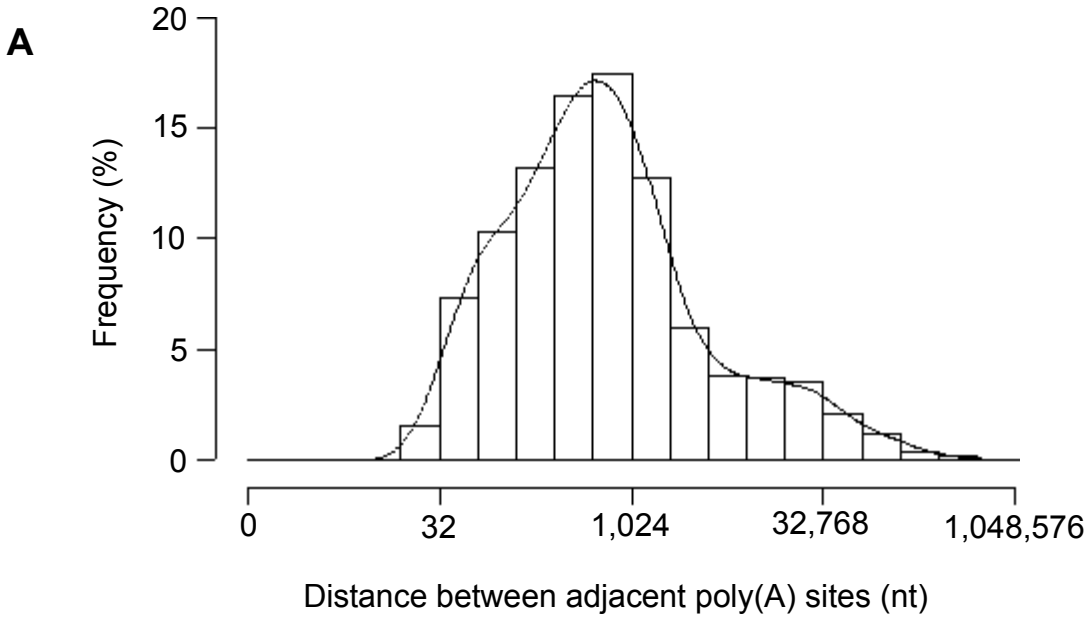
Supplementary Figure 3. Distance between PAS and poly(A) site. Two PAS hexamers were studied. Only human (A) and mouse (B) poly(A) sites with only one associated PAS were used. For each graph, x-axis is the distance (nt) between the PAS and poly(A) site, and y-axis is the number of poly(A) sites.

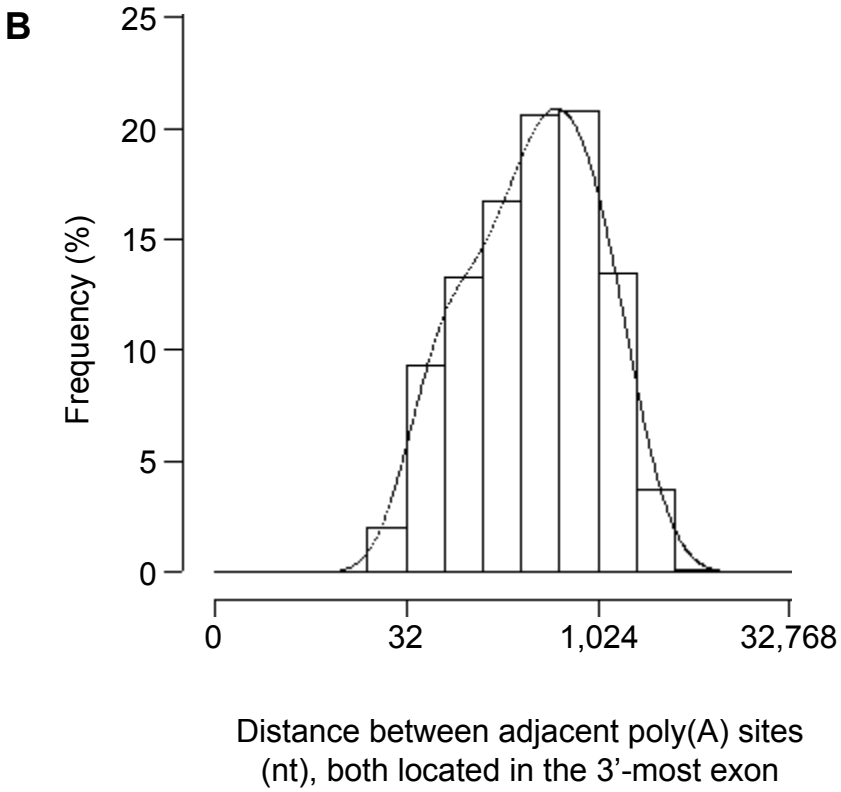
Supplementary Figure 4. Characteristics of different types of poly(A) sites of mouse genes. (A) Usage of various PAS hexamers in different types of poly(A) sites. (B) Cluster analysis of PAS hexamers and poly(A) types. (C) Distribution of the number of cleavage sites per poly(A) site for different types of poly(A) site.

Supplementary Figure 5. Nucleotide composition of terminal sequences containing mouse poly(A) sites. See figure legend of Figure 6 in the main text for details.

Supplementary Figure 6. Nucleotide composition of terminal sequences containing human 1S poly(A) sites. (A) All human terminal sequences containing 1S poly(A) sites; (B) Terminal sequences of genes whose 3' UTR is shorter than 300 nt; (C) Terminal sequence of genes whose 3' UTR is longer than 300 nt.

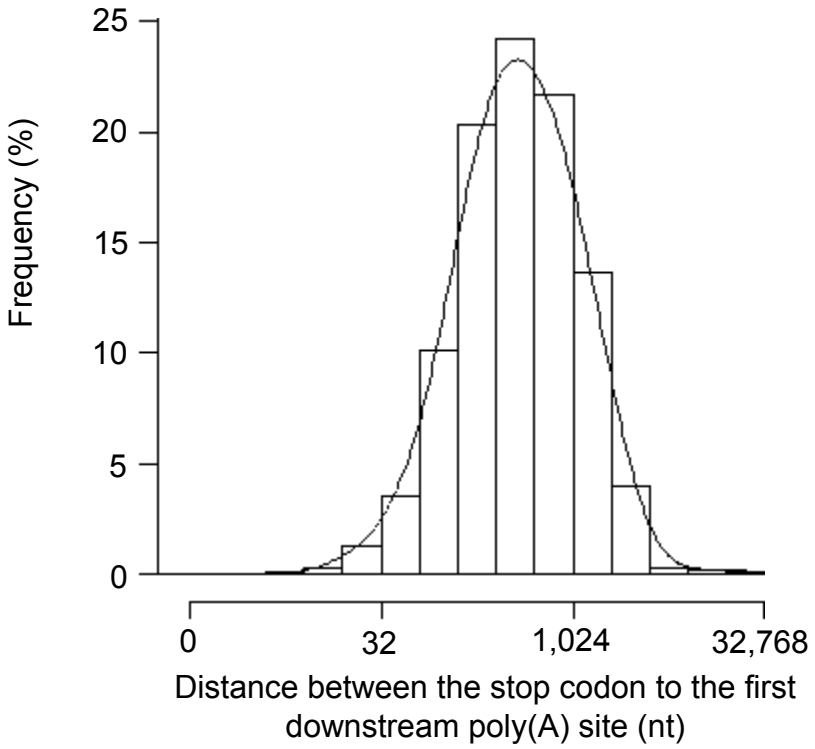
Supplementary Figure 1



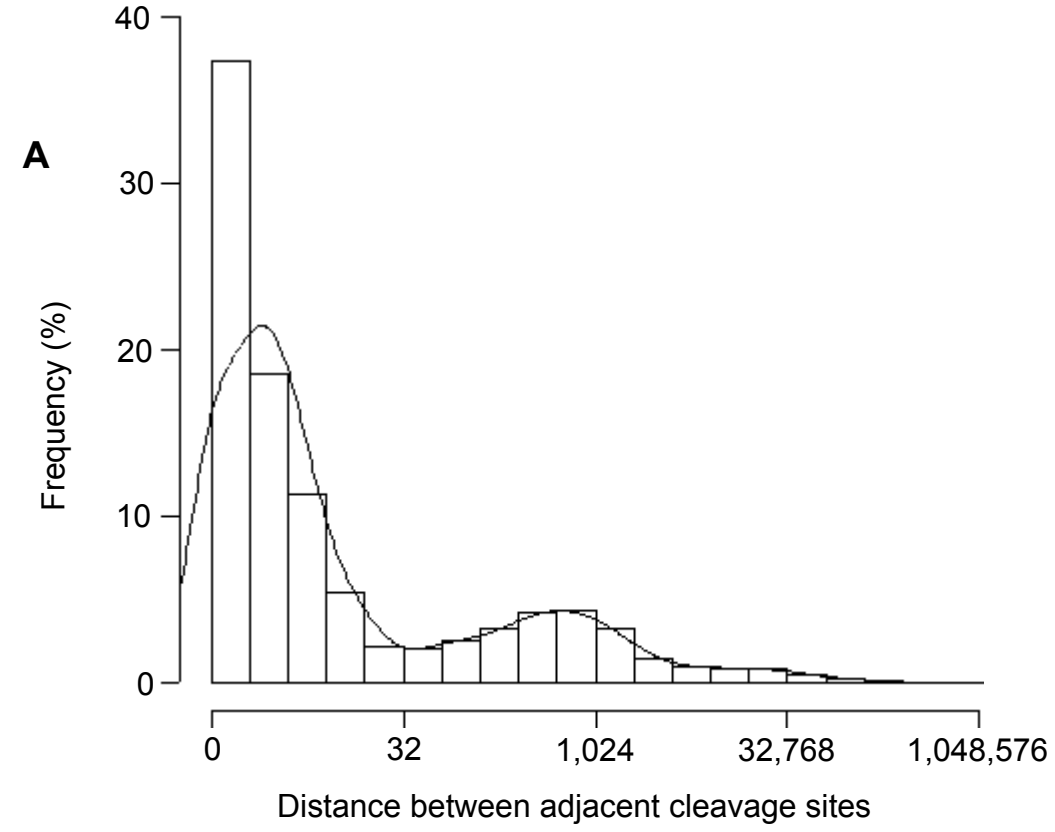


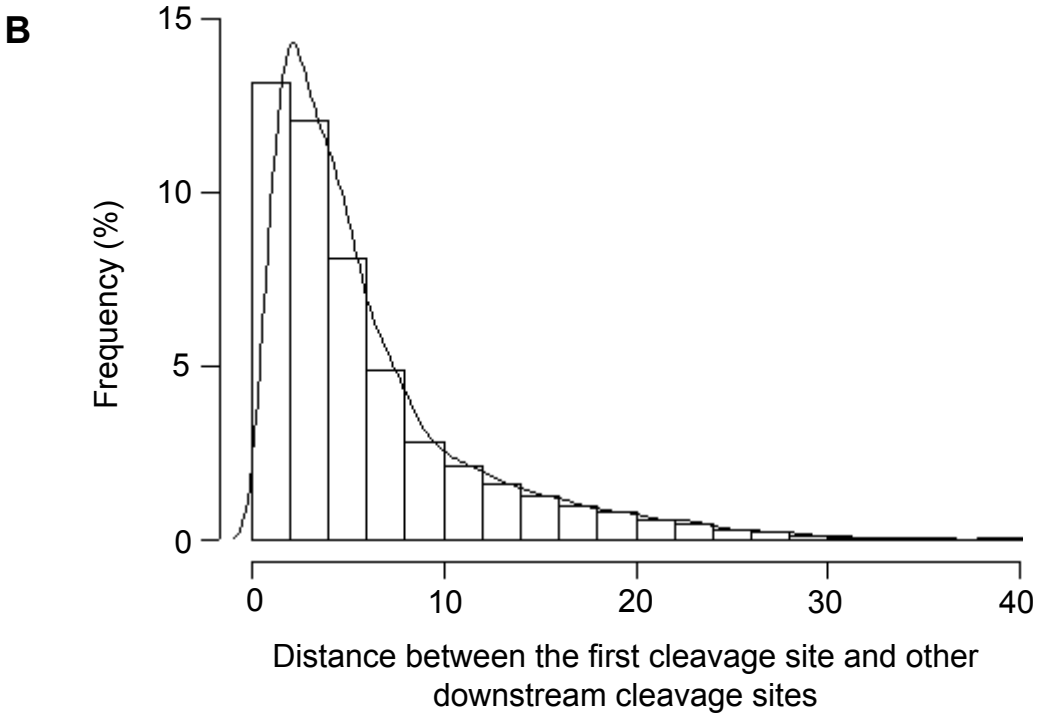
Supplementary Figure 1

C

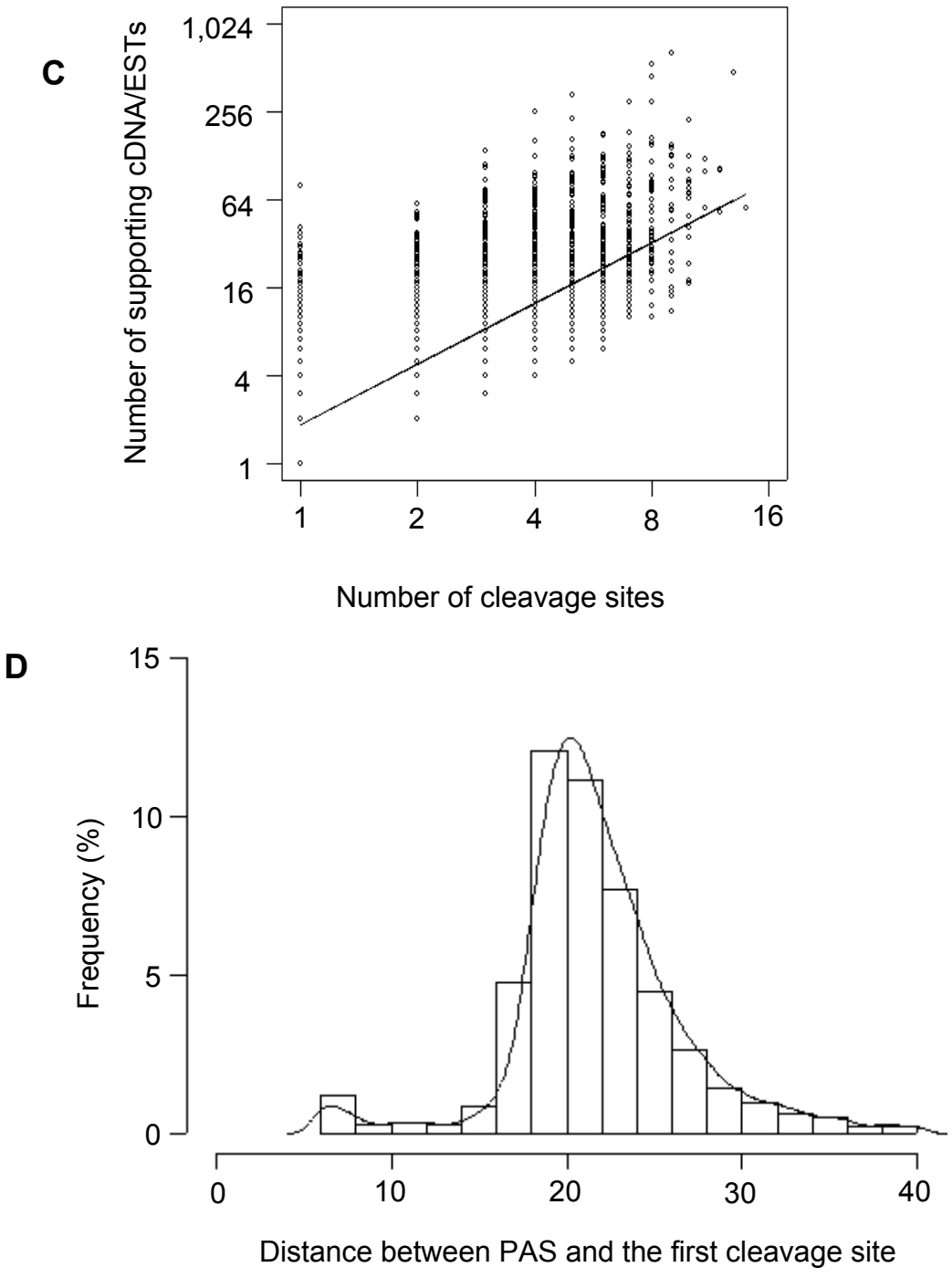


Supplementary Figure 2

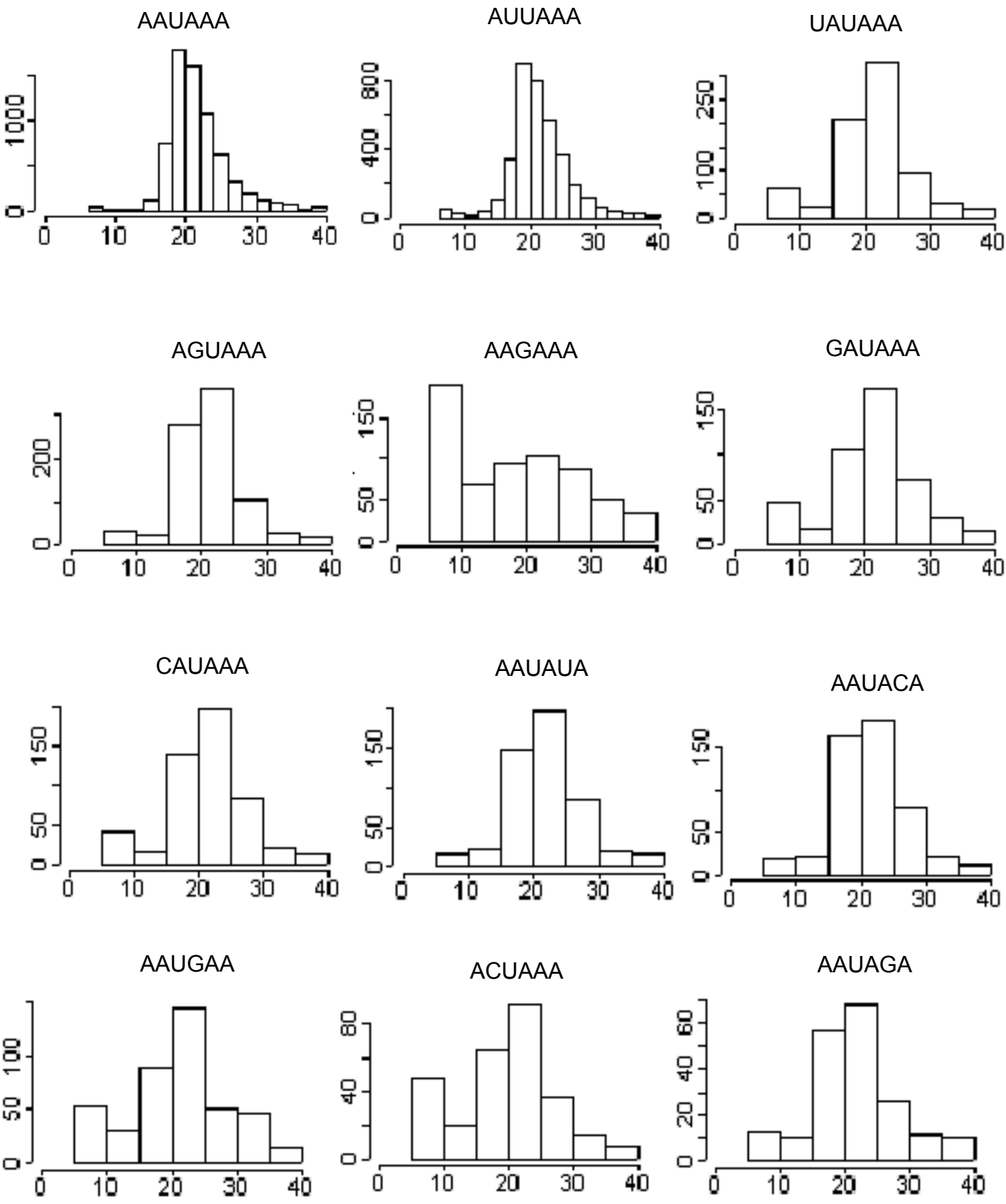




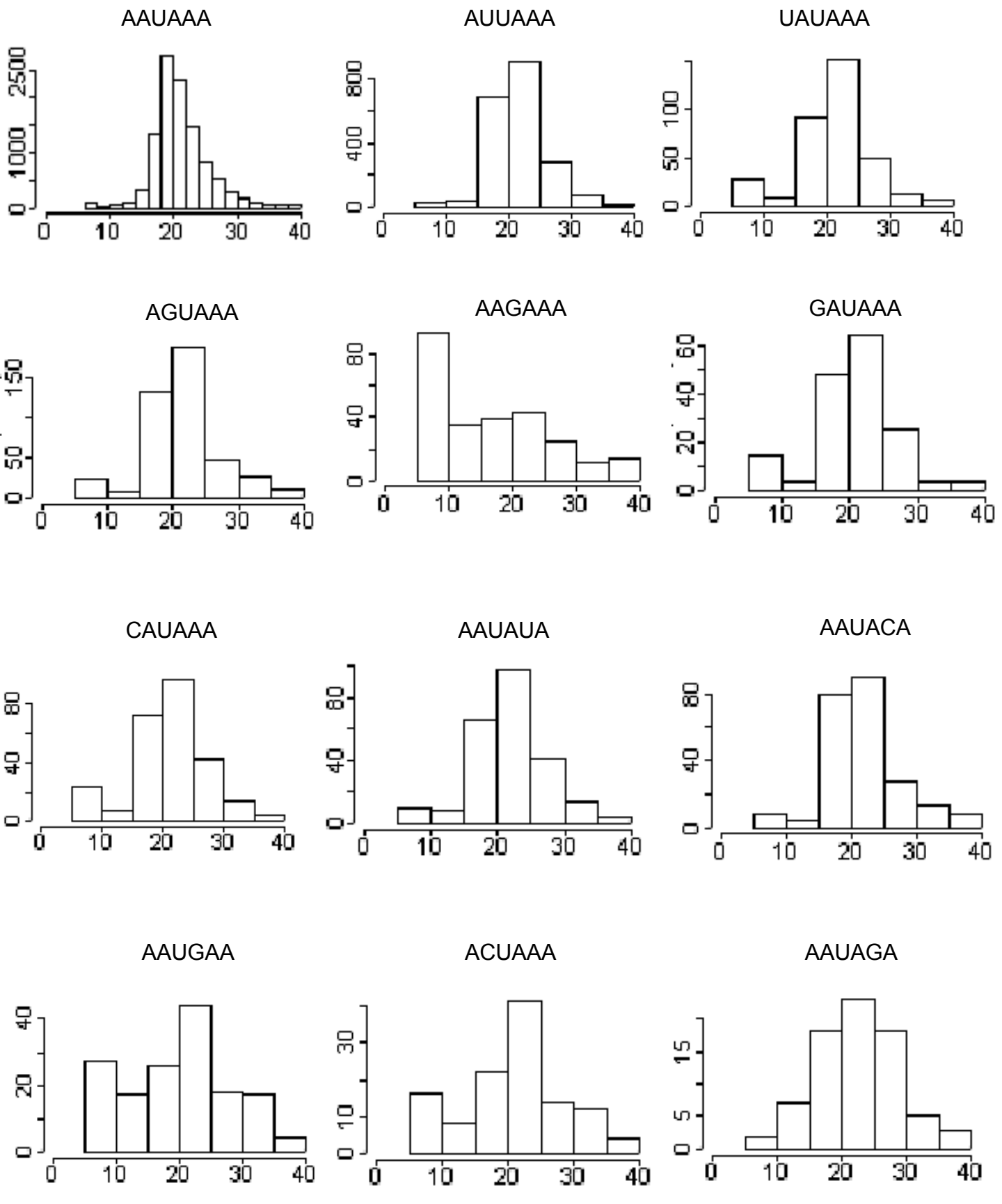
Supplementary Figure 2



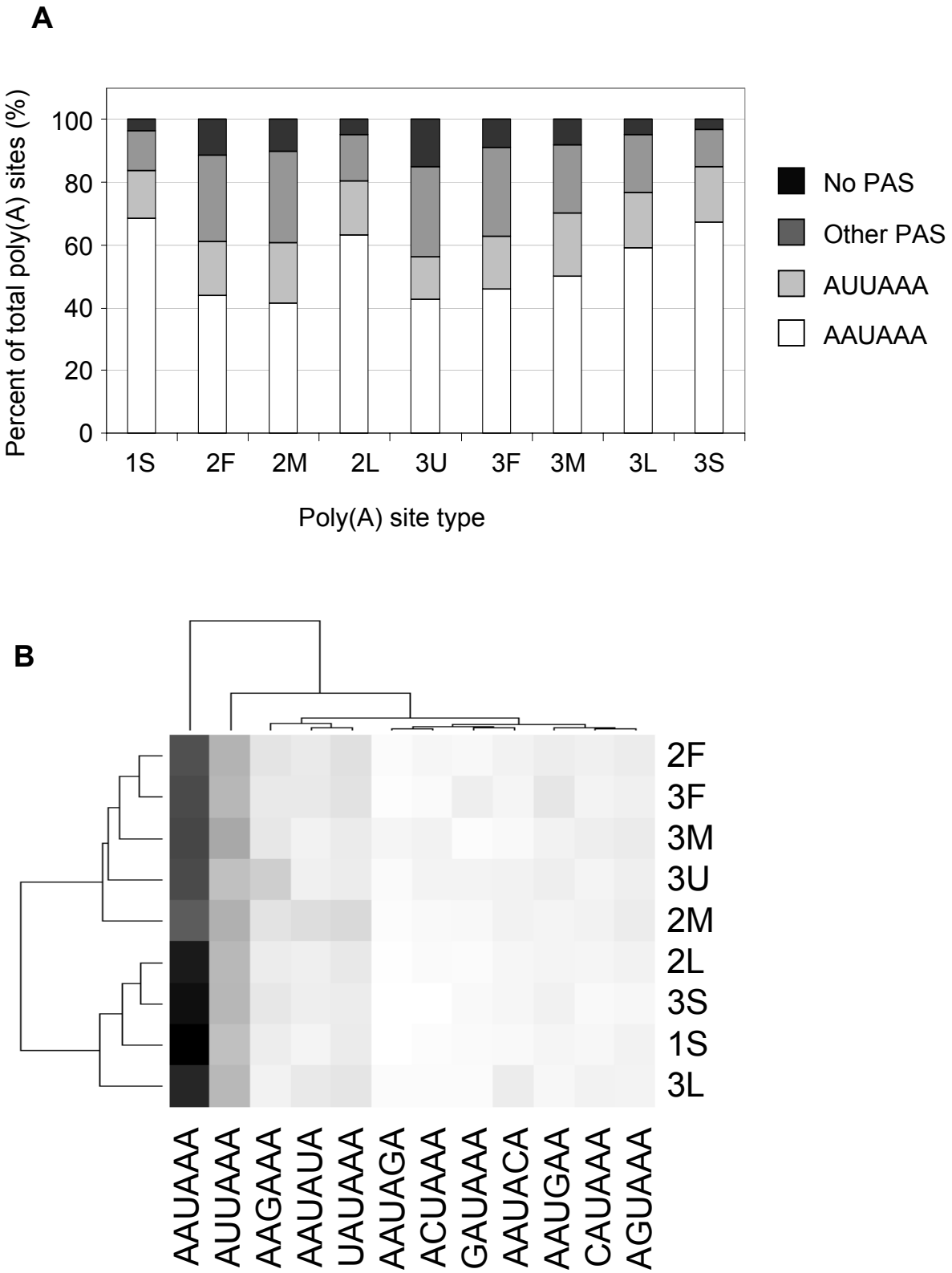
Supplementary Figure 3A



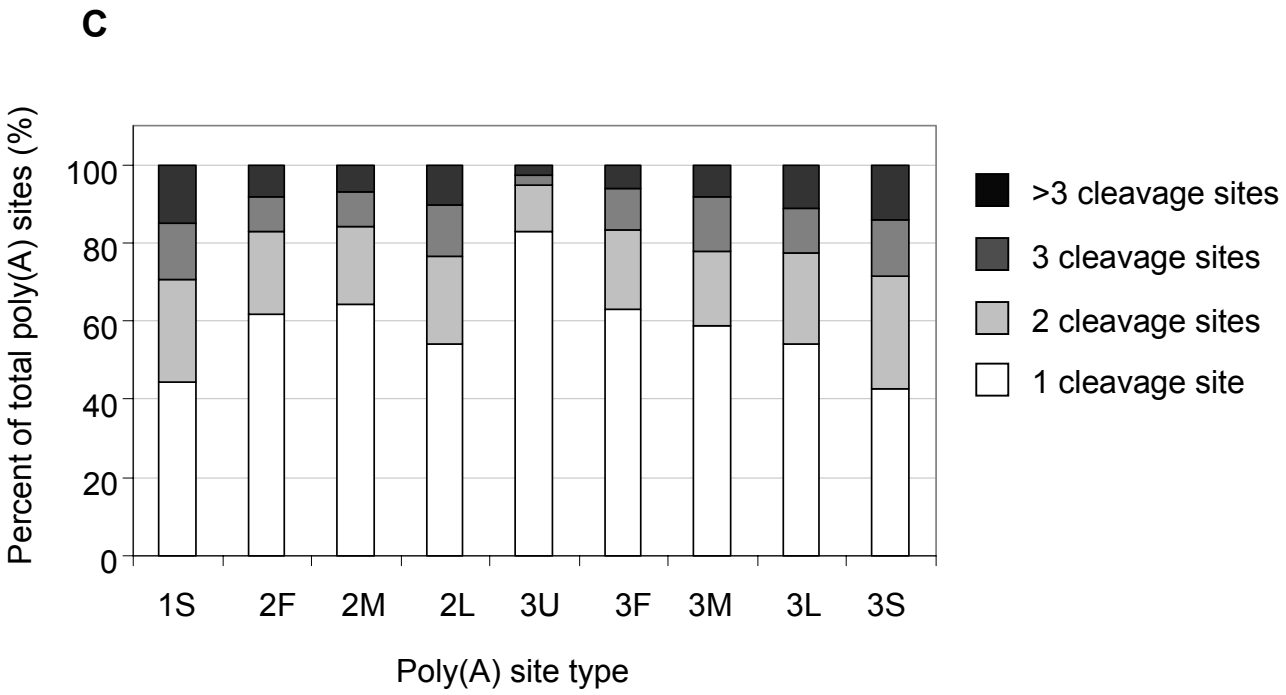
Supplementary Figure 3B



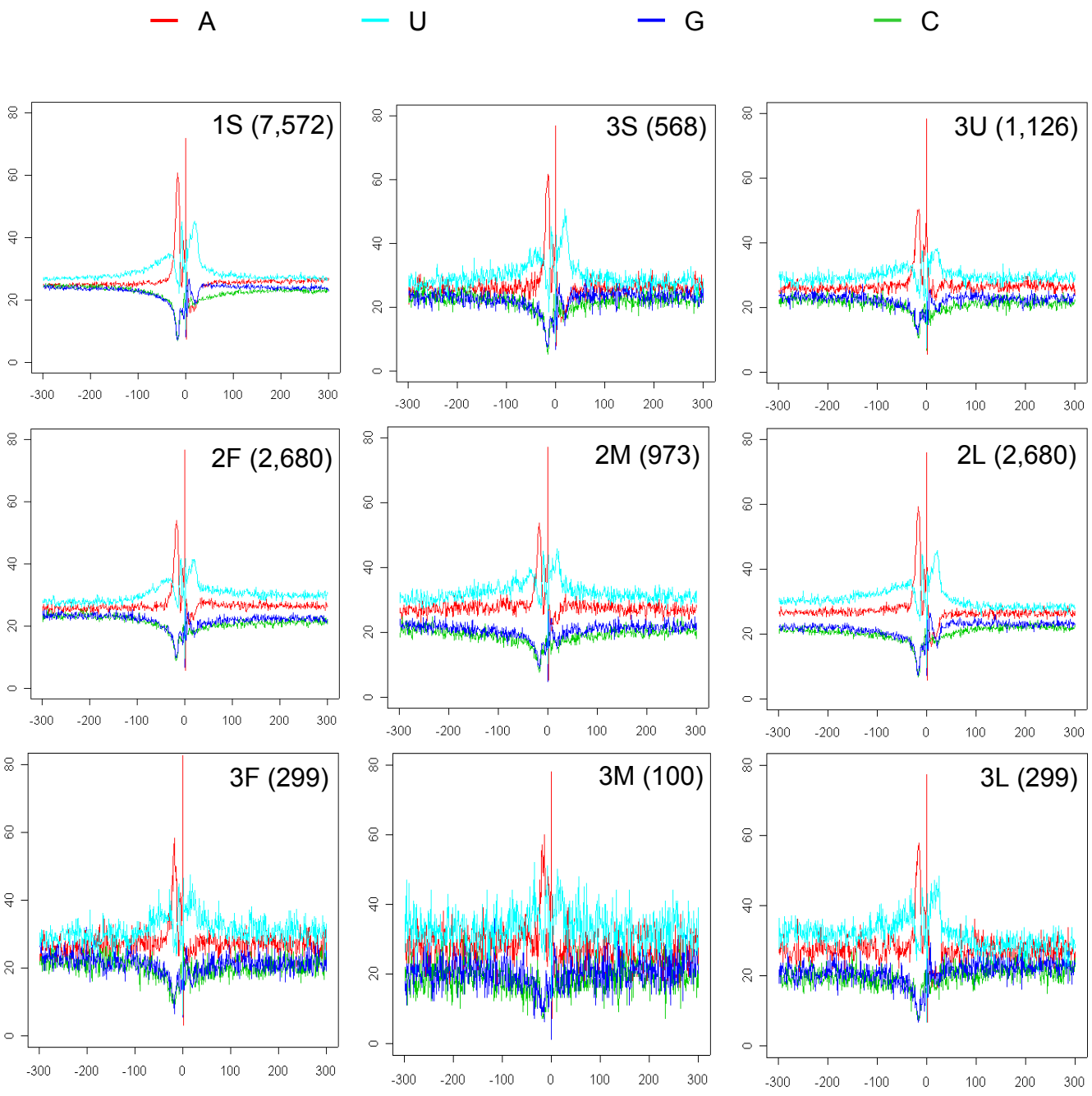
Supplementary Figure 4



Supplementary Figure 4



Supplementary figure 5



Supplementary figure 6