

**Figure Legend (figure overleaf): The top-ten most common low-complexity regions (LCRs) with fLPS P-value  $\leq 1 \times 10^{-6}$ , for the three trees, split into kingdoms.** Within a protein family, LCRs that are not conserved across kingdoms are represented by grey LCRs. {G}-rich LCRs are labelled red, {GR}-rich gold, {P}-rich green and {A}-rich, with other types conserved across kingdoms coloured purple.

The total counts are on the y-axis and the bias signatures from the fLPS algorithm are on the x-axis. The panels are for the following data:

- (A) The metazoan sequences from the EWSR1/TAF15/FUS tree (EWSR1, etc.);
- (B) The non-metazoan sequences from the EWSR1/TAF15/FUS tree (EWSR1, etc.);
- (C) The metazoan sequences from the TDP43 tree;
- (D) The plant sequences from the TDP43 tree;
- (E) The metazoan sequences from the hnRNPA2 tree;
- (F) The fungi sequences from the hnRNPA2 tree;
- (G) The plant sequences from the hnRNPA2 tree;
- (H) The other (not fungi, metazoan or plant) sequences from the hnRNPA2 tree.

