 **Supplementary Information**

**Fig S1.** The normalised RNA-sequencing dataset's principal component analysis (PCA) plot. The PCA was performed to visualise the overall transcriptomic differences between the control retina and retinoblastoma (RB) tumour samples. Each data point represents a single sample, with the 20 control retina samples depicted in blue and the 50 RB tumour samples in red. The first three principal components (PC1, PC2 and PC3) are shown, capturing the most significant proportion of the variance in the dataset. The plot demonstrates a clear separation between the control and RB tumour samples, indicating distinct gene expression profiles between the two groups.



**Fig S2.** Module-trait relationships and gene significance in the weighted gene co-expression network analysis (WGCNA). (A) Heatmap of module-trait associations. Each row corresponds to a module eigengene, and each column corresponds to a trait (tissue expression in control retina, retinoblastoma (RB) tumours, and dysregulated splicing factors). Each cell contains the corresponding Pearson's correlation coefficient and p-value. The cell colour represents the correlation according to the colour legend, with red indicating a positive correlation and blue indicating a negative correlation. (B) Scatter plot of module membership (MM) versus gene significance (GS) for differentially alternatively spliced (DAS) genes in the black module. The black module shows a higher correlation with DAS events in RB tumours than other modules. MM represents the correlation between a gene's expression profile and the module eigengene. GS measures the association between a gene's expression and the RB tumour phenotype. The plot demonstrates a strong positive correlation between MM and GS in the black module.



**Fig S3**. Sub-network of functional enrichment analysis for genes in the black module. Hub genes, selected based on centrality measures, are highlighted in blue. All nodes are red, indicating that all genes in the black module are upregulated. Node shapes represent differential alternative splicing (DAS) status: circles for downregulated DAS genes, hexagons for upregulated DAS genes, and diamonds for non-significant DAS genes. Downward triangles represent enriched pathways.



**Fig S4**. Protein-protein interaction (PPI) network of splicing factors and differentially alternatively spliced (DAS) hub genes. The network visualises the interactions between splicing factors(Triangle) and DAS hub genes(circles) identified in the analysis. Nodes represent proteins encoded by the genes, and edges represent physical and functional interactions between the proteins. The genes highlighted with red borders have the highest degree of centrality (i.e., the highest number of interconnected nodes). They are connected to the splicing factors HNRNPA1 and ILF2, suggesting their potential role in regulating alternative splicing events.