Mapping the GlycoRNA-Protein Interactome: Insights into Glycan-Modified RNA Binding Partners and Functional Networks

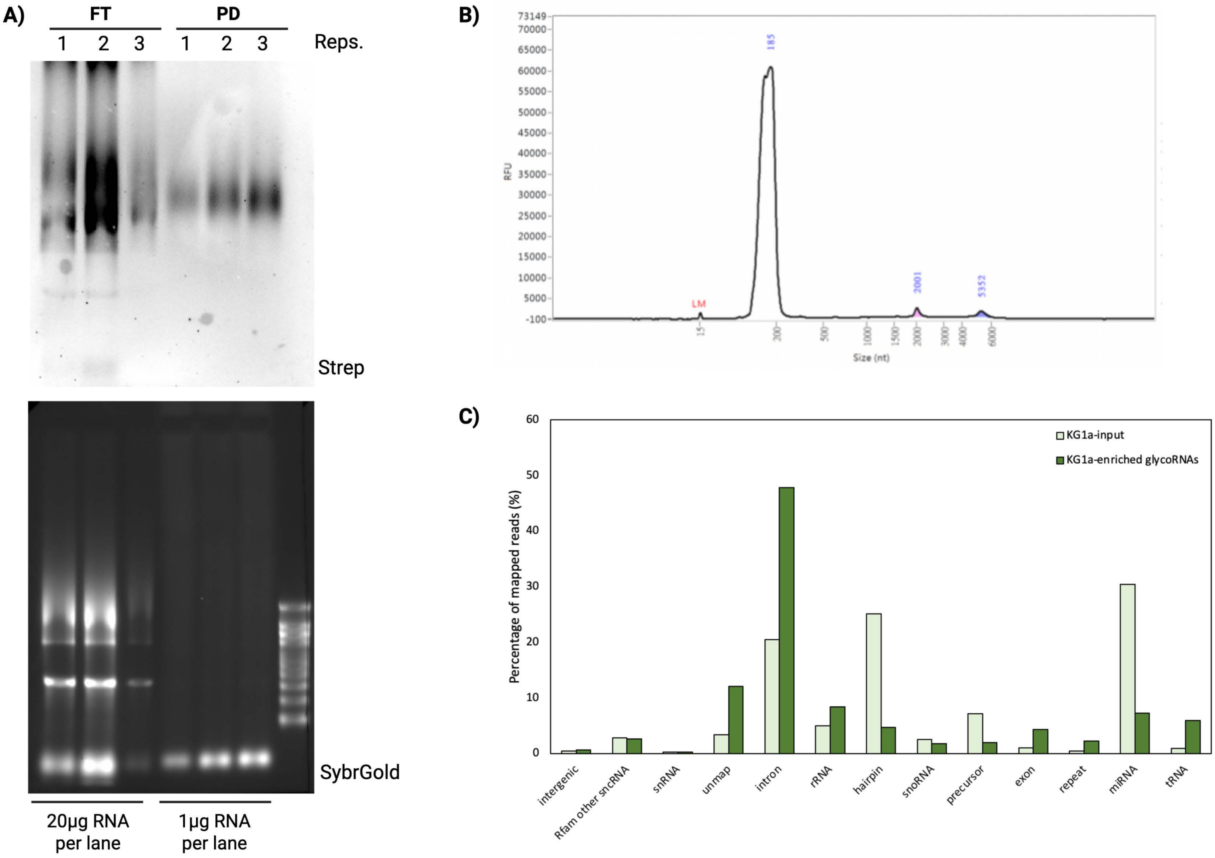
Ioannis Isaioglou[1†], Jing Kai[1†], Jasmeen S. Merzaban[1*]

[1] Bioscience Program, Biological and Environmental Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia

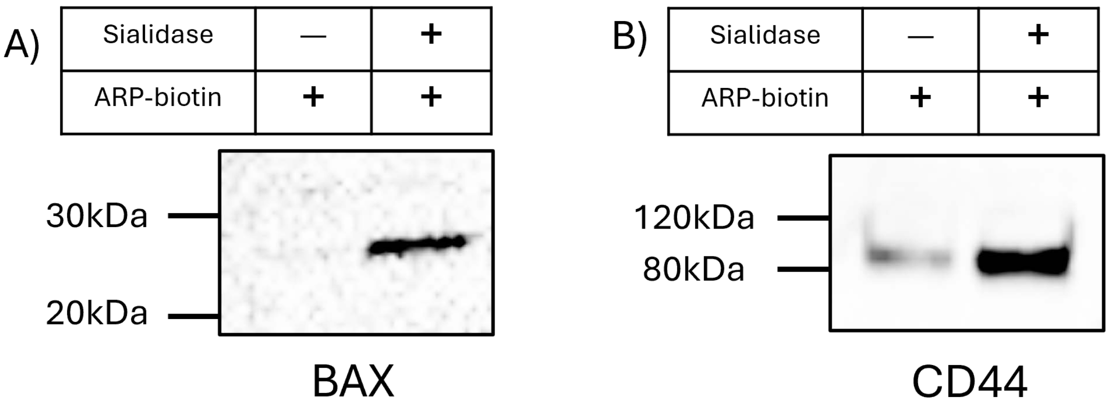
† Authors contributed equally

* Corresponding author, jasmeen.merzaban@kaust.edu.sa

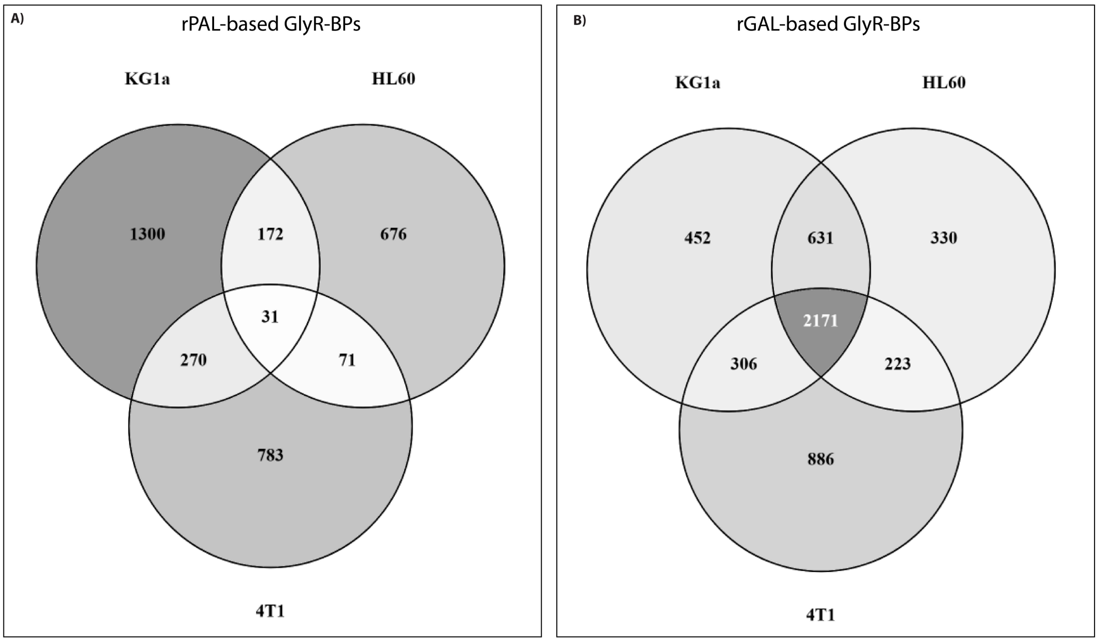
Supplementary Figures



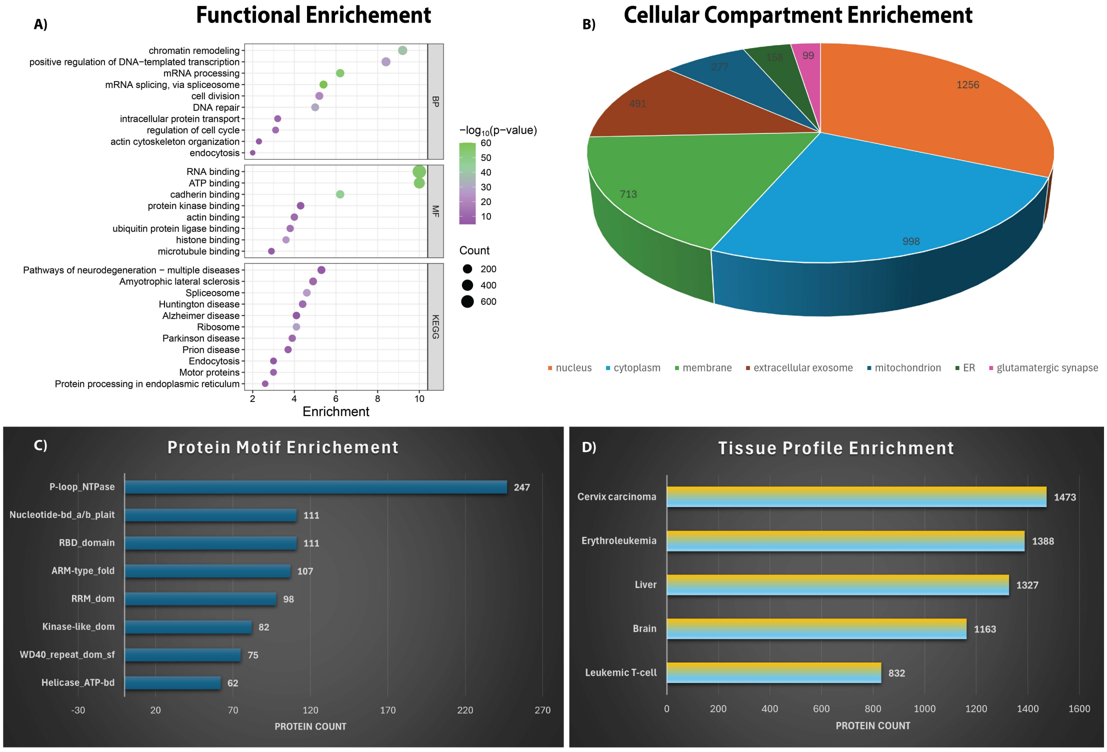
Supplementary Figure S1. rPAL-labeled (sialo-)glycoRNAs in KG1a cells are primarily small non-coding RNAs. (A) rPAL-labelled glycoRNAs from KG1a cells were enriched by streptavidin-beads and both flow-through (FT) and pull-down (PD) samples were analyzed by RNA blotting. Pull-down samples showed strong enrichment in small RNA regions (<200nt). Reps, replicates; Strep, streptavidin-IR800. **(B)** Representative Bioanalyzer trace showing size distribution of purified pull-down glycoRNAs, with prominent peak at ~185 nt. **(C)** Classification of small RNA sequencing reads from rPAL-enriched glycoRNAs compared to input control (<200nt fractionated total RNA). Data represent the average of four biological replicates (n=4).



Supplementary Figure 2. Western blot validation of proteomics data. Pull-downs of KG1a cell lysate using glycoRNAs, with and without sialidase treatment, were analyzed by Western blot. **(A)** BAX was detected only in the untreated sample and was lost upon sialidase treatment, confirming sialic acid-dependent interaction. **(B)** CD44 was present in both untreated and treated pull-downs but showed higher abundance in untreated samples, consistent with proteomics data (protein score 464, 15 peptides vs. 62, 1 peptide in treated sample). Proteins identified with a single peptide in the sialidase-treated sample were excluded from enrichment analyses presented in **Figures 3-6**.



Supplementary Figure 3. Venn diagram analysis of glycoRNA-binding proteins. Comparison of rPAL-based **(A)** and rGAL-based **(B)** glycoRNA-binding protein datasets across species. All Uniprot IDs were converted to gene names to allow cross-species comparison.



Supplementary Figure 4. Functional enrichment of common rGAL-based glycoRNA-binding proteins across three cell lines. (A) Gene Ontology (GO) enrichment of Biological Process (BP), Molecular Function (MF), and KEGG categories for shared rGAL-based glyR-BPs. Top enriched terms are shown. **(B)** GO Cellular Component enrichment showing predominant localization patterns. **(C)** InterPro protein motif enrichment for rGAL-based glyR-BPs. **(D)** Tissue enrichment analysis (UP_Tissue database) showing the most significantly enriched tissues. Top categories are shown here.