



# Ago-associated human usRNAs and other similar small RNAs

Zihua Li<sup>1,2</sup>, Sang Woo Kim<sup>1</sup>, Yuefeng Lin<sup>1</sup>, Patrick S. Moore<sup>2</sup>, Yuan Chang<sup>2</sup>, and Bino John<sup>1,2,\*</sup>

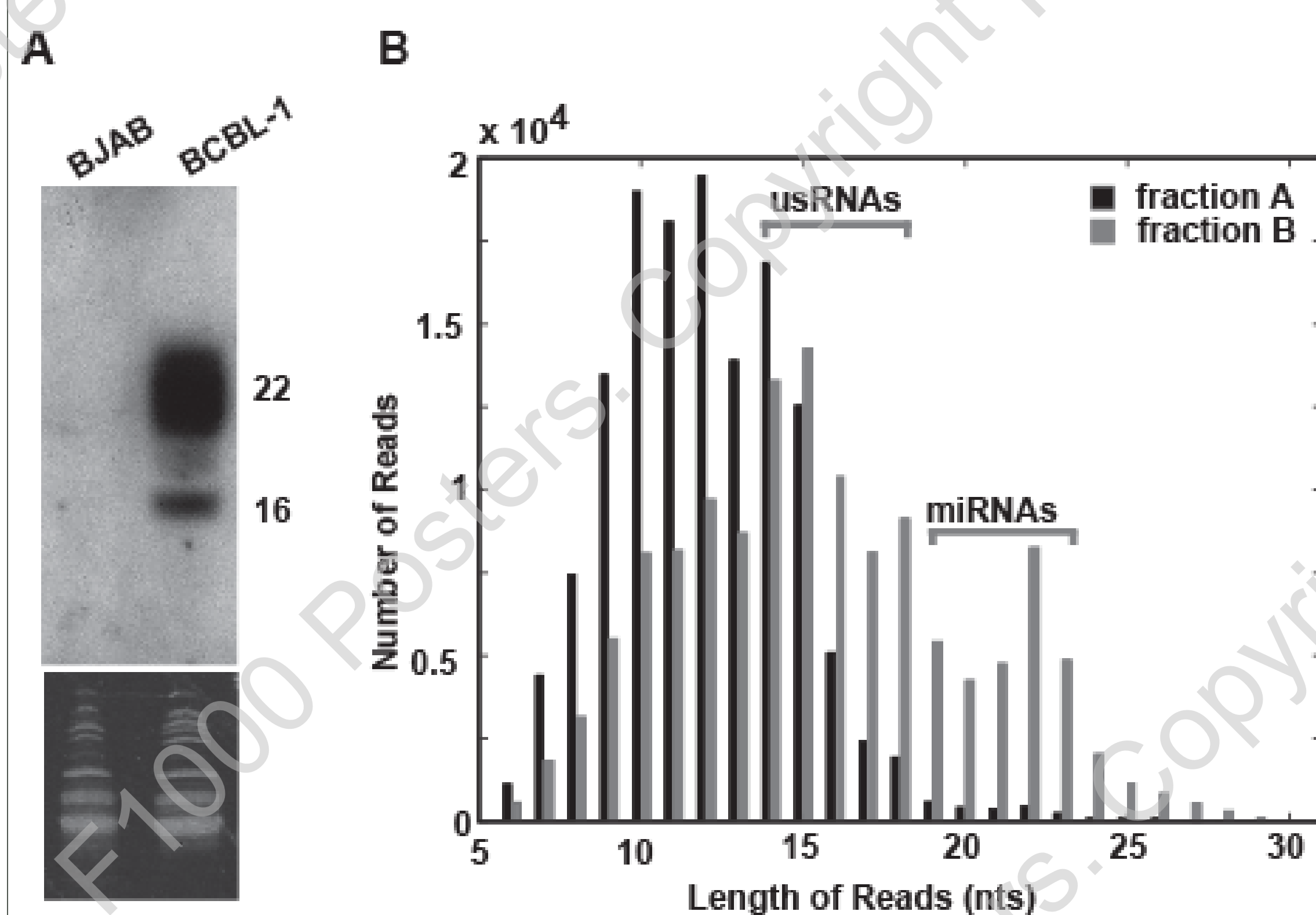
<sup>1</sup>Department of Computational Biology, University of Pittsburgh School of Medicine, & <sup>2</sup>Molecular Virology Program, Hillman Cancer Center, Pittsburgh, USA

bin/joel  
if (Svar =  
s/S+S//;  
#M  
# propose

## Abstract

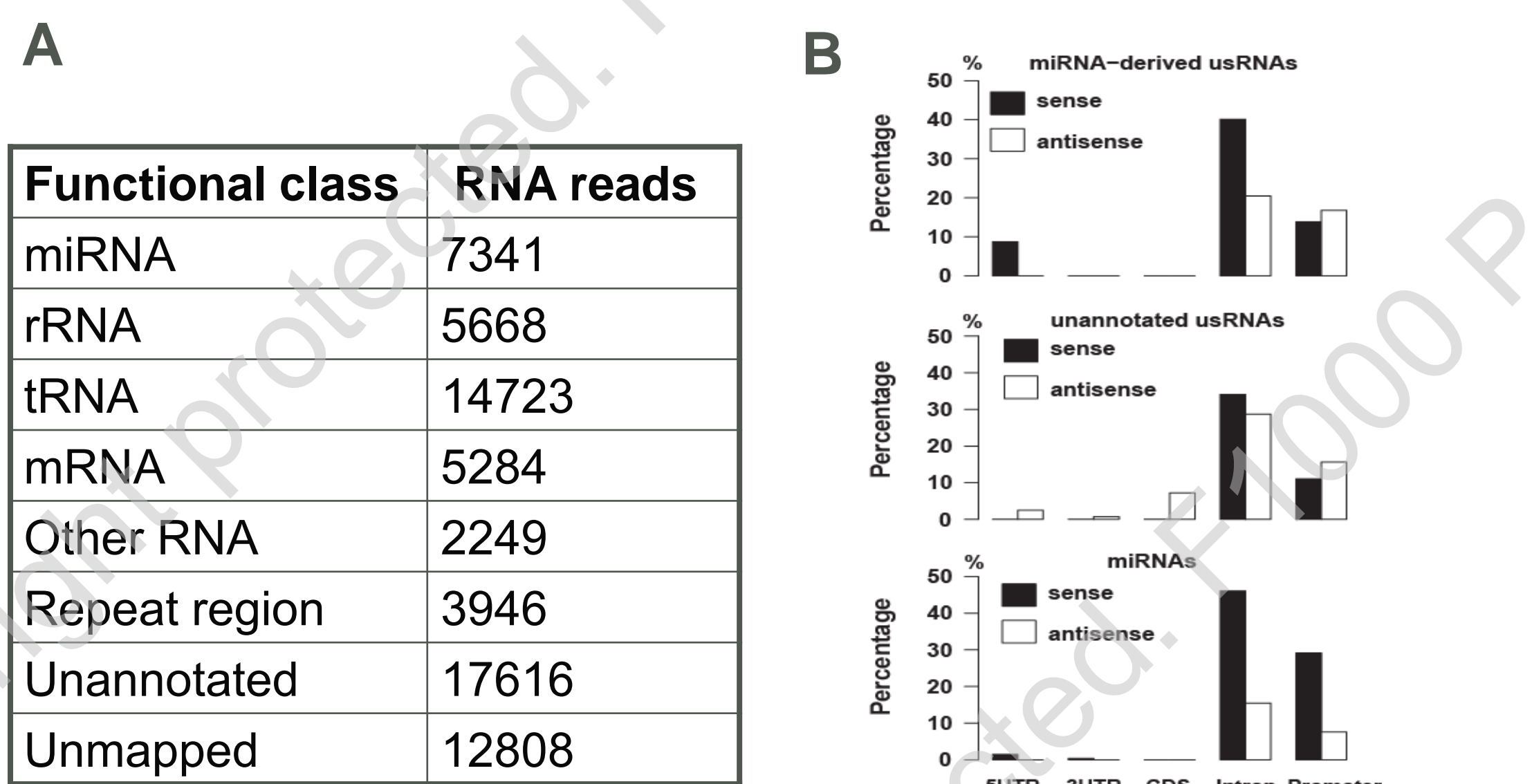
We report on the characterization of several sub classes of unusually small (~15 nts) RNAs (usRNAs). Our observations began with the identification of a Kaposi sarcoma-associated herpesvirus (KSHV) us-K12-1 RNA (17 bases) that is as effective as the K12-1 miRNA in regulating human RAD21. High-throughput sequencing reveals a diverse set of human miRNA-derived usRNAs and other non-miRNA-derived usRNAs. Human miRNA-derived usRNAs preferentially match to 5' ends of miRNAs, and are also more likely to associate with the siRNA effector protein Ago2 than Ago1. Many non-miRNA-derived usRNAs associate with Ago proteins and also frequently contain C-rich 3'-specific motifs that are highly position specific, and are overrepresented in comparison to piRNAs and TSSa-RNAs. We further demonstrate that some non-miRNA-derived usRNAs have development-related expression patterns. In summary, our observations suggest that approximately 30% of usRNAs could have evolved to participate in biological processes including gene-silencing.

## Figure 1: Identification of usRNAs



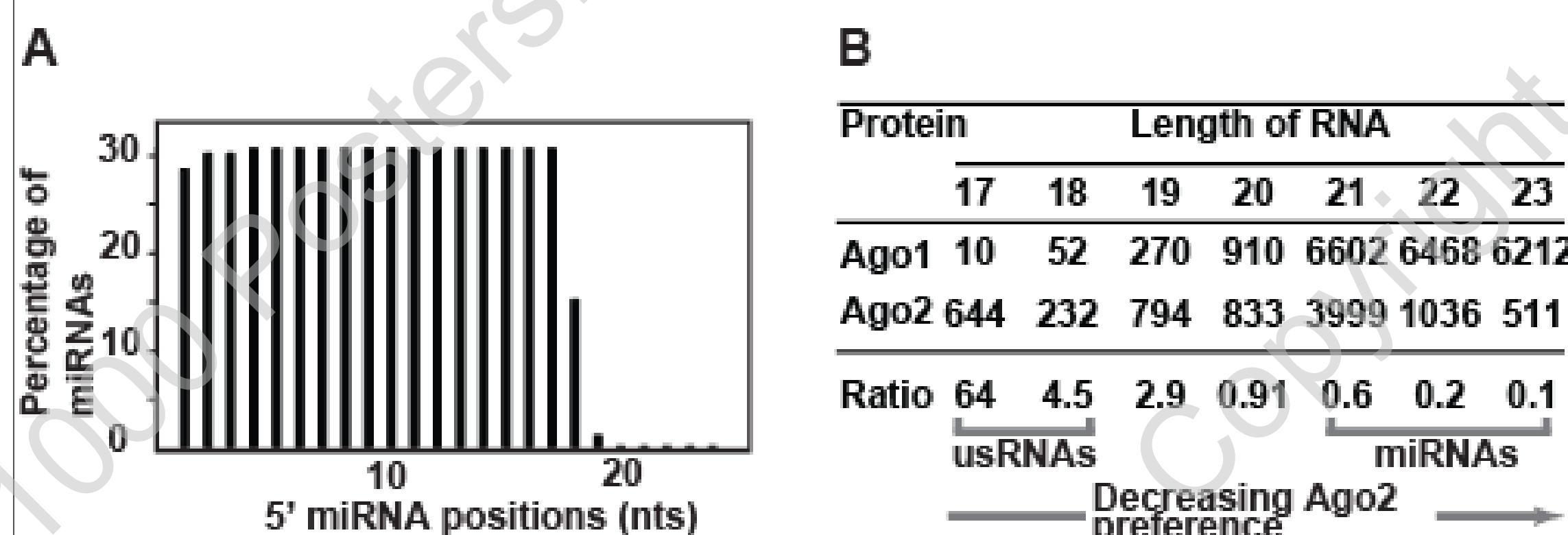
**Figure 1: Detection of us-k12-1 and other putative small RNAs based on two independent RNA fractions A (8–18 nts), and B (18–30 nts).** (A) Northern blot detects K12-1 (~22 nts) and us-K12-1 (~16 nts) in KSHV-positive cells (BCBL-1) but not in KSHV negative cells (BJAB). Ethidium bromide staining (below) confirms equal loading. (B) The number of sequence reads for various sizes of RNAs.

## Figure 2: Annotation of usRNAs



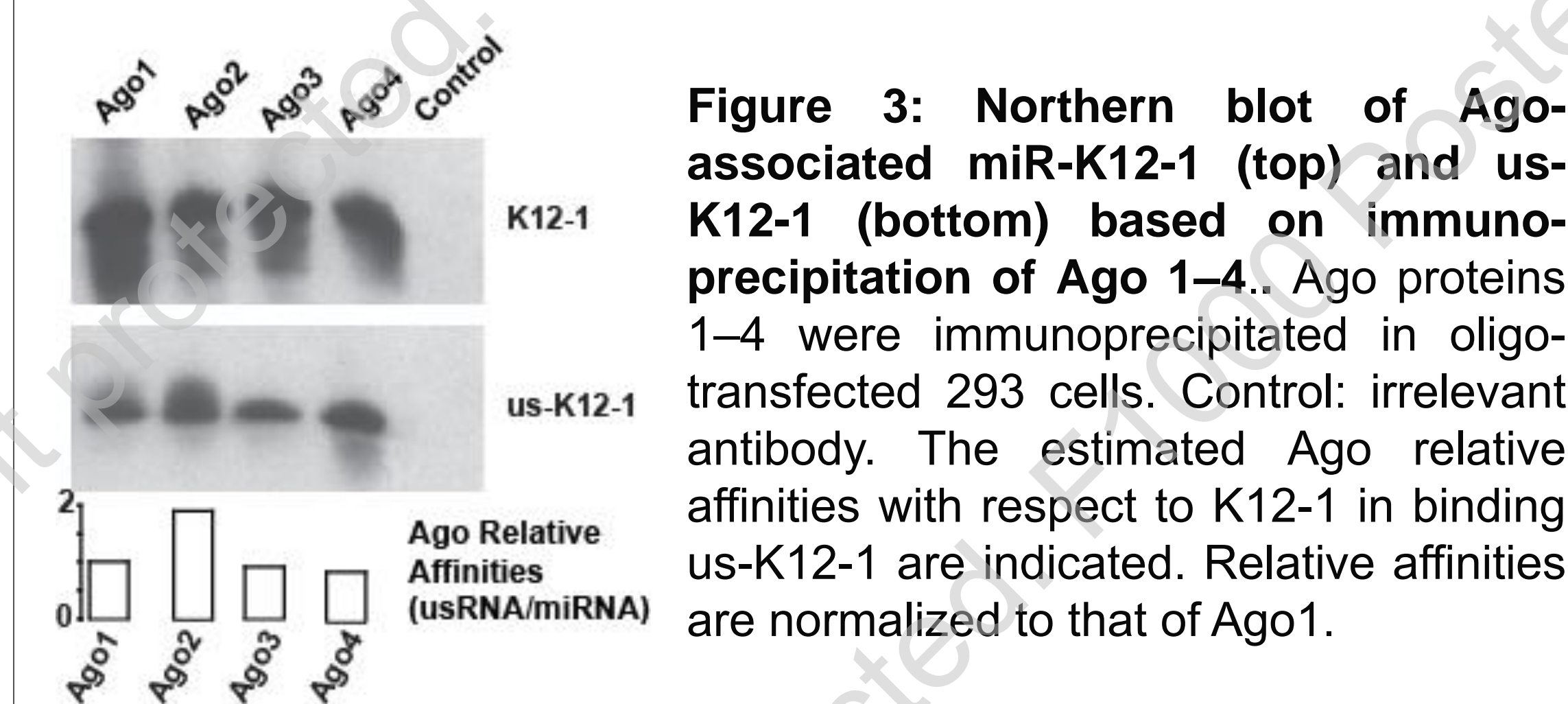
**Figure 2: Annotation of usRNAs.** (A) Summary of the number of usRNA reads derived from different RNA classes. Unannotated: usRNAs mapping to genome regions without annotation. Unmapped: sequences that do not reliably map to human or KSHV genome. (B) Percentage of distinct miRNA-derived usRNAs (upper panel), unannotated usRNAs (middle panel), and miRNAs (bottom panel) mapping to different genomic regions. Only those usRNAs with unique genomic mapping locations were considered. UTR: untranslated region. CDS: coding sequence. Promoter: 3000 nts upstream of transcription start sites. Note that in comparison to miRNAs, significantly more usRNAs map to antisense strand of introns.

## Figure 3: Ago-associated small RNA analysis



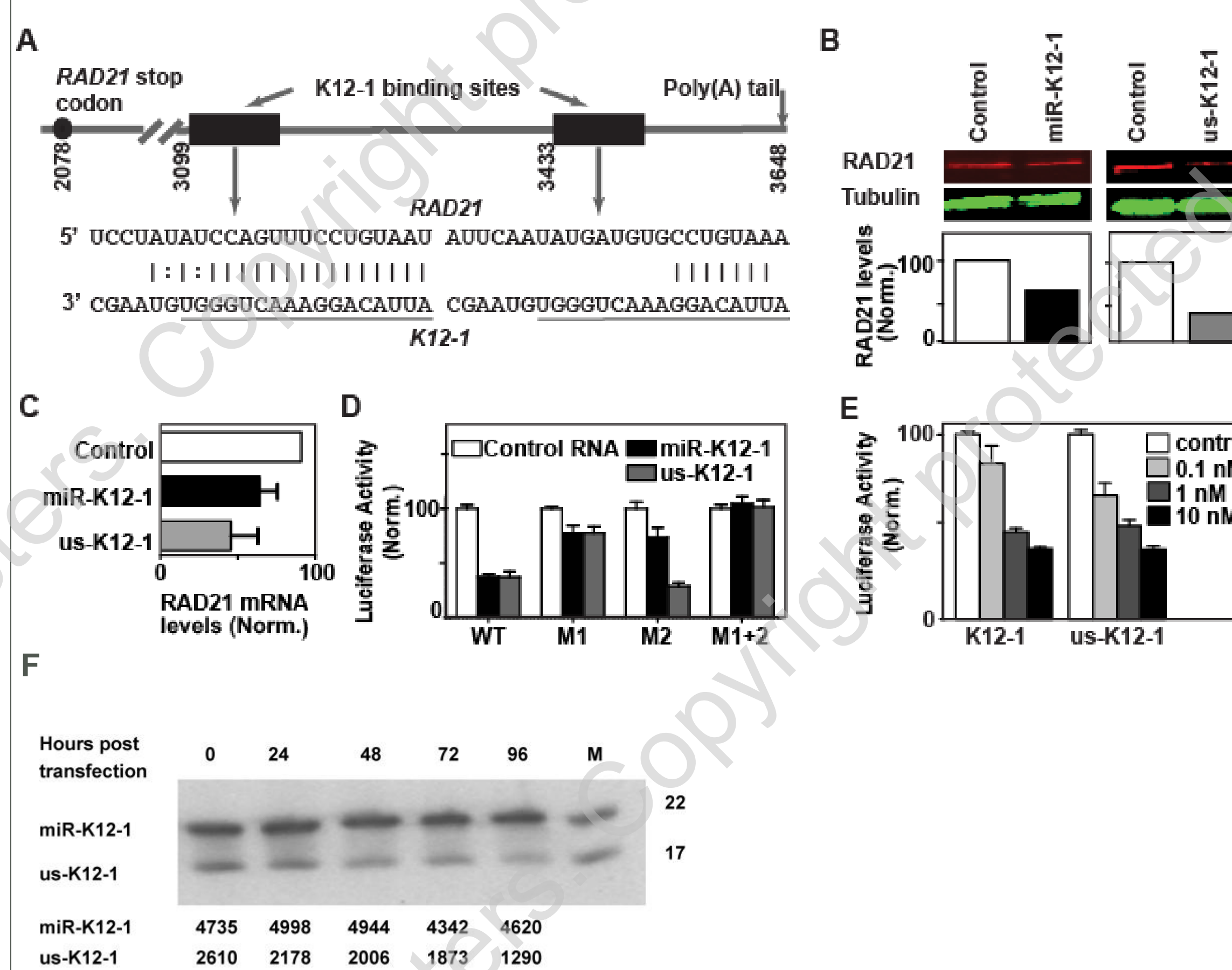
**Figure 2: Characteristics of Ago-associated miRNA-derived usRNAs.** (A) Ago-associated miRNA-derived usRNAs primarily match to 5' positions 1 to 17 of miRNAs and rarely map to 3' ends of miRNAs; percentage of miRNAs that match to usRNAs (17–18 nts) is plotted at each miRNA position. (B) Number of sequence reads associated with Ago1 and Ago2 and their ratios (Ago2 vs. Ago1) indicate preferential association of Ago2 with 17–18 bases long miRNA-derived usRNAs.

## Figure 4: Verification of Ago association



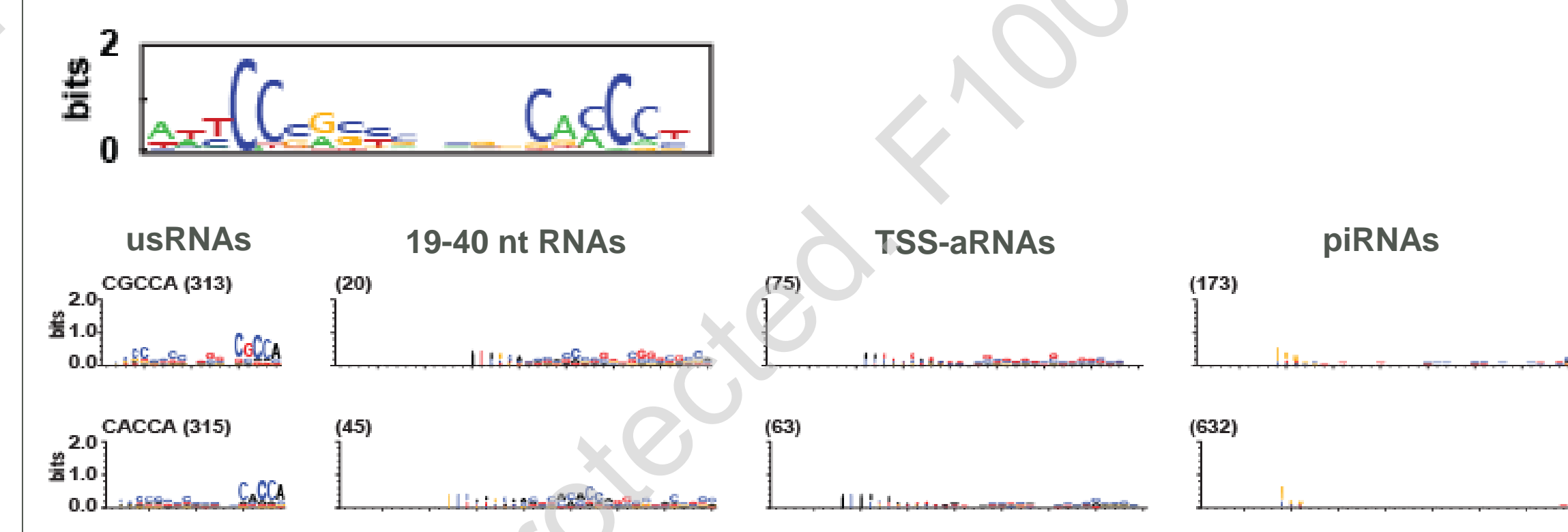
**Figure 3: Northern blot of Ago-associated miR-K12-1 (top) and us-K12-1 (bottom) based on immunoprecipitation of Ago 1–4.** Ago proteins 1–4 were immunoprecipitated in oligo-transfected 293 cells. Control: irrelevant antibody. The estimated Ago relative affinities with respect to K12-1 in binding us-K12-1 are indicated. Relative affinities are normalized to that of Ago1.

## Figure 5: Functional studies



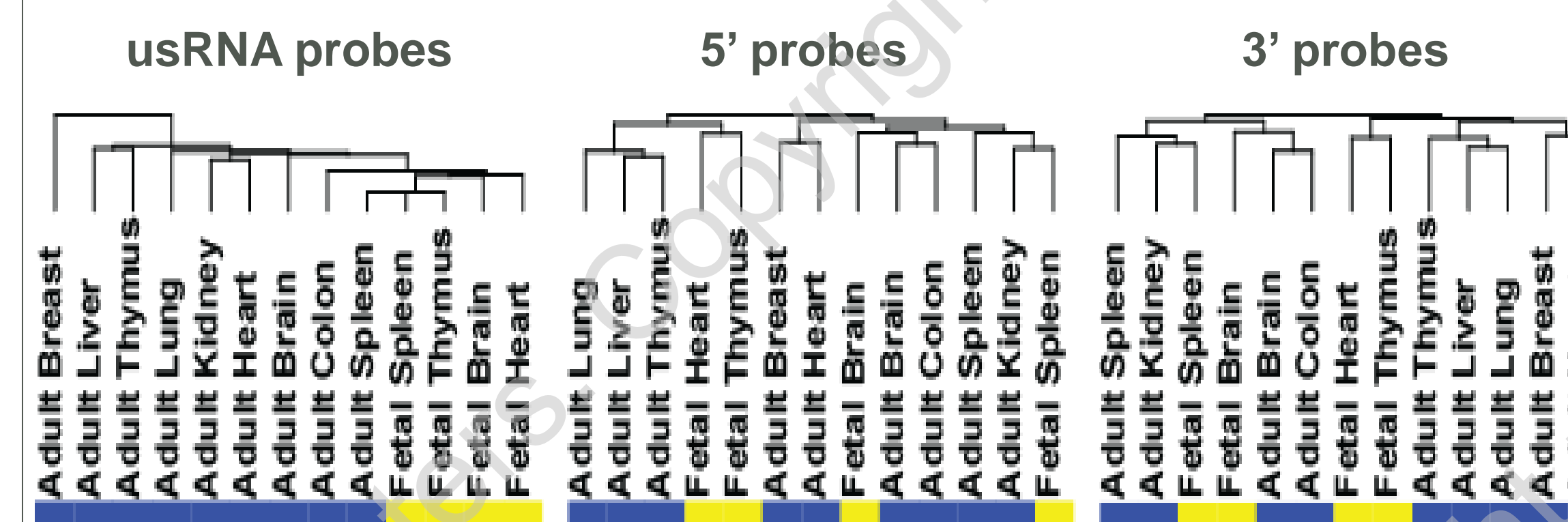
**Figure 4: Both miR-K12-1 and us-K12-1 can downregulate RAD21.** (A) The 3' UTR of the *RAD21* mRNA contains two putative binding sites for miR-K12-1 and us-K12-1 (underline) at positions 3099 (Site 1) and 3433 (Site 2). (B) Immunoblots of endogenous RAD21 after miR-K12-1 or us-K12-1 transfection. (C) Real time RT-PCR of RAD21 transcripts after miR-K12-1 or us-K12-1 transfection. (D) Normalized luciferase activity for wild-type (WT) or mutated (M1, M2, and M1+2) UTR after transfection of K12-1 miRNA or us-K12-1. Note that although mutations in either binding site affect miR-K12-1 activity, us-K12-1 activity is largely unaffected by mutations in the canonical miRNA binding site (Site 2). (E) Dose response of luciferase activities for wild type UTR after miR-K12-1 or us-K12-1 transfection. (F) Half-lives of miR-K12-1 and us-K12-1 after transfection. The relative intensities of the bands are indicated.

## Figure 6: Motif analysis



**Figure 5: Motif analysis of usRNAs.** Upper panel: 16% of Ago-associated non-miRNA-derived usRNAs C-rich motifs at 3' end and at 5' positions (4 & 5). Bottom panel: Motif analysis of the whole usRNA population reveals two C-rich motifs (CGCCA and CACCA), that are overrepresented in usRNAs and are located exactly at their 3' ends. In contrast, such strongly represented motifs are absent in longer RNAs.

## Figure 7: Expression profiling



**Figure 6: Hierarchical classification of adult and fetal tissues based on expression patterns of usRNAs and control regions.** Microarray signals of selected unannotated usRNAs (left panel) are informative to separate fetal (yellow) from adult (blue) tissues. The usRNAs are also significantly more informative than that of their control 5' (middle panel) and 3' (right panel) genomic flanking sequences in classifying tissues of fetal and adult, with p-values of  $10^{-29}$  and  $10^{-42}$ , respectively.

## Discussion

The identification of a viral RNA product (us-K12-1) that is shorter than canonical miRNAs led us to investigate the characteristics of usRNAs. We found that ~30% of the cloned miRNAs produce usRNAs specifically from the 5' ends, which are also associated with Ago proteins and manifest a stronger preference for Ago2 than Ago1. Using us-K12-1 as an example, we demonstrated that miRNA-derived usRNAs can downregulate target genes via base-pairing interactions in 3' UTR of the target gene. In addition to miRNA-derived usRNAs, we found thousands of 15–18 nt base long usRNAs that do not derive from miRNAs. Hundreds of these non-miRNA-derived usRNAs harbor C-rich motifs that are both overrepresented and highly preferred at precise locations at their 3' ends. We also found that many non-miRNA-derived usRNAs associate with Ago proteins and that such Ago-associated usRNAs frequently contain C-rich motifs. Through large scale expression profiling, we found that non-miRNA-derived usRNAs manifest development-related expression patterns. Taken together, the detailed analysis of the diverse group of usRNAs point to a biologic role for many of the previously ignored small RNAs that may or may not derive from longer functional gene transcripts.

## Acknowledgement

We thank Dr. Gunter Meister and Dr. Mikiko C. Siomi for generously providing various Ago Antibodies. This project was supported by funds from the University of Pittsburgh (BJ), NIGMS/NIH grant (BJ), American Cancer Society (BJ), NCI/NIH grants (PSM and YC) and, in part, under a grant from the Pennsylvania Department of Health (PSM and YC).