Extended data

Genome-Wide Dual-Selection Unveils Novel Self-Cleaving Ribozymes in the Human Genome

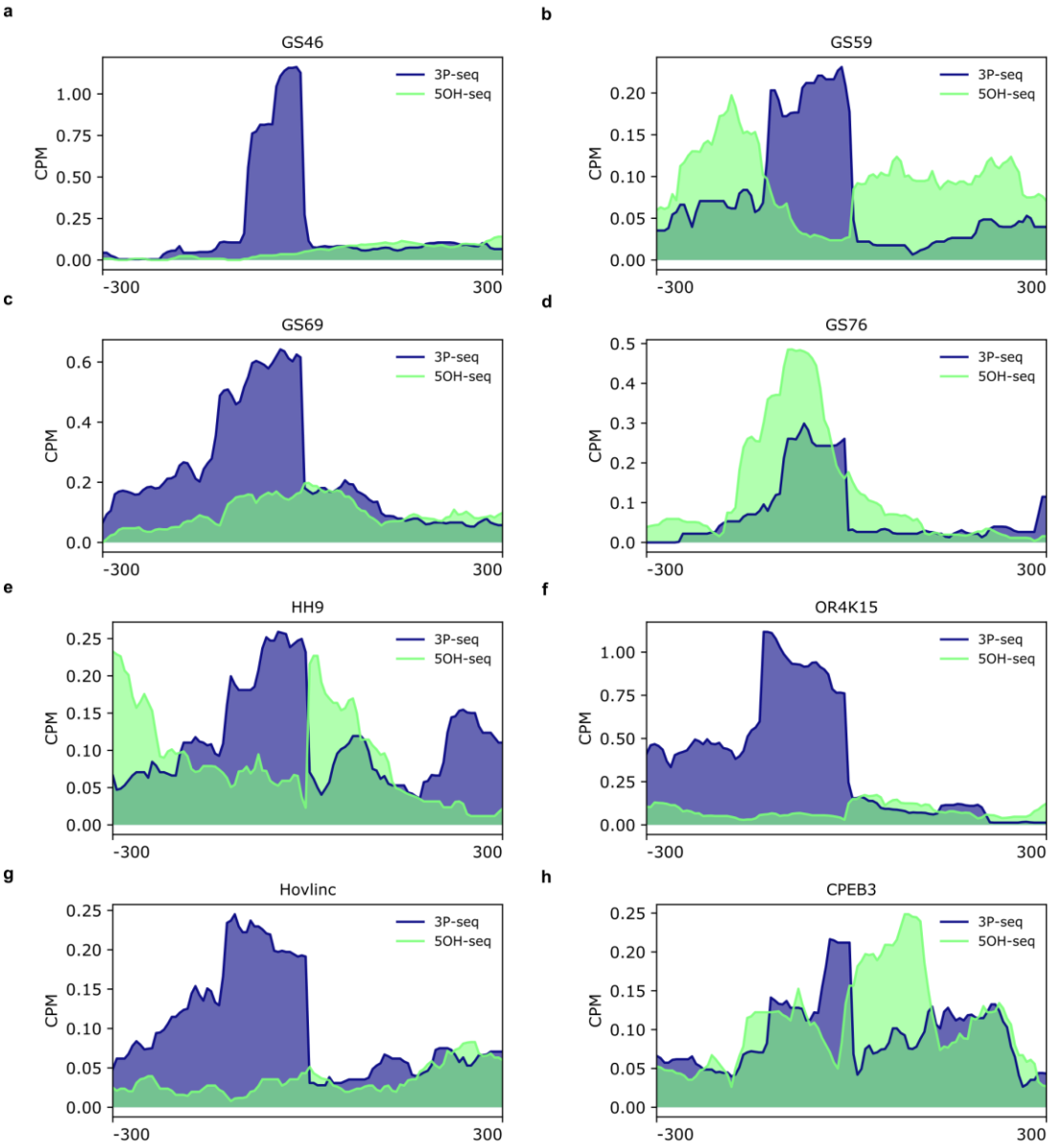
Zhe Zhang[1], Jian Zhan[1,2,3*] and Yaoqi Zhou[1*]

[1]Institute for Systems and Physical Biology, Shenzhen Bay Laboratory, Shenzhen 518107, China

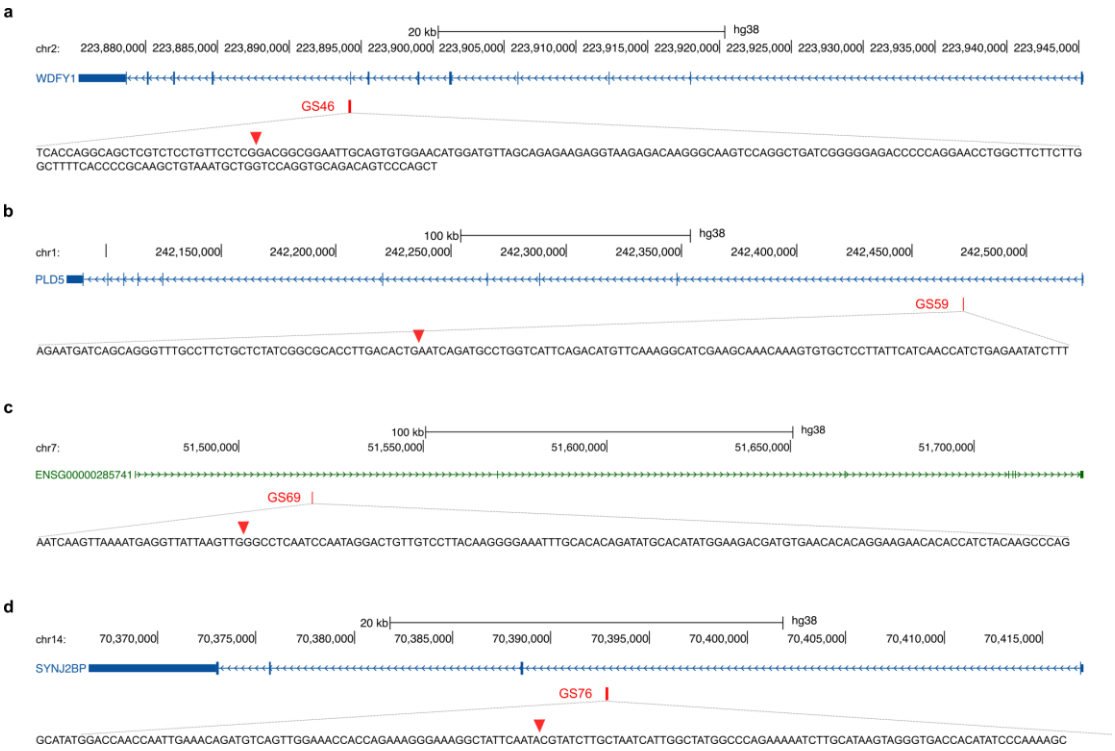
[2]Ribopeutic (Shenzhen) Co., Ltd., Futian, Shenzhen, Guangdong Province, 518000, China

[3]Ribopeutic Inc., Qiantang, Hangzhou, Zhejiang Province, 310018, China

[*] Co-corresponding authors (zhanjian@szbl.ac.cn and zhouyq@szbl.ac.cn)

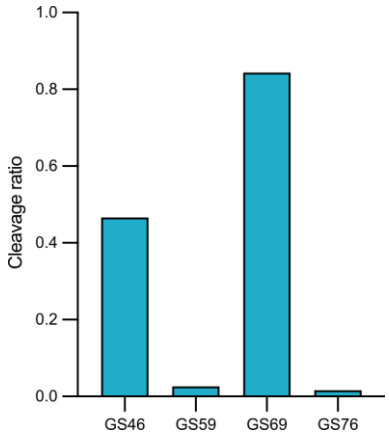


Extend Data Fig. 1 Normalized signal from the two genome-wide selection assays.
5OH-seq (green) and 3P-seq (blue) signals across a 600 nt genomic window centered on the cleavage site of each ribozyme. The y-axis represents counts per million (CPM) of uniquely mapped reads. **a–d**, Four novel ribozymes were identified in this study. **e-f,** Two previously reported ribozymes (HH9 ranked 34th; OR4K15 ranked 56th) were recovered among the top 96 hits in the validation screen. **g-h**, Two known ribozymes ranked lower in the selection, indicating their relatively lower abundance or activity under the experimental conditions.



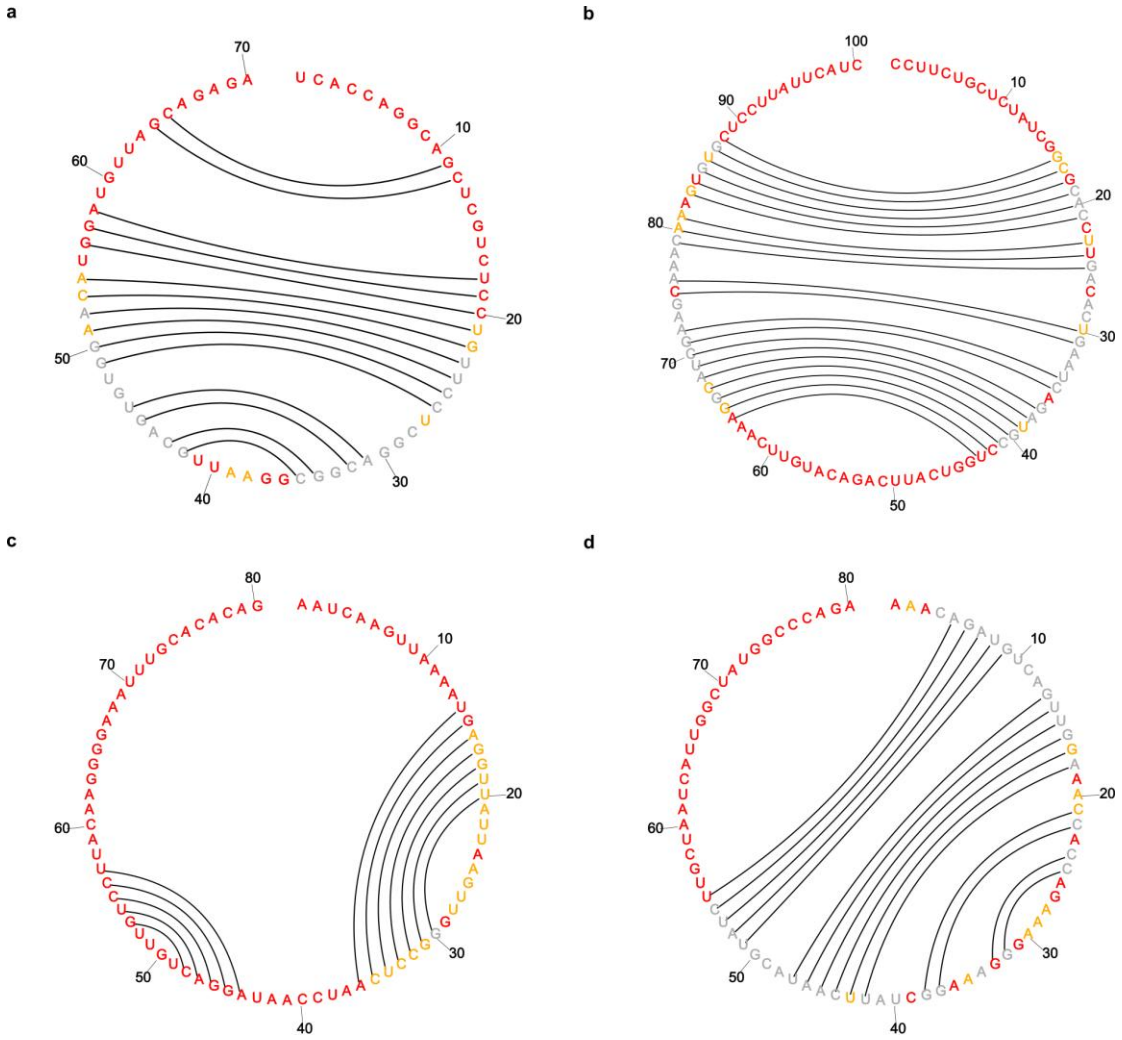
Extend Data Fig. 2 Genomic localization of the four identified ribozymes in the human reference genome (hg38).

Each panel displays a specific genomic region containing one of the ribozymes (GS46, GS59, GS69, GS76). The genomic coordinate scale (hg38) is shown at the top of each panel. Annotated gene models are represented by blue bars, with the gene name indicated at left. The precise location of each ribozyme is marked by a red bar and its corresponding label (e.g., GS46). The relevant genomic DNA sequence spanning the ribozyme locus is shown in black text below each map. **a**, GS46 is located within the eighth exon of the WDFY1 gene on chromosome 2. **b**, GS59 resides in the first intron of the PLD5 gene on chromosome 1. **c**, GS69 is situated in an intron of a novel transcript ENSG00000285741on chromosome 7. **d**, GS76 is found within the reverse strand of an intron of the SYNJ2BP gene on chromosome 14.



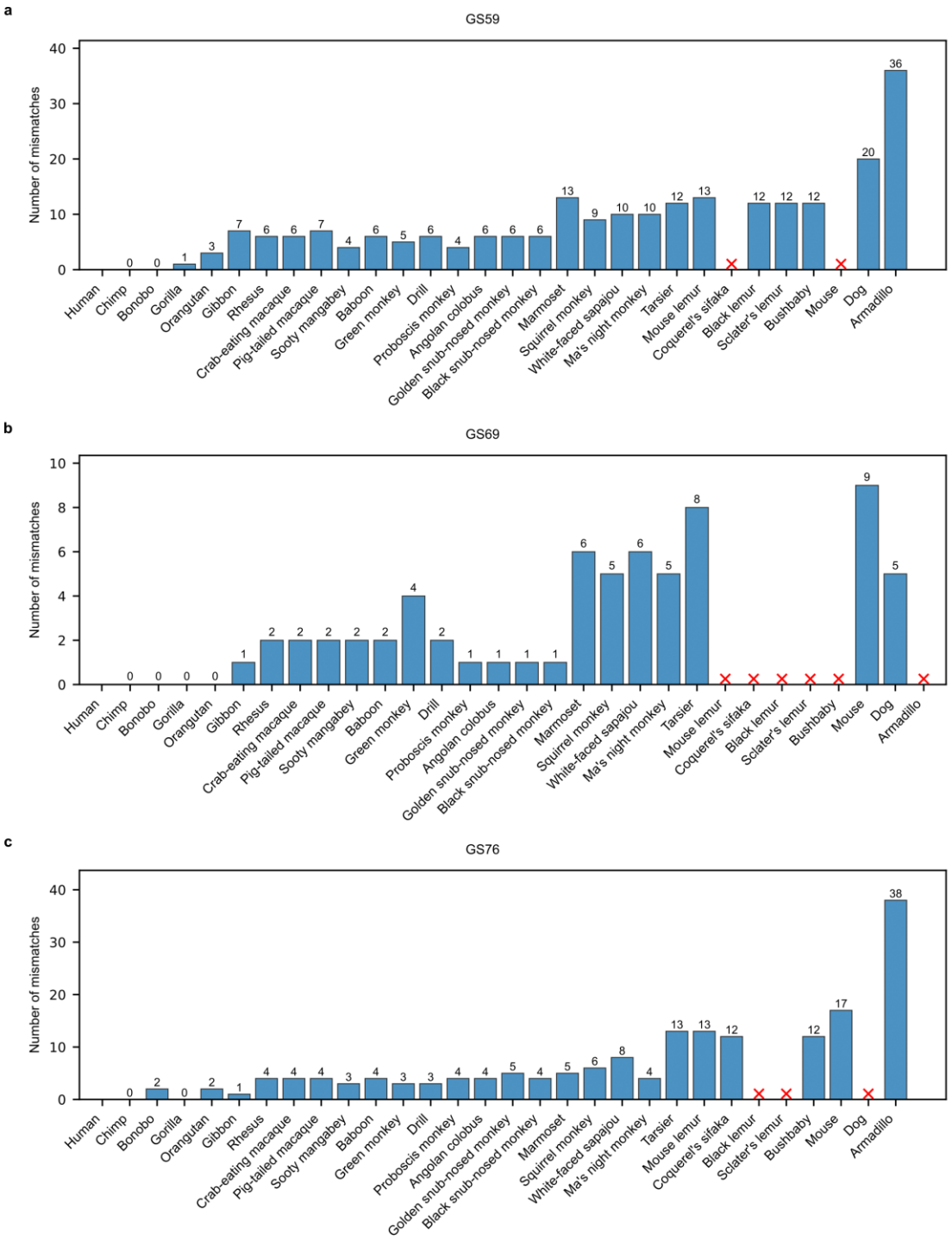
Extend Data Fig. 3 Cleavage ratios of the four novel ribozymes in the mutational scanning experiment.

The cleavage fraction (y-axis) for each wild-type ribozyme was calculated from mutational scanning RNA-seq data as: (number of cleavage reads) / (cleavage reads + uncleaved reads). Values are as follows: GS46, 46.6%; GS59, 2.5%; GS69, 84.4%; GS76, 1.6%. The self-cleaving ribozyme GS69 shows the highest activity, while GS59 and GS76 display minimal cleavage under the tested conditions.



Extend Data Fig. 4 Predicted secondary structures of the 4 truncated ribozymes.

Secondary structures were computationally modeled using RNAstructure, with pseudo-SHAPE reactivity values from mutational scanning data applied as folding constraints. (**a**) GS46, (**b**) GS59, (**c**) GS69, and (**d**) GS76. The circular layout displays nucleotide sequences around the perimeter; lines inside the circle represent predicted base-pairing interactions. Nucleotide colors correspond to pseudo-SHAPE reactivity values: red (≥0.85), orange (0.4–0.85), and grey (no data), following the standard RNAstructure coloring scheme. The overall structure highlights regions of high base-pairing stability supported by mutational fitness.



Extend Data Fig. 5 Conservation of novel ribozyme sequences across vertebrate species. Number of nucleotide mismatches in sequence homologs of ribozymes GS59 (**a**), GS69 (**b**), and GS76 (**c**) relative to the human reference. Data are derived from the UCSC 30 primates track; non-primate vertebrates (e.g., dog, armadillo) are included for broader comparison. Blue bars indicate the mismatch count for each species; a red "X" denotes cases where a homologous sequence was not detected. Lower mismatch counts reflect higher evolutionary conservation. Note: GS46 conservation is shown separately (see Fig. 6). The overall pattern indicates strong conservation in primates, especially in human and chimpanzee, and greater divergence in more distant vertebrates.