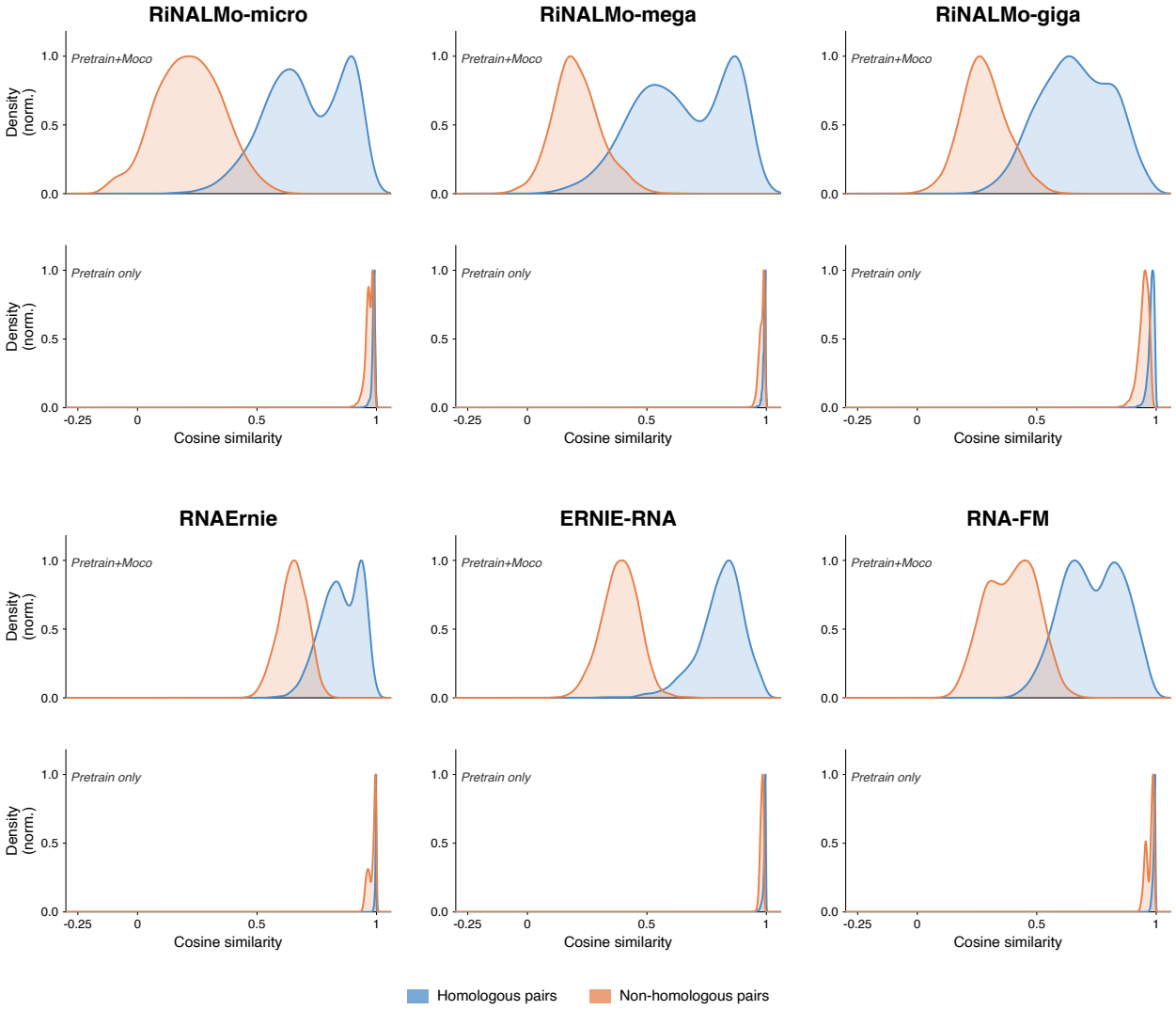
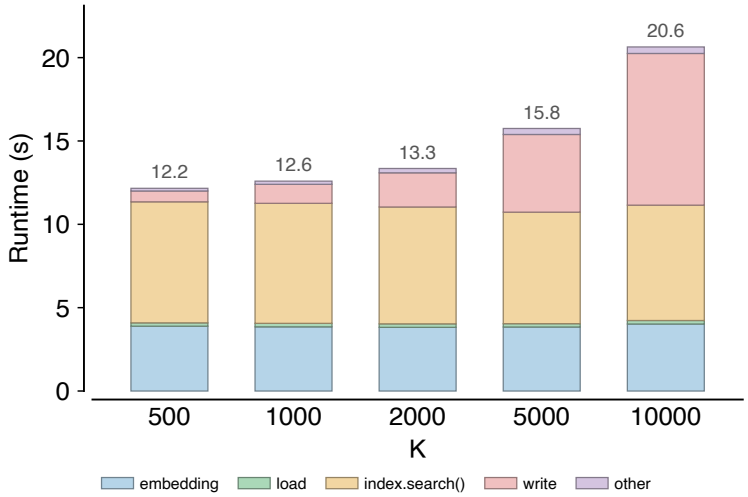
Supplementary Materials

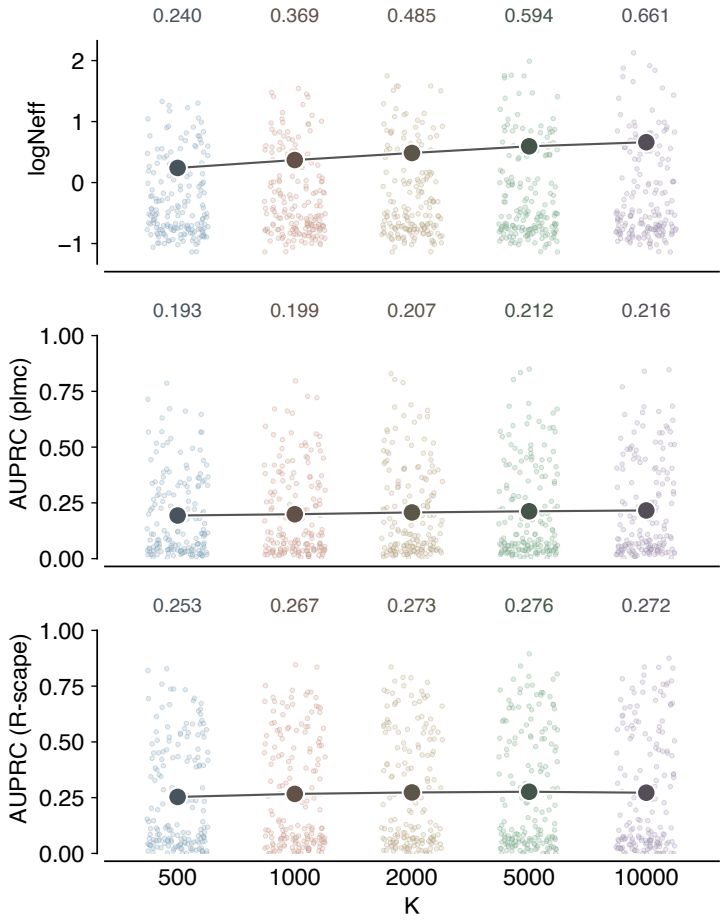
RNAflash: Fast and Sensitive Structure-aware RNA Homology Detection and Alignment



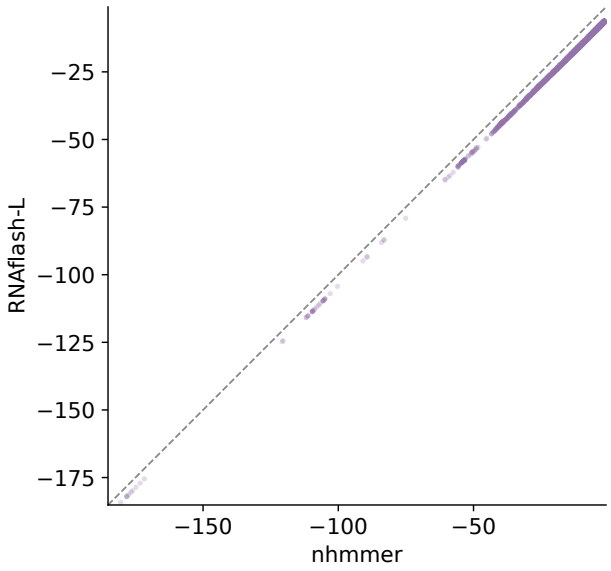
Supplementary Figure 1: Kernel density estimates (KDEs) of the pairwise cosine similarity between sequence embeddings, using different backbones.



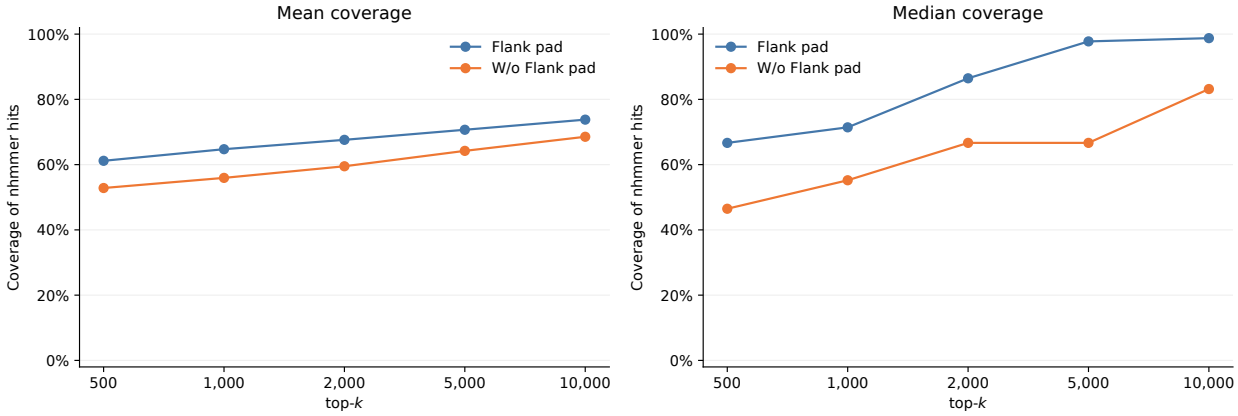
Supplementary Figure 2: Retrieval time for different retrieval depth of RNAflash-R. Compared with runtime of RNAflash-A, the retrieval time did not change significantly with the increase of retrieval depth. The rise in retrieval time was mainly due to the larger volume of data read and written.



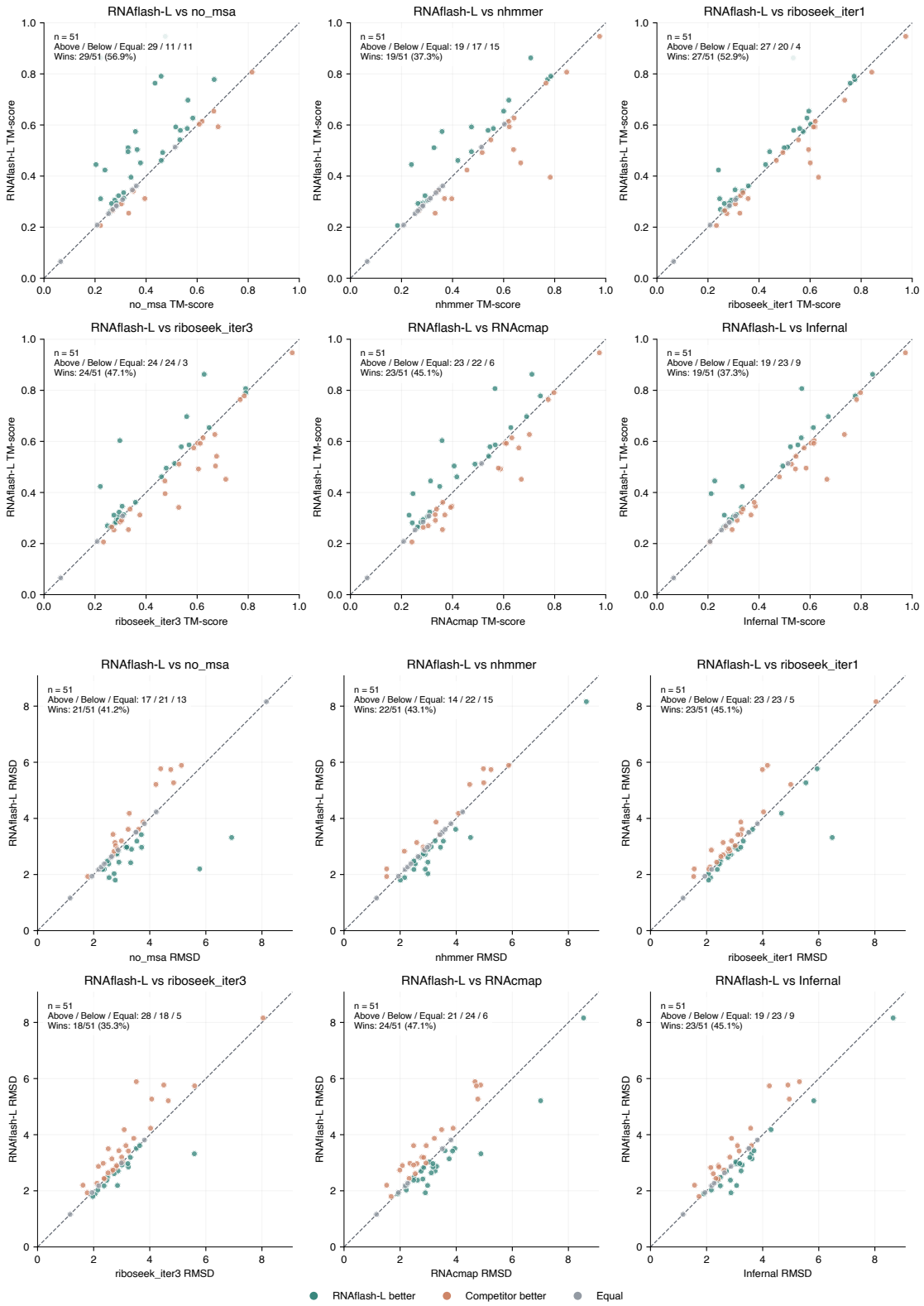
Supplementary Figure 3: Effect of retrieval depth k on RNAflash-L MSA quality. logNeff increases with depth, while AUPRC gains largely plateau at $k = 5000$, supporting this setting.



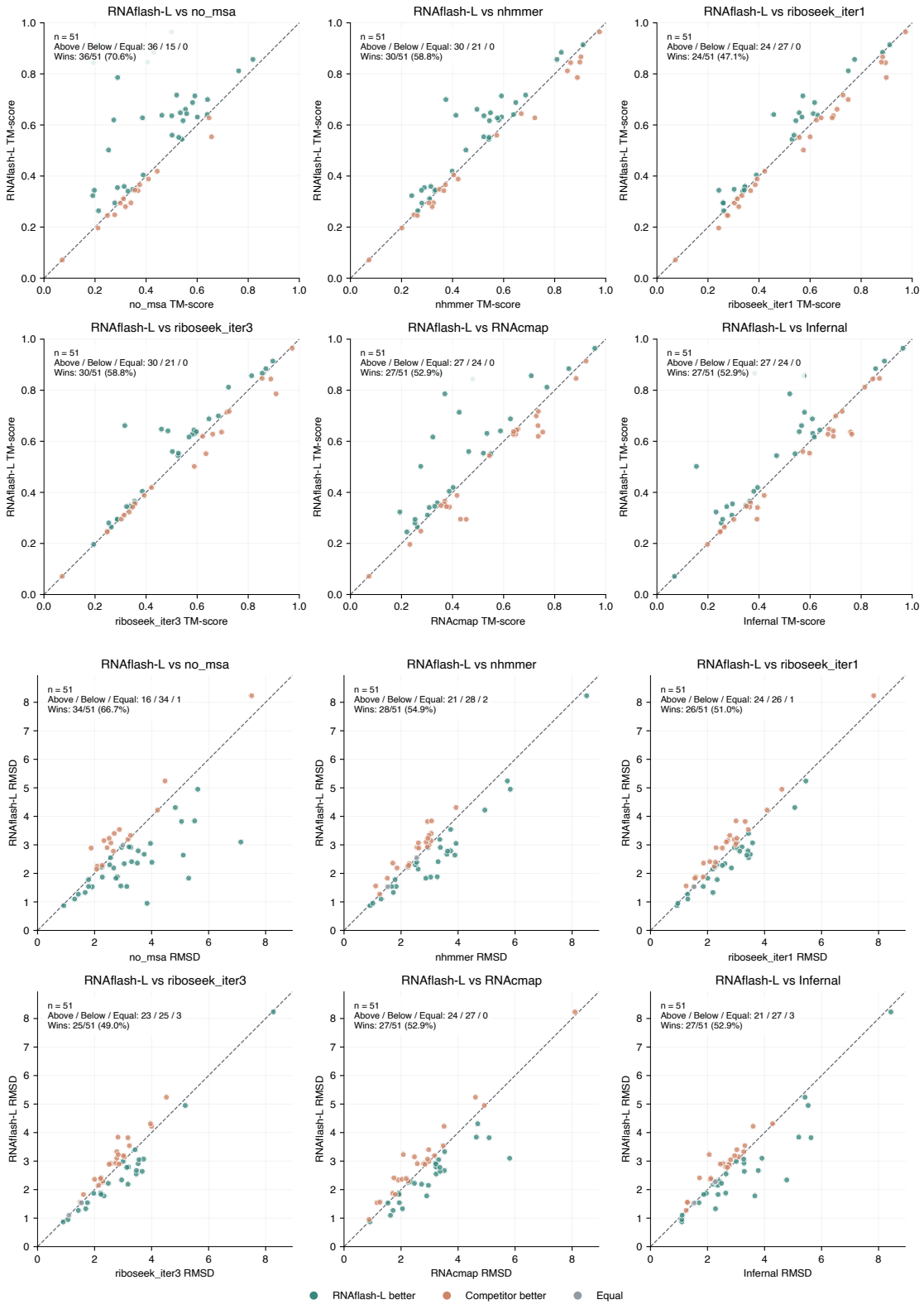
Supplementary Figure 4: Comparison of E-values between nhmmer alone and RNAflash-L in log-10 scale. Each point represents a hit shared by both methods for the same query, matched by target identifier, alignment coordinates, and aligned residues.



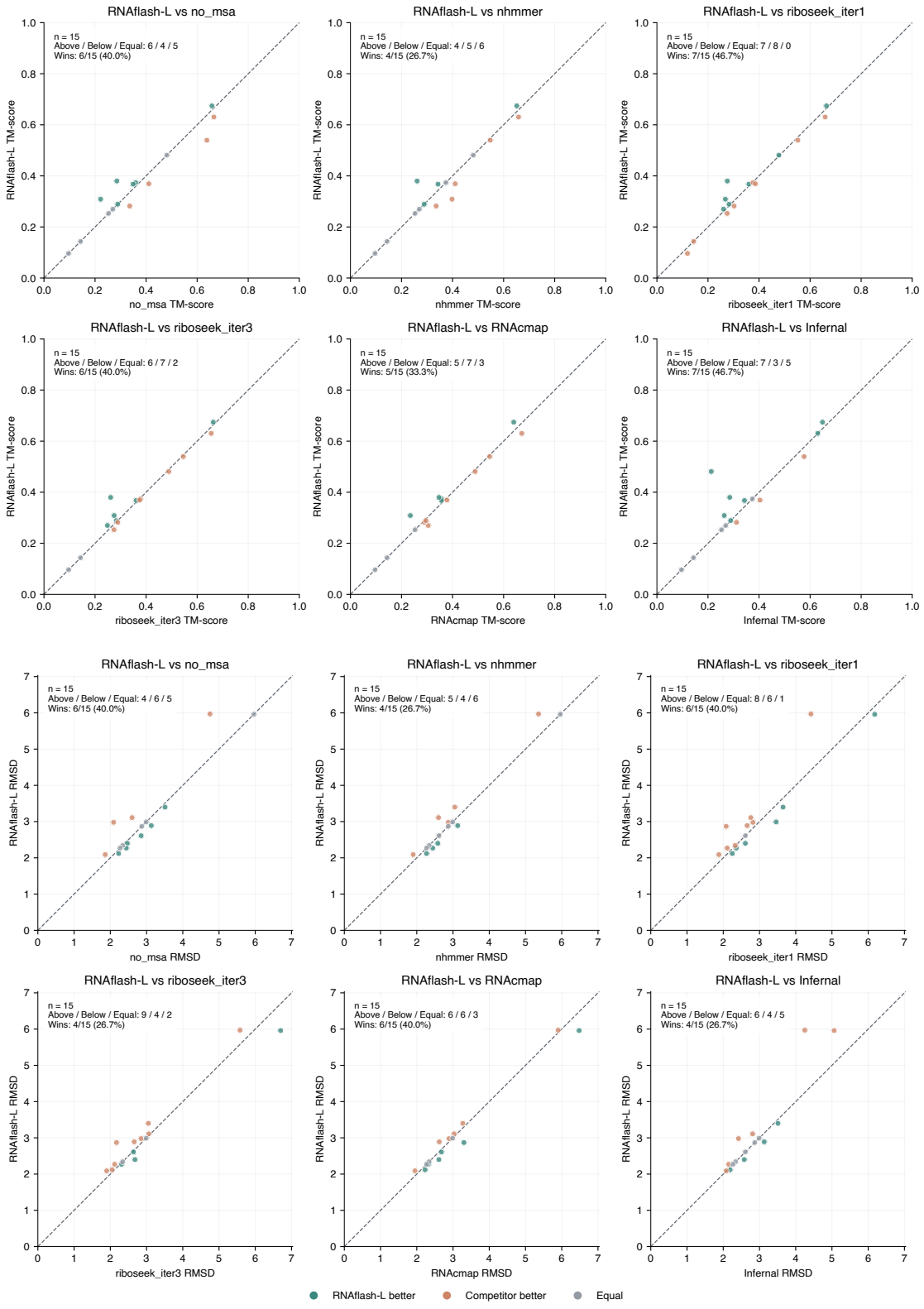
Supplementary Figure 5: Flank augmentation improves the recovery of nhmmer hits across retrieval depths.



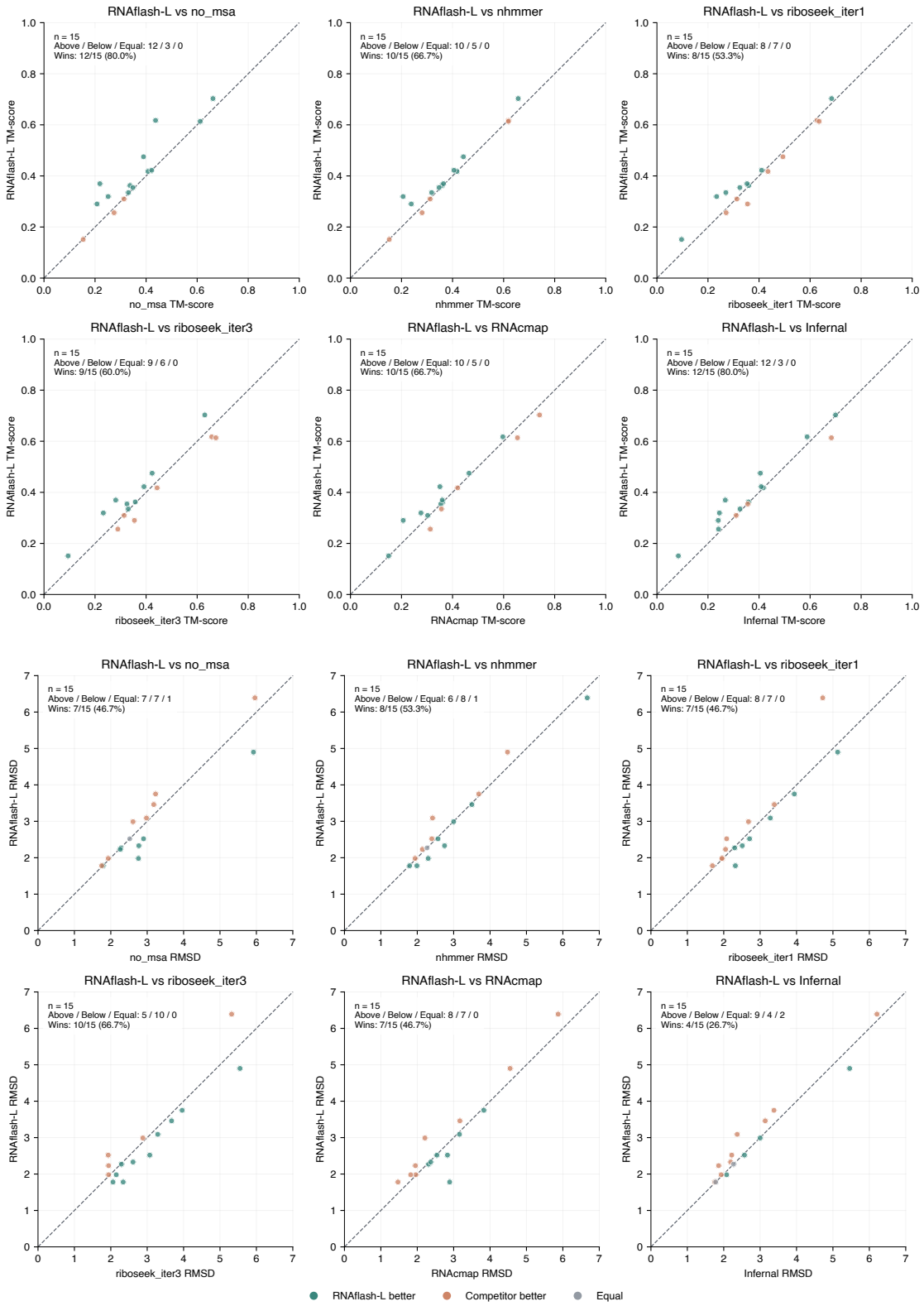
Supplementary Figure 6: Pairwise comparison of RNAflash-L and six benchmark searching methods for RNA structure prediction using AlphaFold3, running on RNAsolo dataset. Each point represents one RNA target, with scores obtained using RNAflash-L on the y-axis and the indicated comparison method on the x-axis.



Supplementary Figure 7: Pairwise comparison of RNAflash-L and six benchmark searching methods for RNA structure prediction using Protenix-v2, running on RNAsolo dataset.



Supplementary Figure 8: Pairwise comparison of RNAflash-L and six benchmark searching methods for RNA structure prediction using AlphaFold3, running on Foldbench dataset.



Supplementary Figure 9: Pairwise comparison of RNAflash-L and six benchmark searching methods for RNA structure prediction using Protenix-v2, running on Foldbench dataset.

Supplementary Table 1: Tertiary structure prediction performance on RNAsolo single-RNA dataset.

	AlphaFold3		Protenix-v2	
MSA source	TM-score	RMSD	TM-score	RMSD
No MSA	0.386	3.309	0.412	3.208
nhmmer	0.450	3.214	0.507	2.877
Riboseek (iter=1)	0.449	3.135	0.526	2.748
Riboseek (iter=3)	**0.456**	**3.034**	0.515	2.770
Infernal	0.448	3.149	0.502	2.892
RNAcmap	0.445	3.154	0.491	2.860
RNAflash-L	0.453	3.189	**0.532**	**2.714**

Supplementary Table 2: Tertiary structure prediction performance on Foldbench single-RNA dataset.

	AlphaFold3		Protenix-v2	
MSA source	TM-score	RMSD	TM-score	RMSD
No MSA	0.364	2.957	0.358	2.989
nhmmer	**0.367**	3.016	0.383	2.927
Riboseek (iter=1)	0.360	2.945	0.391	2.847
Riboseek (iter=3)	0.356	3.007	0.386	3.000
Infernal	0.340	**2.884**	0.375	**2.814**
RNAcmap	0.360	3.125	0.394	2.862
RNAflash-L	0.364	3.085	**0.400**	2.931

Supplementary Table 3: Descriptions of the 174 queries in the RNAsolo dataset. Remarks indicate membership in the supplied lists: U, unpaired (i.e., not used for covariance and coupling analyses); R, Rfam unassigned; 3D, used for tertiary structure prediction (i.e., released after 2022). A dash indicates that none applies.

Name	Chain	Description	Remarks
1ET4	A	RNA APTAMER, 35-MER	R
1FIR	A	HIV-1 REVERSE TRANSCRIPTION PRIMER TRNA(LYS3)	—
1FOQ	B	BACTERIOPHAGE PHI29 PROHEAD RNA	—
1KH6	A	JIIIabc	—
1L2X	A	RNA pseudoknot	—
1NBS	B	RIBONUCLEASE P RNA	—
1Q96	A	sarcin/ricin 28S rRNA	R
1Y26	X	Vibrio vulnificus A-riboswitch	—
1Y27	X	Bacillus subtilis xpt	—
2IL9	A	Ribosomal Binding Domain of the IRES RNA	—
2N0R	A	RNA (5'-R(*GP*CP*UP*GP*AP*GP*CP*UP*CP*GP*AP*AP*AP*GP*AP*GP*CP*AP*AP*UP*GP*AP*UP*GP*UP*C)-3')	—
2QUS	A	Hammerhead ribozyme	—
2QWY	C	SAM-II riboswitch	—
3D0U	A	Lysine Riboswitch RNA	—
3D2V	A	TPP-specific riboswitch	—
3DIL	A	RNA (174-MER)	—
3E5C	A	SMK box (SAM-III) Riboswitch	—
3F2X	X	FMN riboswitch	—
3IVN	A	A-riboswitch	—
3K1V	A	PreQ1 riboswitch	—
3LA5	A	Adenosine RIboswitch	—
3NPQ	A	S-ADENOSYLHOMOCYSTEINE RIBOSWITCH	—
3OXE	A	domain II of glycine riboswitch	—
3PDR	X	M-box Riboswitch RNA	—
3Q3Z	V	c-di-GMP-II riboswitch	—
3RG5	B	tRNA(Sec)	—
3RKF	C	Guanine riboswitch	—
3SKI	A	RNA (68-MER)	—
3SKL	B	RNA (66-MER)	—
3SUX	X	Riboswitch	—
4AOB	A	SAM-I RIBOSWITCH	—
4ENC	A	Fluoride riboswitch	—
4FEN	B	A24U/U25A/A46G mutant of the B. subtilis xpt-pbuX guanine riboswitch aptamer domain	—
4FRG	B	cobalamin riboswitch aptamer domain	—
4FRN	B	Cobalamin riboswitch aptamer domain	—
4GMA	Z	Adenosylcobalamin riboswitch	—
4JF2	A	PreQ1-II Riboswitch	—
4JRC	B	Distal Stem I region of the glyQS T box leader RNA	—

Continued on next page

Supplementary Table 3 (continued)

Name	Chain	Description	Remarks
4KQY	A	YitJ S box/SAM-I riboswitch	—
4LVW	A	THF riboswitch	—
4LX6	A	Mutated adenine riboswitch aptamer	—
4MGN	A	glyQS T box riboswitch	—
4NLF	A	23S ribosomal RNA Sarcin Ricin Loop	—
4OQU	A	SAM-I/IV riboswitch	—
4P5J	A	Turnip yellow mosaic virus mRNA for the coat protein	—
4P8Z	A	Didymium iridis partial IGS, 18S rRNA gene, I-DirI gene and partial ITS1	—
4PCJ	A	trCUG-3('5)	—
4PLX	A	Core ENE hairpin and A-rich tract from MALAT1	—
4PQV	A	XRN1-resistant flaviviral RNA	—
4QLM	A	RNA (108-MER)	—
4R4V	A	VS ribozyme RNA	—
4RBQ	A	U-Helix RNA from Trypanosome editing	U
4RGE	C	env22 twister ribozyme	—
4RUM	A	NiCo riboswitch RNA	—
4WFL	A	RNA	—
4XNR	X	Vibrio Vulnificus Adenine Riboswitch	—
4XWF	A	pfl RNA	—
4Y1M	B	E. coli yybP-ykoY riboswitch	—
4YAZ	R	RNA (84-MER)	—
4ZNP	A	pfI Riboswitch	—
5BJO	E	RNA (36-MER)	—
5BTM	A	RNA (55-mer)	—
5BTP	B	RNA (62-MER)	—
5D5L	A	PreQ1-II riboswitch	—
5FJC	A	SAM-I RIBOSWITCH	—
5KPY	A	5-hydroxytryptophan RNA aptamer	R
5LYS	B	7SK RNA	—
5M0H	A	ASH1 E3 (42 nt-TL/TLR)	R
5ML7	B	23S ribosomal RNA	—
5MWI	A	RNA G(CUG)8C duplex	U
5NWQ	A	Guanidine III riboswitch	—
5SWD	B	Vibrio vulnificus strain 93U204 chromosome II, adenine riboswitch aptamer domain	—
5T83	A	RNA (95-MER)	—
5TPY	A	RNA (71-MER)	—
5U3G	B	ykkC riboswitch	—
5V3F	A	RNA (31-MER)	—
6C27	A	SAM-III riboswitch	—
6CB3	B	RNA (99-MER)	—
6CK5	A	PRPP riboswitch	—

Continued on next page

Supplementary Table 3 (continued)

Name	Chain	Description	Remarks
6CU1	A	YrlA effector-binding module	—
6D3P	A	RNA (45-MER)	—
6DLR	A	PRPP Riboswitch	—
6DME	A	ppGpp Riboswitch	—
6DVK	H	RNA (95-MER)	R
6E7L	A	RNA (5'-R(*GP*GP*GP*CP*UP*GP*CP*AP*CP*UP*UP*CP*GP*GP*UP*GP*CP*UP*GP*CP*CP*C)-3')	U
6E8S	A	iMango-III aptamer	R
6E8T	D	RNA (35-MER)	R
6E8U	B	RNA (37-MER)	R
6FZ0	A	metY SAM V (53-MER)	—
6GYV	A	Lariat-capping ribozyme	—
6H0R	A	SRS2 fragment of Rgs4 3' UTR, RNA (31-MER)	—
6JQ6	U	RNA (81-MER)	—
6JXM	B	RNA (97-mer)	—
6L0Y	A	RNA (5'-R(*GP*GP*UP*AP*GP*AP*CP*GP*CP*UP*UP*CP*GP*GP*CP*GP*UP*UP*UP*GP*CP*C)-3')	U
6MJ0	A	RNA (101-MER)	—
6N2V	A	X. oryzae Mn riboswitch optimized construct	—
6N5P	A	RNA (127-MER)	—
6OL3	C	Adenovirus Virus-Associated (VA) RNA I apical and central domains	—
6P2H	A	RNA (69-MER)	—
6PMO	A	T-box riboswitch discriminator	—
6PMO	B	tRNA-Gly	—
6PRV	A	23S rRNA	—
6SVS	A	RNA (79-mer)	—
6TFF	A	Chains: A	—
6UES	A	RNA (119-MER)	—
6UFM	A	RNA (77-MER)	—
6UFM	B	RNA (98-MER)	—
6UGG	B	tRNAasp	—
6UGI	A	tRNA(Asp) acceptor stem/T stem-loop	—
6VMY	A	B. subtilis cobalamin riboswitch	—
6XKO	A	Class III PreQ1 riboswitch	—
6YMC	A	RNA (26-MER)	3D
7D7W	A	18GAAA (52-MER)	—
7D81	A	832GAAA (50-MER)	—
7ELQ	B	NMT1_Mn (46-MER)	U
7EOG	A	Pepper (49-MER)	—
7JRR	A	RNA (50-MER)	R
7K16	P	RNA (51-MER)	—
7KVT	B	Squash RNA aptamer bound to DFHBI-1T with iridium (III) ions	R, 3D
7L0Z	G	RNA (69-MER)	R

Continued on next page

Supplementary Table 3 (continued)

Name	Chain	Description	Remarks
7LYF	A	RNA promoter stem-loop A fused with tRNA	—
7LYG	A	RNA promoter stem-loop A fused with tRNA	—
7MKT	A	RNA (5'-R(*GP*UP*GP*UP*GP*UP*GP*UP*GP*UP*GP*UP*GP*UP*GP*UP*GP*UP*GP*UP*G)-3')	U, 3D
7MKY	A	RNA (66-MER)	—
7MLW	F	Guanidine-I riboswitch	3D
7OAX	A	Chili RNA Aptamer	—
7REX	B	RNA (34-MER)	—
7SXP	A	NRAS mRNA	U, R, 3D
7TZS	X	RNA (80-MER)	3D
7U4A	A	RNA (70-MER)	3D
7UQ6	B	RNA (86-MER)	3D
7UR5	v	allo-tRNAUTu1	3D
7UZ0	A	AntiT-tRNA flip UCCA	3D
7VFT	A	RNA (5'-R(*GP*GP*GP*CP*CP*AP*GP*CP*AP*GP*CP*AP*GP*CP*AP*GP*CP*AP*GP*GP*UP*CP*C)-3')	U, R, 3D
7VNV	A	RNA (78-MER)	3D
7WIE	V	RNA (50-MER)	3D
8D28	B	RNA (33-MER)	3D
8EYU	A	RNA (49-MER)	3D
8F4O	B	TPP riboswitch aptamer domain	3D
8FZA	A	PreQ1 Riboswitch (30-MER)	R, 3D
8FZA	B	PreQ1 Riboswitch (30-MER)	3D
8HBA	B	RNA (55-MER)	R, 3D
8HZE	B	RNA (36-MER)	R, 3D
8HZL	B	RNA (84-MER)	R, 3D
8ITS	A	RNA (46-MER)	R, 3D
8K0Y	A	RNA (56-MER)	3D
8K15	C	RNA (470-MER)	3D
8K1E	A	RNA (58-MER)	3D
8K7W	A	RNA (49-MER)	R, 3D
8KEB	A	RNA (72-MER)	3D
8S95	C	Lysine tRNA scaffold,Poliovirus cloverleaf RNA	3D
8SP9	C	tRNA scaffold,CVB3 Cloverleaf RNA	3D
8SYK	A	RNA device 43 truncation mutant 3 (U100C)	3D
8TJU	A	DNA/RNA (416-MER)	3D
8TJV	A	RNA (417-MER)	3D
8TJX	N	RNA (440-MER)	3D
8U5Z	A	Mango II variant	R, 3D
8UO6	A	HIV-1 Rev Response Element Stem-Loop II with tRNA scaffold	3D
8UO6	B	HIV-1 Rev Response Element Stem-Loop II with tRNA scaffold	3D
8V1H	A	pre-mascRNA	3D
8V1I	B	mascRNA	3D

Continued on next page

Supplementary Table 3 (continued)

Name	Chain	Description	Remarks
8VY1	B	Chains: A,B,C	3D
8XTS	A	RNA (814-MER)	3D
8XZL	A	RNA (53-MER)	3D
8ZAU	A	RNA (69-MER)	3D
9BZC	A	RNA (87-MER)	3D
9C6J	A	Group IIC intron embedded with the TPP riboswitch	3D
9C6K	X	TPP riboswitch embedded in an RNA scaffold bound to thiamine pyrophosphate	3D
9CXF	A	RNA (200-MER)	3D
9DE7	A	RNA (58-MER)	3D
9DXL	A	RhoBAST aptamer	3D
9FN2	B	SAMURI-SAM	R, 3D
9FN3	B	SAMURI-ProSeDMA	R, 3D
9G7C	A	raiA (224-MER)	3D