

S1. Data used

Table 1: CLASH and CLEAR-CLIP data before and after processing

	miRNAs	Targets	Interactions	miRNAs (after processing)	Targets (after processing)	Interactions (after processing)
CLASH	399	7390	18514	399	7181	18041
CLEAR-CLIP	542	32569	32712	451	9449	20094

Table 2: Training and test data

	miRNAs	Interactions	Training data	Test data
Target-enriched CLASH dataset (with at least 50 interactions for each miRNA)	77	15390	12312	3078
Energy-filtered CLASH dataset (with at least 15 mol/kcal and 20 interactions for each miRNA)	122	16209	12967	3242

Table 3: 10-fold Cross validation results on the training data using both individual and general models

	Training data	Test data	F1 (individual model)	F1 (general model)	Recall (individual model)	Recall (general model)
Target-enriched dataset	11080	1232	0.9292	0.9573	0.8677	0.9180
	11080	1232	0.9198	0.9621	0.8515	0.9269
	11081	1231	0.9337	0.9625	0.8757	0.9277
	11081	1231	0.9258	0.9638	0.8619	0.9301
	11081	1231	0.9244	0.9668	0.8595	0.9358
	11081	1231	0.9333	0.9655	0.8749	0.9334
	11081	1231	0.9268	0.9673	0.8635	0.9366
	11081	1231	0.9221	0.9594	0.8554	0.9220
	11081	1231	0.9272	0.9616	0.8643	0.9261
	11081	1231	0.9356	0.9563	0.8790	0.9163
Energy-filtered dataset	11670	1293	0.9210	0.9477	0.8535	0.9005
	11670	1286	0.9201	0.9460	0.8520	0.8975
	11670	1286	0.9078	0.9485	0.8311	0.9021
	11670	1290	0.9169	0.9591	0.8466	0.9214
	11670	1286	0.9219	0.9485	0.8551	0.9021
	11670	1285	0.9124	0.9587	0.8389	0.9206
	11670	1290	0.9196	0.9523	0.8512	0.9090
	11671	1282	0.9068	0.9459	0.8295	0.8974
	11671	1296	0.9195	0.9540	0.8511	0.9120
	11671	1284	0.9155	0.9561	0.8441	0.9159

S2. Example of the MDPS dynamic programming strategy

```

MED12L  UGCUGC-AGCUCCCUGCCCGGCGCA
          |||||  |||  ||  ||||  ||||
hsa-miR-484  ACGACAAUCG-GGAUCGGG--GCGU
  
```

Fig 1: MDPS alignment between mRNA transcript and miRNA sequence

In order to align a miRNA and target sequence using MDPS strategy, we considered a dynamic programming sequence alignment algorithm mentioned in the manuscript (Section 2.3). Given a *miRNA* ($1 \dots i$) and *target* ($1 \dots j$) the scores $S[i, j, k]$ are calculated using a pretrained position weight matrix (w) and a pretrained state transition matrix (t). Using the definitions mentioned in the manuscript (Section 2.3), here we explain how $S[i, j, k]$ score table is calculated using MDPS strategy with an alignment example between hsa-miR-484 and MED12L mRNA transcript mentioned in the CLASH data (Fig S1). We show how we calculated $S[2, 2, 0]$ (Table 4) step by step.

$S[2,2,0]$ denotes the alignment between the first two nucleotides of the miRNA and the target sequences that ends with a $\{M, N \text{ or } W\}$. In order to calculate $S[2,2,0]$, we also need to calculate $S[1,1, k]$, $k \in \{0,1,2\}$. Now from the initializations mentioned in the manuscript (Section 2.3),

$$S[1, 1, 0] = \log(w(\text{state}(1, 1), 1)) = \log(w(M, 1)) = -0.93$$

$$S[1, 1, 1] = \log(w(B_y, 1)) = -4.06$$

$$S[1, 1, 2] = -\infty$$

Now using the equations stated in the manuscript (Section 2.3) we can calculate $S[2,2,0]$,

$$S[2, 2, 0] = \log(w(\text{state}(2, 2), 2)) + \max \begin{cases} S[1,1,0] + \log(t(\text{state}(1,1), \text{state}(2,2))) \\ S[1,1,1] + \log(t(B_y, \text{state}(2,2))) \\ S[1,1,2] + \log(t(B_x, \text{state}(2,2))) \end{cases}$$

$$\Rightarrow S[2, 2, 0] = \log(w(M, 2)) + \max \begin{cases} S[1,1,0] + \log(t(M, M)) \\ S[1,1,1] + \log(t(B_y, M)) \\ S[1,1,2] + \log(t(B_x, M)) \end{cases}$$

$$\Rightarrow S[2, 2, 0] = -0.54 + \max \begin{cases} -0.93 + (-0.34) \\ -4.06 + (-1.44) \\ -\infty + (-1.52) \end{cases}$$

$$\Rightarrow S[2, 2, 0] = -1.81 \text{ calculated from } S[1,1,0] \text{ as a traceback.}$$

The full alignment score table is shown in Table 5, where the traceback path is also highlighted.

S3. Performance of MDPS

Table 6: Performance comparison between individual and general MDPS models for the two training datasets.

	Pos	F1	Recall	TP	FN
Target-enriched individual model	2396	0.9337	0.8756	2098	298
Target-enriched general model	2396	0.9664	0.9349	2240	156
Energy-filtered individual model	2604	0.9242	0.8591	2237	367
Energy-filtered general model	2604	0.9561	0.9159	2385	219

Table 7: Performance comparison between external tools (MiRanda, RNA22 and TargetScan) alone and the external tools combined with MDPS (target-enriched).

	Positives	Predicted Positives	F1	Precision	Recall	TP	FP
MiRanda on CLASH	3078	20928	0.0463	0.0266	0.1806	556	20372
MiRanda + MDPS on CLASH	3078	15953	0.0551	0.0328	0.1702	524	15429
MiRanda on CLEAR-CLIP	20081	1857429	0.0017	0.0009	0.0811	1629	1855800
MiRanda + MDPS on CLEAR-CLIP	20081	1491514	0.0020	0.0010	0.0753	1513	1490001
RNA22 on CLASH	3078	91134	0.0110	0.0057	0.1680	517	90617
RNA22 + MDPS on CLASH	3078	68216	0.0137	0.0072	0.1592	490	67726
RNA22 on CLEAR-CLIP	20084	3524462	0.0004	0.0002	0.0367	738	3523724
RNA22 + MDPS on CLEAR-CLIP	20084	2603740	0.0005	0.0003	0.0333	669	2603071
TargetScan on CLASH	3078	38240	0.0220	0.0119	0.1475	454	37786
TargetScan + MDPS on CLASH	3078	29700	0.0265	0.0146	0.1413	435	29265
TargetScan on CLEAR-CLIP	20078	3925145	0.0012	0.0006	0.1134	2277	3922868
TargetScan + MDPS on CLEAR-CLIP	20078	3080990	0.0014	0.0007	0.1053	2114	3078876

Table 8: Performance comparison between external tools (MiRanda, RNA22 and TargetScan) alone and the external tools combined with MDPS (energy-filtered)

	Positives	Predicted Positives	F1	Precision	Recall	TP	FP
MiRanda on CLASH	3242	35612	0.0324	0.0177	0.1943	630	34982
MiRanda + MDPS on CLASH	3242	27223	0.0382	0.0214	0.1795	582	26641
MiRanda on CLEAR-CLIP	20081	1857429	0.0017	0.0009	0.0811	1629	1855800
MiRanda + MDPS on CLEAR-CLIP	20081	1431609	0.0020	0.0010	0.0725	1455	1430154
RNA22 on CLASH	3242	121169	0.0088	0.0045	0.1693	549	120620
RNA22 + MDPS on CLASH	3242	89576	0.0110	0.0057	0.1573	510	89066
RNA22 on CLEAR-CLIP	20084	3524462	0.0004	0.0002	0.0367	738	3523724
RNA22 + MDPS on CLEAR-CLIP	20084	2596344	0.0005	0.0003	0.0329	660	2595684
TargetScan on CLASH	3242	67982	0.0146	0.0076	0.1604	520	67462
TargetScan + MDPS on CLASH	3242	51784	0.0180	0.0096	0.1527	495	51289
TargetScan on CLEAR-CLIP	20078	3925145	0.0012	0.0006	0.1134	2277	3922868
TargetScan + MDPS on CLEAR-CLIP	20078	3126308	0.0013	0.0007	0.1047	2103	3124205