

Table 1 Brief network framework description of several pre-selected models.

Model_Name	Description
CnnCrispr	The final model including all units we mentioned
CnnCrispr_NoLSTM	Without biLSTM layer
CnnCrispr_Conv_LSTM	Reverse the order of the convolution layer and the recurrent layer
CnnCrispr_NoBatchNor	Without Batch Normalization layer
CnnCrispr_NoDropout	Without Dropout layer

CnnCrispr is the benchmark model. On this basis, some parts were removed to obtain the comparison models. With the exception of parts mentioned in the description, the structure of the contrast model was completely consistent with the benchmark model. The CnnCrispr_Conv_LSTM model replaced the sequence of convolutional layers and recurrent layer to compare the influence of network sequence on the prediction results.

Table 2 The prediction results in Model Selection.

Test Set	Model	Recall	ROC_AUC	PRC_AUC
Total test set	CnnCrispr	0.857	0.975	0.679
	CnnCrispr_NoLSTM	0.611	0.987	0.651
	CnnCrispr_Conv_LSTM	0.643	0.986	0.67
	CnnCrispr_NoBatchNor	-	0.5	0.504
	CnnCrispr_NoDropout	0.810	0.985	0.625
Hek293t test set	CnnCrispr	0.864	0.971	0.686
	CnnCrispr_NoLSTM	0.631	0.988	0.658
	CnnCrispr_Conv_LSTM	0.660	0.988	0.694
	CnnCrispr_NoBatchNor	-	0.5	0.504
	CnnCrispr_NoDropout	0.816	0.988	0.636
K562 test set	CnnCrispr	0.826	0.995	0.688
	CnnCrispr_NoLSTM	0.522	0.985	0.589
	CnnCrispr_Conv_LSTM	0.565	0.981	0.57
	CnnCrispr_NoBatchNor	-	0.5	0.503
	CnnCrispr_NoDropout	0.783	0.973	0.597

CnnCrispr had the best comprehensive performance in the three test sets, and the calculated recall value was higher than other preselected models.

Table 3 Performance comparison with states-of-the-art models under test pattern 1.

Test Set	Model	auROC	auPRC	Pearson value	Spearman value
Total test set	CnnCrispr	0.975	0.679	0.682	0.154

	CFD	0.942	0.316	0.343	0.140
	MIT	0.77	0.044	0.150	0.085
	CNN_std	0.947	0.208	0.321	0.141
	DeepCrispr	0.981	0.497	-	0.133
	CnnCrispr	0.971	0.686	0.712	0.160
Hek293t test set	CFD	0.936	0.318	0.371	0.143
	MIT	0.756	0.048	0.153	0.084
	CNN_std	0.939	0.204	0.330	0.144
	DeepCrispr	0.984	0.521	-	0.136
	CnnCrispr	0.995	0.688	0.426	0.134
K562 test set	CFD	0.965	0.322	0.336	0.128
	MIT	0.814	0.033	0.057	0.086
	CNN_std	0.983	0.287	0.319	0.132
	DeepCrispr	0.953	0.41	-	0.126

We downloaded the prediction models of CFD, MIT and CNN_std from relevant websites and obtained the prediction results on the same test set as CnnCrispr. Since the training process of CnnCrispr was consistent with DeepCrispr's, we directly used the test results in **additional file 2** given by DeepCrispr for performance comparison. We bolded out the optimal values under the same conditions in the table.

Table 4 Different ratios indicate the possible relationship between three words, w_i, w_j, w_k .

$ratio_{i,j,k}$	w_i is related to w_k	w_i is not related to w_k
w_j is related to w_k	Tends to 1	Very small and tends to 0
w_j is not related to w_k	Greater than 1	Tends to 1

