

Table S3. Results of inter-dataset training and testing of the proposed method for the identification of DS-related homologs

Training set	Testing set	AUC	MCC	Sensitivity	Specificity
Dataset M	Dataset L	0.958	0.901	0.896	0.989
Dataset L	Dataset M	0.969	0.801	0.937	0.954

Only DS-related homologs were used as positive data in this experiment, in which common homologs and non-homologs were both regarded as negative data.