Table 1: False positives and false negatives identified across models, E-value types, and annotation methods (Pfam and InterPro).

Model	E-value Type	False Negatives	False Positives
		Only Pfam-based Evaluation	
Structure	Single-domain	D3GGZ8, O62247, A0A1Q1NL17, Q8WPG5	P84555
Structure	Full-sequence	D3GGZ8, O62247, A0A1Q1NL17, Q8WPG5	NA
Sequence	Single-domain	D3GGZ8, O62247, A0A1Q1NL17, Q8WPG5, Q11101	P84555
Sequence	Full-sequence	D3GGZ8, O62247, A0A1Q1NL17, Q8WPG5	NA
		Including InterPro-based Evaluation	
Structure	Single-domain	Q8MVZ2, Q8MVZ3, P36235, D3GGZ8	P84556, P85040, P85039, PODJ63, PODM47, P0DV02, P83604, P0DV07
Structure	Full-sequence	Q8MVZ2, Q8MVZ3, P36235, D3GGZ8	P84556, P85040, P85039, PODJ63
Sequence	Single-domain	Q8MVZ2, Q8MVZ3, P36235, D3GGZ8	P84556, P85040, P85039, PODJ63
Sequence	Full-sequence	Q8MVZ2, Q8MVZ3, P36235, D3GGZ8	P84556, P85040, P85039, PODJ63

Table 2: Confusion matrices and average MCC values for each model and E-value type, divided by annotation method.

Model	E-value Type	TP	TN	FP	FN	Avg MCC
		Only Pfam-based Evaluation				
Structure	Single-domain	364	572571	1	4	0.993182680
Structure	Full-sequence	364	572572	0	4	0.994546894
Sequence	Single-domain	363	572571	1	5	0.993180816
Sequence	Full-sequence	364	572572	0	4	0.994546894
		Including InterPro-based Evaluation				
Structure	Single-domain	376	572552	8	4	0.984334276
Structure	Full-sequence	376	572556	4	4	0.989064052
Sequence	Single-domain	376	572556	4	4	0.989064052
Sequence	Full-sequence	376	572556	4	4	0.989064052