Supplementary: Context awareness and embedding for biomedical event extraction

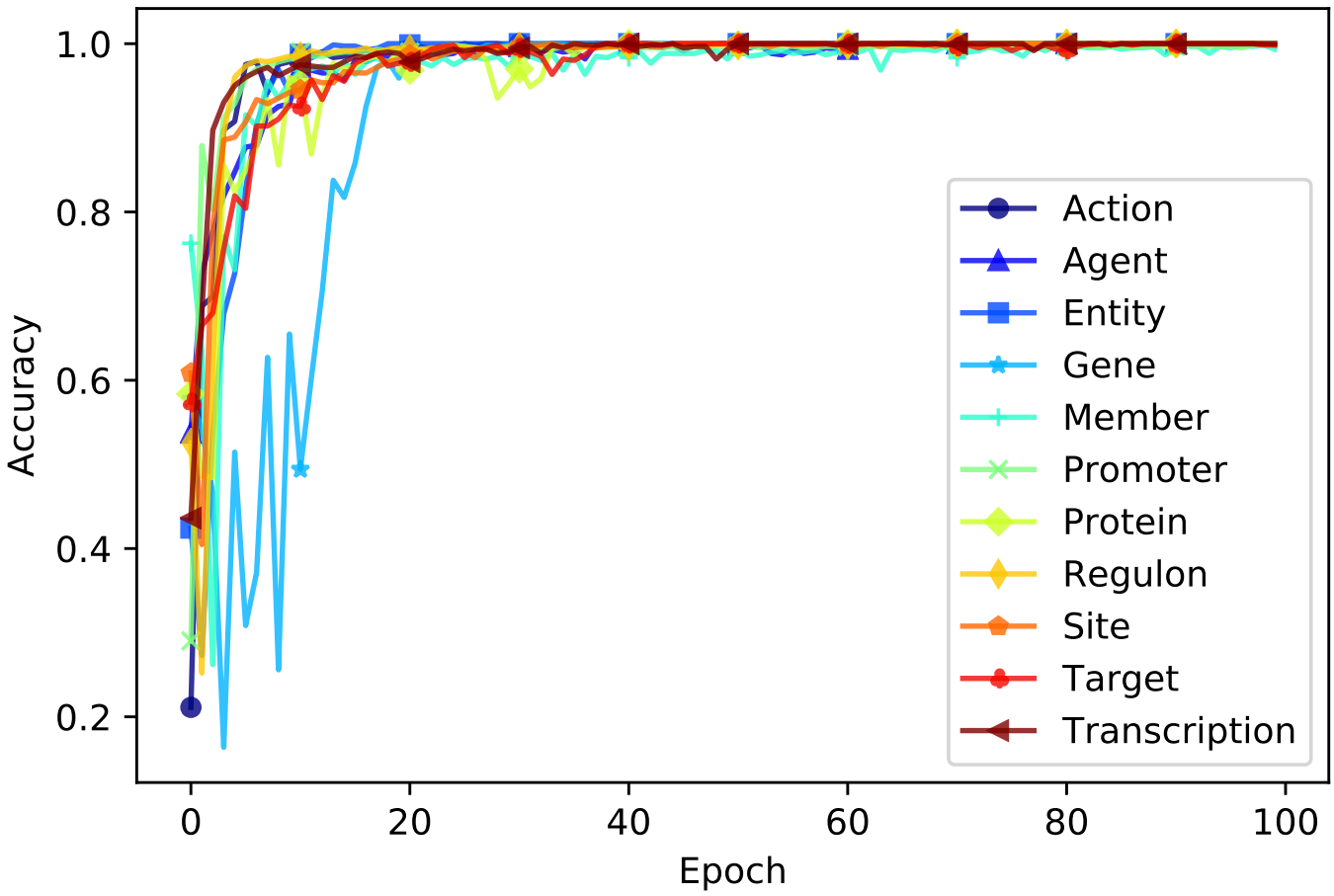


Fig S1: Model Accuracy of VecEntNet on BioNLPST-BGI dataset over training and testing process.

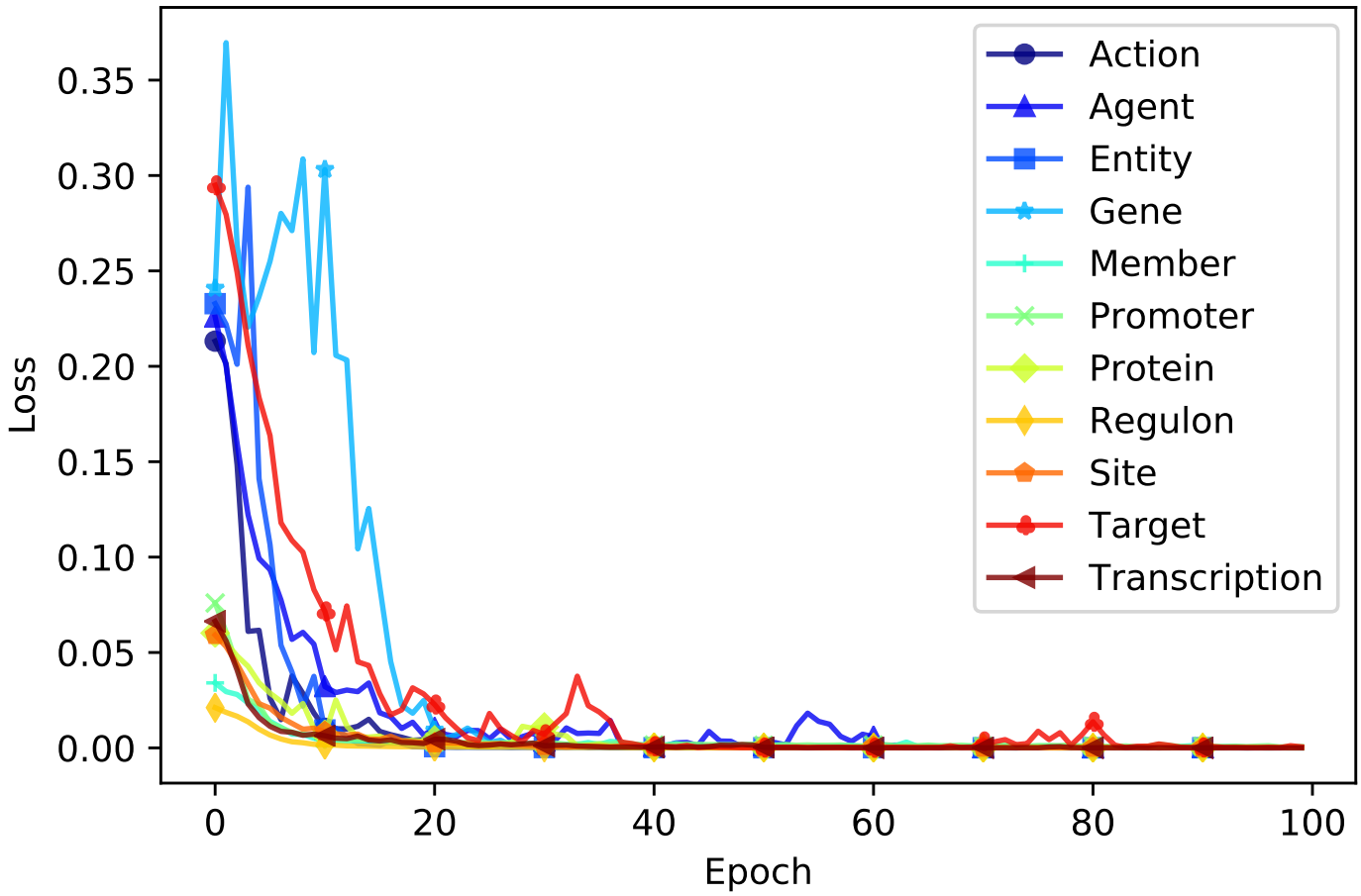


Fig S2: Model Loss of VecEntNet on BioNLPST-BGI dataset over training and testing process.

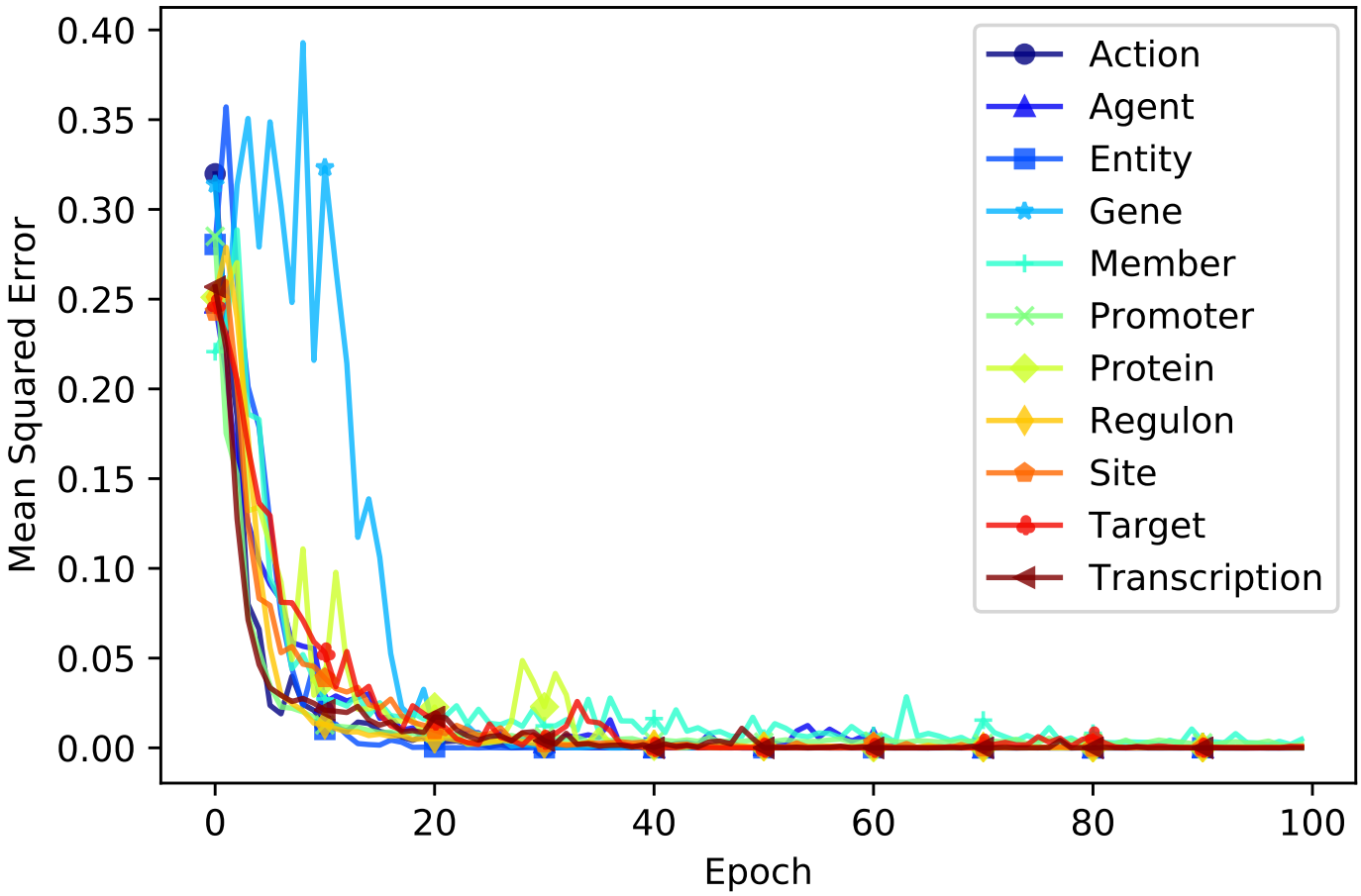


Fig S3: Mean Squared Error of VecEntNet on BioNLPST-BGI dataset over training and testing process.

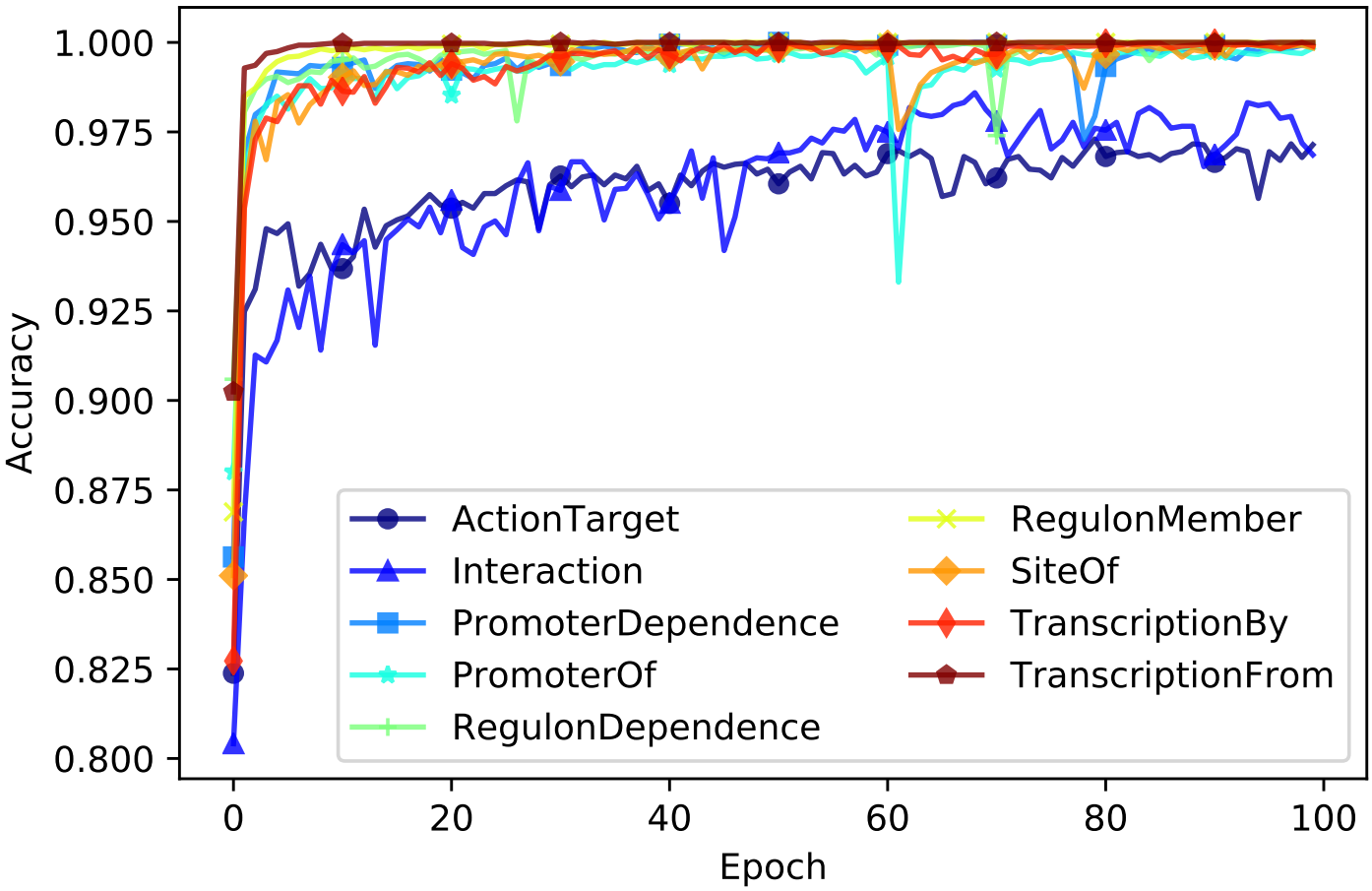


Fig S4: Model Accuracy of VeComNet on BioNLPST-BGI dataset over training and testing process.

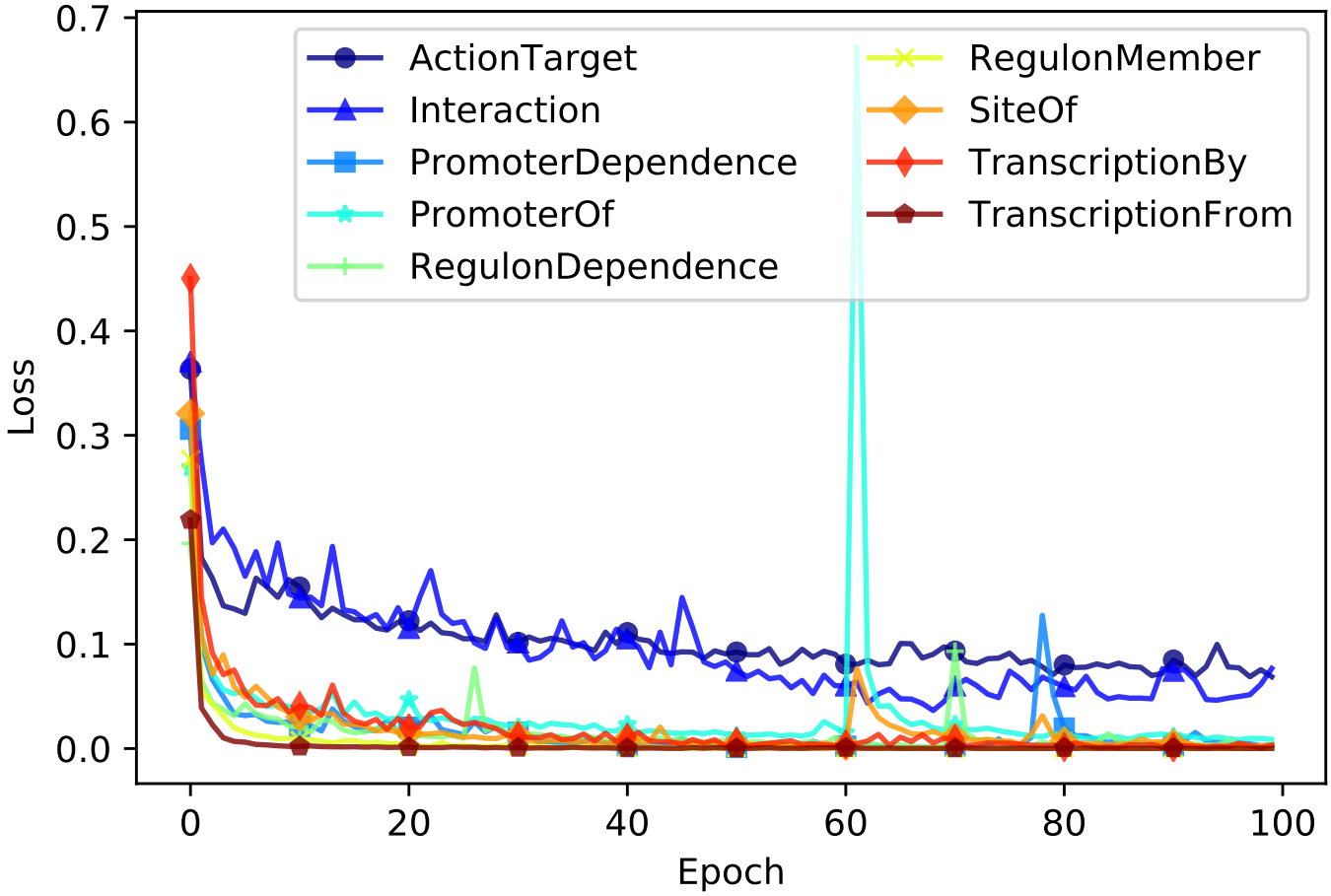


Fig S5: Model Loss of VeComNet on BioNLPST-BGI dataset over training and testing process.

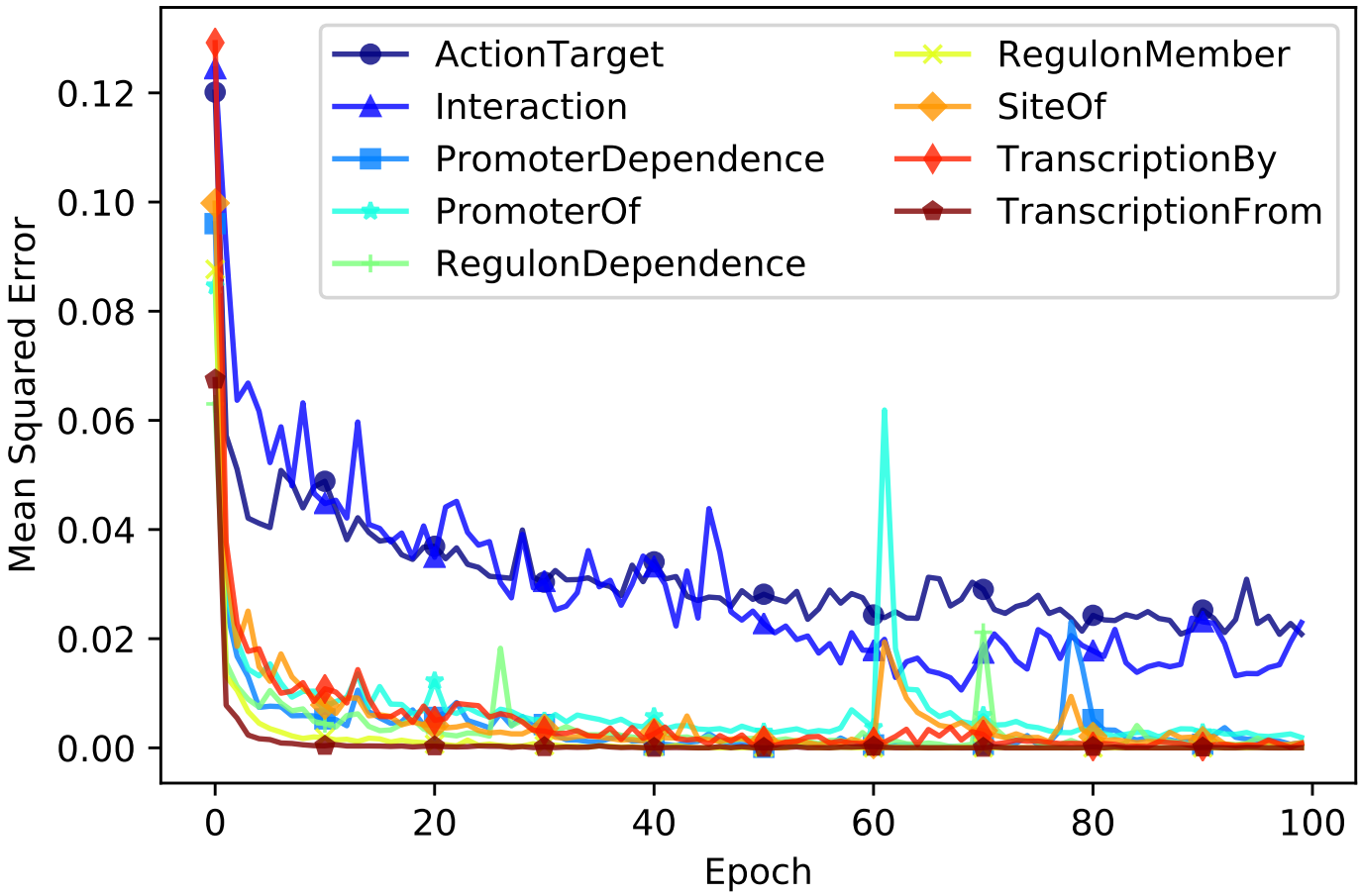


Fig S6: Mean Squared Error of VeComNet on BioNLPST-BGI dataset over training and testing process.

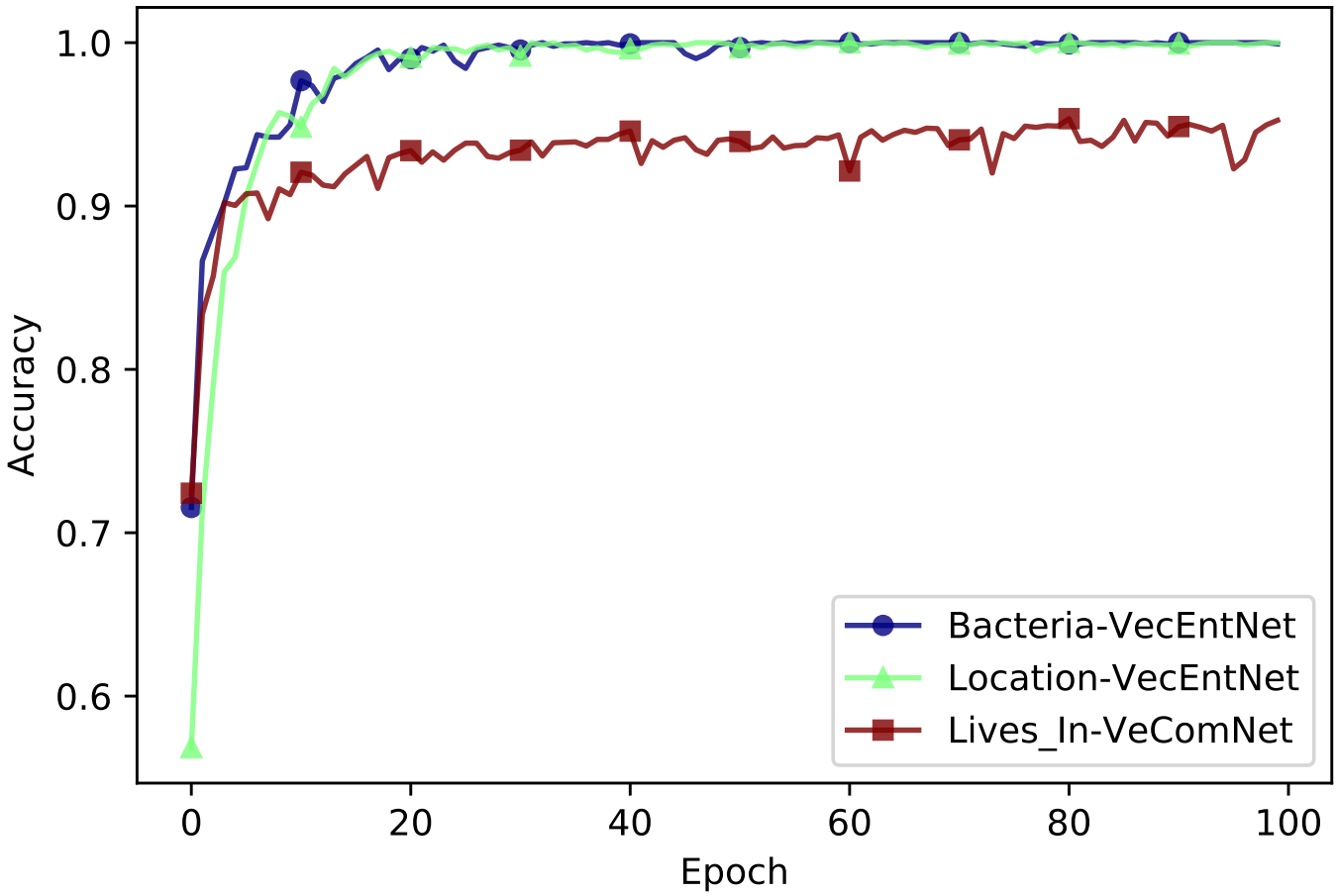


Fig S7: Model Accuracy of VecEntNet and VeComNet on BioNLPST-BB dataset over training and testing process.

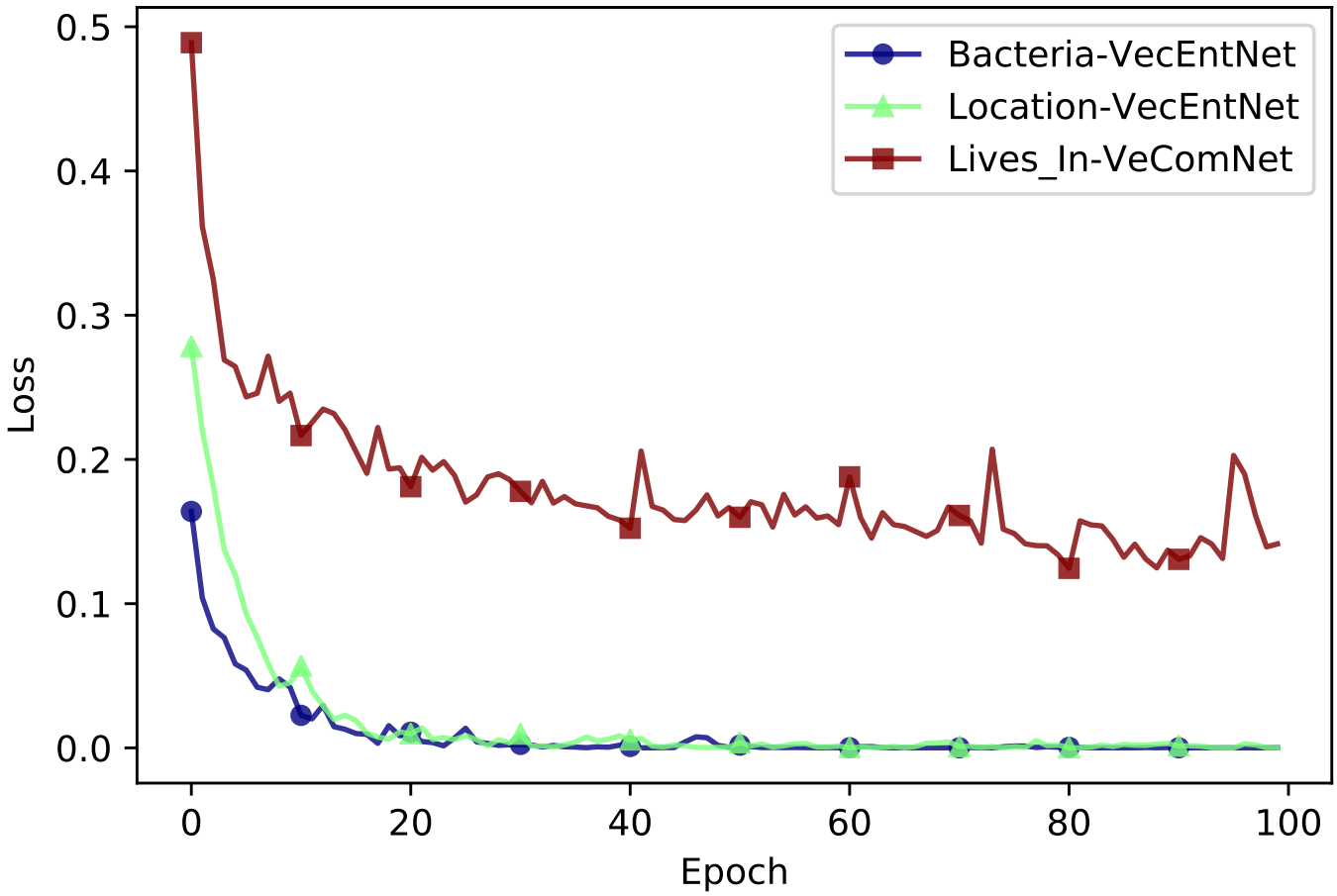


Fig S8: Model Loss of VecEntNet and VeComNet on BioNLPST-BB dataset over training and testing process.

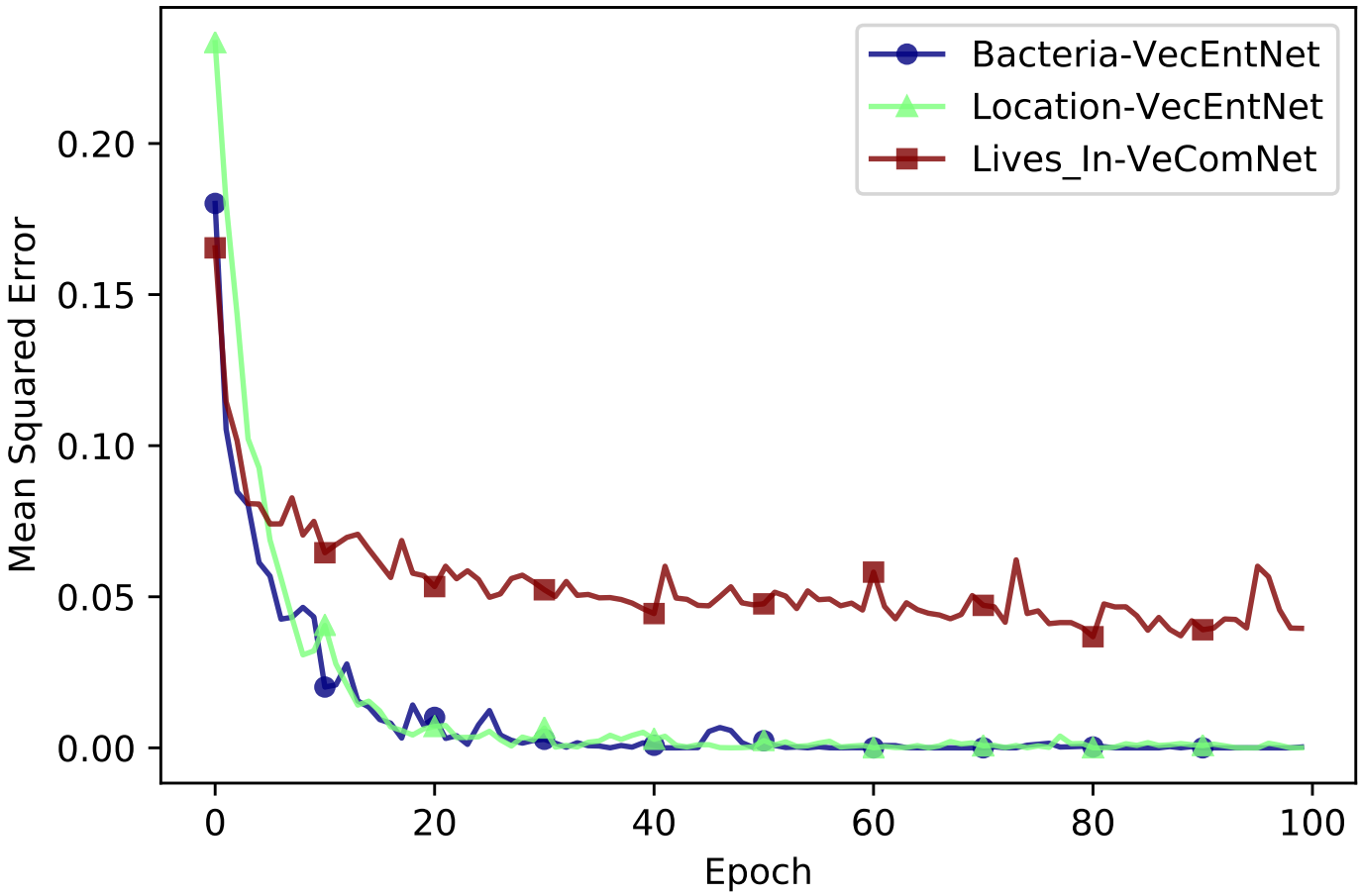


Fig S9: Mean Squared Error of VecEntNet and VeComNet on BioNLPST-BB dataset over training and testing process.

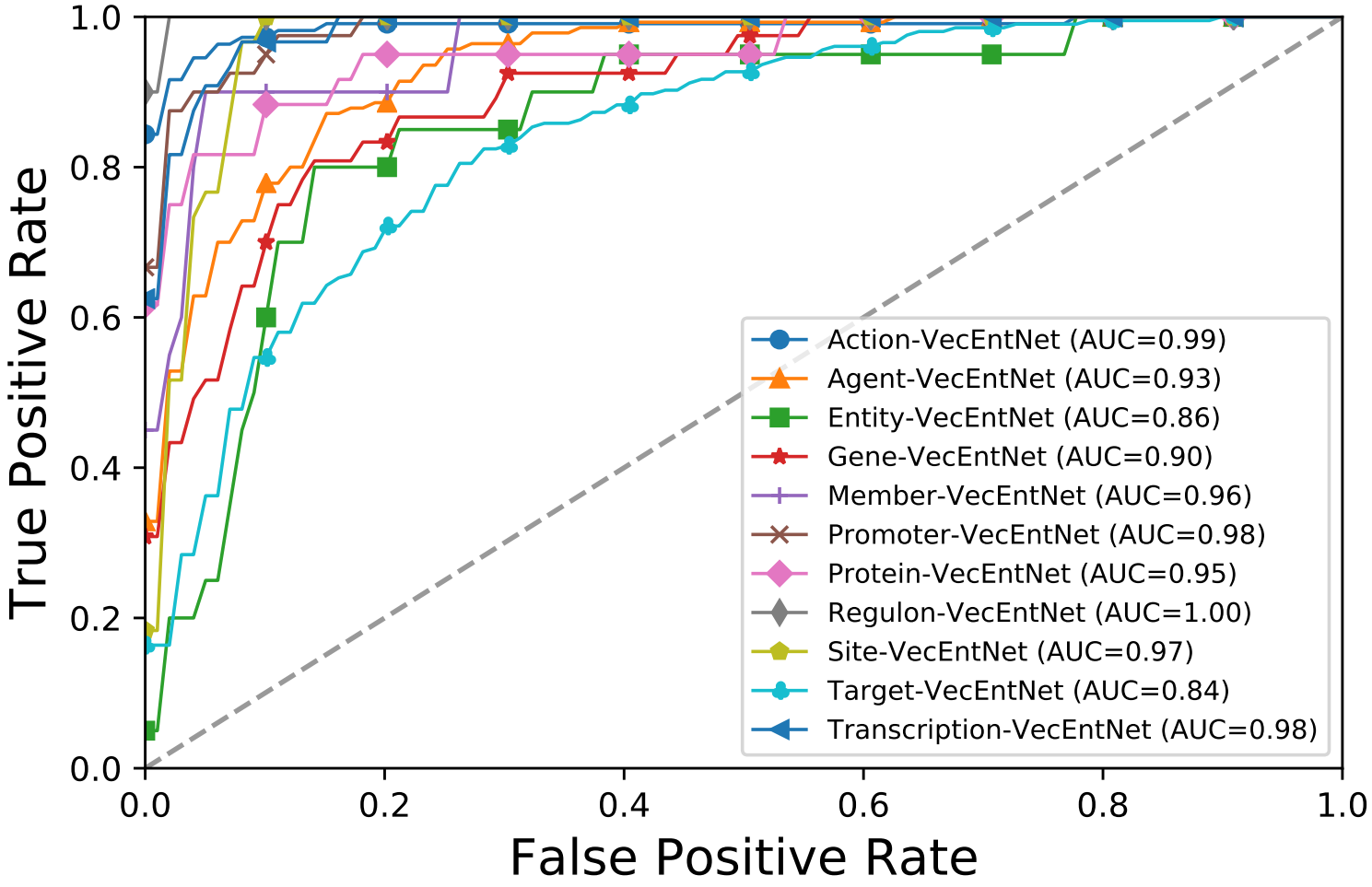


Fig S10: Micro-average ROC curves of VecEntNet on BioNLPST-BGI dataset.

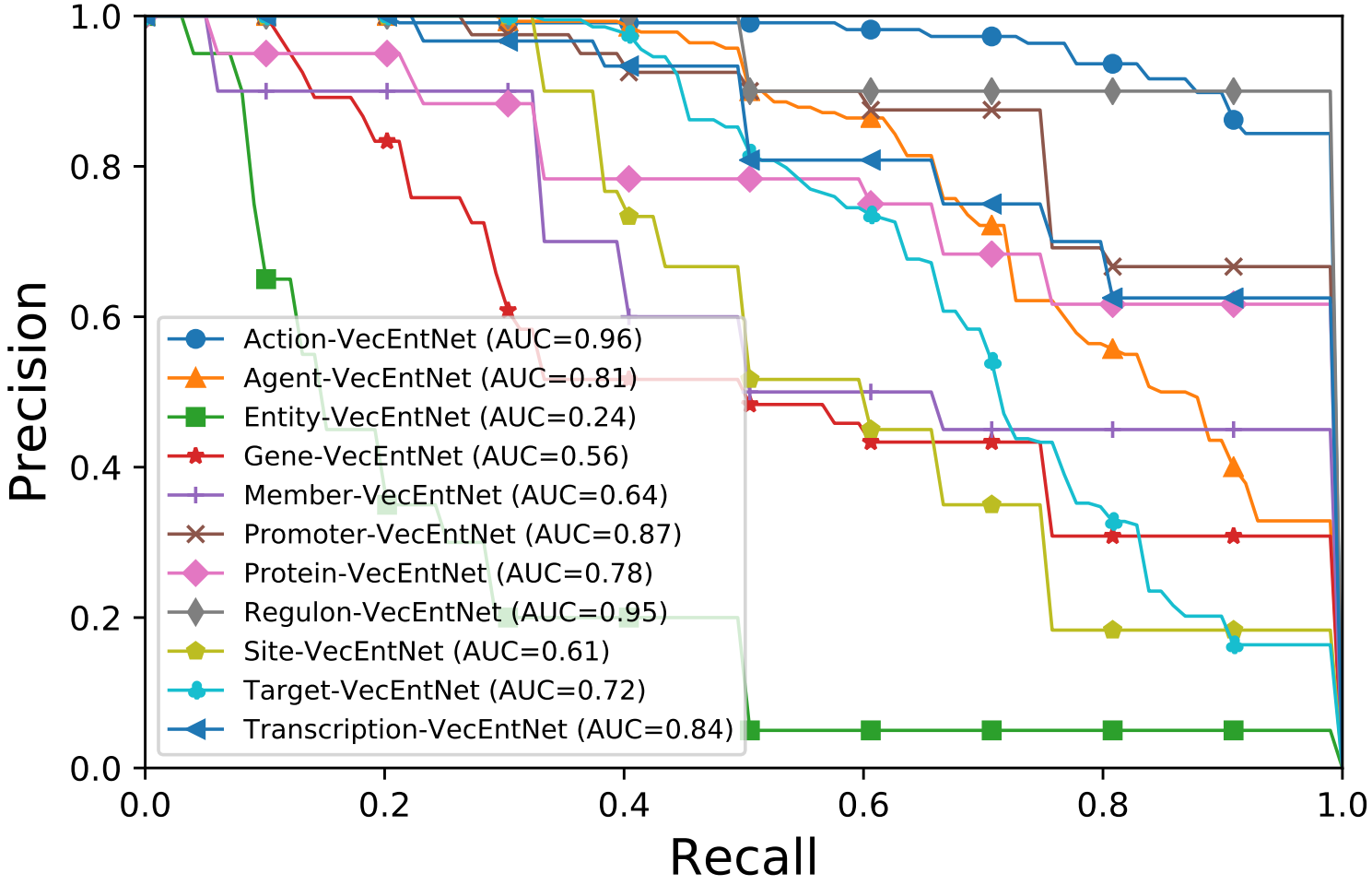


Fig S11: Micro-average PRC curves of VecEntNet on BioNLPST-BGI dataset.

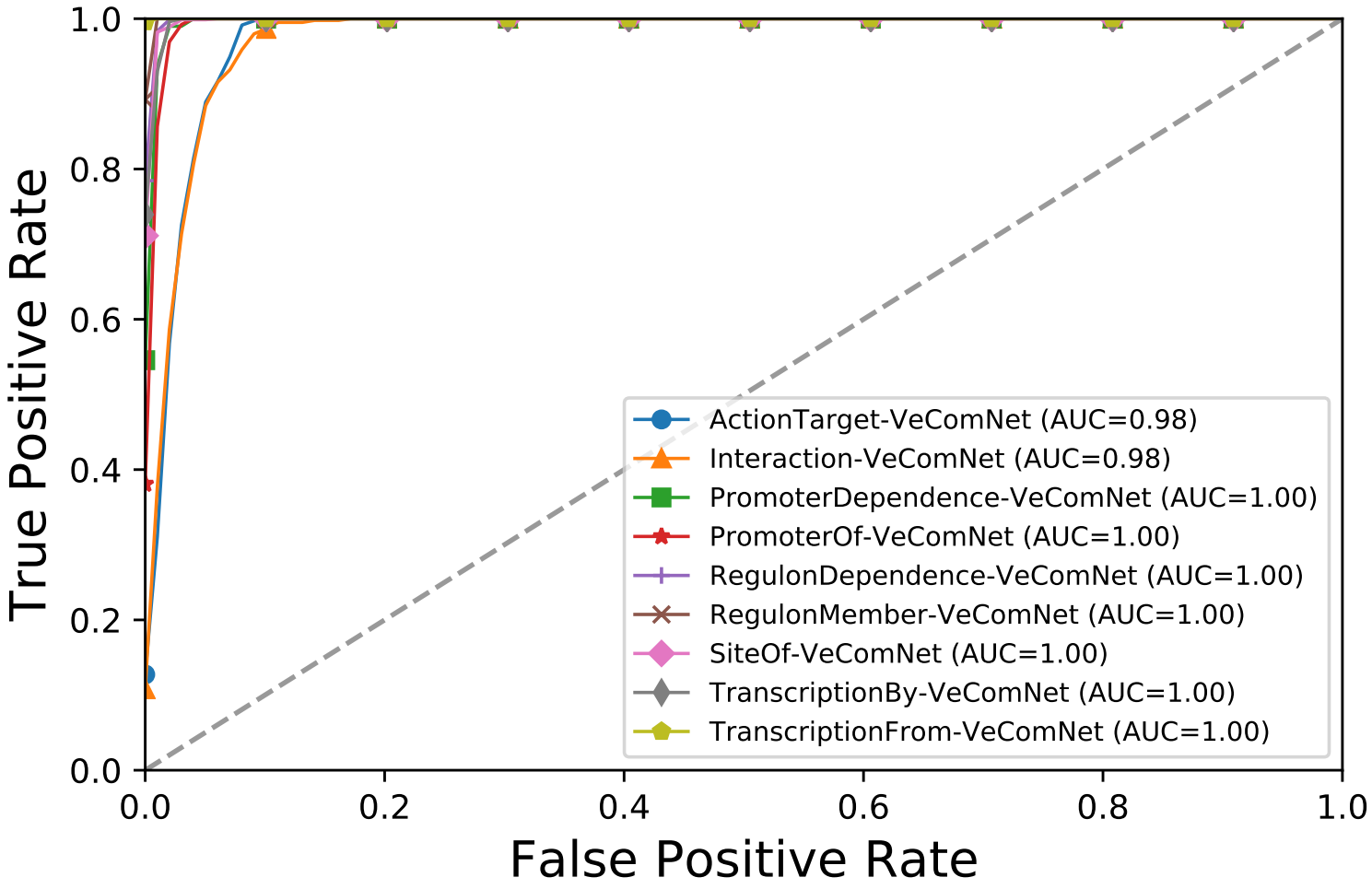


Fig S12: Micro-average ROC curves of VeComNet on BioNLPST-BGI dataset.

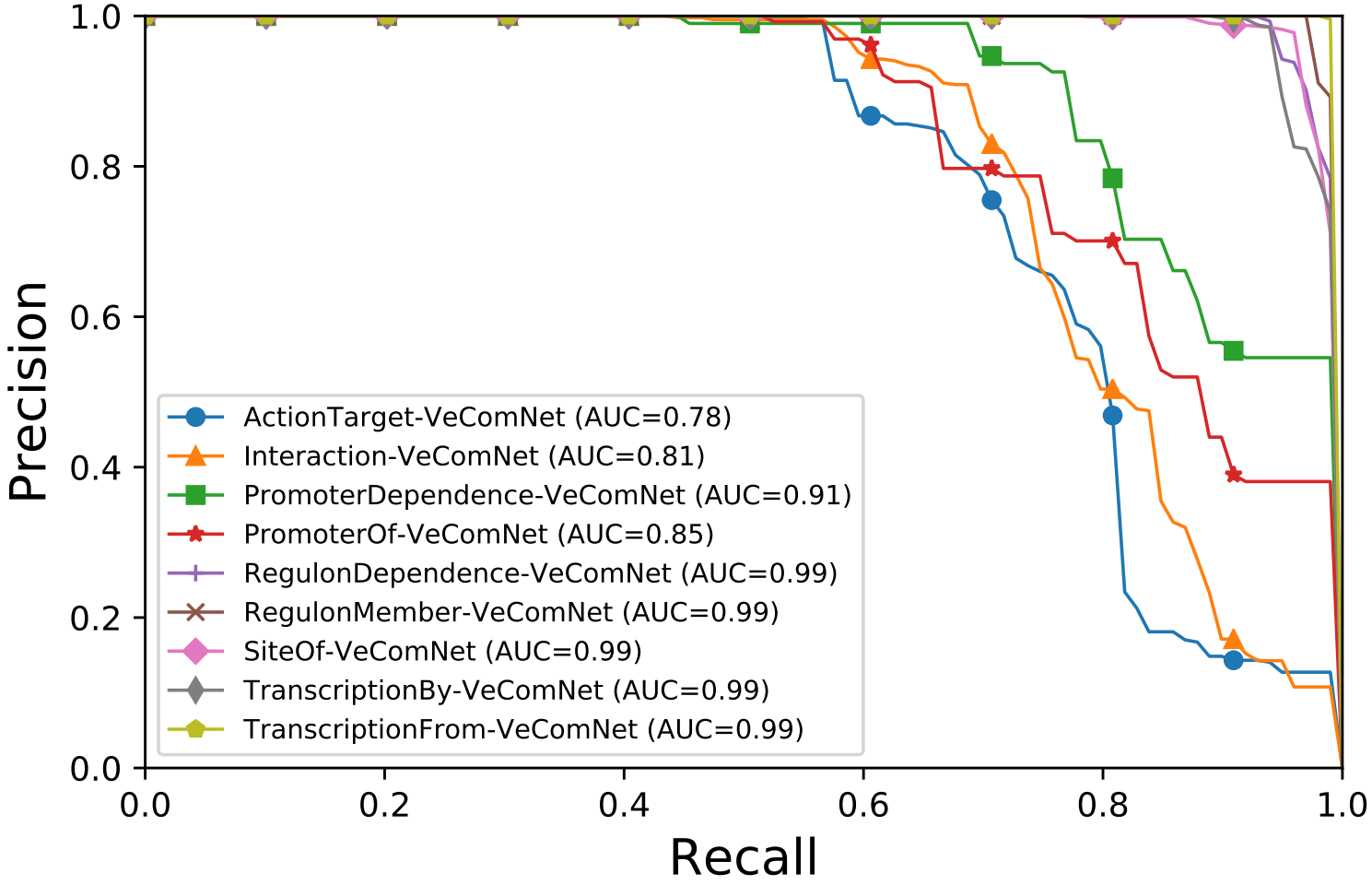


Fig S13: Micro-average PRC curves of VeComNet on BioNLPST-BGI dataset.

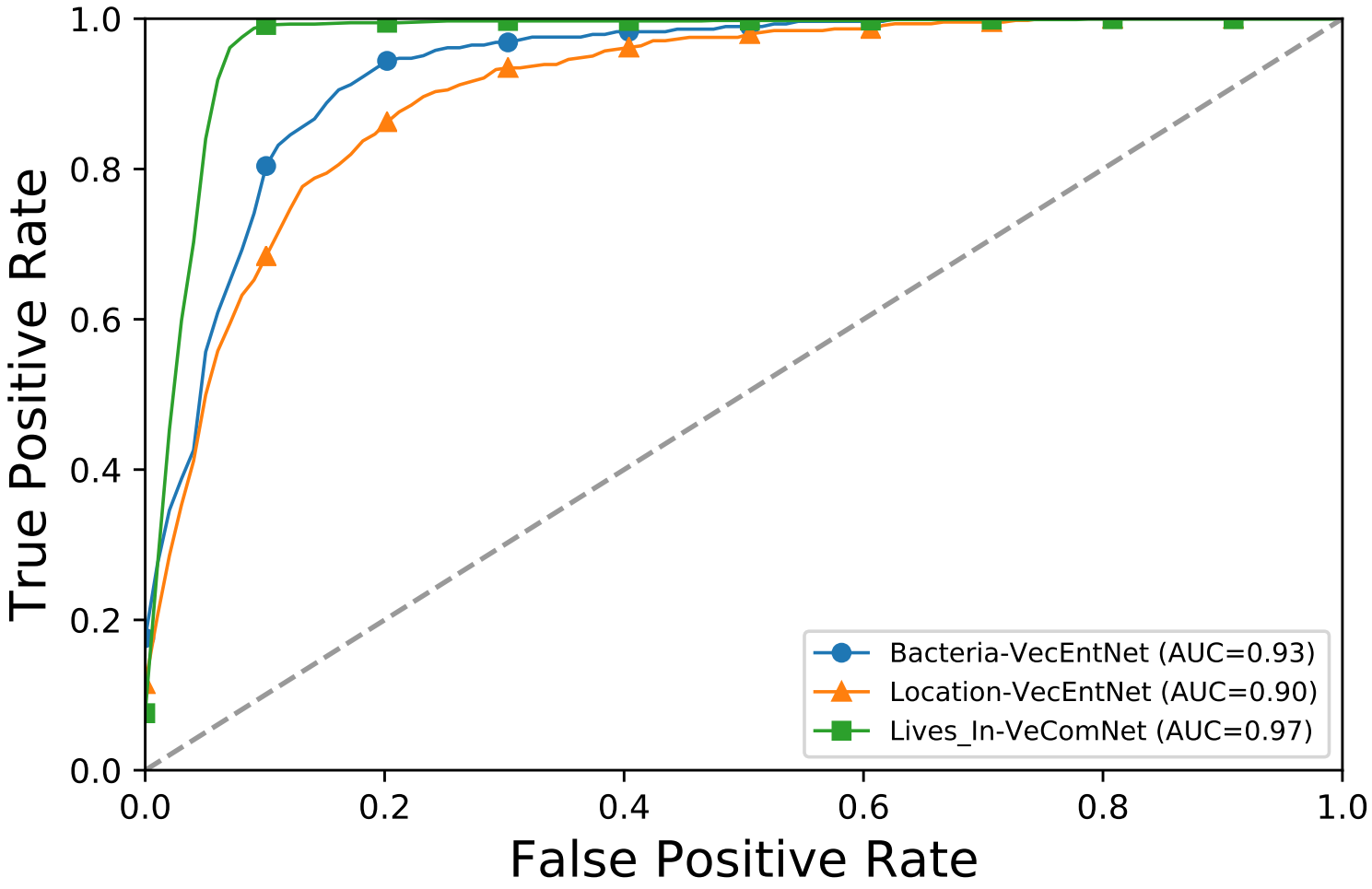


Fig S14: Micro-average ROC curves of VecEntNet and VeComNet on BioNLPST-BB dataset.

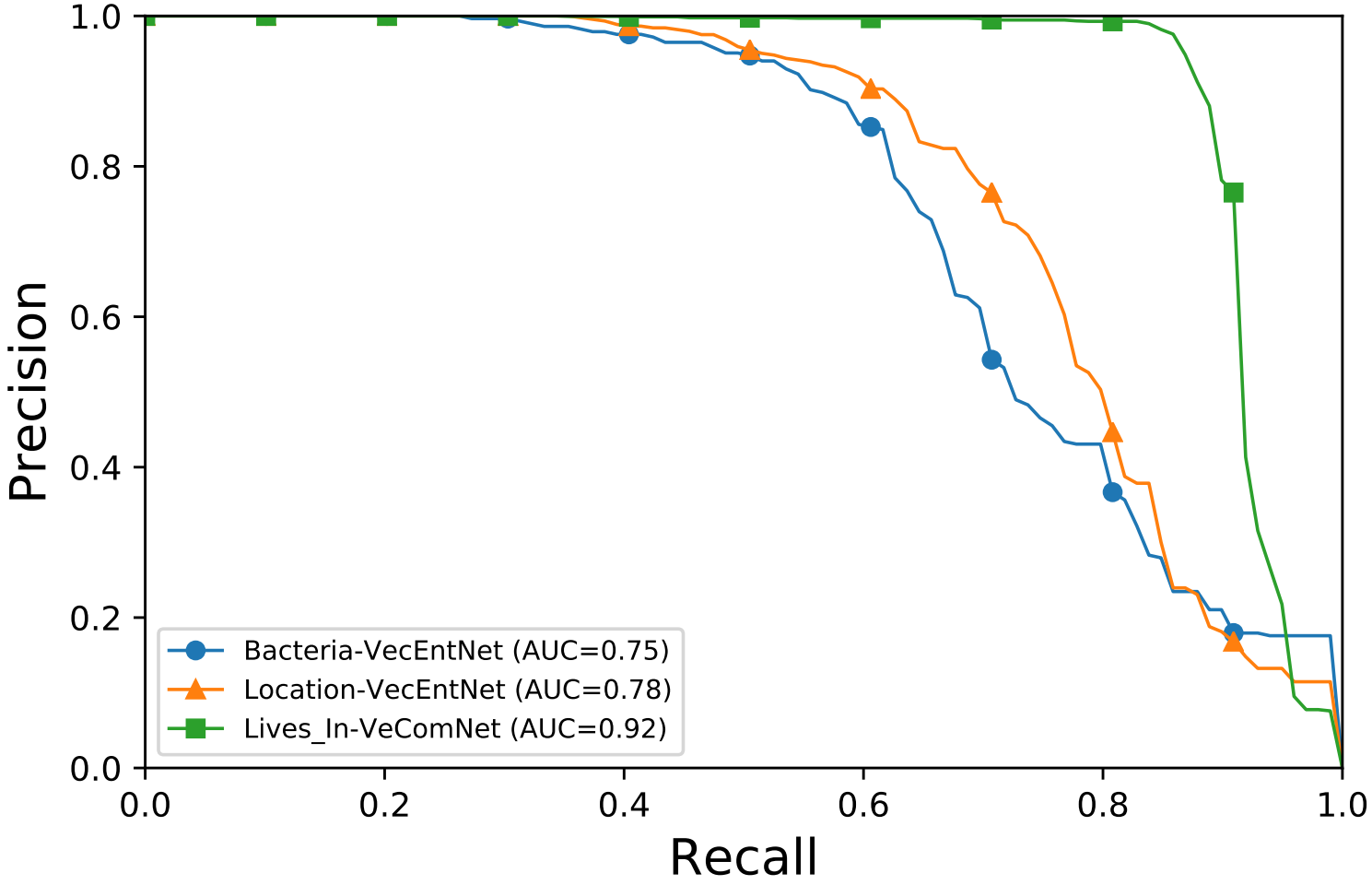


Fig S15: Micro-average PRC curves of VecEntNet and VeComNet on BioNLPST-BB dataset.

Table S1: Hyper-parameters used in VecEntNet

VecEntNet	
context window size	10
LSTM hidden/output units	128
MLP input units	256
MLP hidden units	128
Batch size	32
Epoch	10

Table S2: Hyper-parameters used in VeComNet

VeComNet	
MLP input units	128
MLP hidden units	64
Batch size	32
Epoch	10

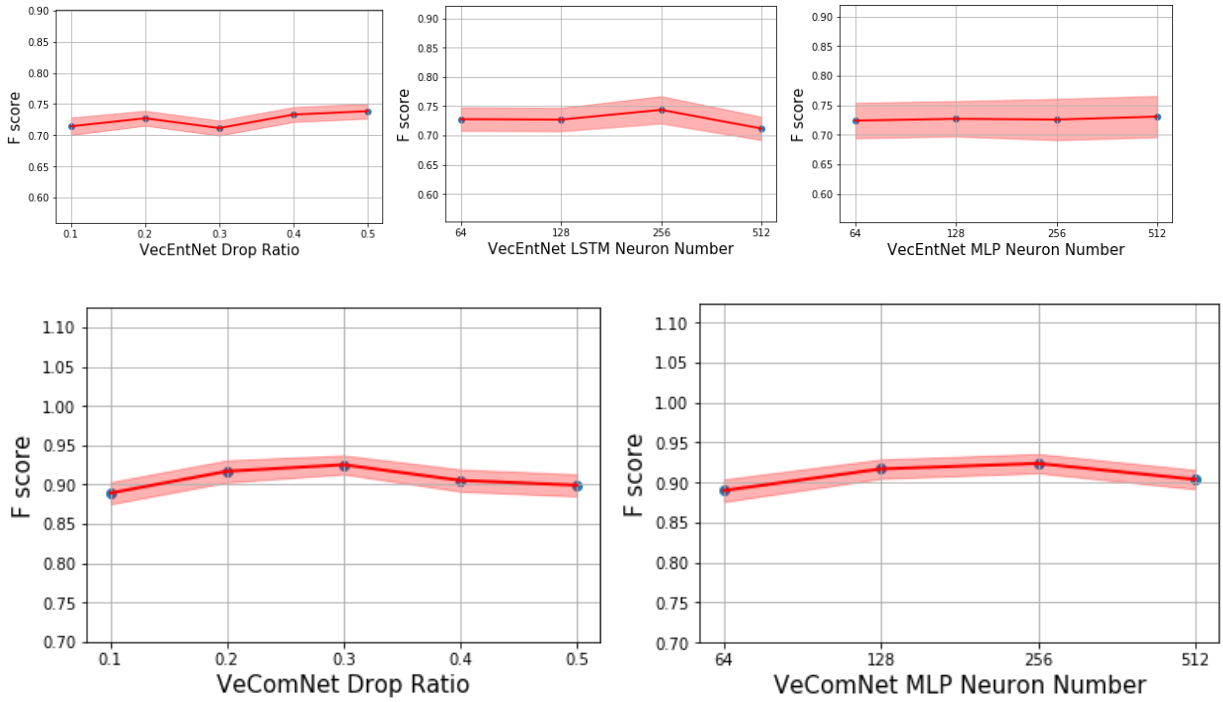
Table S3: Performance of VecEntNet on BioNLPST-BGI dataset with regards to different argument types under 10-fold cross validation

	Action	Agent	Entity	Gene	Member	Promoter	Protein	Regulon	Site	Target	Transcription
# of Positive Training Samples	108	140	15	36	15	38	29	10	29	185	31
Accuracy	0.97	0.83	0.86	0.91	0.97	0.97	0.92	0.99	0.94	0.77	0.97
Precision	0.92	0.56	0.09	0.32	0.48	0.70	0.52	0.68	0.45	0.61	0.60
Recall	0.93	0.93	0.65	0.55	0.80	0.90	0.72	1.00	0.80	0.73	0.98
F score	0.92	0.70	0.15	0.37	0.54	0.77	0.50	0.77	0.52	0.66	0.75
Train time (s)	603.73	696.15	512.58	710.31	677.00	469.79	595.32	391.12	591.96	660.29	610.50
Test time (s)	1.65	1.59	1.56	1.79	2.15	0.89	1.02	0.82	0.94	1.87	1.09

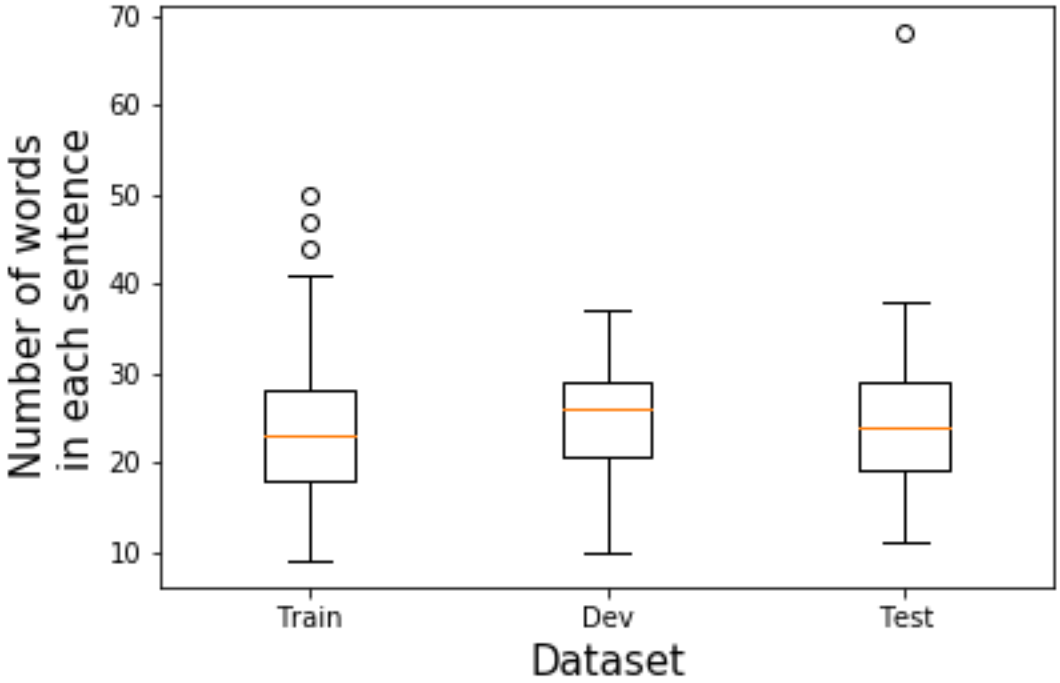
Table S4: Performance of VeComNet on BioNLPST-BGI dataset with regards to different event types under 10-fold cross validation

	Action-Target	Interaction	Promoter-Dependence	PromoterOf	Regulon-Dependence	Regulon-Member	SiteOf	Transcrip-tionBy	Transcrip-tionFrom
# of Positive Training Samples	108	126	32	36	11	15	17	25	14
Accuracy	0.93	0.91	0.98	0.98	0.99	0.99	0.98	0.97	0.99
Precision	0.70	0.73	0.82	0.79	0.97	0.99	0.95	0.60	0.99
Recall	0.91	0.82	0.84	0.78	0.99	0.99	0.98	0.98	0.99
F score	0.79	0.77	0.82	0.76	0.98	0.99	0.97	0.75	0.99
Train time (s)	4.99	5.04	5.00	5.03	5.55	5.67	6.18	5.69	5.78
Test time (s)	0.15	0.15	0.16	0.16	0.16	0.17	0.16	0.16	0.14

Parameter Analysis



Word count in each sentence



Spearman's Rank Correlation

BioNLPST-BB

Argument 0	P-value	Argument 1	P-value	Event	P-value
0.62	1.21E-12	0.60	1.60E-12	0.88	2.07E-136

BioNLPSTBGI

Argument 0	P-value	Argument 1	P-value	Argument 2	P-value	Argument 3	P-value
0.91	2.79E-13	0.62	1.21E-03	0.79	2.15E-02	0.66	1.39E-02
Argument 4	P-value	Argument 5	P-value	Argument 6	P-value	Argument 7	P-value
0.68	3.02E-02	0.77	1.58E-07	0.63	2.53E-03	0.80	2.74E-06
Argument 8	P-value	Argument 9	P-value	Argument 10	P-value		
0.75	4.22E-03	0.70	2.68E-02	0.67	1.39E-03		
Event 0	P-value	Event 1	P-value	Event 2	P-value	Event 3	P-value
0.77	1.01E-50	0.72	1.01E-50	0.74	1.20E-43	0.82	1.66E-44
Event 4	P-value	Event 5	P-value	Event 6	P-value	Event 7	P-value
0.77	1.62E-26	0.97	5.82E-191	0.99	5.37E-262	0.96	5.15E-191
Event 8	P-value	Event 10	P-value				
0.97	6.69E-191	1.00	4.18E-31				

Hierarchical relationship among different event types, argument types and entity types

Dataset: BioNLPST-BGI

Event	Argument	Entity
ActionTarget	Action	Action Expression Transcription
	Target	Gene GeneComplex GeneFamily Promoter Protein Regulon
Interaction	Agent	Gene GeneProduct PolymeraseComplex Protein
	Target	Gene GeneComplex GeneProduct PolymeraseComplex Protein
PromoterDependence	Agent	Promoter
	Target	Gene Protein
PromoterOf	Agent	Promoter
	Target	Gene GeneComplex GeneFamily Protein
RegulonDependence	Agent	Regulon
	Target	Gene Protein
RegulonMember	Member	Gene GeneProduct PolymeraseComplex Protein
	Regulon	Regulon
SiteOf	Site	Gene GeneComplex Promoter
	Target	Site
TranscriptionBy	Polymerase	PolymeraseComplex Protein
	Transcription	Transcription
TranscriptionFrom	Process	Expression Transcription
	Promoter	Promoter Site

Hierarchical relationship among different event types, argument types and entity types

Dataset: BioNLPST-BB

Event	Argument	Entity
Lives_In	Bacteria	Bacterium
	Location	Habitat
		Geographical