

A

IOWA STATE UNIVERSITY

RNA-Protein Interaction Prediction (RPSeq)
Dobts and Honavar laboratories

Home [INPUT SEQUENCES](#)

About
RPSeq is a web-based tool for predicting RNA-protein interactions. It takes RNA and protein sequences as input and outputs a probability score indicating the likelihood of interaction.

Datasets
The RPSeq database contains over 100,000 entries from various sources, including PDB, UniProt, and other protein databases.

Related Links
[RPSeq Home](#) | [UniProt](#) | [PDB](#) | [Protein Data Bank](#) | [Protein Interaction Databases](#)

References
Guttinger S, Hahn R, Koller T, et al. (2011) A general method for the identification of RNA-protein interactions. *Nature* 473: 332-336.

Funding
NIH, NSF, and Iowa State University.

Contact Us
Amit Dobts, PhD | amit.dobts@iastate.edu | <http://rpseq.iastate.edu>

Links:

- [Dobts Lab, Science](#)
- [Honavar lab, Bioinformatics and Computational Biology](#)
- [Iowa State University](#)
- [Center for Computational Intelligence, Learning & Discovery](#)
- [Department of Genetics, Development and Cell Biology](#)

   

Input Sequences

Protein:
MDMTDGCQFPSPSEFYEGSCIPSPEDEFQDFQEPRVAAGFAHKAELOQGSDEEHVRAPTG
HHQAGHCLMWACKACKRKSTTMDRRKAATMRERRRLKKVNAQAFETLKRCTTNPQRLPK
VEILRNARYIESLQELLREQVENYSLPGQSCSEPTSTSNCSDGMPECNSPVWSRKNS
SFDSIYCPDVSNACADKSSVSSLDCLSIVDRITSTEPSELALQDTASLSPATSANSQ
ATPGPSSSRLIYHV

RNA:
GGGGGCAGCUCCGCCCUCCCCGAUGGCUGUAUUAACGCAGCGCGAGCGGGCUCGGUC
UUUCGGUCCCGGCCUJGGUGAGGUCCUCGUGGGAGGGUGUCAGGAUGGUCCCCGGU
CCCCUGAGCUUUCGGCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
AGCAGUUCUCAGGG
CAAAGGUCUCUGGGCCUGCCGACCCGCCCCCACUGCCCAGCCUGCCUGGGACACAAGCG
CGAUCCACCCGAGCCGGUCCCCCACUGCCCAGCCUGCCUGGGACACAAGCG
GUACGUCCCCGGCAUUGAUCGGAGGGGGCCUGGCCACGUGGGCGAGAAAGACGGAGGU
CAGGCCUGUCUUUQUGUGAGCUGGCCGGGGCCUCCCCAGUGCAGCCCUUCCUGUGGCCUU
AGGGUUGUGCCAGUACUUUAAGAGGUACACAGAAGGGUAAGGAACUCUCCAUAAA
CCAGAGAGGAAACACUACCUUCUUCUUGGAUCCUGUCUGGAGUACAAAUGAAC
GCUUCAUAAAACUUGAUGUGCCAUUJGGGUUCUGGCCAUUGGUAGGUGUAGAU
AGGAUUAUUAAGAAAAGAAAACGUAGGCCUACCCUGGUUUAUUUUGGUUGGAGGAGAA
CAGAAGACAGAAUACAUUJGGGUUUCUAGAUGAGUAGGUAGGAGGAGGU
AUAGUUGAGUAUGUUCCACUUUAGUGAAACAUUUUUAUUCUCAUAGGUAGAAA
AAGAAACCCGUGACUGAAUGGCUCCAAGGUGAAGUACUCUUGCAUGUAGUGACCAG
GCCUGUACUUCACUUUAGGGAUAGUAUCGAAGCUGAAUAGGUAGUGAGGGCAUG
AGUUUCCUUGAAGGGCCUGGUACUGAUUACCCUGAAGGUAGGUAGGUAGGU
CAAUGUACUAGGAUCCUGGUCCCCACCUUCAUGUCCACUGGCCUAGAACAGAU
GCAUGGCUAGAACAGAUCCUGGUCCCCACCUUCAUGUCCACUGGCCUAGAACAGAU
AUGGUACUGGUUACCCUGGUUACUGGUAGGUAGGUAGGUAGGUAGGU
GAGGUGUGUUAACCUGUUAUUGUUGCCUAGAGUAUUUCCAUAAUCUUUAAG
CUUCCUGUGUCAUUCU

Interaction probabilities

Prediction using RF classifier: 0.9
Prediction using SVM classifier: 0.94

What do these probabilities mean?
Interaction probabilities generated by RPSeq range from 0 to 1. In performance evaluation experiments, predictors with probabilities > 0.5 were considered "positive," i.e., indicating that the corresponding RNA and protein are likely to interact. Using this threshold, accuracies of the classifiers ranged from 87 - 90% in cross-validation evaluation experiments on benchmark datasets. When classifiers were tested on independent (blind) datasets of RPIs, accuracies of the classifiers ranged from 57 – 99%.

Please see [About/FAQs](#) for additional details.

IOWA STATE UNIVERSITY
Biosciences & Biofuels
RPSeq v1.0 | Last updated: 03/01/2011
Data Loc: (515)-294-4001 | <http://rpseq.iastate.edu>
Copyright © 2011 Dobts Lab. All rights reserved.

B**Input Sequences****Protein:**

MDMTDGCQFPSPSEFYEGSCIPSPEDEFQDFQEPRVAAGFAHKAELOQGSDEEHVRAPTG
HHQAGHCLMWACKACKRKSTTMDRRKAATMRERRRLKKVNAQAFETLKRCTTNPQRLPK
VEILRNARYIESLQELLREQVENYSLPGQSCSEPTSTSNCSDGMPECNSPVWSRKNS
SFDSIYCPDVSNACADKSSVSSLDCLSIVDRITSTEPSELALQDTASLSPATSANSQ
ATPGPSSSRLIYHV

RNA:

GGGGGCAGCUCCGCCCUCCCCGAUGGCUGUAUUAACGCAGCGCGAGCGGGCUCGGUC
UUUCGGUCCCGGCCUJGGUGAGGUCCUCGUGGGAGGGUGUCAGGAUGGUCCCCGGU
CCCCUGAGCUUUCGGCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
AGCAGUUCUCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
CAAAGGUCUCUGGGCCUGCCGACCCGCCCCCACUGCCCAGCCUGCCUGGGACACAAGCG
CGAUCCACCCGAGCCGGUCCCCCACUGCCCAGCCUGCCUGGGACACAAGCG
GUACGUCCCCGGCAUUGAUCGGAGGGGGCCUGGCCACGUGGGCGAGAAAGACGGAGGU
CAGGCCUGUCUUUQUGUGAGCUGGCCGGGGCCUCCCCAGUGCAGCCCUUCCUGUGGCCUU
AGGGUUGUGCCAGUACUUUAAGAGGUACACAGAAGGGUAAGGAACUCUCCAUAAA
CCAGAGAGGAAACACUACCUUCUUCUUGGAUCCUGUCUGGAGUACAAAUGAAC
GCUUCAUAAAACUUGAUGUGCCAUUJGGGUUCUGGCCAUUGGUAGGUGUAGAU
AGGAUUAUUAAGAAAAGAAAACGUAGGCCUACCCUGGUUUAUUUUGGUUGGAGGAGAA
CAGAAGACAGAAUACAUUJGGGUUUCUAGAUGAGUAGGUAGGAGGAGGU
AUAGUUGAGUAUGUUCCACUUUAGUGAAACAUUUUUAUUCUCAUAGGUAGAAA
AAGAAACCCGUGACUGAAUGGCUCCAAGGUGAAGUACUCUUGCAUGUAGUGACCAG
GCCUGUACUUCACUUUAGGGAUAGUAUCGAAGCUGAAUAGGUAGUGAGGGCAUG
AGUUUCCUUGAAGGGCCUGGUACUGAUUACCCUGAAGGUAGGUAGGUAGGU
CAAUGUACUAGGAUCCUGGUCCCCACCUUCAUGUCCACUGGCCUAGAACAGAU
GCAUGGCUAGAACAGAUCCUGGUCCCCACCUUCAUGUCCACUGGCCUAGAACAGAU
AUGGUACUGGUUACCCUGGUUACUGGUAGGUAGGUAGGUAGGUAGGU
GAGGUGUGUUAACCUGUUAUUGUUGCCUAGAGUAUUUCCAUAAUCUUUAAG
CUUCCUGUGUCAUUCU

Interaction probabilities

Prediction using RF classifier: 0.9

Prediction using SVM classifier: 0.94

What do these probabilities mean?

Interaction probabilities generated by RPSeq range from 0 to 1. In performance evaluation experiments, predictions with probabilities > 0.5 were considered "positive," i.e., indicating that the corresponding RNA and protein are likely to interact. Using this threshold, accuracies of the classifiers ranged from 87 - 90% in cross-validation evaluation experiments on benchmark datasets. When classifiers were tested on independent (blind) datasets of RPIs, accuracies of the classifiers ranged from 57 – 99%.