A **B**

RNA-Protein Interaction Prediction (RPISeq)
Dobbs and Honavar Laboratories

Input Sequences

Protein:
MDMTDGCQFSPSEYFYEGSCIPSPEDEFGDQFEPRVAAFGAHKAELQGSDDEEHVRAPTG
HHQAGHCLMWACKACKRKSTTMDRRKAATMRERRRLKKVNQAFETLKRCTTTNPNQRLPK
VEILRNAIRYIESLQELLREQVENYYSLPGQSCSEPTSPTSNCSDGMPECNSPVWSRKNS
SFDSIYCPDVSNACAADKSSVSSLDCLSSIVDRITSTEPSELALQDTASLSPATSANSQP
ATPGPSSSRLIYHVL

RNA:
GGGGGCAGCUCCGCCCCUCUCCCGGAUGGCUGUAUUAACGCAGCGCGCAGCGGCUCGGUC
UUUCGGUCCCGGCCUUGGUGGUGAGUGUCCUCGUGGGAGGGUGUCAGGAUGUCCCCGUGU
CCCCUCGAGCUUUGCCCGCCCCGAGCUCAGCUGCUCCCGGCGGCCAGAACGGAGGGGUAG
AGCAGUUCUCAGCGGCCCCGCGGGGCGGGCGGCAGUGUCCAGCCCCGGGGCUGCGCGAGACA
CAAAGUCCUCUGUGCCGUCCGACCCGGGCACCCUGACCACCGCCCUUUCCCCUCUGCAGGC
CGAUCCACCCGAGCCGGUCCCCACUGCCCAGCCGUGCUGCUGCCUCGGACACAAGCGCAG
GUACGUCCCGGCAUUGAUCGGAGGCAGGCCCUGGCCCACGUGGCCGAGAAGACGGAGGUUC
CAGGCCUGUCUUUGUUGAGCUGCGCGCCGGGCCUCCCCAGUCGCCCUUUCCUGUGCCCUUUU
AAGGUUGUCGCAGUACUUUUAAGAGGCUAACACAGAAGGGUAAAGGAACUCUCCAUAAAAC
CCAGAGAGGAAAACUACACCUCCUCUUUUGGAUCCUGUCUGGGAGUCACAAAUGAACCAAAA
GCUUCAUUUAAAACUUGAAUGUGCGCAUUUUAGGUUCUGGCGAUUGCUAGGUAGUGAUAU
AGGAUUUUUAAAGAAAAGAAAACGUAGCCUACCUGGUUUAAUUUUGGUUGGAGGAGAAAGUU
CAGAAGACAAGAAUACAUUUUGCUUUUCUAGAAGUGAGAGAUUUUUAGGUCAGCUGGAAAGGAU
AUAGUUGAGUAUGUUCCACUUAUGUGAAACAUUUUUUAUUCUUUUCAUAGGUUGAAAACUAC
AAGAAACCCGUGACUGAAUGGCUCCAAGGUGAAGUACUCUCUGGCAUGUAGUGACCAGGUUU
GCCUGUAUUCCACUUUAGGAUUAGUAUCGAAGCUGAAAUAGGUGAAGUGAGGGCCAUGUG
AGUUUCCUUGAAUGGCCUUGUAACUGAUUACCCUGAAGGUAGUUUUGCUAAAGAUUUUAAUU
CAAUGUACUAGGAUCCUGCCCCACCCCUUCUUCAUGUCCCACUGCCCUAGAAUCAGAUCU
GCAUGGCUAUGACAUGCUCGCCGUUUUAAAGAUGCAAAGAUUUUCAUGCUAGAUCAUUAGA
AUGUAUCUGGACUUCGUUAGGACAAGUUCUUUUGCUAGUAACCAGUCACGUGCUAGCAAUU
GAGGUGGUUGUUUAACCUGUAUAAUUGUUGCCUAUGAGUAAUUUUUCCAUAAAUCUUUAAUUAAAGG
CUUCCUGUGUCAAUUCUA

Interaction probabilities

Prediction using RF classifier   0.9
Prediction using SVM classifier 0.94

What do these probabilities mean?
Interaction probabilities generated by RPISeq 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%.

Please see About/FAQs for additional details.