**Supplementary Table Legends**

**Table S1:** Statistic of homologous protein present in the training and testing sets as well as from the blind datasets.

**Table S2:** Positive interactions dataset used in this study to build optimal model for prediction. The positive interactions dataset used in the study were obtained from VirusMINT.

**Table S3:** Negative interactions dataset used in this study to build optimal model for prediction. The negative interactions dataset used in the study were chosen using random protein pairs which are not found in interacting protein pairs.

**Table S4:** Positive interactions dataset used in this study as a positive blind dataset. The positive blind dataset used in the study were obtained from VirusMINT.

**Table S5:** Negative interactions dataset used in this study as a negative blind dataset. The negative blind dataset used in the study were chosen using random protein pairs which are not found in interacting protein pairs (positive blind dataset).

**Table S6:** All 44 input features

**Table S7:** SVM performance measures based on different subsets of features. Optimal parameters were used for respective subset of features.

**Table S8:** Several SVM kernel-wise performance measures (sensitivity and specificity) on different models. Optimal parameters and threshold were used for respective kernel. In Model 1, 1st, 2nd, 3rd and 4th folds were used for training and 5th fold was kept for testing. In Model 2, 1st, 2nd, 3rd and 5th folds were used for training and 4th fold was left out for testing. In Model 3, 1st, 2nd, 4th, 5th folds were used for training and 3rd fold for testing. In Model 4, 1st, 3rd, 4th, 5th folds were used for training and 2nd fold was used for testing. In Model 5, 2nd, 3rd, 4th, 5th folds were used for training and 1st fold was kept aside for testing.

(13KB, xlsx)

**Table S9:** Different parameters used in Random Forest using WEKA.

(15KB, xlsx)

**Table S10:** Predicted scores of HBV-human protein-protein association by proposed optimal model.

(1.32MB, xlsx)

**Table S11:** HEV-human protein-protein association predicted scores by proposed optimal model.

(2.01MB, xlsx)

**Table S12:** GO enrichment analysis on interacting human protein partners of HBV X proteins using DAVID server. Significant biological process terms were chosen by PValue< 0.05**.**

(18KB, xlsx)

**Table S13:** GO enrichment analysis on interacting human protein partners of HBV C proteins using DAVID server. Significant biological process terms were chosen by PValue< 0.05**.**

(21KB, xlsx)

**Table S14:** GO enrichment analysis on interacting human protein partners of HBV P proteins using DAVID server. Significant biological process terms were chosen by PValue< 0.05**.**

(28KB, xlsx)

**Table S15:** GO enrichment analysis on interacting human protein partners of HEV proteins using DAVID server. Significant biological process annotation terms were filter by FDR (false discovery rate)< 0.05.

(14KB, xlsx)

**Table S16:** GO enrichment analysis on interacting human protein partners of HEV ORF1 (Genotype 1) proteins using DAVID server. Significant biological process terms were chosen by PValue< 0.05**.**

(44KB, xlsx)

**Table S17:** GO enrichment analysis on interacting human protein partners of HEV ORF2 (Genotype 1) proteins using DAVID server. Significant biological process terms were chosen by PValue< 0.05**.**

(16KB, xlsx)

**Table S18:** GO enrichment analysis on interacting human protein partners of HEV ORF3 (Genotype 1) proteins using DAVID server. Significant biological process terms were chosen by PValue< 0.05**.**

(16KB, xlsx)

**Table S19:** GO enrichment analysis on interacting human protein partners of HEV ORF1 (Genotype 4) proteins using DAVID server. Significant biological process terms were chosen by PValue< 0.05**.**

(21KB, xlsx)

**Table S20:** GO enrichment analysis on interacting human protein partners of HEV ORF2 (Genotype 4) proteins using DAVID server. Significant biological process terms were chosen by PValue< 0.05**.**

(16KB, xlsx)

**Table S21:** GO enrichment analysis on interacting human protein partners of HEV ORF3 (Genotype 4) proteins using DAVID server. Significant biological process terms were chosen by PValue< 0.05**.**

(16KB, xlsx)

**Table S22:** List of experimentally validated sRNAs of SLT2 with name, source of identification, start and end position.

**Table S23:** List of experimentally validated 182 sRNAs of SLT2 with name, source of identification, start position, end position, length and sequence.

**Table S24:** Negative set1 of SLT2 with name, start position, end position, length and sequence.

**Table S25:** Negative set2 of SLT2 with name, start position, end position, length and sequence.

**Table S26:** Negative set3 of SLT2 with name, start position, end position, length and sequence.

**Table S27:** Negative set4 of SLT2 with name, start position, end position, length and sequence.

**Table S28:** Negative set5 of SLT2 with name, start position, end position, length and sequence.

**Table S29:** Negative set6 of SLT2 with name, start position, end position, length and sequence.

**Table S30:** Negative set7 of SLT2 with name, start position, end position, length and sequence.

**Table S31:** Negative set8 of SLT2 with name, start position, end position, length and sequence.

**Table S32:** Negative set9 of SLT2 with name, start position, end position, length and sequence.

**Table S33:** Negative set10 of SLT2 with name, start position, end position, length and sequence.

**Table S34:** Nucleotide features were significantly different in sRNAs rather than non-sRNAs and vice-versa.

**Table S35:** Performance comparison of different kernel of SVM. Optimal parameter sets were used for respective kernel.

**Table S36:** Prediction score of experimentally verified sRNAs of S. Typhimurium LT2 (SLT2) using sliding window based method.

**Table S37:** Prediction score of experimentally verified sRNAs of S. Typhi ty2 using sliding window based method.

**Table S38:** Prediction score of experimentally verified sRNAs of E. coli K-12 using sliding window based method.

**TableS39:** All the curated infectious diseases-associated human genes from DisGeNET.

**Table S40:** All the mapped gene name to uniprotid using mapping table of DisGeNET.

**Table S41:** Positive dataset for 10-fold cross-validation.

**Table S42:** Positive blind dataset (not used in training or testing of 10-fold cross-validation techniques for developing the prediction model).

**Table S43:** All the disease-associated human reviewed proteins in DisGeNET.

**Table S44:** All the reviewed human proteins collected from UniProtKB dated 12/01/2018.

**Table S45:** All the reviewed human proteins not associated with any diseases.

**Table S46:** Negative dataset for 10-fold cross-validation.

**Table S47:** Negative blind dataset (not used in training or testing of 10-fold cross-validation techniques for developing the prediction model).

**Table S48:** Independent dataset (Befree text mining genes from DisGeNET associated with infectious diseases).

**Table S49:** All human protein-protein interactions (PPIs) from Human Protein Reference Database (HPRD) (Release 9).

**Table S50:** All unique human in HPRD (Release 9).

**Table S51:** All the mapped human protein-protein interactions (PPIs) in uniprot id format.

**Table S52:** All the mapped unique human proteins in uniprot.

**Table S53:** 9 topological properties of protein-protein interaction networks using HPRD PPIs dataset.

**Table S54:** Features wise performance measures on disease and non-disease associated proteins dataset using deep neural network classifier.

**Table S55:** Mixed features based performance on disease and non-disease associated proteins dataset.

**Table S56:** 10 selected features for normalized and filtered PAAC and Network properties.

**Table S57:** 16 selected features for PAAC and Network properties.

**Table S58:** Selected features wise performance measures using different classifier.

**Table S59:** Prediction result on independent dataset.

**Table S60:** Top 100 proteins (genes) are predicted by our proposed DNN based method.

**Table S61:** Significantly enriched disease-ontology terms for top 100 proteins (genes) based on Genetic Association Database (GAD).

**Table S62:** Significantly enriched gene-ontology biological process terms for top 100 proteins (genes).

**Table S63:** Bacterial targeted reviewed human proteins human proteins.

**Table S64:** Viral targeted reviewed human proteins human proteins.

**Table S65:** The top GO IDs for bacterial and viral targeted human proteins.

**Table S66:** Features wise performance measure.

**Table S67:** Selelected features wise performance measure.

**Table S68:** Probability score of top 100 bacteria targeted human proteins**.**

**Table S69:** Probability score of top 100 virus targeted human proteins.