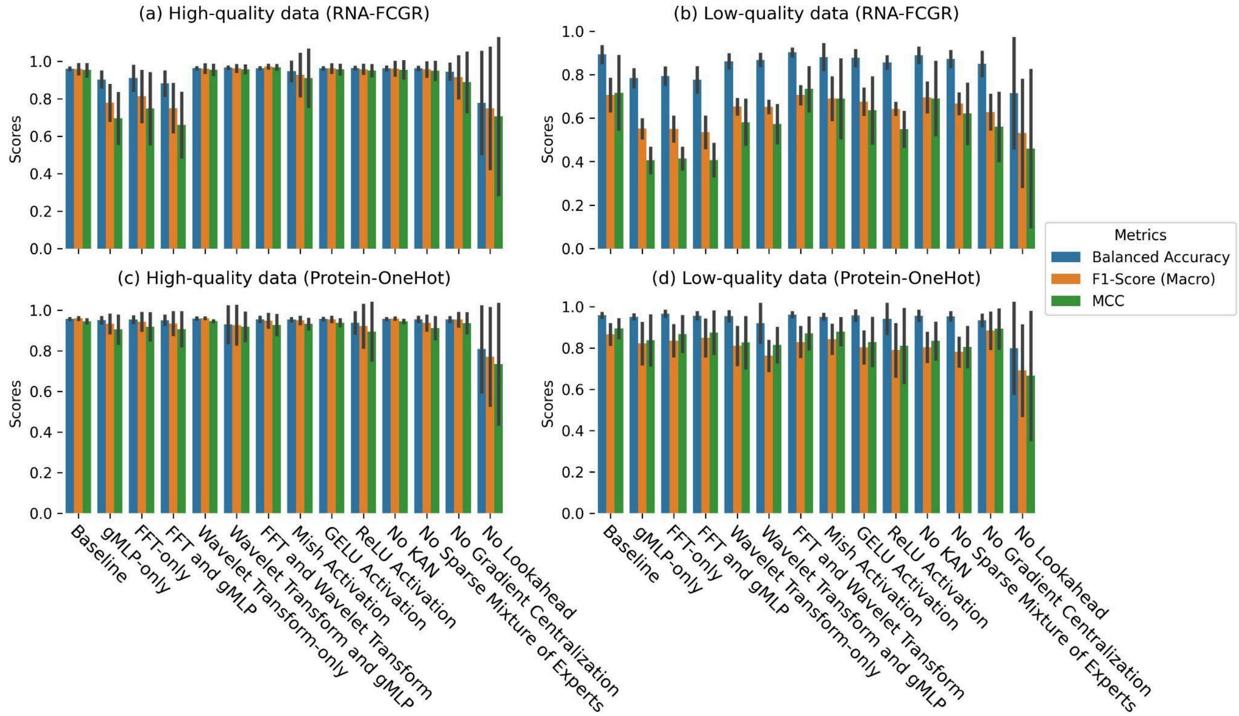
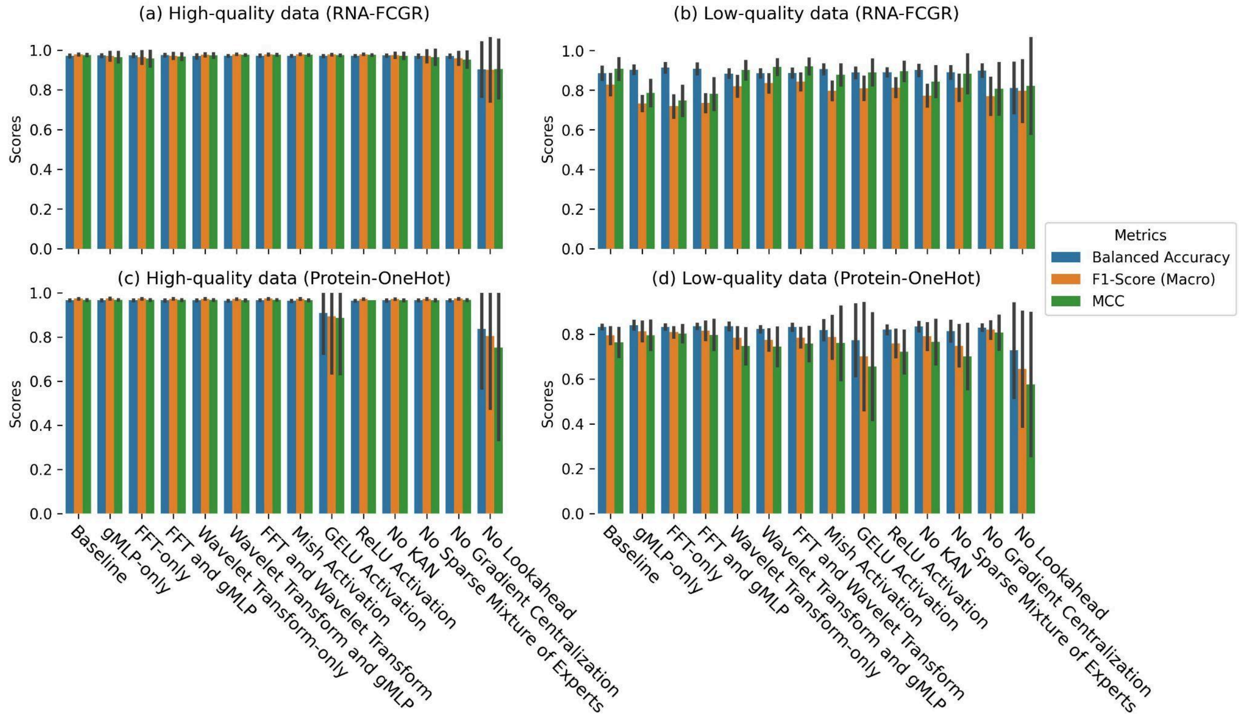
# Supplementary Materials

**Figure S1:** The generalization performance of WaveSeekerNet for host source prediction was evaluated on the HA segment using various hyperparameter settings. The Balanced Accuracy, F1-score (Macro Average), and Matthews Correlation Coefficient (MCC) are reported for high- (a) and low- (b) quality FCGR representation of RNA sequences. Scores for tests on the dataset constructed from high-quality and low-quality one-hot encoded protein sequences are reported in panels (c) and (d), respectively.



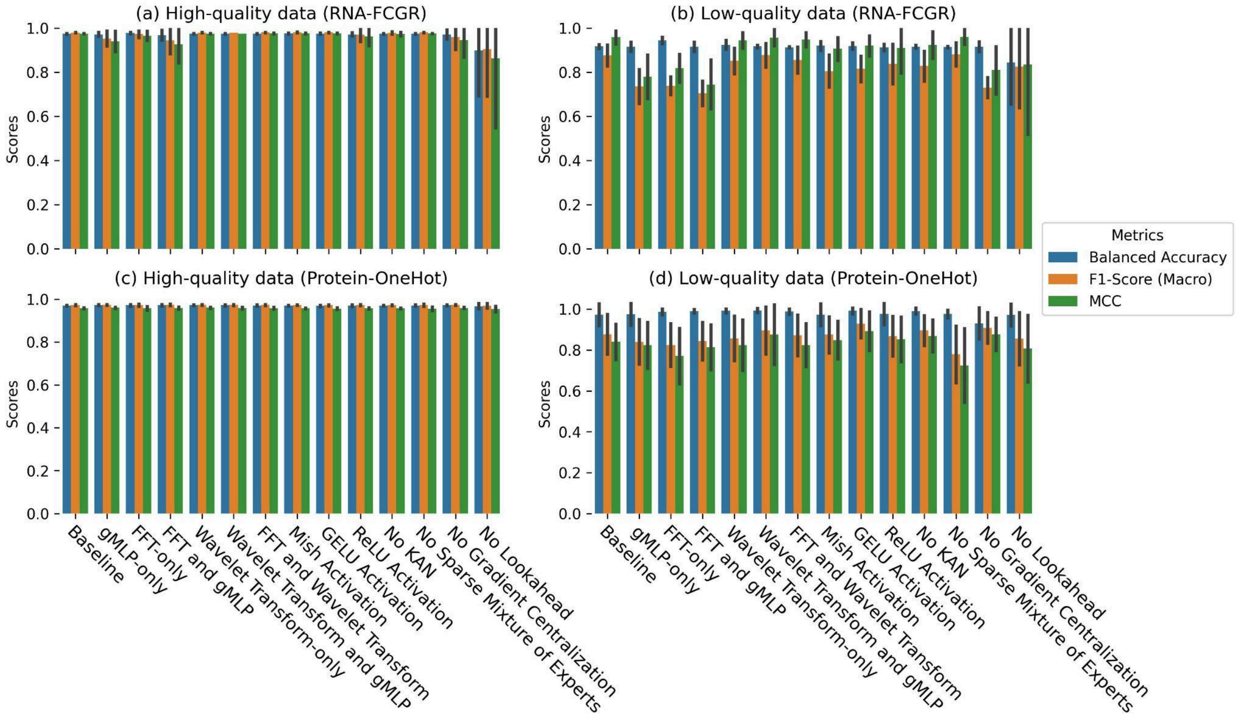
## Supplementary Materials

**Figure S2:** The generalization performance of WaveSeekerNet for host source prediction was evaluated on the NA segment using various hyperparameter settings. The Balanced Accuracy, F1-score (Macro Average), and Matthews Correlation Coefficient (MCC) are reported for high- (a) and low- (b) quality FCGR representation of RNA sequences. Scores for tests on the dataset constructed from high-quality and low-quality one-hot encoded protein sequences are reported in panels (c) and (d), respectively.



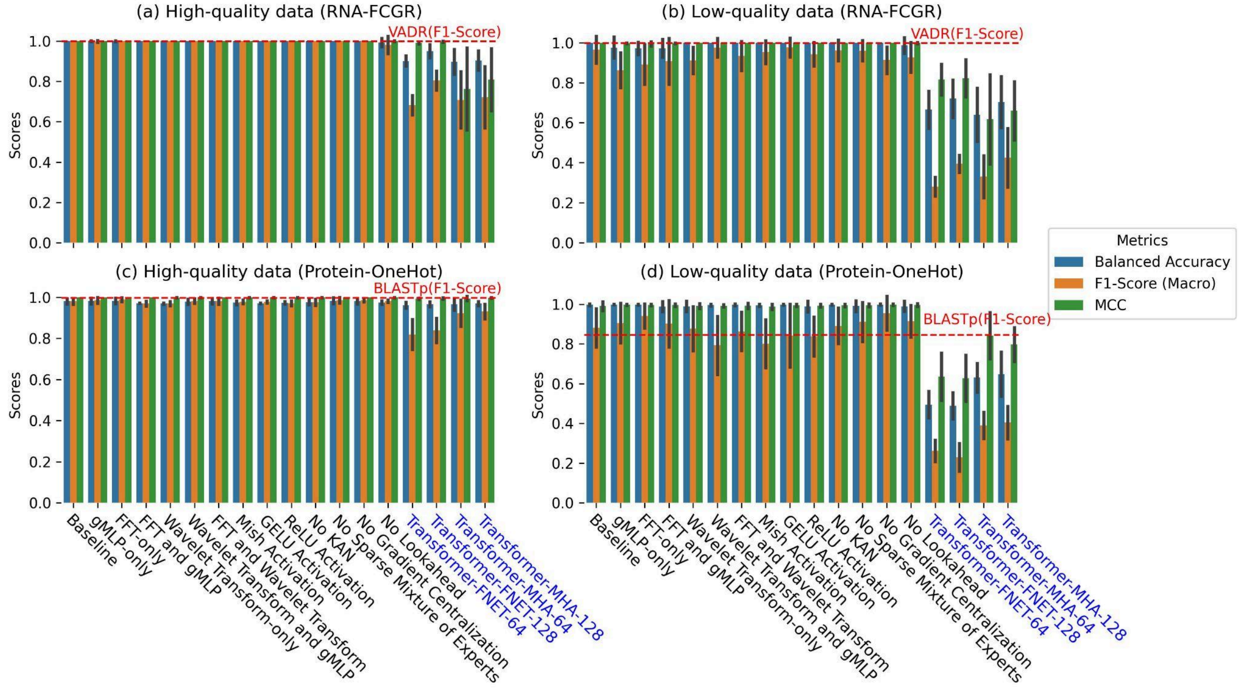
## Supplementary Materials

**Figure S3:** The generalization performance of WaveSeekerNet for host source prediction was evaluated on the combined HA and NA segments (2 channels) using various hyperparameter settings. The Balanced Accuracy, F1-score (Macro Average), and Matthews Correlation Coefficient (MCC) are reported for high- (a) and low- (b) quality FCGR representation of RNA sequences. Scores for tests on the dataset constructed from high-quality and low-quality one-hot encoded protein sequences are reported in panels (c) and (d), respectively.



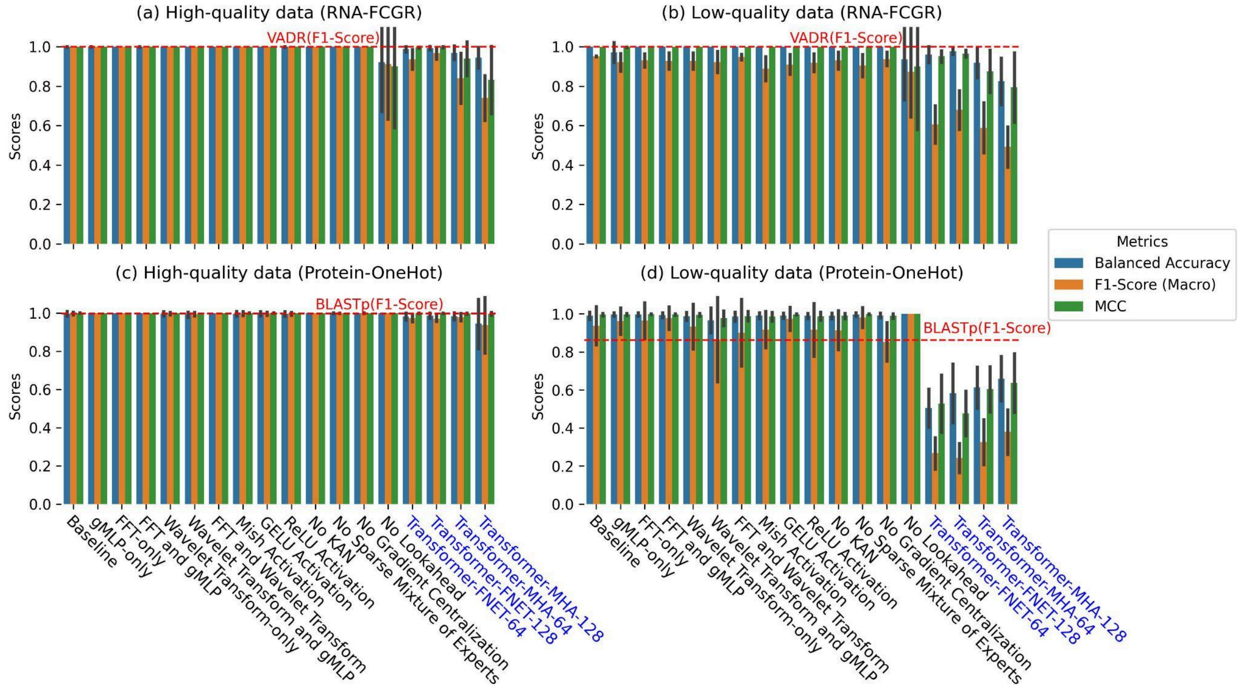
## Supplementary Materials

**Figure S4:** The generalization performance of WaveSeekerNet and Transformer-only for HA subtype prediction using various hyperparameter settings. The Balanced Accuracy, F1-score (Macro Average), and Matthews Correlation Coefficient (MCC) are reported for high- (a) and low- (b) quality FCGR representation of RNA sequences. Panels (c) and (d) show scores for high- and low-quality one-hot encoded protein sequences, respectively.  WaveSeekerNet and Transformer-only using various hyperparameter settings are labeled in Black and Blue, respectively. The F1-score (Macro Average) for VADR and BLASTp are shown as red horizontal lines.



## Supplementary Materials

**Figure S5:** The generalization performance of WaveSeekerNet and Transformer-only for NA subtype prediction using various hyperparameter settings. The Balanced Accuracy, F1-score (Macro Average), and Matthews Correlation Coefficient (MCC) are reported for high- (a) and low- (b) quality FCGR representation of RNA sequences. Panels (c) and (d) show scores for high- and low-quality one-hot encoded protein sequences, respectively. WaveSeekerNet and Transformer-only using various hyperparameter settings are labeled in Black and Blue, respectively. The F1-score (Macro Average) for VADR and BLASTp are shown as red horizontal lines.



## Supplementary Materials

**Figure S6:** Results of NA subtype prediction when tested the best-performing WaveSeekerNet models, Transformer-only models. The performance of baseline WaveSeekerNet is also shown as a point of reference. The boxen plots of Balanced Accuracy, F1-score (Macro Average), and Matthews Correlation Coefficient (MCC) are reported for the high (a) and low (b) quality FCGR representation of RNA sequences. The boxen plots of scores for tests on the dataset constructed from high-quality and low-quality one-hot encoded protein sequences are reported in panels (c) and (d), respectively. Horizontal dash-lines present the F1-score (Macro Average) for VADR and BLASTp. The performance of the WaveSeekerNet models shows little variance and overlaps with that of VADR.

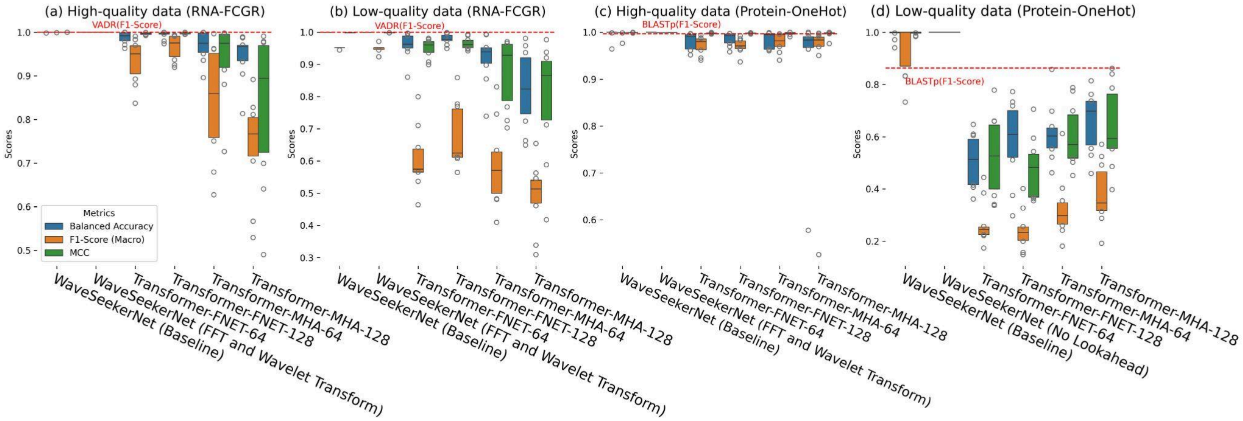
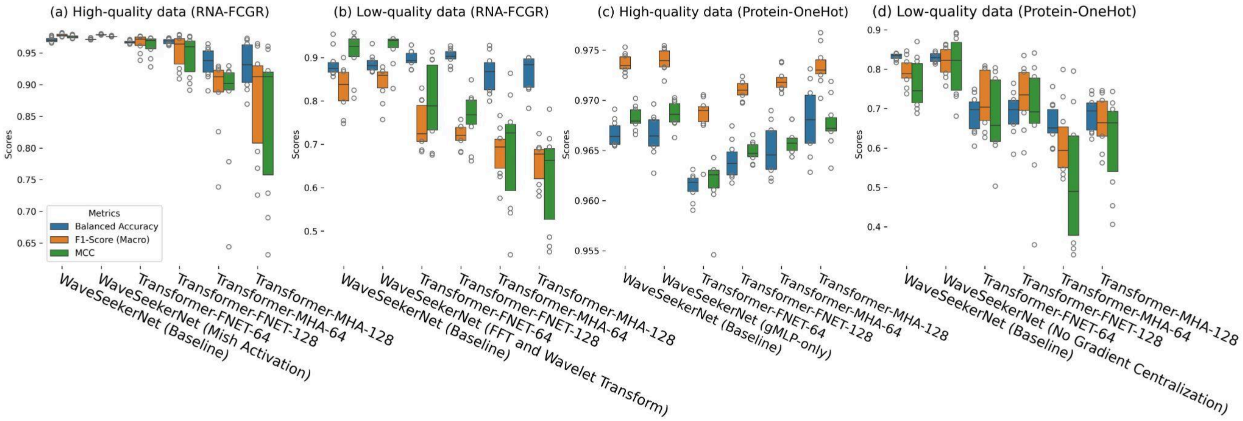


**Figure S7:** Results of host source prediction using the NA segment when tested the best-performing WaveSeekerNet models, Transformer-only models. The performance of baseline WaveSeekerNet is also shown as a point of reference. The boxen plots of Balanced Accuracy, F1-score (Macro Average), and Matthews Correlation Coefficient (MCC) are reported for the high (a) and low (b) quality FCGR representation of RNA sequences. The boxen plots of scores for tests on the dataset constructed from high-quality and low-quality one-hot encoded protein sequences are reported in panels (c) and (d), respectively.


(a) High-quality data (RNA-FCGR)
Scores
Metrics
Balanced Accuracy
F1-Score (Macro)
MCC
WaveSeekerNet (Baseline)
WaveSeekerNet (Mish Activation)
Transformer-FNET-64
Transformer-FNET-128
Transformer-MHA-64
Transformer-MHA-128
(b) Low-quality data (RNA-FCGR)
Scores
WaveSeekerNet (Baseline)
WaveSeekerNet (FFT and Wavelet Transform)
Transformer-FNET-64
Transformer-FNET-128
Transformer-MHA-64
Transformer-MHA-128
(c) High-quality data (Protein-OneHot)
Scores
WaveSeekerNet (Baseline)
WaveSeekerNet (gMLP-only)
Transformer-FNET-64
Transformer-FNET-128
Transformer-MHA-64
Transformer-MHA-128
(d) Low-quality data (Protein-OneHot)
Scores
WaveSeekerNet (Baseline)
WaveSeekerNet (No Gradient Centralization)
Transformer-FNET-64
Transformer-FNET-128
Transformer-MHA-64
Transformer-MHA-128