



S Fig 1: HMMERCTTER user interfaces

Shown are single examples obtained from the training (A) and the target phase (B). Blue asterisks are scores of sequences of the complete dataset (either training or combined), green circles are scores of sequences of the group in question. The red line shows the drop in scores among successive sequences, also indicated as Score-drop. The magenta line shows the current threshold and can be moved in order to determine the threshold in the interactive part of the classification. In the training the user must accept or reject the 100% P&R cluster. During classification the user can accept the cluster, return to the former or initial state or add seemingly negatives as indicated for posterior 100% P&R testing.