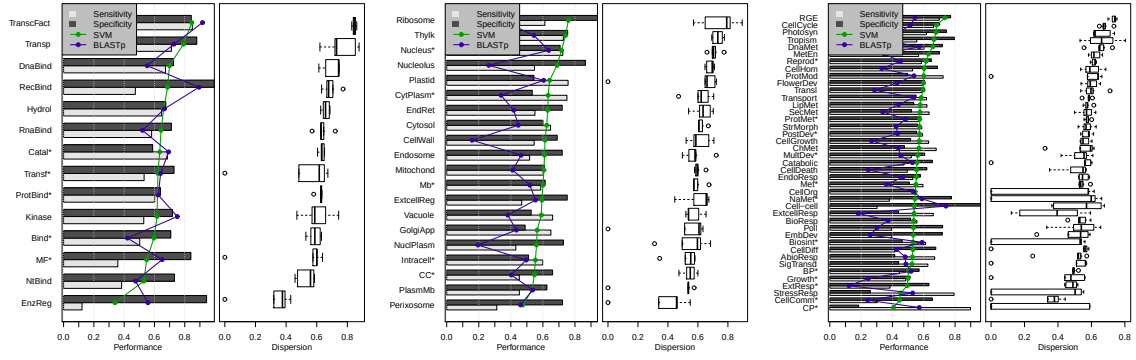
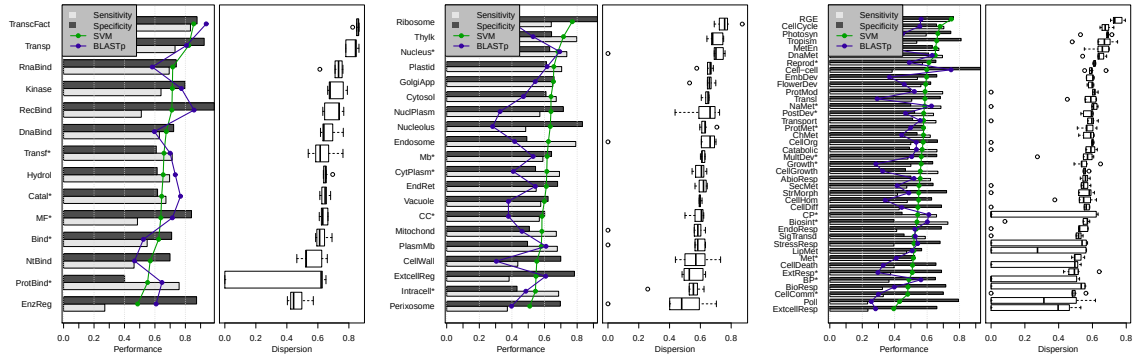


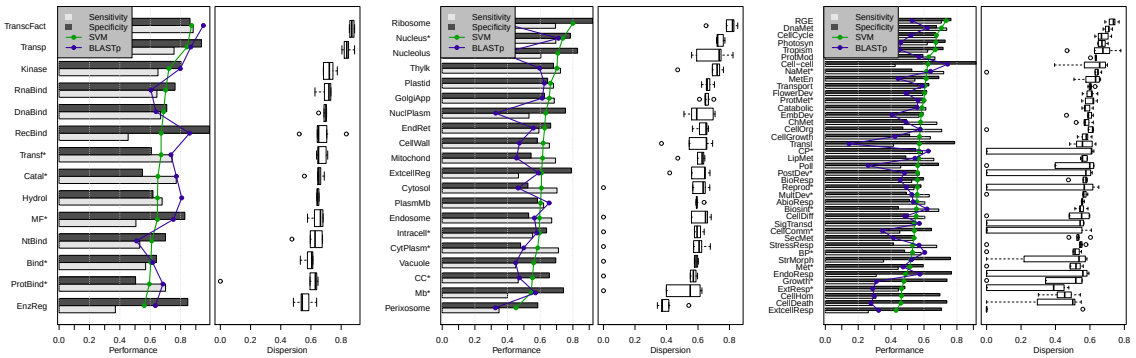
The following figures depict detailed results of predicting each class with the full feature set, for several identity cutoffs. Left plots show sensitivity, specificity and geometric mean (green line) achieved with the five-fold cross-validation procedure, while right plots depicts boxplots for analyzing performance variation throughout the five repetitions. Left plots also depicts the performance of the BLASTP algorithm for comparison purposes (blue line).



(a) Molecular function at 30% cutoff (b) Cellular component at 30% cutoff (c) Biological process at 30% cutoff

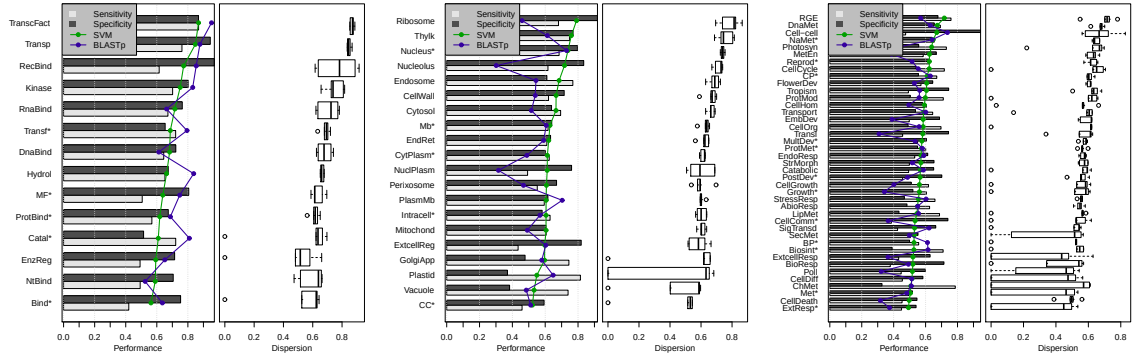


(d) Molecular function at 40% cutoff (e) Cellular component at 40% cutoff (f) Biological process at 40% cutoff

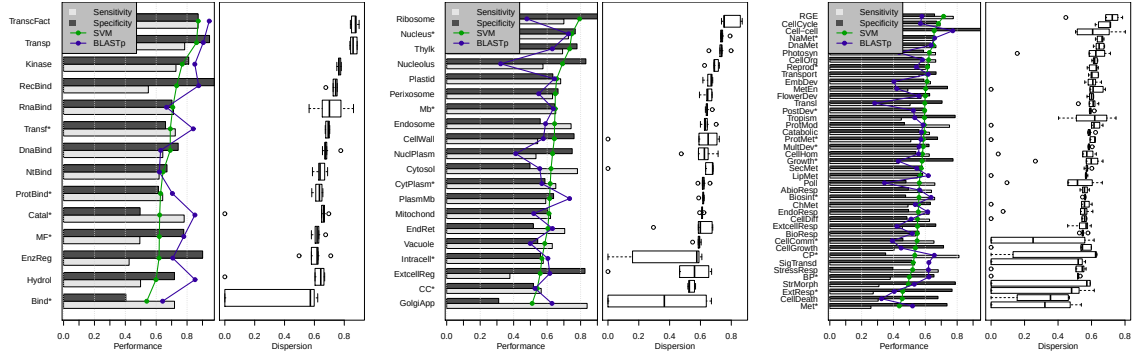


(g) Molecular function at 50% cutoff (h) Cellular component at 50% cutoff (i) Biological process at 50% cutoff

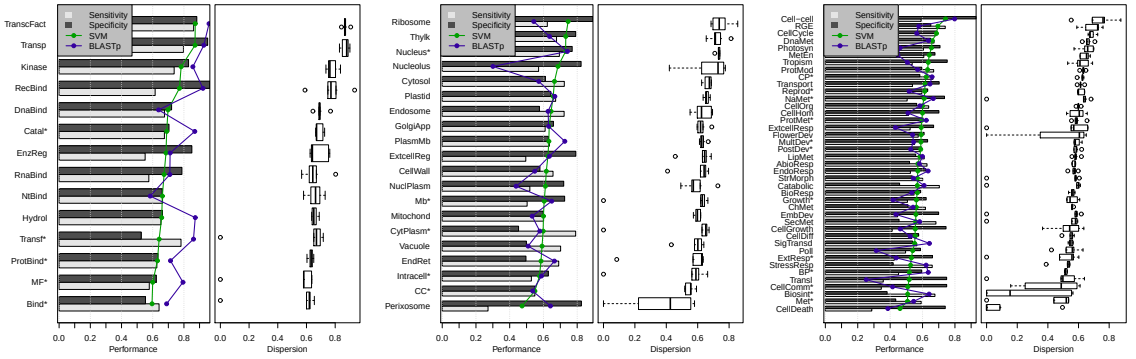
Figure 1: Detailed results for various identity percentage cutoffs



(a) Molecular function at 60% cutoff (b) Cellular component at 60% cutoff (c) Biological process at 60% cutoff



(d) Molecular function at 70% cutoff (e) Cellular component at 70% cutoff (f) Biological process at 70% cutoff



(g) Molecular function at 80% cutoff (h) Cellular component at 80% cutoff (i) Biological process at 80% cutoff

Figure 2: Detailed results for various identity percentage cutoffs