

Feature Overlap Between Top BRCA Models

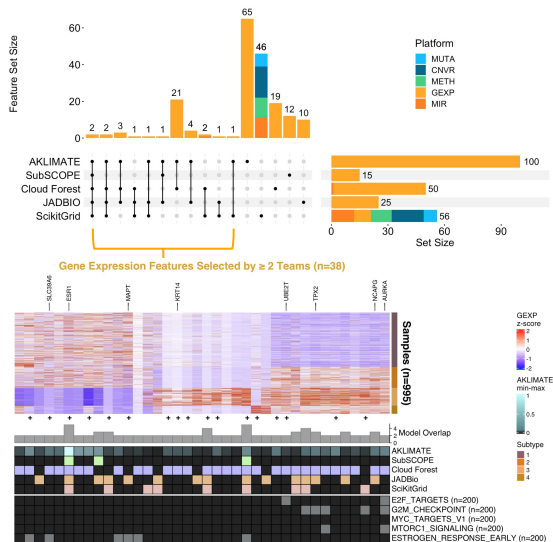
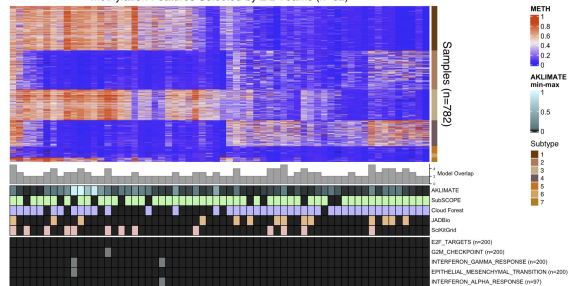
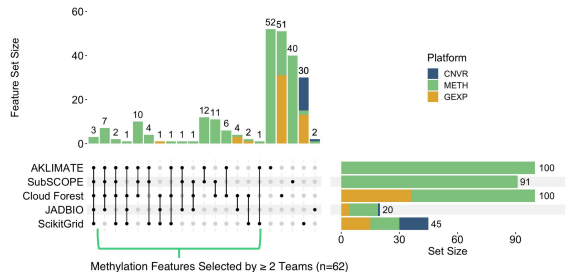
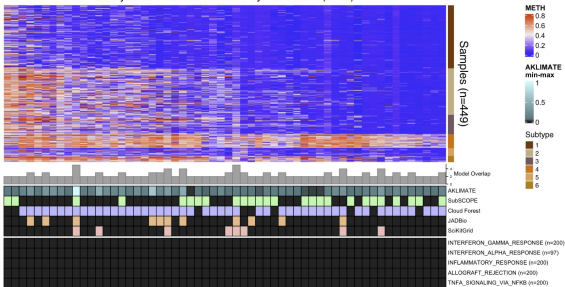
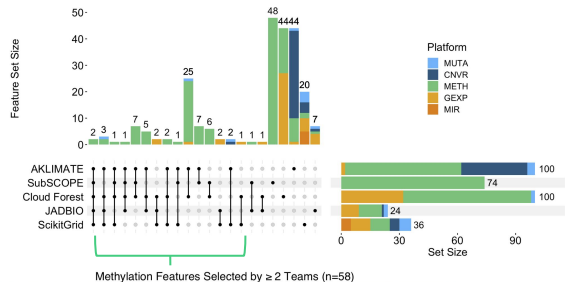


Figure 5 Group best model feature overlap. Overlap of all five data platforms between groups shown in upset plot. The molecular profiles of features common across ≥ 2 groups along with hallmark associations are shown in heatmap. AKLIMATE min-max normalized feature importance scores were calculated across all data platforms together. All other teams show presence/absence in feature set. + indicate PAM50 membership. Gene symbols labeled for features selected by ≥ 3 groups.

Feature Overlap Between Top LGGGBM Models



Feature Overlap Between Top COADREAD Models



Pipeline Highlights

Updated on 5/17/21

Plots created based on:

1. Heatmaps only show features: found in at least 2 teams

Data Used

- TMP_v12_20210228.tar.gz (syn25007874)
- classifier_metrics_20210514.tar.gz(syn25722573)
- aklimite_feature_importance_scores_20200807.tar.gz (syn22294880)

Scaled data: GEXP, MIR

AKLIMATE importance score normalized across all data platforms together, then we look at each data type separately on heatmaps. Thus some platforms may only have features with low importance scores.

Temporarily skipped cohorts: KIRCKICH, LIHCCHOL, UCEC