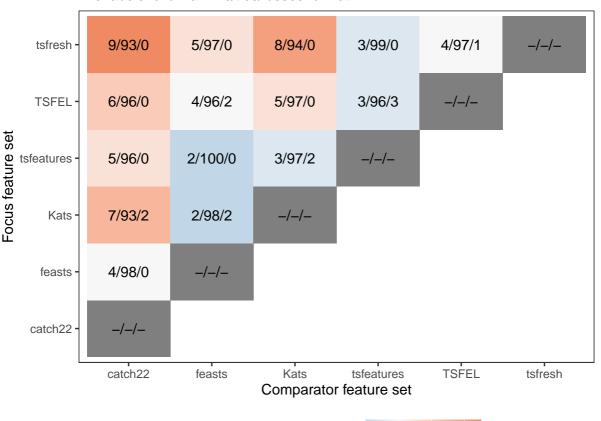Head to head of feature sets over maximum of 102 problems

t−tests between accuracy distributions for each feature set and problem combina and p−values corrected using Holm−Bonferroni method for each pairwise feature Tile labels follow a wins/ties/losses format.