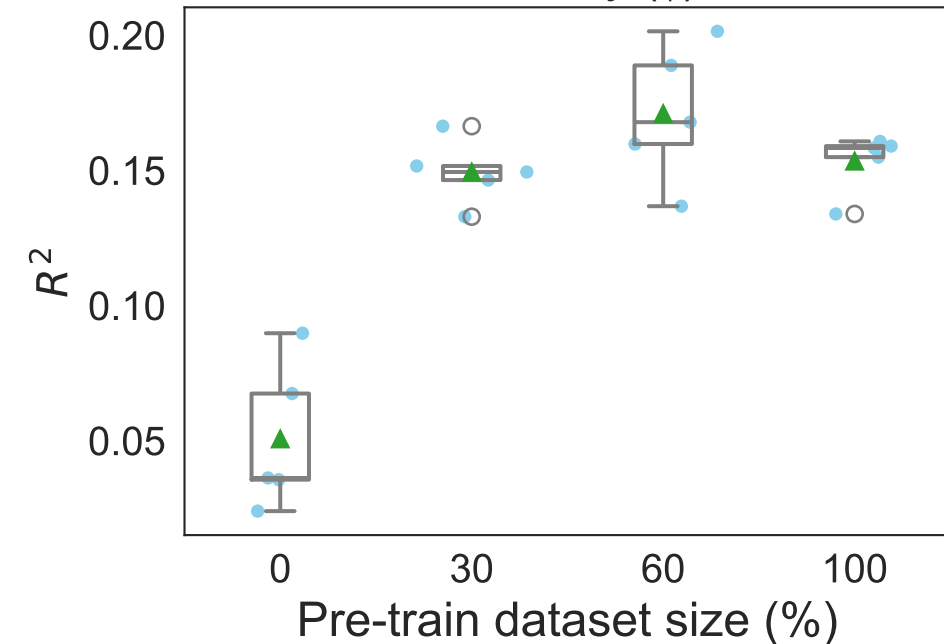
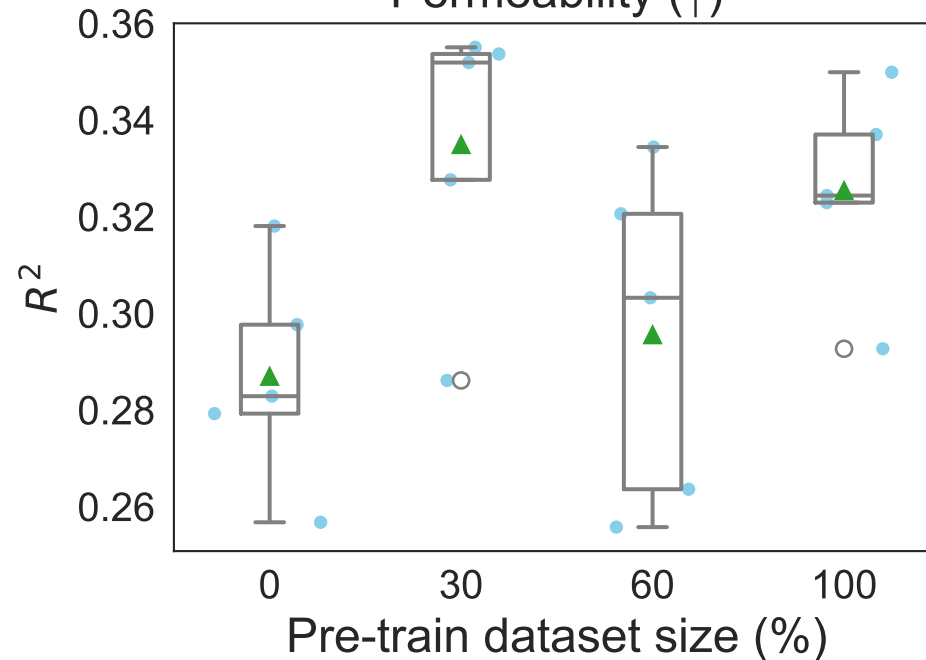


Comparing pre-training dataset sizes / pre-training objective = MLM / split = random

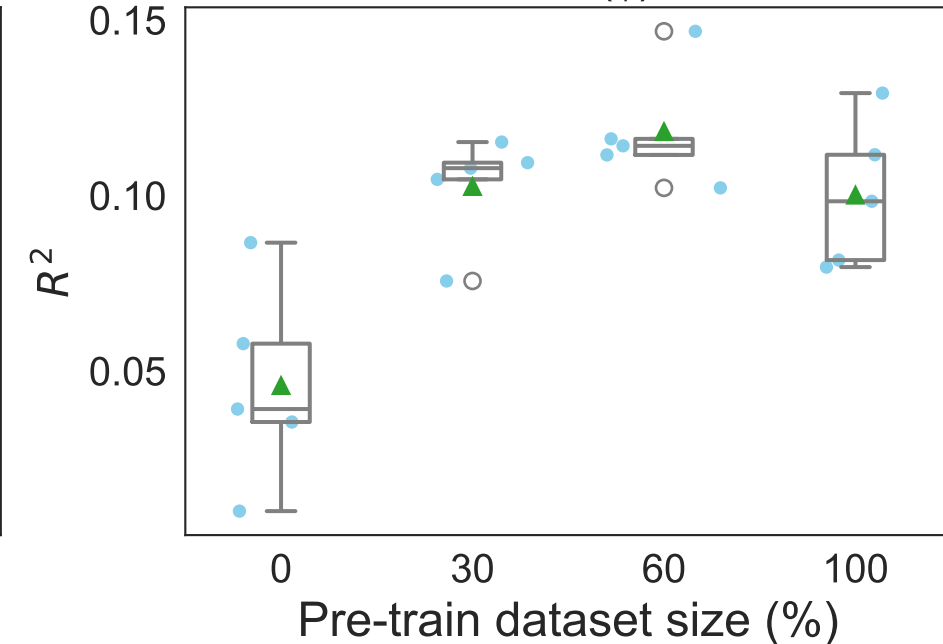
Solubility (\uparrow)



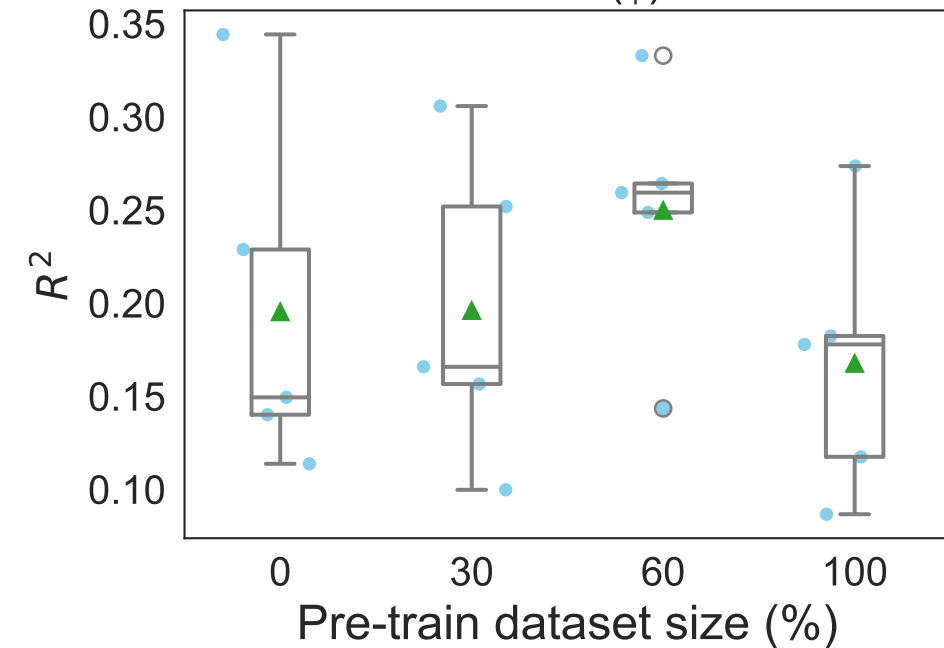
Permeability (\uparrow)



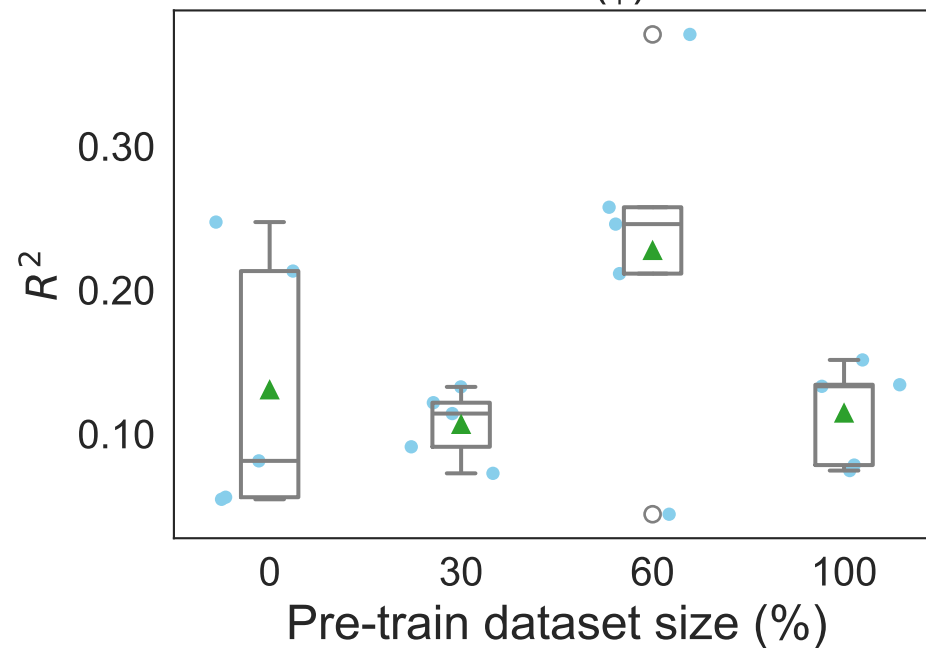
HLM (\uparrow)



hPPB (\uparrow)



rPPB (\uparrow)



RLM (\uparrow)

