

Restructure meta-data as ductal, lobular or normal



Median center data, arraywise



For each data set exclude IMAGE clones with >50% missing data



Gene filtering with T-test



Exclude IMAGE clones with >40% missing data

for each



Test



Shuffled 1

Shuffled 2



Supervised re-sampling



Unsupervised re-sampling



Ranksum test: p_t



Ranksum test: pr_1, pr_2



Kolmogorov-Smirnov Test: kst, ksr



Determine cut off value according to FDR_{observed}



Select significant genes