Enhancer Predictions

P2-TECH

Probability of peak scores
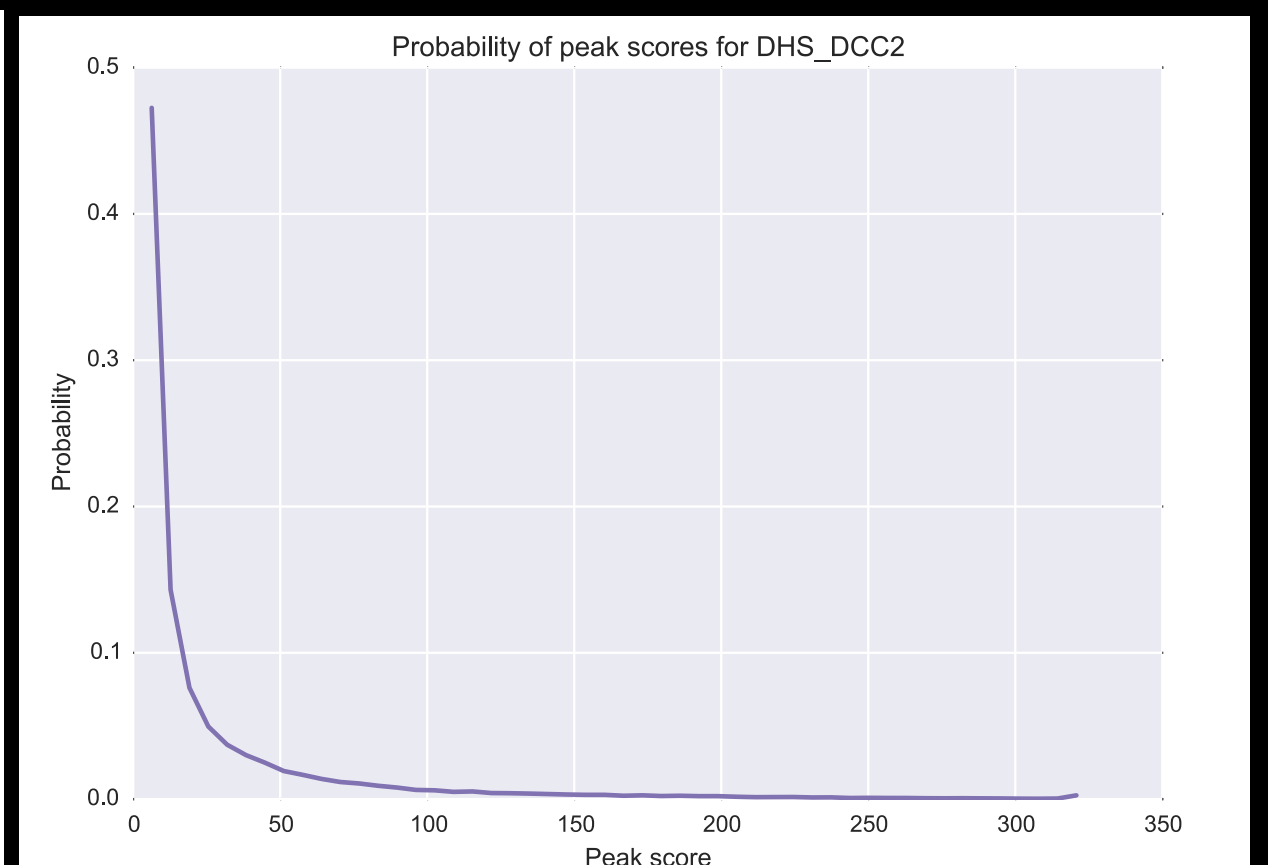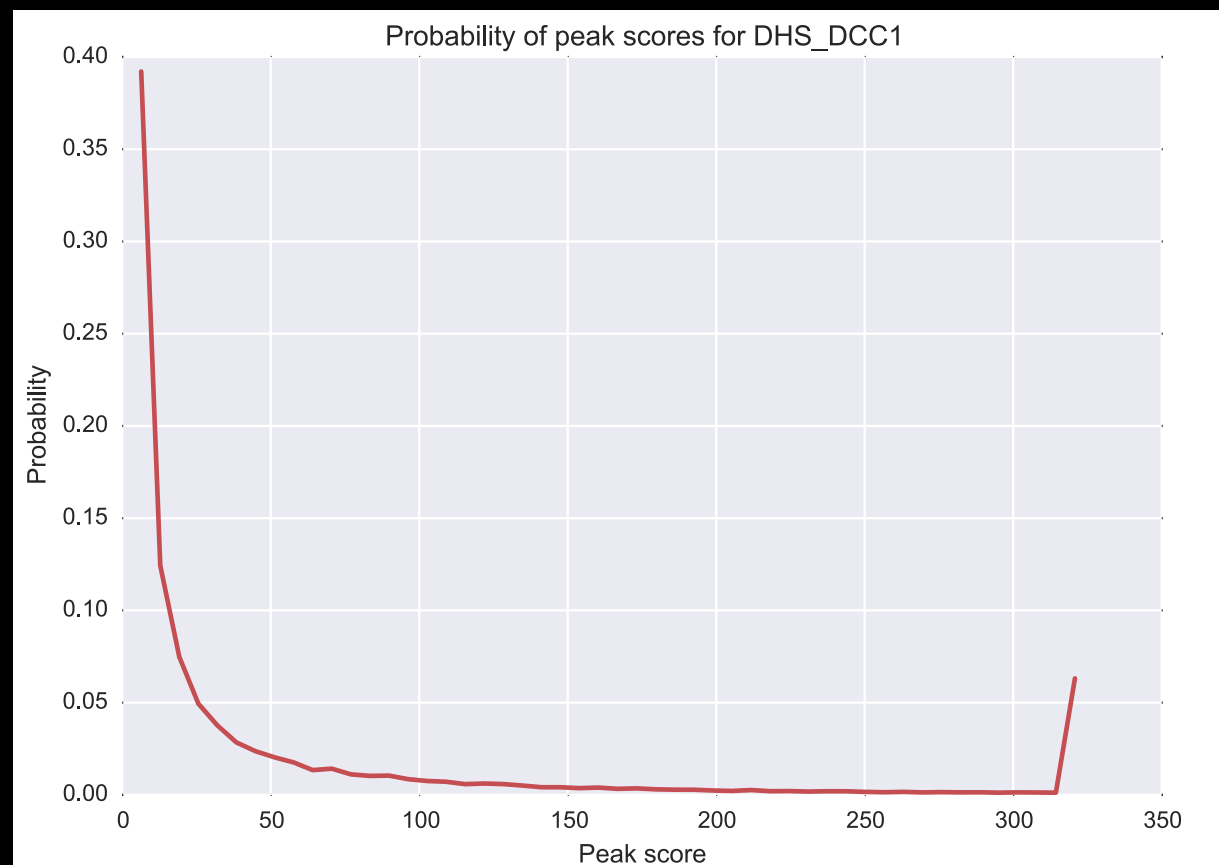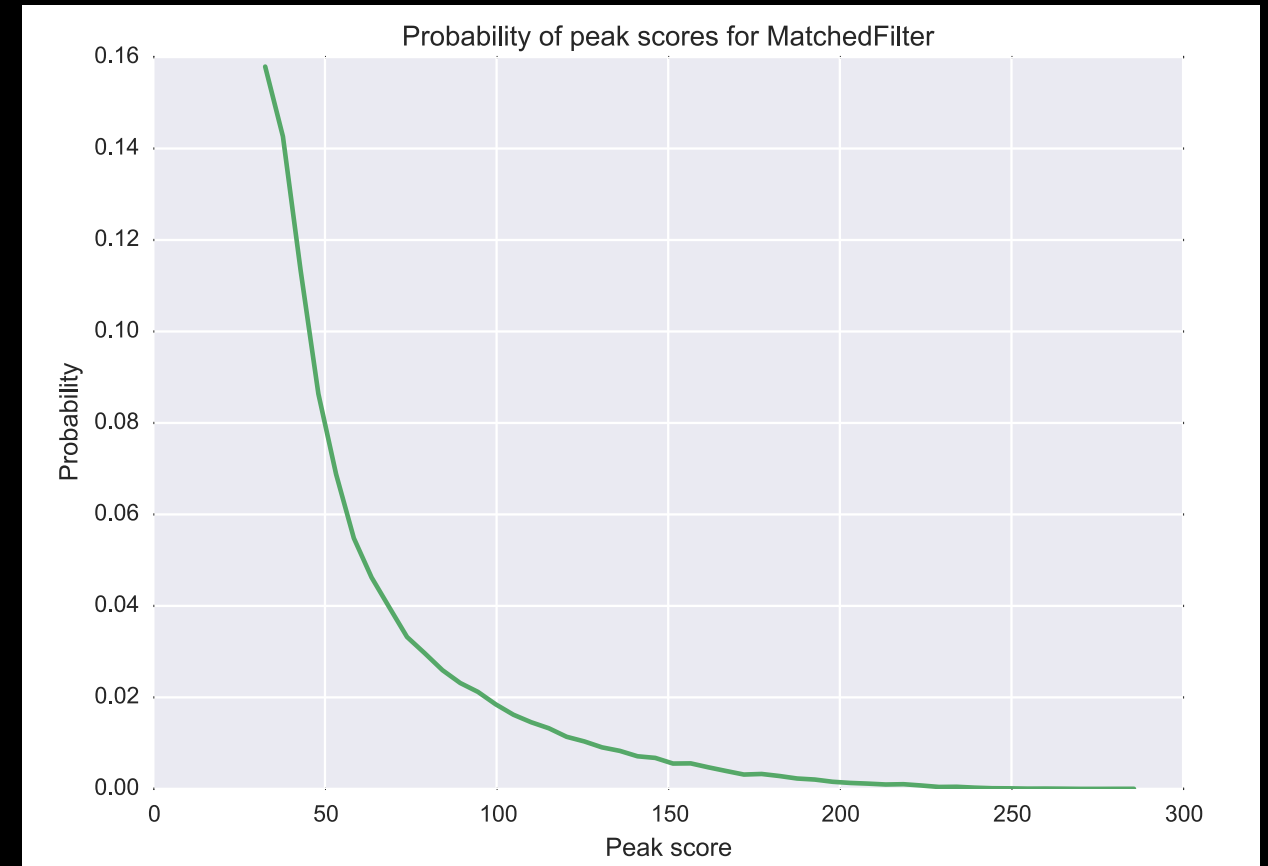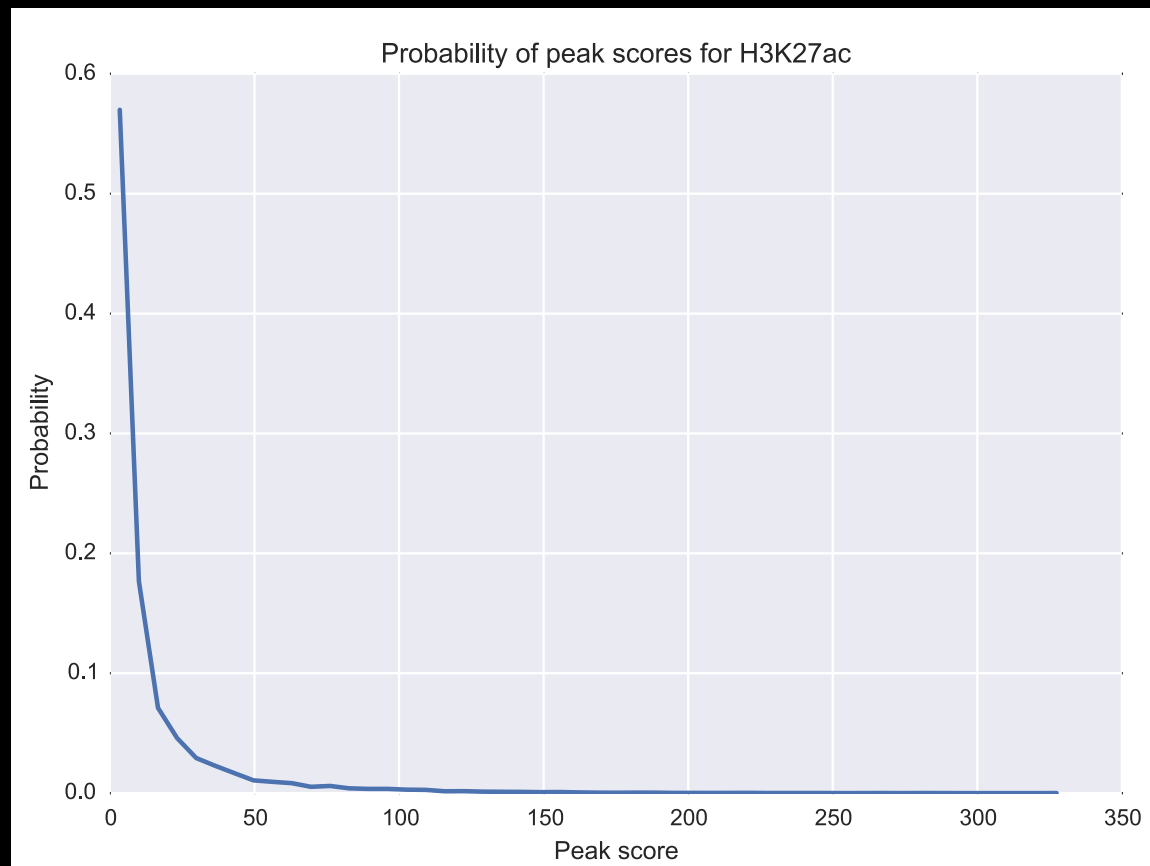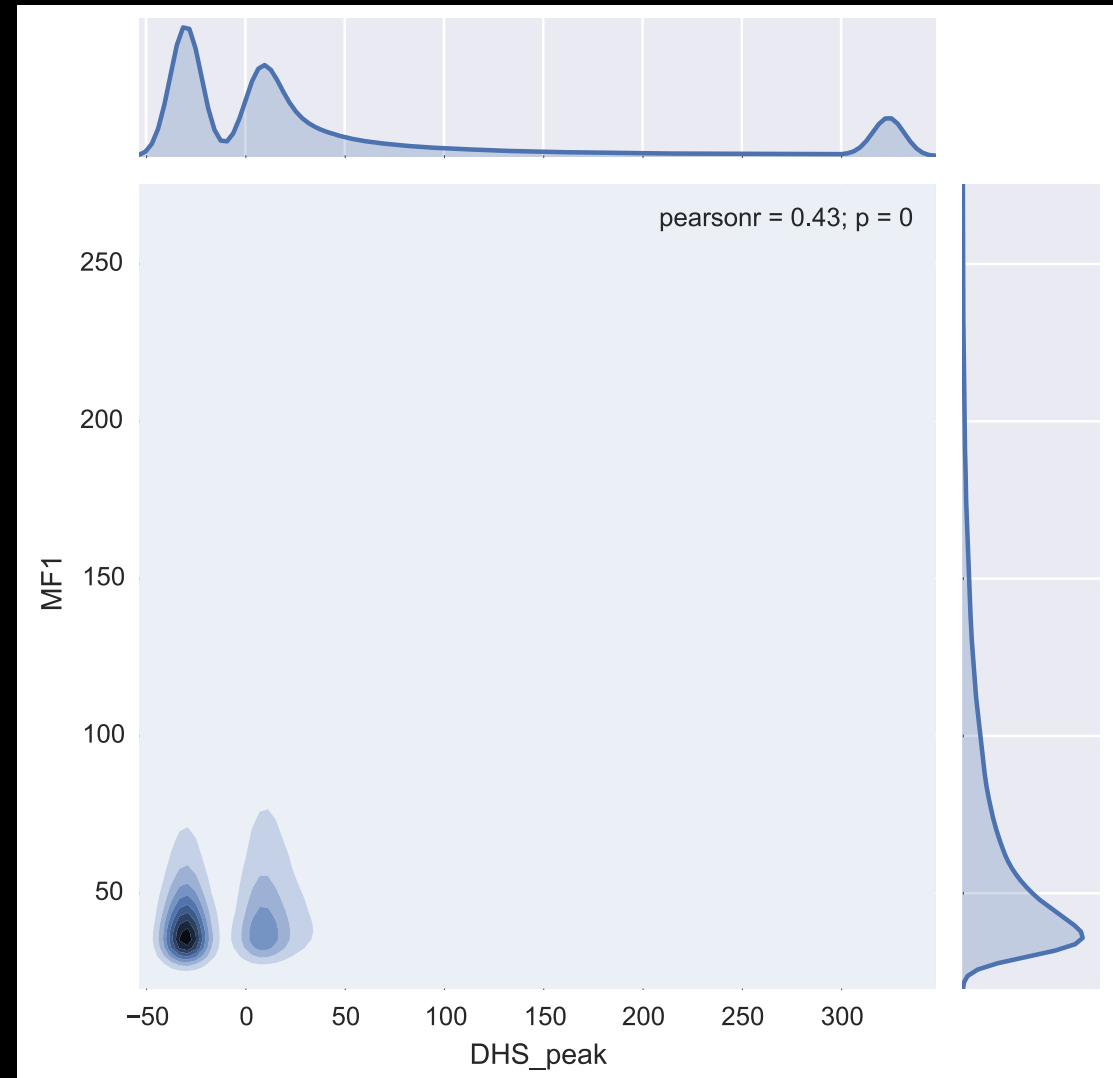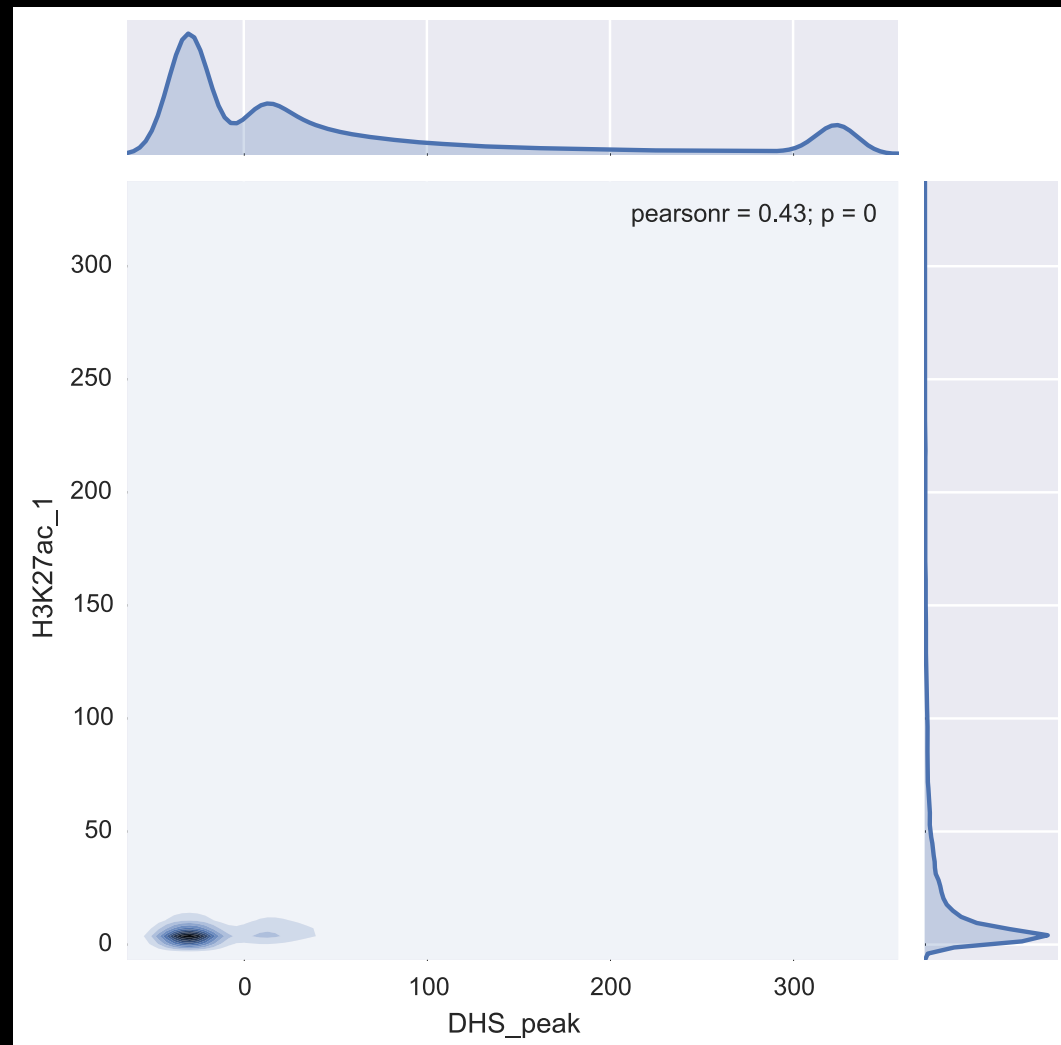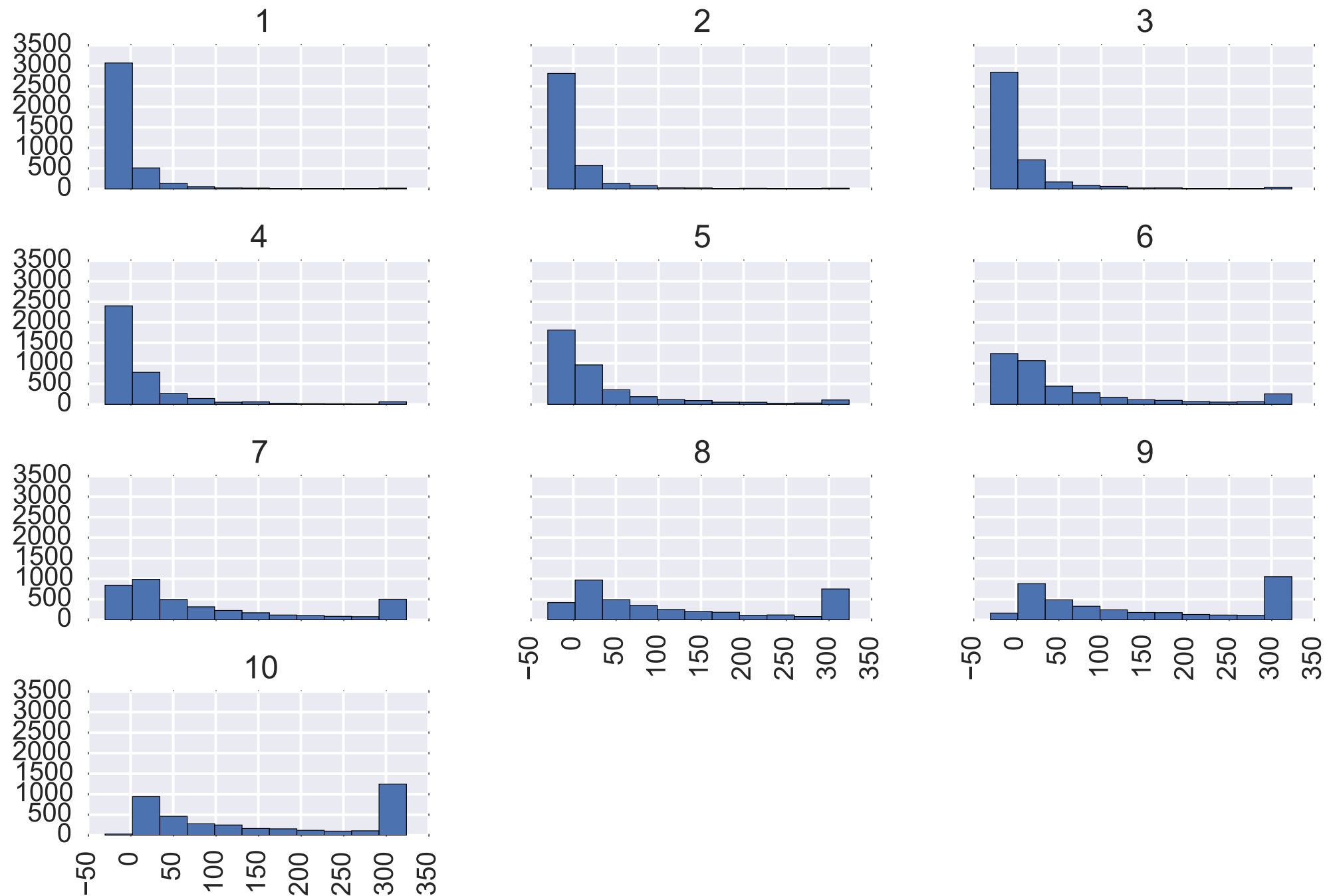
# Probability density of peak scores (when they intersect and they do not)

# Higher scoring histone peaks tend to be associated with higher scoring DHS peaks (but the relationship is not one-to-one)


Histogram of DHS peak (q-value) based on H3K27ac_1 ranking - better ranking => higher category

Similar trend with MF predictions (but 3 times more predictions)

Histogram of DHS peak (q-value) based on MF1 ranking - better ranking => higher category
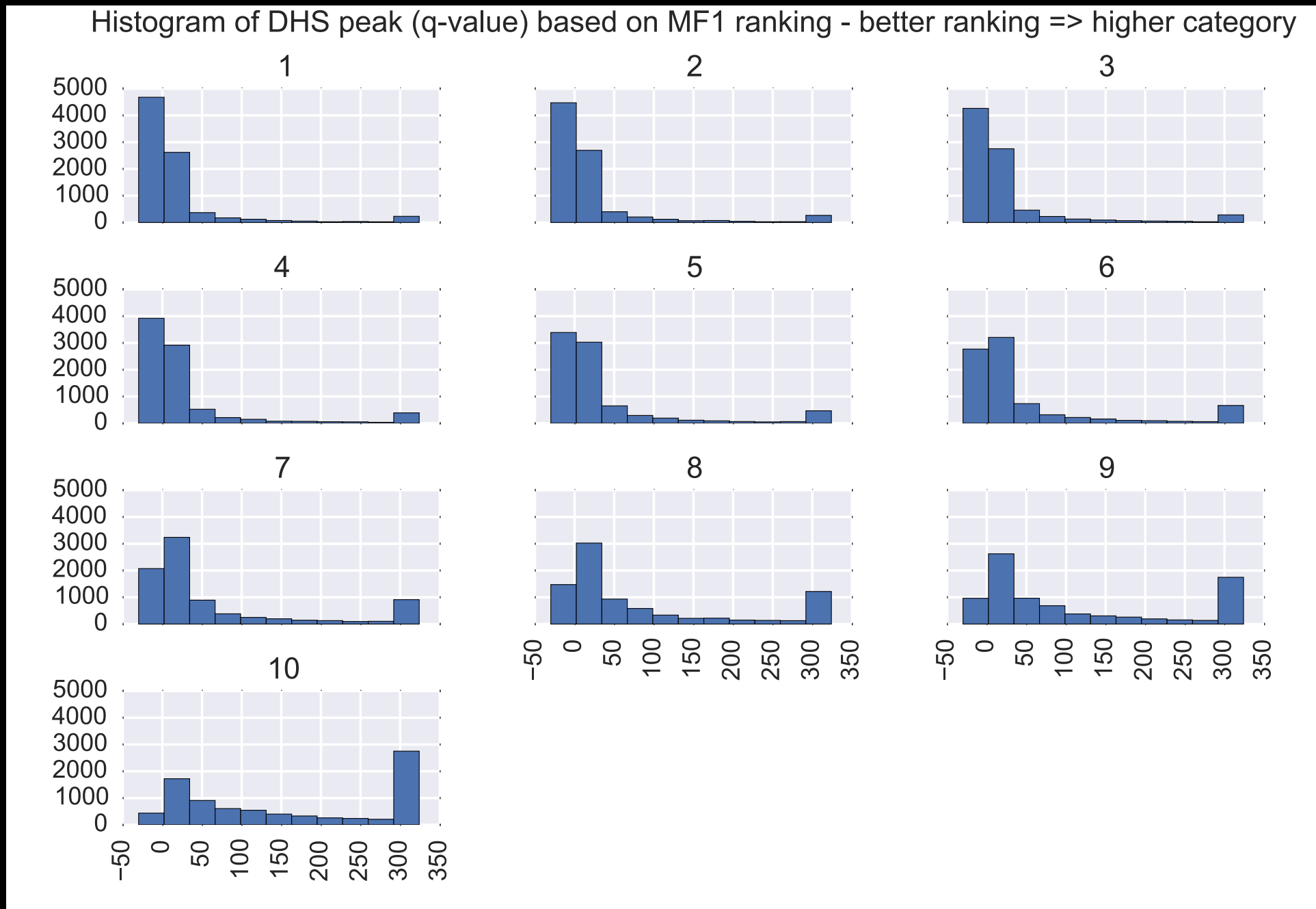
A lot of DHS peaks do not intersect with H3K27ac peaks (nearly double intersect with MF predictions but still a lot of DHS peaks out of MF predictions)

Histogram of DHS peaks not intersecting with H3K27ac_1 (145143 peaks)

Histogram of DHS peaks not intersecting with MF1 (112410 peaks)
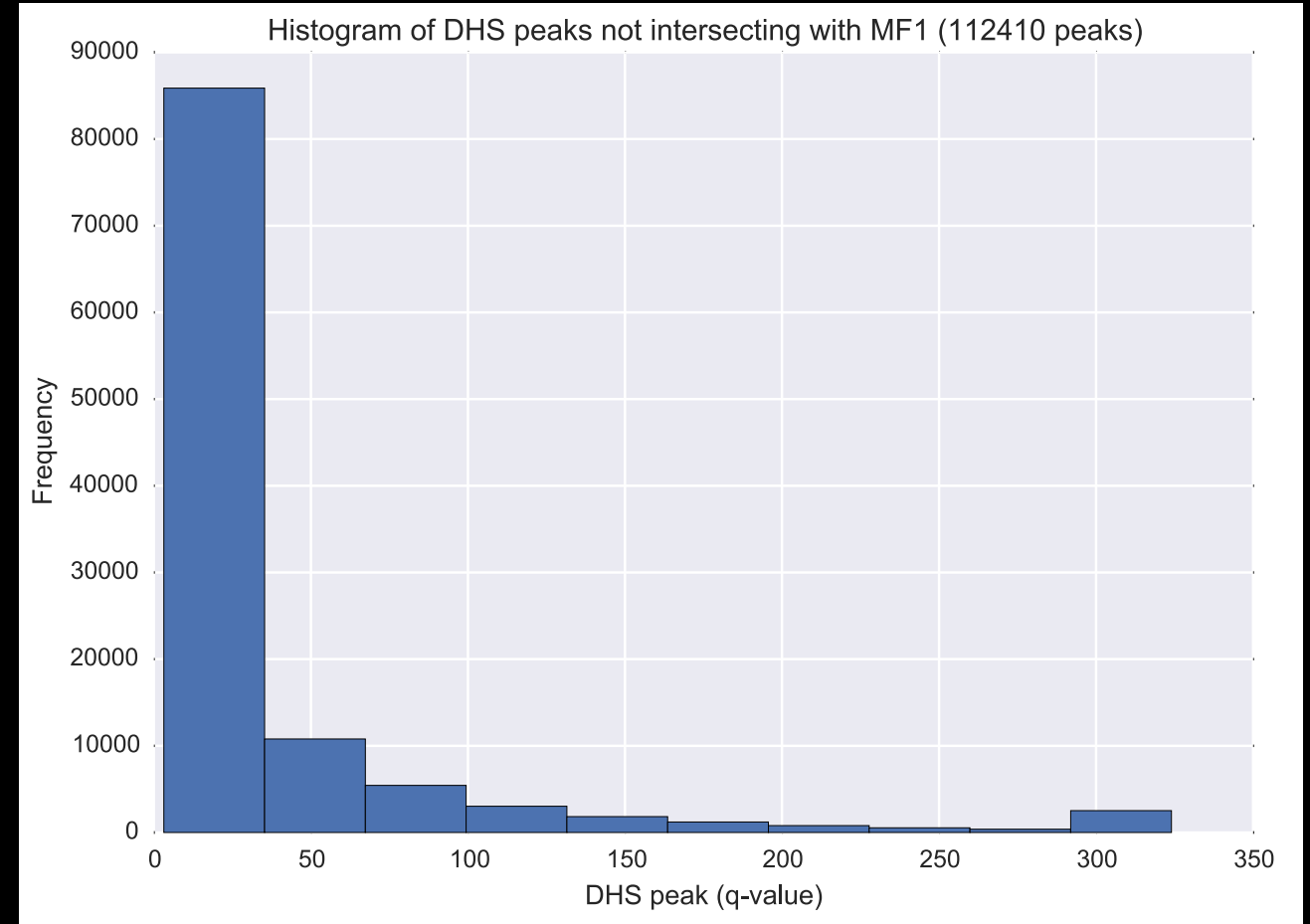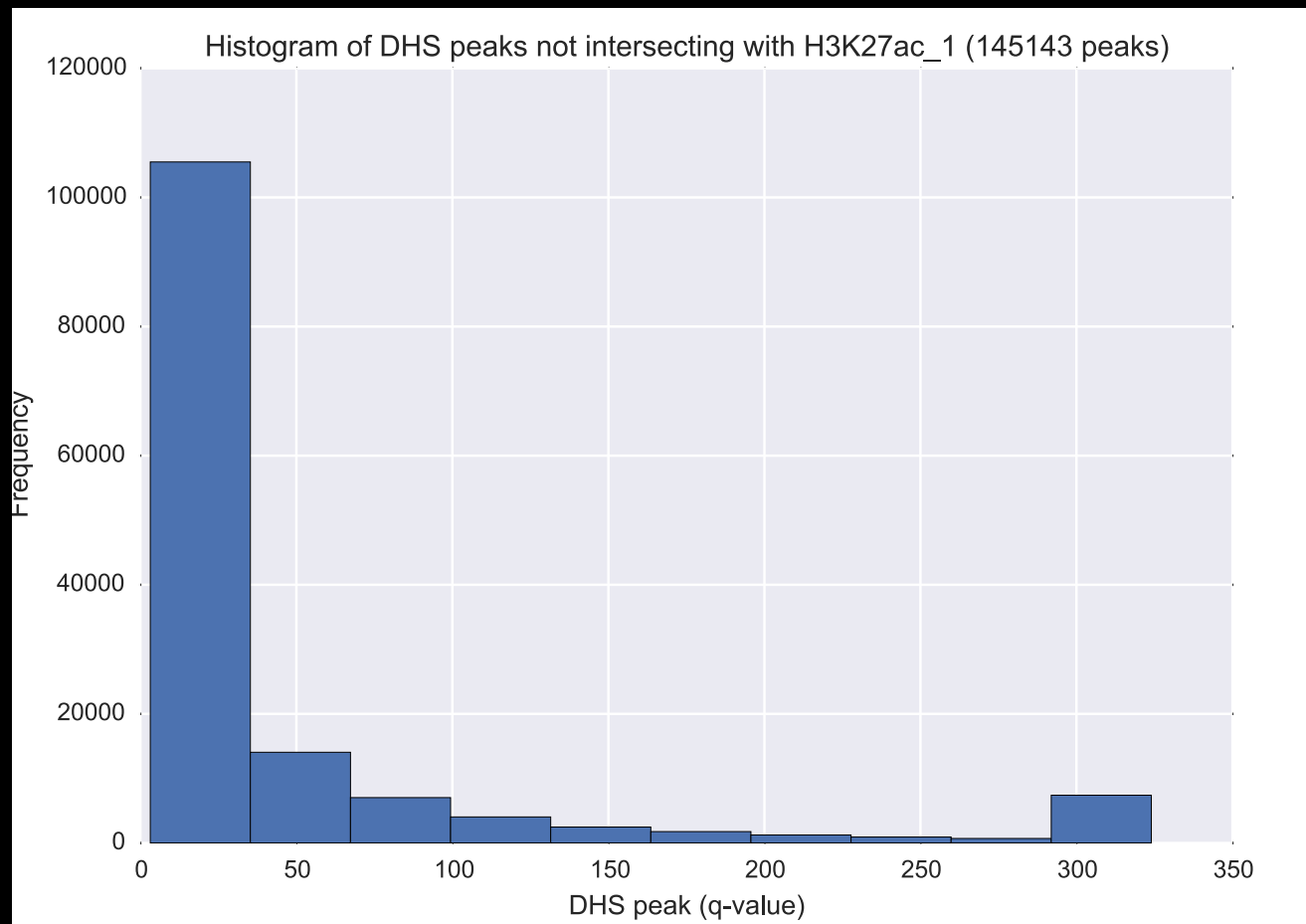
This is a slide with some trends but this is harder to be confident about - when looking at signal tracks
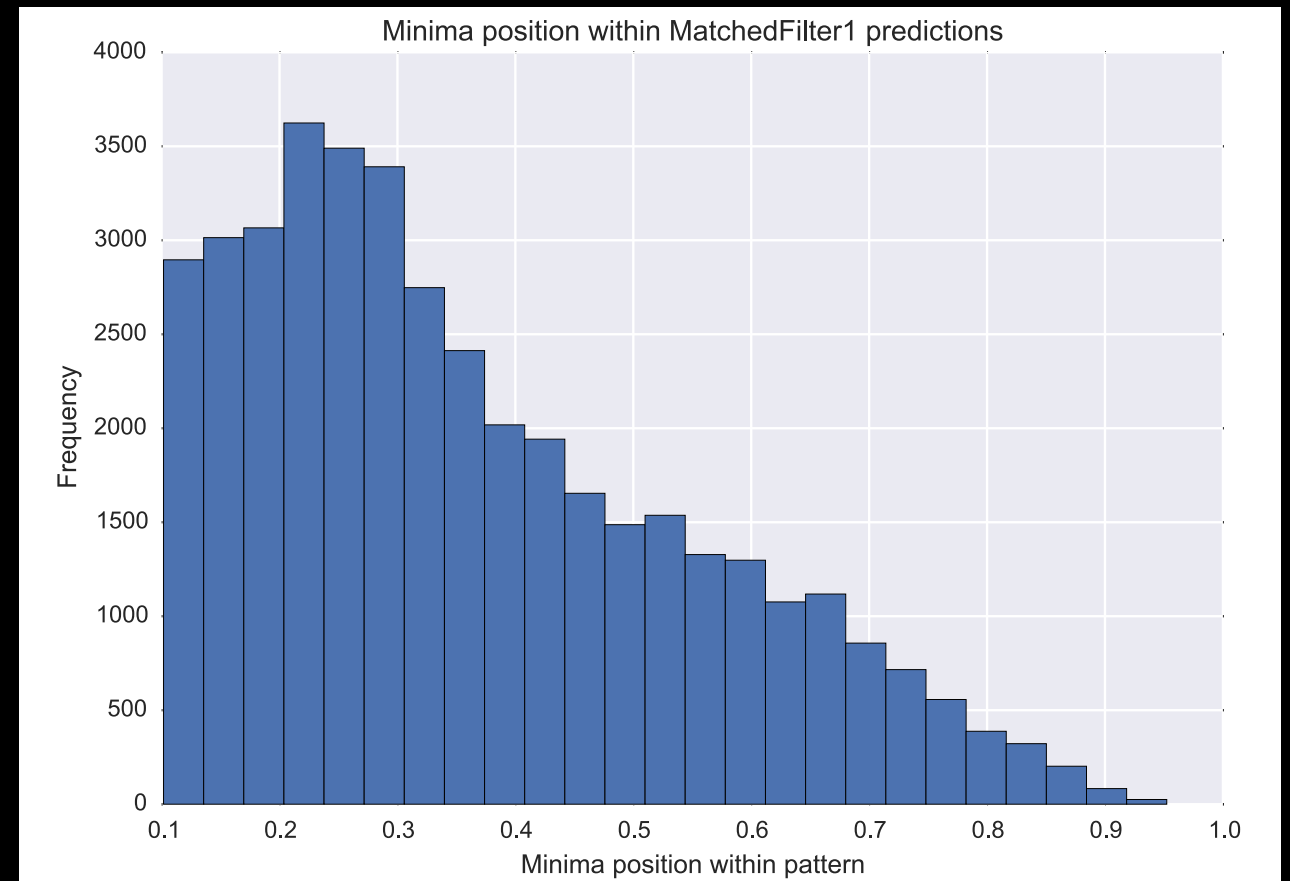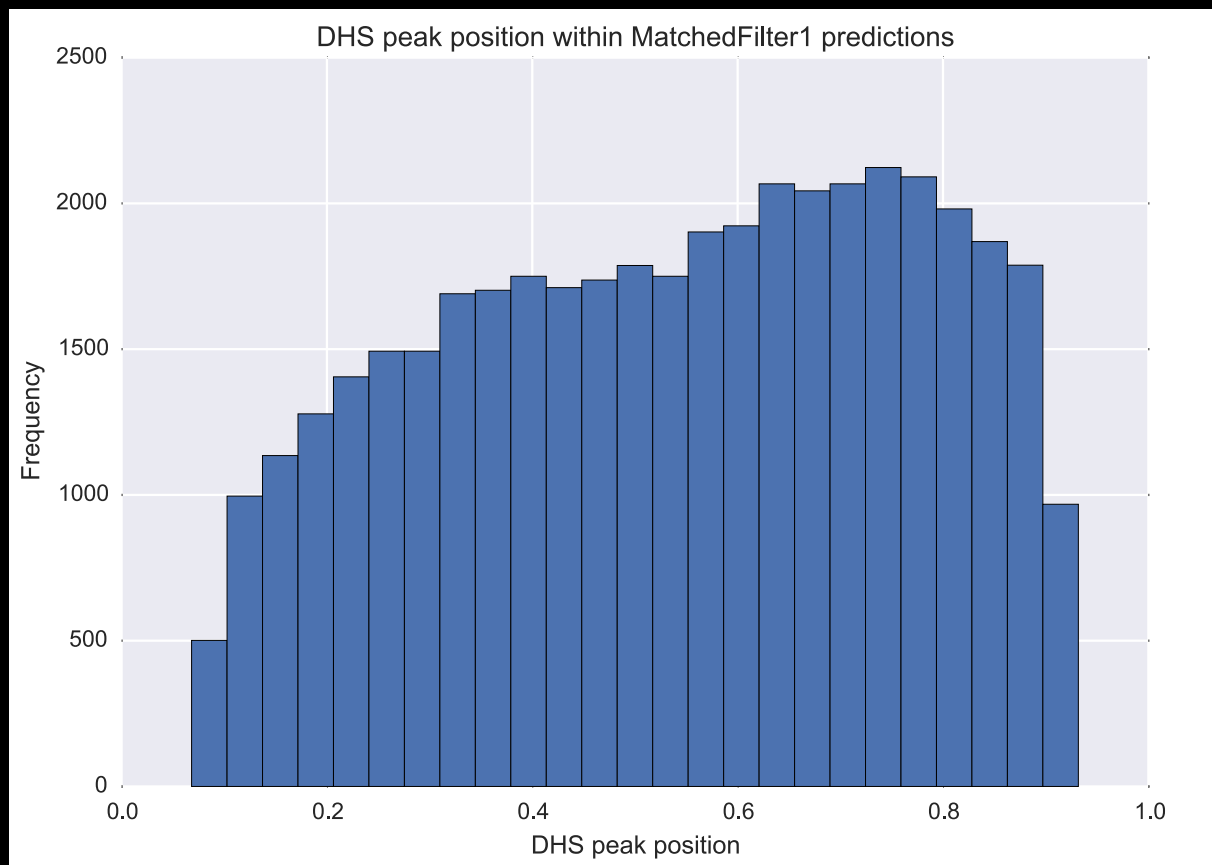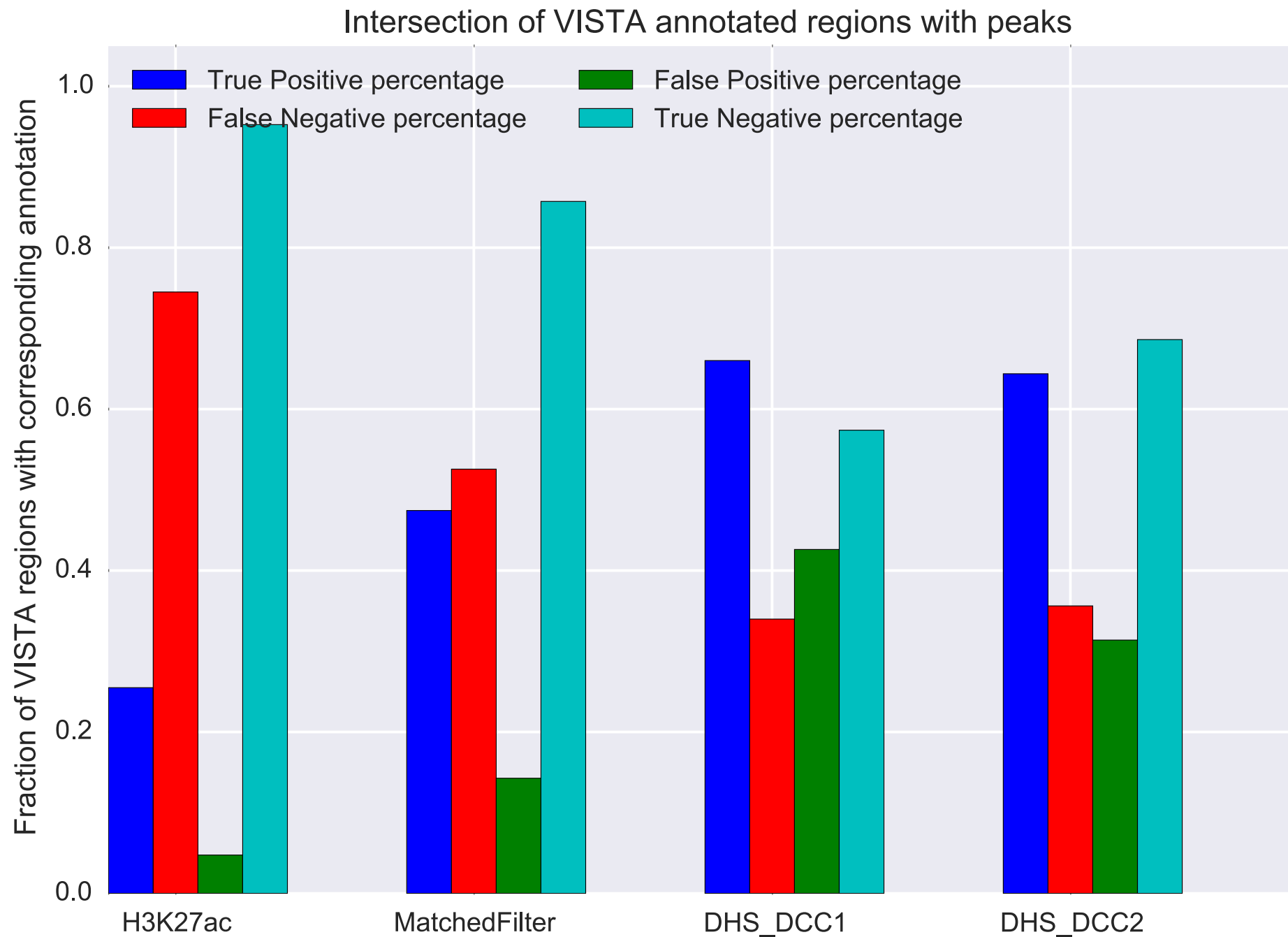

Hints towards creating a double matched filter over both signals simultaneously (but this will require time and we don't know how much we will gain in performance).
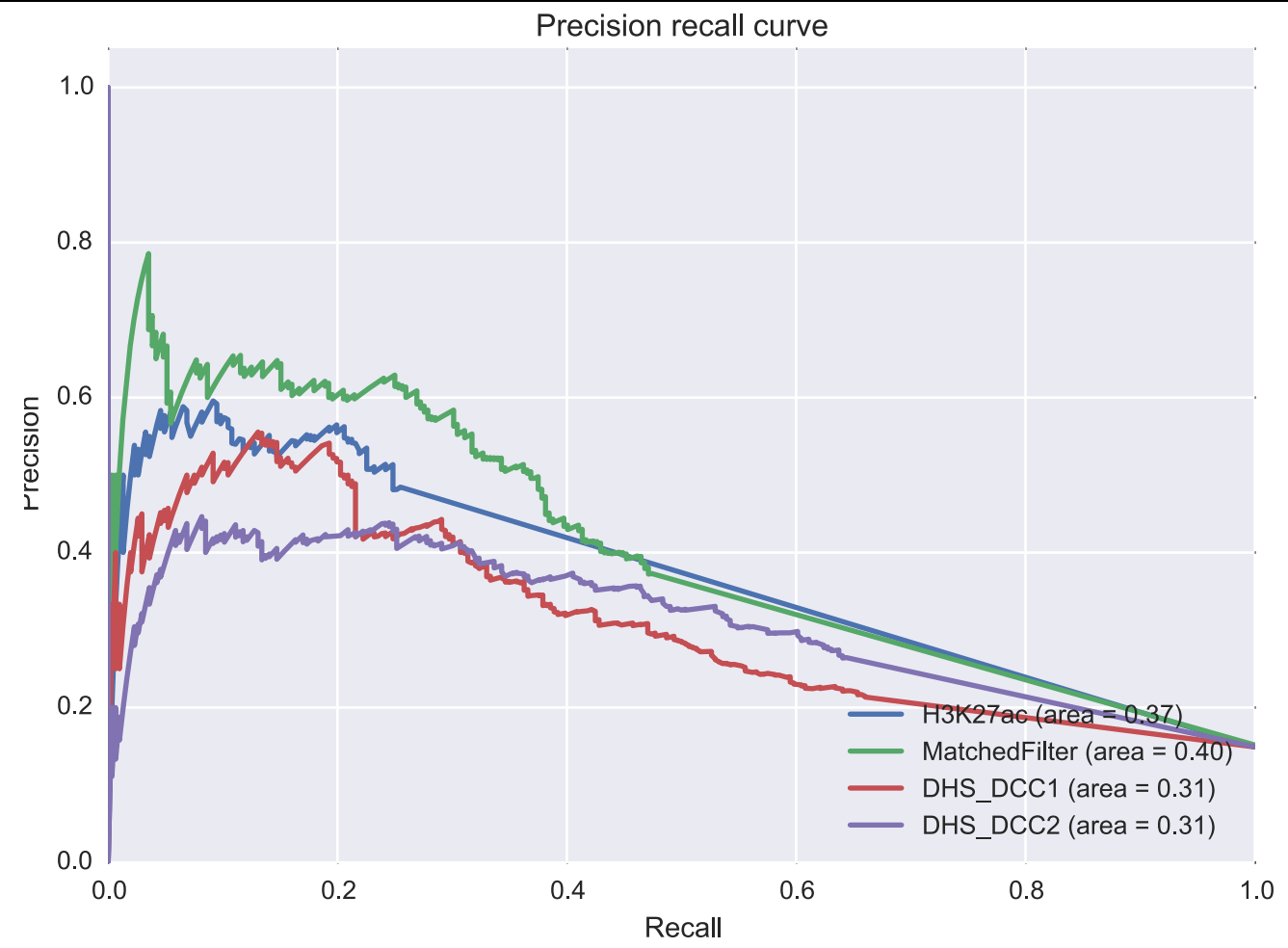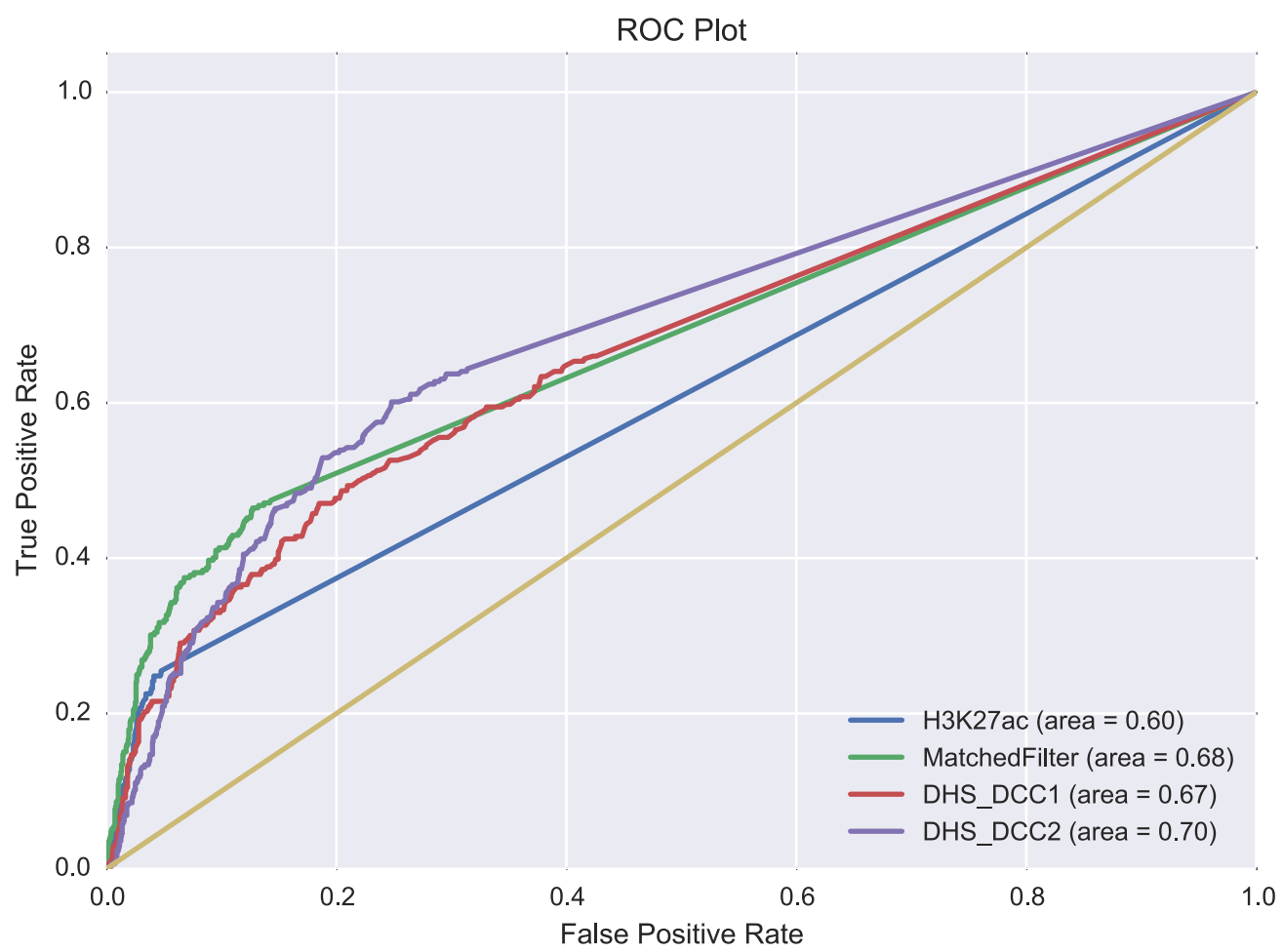Alternative proposal is to just ensemble rank the intersection of DHS peaks and MF predictions (quick solution but minimal improvement).

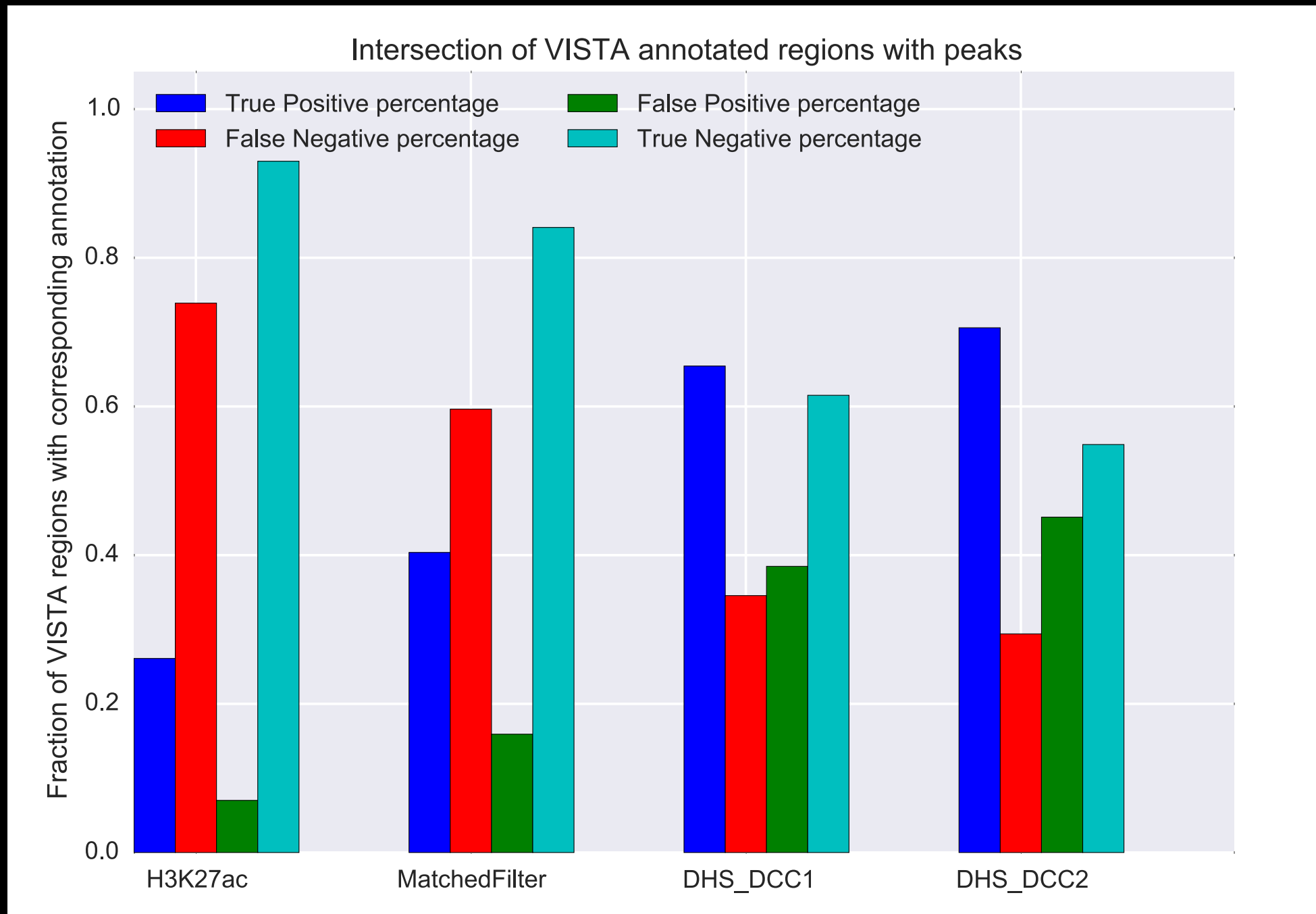Can look up figures for all 4 tissue but very similar results
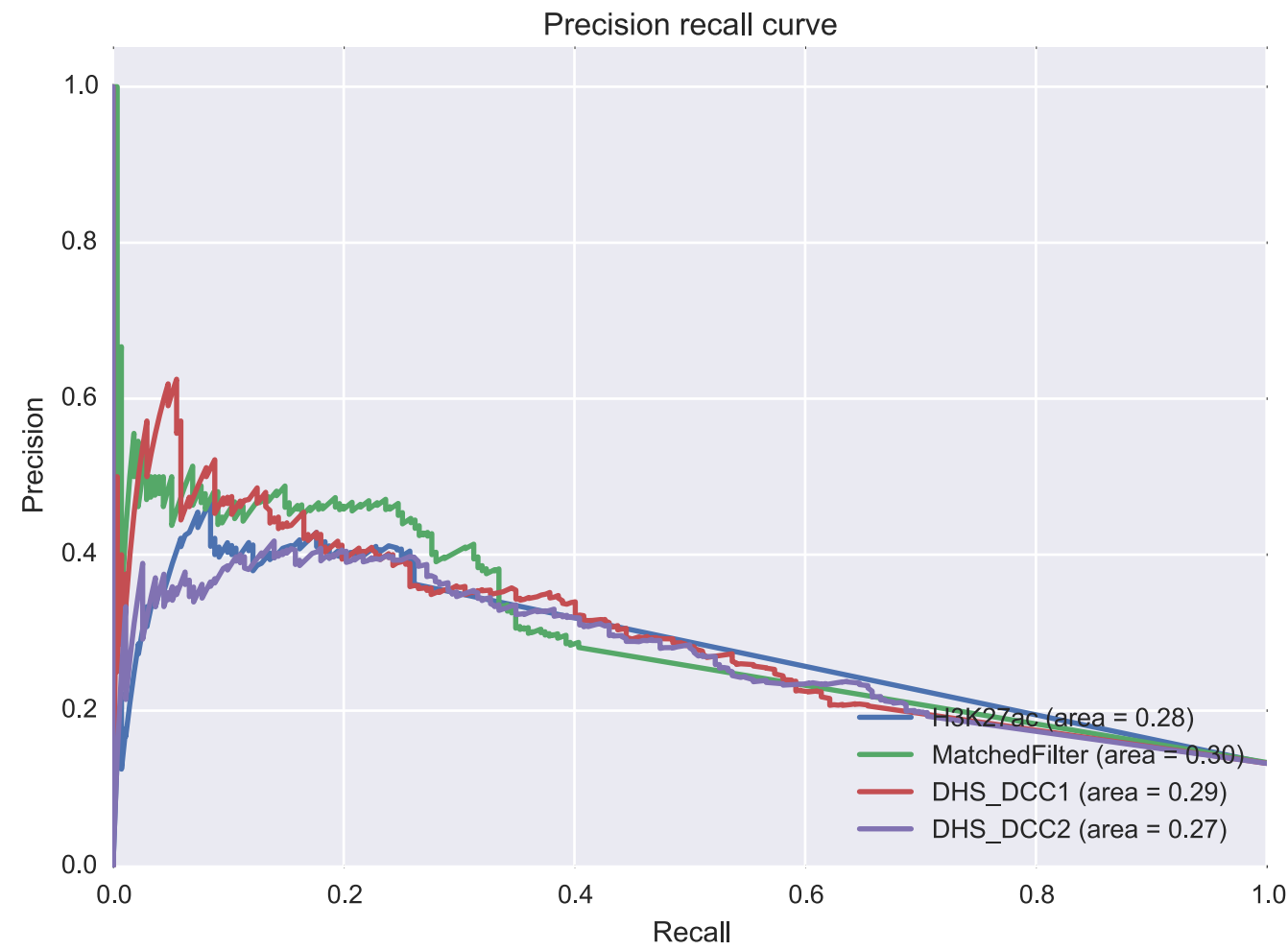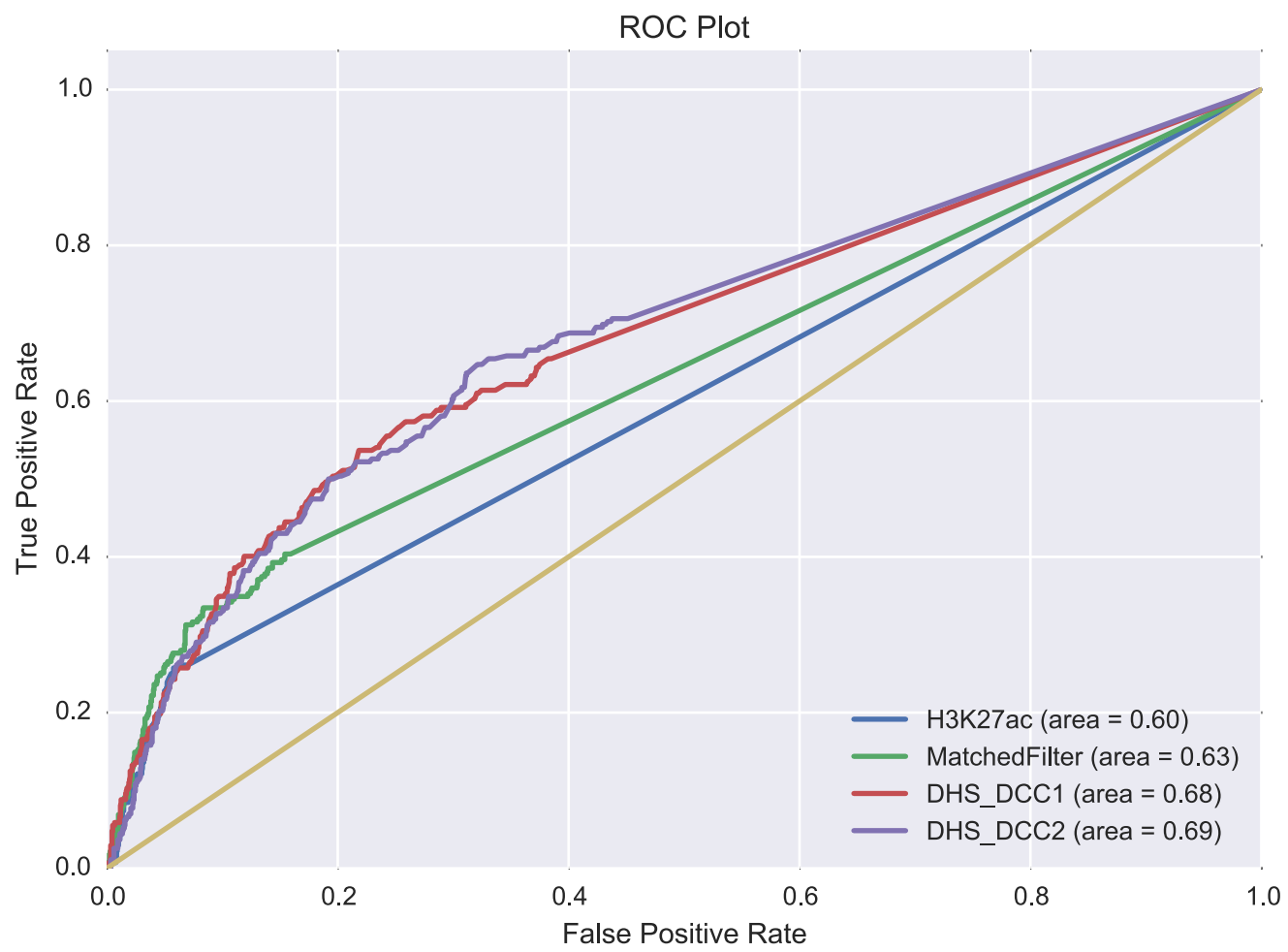
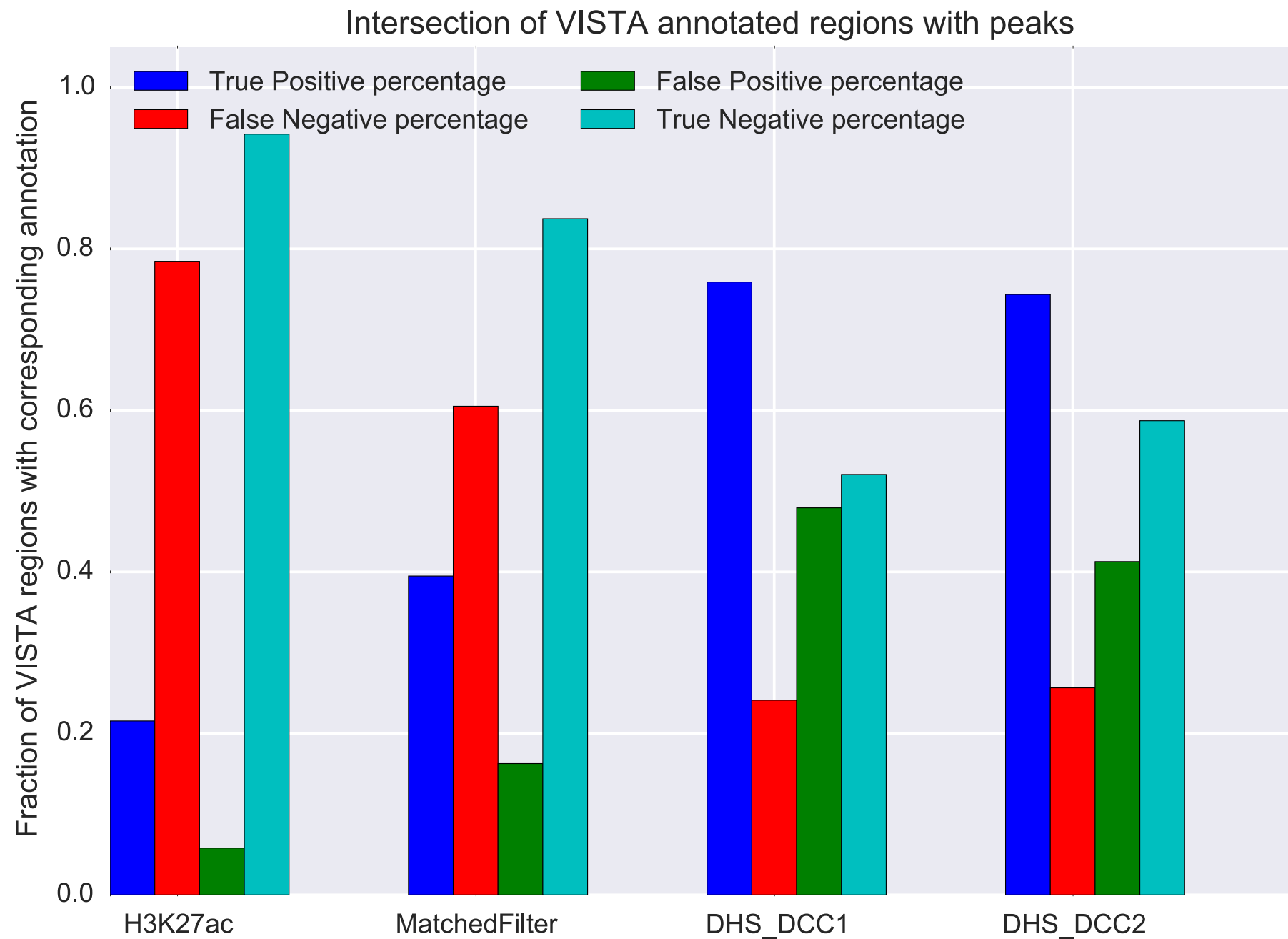Performance on midbrain
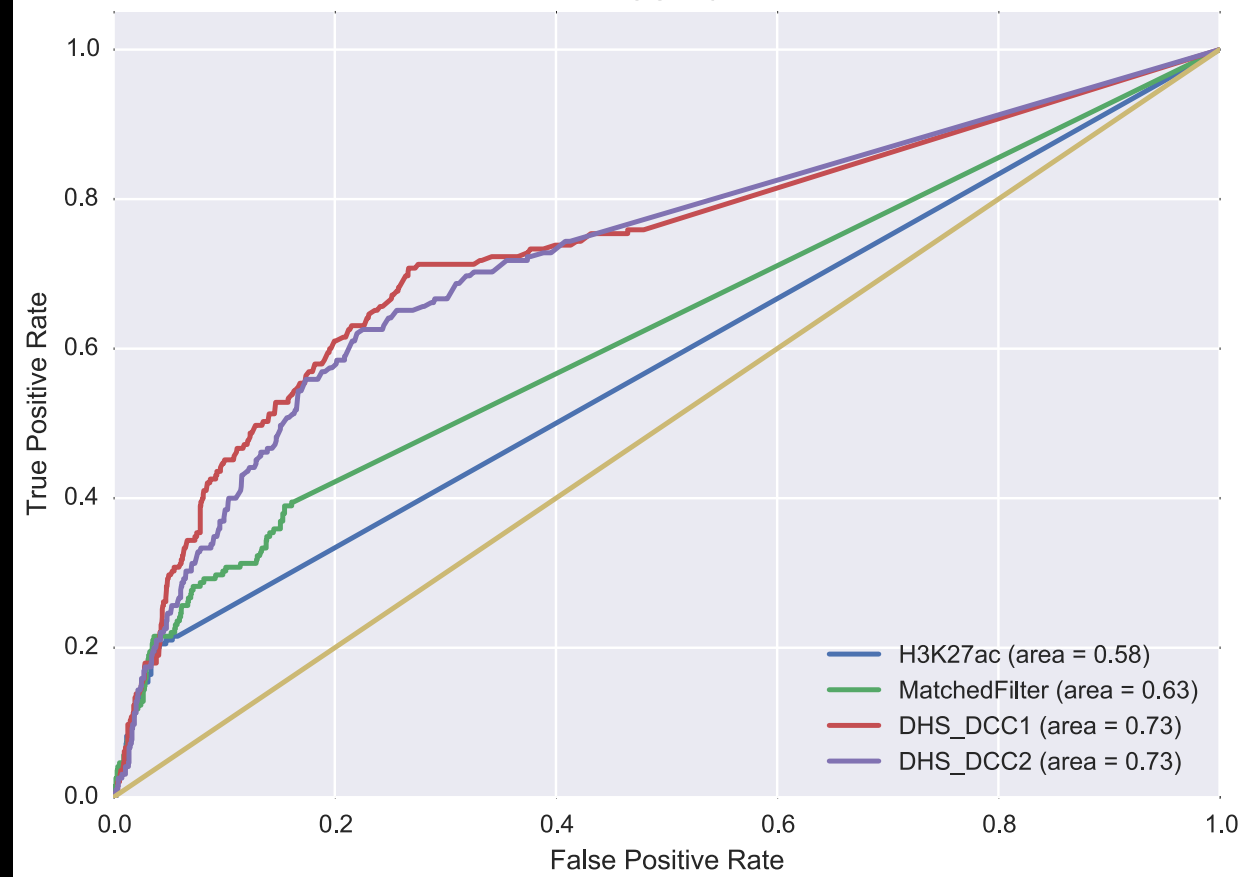
Performance on midbrain

Performance on hindbrain

Intersection of VISTA annotated regions with peaks

Performance on hindbrain

ROC Plot

H3K27ac (area = 0.60)
MatchedFilter (area = 0.63)
DHS_DCC1 (area = 0.68)
DHS_DCC2 (area = 0.69)

Precision recall curve

H3K27ac (area = 0.28)
MatchedFilter (area = 0.30)
DHS_DCC1 (area = 0.29)
DHS_DCC2 (area = 0.27)

Performance on neural tube

Intersection of VISTA annotated regions with peaks

Performance on neural tube
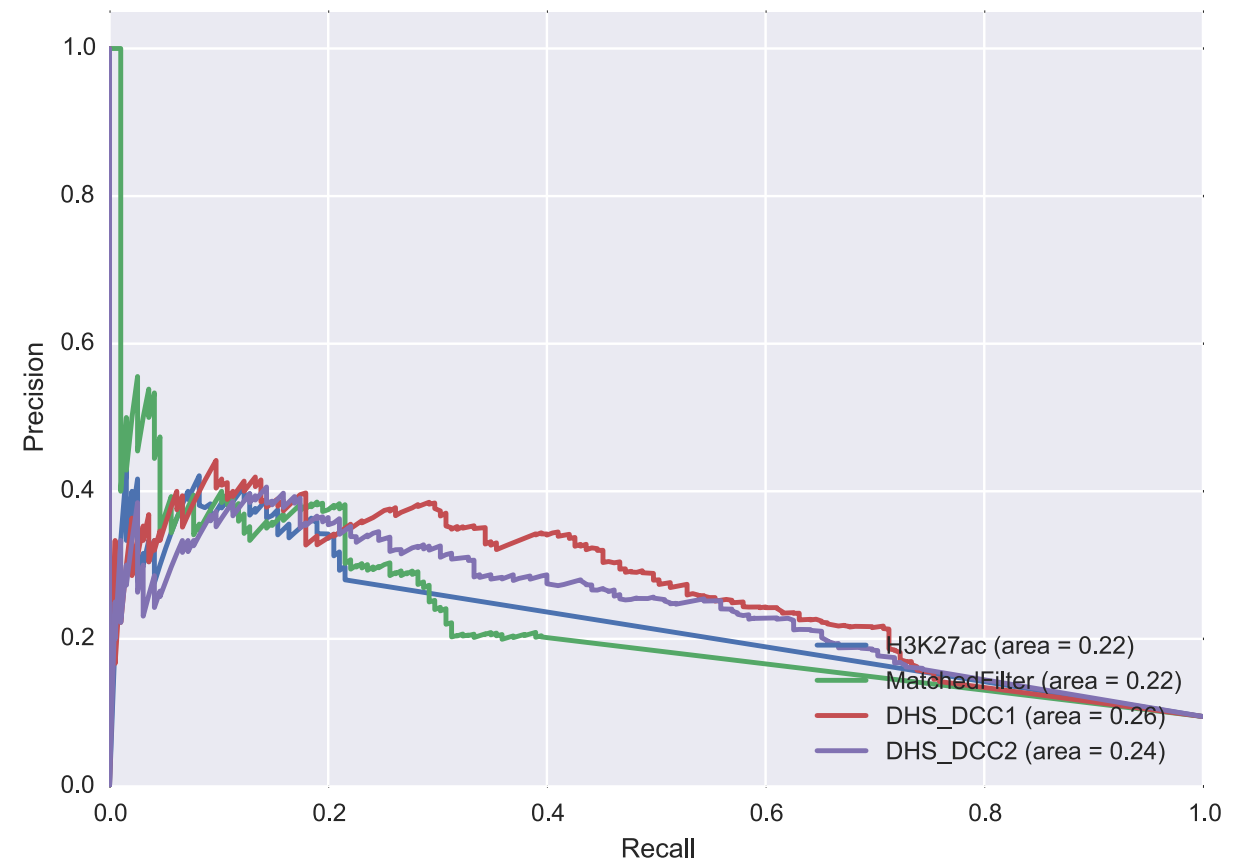
Performance on limb

Performance on limb