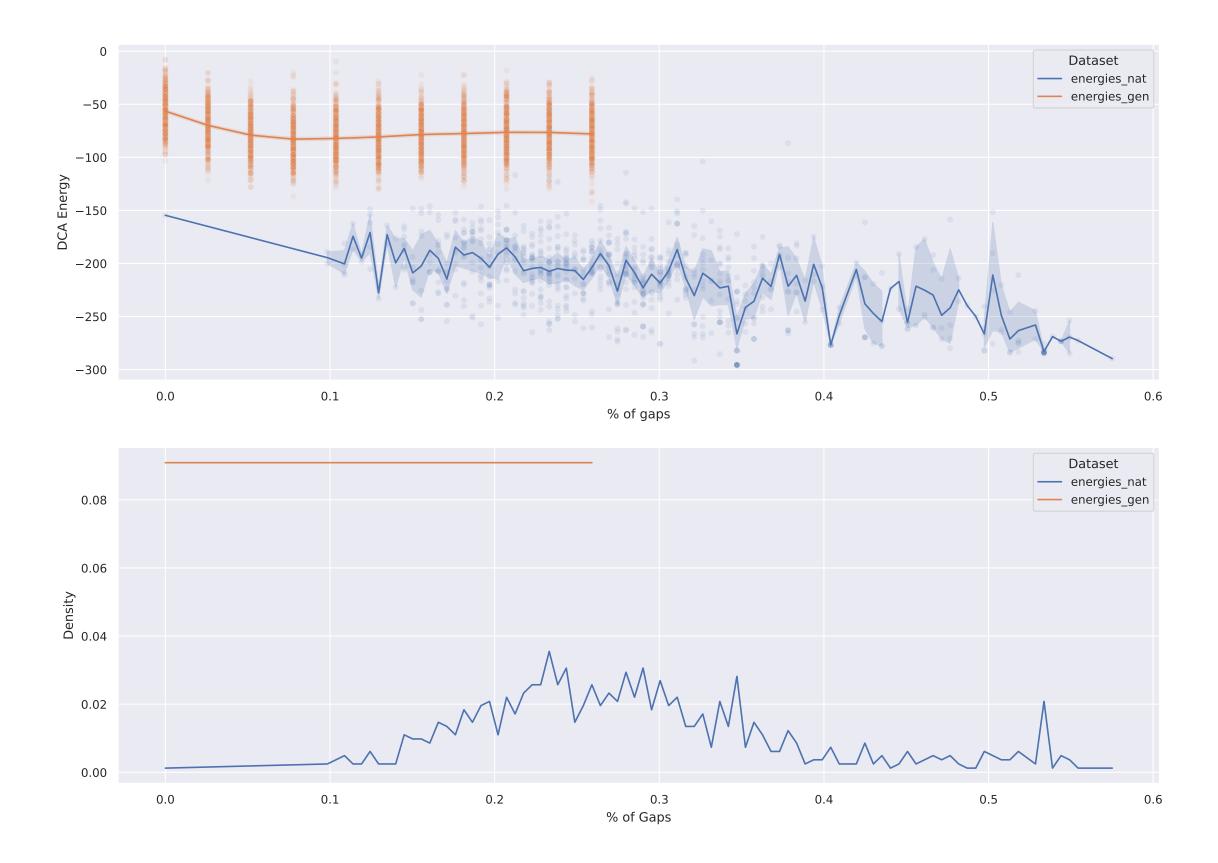


Jaccard distribution score depending on size with k=50.18 0.17 Jaccard distance score 0.15 0.14 150 170 160 180 190

Sequence size (nuc)



Similarity score between binfile and natural data

