Big Data Set : Biggest gap threshold = 6 | foldChange = 21.607 **Dimensions** 1000000 -2(Original data) threshold

Frequency of similarities Original Random 100000 -Frequency 1000 -10 -120 -Similarity