Big Data Set : Gap threshold = $54 \mid foldChange = -5.308$ 600000 -**Dimensions** 2(Original data) - 5 threshold

Frequency of similarities Original Random Frequency 1000 -10 -. 8 Similarity