Big Data Set : Gap threshold =  $79 \mid foldChange = -1.824$ 5500 -5000 4500 **Dimensions** 2(Original data) <del>--</del> 5 4000 3500 3000 70 65 60 . 75 80 threshold

Frequency of similarities Original Random Frequency 1000 -10 · Similarity