Use your functions to calculate the pairwise Jaccard similarity for this set of 4 experiments, then answer the following questions:

1. Which two interval sets are the most similar? (10 pts)

CTCF peaks from HeLaS3 and HepG2 are the most similar.

1. Which two interval sets are the most different? (10 pts)

HelaS3 Jun peaks and HepG2 CTCF peaks are the most different.

1. Based on these results, which factor, CTCF or Jun, would you predict varies more across cell types? (10 pts)

Jun varies more across cell types. The two CTCF peaks have much higher Jaccard score which indicates more overlap of peaks than Jun among different cell types.

1. Based on these results, do the genomic locations found by ChIP-seq experiments depend more on the cell-type, or on the transcription factor being assayed? (20 pts)

The genomic locations found by ChIP-seq experiments depend more on the transcription factor. Because from the pairwise Jaccard scores, we know that the two CTCF peaks have the highest Jaccard score, means the peaks of one transcription factor from different cell types are more consistent than the different transcription factors from one cell type. So, even in different cell types, a transcription factors tends to bind on the same locations. However, the binding locus consistency varies in different transcription factors. Jun binding sites is not as consistent as CTCF in different cell types.