

I have downloaded some files from <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE63525> there are lots of intrachromosomal data in there, corresponding to (I think!) 11 cell lines. During the correspondence that I had with University of Missouri (Mr. Oluwadare), he told me that I have to use the GM12878 data set in there, which is much larger than the original [?] dataset. There are two inter- and intra-chromosomal data sets there:

1. GSE63525\_GM12878\_primary\_interchromosomal\_contact\_matrices.tar.gz  
GSE63525\_GM12878\_primary\_intrachromosomal\_contact\_matrices.tar.gz
2. GSE63525\_GM12878\_combined\_interchromosomal\_contact\_matrices.tar.gz  
GSE63525\_GM12878\_combined\_intrachromosomal\_contact\_matrices.tar.gz

I wonder which ones I should use at the moment.

## **Things to consider**

1. effect of multiple pearson normalizations
2. revised SCN algorithm