Main Tasks

* Obtain and process data from Gene Expression Omnibus
* Understand and copy Kingsford’s methodology & code
  + Read the readme file
  + Compute similarity scores (relative entropy metric)
  + Repeat this for all candidate pairwise region in genome
  + Statistics, p-model, giving a p-value for each each TAD across experimental conditions
  + Produce some visualizations...
  + Report

java -jar juicer\_tools\_1.22.01.jar dump observed KR/your\_data.hic chr1 chr1 BP 1000000 > output.txt

C:\Users\simon\OneDrive\CMU MSCB\Spring 2024\bioinfo\project\data\GSE237722\_ENCFF203NXG\_haplotype-specific\_contact\_matrix\_GRCh38.hic.hic

java -jar juicer\_tools\_1.22.01.jar dump observed KR ../data/GSE237722\_ENCFF203NXG\_haplotype-specific\_contact\_matrix\_GRCh38.hic chr1 chr1 BP 1000000 > output.txt

Most recent one below

java -jar juicer\_tools\_1.22.01.jar dump ../data/GSE237722\_ENCFF203NXG\_haplotype-specific\_contact\_matrix\_GRCh38.hic chr1 chr1 BP 1000000 > ../data/output.txt

java -jar juicer\_tools.1.6.2\_windows\_jcuda.0.8.jar dump ../data/GSE237722\_ENCFF203NXG\_haplotype-specific\_contact\_matrix\_GRCh38.hic chr1 chr1 BP 1000000 > ../data/output2.txt

java -jar juicer\_tools.1.6.2\_windows\_jcuda.0.8.jar dump observed ../data/GSE237722\_ENCFF203NXG\_haplotype-specific\_contact\_matrix\_GRCh38.hic chr1 chr1 BP 1000000 > ../data/output2.txt

java -jar juicer\_tools.1.6.2\_windows\_jcuda.0.8.jar dump observed ../data/GSE237722\_ENCFF776XCM\_contact\_matrix\_GRCh38.hic chr1 chr1 BP 1000000 > ../data/output2.txt

java -jar juicer\_tools.1.6.2\_windows\_jcuda.0.8.jar dump ../data/ENCFF314KSN.hic chr1 chr1 BP 1000 > ../data/output3.txt

java -jar juicer\_tools.2.20.00.jar dump ../data/ENCFF314KSN.hic chr1 chr1 BP 1000 > ../data/output3.txt

java -jar juicer/juicer\_tools.1.6.2\_windows\_jcuda.0.8.jar dump data/ENCFF314KSN.hic chr1 chr1 BP 1000 > data/output4.txt

java -jar juicer/juicer\_tools.1.6.2\_windows\_jcuda.0.8.jar dump observed KR data/ENCFF314KSN.hic chr3 chr3 BP 100 > data/output5.txt

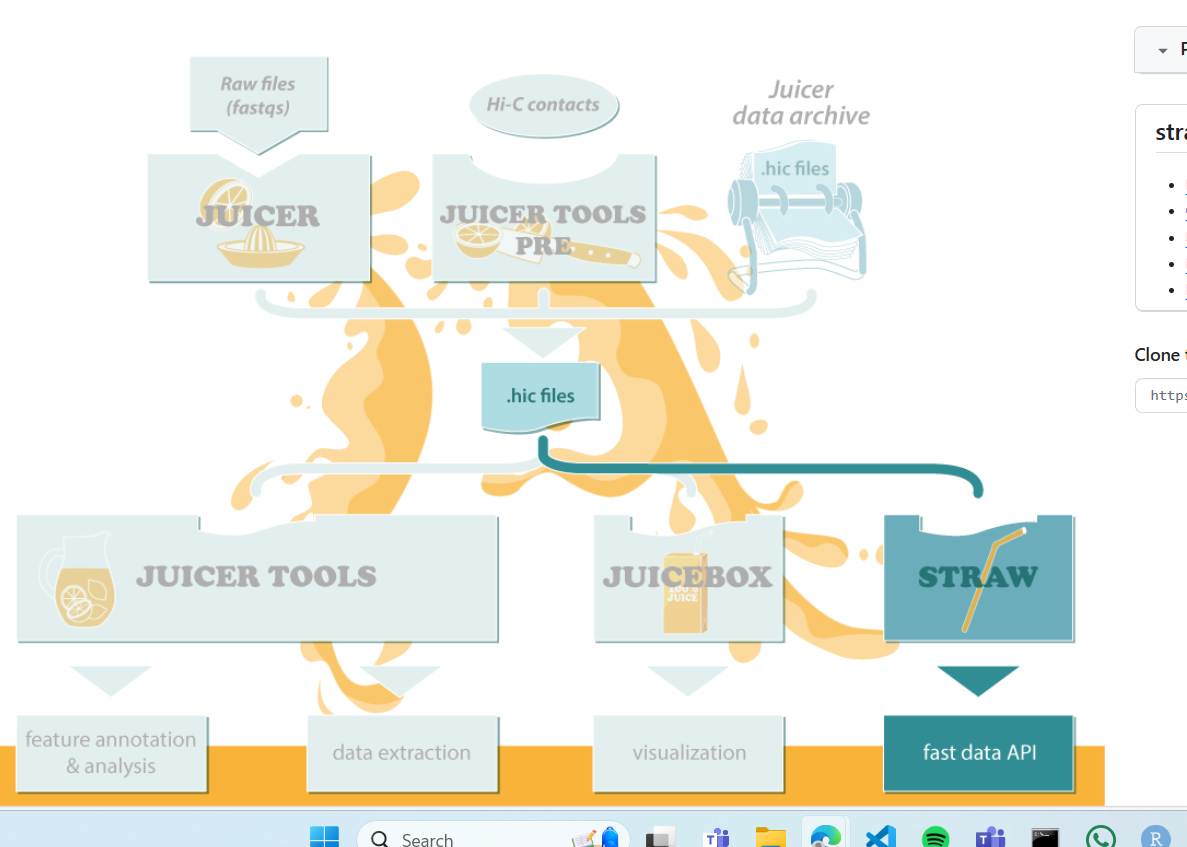
Installing HiCExplorer for windows [deeptools/HiCExplorer: HiCExplorer is a powerful and easy to use set of tools to process, normalize and visualize Hi-C data. (github.com)](https://github.com/deeptools/HiCExplorer)

hicConvertFormat -m matrix.hic --inputFormat hic --outputFormat cool -o matrix.cool --resolutions 10000

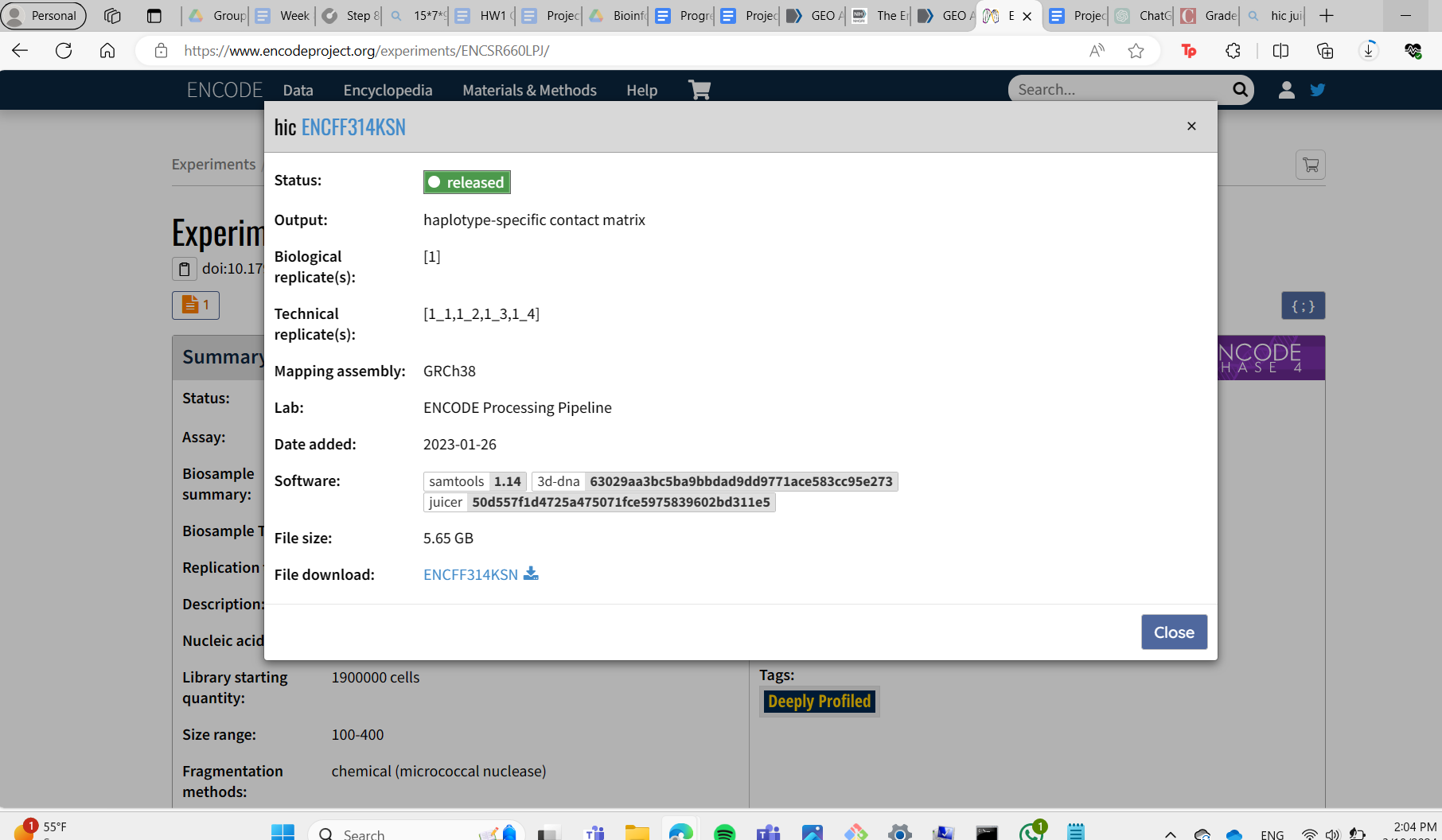
hicConvertFormat -m data/ENCFF314KSN.hic --inputFormat hic --outputFormat cool -o data/matrix.cool --resolutions 10000

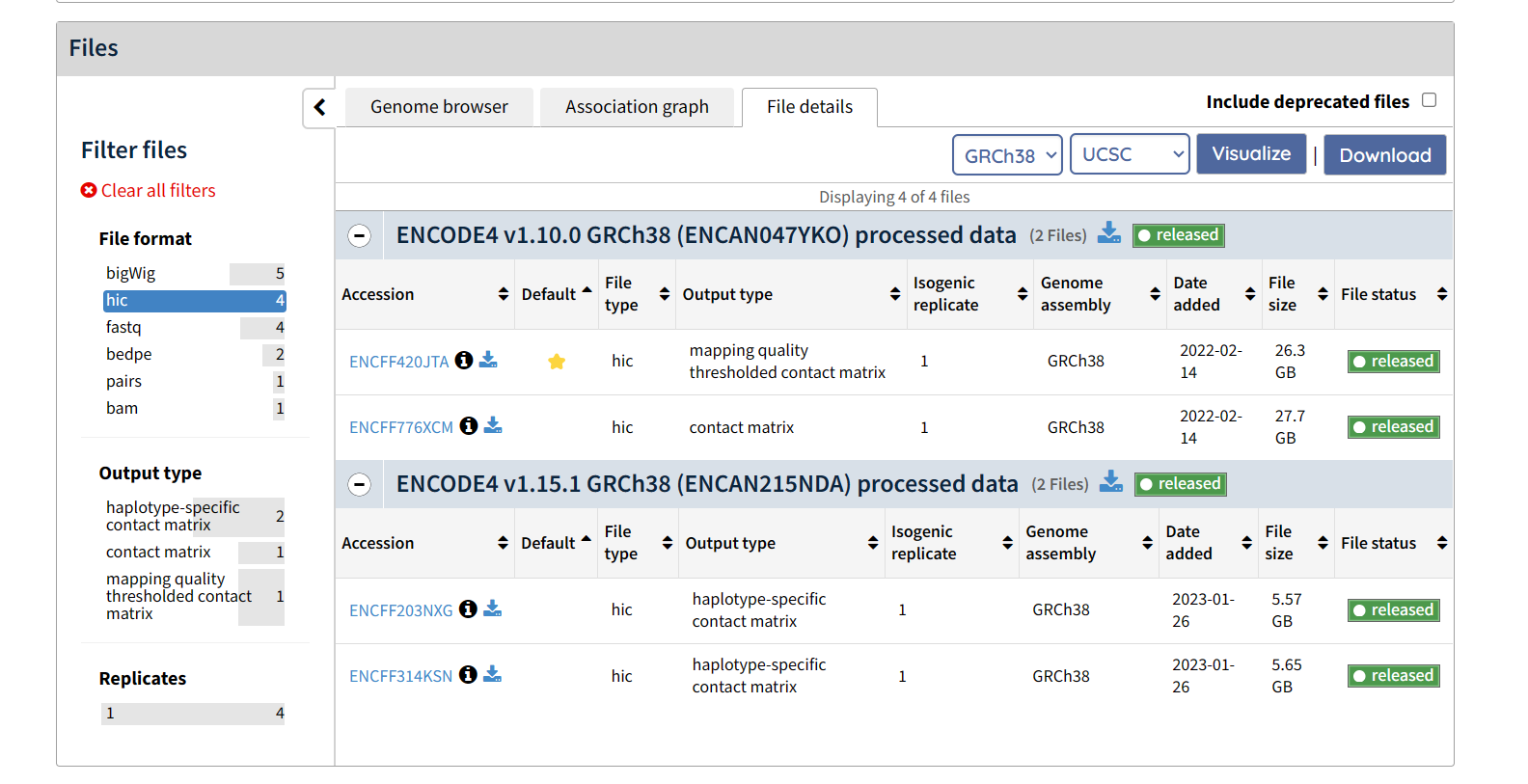
hicConvertFormat -m ../data/ENCFF314KSN.hic --inputFormat hic --outputFormat cool -o ../data/matrix.cool --resolutions 10000

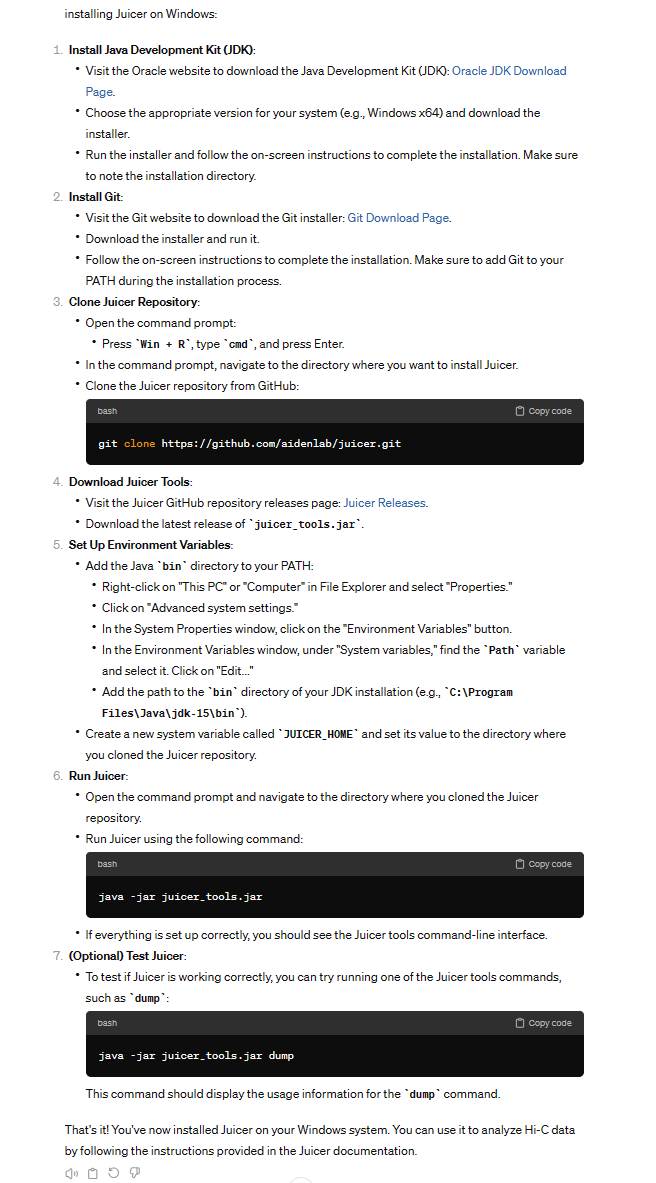
* Go their Github page
* Clone it, cd into it, then py setup.py install
* Then import ‘HiCExplorer’, should work...



Links

* [GEO Accession viewer (nih.gov)](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE237722)
* [GEO Accession viewer (nih.gov)](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM7646184)
* [The Hitchhiker's Guide to Hi-C Analysis: Practical guidelines - PMC (nih.gov)](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4347522/)
* [ENCSR660LPJ – ENCODE (encodeproject.org)](https://www.encodeproject.org/experiments/ENCSR660LPJ/)
  + 





| ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012 Sep 6;489(7414):57-74. PMID: [22955616](https://www.ncbi.nlm.nih.gov/pubmed/22955616) |
| --- |
|  |

