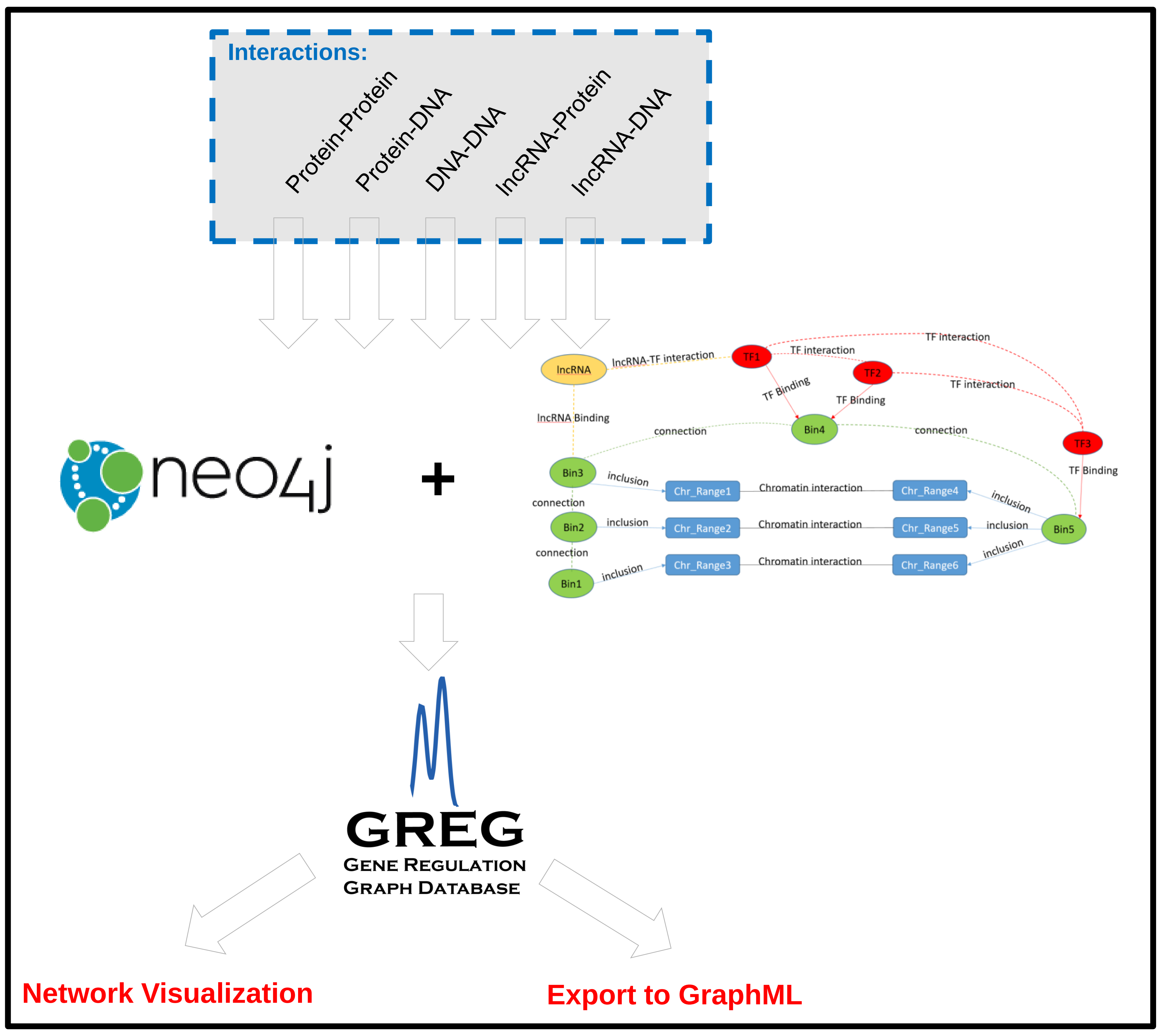
GREG
GENE REGULATION
GRAPH DATABASE

Studying transcriptional regulation using Graph Databases

Songqing Mei, Chengshu Xie, and Antonio Mora
Joint School of Life Sciences, Guangzhou Medical University and GIBH (Chinese Academy of Sciences)

GREG (The Gene Regulation Graph Database) is a graph database written in the Neo4J language.

Its goal is to allow genomic researchers to see in a graphical way all the known **interactions between proteins, lncRNAs, and DNA for a given transcription factor, lncRNA, or a given genomic region** (i.e., its "regulatory landscape").



With GREG...

1. **Find all the interactions for your query TF, lncRNA, or genomic region**! (No need to collect and merge information from different databases).
2. **Get the output as a graph**! (See the complete picture of your query regulation mechanism as a graph. Export the result as GraphML format).
3. **Build advanced graphical queries**! (Ask GREG to **extend** your query to the neighbors of a given node, or ask for the **shortest path** between two given nodes, as well as several other graph-related queries).
4. **Get networks for either one cell type or multiple cell types**! (explore differences between cell types or infer missing data).

EXAMPLES

Step1 Choose Cell lines(Default:All Cell Lines)

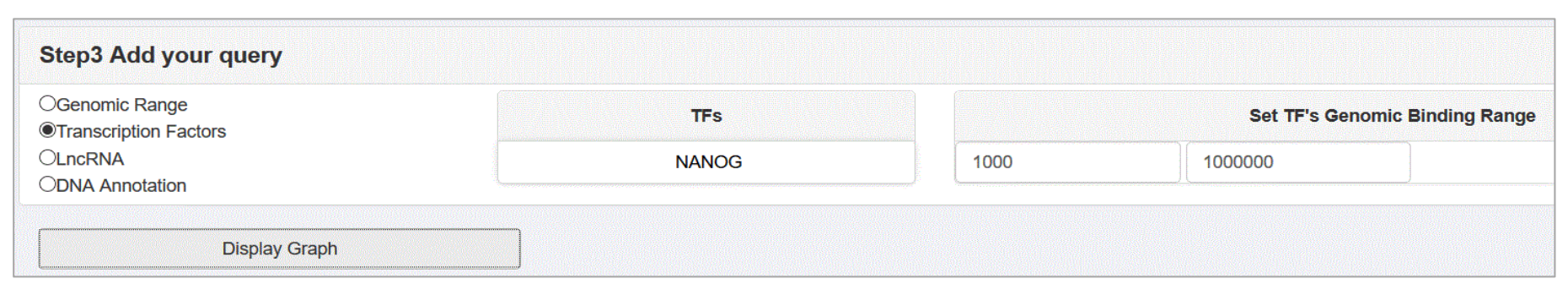
Chosen:H1ESC

Cell Type	Cell Name
All	H1ESC
Stem Cells	IPS6.9
Cancer Cells	IPS19.11
Other	

Step2 Choose chromosome (Default: Chr1 with small bins)

Genome Bin size	Chromosome
Small Bins(2kbp)	chr1

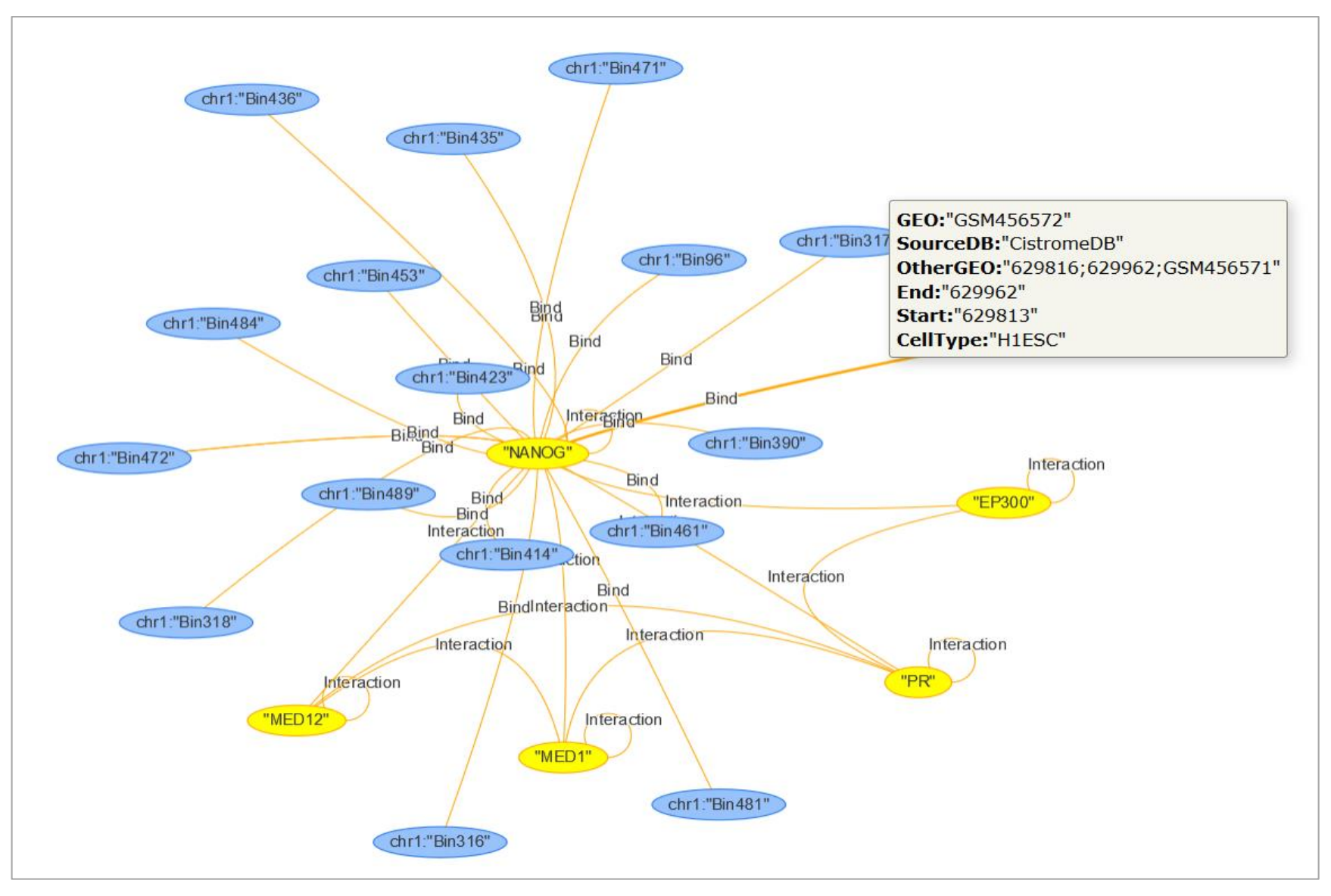
Find interactions for your query TF or lncRNA



Step3 Add your query

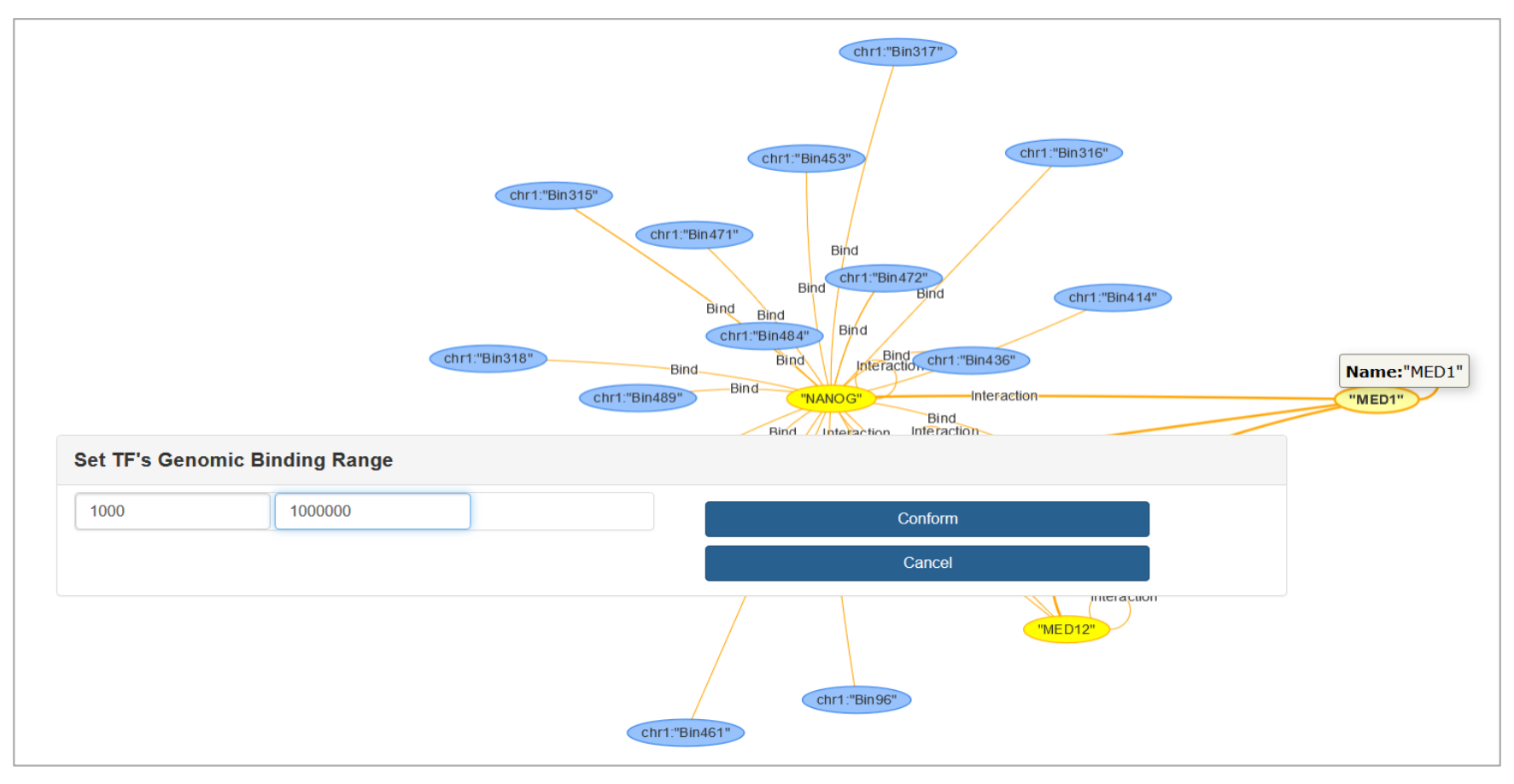
○Genomic Range
●Transcription Factors
○LncRNA
○DNA Annotation

TFs	Set TF's Genomic Binding Range	
NANOG	1000	1000000



GEO:"GSM456572"
SourceDB:"CistromeDB"
OtherGEO:"629816;629962;GSM456571"
End:"629962"
Start:"629813"
CellType:"H1ESC"

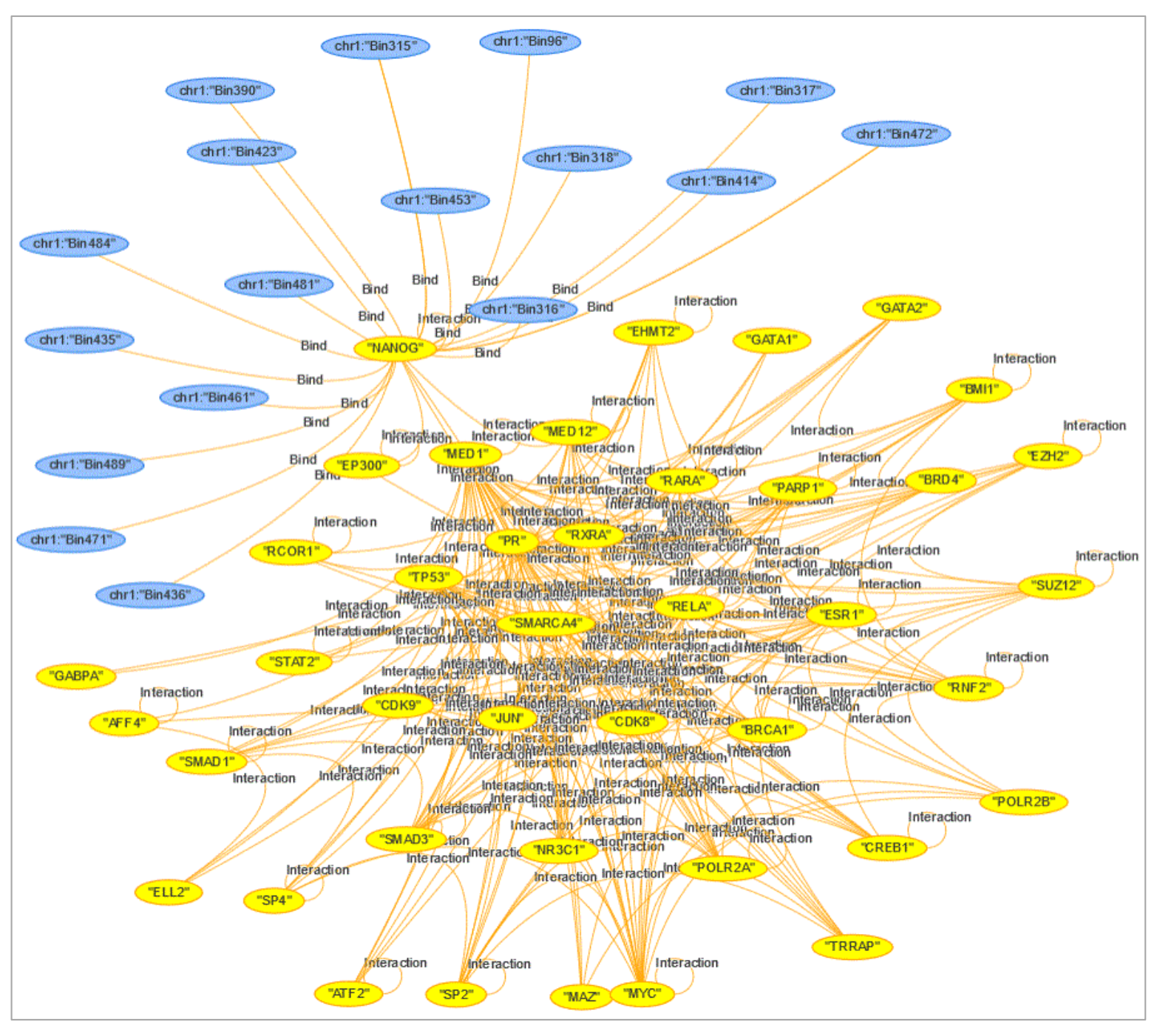
Select a transcription factor and the genomic region of your interest. Get the network with all the interacting TFs, lncRNAs, and DNA ranges for that region.



Name:"MED1"

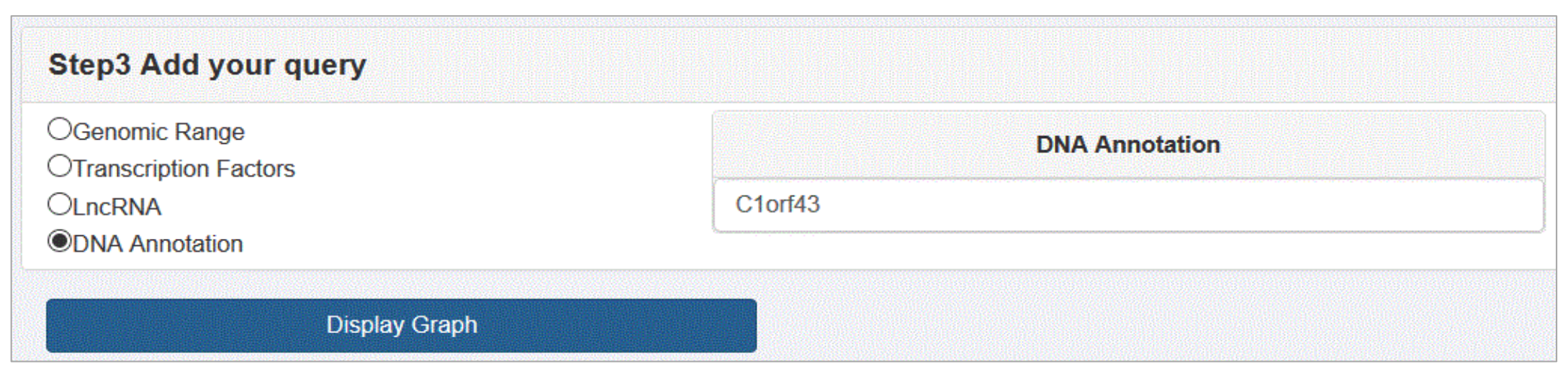
Set TF's Genomic Binding Range

1000 | 1000000



Expand your search: Double-click on any node and find its interactors.

Find interactions for DNA annotations

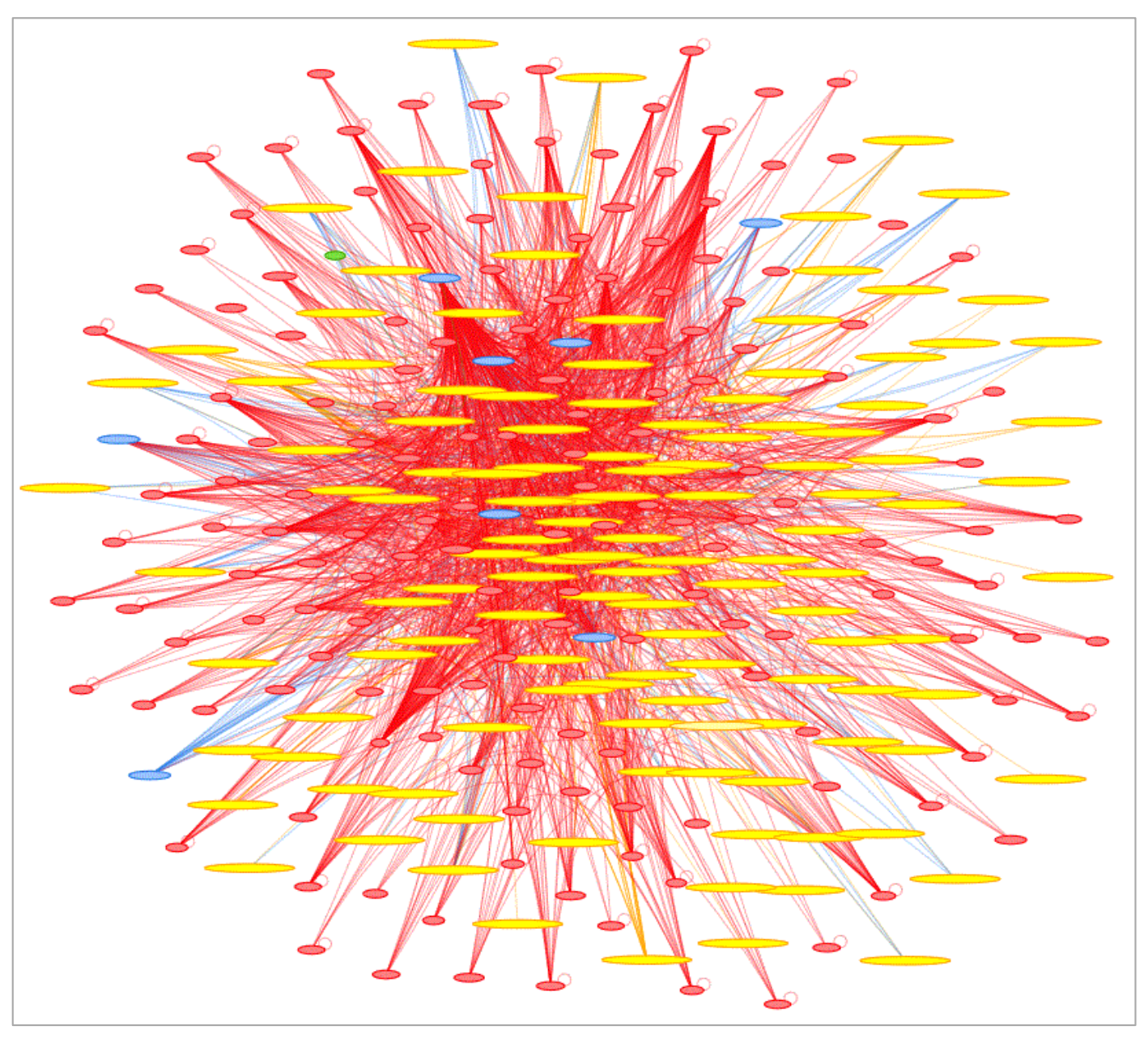


Step3 Add your query

○Genomic Range
○Transcription Factors
○LncRNA
●DNA Annotation

DNA Annotation
C1orf43

Select "DNA annotation" and input the gene name associated to a given DNA region.



Get the network with all the interacting TFs, lncRNAs, and DNA ranges for that region.

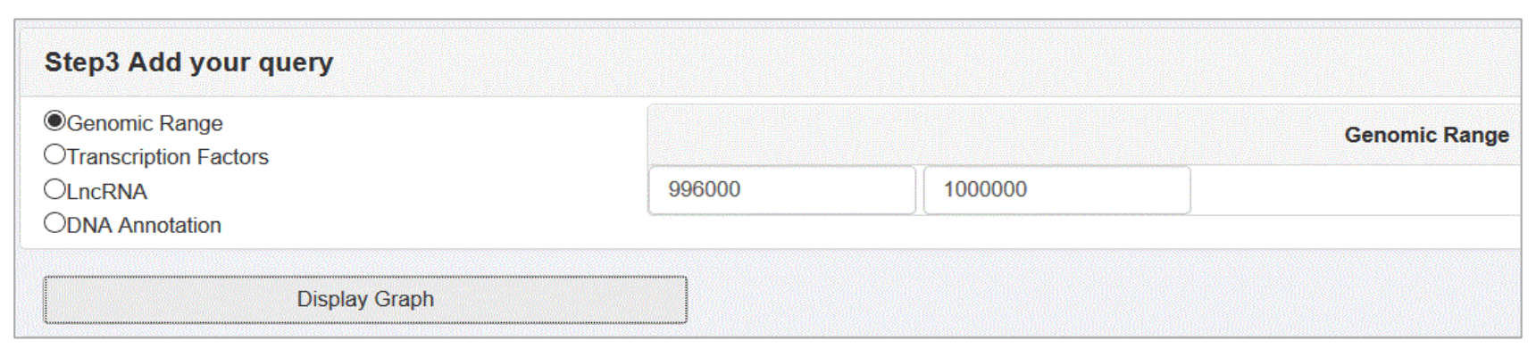
Test GREG now:

http://210.38.57.54:8080/index.html

GREG
GENE REGULATION
GRAPH DATABASE

Acknowledgments: Shaurya Jauhari, Rajni Jauhari

Find interactions for a chosen genomic region

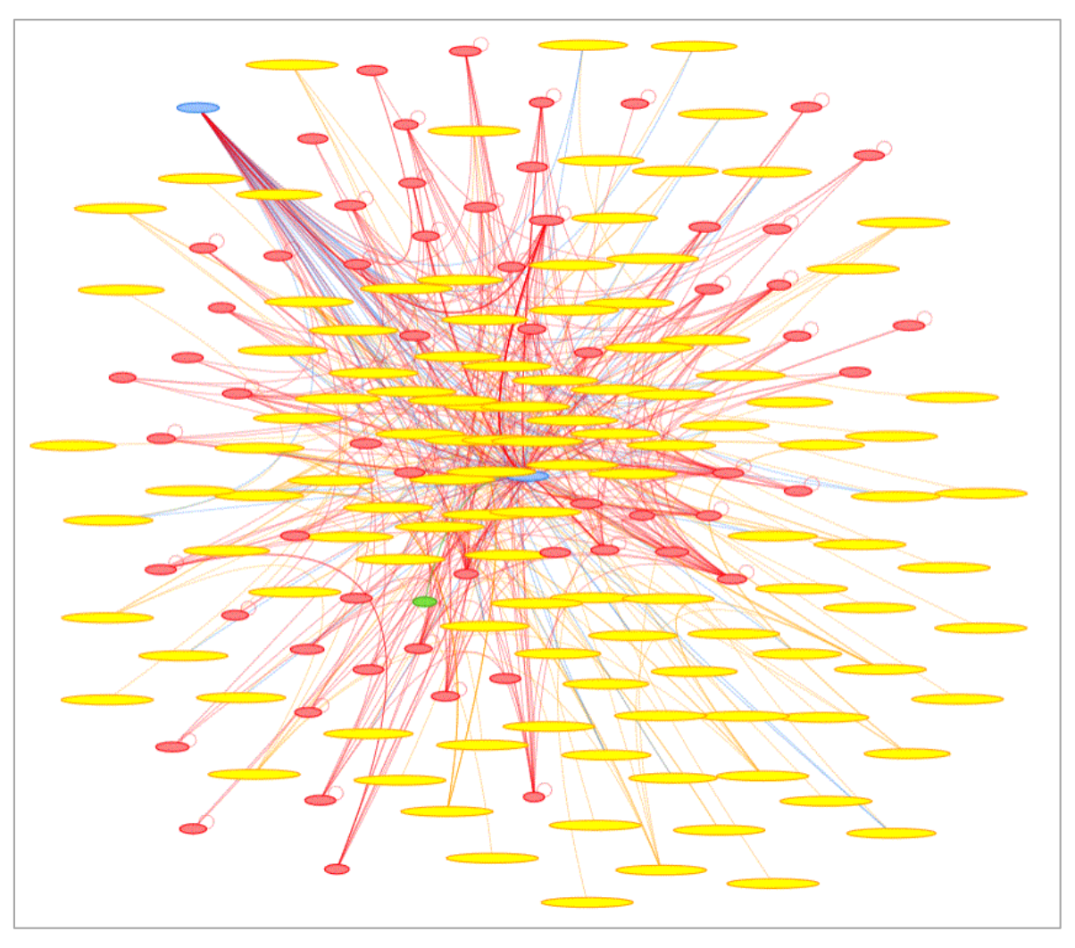


Step3 Add your query

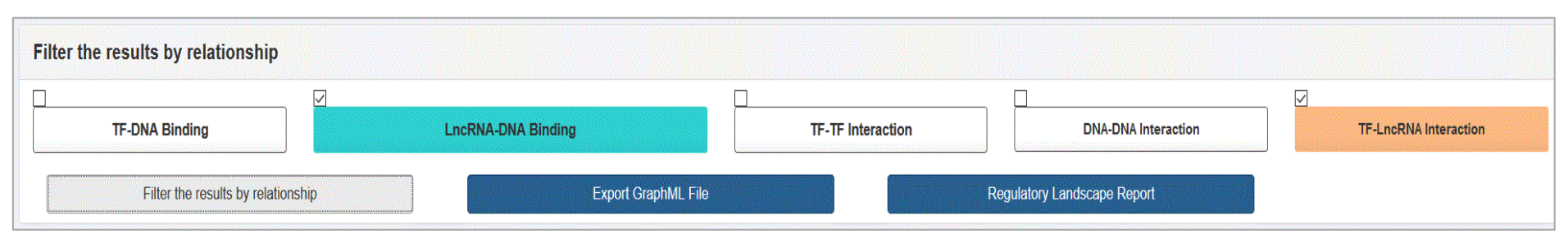
●Genomic Range
○Transcription Factors
○LncRNA
○DNA Annotation

Genomic Range	
996000	1000000

Select "Genomic Range" and input the coordinates of a DNA region for the previously chosen chromosome



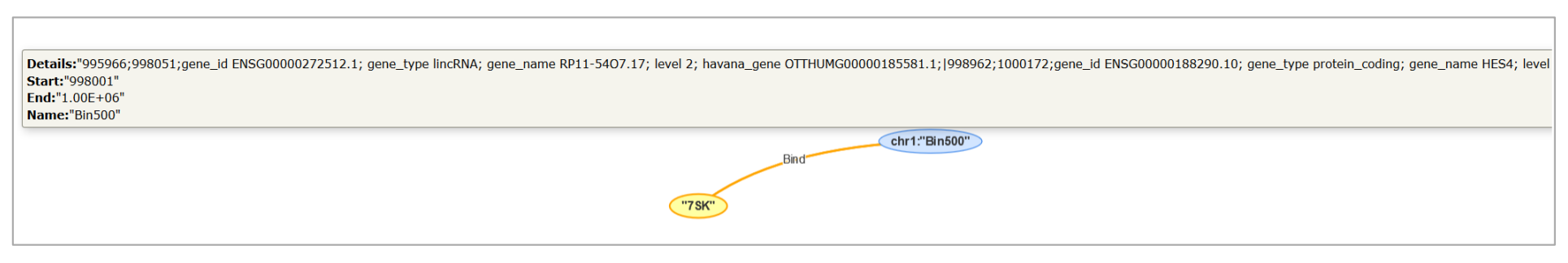
Get the network with all the interacting TFs, lncRNAs, and DNA ranges for that region.



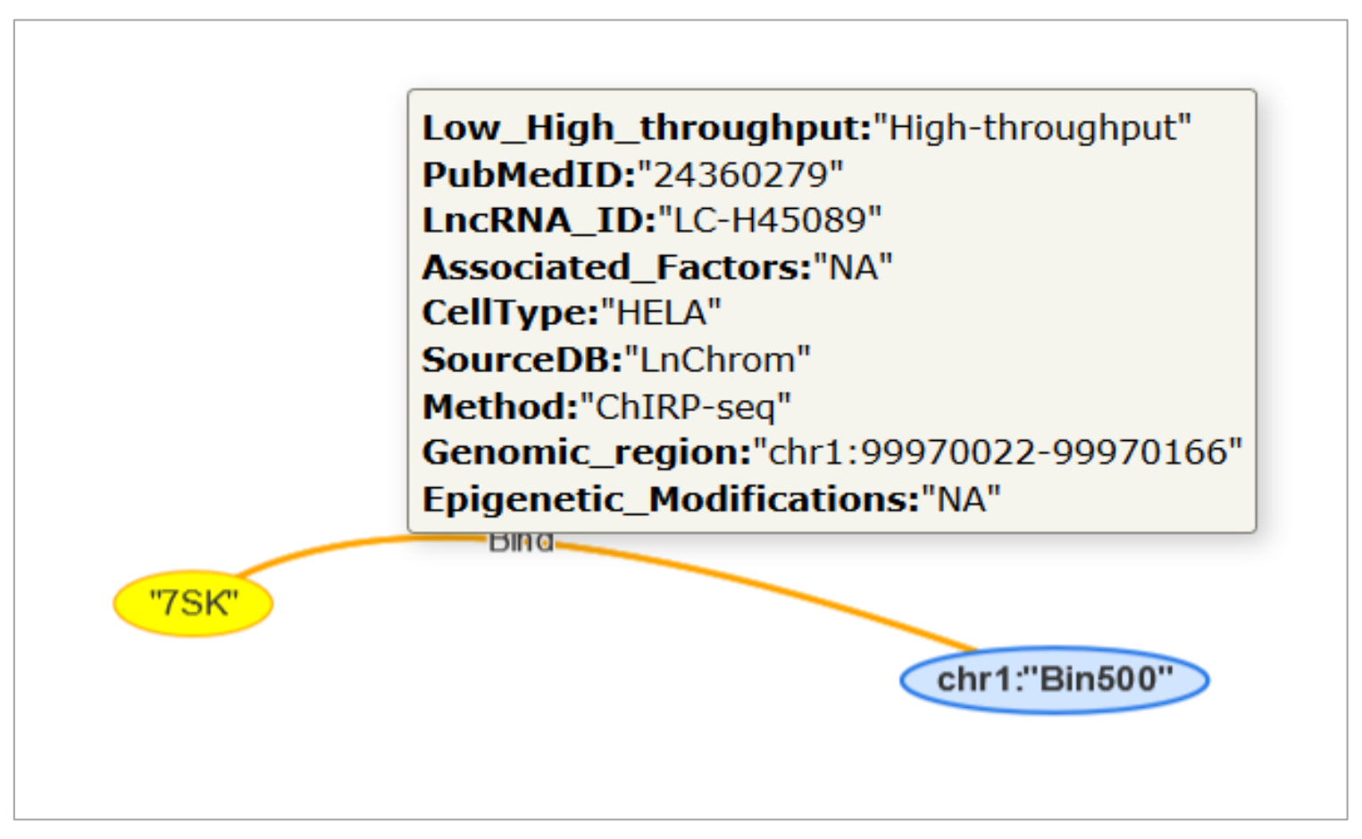
Filter the results by relationship

☐ TF-DNA Binding
☑ LncRNA-DNA Binding
☐ TF-TF Interaction
☐ DNA-DNA Interaction
☑ TF-LncRNA Interaction

Filter your results according to specific relationships (example, lncRNA-TF and lncRNA-DNA).



Details:"995966;998051;gene_id ENSG00000272512.1; gene_type lincRNA; gene_name RP11-54O7.17; level 2; havana_gene OTTHUMG00000185581.1;|998962;1000172;gene_id ENSG00000188290.10; gene_type protein_coding; gene_name HES4; level
Start:"998001"
End:"1.00E+06"
Name:"Bin500"



Low_High_throughput:"High-throughput"
PubMedID:"24360279"
LncRNA_ID:"LC-H45089"
Associated_Factors:"NA"
CellType:"HELA"
SourceDB:"LnChrom"
Method:"ChIRP-seq"
Genomic_region:"chr1:99970022-99970166"
Epigenetic_Modifications:"NA"

Examine the data associated to each node and interaction