# Networks 

Gerstein lab

AGGAAGGGAAAGCGCAAGAGAGAGCGCACACGCACACACCCGCCGCGCGCACTCGCGCACGGAC CCGCACGGGGACAGCTCGGAAGTCATCAGTTCCATGGGCGAATGCTGCTGCTGGCGAGATGTCTGC TGCTAGTCCTCGTCTCCTCGCTGCGGTATGCTCGGGACTGGCGTGCGGACCGGGCAGGGGGTTCGG GAAGAGGAG GGCCGA GAAGACCCT/ OOO TTAAGGA ACTCACCCCC GATGACTCAC CCGAACACTI CCACCCACC GAGAAGTTAC AGTCTTGCTA GGGGGCGTC AAACATGAAC $\qquad$ CAGGCT GAGAAG :TCCTTCC :CAGCTG [TTGATG TCACTTT :TCCCAC :TCAGAG CCCCCACGTTTCCCGTTGCCTCTGTGGTTTTCTTTCCACCACTACCCCCACCCTGCACCTCCCCACC GGCGGTTCTGACAAAAGGCAGGGGCTCCAGGGGAGAC


TCGGAAGGGAGAGGAATGGGAAATGGTCd TAAATCTGGACCGCGAGTGTGGACGCGCG CTGCACGGGCCATGGGGTGCCCGTTGCGT GCGCCCATCCTCTCGTCACCCTCACTCCC CGAGAGGAGGGAGGCACTGCGGAGACAG AGTGGGTCAGCTTCTGACCTGCTCTGCGG CCTTCCCAGGGCTTGAGCCTTGCAGCCCT

:CAGGGTCCCTCGGAACC CCGAGGTCCCAGAGCCAG GGGTCCTGGGCTCGCACA ATTTATTAGCTAGCGTACTA

## ENCODE: Encyclopedia of DNA Elements



# Network is a powerful framework for data integration 



## Common biological networks



Protein-protein Interaction networks


Metabolic pathway networks


Undirected

TF-target-gene Regulatory networks

miRNA-target networks


Directed

## Example: 3D organization of genome


"We finished the genome map, now we can't figure out how to fold it."

# Example: 3D organization of genome 



Topologically associating domains (TADs)


Network communities


## Example: 3D organization of genome



No description ar website provided.

| (-) 30 cammits |  | IV 1 branch | $\bigcirc$ releases | 12 contributor |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Branch: master ${ }^{\text {r }}$ | New pull request |  |  | Find file | Clone or downlosd - |
| (1) Kocn-Kiu Yan of group gearstein fix broken functions |  |  |  | lest comm | cat3b9e 26 days ago |
| E data |  |  | fix broken functions |  | 26 days ago |
| E MrTADFinder.j |  |  | fix broken functions |  | 26 deys ago |
| E README.md |  |  | new distance |  | 3 months ago |
| E run_MrTADFinder.jl |  |  | fix broken functions |  | 26 days ago |

## ExEADME.md

## MrTADFinder

MrTADFinder aims to identify topologically associating domains (TADs) in multiple resolutions.

## INPUT FILES:

MrTADFinder takes an whole-genome-to-whole genome contact map as an input. The contact map should be a

## Disruption of domain structure causes congenital disease <br> 

(A)
$2 \mathrm{Mb} \mid$ $\qquad$ - hg19

## Example: Transcriptional regulatory network



Bacteria
E. coli: 100 TFs, 4000 genes


Single-celled eukaryotes
Baker's yeast: 200 TFs, 6000 genes

Complex eukaryotes Human: 1400 TFs, 20000 genes

# Example: Hierarchical organization of human transcriptional regulatory network 



Distributions of the feature across levels are significantly different

$\bar{f}_{\text {MID }}$
MIDDLE

## $\bar{f}_{\mathrm{BOT}}$ BOTTOM






49 TOP

Gerstein $\mathrm{M}^{*}$, Kundaje A*, Hariharan M*, Landtl 1 S*. Yan KK* et al. Nature. 2012

## Example: Network comparison



Yan et. al. Gerstein Cell Systems 2016

## Example: human transcriptional regulatory network versus chain-ofcommand hierarchy



Yan et. al. Gerstein Cell Systems 2016

## Example: Protein-protein interaction network versus software package dependency network



## Example: Collaboration network in a scientific consortium



