

Powered by:  nextflow

```
$ nextflow run QuackenbushLab/tcga-data-nf \
-c my-config.conf
```

tcga-data-nf



download

Downloaded
Raw Data

full

prepare

Preprocessed
Data

analyze

Output
Networks

recount3

TCGA Gene Expression
GTEx Gene Expression

GDC

Mutations
Methylation
CNV

NetSciDataCompanion

<https://github.com/QuackenbushLab/NetworkDataCompanion>

Filtering

- Tumor Purity
- Gene Expression
- Duplicate Samples
- ...

Conversion

- sample-ID
- gene names
- ...

Methylation specific

- epigenetic scores
- beta - M values
- probe maps
- ...

Normalization

Batch correction

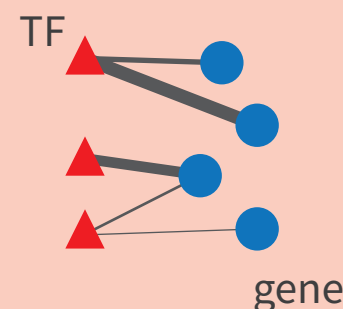


netZoo

<https://github.com/netZoo/>

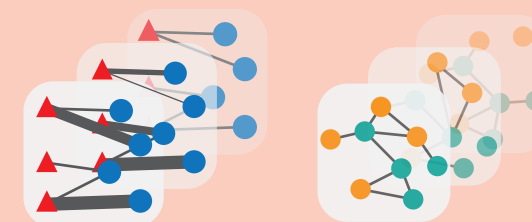
PANDA

Gene Regulatory Network



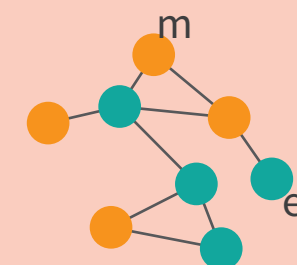
LIONESS

Sample-specific Networks



DRAGON

Multomics association networks



Textual configuration files

★ download.json
List of tumor types and modalities

▲ parameters.conf
Filters
Normalization
Output folders
Execution
...

● metadata.csv
ID, fata_file, tumor_type
lung, tumor1.rds, LUAD
colon, tumor2.rds, COAD
...

Data for 10 common cancers

Multi-omics data

<https://tcga-data-nf-precomputed.s3.us-east-2.amazonaws.com/raw-data/firstround-20221102/>

Pre-computed networks

<https://grand.networkmedicine.org/cancers/>

Examples

<https://github.com/QuackenbushLab/tcga-data-supplement>