

# ConnectOR-optimized

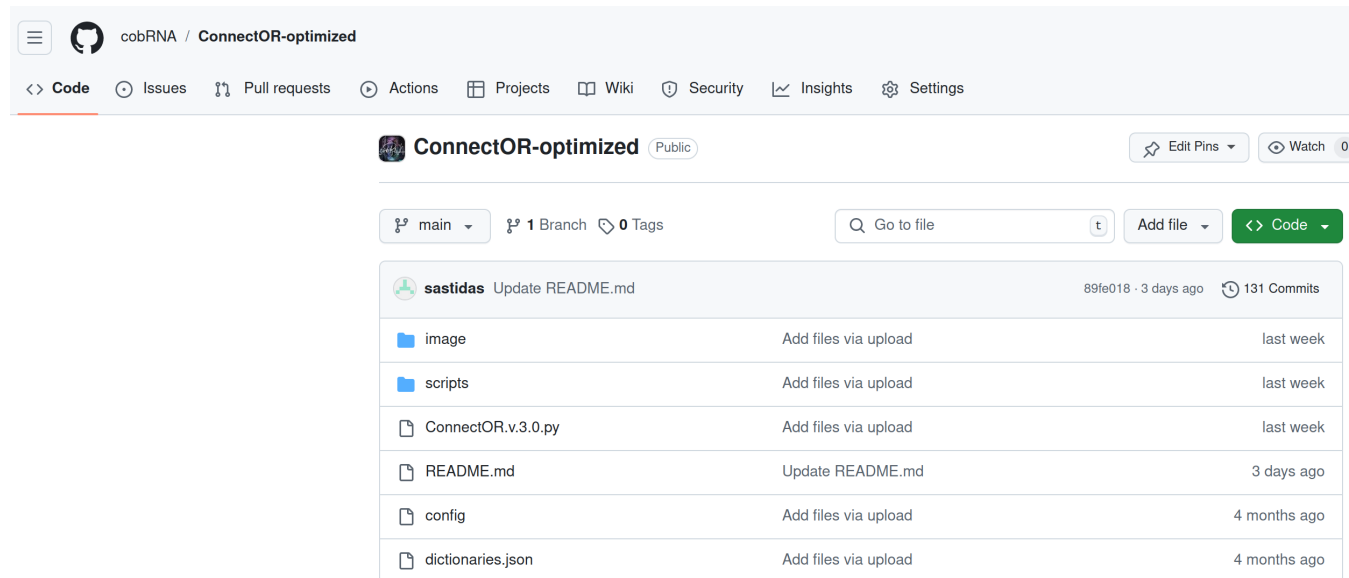
Optimized & modified

by

barbara uszczyńska-ratajczak & sasti Gopal Das

Download github link

<https://github.com/cobRNA/ConnectOR-optimized>



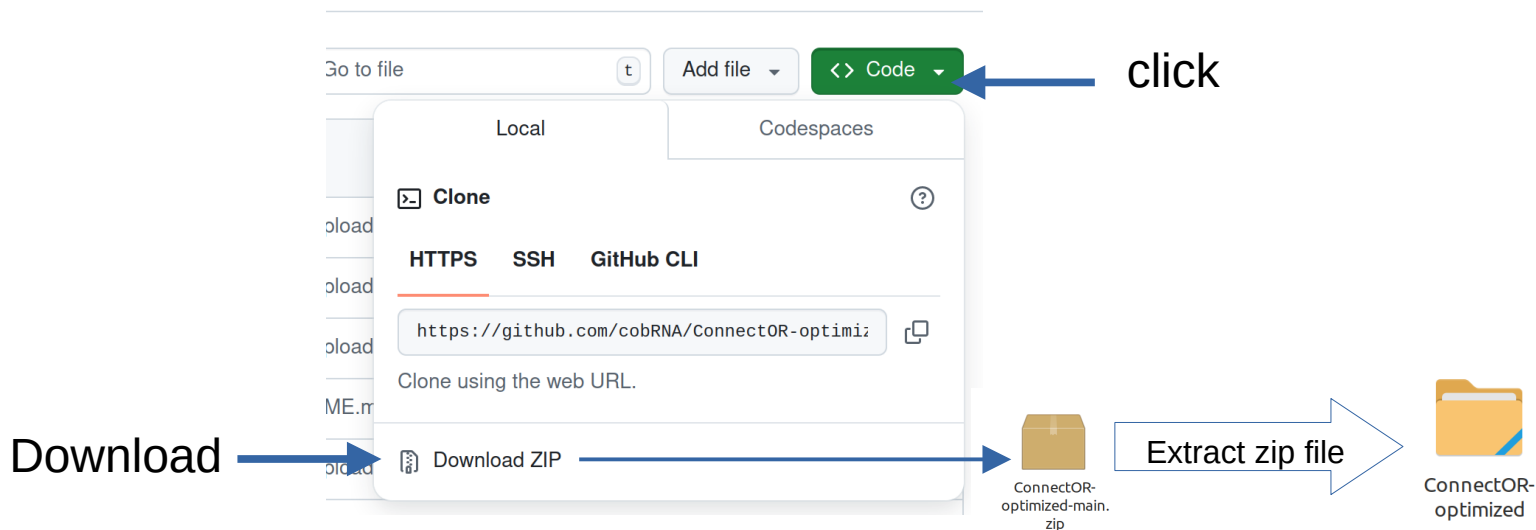
The screenshot shows the GitHub repository page for 'cobRNA / ConnectOR-optimized'. The repository is public and has 1 branch (main) and 0 tags. The file list includes:

File/Folder	Commit Message	Commit Date	Commits
image	Add files via upload	last week	
scripts	Add files via upload	last week	
ConnectOR.v.3.0.py	Add files via upload	last week	
README.md	Update README.md	3 days ago	131
config	Add files via upload	4 months ago	
dictionaries.json	Add files via upload	4 months ago	

# Download

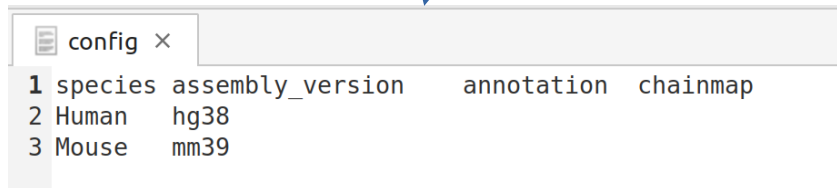
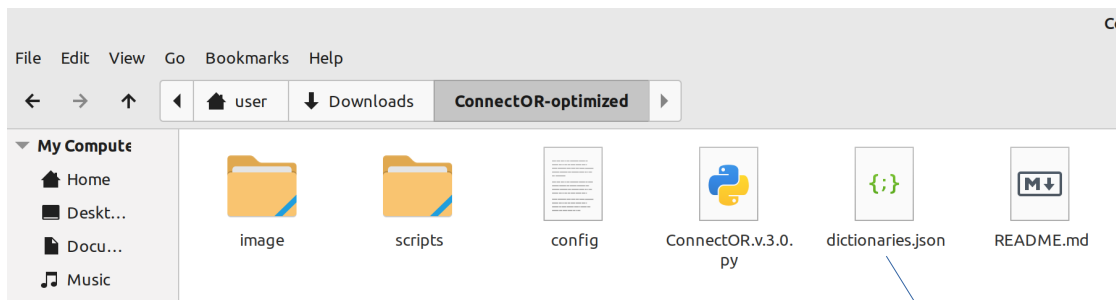
```
(base) user@user-X570S-AERO-G:~/Downloads$ git clone https://github.com/cobRNA/ConnectOR-optimized
Cloning into 'ConnectOR-optimized'...
remote: Enumerating objects: 385, done.
remote: Counting objects: 100% (84/84), done.
remote: Compressing objects: 100% (59/59), done.
remote: Total 385 (delta 64), reused 25 (delta 25), pack-reused 301 (from 1)
Receiving objects: 100% (385/385), 18.87 MiB | 29.06 MiB/s, done.
Resolving deltas: 100% (175/175), done.
```

OR





ConnectOR-  
optimized



OR

```
1 species assembly_version annotation chainmap
2 Human hg38 GTFs/genCODE.v47.primary_assembly.annotation.gtf.gz chainmaps/hg38ToMm39.over.chain.gz
3 Mouse mm39 GTFs/genCODE.vM36.primary_assembly.annotation.gtf.gz chainmaps/mm39ToHg38.over.chain.gz
```

```
{"gtfs": {"hg19": "ftp://ftp.ebi.ac.uk/pub/databases/genCODE/GenCODE_human/release_19/genCODE.v19.annotation.gtf.gz",
"hg38": "ftp://ftp.ebi.ac.uk/pub/databases/genCODE/GenCODE_human/release_32/genCODE.v32.annotation.gtf.gz",
"mm9": "ftp://ftp.ebi.ac.uk/pub/databases/genCODE/GenCODE_mouse/release_M9/genCODE.vM9.annotation.gtf.gz",
"mm10": "ftp://ftp.ebi.ac.uk/pub/databases/genCODE/GenCODE_mouse/release_M22/genCODE.vM22.annotation.gtf.gz",
"mm39": "https://ftp.ebi.ac.uk/pub/databases/genCODE/GenCODE_mouse/release_M30/genCODE.vM30.annotation.gtf.gz"
},
```

```
"gtfs_ensembl_r96": {"mm10": "ftp://ftp.ensembl.org/pub/release-98/gtf/mus_musculus/Mus_musculus.GRCh38.98.gtf.gz",
"mm9": "ftp://ftp.ensembl.org/pub/release-65/gtf/mus_musculus/Mus_musculus.NCBI37.65.gtf.gz",
"danrer11": "ftp://ftp.ensembl.org/pub/release-98/gtf/danio_rerio/Danio_rerio.GRCz11.98.gtf.gz",
"danrer7": "ftp://ftp.ensembl.org/pub/release-86/gtf/danio_rerio/Danio_rerio.GRCz10.86.gtf.gz",
"hg19": "ftp://ftp.ensembl.org/pub/grch37/release-98/gtf/homo_sapiens/Homo_sapiens.GRCh37.87.gtf.gz",
"hg38": "ftp://ftp.ensembl.org/pub/release-98/gtf/homo_sapiens/Homo_sapiens.GRCh38.98.gtf.gz"}},
```

```
"chain_maps": { "hg18": {"brafl01": "https://hgdownload-test.gi.ucsc.edu/goldenPath/hg18/liftOver/hg18ToBraFl01.over.chain.gz",
"hg19": {"danrer11": "https://hgdownload.soe.ucsc.edu/goldenPath/hg19/liftOver/hg19ToDanRer11.over.chain.gz",
"hg38": "https://hgdownload.soe.ucsc.edu/goldenPath/hg19/liftOver/hg19ToHg38.over.chain.gz",
"mm10": "https://hgdownload.soe.ucsc.edu/goldenPath/hg19/liftOver/hg19ToMm10.over.chain.gz",
"mm9": "https://hgdownload.soe.ucsc.edu/goldenPath/hg19/liftOver/hg19ToMm9.over.chain.gz",
"danrer10": "https://hgdownload.soe.ucsc.edu/goldenPath/hg19/liftOver/hg19ToDanRer10.over.chain.gz",
"danrer7": "https://hgdownload.soe.ucsc.edu/goldenPath/hg19/liftOver/hg19ToDanRer7.over.chain.gz",
"hg18": "https://hgdownload-test.gi.ucsc.edu/goldenPath/hg18/liftOver/hg18ToBraFl01.over.chain.gz"
}
}
```

# Prerequisites:

## Prerequisites:

```
(base) user@user-X570S-AERO-G:~/Downloads$ python --version
Python 3.10.12
(base) user@user-X570S-AERO-G:~/Downloads$
```

Python3: (install Python 3)

```
(base) user@user-X570S-AERO-G:~/Downloads$ pip show wget
Name: wget
Version: 3.2
```

```
install Packages

pip3 install wget
pip3 install pandas
```

```
(base) user@user-X570S-AERO-G:~/Downloads$ pip show pandas
Name: pandas
Version: 2.0.3
```

R:

```
> packageVersion("ggplot2")
[1] '3.5.0'
```

```
install Packages

ggplot2
tidyr
dplyr
```

```
> library(ggplot2)
packageVersion("tidyr")
[1] '1.3.1'
```

```
> library(ggplot2)
packageVersion("dplyr")
[1] '1.1.4'
```

install BedTools

```
sudo apt-get install bedtools
```

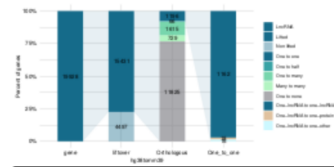
```
(base) user@user-X570S-AERO-G:~/Downloads$ bedtools --version
bedtools v2.31.0
(base) user@user-X570S-AERO-G:~/Downloads$
```

# Run ConnectOR

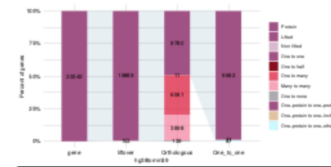
python3 ConnectOR.v.3.0.py

```
(base) user@user-X570S-AERO-G:~/Downloads$ cd ConnectOR-optimized/
(base) user@user-X570S-AERO-G:~/Downloads/ConnectOR-optimized$ python3 ConnectOR.v.3.0.py
Checking Config_file...
100% [.....] 49661267 / 49661267 gencode.v43.annotation.gtf.gz downloaded succesfully..
100% [.....] 73017320 / 73017320 hg38ToMm39.over.chain.gz downloaded succesfully..
100% [.....] 29299972 / 29299972 gencode.vM32.annotation.gtf.gz downloaded succesfully..
100% [.....] 70905409 / 70905409 mm39ToHg38.over.chain.gz downloaded succesfully..
Config_file is correct..
Generating geneID - geneName - geneBiotype - transcripts maps...
  Maps for hg38... generating
  Maps for mm39... generating
  Maps for mm39... generated
  Maps for hg38... generated
Generating BED files for exons and genes...
  Exons BEDs for hg38... generating
  Exons BEDs for mm39... generating
  Genes BEDs for hg38... generating
  Genes BEDs for mm39... generating
  BEDs for mm39... generated
  BEDs for hg38... generated
LiftOver features between species...
  hg38 exons to mm39... mapping
  mm39 exons to hg38... mapping
Reading liftover chains
Reading liftover chains
Mapping coordinates
Mapping coordinates
  mm39 exons to hg38... done
  hg38 exons to mm39... done
Intersecting LiftOver...
  Human exons to Mouse... done
  Mouse exons to Human... done
Predicting orthologues...
  Human exonStrand to Mouse... done
  Mouse exonStrand to Human... done
completed
```

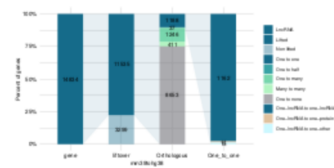
# ConnectOR result



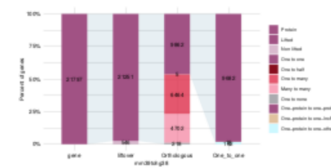
hg38tomm39\_lncRNA\_Orthologous.pdf  
5,5 kB  
Fri 13 Sep 2024 02:49:00 PM CEST



hg38tomm39\_protein\_Orthologous.pdf  
5,5 kB  
Fri 13 Sep 2024 02:49:01 PM CEST



mm39toh38\_lncRNA\_Orthologous.pdf  
5,5 kB  
Fri 13 Sep 2024 02:49:01 PM CEST



mm39tohg38\_protein\_Orthologous.pdf  
5,6 kB  
Fri 13 Sep 2024 02:49:02 PM CEST