

# GitHub - cadms/resfinder

---

 [github.com/cadms/resfinder](https://github.com/cadms/resfinder)

ResFinder identifies acquired antimicrobial resistance genes in total or partial sequenced isolates of bacteria.

## Content of the repository

---

1. run\_resfinder.py - Use this script to run ResFinder
2. tests/data - Contains fasta and fastq data for testing. More information in the "Test data" section
3. scripts/ - All scripts in this directory is unsupported but has been uploaded as they may be useful
4. cge/ - ResFinder code
5. dockerfile - Used to build ResFinder docker image (See Docker section near the end)

## Installation

---

The installation described here will first install the actual ResFinder software, then the dependencies, and finally the databases. A more detailed breakdown of the installation is provided below:

1. Install ResFinder tool
2. Install python modules
3. Install BLAST (optional)
4. install KMA (optional)
5. Download ResFinder database
6. Download PointFinder database
7. Index databases with KMA (if installed)
8. Test installation

A small script has been written to automate this process. It is available from the scripts directory and is named install\_resfinder.sh. It is very simple and might not work in all environments. It is only meant as a supplement and no support will be provided for any scripts in this directory. However, specific suggestions (with code) for improvement is very welcome.

## ResFinder tool

---

Setting up ResFinder script and database

```
# Go to wanted location for resfinder
cd /path/to/some/dir

# Clone the latest version and enter the resfinder directory
git clone https://git@bitbucket.org/genomicepidemiology/resfinder.git
cd resfinder
```

## Dependencies:

---

Depending on how you plan to run ResFinder BLAST and KMA can be optional. BLAST is used to analyse assemblies (ie. FASTA files). KMA is used to analyse read data (ie. FASTQ files).

## Python modules: Tabulate, BioPython, CGECore and Python-Git

---

To install the needed python modules you can use pip

```
pip3 install tabulate biopython cgecore gitpython python-dateutil
```

For more information visit the respective website

```
https://bitbucket.org/astanin/python-tabulate
https://biopython.org
https://bitbucket.org/genomicepidemiology/cge_core_module
https://gitpython.readthedocs.io/en/stable/index.html
```

## BLAST (optional)

---

If you don't want to specify the path of blastn every time you run ResFinder, make sure that blastn is in your PATH.

Blastn can be obtained from:

```
ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/
```

## KMA (optional)

---

The instructions here will install KMA in the default location ResFinder uses. KMA can be installed in another location but the path to KMA will then need to be specified every time you run ResFinder unless you add the kma program to your PATH.

```
# Go to the directory in which you installed the ResFinder tool
cd /path/to/some/dir/resfinder
cd cge
git clone https://bitbucket.org/genomicepidemiology/kma.git
cd kma && make
```

## Databases

---

This section describes how to install the databases at the ResFinder default locations. The database locations can be changed, but must then be specified to ResFinder at run time.

### ResFinder database

---

```
# Go to the directoy in which you installed the ResFinder tool
cd /path/to/some/dir/resfinder
git clone
https://git@bitbucket.org/genomicepidemiology/resfinder_db.git
db_resfinder
```

### PointFinder database

---

```
# Go to the directoy in which you installed the ResFinder tool
cd /path/to/some/dir/resfinder
git clone
https://git@bitbucket.org/genomicepidemiology/pointfinder_db.git
db_pointfinder
```

## Indexing databases with KMA

---

If you have KMA installed you either need to have the kma\_index in your PATH or you need to provide the path to kma\_index to INSTALL.py

**NOTE:** The documentation given here describes the procedure for the ResFinder database, but the procedure is identical for the PointFinder database. **PointFinder database documentation:**

[[https://bitbucket.org/genomicepidemiology/pointfinder\\_db](https://bitbucket.org/genomicepidemiology/pointfinder_db)]

a) Run INSTALL.py in interactive mode

```
# Go to the database directory
cd path/to/db_resfinder
python3 INSTALL.py
```

If kma\_index was found in your path a lot of indexing information will be printed to your terminal, and will end with the word "done".

If kma\_index wasn't found you will recieve the following output:

```
KMA index program, kma_index, does not exist or is not executable
Please input path to executable kma_index program or choose one of the
options below:
```

1. Install KMA using make, index db, then remove KMA.
2. Exit

You can now write the path to kma\_index and finish with or you can enter "1" or "2" and finish with .

If "1" is chosen, the script will attempt to install kma in your systems default temporary location. If the installation is successful it will proceed to index your database, when finished it will delete the kma installation again.

b) Run INSTALL.py in non\_interactive mode

```
# Go to the database directory
cd path/to/resfinder_db
python3 INSTALL.py /path/to/kma_index non_interactive
```

The path to kma\_index can be omitted if it exists in PATH or if the script should attempt to do an automatic temporary installation of KMA.

c) Index database manually (not recommended)

It is possible to index the databases manually, but is generally not recommended as it is more prone to error. If you choose to do so, be aware of the naming of the indexed files.

This is an example of how to index the ResFinder database files:

```

# Go to the resfinder database directory
cd path/to/resfinder_db
# create indexing directory
mkdir kma_indexing
# Index files using kma_index
kma_index -i db_resfinder/fusidicacid.fsa -o
db_resfinder/kma_indexing/fusidicacid
kma_index -i db_resfinder/phenicol.fsa -o
db_resfinder/kma_indexing/phenicol
kma_index -i db_resfinder/glycopeptide.fsa -o
db_resfinder/kma_indexing/glycopeptide
kma_index -i db_resfinder/trimethoprim.fsa -o
db_resfinder/kma_indexing/trimethoprim
kma_index -i db_resfinder/oxazolidinone.fsa -o
db_resfinder/kma_indexing/oxazolidinone
kma_index -i db_resfinder/tetracycline.fsa -o
db_resfinder/kma_indexing/tetracycline
kma_index -i db_resfinder/quinolone.fsa -o
db_resfinder/kma_indexing/quinolone
kma_index -i db_resfinder/nitroimidazole.fsa -o
db_resfinder/kma_indexing/nitroimidazole
kma_index -i db_resfinder/fosfomycin.fsa -o
db_resfinder/kma_indexing/fosfomycin
kma_index -i db_resfinder/aminoglycoside.fsa -o
db_resfinder/kma_indexing/aminoglycoside
kma_index -i db_resfinder/macrolide.fsa -o
db_resfinder/kma_indexing/macrolide
kma_index -i db_resfinder/sulphonamide.fsa -o
db_resfinder/kma_indexing/sulphonamide
kma_index -i db_resfinder/rifampicin.fsa -o
db_resfinder/kma_indexing/rifampicin
kma_index -i db_resfinder/colistin.fsa -o
db_resfinder/kma_indexing/colistin
kma_index -i db_resfinder/beta-lactam.fsa -o
db_resfinder/kma_indexing/beta-lactam
# Go to the pointfinder database directory
cd path/to/pointfinder_db
# Index files using kma_index
kma_index -i db_pointfinder/campylobacter/*.fsa -o
db_pointfinder/campylobacter/campylobacter
kma_index -i db_pointfinder/escherichia_coli/*.fsa -o
db_pointfinder/escherichia_coli/escherichia_coli
kma_index -i db_pointfinder/enterococcus_faecalis/*.fsa -o
db_pointfinder/enterococcus_faecalis/enterococcus_faecalis
kma_index -i db_pointfinder/enterococcus_faecium/*.fsa -o
db_pointfinder/enterococcus_faecium/enterococcus_faecium
kma_index -i db_pointfinder/neisseria_gonorrhoeae/*.fsa -o
db_pointfinder/neisseria_gonorrhoeae/neisseria_gonorrhoeae
kma_index -i db_pointfinder/salmonella/*.fsa -o
db_pointfinder/salmonella/salmonella
kma_index -i db_pointfinder/mycobacterium_tuberculosis/*.fsa -o
db_pointfinder/mycobacterium_tuberculosis/mycobacterium_tuberculosis

```

## Test ResFinder intallation

---

(This will not function with the docker installation.) If you did not install BLAST, test 1 and 3 will fail. If you did not install KMA, test 2 and 4 will fail. The 4 tests will in total take approximately take 5-60 seconds, depending on your system.

```
# Go to the directoy in which you installed the ResFinder tool
cd /path/to/some/dir/resfinder
```

```
# For seeing the unittest options for running the tests
python3 tests/functional_tests.py -h
```

```
# In case you need to point blastn or kma, or the resfinder or the
pointfinder databases as they are not in the places indicated above,
see the optional arguments for the test by:
python3 tests/functional_tests.py -res_help
```

```
#Which outputs:
```

```
usage: functional_tests.py [-res_help] [-db_res DB_PATH_RES] [-b
BLAST_PATH]
                        [-k KMA_PATH] [-db_point DB_PATH_POINT]
```

```
Options:
```

```
-res_help, --resfinder_help
-db_res DB_PATH_RES, --db_path_res DB_PATH_RES
                        Path to the databases for ResFinder
-b BLAST_PATH, --blastPath BLAST_PATH
                        Path to blastn
-k KMA_PATH, --kmaPath KMA_PATH
                        Path to KMA
-db_point DB_PATH_POINT, --db_path_point DB_PATH_POINT
                        Path to the databases for PointFinder
```

```
# Run tests
python3 tests/functional_tests.py
```

```
# Output from successful tests
```

```
....
-----
Ran 4 tests in 8.263s

OK
```

## Test data

---

Test data can be found in the sub-dierectory /tests/data

## Usage

---

You can run resfinder command line using python3.

**NOTE:** Species should be entered with their full scientific names (e.g. "escherichia coli"), using quotation marks, not case sensitive. An attempt has been made to capture some deviations like "ecoli" and "e.coli", but it is far from all deviations that will be captured.

```

# Example of running resfinder
python3 run_resfinder.py -o path/to/outdir -s "Escherichia coli" -l
0.6 -t 0.8 --acquired --point -ifq test_isolate_01_*

# The program can be invoked with the -h option
usage: run_resfinder.py [-h] [-ifa INPUTFASTA]
                        [-ifq INPUTFASTQ [INPUTFASTQ ...]] [-scripts
SCRIPTS]
                        [-o OUT_PATH] [-b BLAST_PATH] [-k KMA_PATH]
                        [-s SPECIES] [-l MIN_COV] [-t THRESHOLD]
                        [-db_res DB_PATH_RES] [-db_res_kma
DB_PATH_RES_KMA]
                        [-d DATABASES] [-acq] [-c] [-db_point
DB_PATH_POINT]
                        [-g SPECIFIC_GENE [SPECIFIC_GENE ...]] [-u]

optional arguments:
  -h, --help                show this help message and exit
  -ifa INPUTFASTA, --inputfasta INPUTFASTA
                            Input fasta file.
  -ifq INPUTFASTQ [INPUTFASTQ ...], --inputfastq INPUTFASTQ
[INPUTFASTQ ...]
                            Input fastq file(s). Assumed to be single-end
fastq if
                            only one file is provided, and assumed to be
paired-
                            end data if two files are provided.
  -o OUT_PATH, --outputPath OUT_PATH
                            All output will be stored in this directory.
  -b BLAST_PATH, --blastPath BLAST_PATH
                            Path to blastn
  -k KMA_PATH, --kmaPath KMA_PATH
                            Path to kma
  -s SPECIES, --species SPECIES
                            Species in the sample

                            Available species:
Campylobacter, Campylobacter jejuni, Campylobacter coli,
                            Enterococcus faecalis,
Enterococcus faecium, Escherichia coli, Helicobacter pylori,
                            Klebsiella,
Mycobacterium tuberculosis, Neisseria gonorrhoeae,
                            Plasmodium falciparum,
Salmonella, Salmonella enterica, Staphylococcus aureus
                            -s "Other" can be used
for metagenomic samples or samples with unknown species.
  -db_res DB_PATH_RES, --db_path_res DB_PATH_RES
                            Path to the databases for ResFinder
  -db_res_kma DB_PATH_RES_KMA, --db_path_res_kma DB_PATH_RES_KMA
                            Path to the ResFinder databases indexed with
KMA.
                            Defaults to the 'kma_indexing' directory
inside the
                            given database directory.
  -d DATABASES, --databases DATABASES
                            Databases chosen to search in - if none is
specified

```



all is used

-acq, --acquired Run resfinder for acquired resistance genes

-l MIN\_COV, --min\_cov MIN\_COV  
Minimum (breadth-of) coverage of

ResFinder

Valid interval:

0.00-1.00

-t THRESHOLD, --threshold THRESHOLD

Threshold for identity of ResFinder

Valid interval: 0.00-1.00

-c, --point Run pointfinder for chromosomal mutations

-db\_point DB\_PATH\_POINT, --db\_path\_point DB\_PATH\_POINT  
Path to the databases for PointFinder

-g SPECIFIC\_GENE [SPECIFIC\_GENE ...]  
Specify genes existing in the database to

search for -  
if none is specified all genes are included in

the  
search.

-u, --unknown\_mut Show all mutations found even if in unknown to

the  
resistance database

-l\_p MIN\_COV\_POINT, --min\_cov\_point MIN\_COV\_POINT  
Minimum (breadth-of) coverage of

Pointfinder. If None  
is selected, the minimum coverage of

ResFinder will be  
used.

-t\_p THRESHOLD\_POINT, --threshold\_point THRESHOLD\_POINT

Threshold for identity of Pointfinder. If None is  
selected, the minimum coverage of ResFinder will be  
used.

## Web-server

---

A webserver implementing the methods is available at the [CGE website](https://cge.cbs.dtu.dk/services/ResFinder/) and can be found here: <https://cge.cbs.dtu.dk/services/ResFinder/>

## Install ResFinder with Docker

---

If you would like to build a docker image with ResFinder, make sure you have cloned the ResFinder directory as well as installed and indexed the databases: `db_pointfinder` and `db_resfinder`. Then run the following commands:

```
# Go to ResFinder directory
cd path/to/resfinder
# Build docker image with name resfinder
docker build -t resfinder .
```

When running the docker make sure to mount the **db\_resfinder** and the **db\_pointfinder** with the flag -v, as shown in the examples below.

You can test the installation by running the docker with the test files:

```
cd path/to/resfinder/
mkdir results

# Run with raw data (this command mounts the results to the local
directory "results")
docker run --rm -it -v $(pwd)/db_resfinder:/usr/src/db_resfinder -v
$(pwd)/results:/usr/src/results resfinder -ifq
/usr/src/tests/data/test_isolate_01_1.fq
/usr/src/tests/data/test_isolate_01_2.fq -acq -db_res
/usr/src/db_resfinder -o /usr/src/results

# Run with assembled data (this command mounts the results to the
local directory "results")
docker run --rm -it -v $(pwd)/db_resfinder:/usr/src/db_resfinder -v
$(pwd)/results:/usr/src/results resfinder -ifa
/usr/src/tests/data/test_isolate_01.fa -acq -db_res
/usr/src/db_resfinder -o /usr/src/results
```