

BIOL60201 Programming Skills  
Project 2 Man Page

**NAME**

lookUpTable.py

**DESCRIPTION**

After the filtering and merging of 3 separate data files containing miRNA to locus information, locus to gene information and gene to disease information, the script creates a lookup database which allows the user to query associations between miRNA, genes and diseases.

**The database is searchable using the following core options:**

- h** Help statement which lists all available options as well as their description.
- mirna** Searches the database for all genes and diseases associated with a specific miRNA.
- disease** Searches the database for all miRNA and genes associated with a specific disease.
- sm** Searches the database for all miRNA that match a given part of a miRNA.
- sd** Searches the database for all diseases that match a given part of a disease.

**Extra options:**

- geneConf** Is an integer value used in tandem with --mirna. Filters out the genes and diseases that have a confidence score smaller than or equal to the input. The minimum value for -geneConf is 80 and the maximum is 100.
- locusConf** Is an integer value used in tandem with --disease. Filters out the miRNA and genes that have a predicted score less than or equal to the input. The minimum value for -locusConf is 80 and the maximum is 100.
- score** is a Boolean value true/false used in tandem with --mirna and --disease. By default -score is set to true which would output the miRNAPredScore or score columns alongside the results when used with --mirna and --disease respectively.

**Notes:**

- Only 1 argument mentioned in the core options sections could be used at time. An error message will be displayed otherwise.
- Misusing an argument mentioned in the extra options section would result in an error message.
- Using only an argument mentioned in the extra options section above would result in an error message highlighting the need to have an argument from the core options section.
- Double quotes (") should be used when writing a disease name which is more than 1 word in length.

## EXAMPLES

o Python lookUpTable.py --mirna hsa-miR-4699-3p

geneSymbol	diseaseName	Score
ZIC2	holoprosencephaly 5	90
CILK1	endocrine-cerebroosteodysplasia	90
FGFR2	pfeiffer syndrome	100

o Python lookUpTable.py --mirna hsa-miR-4699-3p -score false

geneSymbol	diseaseName
ZIC2	holoprosencephaly 5
CILK1	endocrine-cerebroosteodysplasia
FGFR2	pfeiffer syndrome

o Python lookUpTable.py --mirna hsa-miR-4699-3p -geneConf 90

geneSymbol	diseaseName	Score
FGFR2	pfeiffer syndrome	100

## Notes:

- The tables above do not represent the entire results.
- The script would behave similarly using --disease instead of --mirna and -locusConf instead of -geneConf whilst searching for a disease.

o Python lookUpTable.py --sm 145

['hsa-miR-145-5p', 'hsa-miR-3145-3p', 'hsa-miR-145-3p', 'hsa-miR-3145-5p']

Note: the script would behave similarly using --sd instead of --sm whilst searching for part of a disease.

## COMPATIBILITY

This program is compatible with Python3.

## AUTHOR

Jospin Al Deek, Student ID: 10965177.