mtR_find

Purpose:

mtR_find is a tool for identification and annotation of sequences mapping to mitochondrial genomes. To run the script on adapter-trimmed FASTQ files, the following dependencies are required: PYTHON pandas module, cutadapt, and bowtie.

Usage:

mtR_find.py <species_name> <sRNA or lncRNA> [--FASTA path/to/mitochondrial genome.fa file] [-GTF path/to/gtf file] [--graphical_output yes/no][--output_path path/to/folder] [--input_path path/to/folder] [--files list of files]

Required arguments are indicated within <> and one of the allowed options has to be entered compulsorily. Optional arguments are indicated within square brackets. If the optional arguments are not specified, then the default values will be used. Required and optional arguments are indicated below:

Required arguments	
Argument	**Allowed Options**
species_name	1) *Homo sapiens* - hsa 2) *Danio rerio* - dre 3) *Mus musculus* - mmu 4) *Rattus norvegicus* - rno 5) *Xenopus* sp. - xen 6) *Gallus gallus* – gga 7) Other species – other

Optional arguments	
Argument	**Default Option**
--FASTA	None. If species_name is "other", then the path to the mitochondrial genome file has to be specified.
--index	None. If a bowtie index file of the mitochondrial genome is specified, it will be used to map the small RNAs or lncRNAs.
---GTF	None. The path to the GTF (General Transfer Format) file containing the annotation of mitochondrial genes has to be specified. The GTF file format has to be downloaded from ENSEMBL or has to meet the requirement of ENSEMBL GTF file format (https://www.ensembl.org/info/website/upload/gff.html)
-- graphical_output	No. If graphical output of the basic plots has to be generated, the user has to specify "yes" under graphical_output.
--input_path	Current working directory.
--output_path	Current working directory.
--files	None. If the files are in different locations, the absolute path of the files can be specified. Note: if –input_path is specified, --files cannot be specified. If –files argument is specified, files in the current working directory will not be analyzed, even though the output directory will be the current working directory – unless a different output path is specified using –output_path argument.