

starting to denoise

['--help']

Displaying help

-h --help display help

Input file options:

--csv_input [path] input file path in csv format

--fasta_input [path] input file path in fasta format

--fastq_input [path] input file path in fastq format

--joining_file [path] file path of an info output from DnoiseE. This option allows to use the information of previous runs of DnoiseE to return different joining criteria outputs without running the program again

-n --count_name [size/reads/count...] count name column 'size' by default

-p --sep [1/2/3] separator in case of csv input file

1=' ' (tab)

2=','

3=';'

-q --sequence [sequence/seq...] sequence column name, 'sequence' by default

-s --start_sample_cols [number] first sample column (1 == 1st col) if not given, just one column with total read counts expected (see README.md)

-z --end_sample_cols [number] last sample column (n == nst col) if not given, just one column with total read counts expected (see README.md)

Output file options:

--csv_output [path] common path for csv format

--fasta_output [path] common path for fasta format

-j --joining_criteria [1/2/3/4]

(default r_d criterion)

1-> will join by the lesser [abundance ratio / beta(d)]

2-> will join by the lesser abundance ratio (r criterion)

3-> will join by the lesser distance (d) value (d criterion)

4-> will provide all joining criteria in three different

outputs (all)

Other options:

-a --alpha [number] alpha value, 5 by default

-c --cores [number] number of cores, 1 by default

-e --entropy [number,number,number] entropy values (or any user-settable measure of variability) of the different codon positions. If -y is enabled and no values are given, default entropy values are computed from the data

-g --get_entropy get only entropy values from a dataset

-m --modal_length [number] when running DnoiseE with entropy correction, sequence length expected can be set, if not, modal_length is used and only sequences with modal_length + or - 3*n are accepted

-r --min_abund [number] minimum abundance filtering applied at the end of analysis, 1 by default

-u --unique_length only modal length is accepted as sequence length when running with entropy correction

-w --within_MOTU [MOTU/motu/...] MOTU column name. This option allows to run DnoiseE within MOTU. Is only available for --csv_input and --csv_output

-x --first_nt_codon_position [number] as DnoiseE has been developed for COI sequences amplified with Leray-XT primers, default value is 3 (i.e., the first nucleotide in the sequences is a third codon position)

-y --entropy_correction a distance correction based on entropy is performed (see <https://github.com/adriantich/DnoiseE>). If not enabled no correction for entropy is performed (corresponding to the standard Unoise formulation)