random_mer

Purpose: random_mer identifies random_mer sequences from single end adapter trimmed FASTQ files. To run the script on raw FASTQ files, the following dependencies are required: PYTHON pandas module, cutadapt, blast and bowtie.

Usage:

random_mer.py < path to genome FASTA file>[-- mapping_stats] [-- index] [-- input_path] [-- output_path] [-- files]

Required arguments are indicated within <> and one of the allowed options have 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. Detailed information regarding required and optional arguments are indicated below

Required arguments	
Argument	**Allowed Options**
--genome	The path to the genome FASTA have to be specified

Optional arguments	
Argument	**Default Option**
--input_path	Current working directory. If the FASTQ files are not in the current working directory, the path to input FASTQ files can be specified here.
--output_path	Current working directory.

--mapping_stats	Yes. If user does not want to compute the % of mappable reads, they can use the option "NO". This will skip the bowtie from mapping the random mer trimmed reads against the genome. Hence the % of mappable reads shows as "na"
--index	If –mapping_stats value is "Yes", then the path to bowtie index has to be specified for bowtie mapping.
--files	None. If the files are located 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.