adapt_find

Purpose: adapt_find identifies adapters sequences from single-end raw sequencing files in FASTQ format. To run the script on raw FASTQ files, the following dependencies are required: PYTHON pandas module, scipy, numpy, cutadapt, blast, and bowtie.

Usage:

adapt_find.py <sequencing platform> [-- min_len] [-- max_len] [-- index] [-- input_path] [-- output_path] [-- 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 by a user, then the default values will be used. Detailed information regarding required and optional arguments are indicated below.

Required arguments	
Argument	**Allowed Options**
Sequencing platform	1) ILLUMINA – for ILLUMINA sequencing technology 2) SOLID - for SOLID sequencing technology 3) ION_TORRENT – for Ion Torrent sequencing technology 4) 454 – for Roche 454 Sequencing technology

Optional arguments	
Argument	**Default Option**
--min_len	15

--max_len	50
--index	None. If the path to the bowtie genome index files is specified, the percentage of number of reads mapping to the genome after adapter trimming will be reported.
--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. Users can also specify a different output folder path.
--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 option 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.