

General parser scripts

Scripts found on fourierseq which are useful for parsing and filtering through fasta, fastq files and mapper output files. Scripts for base space, colorspace and mock base space conversion are also included.

Mapreads and SREK related parsers

- mapreads_interpreter: Converts a mapreads output file into a tab-delimited info file.
 - Inputs: mapreads output file; reference fasta file
 - Outputs: Info file with readid\tgi#\tmismatches\tdirection\tstartlocation\tstart%\tend%\tcoverage%\tgenedescription\tgenelength
- mapreads_interpreter_SREK: Converts a SREK mapping output file into a tab-delimited info file.
 - Inputs: SREK output file (after extension); reference fasta file
 - Outputs: Info file with readid\tgi#\tmismatches\tdirection\tstartlocation\tstart%\tend%\tcoverage%\tgenedescription\tgenelength\tfragmentlength* mapreads_select_mismatches: Will filter info file by number of mismatches
 - Inputs: info file generated by mapreads_interpreter; mismatch cutoff
 - Outputs: info file filtered to include only mappings with mismatches less than or equal to user specified cutoff* mapreads_select_by_length: Will filter info filter by length
 - Inputs: info file generated by mapreads_interpreter_SREK; minimum length; maximum length
 - Outputs: info file filtered to include only results with mapping length within user specified cutoff

Bowtie related parsers

- bowtie_interpreter: Converts a bowtie mapping output file into a tab-delimited info file.
 - Inputs: bowtie output file; reference fasta file
 - Outputs: Info file with readid\tgi#\tmismatches\tstart%\tend%\tcoverage%\tgenedescription\tgenelength; start locations for each mapping
- bowtie_select_mismatches: Will filter info file by number of mismatches
 - Inputs: info file generated by bowtie; mismatch cutoff
 - Outputs: info file filtered to include only mappings with mismatches less than or equal to user specified cutoff* bowtie_filter_all_repeats: Will filter the bowtie mapping output file to exclude poly-N reads (within a limit set by the user).
 - Inputs: bowtie output file; poly-N limit (a read with poly-N regions longer than or equal to the limit are excluded).
 - Outputs: filtered bowtie output file.