Bowtie Additional Commands

Options

Input

The query input files (specified either as `<m1>` and `<m2>`, or as `<s>`) are FASTQ files (usually having extension `.fq` or `.fastq`). This is the default. See also: `--solexa-quals` and `--integer-quals`.

- The query input files (specified either as `<m1>` and `<m2>`, or as `<s>`) are FASTA files (usually having extension `.fa`, `.mfa`, `.fna` or similar). All quality values are assumed to be 40 on the Phred quality scale.

- Raw files: one sequence per line, without quality values or names. All quality values are assumed to be 40 on the Phred quality scale.

- The query sequences are given on command line.

  - I.e. `<m1>`, `<m2>` and `<singles>` are comma-separated lists of reads rather than lists of read files.

  - Align in colorspace. Read characters are interpreted as colors. The index specified must be a colorspace index (i.e. built with `bowtie-build --C`, or `bowtie` will print an error message and quit. See Colorspace alignment for more details.

  - Comma-separated list of files containing quality values for corresponding unpaired CSFASTA reads. Use in combination with `--C` and `--f`, `--integer-quals` is set automatically when `--Q`/`--quals` is specified.

- Comma-separated list of files containing quality values for corresponding CSFASTA #1 mates. Use in combination with `--C`, `--f`, and `--integer-quals` is set automatically when `--Q1` is specified.

- Comma-separated list of files containing quality values for corresponding CSFASTA #2 mates. Use in combination with `--C`, `--f`, and `--integer-quals` is set automatically when `--Q2` is specified.

- Skip (i.e. do not align) the first `<int>` reads or pairs in the input.

- Only align the first `<int>` reads or read pairs from the input (after the `--skip` reads or pairs have been skipped). Default: no limit.

- Trim `<int>` bases from high-quality (left) end of each read before alignment (default: 0).
Trim <int> bases from low-quality (right) end of each read before alignment (default: 0).

Input qualities are ASCII chars equal to the Phred quality plus 33. Default: on.

Input qualities are ASCII chars equal to the Phred quality plus 64. Default: off.

Convert input qualities from Solexa (which can be negative) to Phred (which can't). This is usually the right option for use with (unconverted) reads emitted by GA Pipeline versions prior to 1.3. Default: off.

Same as --phred64-quals. This is usually the right option for use with (unconverted) reads emitted by GA Pipeline version 1.3 or later. Default: off.

Quality values are represented in the read input file as space-separated ASCII integers, e.g., 40 40 30 40..., rather than ASCII characters, e.g., II?I.... Integers are treated as being on the Phred quality scale unless --solexa-quals is also specified. Default: off.

Alignment

Report alignments with at most <int> mismatches. -e and -l options are ignored and quality values have no effect on what alignments are valid. -v is mutually exclusive with -n.

Maximum number of mismatches permitted in the "seed", i.e. the first L base pairs of the read (where L is set with -l/--seedlen). This may be 0, 1, 2 or 3 and the default is 2. This option is mutually exclusive with the -v option.

Maximum permitted total of quality values at all mismatched read positions throughout the entire alignment, not just in the "seed". The default is 70. Like Maq, bowtie rounds quality values to the nearest 10 and saturates at 30; rounding can be disabled with --nomaqround.

The "seed length"; i.e., the number of bases on the high-quality end of the read to which the -n ceiling applies. The lowest permitted setting is 5 and the default is 28. bowtie is faster for larger values of -l.

Maq accepts quality values in the Phred quality scale, but internally rounds values to the nearest 10, with a maximum of 30. By default, bowtie also rounds this way. --nomaqround prevents this rounding in bowtie.

The minimum insert size for valid paired-end alignments. E.g. if -I 60 is
specified and a paired-end alignment consists of two 20-bp alignments in the appropriate orientation with a 20-bp gap between them, that alignment is considered valid (as long as \(-x\) is also satisfied). A 19-bp gap would not be valid in that case. If trimming options \(-3\) or \(-5\) are also used, the \(-r\) constraint is applied with respect to the untrimmed mates. Default: 0.

The maximum insert size for valid paired-end alignments. E.g. if \(-x 100\) is specified and a paired-end alignment consists of two 20-bp alignments in the proper orientation with a 60-bp gap between them, that alignment is considered valid (as long as \(-I\) is also satisfied). A 61-bp gap would not be valid in that case. If trimming options \(-3\) or \(-5\) are also used, the \(-x\) constraint is applied with respect to the untrimmed mates, not the trimmed mates. Default: 250.

The upstream/downstream mate orientations for a valid paired-end alignment against the forward reference strand. E.g., if \(--fr\) is specified and there is a candidate paired-end alignment where mate1 appears upstream of the reverse complement of mate2 and the insert length constraints are met, that alignment is valid. Also, if mate2 appears upstream of the reverse complement of mate1 and all other constraints are met, that too is valid. \(--rf\) likewise requires that an upstream mate1 be reverse-complemented and a downstream mate2 be forward-oriented. \(--ff\) requires both an upstream mate1 and a downstream mate2 to be forward-oriented. Default: \(--fr\) when \(-C\) (colorspace alignment) is not specified, \(--ff\) when \(-C\) is specified.

If \(--nofw\) is specified, bowtie will not attempt to align against the forward reference strand. If \(--norc\) is specified, bowtie will not attempt to align against the reverse-complement reference strand. For paired-end reads using \(--fr\) or \(--rf\) modes, \(--nofw\) and \(--norc\) apply to the forward and reverse-complement pair orientations. I.e. specifying \(--nofw\) and \(--fr\) will only find reads in the R/F orientation where mate 2 occurs upstream of mate 1 with respect to the forward reference strand.

The maximum number of backtracks permitted when aligning a read in \(-n 2\) or \(-n 3\) mode (default: 125 without \(--best\), 800 with \(--best\)). A "backtrack" is the introduction of a speculative substitution into the alignment. Without this limit, the default parameters will sometimes require that bowtie try 100s or 1,000s of backtracks to align a read, especially if the read has many low-quality bases and/or has no valid alignments, slowing bowtie down significantly. However, this limit may cause some valid alignments to be missed. Higher limits yield greater sensitivity at the expensive of longer running times. See also: \(--y/\--\)
tryhard.
For paired-end alignment, this is the maximum number of attempts bowtie will make to match an alignment for one mate up with an alignment for the opposite mate. Most paired-end alignments require only a few such attempts, but pairs where both mates occur in highly repetitive regions of the reference can require significantly more. Setting this to a higher number allows bowtie to find more paired-end alignments for repetitive pairs at the expense of speed. The default is 100. See also: -y/--tryhard.

Try as hard as possible to find valid alignments when they exist, including paired-end alignments. This is equivalent to specifying very high values for the --maxbts and --pairtries options. This mode is generally much slower than the default settings, but can be useful for certain problems. This mode is slower when (a) the reference is very repetitive, (b) the reads are low quality, or (c) not many reads have valid alignments.

The number of megabytes of memory a given thread is given to store path descriptors in --best mode. Best-first search must keep track of many paths at once to ensure it is always extending the path with the lowest cumulative cost. Bowtie tries to minimize the memory impact of the descriptors, but they can still grow very large in some cases. If you receive an error message saying that chunk memory has been exhausted in --best mode, try adjusting this parameter up to dedicate more memory to the descriptors. Default: 64.

**Reporting**

Report up to <int> valid alignments per read or pair (default: 1). Validity of alignments is determined by the alignment policy (combined effects of -n, -v, -l, and -e). If more than one valid alignment exists and the --best and --strata options are specified, then only those alignments belonging to the best alignment "stratum" will be reported. Bowtie is designed to be very fast for small -k but bowtie can become significantly slower as -k increases. If you would like to use Bowtie for larger values of -k, consider building an index with a denser suffix-array sample, i.e. specify a smaller -o/--offrate when invoking bowtie-build for the relevant index (see the Performance tuning section for details).

Report all valid alignments per read or pair (default: off). Validity of alignments is determined by the alignment policy (combined effects of -n, -v, -l, and -e). If more than one valid alignment exists and the --best and --
strata options are specified, then only those alignments belonging to the best alignment "stratum" will be reported. Bowtie is designed to be very fast for small -k but bowtie can become significantly slower if -a/--all is specified. If you would like to use Bowtie with -a, consider building an index with a denser suffix-array sample, i.e. specify a smaller -o/--offrate when invoking bowtie-build for the relevant index (see the Performance tuning section for details).

Suppress all alignments for a particular read or pair if more than <int> reportable alignments exist for it. Reportable alignments are those that would be reported given the -n, -v, -l, -e, -k, -a, --best, and --strata options. Default: no limit. Bowtie is designed to be very fast for small -m but bowtie can become significantly slower for larger values of -m. If you would like to use Bowtie for larger values of -k, consider building an index with a denser suffix-array sample, i.e. specify a smaller -o/--offrate when invoking bowtie-build for the relevant index (see the Performance tuning section for details).

Behaves like -m except that if a read has more than <int> reportable alignments, one is reported at random. In default output mode, the selected alignment's 7th column is set to <int>+1 to indicate the read has at least <int>+1 valid alignments. In -S/--sam mode, the selected alignment is given a MAPQ (mapping quality) of 0 and the XM:I field is set to <int>+1. This option requires --best; if specified without --best, --best is enabled automatically.

Make Bowtie guarantee that reported singleton alignments are "best" in terms of stratum (i.e. number of mismatches, or mismatches in the seed in the case of -n mode) and in terms of the quality values at the mismatched position(s). Stratum always trumps quality; e.g. a 1-mismatch alignment where the mismatched position has Phred quality 40 is preferred over a 2-mismatch alignment where the mismatched positions both have Phred quality 10. When --best is not specified, Bowtie may report alignments that are sub-optimal in terms of stratum and/or quality (though an effort is made to report the best alignment). --best mode also removes all strand bias. Note that --best does not affect which alignments are considered "valid" by bowtie, only which valid alignments are reported by bowtie. When --best is specified and multiple hits are allowed (via -k or -a), the alignments for a given read are guaranteed to appear in best-to-worst order in bowtie's output. bowtie is somewhat slower when --best is specified.

If many valid alignments exist and are reportable (e.g. are not disallowed via the -k option) and they fall into more than one alignment "stratum", report only those alignments that fall into the best stratum. By default, Bowtie
reports all reportable alignments regardless of whether they fall into multiple strata. When `--strata` is specified, `--best` must also be specified.

**Output**

- `-t/--time`
  Print the amount of wall-clock time taken by each phase.

- `-B/--offbase <int>`
  When outputting alignments, number the first base of a reference sequence as `<int>`. Default: 0.

- `--quiet`
  Print nothing besides alignments.

- `--refout`
  Write alignments to a set of files named `refXXXXX.map`, where `XXXXX` is the 0-padded index of the reference sequence aligned to. This can be a useful way to break up work for downstream analyses when dealing with, for example, large numbers of reads aligned to the assembled human genome. If `<hits>` is also specified, it will be ignored.

- `--refidx`
  When a reference sequence is referred to in a reported alignment, refer to it by 0-based index (its offset into the list of references that were indexed) rather than by name.

- `--al <filename>`
  Write all reads for which at least one alignment was reported to a file with name `<filename>`. Written reads will appear as they did in the input, without any of the trimming or translation of quality values that may have taken place within bowtie. Paired-end reads will be written to two parallel files with `_1` and `_2` inserted in the filename, e.g., if `<filename>` is `aligned.fq`, the #1 and #2 mates that fail to align will be written to `aligned_1.fq` and `aligned_2.fq` respectively.

- `--un <filename>`
  Write all reads that could not be aligned to a file with name `<filename>`. Written reads will appear as they did in the input, without any of the trimming or translation of quality values that may have taken place within Bowtie. Paired-end reads will be written to two parallel files with `_1` and `_2` inserted in the filename, e.g., if `<filename>` is `unaligned.fq`, the #1 and #2 mates that fail to align will be written to `unaligned_1.fq` and `unaligned_2.fq` respectively. Unless `--max` is also specified, reads with a number of valid alignments exceeding the limit set with the `-m` option are also written to `<filename>`.

- `--max <filename>`
  Write all reads with a number of valid alignments exceeding the limit set with the `-m` option to a file with name `<filename>`. Written reads will appear as they did in the input, without any of the trimming or translation of quality values that may have taken place within bowtie. Paired-end reads will be written to two parallel files with `_1` and `_2` inserted in the filename, e.g., if `<filename>` is `max.fq`, the #1 and #2 mates that exceed
the \texttt{--m} limit will be written to \texttt{max\_1.fq} and \texttt{max\_2.fq} respectively. These reads are not written to the file specified with \texttt{--un}.

Suppress columns of output in the default output mode. E.g. if \texttt{--suppress} \texttt{1,5,6} is specified, the read name, read sequence, and read quality fields will be omitted. See Default Bowtie output for field descriptions. This option is ignored if the output mode is \texttt{-S|--sam}.

\texttt{--fullref} Print the full reference sequence name, including whitespace, in alignment output. By default \texttt{bowtie} prints everything up to but not including the first whitespace.