Release Notes

BCL Convert v4.0.5

INTRODUCTION

These Release Notes detail the latest release of BCL Convert, including known issues.

BCL Convert converts per cycle binary data output by Illumina sequencers containing basecall files and quality scores to per read FASTQ files.

NEW FEATURES

- Per tile and per cycle primary statistics output to report files
  - Adapter_Cycle_Metrics: provides, for each cycle and sample, the number of reads mapping to the sample where the adapter was detected beginning at that cycle.
  - Demultiplex_Tile_Stats.csv: includes all columns that exist in the Demultiplex_Stats.csv output file but at the per sample and per tile level
  - Quality_Tile_Stats.csv: includes all columns that exist in the Quality_Stats.csv output file but at the per sample and per tile level
- Support Sample_Name column in the Sample Sheet Data section
  - Only allowed when –sample-name-column-enabled is enabled
  - Must be specified for every sample when enabled
  - Fastq filename will be named according to Sample_Name in sample sheet.
  - Fastq files will be output to subdirectory name by Sample_ID when –sample-name-column-enabled is enabled and –-bcl-sampleproject-subdirectories is enabled
  - Reports will include Sample_Name and Sample_Project values when these features are enabled on the command line.
- Allow legacy FindAdaptersWithIndels sample sheet setting (default off)
  - Provides identical output to bcl2fastq2 2.20 for default adapter trim settings.
- Allow for no sample sheet to be provided using the –no-sample-sheet command
  - Default is disabled (sample sheet required)
  - When enabled, all sequences will go to the ‘Undetermined’ fastq files
  - Cannot be enabled with any of the following options
    - bcl-sampleproject-subdirectories
    - sample-name-column-enabled
    - bcl-only-matched-reads
    - num-unknown-barcodes-reported
    - bcl-validate-sample-sheet-only
    - sample-sheet
- Can specify fastq gzip compression level
  - Command line option is --fastq-gzip-compression-level and default is 1
  - 0 through 9 are allowed to be specified
- Corruption detection for BCI input files
  - In strict mode, will abort conversion with an error
  - In robust mode, will skip processing lane

RESOLVED ISSUES IN V4.0.5

- Fix a crash that can occur when using per-sample-settings with higher sample counts in a lane, due to a hash-table pre-size using a signed integer as input that is overflowed.

- Fix BCL FASTQ file paths in the fastq_list.csv when ORA-interleaved compression format is used. Previously the fastq_list.csv file contained two files (instead of one single interleaved file) under the "Read1File" and "Read2File" columns and the files were not named correctly.
- Some versions of RTA3 outputs CBCL with 0 qual + nonzero base for 0/"N#". In other words, masked nibbles (0 qual) do not zero out the base. bcl2fastq2 has a masking step, but DRAGEN™ and bcl-convert did not. This fix adds masking so that DRAGEN™ and bcl-convert matches bcl2fastq2 for those RTA3 outputs.
- Fix for no-sample-sheet setting that omits index sequences from fastq headers.
- Fix to remove BCL conversion thread settings limit of 64, a regression in v3.10, to allow runs on high core count systems.
- Fix a crash due to threading error on a 150k sample dataset.
- Fix for Index_Hopping_Counts.csv containing incorrect Same_{Name,Project}.

Resolved Issues in v4.0.3

- --tiles command line option did not work correctly with bgzf inputs, resulting in fastq files with incorrect tile header. This will now be correct.
- Barcode Collision Error and Solution: Previous versions of BCL Convert allowed a conversion to continue if either index (i7 or i5) had sufficient hamming distance from all other samples in the lane. BCL Convert 3.10.5 uses stricter barcode collision logic to support increased high-throughput and complex sample pooling. Each index in a dual setup must individually meet the hamming distance requirements set by the BarcodeMismatchesIndex# value. If either i7 or i5 does not meet the hamming distance requirements the program will error. For default BarcodeMismatchesIndex1 and BarcodeMismatchesIndex2:
    - Barcodes must differ by at least three bases.
    - If any two samples in i7 differ by fewer than 3 bases, an error is produced and the run will not proceed, regardless of their i5 values.
    - If any two samples in i5 differ by fewer than 3 bases, an error is produced regardless of their i7 values.

  If you receive errors with current versions of DRAGEN or BCL Convert, lower the mismatch tolerance for the index producing the error by using the BarcodeMismatchesIndex1 or BarcodeMismatchesIndex2 sample sheet settings.

Known Issues

- If a directory is specified as input to '--sample-sheet', BCL Convert will hang at the beginning of a run while trying to copy that path as a file to <outdir>/Reports/SampleSheet.csv
- BCL Convert does not validate when “Logs” or “Reports” is provided for a Sample_Project, and the software will be unable to create the subdirectories if these string are provided.
- BCL Convert does not support the --first-tile-only option being specified for SP flow cells, but the new ***–tiles*** option can be used as a substitute.
- Does not error when no tile list exists in the RunInfo.xml file and –tile or –exclude-tiles is specified in the command line

- ORA Compression cannot be specified on the command line when –no-sample-sheet option is used. Therefore gzip compression is the only type of compression supported when no sample sheet is provided

RELEASE HISTORY

Revision	Release Reference	Originator	Description of Change
00	1085822	Daniel Tracy	Initial release