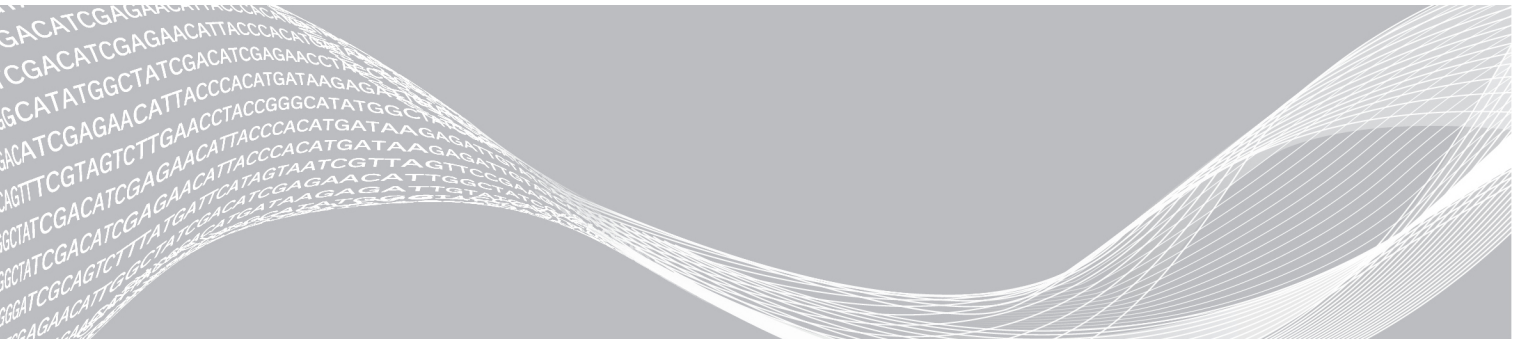
illumina®

Local Run Manager
DNA GenerateFASTQ Dx Analysis Module

Workflow Guide for MiSeqDx

FOR IN VITRO DIAGNOSTIC USE



Document # 200015661 v01

May 2022

FOR IN VITRO DIAGNOSTIC USE

ILLUMINA PROPRIETARY

Overview

The Local Run Manager DNA GenerateFASTQ Dx module first demultiplexes indexed reads. If present, DNA GenerateFASTQ Dx generates intermediate output files in the FASTQ file format, and then exits the workflow. No alignment or further analysis is performed. FASTQ files are required input for analysis with third-party analysis tools.

The Local Run Manager DNA GenerateFASTQ Dx module can be run on Local Run Manager v3.1.0 (or later) and is compatible with Windows 10. The analysis module supports sequencing to analysis for the Illumina DNA Prep with Enrichment Dx assay.

About This Guide

This guide provides instructions for setting up run parameters for sequencing and analysis for the DNA GenerateFASTQ Dx analysis module. Use of the software requires basic knowledge of the current Windows operating system and web browser-based user interface. For information about the Local Run Manager dashboard and system settings, refer to the *Local Run Manager Software Reference Guide for MiSeqDx (document # 200003931)*.

Enter Run Information

Set Parameters

1 Log in to Local Run Manager.

2 Select **Create Run**, and then select **DNA GenerateFASTQ Dx.**

3 Enter a unique run name that identifies the run from sequencing through analysis (40 characters or less).
The run name can contain alphanumeric characters, spaces, and the special characters `.~!@#$%-_{}.
You cannot use a name from a prior run.

4 **[Optional]** Enter a run description to help identify the run (150 characters or less).
The run description can contain alphanumeric characters, spaces, and the following special characters:
`.~!@#$%-_{}.

5 Configure the following run settings:
- Index Plate—Select the index plate layout used during library preparation. You can select from Index Set A, Index Set B, and Index Set AB. Refer to the *Illumina DNA Prep With Enrichment Dx Package Insert* for information on the index plate layouts.
  Index sets A and B contain 96 samples and the corresponding unique dual primers (UDPs). Index set AB contains 192 samples and the corresponding UDPs.
- Read Type—Select single read or paired-end. The default read type is paired-end.
- Read Lengths—Enter the read length. The default read length is 151.

6 Under Module-Specific Settings, set the Adapter Trimming option.
Adapter trimming is enabled by default.

7 Select the number of samples to be sequenced. The selected number of samples includes autopopulated UDP recommendations. If you do not want to use UDP recommendations, select **Custom**.

If the number of samples you are sequencing is not included in the drop-down list, select the nearest number of samples. Make sure that the selected number is less than the number being sequenced and to add additional UDPs as needed. For example, to test 18 samples, select the 16 samples option.

Specify Samples for the Run

Specify samples for the run using one of the following options.

- **Enter samples manually**—Use the blank table on the Create Run screen.
- **Import samples**—Navigate to an external file in a comma-separated values (*.csv) format. A template is available for download on the Create Run screen.

Enter Samples Manually

1 Enter a unique sample ID in the Sample ID tab. Use alphanumeric characters and/or dashes (40 characters or less).
The sample ID and corresponding sample description and UDP position are highlighted in blue to indicate that the sample is entered.

2 [Optional] To select positive and negative control samples, right-click on sample wells.

3 [Optional] Enter a sample description in the sample Description tab. The sample description can contain alphanumeric characters, periods, and the special characters `~!@#$%-_{}. Spaces are not permitted.
If the sample ID associated with the sample description is used again in a later run, the initial sample description is overwritten.

4 Modify the recommended UDP positions as needed. The suggested sample well positions are highlighted in yellow, purple, orange, and pink.
If using suggested sample wells, the software autopopulates UDP index adapters that meet diversity index requirements. If the number of samples you selected is not the exact number of samples you are testing, make sure to select UDP index adapters for the extra wells.

5 [Optional] Select **Export Samples** to export the sample information file.

6 Select **Save Run**.

Import Sample Sheet

You can import sample information from a sample information file previously exported from the DNA GenerateFASTQ Dx module using the Export Samples feature or a template file, which can be generated by selecting **Template** on the Create Run screen. See *Enter Samples Manually* on page 4 for instructions on how to create and export sample information.

The template file does not include the autopopulated UDP recommendations.

To edit the template file:

1 Select **Template** on the Create Run screen to make a new plate layout. The template file contains the correct column headings for import. Edit the file as follows.

   a Open the sample sheet in a text editor.
   b Enter the required sample information.
   c Save the file in comma-separated values (*.csv) format. Make sure that the sample IDs are unique.

To import sample information:

2 Select **Import Samples**, and then select the CSV file.

3 [Optional] Select **Export** to export sample information to an external file.

4 Select **Save Run**.

Edit a Run

For instructions on editing the information in your run before sequencing, refer to *Local Run Manager Software Reference Guide for MiSeqDx (document # 200003931)*.

Analysis Methods

The DNA GenerateFASTQ Dx analysis module performs the following analysis steps and then writes analysis output files to the Alignment folder.

- Demultiplexes index reads
- Generates FASTQ files

Demultiplexing

Demultiplexing compares each Index Read sequence to the index sequences specified for the run. No quality values are considered in this step.

Index reads are identified using the following steps:

- Samples are numbered starting from 1 based on the order they are listed for the run.
- Sample number 0 is reserved for clusters that were not assigned to a sample.
- Clusters are assigned to a sample when the index sequence matches exactly or when there is up to a single mismatch per Index Read.

FASTQ File Generation

After demultiplexing, the software generates intermediate analysis files in the FASTQ format, which is a text format used to represent sequences. FASTQ files contain reads for each sample and the associated quality scores. Any controls used for the run and clusters that did not pass filters are excluded.

Each FASTQ file contains reads for only one sample, and the name of that sample is included in the FASTQ file name. FASTQ files are the primary input for alignment.

View Run and Results

1 From the Local Run Manager dashboard, select the run name.

2 From the Run Overview tab, review the sequencing run metrics.

3 To change the analysis data file location for future requeues of the selected run, select the **Edit** icon, and edit the output run folder file path.
You cannot edit the name of the output run folder.

4 [Optional] Select the **Copy to Clipboard** to copy the output run folder file path.

5 Select the Sequencing Information tab to review run parameters and consumables information.

6 Select the Samples & Results tab to view the analysis report.
   - If analysis was requeued, select the appropriate analysis from the Select Analysis drop-down list.
   - From the left navigation bar, select a sample ID to view the report for another sample.

7 [Optional] Select the Copy to Clipboard to copy the Analysis Folder file path.

Results Report

Results are summarized on the Samples and Results tab.

Samples

Table 1 Samples Table

Column Heading	Description
Sample ID	The sample ID provided when the run was created.
Plate	The plate provided with the index plate when the run was created. The column is only shown if the index plate AB is selected.
Index Well	The index well provided with the sample well location when the run was created.
Description	The sample description provided when the run was created.
UDP	The UDP used with the sample.
Control	The positive or negative control used with the sample.

Indexing

Table 2 Indexing Table

Column Heading	Description
Index Number	An assigned ID based on the order that samples are listed in the sample table.
Sample ID	The sample ID provided when the run was created.
UDP	The UDP used with the sample.
% Reads Identified (PF)	The percentage of reads that passed filters.

Analysis Output Files

The following analysis output files are generated for the DNA GenerateFASTQ Dx analysis module.

File Name	Description
Demultiplexing (*.demux)	Intermediate files containing demultiplexing results.
FASTQ (*.fastq.gz)	Intermediate files containing quality scored base calls. FASTQ files are the primary input for the alignment step.

Demultiplexing File Format

The process of demultiplexing reads the index sequence attached to each cluster to determine from which sample the cluster originated. The mapping between clusters and sample number is written to a demultiplexing (*.demux) file for each tile of the flow cell.

The demultiplexing file naming format is s_1_X.demux, where X is the tile number.

Demultiplexing files start with a header:

- Version (4 byte integer), currently 1
- Cluster count (4 byte integer)

The remainder of the file consists of sample numbers for each cluster from the tile.

When the demultiplexing step is complete, the software generates a demultiplexing file named DemultiplexSummaryF1L1.txt.

- In the file name, F1 represents the flow cell number.
- In the file name, L1 represents the lane number.
- Demultiplexing results in a table with one row per tile and one column per sample, including sample 0.
- The most commonly occurring sequences in index reads.

FASTQ File Format

FASTQ is a text-based file format that contains base calls and quality values per read. Each record contains 4 lines:

- The identifier
- The sequence
- A plus sign (+)
- The Phred quality scores in an ASCII + 33 encoded format

The identifier is formatted as:

@Instrument:RunID:FlowCellID:Lane:Tile:X:Y ReadNum:FilterFlag:0:SampleNumber

Example:

```
@SIM:1:FCX:1:15:6329:1045 1:N:0:2
TCGCACTCAACGCCCTGCATATGACAAGACAGAATC
+
<>;##=><9=AAAAAAAAAA9#:<#<;<<<????#=
```

Supplementary Output Files

The following output files provide supplementary information, or summarize run results and analysis errors. Although these files are not required for assessing analysis results, they can be used for troubleshooting purposes. All files are located in the Alignment folder unless otherwise specified.

File Name	Description
AdapterTrimming.txt	Lists the number of trimmed bases and percentage of bases for each tile. This file is present only if adapter trimming was specified for the run.
AnalysisLog.txt	Processing log that describes every step that occurred during analysis of the current run folder. This file does not contain error messages. Located in the root level of the run folder.
AnalysisError.txt	Processing log that lists any errors that occurred during analysis. This file will be empty if no errors occurred. Located in the root level of the run folder.
CompletedJobInfo.xml	Written after analysis is complete, contains information about the run, such as date, flow cell ID, software version, and other parameters. Located in the root level of the run folder.
Checksum.csv	Contains the file names and unique checksum values for determined and undetermined FASTQ files, BCL files, and the SampleSheetUsed.csv file.
DemultiplexSummaryF1L1.txt	Reports demultiplexing results in a table with 1 row per tile and 1 column per sample.
GenerateFASTQRunStatistics.xml	Contains summary statistics specific to the run. Located in the root level of the run folder.

Analysis Folder

The analysis folder holds the files generated by the Local Run Manager software.

The relationship between the output folder and analysis folder is summarized as follows:

- During sequencing, Real-Time Analysis (RTA) populates the output folder with files generated during image analysis, base calling, and quality scoring.
- RTA copies files to the analysis folder in real time. After RTA assigns a quality score to each base for each cycle, the software writes the file RTAComplete.xml to both folders.
- When the file RTAComplete.xml is present, analysis begins.
- As analysis continues, Local Run Manager writes output files to the analysis folder, and then copies the files back to the output folder.

Alignment Folders

Each time that analysis is requeued, the Local Run Manager creates an alignment folder named **Alignment_N**, where N is a sequential number.

Folder Structure

Data

Alignment_## or Alignment_Imported_##

[Timestamp of Run]

DataAccessFiles

Fastq

- FastqSummaryF1L1.txt
- Sample1_S1_L001_R1_001.fastq.gz
- Sample2_S2_L001_R2_001.fastq.gz
- Undetermined_S0_L001_R1_001.fastq.gz
- Undetermined_S0_L001_R2_001.fastq.gz

Logging

- BuildFastq0.stdout.txt
- BuildFastq1.stdout.txt
- commands.txt

Plots

- AdapterCounts.txt
- AdapterTrimming.txt
- AnalysisError.txt
- AnalysisLog.txt
- Checkpoint.txt
- Checksum.csv
- CompletedJobInfo.xml
- DemultiplexSummaryF1L1.txt

- GenerateFASTQRunStatistics.xml
- SampleSheetUsed.csv

Base Calling and Index Diversity

When samples are sequenced on the MiSeqDx instrument, base calling determines a base (A, C, G, or T) for every cluster of a given tile, or imaging area on the flow cell, at a specific cycle. The MiSeqDx instrument uses four-channel sequencing, which requires four images to encode the data for four DNA bases, two from the red channel and two from the green channel.

The process for base calling index reads differs from base calling during other reads.

When selecting indexes during run creation a low diversity warning will appear if the indexes do not meet diversity requirements. To prevent the low diversity warning, select index sequences that provide signal in both channels for every cycle.

- Red channel—A or C
- Green channel—G or T

This base calling process ensures accuracy when analyzing low-plex samples. For more information about the sequences of your indexes, refer to the *Illumina DNA Prep With Enrichment Dx Package Insert*.

During run creation in Local Run Manager you will choose the number of samples to be tested. Suggested index combinations that meet index diversity requirements are auto populated by the software. Although you are not required to use the suggested UDP index combinations, it is recommended.

Revision History

Document	Date	Description of Change
Document # 200015661 v01	May 2022	Added address for Australian sponsor. Clarified sample description limitation.
Document # 200015661 v00	Feb 2022	Initial release

Technical Assistance

For technical assistance, contact Illumina Technical Support.

Website: www.illumina.com
Email: techsupport@illumina.com

Illumina Customer Support Telephone Numbers

Region	Toll Free	Regional
North America	+1.800.809.4566	
Australia	+1.800.775.688	
Austria	+43 800006249	+43 19286540
Belgium	+32 80077160	+32 34002973
China	400.066.5835	
Denmark	+45 80820183	+45 89871156
Finland	+358 800918363	+358 974790110
France	+33 805102193	+33 170770446
Germany	+49 8001014940	+49 8938035677
Hong Kong, China	800960230	
Ireland	+353 1800936608	+353 016950506
Italy	+39 800985513	+39 236003759
Japan	0800.111.5011	
Netherlands	+31 8000222493	+31 207132960
New Zealand	0800.451.650	
Norway	+47 800 16836	+47 21939693
Singapore	+1.800.579.2745	
South Korea	+82 80 234 5300	
Spain	+34 911899417	+34 800300143
Sweden	+46 850619671	+46 200883979
Switzerland	+41 565800000	+41 800200442
Taiwan, China	00806651752	
United Kingdom	+44 8000126019	+44 2073057197
Other countries	+44.1799.534000	

Safety data sheets (SDSs)—Available on the Illumina website at support.illumina.com/sds.html.

Product documentation—Available for download from support.illumina.com.

Document # 200015661 v01

Illumina
5200 Illumina Way
San Diego, California 92122 U.S.A.
+1.800.809.ILMN (4566)
+1.858.202.4566 (outside North America)
techsupport@illumina.com
www.illumina.com

IVD

EC REP

Illumina Netherlands B.V.
Steenoven 19
5626 DK Eindhoven
The Netherlands

Australian Sponsor
Illumina Australia Pty Ltd
Nursing Association Building
Level 3, 535 Elizabeth Street
Melbourne, VIC 3000
Australia

FOR IN VITRO DIAGNOSTIC USE

illumina®