

DRAGEN Microbial Amplicon for MiSeq i100 v1.1.7

Table of Contents

1	App Highlights	3
1.1	Overview	3
1.2	Input	3
1.3	Pipeline steps	4
1.4	Output	4
1.5	Important Notes	4
2	Pipeline Configuration	5
2.1	Reference FASTA (CustomFasta)	6
2.2	Reference BED (CustomBed).....	6
2.3	PCR primer definition (CustomPrimerDefinition)	6
3	Modes of Operation	7
3.1	Cloud and Local Execution:	7
3.2	Requeues:	7
4	Output Files	8
4.1	Summary	8
4.2	Toplevel DRAGEN.....	8
4.3	Workflow Level Output	8
5	How To Install DRAGEN and DRAGEN Applications	10
5.1	Install DRAGEN Versions:.....	10
5.2	Uninstall DRAGEN Versions:.....	10
5.3	Application Installation:	10
5.4	View Application Settings:	11
5.5	Uninstall Applications	11
	Release History	12

1 App Highlights

1.1 Overview

DRAGEN Microbial Amplicon is a software application designed to analyze sequencing data from amplicon library preps (both DNA and RNA) on microbiological samples, with an emphasis on viruses. Illumina sequencing reads are processed to generate consensus sequences that represent a best estimate of the population of viral sequences in each sample. Where appropriate, these consensus sequences are further analyzed by the phylogenetic analysis tools Nextclade and/or Pangolin to provide an identification of the clade or lineage of each sequence.

1.2 Input

Data can be provided in one of the following ways:

- Samples / biosamples with FASTQ datasets (see details in library preparation documents)
- A project containing one or more samples / biosamples with FASTQ datasets
 - All samples / biosamples in the selected project will be analyzed

Supported amplicon primer schemes

- Chikungunya
 - [Grubaugh Lab](#)
 - Illumina
- Dengue
 - Serotype 1 - Illumina
 - All serotypes - [DengueSeq from Grubaugh Lab](#)
- Influenza [A/B](#) - Universal
- Mpox
 - Pan-clade - [ARTIC](#)
 - Clade I - Illumina
 - Clade II - [Grubaugh Lab](#)
- RSV
 - [CDC](#)
 - [WCCRRRI](#)
- SARS-CoV-2 - ARTIC
 - [v5.4.2](#)
 - [v5.3.2, v4.1, v4, v3](#)
- Zika - [Grubaugh Lab](#)

Custom genome and primer sets

Users can upload custom files to provide user-defined reference genome set and primer definitions. Multiplexed amplicon panels targeting multiple organisms in the same reaction are supported.

1.3 Pipeline steps

1. Trim and filter reads
2. Remove off-target reads
3. For organisms with multiple candidates, trim primer sequences, perform assembly, cluster contigs, and map contigs to candidate reference genomes, then select reference genomes based on the mapping
4. Align reads to the default reference genome or selected reference genomes using DRAGEN
5. Trim primer sequences in aligned reads based on coordinates
6. Filter out samples with insufficient amplicon coverage
7. Call sequence variants from the alignments and apply them to the corresponding reference genomes to create consensus sequences
8. If applicable, run Nextclade/Pangolin on the consensus sequences

1.4 Output

- Consensus sequences representing a best estimate of targeted sequences
- Tables and plots reporting read counts, coverage, and Nextclade/Pangolin results

1.5 Important Notes

- The sequences are labeled according to the best match in the reference database, which is not exhaustive and the labels should not be taken as definitive for strain-typing. If strain typing is needed, the built-in Nextclade and/or Pangolin tools can be used for supported organisms. Alternatively, a BLAST or similar search of nucleotide databases may provide a more detailed match.
- Because of sequence homology, it is possible that organisms with very few reads will result in the generation of a sequence not present (false positive). Although the de novo assembly step of this software largely mitigates such instances, sequences with very low horizontal coverage (< 5%) should be treated with caution.

2 Pipeline Configuration

DragenMicrobialAmplicon_Settings

Parameter	Required	Description
SoftwareVersion	Yes	The version of the DRAGEN software used to process the DRAGEN 16s Plus pipeline, including conversion to FASTQ, specified using all three integers included in the DRAGEN version name. For example, 4.4.6.
AppVersion	Yes	The version of the workflow-specific application (i.e., DRAGEN Enrichment), using all three integers included in the version name. For example, 1.0.0.
KeepFastQ	Yes	Select whether FASTQs are saved (true) or discarded (false).
ReferencePrimerSetId	Yes	A string containing one of: chik-grubaugh chik-ilmn custom denv-grubaugh denv1-ilmn inf_ab-universal inf_a-universal inf_b-universal mpox_panclade-artic mpox_cladei-ilmn mpox_cladeii-grubaugh rsv-cdc20 rsv-wccrri sc2-artic_v3 sc2-artic_v4 sc2-artic_v4.1 sc2-artic_v5.3.2 sc2-artic_v5.4.2 sc2-none zika-grubaugh
CustomFasta	Conditionally required	A file name as an alphanumeric string with _ or - or . with no spaces allowed. Required if ReferencePrimerSetId is set to custom
CustomBed	No	A file name as an alphanumeric string with _ or - or . with no spaces allowed. Not required but can be provided only if ReferencePrimerSetId is set to custom
CustomPrimerDefinition	No	A file name as an alphanumeric string with _ or - or . with no spaces allowed. Not required but can be provided only if ReferencePrimerSetId is set to custom

DragenMicrobialAmplicon_Data

Parameter	Required	Description
Sample_ID	Yes	Alphanumeric string with _ or - with no spaces allowed

2.1 Reference FASTA (CustomFasta)

This is a required input file if the 'Custom' option is selected for 'Amplicon Primer Set'. This file contains one or more reference sequences as the target for short read alignment and as the basis for generating consensus sequences, but not all provided sequences may be used.

The file can contain sequences from multiple sources. For example, it can contain sequences of multiple segments from different Influenza genomes. In such cases, providing a corresponding Custom Reference BED file is highly recommended to inform the app which sequences should be belong together.

The app generates the required DRAGEN hash tables and other auxiliary files automatically, so there is no need to process the FASTA file with a separate app.

More details can be found here: <https://help.connected.illumina.com/infectious-disease/dragen-microbial-amplicon/dragen-microbial-amplicon/custom/fasta>

2.2 Reference BED (CustomBed)

A BED-like tab-separated value (TSV) file with no header row and with 4 or 5 columns:

1. `accession`: each sequence accession as it appears in Custom Reference FASTA header
2. `start`: start position (always set to 0)
3. `end`: end position (sequence length)
4. `genome`: full name of the virus the sequence belongs to (e.g. Influenza A H1N1)
5. (optional) `segment`: how this sequence is labeled within the virus (e.g. Segment 4 (HA)). Set to 'Full' if the sequence is the full genome

More details can be found here: <https://help.connected.illumina.com/infectious-disease/dragen-microbial-amplicon/dragen-microbial-amplicon/custom/bed>

2.3 PCR primer definition (CustomPrimerDefinition)

If provided, primer definition files are used to:

- Trim primer sequences from input reads to minimize artifacts introduced by primer sequences (e.g. false positive variant called from a mismatch between primer and reference sequences)
- Compute per-amplicon coverage to infer if there is sufficient sample titer for variant calling and consensus sequence generation

More details can be found here: <https://help.connected.illumina.com/infectious-disease/dragen-microbial-amplicon/dragen-microbial-amplicon/custom/pcr-primer-definition>

3 Modes of Operation

3.1 Cloud and Local Execution:

DRAGEN can be scheduled to operate in the following modes:

- Cloud:
 - Run planning:
 - Occurs in the Cloud via the Cloud User Interface (UI), or alternatively, sample sheets can be imported into the Cloud UI
 - Sequencing execution:
 - Runs are identified on the sequencer, and are executed locally on the instrument
 - DRAGEN Execution:
 - Occurs in the Cloud automatically after sequencing is complete
- Local:
 - Run planning:
 - Occurs on the instrument via the Instrument UI, or alternatively, sample sheets can be imported into the Instrument UI
 - Sequencing execution:
 - Runs are executed locally on the instrument
 - DRAGEN Execution:
 - Occurs locally on the instrument, automatically after sequencing is complete

3.2 Requeues:

After a sequencing run is complete, it may be desirable to run or re-run DRAGEN with, for example, an updated configuration. Requeues can occur on the instrument or in the Cloud. For on-instrument requeues, the following pre-requisites are necessary:

- The BCL's must be stored on the instrument
- The system must be idle

4 Output Files

4.1 Summary

- Consensus sequences representing a best estimate of targeted sequences
- Tables and plots reporting read counts, coverage, and Nextclade/Pangolin results

4.2 Toplevel DRAGEN

<run_id>/Analysis/<no>/inputs

- <SampleSheet>.csv

<run_id>/Analysis/<no>/Data

- Secondary_Analysis_Complete.txt
- summary
 - <x.y.z> (Note: DRAGEN version)
 - highlevel_summary.json
 - detailed_summary.json
- AggregateReports
 - report.html
 - report_files
 - *Links to workflow level reports*
- Demux
 - AggregateReports
 - *Links to lower-level reports*
 - Demultiplex_Stats.csv
 - *other stats*
- Dragen<Workflow>
 - Data (including FASTQ) for samples configured for the workflow (see below)
- RunInstrumentAnalyticsMetrics
- logs
 - *high-level logs*

4.3 Workflow Level Output

<run_id>/Analysis/<no>/Data/DragenMicrobialAmplicon

- Dragen<workflow>
 - AggregateReports
 - *Links to workflow specific sample level reports*

- fastq (or ora_fastq)
 - <sample_ID>.S0_Rm_001.fastq.gz or *.fastq.ora (m=1-2)
 - *Additional samples*
 - Reports
 - *Adapter_Metrics.csv*
 - *Quality_Metrics.csv*
 - *Additional metrics files*
- <sample_ID>
 - sample-specific outputs
 - amplicon
 - consensus
 - map_align
 - metrics
 - variant_calling
- logs
 - *Workflow specific logs*

5 How To Install DRAGEN and DRAGEN Applications

5.1 Install DRAGEN Versions:

- When a new DRAGEN version is available, download the DRAGEN installer (*.ires) from the MiSeq i100 Series support page. Save the installer locally or to a network drive.
- Make sure that there are no sequencing runs or on-instrument secondary analysis in progress.
- Select the instrument icon to open the global navigation menu.
- Select Settings, and then select DRAGEN.
- On the Versions tab, select Install version.
- Navigate to the installer, and then select Open.
- Select Install. A message indicates if the installation was successful or failed

5.2 Uninstall DRAGEN Versions:

- Select the instrument icon to open the global navigation menu.
- Select Settings, and then select DRAGEN.
- To uninstall a previous DRAGEN version, do as follows.
 - a. On the Versions tab, select the ellipsis icon in the Actions column.
 - b. Select Uninstall.
 - c. Select Yes, uninstall.
- To uninstall the latest DRAGEN version, do as follows.
 - a. On the Versions tab, select the ellipsis icon in the Actions column.
 - b. Select Uninstall all.
 - c. Select Yes, uninstall all

5.3 Application Installation:

- Download the application (*.iapp) from the MiSeq i100 Series support page. Save the installer to a network drive.
- Select the instrument icon to open the global navigation menu.
- Select Settings, and then select Applications.
- Select Install application.
- Navigate to the application file, and then select Open. After the file uploads, information about the application displays.
- Select Install. After the application installs, you can review the application configuration.

5.4 View Application Settings:

The DRAGEN application provides a default library prep kit, index adapter kit, read information, index information, and permissions. Some applications also provide settings and configuration for secondary analysis.

- Select the instrument icon to open the global navigation menu.
- Select Settings, and then select Applications.
- Select the application to view. After you install an application, the Configuration screen opens automatically.
- Edit any of the following information:
 - Library prep kits
 - Index adapter kits
 - Index reads
 - Read type
 - Index lengths
 - Read length
- Select Save.

5.5 Uninstall Applications

Administrators can uninstall applications.

- Select the instrument icon to open the global navigation menu.
- Select Settings, and then select Applications.
- Select the application to uninstall.
- Select Uninstall.
- Confirm to uninstall the application.

Release History

Revision	Release Reference	Originator	Description of Change
00	1130656	Mark Bilstad	Initial release