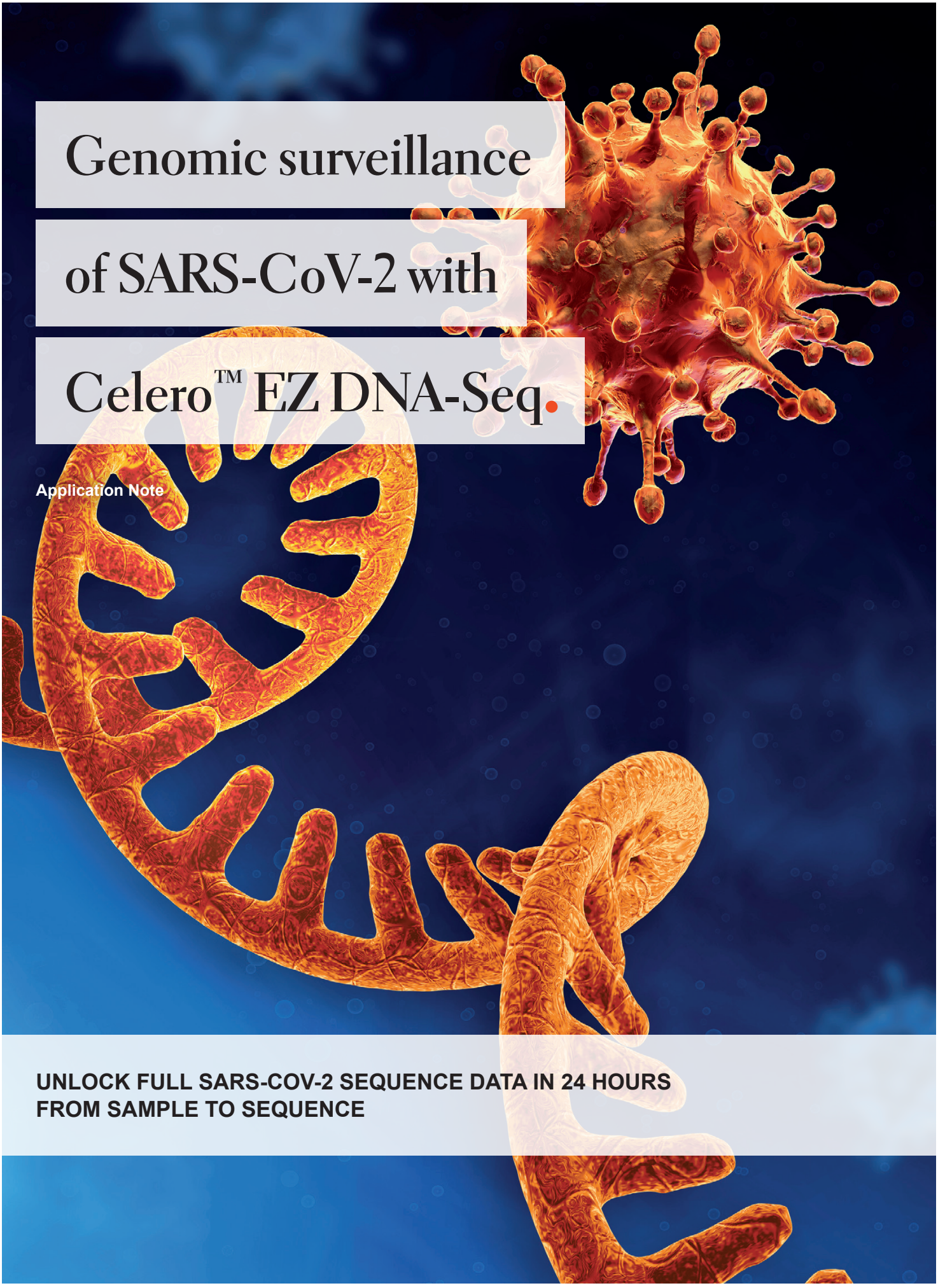
Genomic surveillance of SARS-CoV-2 with Celero™ EZ DNA-Seq.

Application Note



UNLOCK FULL SARS-COV-2 SEQUENCE DATA IN 24 HOURS FROM SAMPLE TO SEQUENCE

INTRODUCTION

The SARS-CoV-2 virus, a strain of the β-coronavirus, contains a single, positive RNA strand comprising approximately 30,000 bases coding a total of 10 genes.[1] The virus is able to enter human cells utilizing the receptor-binding domain that comprises part of the viral spike protein which binds the human angiotensin-converting enzyme 2.[2] Different variants of the SARS-CoV-2 virus have been shown to vary in their infection rates, mortality rates, and severity of COVID-19. Understanding and tracking viral variants has therefore become a key part of the global effort to understand and contain the virus.

In a typical workflow, patient samples are first assayed using qPCR or other qualitative methods to determine the presence or absence of the virus. Following confirmation of a positive result, the positive samples are reflexed to whole-genome sequencing (WGS) to identify the viral variant.

The ARTIC SARS-CoV-2 primer set[3] available from any oligo supplier has become the de facto standard for SARS-CoV-2 whole-genome sequencing. This application note describes the coupling of cDNA amplicon creation using the ARTIC primer set with the Celero-EZ library prep kit to create a single protocol that can go from extracted RNA to final sequence data in just 24 hours. The ARTIC Celero-EZ protocol is also fully automation-ready via DreamPrep on for use in higher throughput applications.

SARS-CoV-2 Variant Analysis with Celero EZ Library Prep

Celero EZ is a rapid, automation-ready solution for the preparation of sequence-ready libraries for Illumina sequencers. Incorporating Tecan's proprietary technologies such as enzymatic fragmentation, and NuQuant library quantification enables the Celero EZ to go from sample to sequencer in one workday, and from sample to sequence data in just 24 hours (Figure 1).

Celero EZ provides an innovative solution to streamline SARS-CoV-2 variant analysis. Utilizing standard ARTIC Primer sets to create cDNA amplicons from positive SARS-CoV-2 samples, and coupling this to the Celero EZ rapid library prep kit for Illumina sequencers allows both the ARTIC cDNA prep and the Celero EZ library prep to be completed on the same day. Allowing loading of libraries onto the sequencer at the end of day one for an overnight sequencing run that provides full sequence data the next morning.

This 24-hour sample to sequence data turnaround is made possible by the incredibly streamlined, user-friendly Celero EZ workflow. Unlike other sample-to-sequence solutions, this setup eliminates post-ligation bead purification and integrates Tecan's proprietary NuQuant library quantification as a substitute for traditional time-consuming or inaccurate methods. The following application note describes this simple 3-step addition only workflow, that results in faster library preparation and reduced hands-on time.

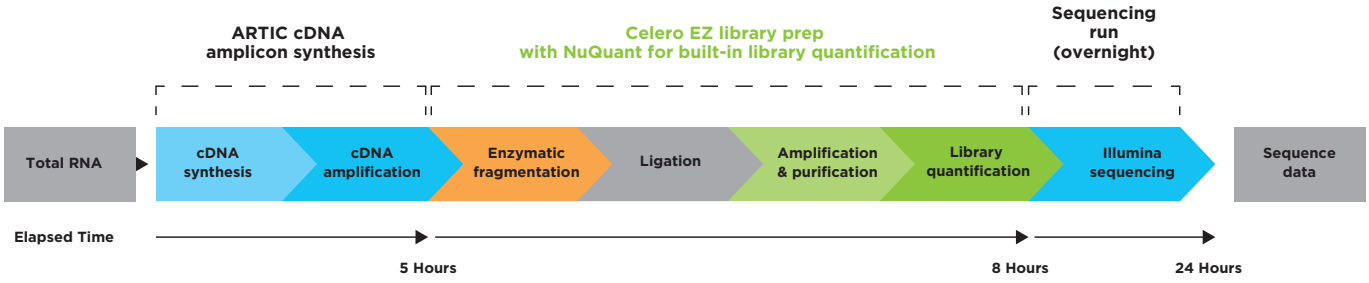


Figure 1: The SARS-CoV-2, ARTIC-Celero-EZ protocol workflow showing 24-hour time elapsed from extracted RNA sample to final sequence data.

MATERIALS AND METHODS

Materials

- Starting material: 10-500 ng total RNA extracted from SARS-CoV-2 positive nasal swab samples
- ARTIC primers. Available from IDT and other oligo suppliers. Sequence available on Github [3]
- Celero EZ DNA-Seq
  Tecan p/n 0569A-A01

Experimental Design and Protocol

This protocol is designed for use with confirmed positive SARS-CoV-2 viral samples from patients with COVID-19. It is assumed that the samples have first been screened for the presence or absence of the virus using qualitative methods such as qPCR. Positive samples can then be reflexed to whole-genome sequencing (WGS) using this set-up to obtain a greater amount of information in respect to the SARS-CoV-2 and critically to enable variant identification and tracking.

Creation of cDNA library using ARTIC primers

Amplification of SARS-CoV-2 Genomic RNA can now be accomplished in a standardized fashion thanks to the ARTIC V3 nCov-2019 primers.[3] This primer set enables direct amplification of the virus using tiled, overlapping, multiplexed primers to create ~400 nucleotide, double-stranded cDNA amplicons covering the entire 30kbp genome of SARS-CoV-2. Amplification using this primer set has been proven to have high sensitivity and to work directly from clinical samples.[4] The widespread adoption of the ARTIC primers also enables direct comparison of data from different institutions and different geographic regions. It is for these reasons that Tecan recommends the use of the ARTIC V3 nCov-2019 primers for the initial amplification step and creation of the amplicon pool. A brief overview of the amplification of SARS-CoV-2 RNA using the ARTIC V3 nCov-2019 primers is shown in figure 2.

Library Preparation with Celero EZ

Celero EZ DNA-Seq provides a rapid, user-friendly, automation-ready solution to streamline WGS library preparation and library quantification for SARS-CoV-2. Featuring a fast, easy-to-use, addition-only workflow that eliminates post-ligation bead purification, resulting in faster library preparation and reduced hands-on time. DNA libraries created with Celero EZ provide higher yields and better uniformity of coverage, ideal for clinical samples with variable quantities of starting material.

ARTIC WORKFLOW

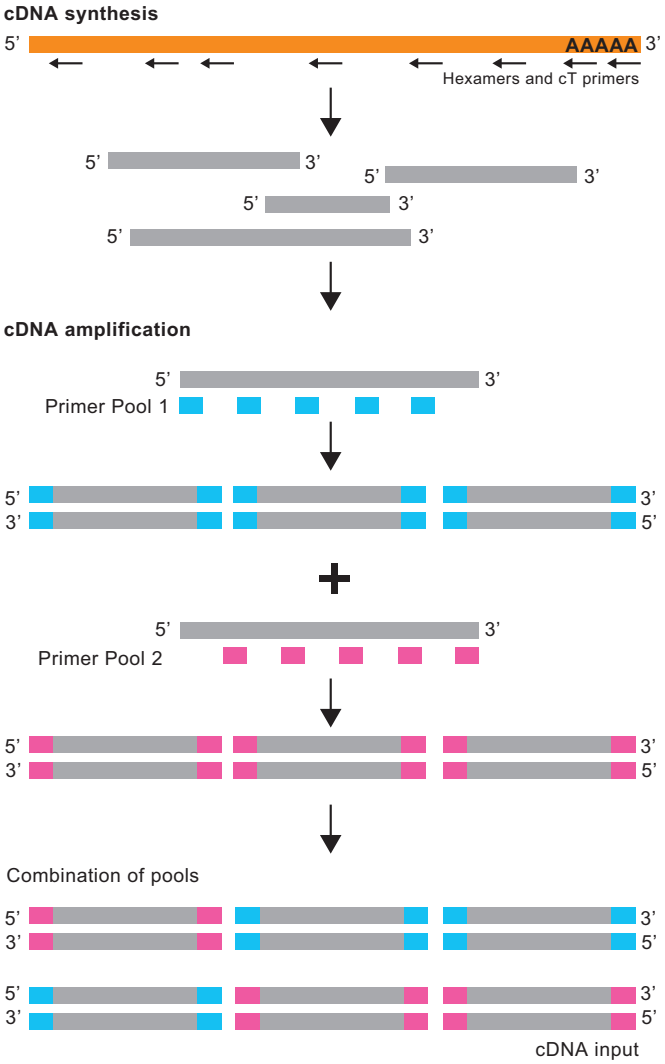


Figure 2: synthesis of overlapping cDNA amplicons from SARS-CoV-2 positive samples using the ARTIC V3 nCov-2019 primers.

The Celero EZ workflow is automation-ready with Tecan's DreamPrep™ NGS automation workflow, providing a robust solution for high-throughput viral monitoring and variant analysis applications.

Celero EZ is also integrated with NuQuant®, a proprietary library quantification method for efficient and accurate quantification of Next-Generation Sequencing (NGS) libraries. NuQuant method eliminates the need for time-consuming or inaccurate library quantification methods like qPCR, fluorometry (e.g. Qubit®), or microfluidic electrophoresis (e.g. Bioanalyzer®) allowing DNA-Seq libraries to be constructed, quantified, and sequencer ready in less than 3 hours.

Benefits of Celero EZ for SARS-CoV-19 Sequencing

- Broad input range from 10-500 ng
- Optimization-free enzymatic fragmentation
- No adaptor or template dilution and no post-ligation purification step leading to reduced hands-on time and faster workflow
- Flexible multiplexing with 384 dedicated dual index or Unique Dual Index (UDI) adaptors
- Integrated NuQuant library quantification to simplify library pooling and accelerate the workflow
- Readily available automation scripts with DreamPrep NGS

RESULTS

The SARS-CoV-2, ARTIC-Celero-EZ protocol showed significant advantages over competitor kits, particularly in relation to hands-on work time and total library prep time. Due to its simple addition only workflow, the complete lack of any adapter or template dilution step, no requirement for post-ligation purification, and rapid library quantification method, the Celero-EZ was the only kit able to go from starting RNA to high-quality library pools in a single workday.

Fast and Accurate Library Quantification with Built-in NuQuant

One of the key features that enables such a rapid turnaround time from the Celero-EZ kit is the built-in NuQuant library quantification which enables fast, accurate quantification of the library prior to loading without the requirement for time-consuming or inaccurate quantification methods required by other library prep kits. Data below from PMI R&D,shows the accuracy of NuQuant compared to other methods of quantification.

Table 1 shows the raw data comparing three different methods of library quantification (NuQuant, Bioanalyzer, and QuBit) to the gold standard iSEQ100 corrected data. The raw readings using each of the methods are shown in the top third of the table. The central section shows the discrepancy between the results from each of the methods used and the iSEQ standard. Positive or negative values indicate whether the error was larger or smaller than the standard value. The last section of the table shows the same discrepancy (error) expressed as a percentage. i.e., if this information had been used to decide how much to load onto the sequencer's flow cell, what would have been the percentage by which the user would have overloaded or under-loaded the sequencer. As demonstrated in the data, NuQuant is the most accurate method of quantifying the library and thus can be relied upon when making loading decisions.

	lib 01	lib 02	lib 03	lib 04	lib 05	lib 06	lib 07	lib 08	lib 09
iSeq100 corrected (Standard)	**67.6**	**160.7**	**67**	**71.8**	**59.5**	**77.7**	**145**	**104.9**	**70.6**
NuQuant	69.4	153.7	66.9	65.6	57.9	72.8	146.4	109.2	71.9
Bioanalyzer	116.2	151.7	131	81.4	69	133	118	165.9	110.5
QuBit-NQ/2	60	171	60	70	55	68.5	156.5	94	67
Discrepancy to standard. The difference between the result for each method of quantification and the iSeq100 result									
NuQuant	-1.8	7	0.1	6.2	1.6	4.9	-1.4	-4.3	-1.3
Bioanalyzer	-48.6	9	-64	-9.6	-9.5	-55.3	27	-61	-39.9
QuBit-NQ/2	7.6	-10.3	7	1.8	4.5	9.2	-11.5	10.9	3.6
% discrepancy to standard. The difference between the result for each method of quantification and the iSeq100 result expressed as a % of the iSeq value									
NuQuant	-2.7	4.4	0.1	8.6	2.7	6.3	-1.0	-4.1	-1.8
Bioanalyzer	-71.9	5.6	-95.5	-13.4	-16.0	-71.2	18.6	-58.2	-56.5
QuBit-NQ/2	11.2	-6.4	10.4	2.5	7.6	11.8	-7.9	10.4	5.1

Table 1. Comparison of NuQuant with two other common library quantification techniques.

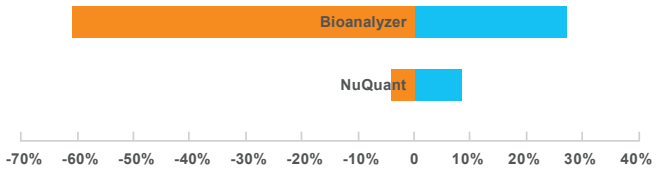


Figure 3: Range of differences in library concentration results over nine library pool samples

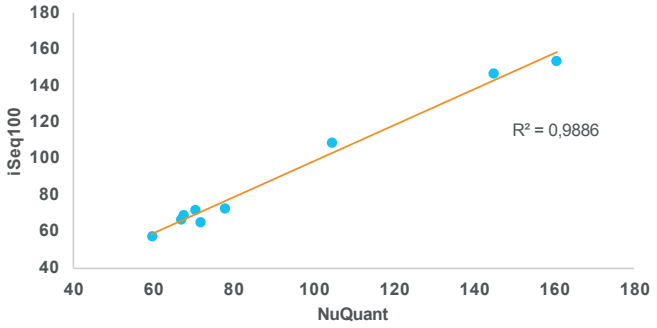


Figure 4: Correlation between NuQuant and iSeq100, the gold standard for library quantification, over the nine library pool samples.

Figure 3 shows the error range across all nine samples from Table 1 calculated as the percentage discrepancy between the quantification method being used NuQuant, or Bioanalyzer and the iSeq100 standard. As shown, the percentage error range for the Bioanalyzer is significantly greater than that obtained from NuQuant.

Figure 4 shows that the measurements for the library pools were highly correlated between the gold standard, iSeq100, and NuQuant (R^2=0.9886).

The NuQuant data while only requiring minutes to obtain using a standard lab fluorimeter is highly accurate and can be relied upon to give the correct loading concentration for each pool onto the sequencer flow cells. The variability seen in the data generated from the QuBit or Bioanalyzer could result in under-loading or overloading of the sequencer with a knock-on effect on the quality of the sequence data.

Quality and reliability of results

For variant analysis, it is critical to have sufficiently high coverage of the entire viral genome to have confidence that any variants have been detected. Depth of coverage (the number of times each region was sequenced) is another important driver of confidence in the sequence data, the more times each base was called the greater the confidence the user can have in the final sequence result.

Library	Raw reads	Insert size (sequenced)	Mean depth of coverage	Genome coverage (%)
Lib 01	652842	179	996.4	99.6
Lib 02	641490	182	937.7	99.6
Lib 03	525352	154	697.7	98.2
Lib 04	659900	202	914.2	99.5
Lib 05	617216	209	916.5	99.6
Lib 06	433208	154	574.9	96.5
Lib 07	707676	171	1009.0	99.5
Lib 08	666150	163	944.4	99.3
Lib 09	742852	182	889.3	97.4

Figure 5: Summary sequence data from 9 library pools created using the SARS-CoV-2, ARTIC-Celero-EZ protocol

Figure 5 shows the sequencing data for the same nine library pools. As shown, the genome coverage is consistently greater than 96.5% with a depth of coverage ranging from 574.9 to 1009, with most of the samples exhibiting genome coverage of 99% or higher. Data demonstrates that even with the speed of the Celero-EZ protocol, this method is still able to produce high-quality results with excellent genome coverage and coverage depth, giving a very high level of confidence in the ability of the method to detect any sequence variants.

SUMMARY

The SARS-CoV-2, ARTIC-Celero-EZ protocol is a fast, accurate method for the whole genome sequence of SARS-CoV-2 positive samples. The results demonstrate that even with the simplicity and speed of the Celero EZ protocol it is still capable of producing sequence data of sufficiently high quality for use in variant analysis and other demanding applications.

- Uses standard ARTIC SARS-CoV-2 primer sets
- Rapid protocol: sample to sequencer in one workday, sample to sequence data in 24 hours
- Robust, user-friendly workflow. Simple 3-step, addition only protocol
- Integrated NuQuant library quantification. Accurately quantify libraries in just minutes
- Fully automation-ready with DreamPrep for higher throughput applications.

The use of proprietary technology in the Celero EZ kit allows a sufficiently truncated library prep workflow for a researcher to complete both the ARTIC cDNA amplification step and the Celero EZ library prep step on the same day, a step up in terms of speed and efficiency from the competitor kits currently available.

ACKNOWLEDGEMENTS

Data provided by a Genomics & Transcriptomics department at an European R&D firm.

REFERENCES

[1] Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1, complete genome. NCBI Reference Sequence: NC_045512.1. Bethesda: National Center for Biotechnology Information; 2020. Available from: https://www.ncbi.nlm.nih.gov/nuccore/NC_045512.1 [cited 2020 May 29].

[2] Hoffmann M., *et al. Cell* 2020. 181(2):271-280.

[3] V3 ARTIC primer sequences on Github: https://github.com/esha-joshi/artic-ncov2019-nextflow/blob/master/primer_schemes/nCoV-2019/V3/nCoV-2019.tsv

[4] https://artic.network/ncov-2019

For Research Use Only. Not for use in diagnostic procedures.

Australia +61 3 9647 4100 **Austria** +43 62 46 89 330 **Belgium** +32 15 42 13 19 **China** +86 21 220 63 206 **France** +33 4 72 76 04 80 **Germany** +49 79 51 94 170 **Italy** +39 02 92 44 790 **Japan** +81 44 556 73 11 **Netherlands** +31 18 34 48 17 4 **Nordic** +46 8 750 39 40 **Singapore** +65 644 41 886 **Spain** +34 93 595 25 31 **Switzerland** +41 44 922 89 22 **UK** +44 118 9300 300 **USA** +1 919 361 5200 **Other countries** +41 44 922 81 11

401943 V1.0, 2021-05

www.tecan.com

TECAN.