Application Note

Efficient and Highly Multiplexed Whole Viral Genome Sequencing of SARS-CoV-2 Samples

Overview

- Coronavirus sequencing improves our understanding and ability to respond to the global COVID-19 pandemic:
  - Tracking pattern mutations
  - Identifying modes of spread
  - Genomic epidemiology
- Multiplexing is required for cost-effective SARS-CoV-2 sequencing
- Combining the widely-used ARTIC v3 RT-PCR protocol with plexWell multiplexed library prep allows for hundreds to thousands of SARS-CoV-2 isolates to be sequenced in a single run

SARS-CoV-2 Sequencing with ARTIC v3 RT-PCR and plexWell Library Preparation Kits

Viral whole-genome sequencing data can assist COVID-19 epidemiological research, particularly when combined with other types of data (e.g., geographical case counts). One obstacle to scaling up sequencing for COVID-19 applications is preparing NGS libraries from 100s or 1,000s of SARS-CoV-2 samples per day to match the throughput required for cost-effective sequencing. The plexWell workflow was designed to specifically address this challenge through scalable and highly multiplexed library preparation, enabling sequencing of up to 2,304 samples in a single sequencing run.

Here, we describe the application of plexWell library preparation downstream of a widely-used RT-PCR protocol (ARTIC). Combining the ARTIC multiplexed PCR method with plexWell provides an efficient and cost-effective way to prepare libraries to sequence large numbers of SARS-CoV-2 samples in the same sequencing run.

Workflow

Our multiplex sample preparation method was developed to easily integrate with existing RT-PCR and next-generation sequencing protocols (**Fig. 1**). Our method provides the missing step in increasing the sequencing throughput required for rapid and cost-efficient testing of SARS-CoV-2 samples. The sample preparation method consists of a 2-part workflow: 1) generation of ARTIC amplicons and 2) plexWell library preparation.

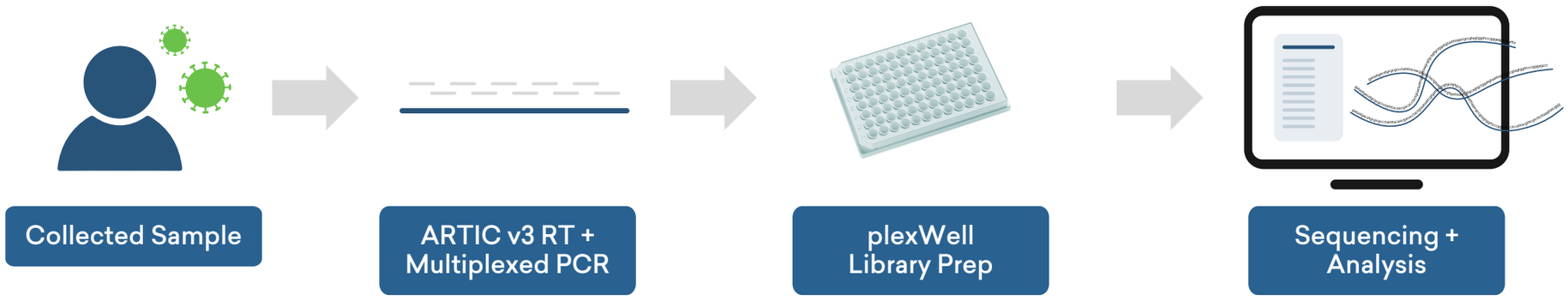

Figure 1. seqWell SARS-CoV-2 Sequencing Workflow Overview

First, the samples containing purified total nucleic acid isolated from human saliva are converted to cDNA using the published ARTIC v3 protocol (http://artic.network/ncov-2019). Each cDNA sample is amplified in two multiplexed PCR reactions using the ARTIC v3 protocol, generating two PCR pools that are then recombined into a single pool following amplification (**Fig. 2**).

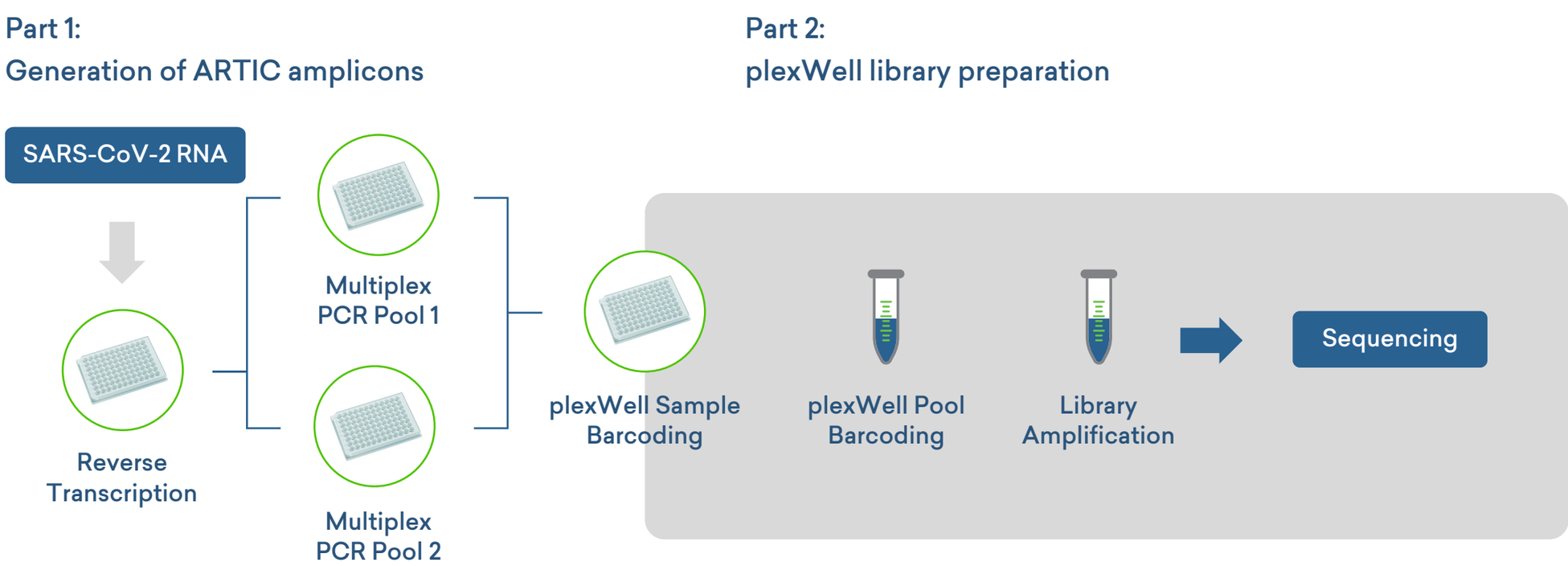

Figure 2. ARTIC v3 and plexWell Combined Method

The samples, each containing the amplicons from two multiplex PCRs, are entered into the plexWell workflow. In the first step, each sample on the plate receives one of 96 sample barcodes, which enables sample tracking within a plate of samples. Then all 96 samples on a plate are pooled into a single tube, purified, and barcoded with a pool barcode prior to amplificaiton to generate Illumina-compatible libraries. A detailed step-by-step process is described in **Fig. 4**.

Takeaways:

- plexWell allows multiplexed libraries to be prepared as a **single pool** after a simple molecular tagging step
- The method **normalizes across a wide range** of input RNA amounts, simplifying the workflow
- 24 pool-specific i5 barcodes increase the number of samples that can be pooled in a sequencing run: 96 i7 x 24 i5 = up to **2,304 unique barcode combinations** per run

Establishing Input Viral RNA Working Range for Typical Samples

An important consideration is determining the optimal input range of the viral sample. We provide the following guidelines to help you achieve the optimal results:

- rRT-PCR positive samples show mean Ct=24.3 [range: 16.9 - 38.4] -- **average viral load of ~1.4 x 10^6 SARS-CoV-2 copies per mL** in patient nasal swabs [Wang et al., 2020].
- Peak viral load of 104 - 107 SARS-CoV-2 copies per mL converts to between **~330 copies and ~3.3 x 106 copies per RT reaction** for the standard ARTIC protocol.
- Purification of total nucleic acid concentrates viral RNA **3- to 4-fold**, translating into **~4.2 x 10^4 input copies per ARTIC RT reaction** for a ”typical” nasal swab sample.

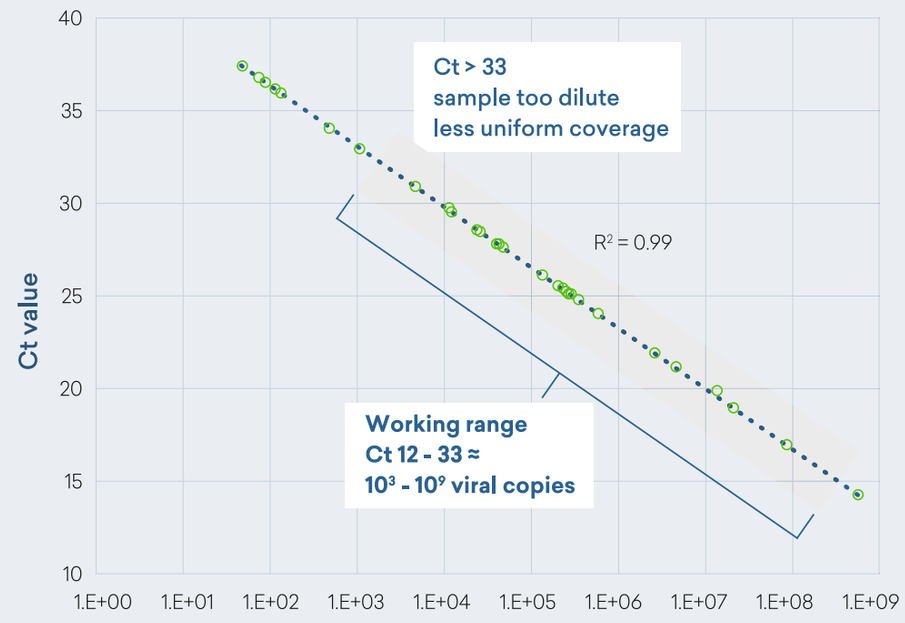

Figure 3. Establishing Input Viral RNA Working Range

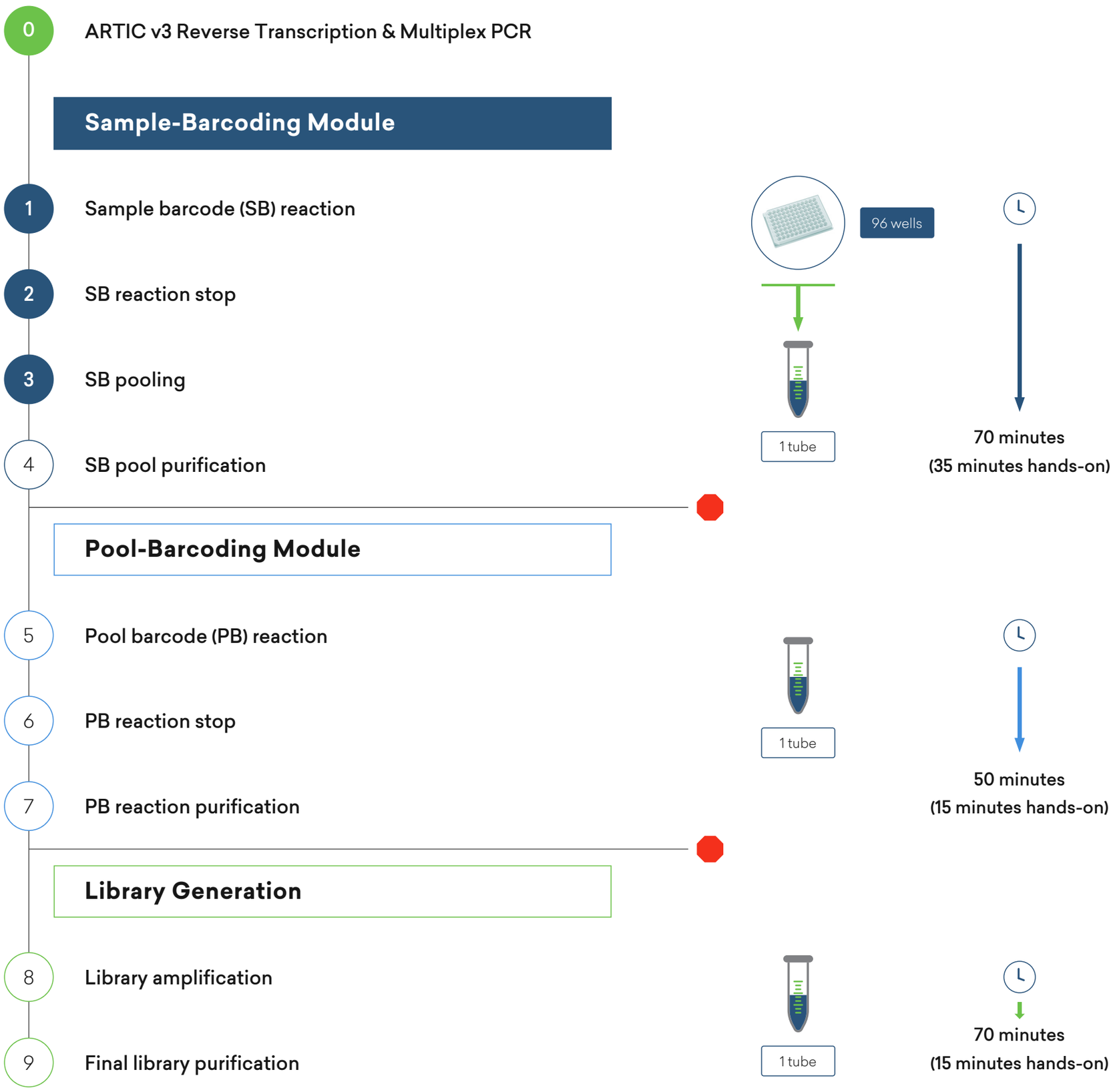

Figure 4. Integrated ARTIC + plexWell 384 Kit Workflow
See detailed protocol at https://seqwell.com/sars-cov-2/

Results

We have used our method to assess the quality of data that can be generated with plexWell. For this purpose we analyzed reference SARS-CoV-2 RNA in the context of purified total nucleic acid isolated from human saliva. The combined sample was converted to cDNA and amplified in two multiplexed PCR reactions. The undiluted PCR products were then converted to an Illumina-compatible library with the plexWell-96 library preparation kit and sequenced.

The average coverage by viral input level was characterized. The analysis showed a high correlation between the total and on-target reads as a function of the amount of input viral RNA used for the ARTIC RT reaction (**Fig. 5**).

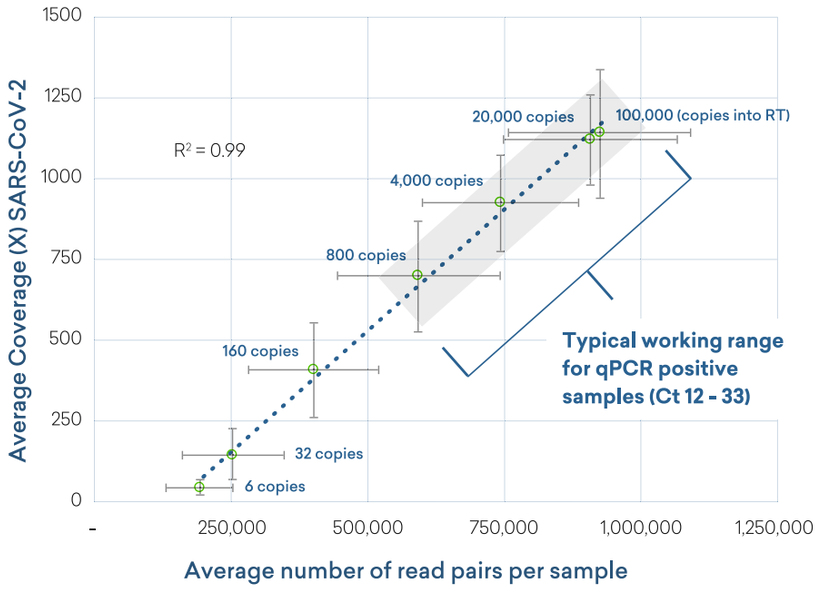

Figure 5. Average Coverage by Viral Input Level in RT Reaction

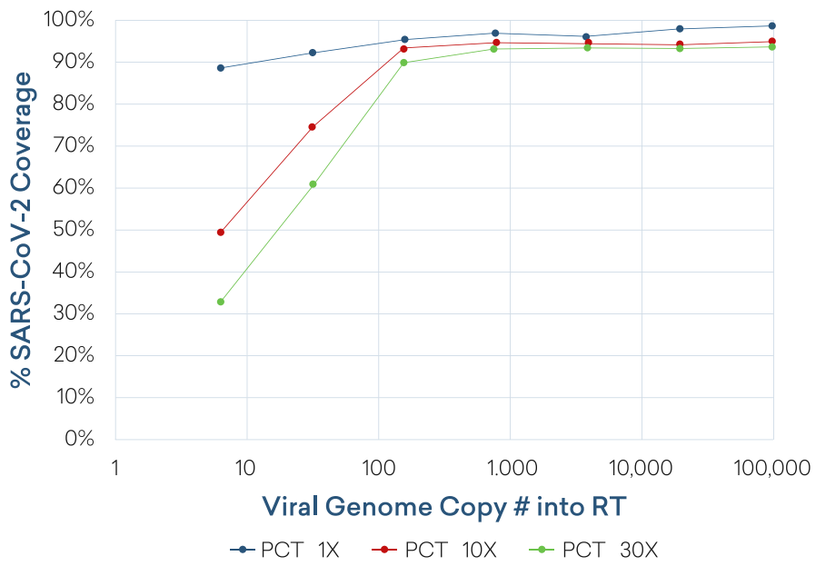

Figure 6. % of SARS-CoV-2 Genome Covered vs Viral Input Level in RT Reaction

Additionally, we looked at the percent genome coverage depending on the viral input level. The results showed that the percentage of the genome covered at >1x, >10x, or >30x is consistent for viral RNA input of >1000 copies (**Fig. 6**). The assay limit of detection is thus set by the lower limit of RNA that produces consistently high coverage of the full genome.

The alignment and coverage rates were assessed over a wide range of input. We saw that alignment rates of >80% are maintained down to <1000 copies of viral genome (**Fig. 7**). Note that the coverage gaps visible at 5kb intervals correspond to amplicon breakpoints in the Twist Bioscience synthetic SARS-CoV-2 standard.

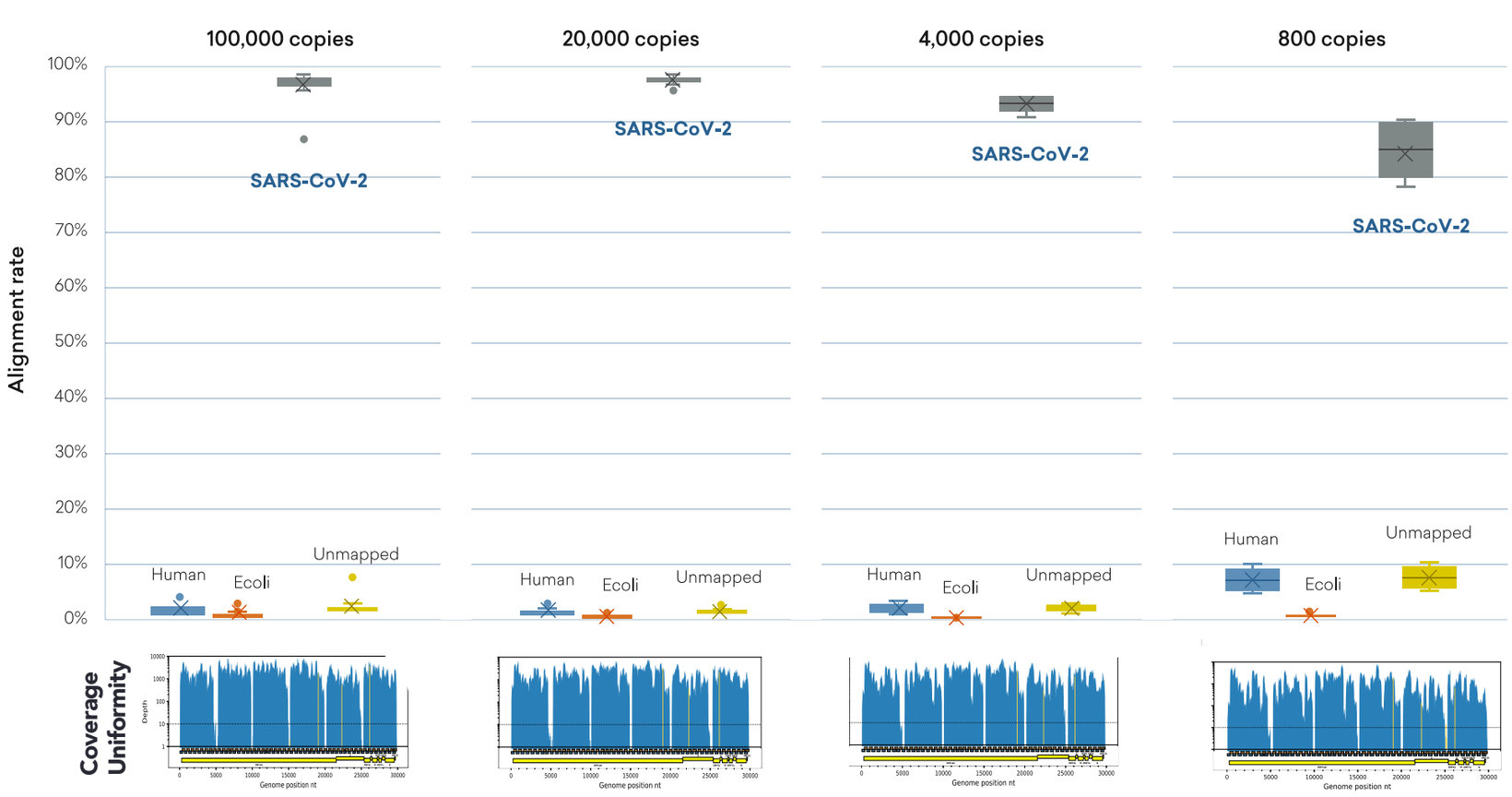

Figure 7. Alignment and Coverage Uniformity Across >100x Range of Viral Input

We looked at the variant calling accuracy at the single-nucleotide resolution for each one of three known variants that differ between the Twist Bioscience and ATCC (Washington State) control strain. The sensitivity and specificity of variant calling at these three variants was shown to be 100% down to 160 total copies of viral input RNA, demonstrating a high variant calling accuracy associated with the method (**Fig. 8**).

Finally, we performed assessment of library demultiplexing as an important step of the sequencing data analysis. An approximate range of 1000-fold amounts of viral RNA (plus negatives) was used to prepare four replicate libraries of 96 samples from the ATCC and Twist Bioscience standards. We saw high demultiplexing uniformity (reads per sample) while coverage maps showed high coverage for the two selected samples (**Fig. 9**). Together, these results demonstrate the high quality of data obtained with multiplex plexWell library sample preparation method.

Figure 8. Variant Calling Accuracy at Different Levels of Viral RNA Input

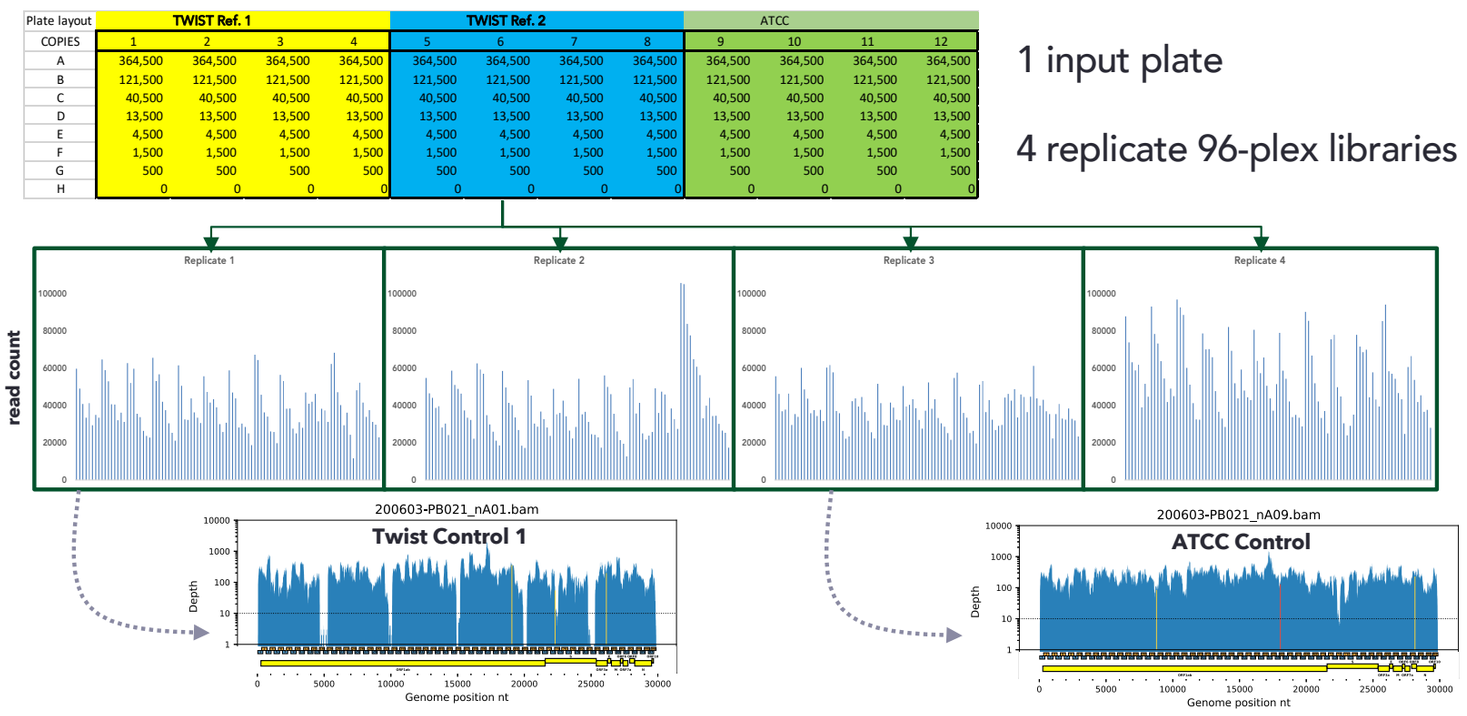

Figure 9. Example Demultiplexing Summary from a 384-plex Library

Summary

Our new plexWell sample preparation method in combination with the widely-used ARTIC RT-PCR protocol greatly simplifies and increases the throughput of multiplex whole genome sequencing of SARS-CoV-2 RNA. It provides high quality results while reducing the time and cost of sequencing efforts. The method provides numerous advantages, including:

- robust performance
- simple workflow
- high multiplexing capacity

This cost-effective solution can provide up to 40% savings compared to other methods and reduce processing time from sample to result. The plexWell workflow is automation-friendly and compatible with a variety of platforms. plexWell enables sequencing of up to 2,304 samples in a single sequencing run and provides a simple workflow capable of meeting the high demand for COVID-19 testing.

Learn more at https://seqwell.com/sars-cov-2/