# Rapid, automated NGS solution to survey the complete SARS-CoV-2 genome for epidemiological investigation

## Workflow and performance of the Ion AmpliSeq SARS-CoV-2 Research Panel on the Genexus Integrated Sequencer

### Introduction

First reported in Wuhan, China, in November 2019, SARS-CoV-2, a novel coronavirus, is now a global crisis. Accurately and efficiently tracking the geographic distribution and evolution of SARS-CoV-2 is essential in order to help slow down the spread of the virus. Due to genetic variation in SARS-CoV-2 strains occurring globally, viral surveillance and epidemiological research allows researchers to better understand the virus and its spread. Next-generation sequencing (NGS) is a powerful tool that allows for rapid identification of thousands of genetic mutations across samples in parallel. Using NGS, researchers can look at the entire SARS-CoV-2 genome in many samples simultaneously to get a more complete picture.

Unfortunately, most NGS workflows require skilled technicians and several days to complete, delaying the return of results. In contrast, the Ion AmpliSeq™ SARS-CoV-2 Research Panel on the Ion Torrent™ Genexus™ Integrated Sequencer provides a rapid and highly automated NGS workflow for analyzing the SARS-CoV-2 genome that enables labs to go from nucleic acid to report in a single day with 5 minutes of hands-on time. This hands-off, set-up-and-go workflow increases reproducibility of results and improves lab efficiency to enable easy adoption of NGS for epidemiological studies. Here we present the use of the Ion AmpliSeq SARS-CoV-2 Research Panel on the Genexus Integrated Sequencer for sequencing of the SARS-CoV-2 genome from positive samples with a range of viral loads.

### Workflow features

- The Genexus Integrated Sequencer enables complete workflow automation from nucleic acid to report for SARS-CoV-2 sequencing in a single day (Figure 1).
- The Ion AmpliSeq SARS-CoV-2 Research Panel provides 99.9% coverage of the SARS-CoV-2 viral genome.
- Ion Torrent™ Genexus™ Software supports a low-titer workflow, allowing for high genomic coverage from as few as 20 copies of viral RNA.
- Genexus Software automates post-sequencing run analysis for variant annotation (SnpEff), genome-assisted (IRMA), or *de novo* (Trinity) sequence assembly.

**Genexus Integrated Sequencer**


1 touchpoint; 5 min of hands-on time
**Total turnaround time: 1 day**

**Company I's NGS systems**


>10 touchpoints; ~7 hr of hands-on time
**Total turnaround time: 3–4 days**

**Figure 1. Comparison of the targeted NGS workflows for SARS-CoV-2 on the Genexus Integrated Sequencer and Company I's NGS system.** The Genexus Integrated Sequencer enables labs to go from nucleic acid to report in a single day with minimal user intervention.

ThermoFisher SCIENTIFIC

## Ion AmpliSeq SARS-CoV-2 Research Panel

The Ion AmpliSeq SARS-CoV-2 Research Panel consists of 2 primer pools targeting 237 amplicons tiled across the SARS-CoV-2 genome, with an additional 5 primer pairs targeting human expression controls. These 237 viral amplicons provide >99% coverage of the SARS-CoV-2 genome (~30 kb), while the additional primer pairs for human expression serve as internal controls for library generation. The SARS-CoV-2 amplicons range from 125 to 275 bp in length. Ion AmpliSeq™ technology is based on multiplex PCR and allows for flexible input amounts of as little as 1 ng of total RNA.

## Materials and methods

To demonstrate the simplified workflow and performance of the Ion AmpliSeq SARS-CoV-2 Research Panel on the Genexus Integrated Sequencer, a synthetic SARS-CoV-2 RNA control (Twist Bioscience, Cat. No. 102019) was used. This control uses GenBank™ database ID MT007544.1 (Wuhan strain) as a reference and includes 3 SNVs and one 10 bp deletion relative to GenBank ID MN908947.3 (Australian strain). The synthetic RNA control was synthesized as 6 nonoverlapping ~5 kb fragments that cover 99.9% of the SARS-CoV-2 genome. The synthetic RNA control was received at a concentration of 1 x $10^6$ copies/µL. Serial dilutions were made and then spiked into 5 ng of Invitrogen™ Human Lung Total RNA (Cat. No. AM7968) to obtain final viral copy numbers ranging from 20 to 200,000 copies in a final volume of 25 µL per sample.

Samples were distributed to two Applied Biosystems™ MicroAmp™ EnduraPlate™ Optical 96-Well Clear Reaction Plates (Cat. No. 4483354) as shown in Table 1, and each plate was sealed with a sheet of adhesive PCR plate foil (Cat. No. AB0626).

**Table 1. Configuration of sample plates.** Each plate has 16 samples, including technical replicates, with viral copy numbers ranging from either 0 to 200 or 0 to 200,000.

| Plate | Technical replicates | Copies of synthetic RNA control |
|---|---|---|
| 1 | 2 | 0 |
| | 4 | 20 |
| | 4 | 80 |
| | 2 | 120 |
| | 4 | 200 |
| 2 | 2 | 0 |
| | 3 | 20 |
| | 3 | 200 |
| | 3 | 2,000 |
| | 3 | 20,000 |
| | 2 | 200,000 |

Sample plates were placed on separate Genexus Integrated Sequencers, along with Ion AmpliSeq SARS-CoV-2 Research Panel primers (**ampliseq.com**), the Ion Torrent™ GX5™ Chip and Genexus™ Coupler (Cat. No. A40269), and consumables from the Ion Torrent™ Genexus™ GX5™ Starter Pack-AS (Cat. No. A40279) to support 16 samples (32 library reactions) across 2 lanes of sequencing per instrument. After entering sample information into the Genexus Software, run plans were created and executed using one of two preinstalled assay definition files: Plate 1 (0 to 200 copies) used the "SARS-CoV-2 Low Titer Research Assay", while Plate 2 (0 to 200,000 copies) used the "SARS-CoV-2 Research Assay".

Total turnaround time from RNA to variant report for each run is shown in Table 2, while Figure 2 displays all the steps automated by the Genexus Integrated Sequencer.

**Table 2. Total turnaround time for SARS-CoV-2 assay on the Genexus Integrated Sequencer.** Total turnaround time includes automated cDNA synthesis, library preparation, template preparation, sequencing, and post-run analysis for 16 samples run across 2 lanes on the GX5 Chip. Post-run analysis time scales based on total reads returned. Faster turnaround times can be achieved with smaller sample batch sizes.

| Assay definition used | Total turnaround time |
|---|---|
| SARS-CoV-2 Low Titer Research Assay | 24 hr 23 min |
| SARS-CoV-2 Research Assay | 28 hr 12 min |

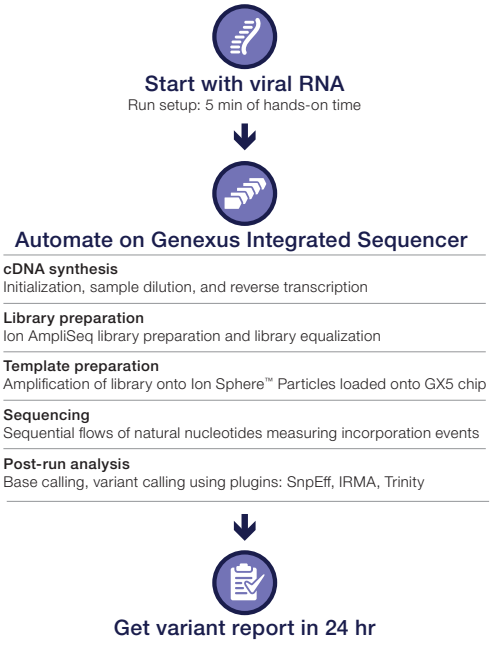


**Figure 2. Automated workflow on the Genexus Integrated Sequencer.** The Genexus Integrated Sequencer automates all steps from cDNA synthesis through post-run analysis. Approximate turnaround times shown are for 16 SARS-CoV-2 research samples run in 2 lanes on the GX5 chip. Post-run analysis time scales based on total reads returned. The next run can be started on the Genexus Integrated Sequencer while analysis from the previous run completes.

## Results

All samples across both runs returned reads mapping to the 5 human expression controls of the Ion AmpliSeq SARS-CoV-2 Research Panel, indicating automated library generation on the Genexus Integrated Sequencer was successful. Reads not mapped to the 5 human expression controls were mapped against the SARS-CoV-2 reference to determine percent base reads on target (Figure 3). For samples with 0 copies of synthetic control RNA, an average of 3.2% of reads mapped to the SARS-CoV-2 reference. The majority of these mapped reads were the same 35 bp fragment found across all 0 copy input samples. For the remaining samples on Plate 1 using the SARS-CoV-2 Low Titer Research Assay, the average percent base reads on target increases with input copy number, from 60.4% at 20 copies to 95.7% at 200 copies. On Plate 2 using the SARS-CoV-2 Research Assay, samples at ≥200 input copies increase from an average percent base reads on target of 95.3% to 99.9%.

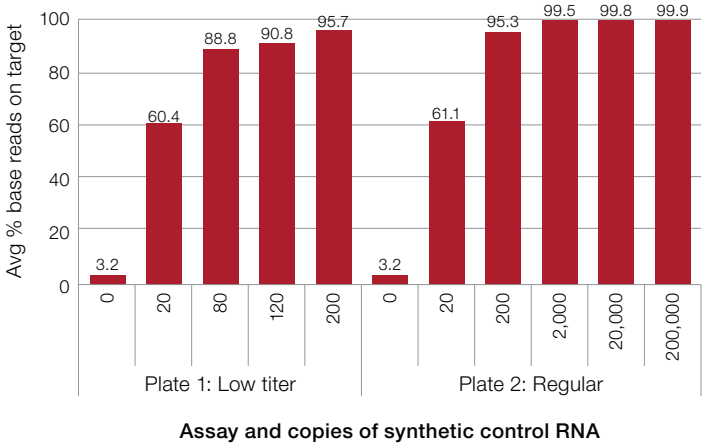


**Figure 3. Average percent base reads on target across both runs.** "Percent base reads on target" refers to reads mapped to the SARS-CoV-2 reference after removing human expression controls. Averages were taken from all replicates for each synthetic RNA control input and assay used.

Average percent genomic coverage results for Plate 1 using the SARS-CoV-2 Low Titer Research Assay definition file are shown in Figure 4, while results for Plate 2 using the SARS-CoV-2 Research Assay definition file are shown in Figure 5.

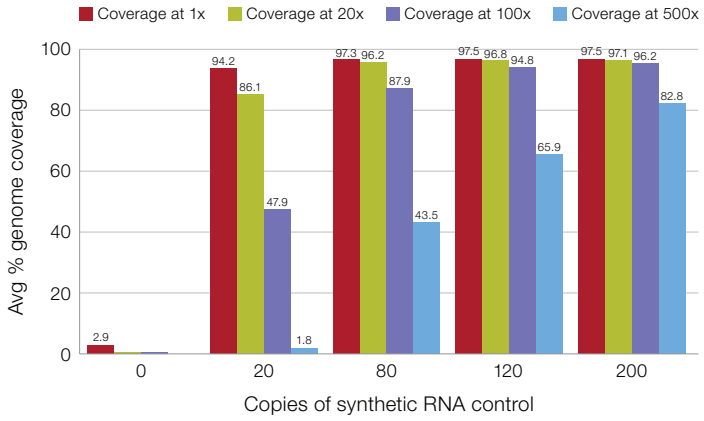


**Figure 4. Average percent genomic coverage at varying depths per copies of input RNA for the SARS-CoV-2 Low Titer Research Assay.** At inputs as low as 20 copies per sample, the Genexus Integrated Sequencer generated 94.2% genomic coverage of the SARS-CoV-2 genome.

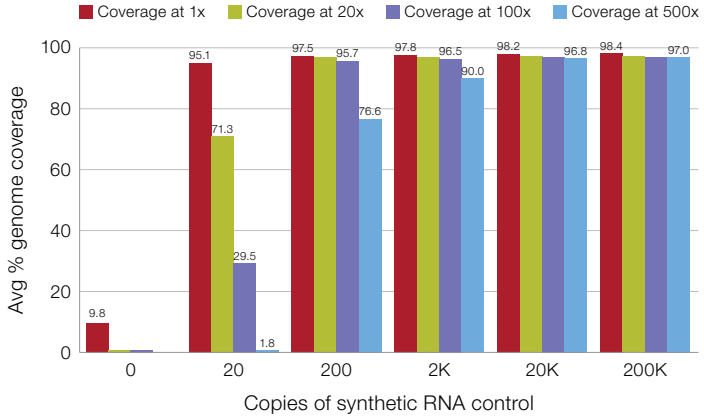


**Figure 5. Average percent genomic coverage at varying depths per copies of input RNA for the SARS-CoV-2 Research Assay.** At inputs as low as 200 copies per sample, the Genexus Integrated Sequencer generated 95.7% genomic coverage at 100x for the SARS-CoV-2 genome.

Variant calling results provided by the SnpEff plugin successfully identified all three SNVs and a 10 bp deletion for all samples with 120 copies of synthetic control RNA. Example output is shown in Figure 6.

| | CHROM | POS | REF | ALT | VARTYPE |
|---|---|---|---|---|---|
| 0 | 2019-nCoV | 19065 | T | C | SNP |
| 1 | 2019-nCoV | 22303 | T | G | SNP |
| 2 | 2019-nCoV | 26144 | G | T | SNP |
| 3 | 2019-nCoV | 29749 | ACGATCGAGTG | A | DEL |

**Figure 6. Annotated SnpEff plugin output.** The results show successful detection of 3 SNVs and 1 deletion in synthetic RNA controls for two variants of SARS-CoV-2: GenBank IDs MT007544.1 and MN908947.3.

ion torrent

As previously described, the Ion AmpliSeq SARS-CoV-2 Research Panel targets 237 amplicons tiled across the SARS-CoV-2 genome, while the synthetic RNA control used for this experiment consists of 6 nonoverlapping 5 kb fragments. Ten of the targets have priming sites on separate fragments of the synthetic control RNA; therefore, for these 10 targets no amplicons were generated (as observed in the coverage overview plot in Figure 7). The bases covered by the 10 missing targets with no overlap by adjacent amplicons total 907 bp, or approximately 3% of the SARS-CoV-2 genome. Starting with full-length or overlapping viral RNA will prevent the amplicon dropouts described above.

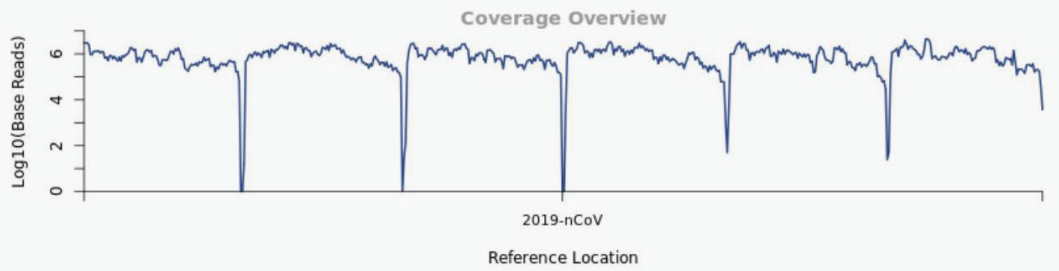


**Figure 7. Genomic coverage using synthetic control RNA with the Ion AmpliSeq SARS-CoV-2 Research Panel.** Coverage plot from 200,000-copy input sample demonstrates high genomic coverage across the entire 29.8 kb SARS-CoV-2 genome, with 5 spikes corresponding to breakpoints between nonoverlapping fragments in the synthetic control RNA.

## Conclusion

Rapid and efficient prediction of patterns of evolution and emergence of SARS-CoV-2 is essential in slowing the spread of the virus. Researchers need an easy-to-adopt solution to quickly and accurately sequence the SARS-CoV-2 genome to understand how the virus is evolving, assist with contact-tracing efforts, and inform vaccine research development. The Genexus Integrated Sequencer, combined with the Ion AmpliSeq SARS-CoV-2 Research Panel, provides a highly automated nucleic acid–to-report NGS workflow in a single day, enabling labs to survey the complete SARS-CoV-2 genome at a speed never possible before. With unmatched ease of use and less operational hands-on time compared to other technologies, this new solution makes the power of NGS accessible to labs that want to easily and quickly adopt the technology for epidemiological studies.

Find out more at **thermofisher.com/coronavirus-genexus**

ThermoFisher SCIENTIFIC