APPLICATION NOTE

Analytical performance of the Respiratory Pathogen ID/AMR Panel

- Robust detection of respiratory pathogens and antimicrobial resistance alleles at low titer across various read lengths and depths
- Highly sensitive, multipathogen detection achieved in samples with high bacterial or viral background
- Strong quantitative correlation across a wide range of pathogen titer between NGS and qPCR using IDbyDNA Explify Platform

Analysis powered by

IDbyDNA

illumina®

Introduction

Respiratory tract infections due to viral, bacterial, and fungal pathogens are a public health concern. While some illnesses are less severe (eg, the common cold), others can be more severe and require hospitalization, or even be fatal (eg, pneumonia).[1,2] Accurate identification of respiratory pathogens can be challenging, particularly in the case of mixed or coinfections. Conventional detection methods can require multiple, sequential assays. Compounding the COVID-19 pandemic with a flu season could increase the need for rapid, accurate, and broad-spectrum pathogen detection solutions.

Next-generation sequencing (NGS) provides an effective way to detect known and emerging respiratory pathogens from various sample types, including those harboring multiple infectious agents, in a single assay. To support the use of NGS for pathogen detection, Illumina offers the Respiratory Pathogen ID/AMR Panel. This panel targets > 280 respiratory pathogens, including viruses, bacteria, and fungi, and > 1200 associated antimicrobial resistance (AMR) alleles. This application note demonstrates the exceptional analytical performance of the Respiratory Pathogen ID/AMR Panel.

Methods

Sample preparation

Extracted genetic material from viral, bacterial, and fungal pathogens was procured from various commercial sources (Table 1). To demonstrate limit of detection for this assay, genome copy number titrations were generated as follows in triplicate: 100K, 10K, 1K, 100, 10, 1, and 0 copies for different respiratory pathogens separately and in combination. Multipathogen samples were also generated in triplicate to test the ability to detect a viral pathogen in a bacterial background and a bacterial pathogen in a viral background. The background pathogen was present at 100K genome copies and the test pathogen was titrated to 10K, 1K, 100, 10, or 1 copy. All titration samples were spiked into a human background that included 10 ng of NA12877 Human Coriell DNA and 10 ng of Universal Human Reference RNA (Invitrogen, Catalog no. 14919). The pathogen samples in human background were reverse transcribed into cDNA using reagents included with the Illumina RNA Prep with Enrichment Kit.

Table 1: Respiratory pathogens detected in this study

Sample	Source
Aspergillus fumigatus	Microbiologics
Influenza A virus	Virapur
Klebsiella oxytoca	ATCC
SARS-CoV-2	Virapur
Streptococcus pneumoniae	Microbiologics
T7 bacteriophage	Virapur

Library preparation

Sequencing-ready libraries were prepared with Illumina RNA Prep with Enrichment (Illumina, Catalog no. 20040536) and IDT for Illumina DNA/RNA UD Indexes (Illumina, Catalog no. 20027213). Illumina RNA Prep with Enrichment uses On-Bead Tagmentation followed by a single hybridization step to generate enriched DNA and RNA libraries in the same tube. This novel approach combines cDNA and DNA samples into a single tube that can be carried straight through into tagmentation with no quantification required.* Libraries were enriched as 1-plex reactions using the Illumina Respiratory Pathogen ID/AMR Panel (Illumina, Catalog no. 20047050).

Sequencing

Prepared libraries were denatured and diluted to a final loading concentration of 10 pM, according to the MiSeq™ System Denature and Dilute Libraries Guide (Document no. 15039740 v10), and sequenced on the MiSeq System at 2 × 75 bp read length using MiSeq v3 reagents, then trimmed to 1 × 75 bp. Libraries were also denatured and diluted to a final loading concentration of 1.8 pM, according to the NextSeq™ 550 System Denature and Dilute Libraries Guide (Document no. 15039740 v10), and sequenced on the NextSeq 550 System at 2 × 149 bp read length using NextSeq 550 High Output (HO) v2.5 reagents, then trimmed to 1 × 149 bp.

* For DNA samples, where cDNA synthesis is not required, the recommended total DNA input for tagmentation is 10-100 ng per sample.

Data analysis

FASTQ sequencing files were input to the IDbyDNA Respiratory Pathogen ID/AMR Panel Platform for analysis. The software can be accessed in BaseSpace™ Sequence Hub.

Results

Single respiratory pathogen detection

The limit of detection of the Respiratory Pathogen ID/AMR Panel was evaluated using a titration experiment, as described (see Methods). Twenty ng total nucleic acid† was input for reverse transcription, libraries were prepared and enriched from the resulting cDNA, and sequencing was performed on the MiSeq System at 1 × 75 bp. Analysis with the Explify Platform successfully identified all tested pathogens, even at low genome copy numbers (Figure 1, Table 2). Each of the four pathogens was detected down to 100 genome copies with except for A. fumigatus, which was detected at 1000 copies (Figure 1, Table 2). Pathogen detection was based on reads per kilobase million (RPKM), a metric that normalizes the number of detected reads to 1M reads, then to the length of the targeted region.

† Inputs higher than 20 ng of DNA/RNA per sample are likely to increase library diversity and coverage, if available (data not shown).

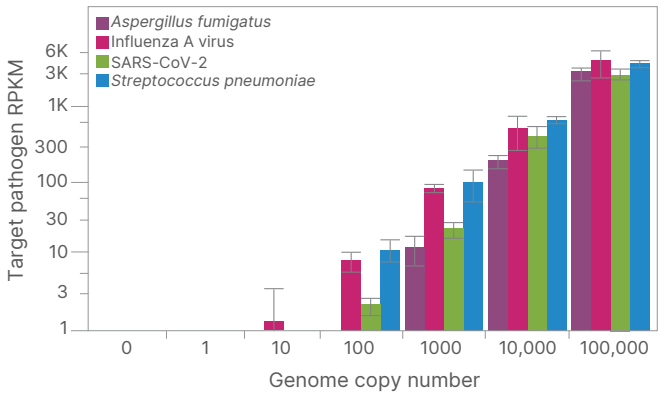

Figure 1: Limit of detection for individual pathogens—Each of the four pathogens was detected down to 100 genome copies, except for *A. fumigatus*. Influenza A virus was detected at 10 genome copies, as measured by RPKM. Libraries were sequenced on the MiSeq System at 1 × 75 bp.

Table 2: Limit of detection for individual pathogens

Genome copy number	RPKM			
	SARS-CoV-2	Influenza A virus	*S. pneumoniae*	*A. fumigatus*
100,000	3023	4225	4010	2659
10,000	425	530	673	175
1000	24.3	83.7	98.7	12.7
100	2.33	9	12.3	0
10	0	1.33	0	0
1	0	0	0	0
0	0	0	0	0

Additionally, SARS-CoV-2 and *S. pneumoniae* libraries were sequenced on the NextSeq 550 System at 2 × 149 bp, trimmed to 1 × 75 bp, downsampled to 2M, 1M, 0.5M, and 0.25M reads, and analyzed with the Explify Platform to demonstrate how read depth affects pathogen detection. SARS-CoV-2 RPKM was consistent for all read depths, demonstrating detection of SARS-CoV-2 down to 100 viral copies (Figure 2). *S. pneumoniae* RPKM was consistent for all read depths, demonstrating detection of *S. pneumoniae* down to 100 viral copies (Figure 3).‡

Multipathogen detection

To evaluate the analytical performance of the Respiratory Pathogen ID/AMR Panel in multipathogen detection, a titration experiment was conducted with all four respiratory pathogen samples combined in equal, diminishing amounts, as described (Methods). Two other multipathogen titration experiments were performed in which a bacterial pathogen was spiked into a viral background, and vice versa, as described (Methods). In the first experiment, *S. pneumoniae* was spiked into SARS-CoV-2, and in the second experiment SARS-CoV-2 was spiked into *S. pneumoniae*. For each of these experiments, 20 ng total nucleic acid was input for reverse transcription. Libraries were prepared from the resulting cDNA, enriched, and sequenced on the MiSeq System. Analysis was performed with the Explify Platform.

All four pathogens present in the multipathogen titration were detected reliably down to 100 genome copies (Table 3). This is the same level of detection seen in the single-pathogen samples, indicating that assay sensitivity is not lost when more than one pathogen is present. Similar results were seen for the viral titration in a bacterial background and bacterial titration in a viral background. In a background of 100K genome copies of SARS-CoV-2, *S. pneumoniae* was detected down to 100 genome copies (Table 4). Likewise, in a background of 100K genome copies of *S. pneumoniae*, SARS-CoV-2 was detected down to 100 genome copies (Table 5).

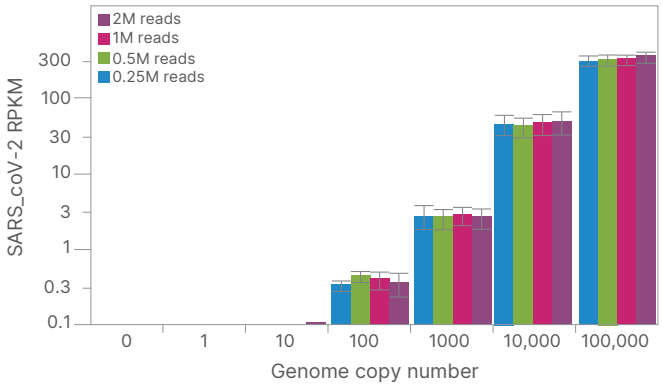

Figure 2: Limit of detection of SARS-CoV-2 at varying read depths—SARS-CoV-2 was detected down to 100 viral copies at all sequencing depths tested. Libraries were sequenced on the NextSeq 550 System at 2 × 149 bp, trimmed to 1 × 75 bp, then downsampled to 2M, 1M, 0.5M, and 0.25M reads.

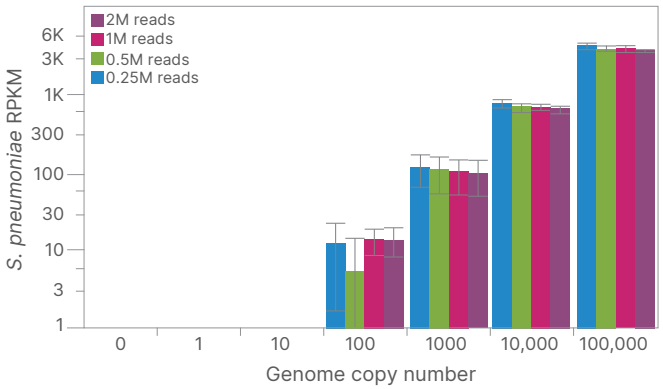

Figure 3: Limit of detection of *S. pneumoniae* at varying read depths—*S. pneumoniae* was detected down to 100 viral copies at all sequencing depths tested. Libraries were sequenced on the NextSeq 550 System at 2 × 149 bp, trimmed to 1 × 75 bp, then downsampled to 2M, 1M, 0.5M, and 0.25M reads.

‡ Illumina recommends a read depth of 0.5-1M reads per sample to ensure accurate pathogen detection. Read depth should be determined by each user.

Table 3: Limit of detection for multipathogen sample

Genome copy number	Confidence score			
	SARS-CoV-2	Influenza A virus	*S. pneumoniae*	*A.fumigatus*
100,000	3.00	3.00	1.54	3.00
10,000	3.00	3.00	1.54	3.00
1000	3.00	2.55	1.54	0.76
100	0.85	1.74	0.95	0.06
10	N/A	0.08	N/A	N/A
1	N/A	N/A	N/A	N/A
0	N/A	N/A	N/A	N/A

Table 4: Bacterial titration in viral background

Genome copy number	Confidence score	
	SARS-CoV-2	*S. pneumoniae*
100,000	3.00	N/A
10,000	N/A	2.70
1000	N/A	2.76
100	N/A	1.49
10	N/A	N/A
1	N/A	N/A

Table 5: Viral titration in bacterial background

Genome copy number	Confidence score	
	S. pneumoniae	SARS-CoV-2
100,000	2.70	N/A
10,000	N/A	3.00
1000	N/A	3.00
100	N/A	0.88
10	N/A	N/A
1	N/A	N/A

AMR detection

To evaluate the analytical performance of the Respiratory Pathogen ID/AMR Panel in AMR detection, a titration experiment was performed with a known multidrug–resistant *Klebsiella pneumoniae* (ATCC, Catalog no. BAA-1705) sample, from 100K copies down to 10 copies. Additionally, samples were run on a NextSeq 550 System at 2 × 149 bp, and then trimmed to 1 × 149 bp, 1 × 101 bp, 1 × 75 bp, and 1 × 50 bp at 1M total reads, to determine the effect of read length on AMR detection. Separately, samples were run on a NextSeq 550 System at 1 × 75 bp and downsampled to 2M, 1M, 0.5M, and 0.25M to determine the effect of read depth on AMR detection.

AMR detection remained fairly consistent down to 1000 genome copies regardless of read length (Table 6) or read depth (Table 7), although some detection was lost at 250K reads and 1 × 50 bp. The recommendation for this workflow is to sequence at 1 × 75 bp with 500K reads or greater per sample.

Table 6: AMR detection as a function of read length

Read length	Spike-in level	*Klebsiella oxytoca*	aadA	SHV-5	sul1	dfrA1	AAC(6')-Ib-cr[1]	AAC(3)-Ia
1 × 149 bp	100,000	3/3	3/3	3/3	3/3	3/3	3/3	3/3
	10,000	3/3	3/3	3/3	3/3	3/3	3/3	3/3
	1000	3/3	3/3	3/3	3/3	2/3	2/3	2/3
	100	2/3	2/3	0/3	0/3	0/3	0/3	0/3
1 × 100 bp	100,000	3/3	3/3	3/3	3/3	3/3	3/3	3/3
	10,000	3/3	3/3	3/3	3/3	3/3	3/3	3/3
	1000	3/3	3/3	3/3	3/3	3/3	1/3	2/3
	100	0/3	0/3	0/3	0/3	0/3	0/3	0/3
1 × 75 bp	100,000	3/3	3/3	3/3	3/3	3/3	3/3	3/3
	10,000	3/3	3/3	3/3	3/3	3/3	3/3	3/3
	1000	3/3	3/3	3/3	3/3	2/3	2/3	1/3
	100	0/3	0/3	0/3	0/3	0/3	0/3	0/3
1 × 50 bp	100,000	3/3	3/3	3/3	3/3	3/3	3/3	3/3
	10,000	3/3	3/3	3/3	3/3	3/3	3/3	3/3
	1000	3/3	1/3	2/3	1/3	1/3	0/3	0/3
	100	0/3	0/3	0/3	0/3	0/3	0/3	0/3

Pathogen Quantification with IDbyDNA

SARS-CoV-2 was used to evaluate the performance of the Respiratory Pathogen ID/AMR Panel in quantifying pathogens present in clinical samples. Triplicates of ten-fold serial dilutions (10-10^8 GE/ml) of inactivated SARS-CoV-2 were prepared in a background of 5 × 10^3 A549 cells/ml and spiked with 2.5 × 10^5 MS2 phage (Microbiologics) as an internal standard. Fifteen clinical residual samples (nasopharyngeal swabs in viral transport media) that previously were shown to be SARS-CoV-2 positive by Panther Fusion SARS-CoV-2 Assay (Hologic, Catalog no. AW-211569-004) (qPCR Ct-values varying from 11.8 –33.9), were also spiked with 2.5 × 10^5 GE/ml MS2 internal standard. Subsequently, DNA and RNA were extracted using the Zymobiomics DNA/RNA MiniKit (Catalog no. R2002). Three aliquots of RNA from each dilution of contrived samples were tested with the CDC Novel Coronavirus (2019-nCoV) Real-Time Reverse Transcriptase (RT) Diagnostic Panel.[3] Resulting Ct values for N1 were used to generate a standard curve (Figure 4).

In addition, one aliquot of RNA from each replicate of the serial dilutions and the clinical samples were reverse transcribed into cDNA, combined with the extracted DNA, and input for library preparation and sequencing, as described (see Methods). Data was analyzed using the Respiratory Pathogen ID/AMR Panel App on BaseSpace Seqence Hub. Viral loads were calculated using the following equation based on RPKM values resulting from analysis:

SARS-CoV-2 (GE/ml) = SARS-CoV-2 (RPKM) * (MS2 (GE/ml)/MS2 (RPKM))

Table 7: AMR detection as a function of read depth

Read depth	Spike-in level	*Klebsiella oxytoca*	aadA	SHV-5	sul1	dfrA1	AAC(6')-Ib-cr[1]	AAC(3)-Ia
2M reads	100,000	3/3	3/3	3/3	3/3	3/3	3/3	3/3
	10,000	3/3	3/3	3/3	3/3	3/3	3/3	3/3
	1000	3/3	3/3	3/3	3/3	2/3	2/3	2/3
	100	1/3	2/3	0/3	0/3	0/3	0/3	0/3
1M reads	100,000	3/3	3/3	3/3	3/3	3/3	3/3	3/3
	10,000	3/3	3/3	3/3	3/3	3/3	3/3	3/3
	1000	3/3	3/3	3/3	3/3	3/3	2/3	1/3
	100	0/3	0/3	0/3	0/3	0/3	0/3	0/3
0.5M reads	100,000	3/3	3/3	3/3	3/3	3/3	3/3	3/3
	10,000	3/3	3/3	3/3	3/3	3/3	3/3	3/3
	1000	3/3	3/3	2/3	2/3	1/3	0/3	0/3
	100	0/3	0/3	0/3	0/3	0/3	0/3	0/3
0.25M reads	100,000	3/3	3/3	3/3	3/3	3/3	3/3	3/3
	10,000	3/3	3/3	3/3	3/3	3/3	3/3	3/3
	1000	3/3	0/3	0/3	0/3	0/3	0/3	0/3
	100	0/3	0/3	0/3	0/3	0/3	0/3	0/3

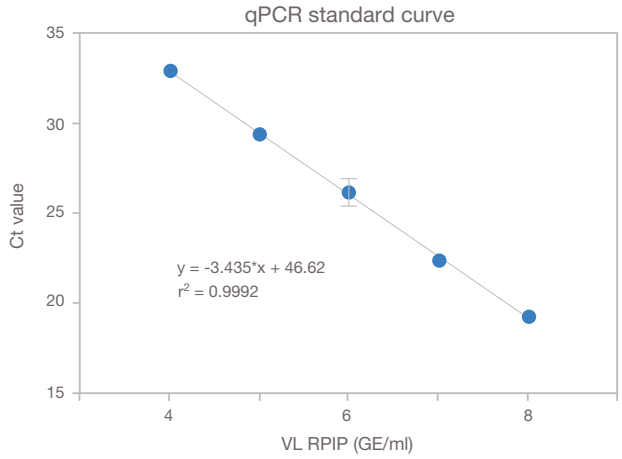

Figure 4: q PCR Standard curve—10-fold serial dilutions of SARS-CoV-2 were used for a qPCR standard curve. Only dilutions where all nine replicates were positive were used.

Comparison of viral loads derived by qPCR and NGS showed significant correlation (Spearman r = 0.87, $p < 0.0001$)) between the two methods (Figure 5). Importantly, these results highlight the advantage of using the Respiratory Pathogen ID/AMR Panel for viral quantification over PCR, in that NGS with the panel does not require generation of a standard curve.

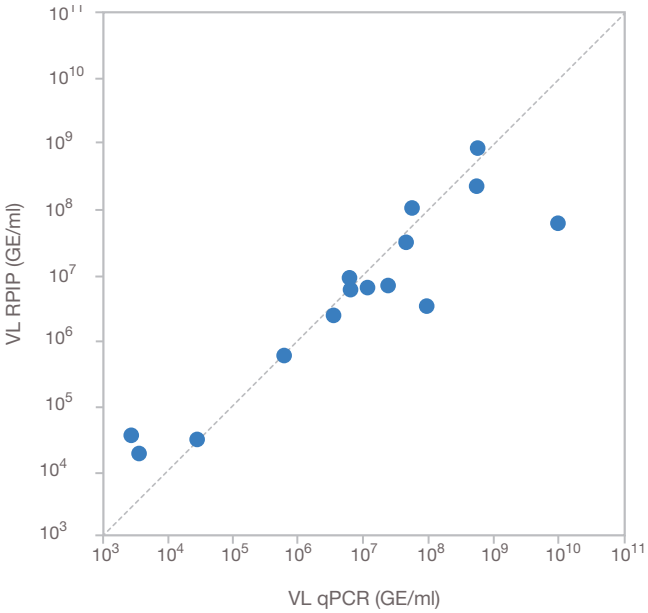

Figure 5: Comparison of SARS-CoV-2 viral loads in clinical samples determined by qPCR and NGS—The two methods shows significant correlation, Spearman r = 0.87.

Summary

The identification and characterization of respiratory pathogens is central to improving public health. NGS is a powerful method for simultaneous, broad-range detection of multiple infectious agents. The Respiratory Pathogen ID/AMR Panel targets > 280 respiratory pathogens, and > 1200 AMR markers in a single assay. This application note demonstrates the exceptional analytical performance of the Respiratory Pathogen ID/AMR Panel. It provides highly sensitive, multipathogen detection in mixed samples with high bacterial or viral background. The panel shows strong quantitative correlation across a wide range of pathogen titer between NGS and qPCR using IDbyDNA Explify Platform, with the added benefit of not requiring generation of a standard curve.

Learn more

To learn more about the Respiratory Pathogen ID/AMR Panel, visit www.illumina.com/products/by-type/sequencing-kits/library-prep-kits/respiratory-pathogen-id-panel.html

References

1. Dasaraju PV, Liu C. Infections of the Respiratory System. *Medical Microbiology*. 4th edition. 1996; Chapter 93.
2. Li Z, Lu G, Meng G. Pathogenic Fungal Infection in the Lung. *Front Immunol*. 2019;10:1524.
3. Centers for Disease Control and Prevention. CDC 2019-Novel Coronavirus (2019-nCoV) Real-Time RT-PCR Diagnostic Panel Instructions for Use. CDC-006-000019, Revision: 06.

illumina®

1.800.809.4566 toll-free (US) | +1.858.202.4566 tel
techsupport@illumina.com | www.illumina.com

M-GL-00007