illumina®

High-accuracy next-generation sequencing with the MiSeq™ i100 Series

Data comparison to the MiSeq System across key MiSeq i100 Series applications

Uniform coverage of microbial communities, regardless of GC content, using microbial genomics methods

Broad detection of microbial pathogens, including bacteria and viruses using Illumina infectious disease panels

Accurate variant calling across different cancer types using Illumina oncology panels

M-GL-02246 v1.0

Introduction

With the MiSeq i100 and MiSeq i100 Plus Sequencing Systems, Illumina continues to set the standard for simplicity and usability in next-generation sequencing (NGS). Breakthrough advancements in system design, XLEAP-SBS™ chemistry, and integrated DRAGEN™ secondary analysis deliver operational simplicity, high data accuracy, and exceptional speed. As part of an end-to-end NGS solution, the MiSeq i100 Series provides same-day results for various applications across microbiology, infectious disease, oncology, and more.

This application note demonstrates that the MiSeq i100 Series delivers data quality that meets or exceeds that of the MiSeq System for key methods, including microbial genomics and targeted gene sequencing studies.

Methods

Microbial genomics

Small whole-genome sequencing (WGS)

Small WGS libraries were prepared from 100 ng commercially available microbial gDNA, including *Escherichia coli* strain MG1655 (ATCC, Catalog no. 700926D-5), *Rhodobacter sphaeroides* strain ATH 2.4.1 (ATCC, Catalog no. 17023D-5), and *Bacillus pacificus* strain NRS 248 (ATCC, Catalog no. 10987D-5) using Illumina DNA Prep (Illumina, Catalog no. 20060060).

Sequencing was performed on the MiSeq i100 Plus System with the MiSeq i100 Series 25M Reagent Kit (300 cycles) (Illumina, Catalog no. 20126568) using the 2 × 151 bp run configuration, 24-plex with 1% PhiX Control spike-in. For comparison, the same libraries were sequenced on the MiSeq System with the MiSeq Reagent Kit v3 using a 2 × 151 bp run configuration with 1% PhiX Control spike-in.

Secondary analysis was performed using the DRAGEN Small Whole-Genome Sequencing app v4.3.13 for reference-based mapping of microbial genomes.

16S rRNA sequencing

Libraries were prepared from 5 ng of input gDNA from 20 Strain Even Mix Genomic Material (ATCC, Catalog no. MSA-1002) or 20 Strain Staggered Mix Genomic Material (ATCC, Catalog no. MSA-1003) using the Illumina 16S Metagenomic Sequencing Library Preparation amplicon-based protocol.

Sequencing was performed on the MiSeq i100 Plus System with the MiSeq i100 Series 25M Reagent Kit (600 cycles) (Illumina, Catalog no. 20126566) using the 2 × 301 bp run configuration, 96-plex. For comparison, the same libraries were sequenced on the MiSeq System with the MiSeq Reagent Kit v3 using a 2 × 301 bp run configuration.

Data analysis was performed using the 16S Metagenomics app v1.1.3 for taxonomic classification, relative abundance, and HTML visualizations.

Infectious disease panels

Libraries were prepared from commercially available controls and raw wastewater samples to profile various infectious microbes, including respiratory pathogens, urinary pathogens, and influenza viruses (Table 1).

Sequencing was performed on the MiSeq i100 Plus System with the MiSeq i100 Series 25M Reagent Kit (300 cycles) (Illumina, Catalog no. 20126568) using the 2 × 151 bp run configuration for the Viral Surveillance Panel v2 and Illumina Microbial Amplicon Prep and the 1 × 151 bp run configuration for the Respiratory Pathogen ID/AMR Enrichment Panel Kit and the Urinary Pathogen ID/AMR Enrichment Kit. For comparison, the same libraries were sequenced on the MiSeq System with the MiSeq Reagent Kit v3 using 2 × 151 bp and 1 × 151 bp run configurations.

Data analysis was performed using the DRAGEN Microbial Enrichment Plus app v1.1.0 or DRAGEN Microbial Amplicon app v4.3.6 for microorganism detection and viral consensus sequence generation.

Table 1: Sample and library preparation using infectious disease panels

Sample	Source	Input	Panel	Source
Raw wastewater samples	Collected from wastewater treatment plants by Wisconsin State Laboratory of Hygiene and from student dormitories by Colorado State University	8.5 µl (≥ 100 ng extracted total nucleic acids)	Viral Surveillance Panel v2	Illumina, Catalog no. 20087932
NATtrol Respiratory Panel 2.1 (RP2.1) Controls	Zeptometrix, Catalog no. NATRPC2.1-BIO	8.5 µl (extracted RNA)	Respiratory Pathogen ID/AMR Enrichment Panel Kit	Illumina, Catalog no. 20047050
Microbial Community Standard	ZymoBIOMICS, Catalog no. D6300	30 µl (extracted DNA)	Urinary Pathogen ID/AMR Enrichment Kit	Illumina, Catalog no. 20090309
Genomic RNA from Influenza A virus (H1N1) strain A/Virginia/ ATCC2/2009	ATCC, Catalog no. VR-1737D	Extracted RNA with Ct values < 30	Illumina Microbial Amplicon Prep	Illumina, Catalog no. 20097857
Genomic RNA from Influenza A virus strain A/Hong Kong/8/68	ATCC, Catalog no. VR-1679D			
Quantitative Genomic RNA from Influenza A virus (H1N1) strain A/ PR/8/34	ATCC, Catalog no. VR-95DQ			
Quantitative Genomic RNA from Influenza A virus (H3N2) strain A/Wisconsin/15/2009	ATCC, Catalog no. VR-1882DQ			
Genomic RNA from Influenza B virus strain B/Lee/40	ATCC, Catalog no. VR-1535D			
Genomic RNA from Influenza B virus strain B/Taiwan/2/62	ATCC, Catalog no. VR-1735D			
Genomic RNA from Influenza B virus (BY) B/ Massachusetts/2/2012	ATCC, Catalog no. VR-1813D			
Quantitative Genomic RNA from Influenza B virus strain B/ Florida/4/2006	ATCC, Catalog no. VR-1804DQ			

Oncology panels

Libraries were prepared from commercially available samples using the Pillar® oncoReveal™ Myeloid Panel (Illumina, Catalog no. HDA-MY-1001-24), the Pillar oncoReveal BRCA1 & BRCA2 + CNV Panel (Illumina, Catalog no. HDA-BR-1003-24), and the TruSight™ RNA Pan-Cancer Panel Set A (Illumina, Catalog no. RS-303-1002) (Table 2).

Sequencing was performed on the MiSeq i100 Plus System with MiSeq i100 Series 25M Reagent Kit (300 cycles) (Illumina, Catalog no. 20126568) with run configurations of 2 × 151 bp for Pillar oncoReveal panels and 2 × 76 bp for the TruSight RNA-Pan-Cancer Panel. For comparison, the same libraries were sequenced on the MiSeq System with the MiSeq Reagent Kit v3 using 2 × 151 bp and 2 × 76 bp run configurations.

Data analysis was performed using the DRAGEN Amplicon app for Pillar oncoReveal libraries. For analysis of TruSight RNA Pan-Cancer libraries, the BaseSpace™ RNA-Seq Alignment tool was used.

Table 2: Sample and library preparation using oncology panels

Sample	Source	DNA input	Panel	Source
Seraseq Myeloid Mutation DNA Mix	SeraCare, Catalog no.0710-0408	20 ng	Pillar oncoReveal Myeloid Panel	Illumina, Catalog no. HDA-MY-1001-24
Admixture of NA12877 and NA12878	Coriell Institute for Medical Research	20 ng		
Mimix BRCA Germline I, gDNA Reference Standard	Horizon Discovery, Catalog no. HD793	20 ng	Pillar oncoReveal BRCA 1 & BRCA 2 + CNV Panel	Illumina, Catalog no. HDA-BR-1003-24
Mimix BRCA Germline II, gDNA Reference Standard	Horizon Discovery, Catalog no. HD794	20 ng		
Mimix BRCA Somatic Multiplex I, gDNA Reference Standard	Horizon Discovery, Catalog no. HD795	20 ng		
NA12878	Coriell Institute for Medical Research, NIST ID HG001	20 ng		
NA24385	Coriell Institute for Medical Research, NIST ID HG002	20 ng		
NA24149	Coriell Institute for Medical Research, NIST ID HG003	20 ng		
NA24143	Coriell Institute for Medical Research, NIST ID HG004	20 ng		
NA24631	Coriell Institute for Medical Research, NIST ID HG005	20 ng		
Universal Human Reference RNA	Thermo Fisher Scientific, Catalog no. QS0639	50 ng	TruSight RNA Pan-Cancer Panel Set A	Illumina, Catalog no. RS-303-1002
Mimix Pan-Cancer 6-Fusion Panel, FFPE Reference Standard	Horizon Discovery, Catalog no. HD834	50 ng		
MCF7 human breast cancer cell line	ATCC, Catalog no. HTB-22	50 ng		
K-562 human leukemia cell line	ATCC, Catalog no. CCL-243	50 ng		

Results

Microbial genomics

Small WGS

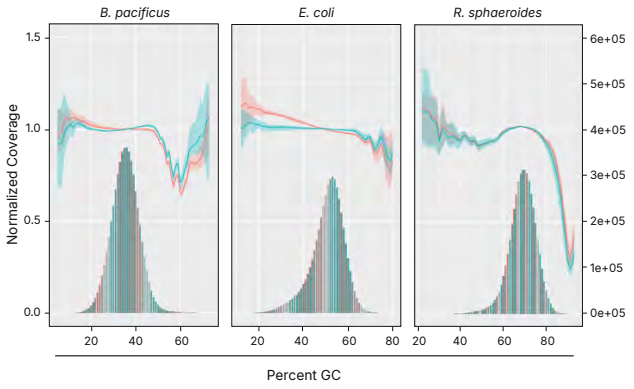
Sequencing run metrics were evaluated, including percent bases above Q30 and error rate. Both systems delivered high-quality data. The MiSeq i100 Plus delivered a significantly reduced error rate compared to the MiSeq System (Table 3).

Table 3: Sequencing run metrics for sWGS

Metric	MiSeq System	MiSeq i100 Plus System
Run configuration	2 × 151 bp	2 × 151 bp
Read 1 bases ≥ Q30	96.88%	96.30%
Read 2 bases ≥ Q30	94.07%	96.75%
Read 1 error rate	0.72%	0.13%
Read 2 error rate	0.78%	0.23%

To assess coverage performance across bacterial species with a range of low-, medium-, and high-GC content, normalized coverage data from the MiSeq i100 Plus and MiSeq Systems were plotted against reference genome content by GC percentage. Both systems show even coverage levels and high-quality data metrics across all microbial species tested, regardless of GC content (Figure 1 and Table 4). These data demonstrate that sWGS results on the MiSeq i100 Series are equivalent to MiSeq System performance.

Figure 1: Coverage across genomes of varying GC content



The MiSeq i100 Plus System (blue lines) provides consistent and comparable read coverage across microbial genomes of varying GC content, as compared to the MiSeq System (red lines). The bell curve traces at the bottom of each plot show the actual GC composition of each microbial species.

Table 4: WGS alignment metrics for genomes of varying GC content

	B. pacificus		*E. coli*		*R. sphaeroides*	
Sequencing system	MiSeq System	MiSeq i100 Plus System	MiSeq System	MiSeq i100 Plus System	MiSeq System	MiSeq i100 Plus System
Coverage uniformity	98.62%	98.58%	99.13%	99.04%	99.16%	99.07%
% mapped bases	96.54%	97.72%	95.41%	97.56%	95.21%	97.39%
% mapped reads	99.75%	99.88%	99.75%	99.93%	97.85%	97.94%
% mismatched reads	0.18%	0.05%	0.21%	0.04%	0.19%	0.05%

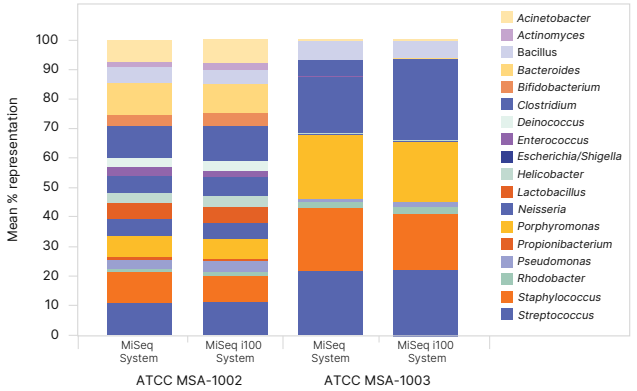
16S rRNA sequencing

Analysis of the 16S sequencing results identified all expected members of the bacterial community and showed comparable results between the MiSeq i100 Plus System and the MiSeq System (Figure 2). These results demonstrate the parity of performance between the MiSeq i100 Plus and MiSeq Systems for 16S metagenomics applications.

Infectious disease panels

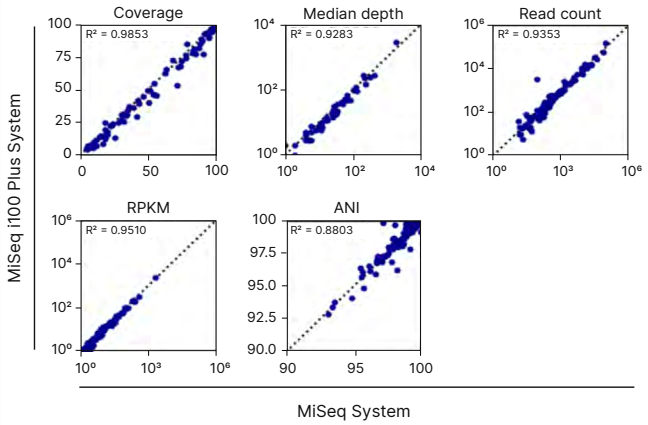
Analysis of results from sequencing wastewater samples with the Viral Surveillance Panel v2 showed similar performance between the MiSeq i100 Plus and MiSeq Systems in detection and coverage of viral genomes with highly concordant sequencing metrics (Figure 3 and Table 5).

Figure 2: Genus representation across systems



Analysis of microbial composition of ATCC samples with the MiSeq i100 Plus System demonstrates similar genus-level classification obtained by sequencing on the MiSeq System.

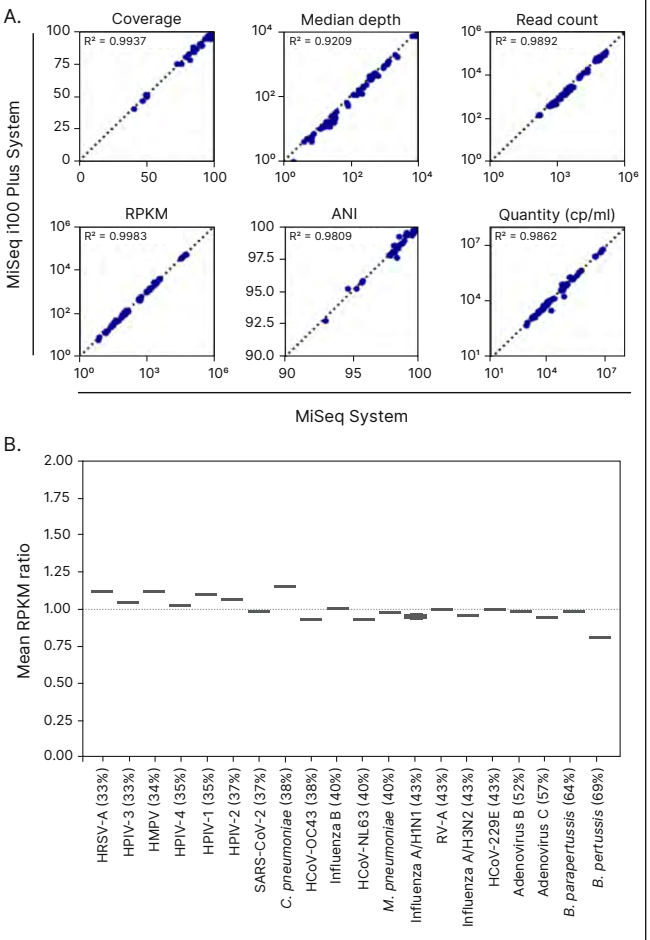
Figure 3: Viral Surveillance Panel v2 metrics



The MiSeq i100 Plus System shows highly concordant results with the MiSeq System for various sequencing metrics, including coverage, median depth, read count, targeted reads mapped per kilobase of targeted sequence per million quality-filtered reads (RPKM), and average nucleotide identity (ANI).

Analysis of results from sequencing RP2.1 Controls with the Respiratory Pathogen ID/AMR Panel showed similar performance between the MiSeq i100 Plus and MiSeq Systems in detection and quantification of microbes with highly concordant sequencing metrics (Figure 4A and Table 6). Of note, there was no observable difference in performance between the two systems, regardless of targeted GC content of the microbe analyzed (Figure 4B).

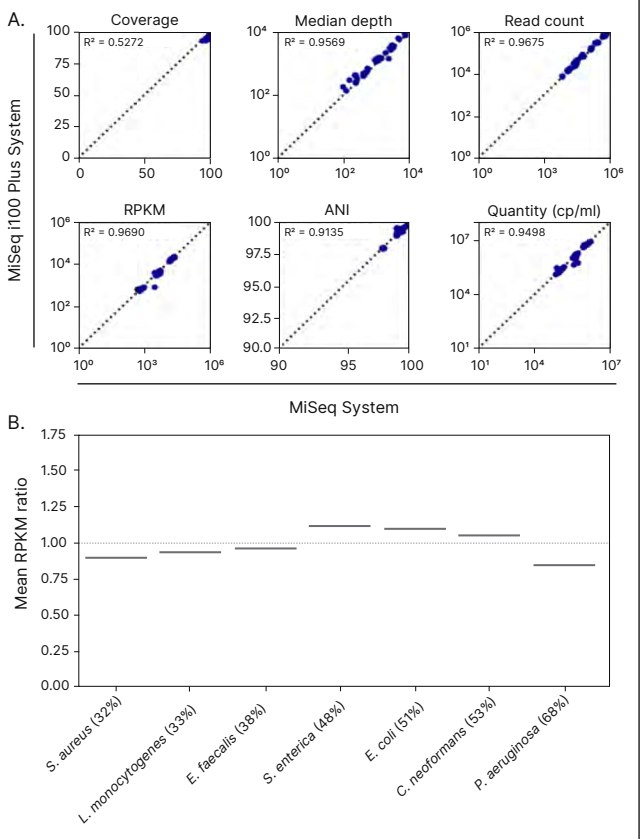
Figure 4: Respiratory Pathogen ID/AMR Panel metrics



(A) The MiSeq i100 Plus System shows highly concordant results with the MiSeq System for various sequencing metrics, including coverage, median depth, read count, RPKM, ANI, and quantity. (B) The mean RPKM ratio for microbes with varying levels of GC content (percentages in parentheses after each sample name) shows little to no difference between systems.

Analysis of results from sequencing the Microbial Community Standard with the Urinary Pathogen ID/AMR Panel showed similar performance between the MiSeq i100 Plus and MiSeq Systems in detection and quantification of microbes with highly concordant sequencing metrics (Figure 5A and Table 7). There was no observable difference in performance between the two systems, regardless of targeted GC content of the microbe analyzed (Figure 5B).

Figure 5: Urinary Pathogen ID/AMR Panel metrics



(A) The MiSeq i100 Plus System shows highly concordant results with the MiSeq System for various sequencing metrics. (B) The mean RPKM ratio for microbes with varying levels of GC content (percentages in parentheses after each sample name) shows little to no difference between systems.

Table 5: Viruses detected with at least 25% genomic coverage using the Viral Surveillance Panel v2

Microorganism	MiSeq System	MiSeq i100 Plus System	Microorganism	MiSeq System	MiSeq i100 Plus System
Aichi virus 1 (AIV-A1)	1	1	Human adenovirus F	4	4
Human adenovirus A	1	1	Human bocavirus (HBoV)	4	4
Human papillomavirus 69 (HPV69)	1	1	Mamastrovirus 1 (MAstV1)	4	4
Rhinovirus A (RV-A)	1	1	Mamastrovirus 8 (MAstV8)	4	4
Rhinovirus C (RV-C)	1	1	Mamastrovirus 9 (MAstV9)	5	5
Human polyomavirus 6 (HPyV6)	2	1	Sapovirus	5	5
Norovirus GII	1	2	Merkel cell polyomavirus (MCPyV)	5	6
Human papillomavirus 53 (HPV53)	2	2	BK polyomavirus (BKPyV)	6	6
Rotavirus A (RVA)	2	2	Human coronavirus OC43 (HCoV_ OC43)	6	6
Norovirus GI	3	2	JC polyomavirus (JCPyV)	6	6
Mamastrovirus 6 (MAstV6)	3	3	**Total detection**	**67**	**67**

Table 6: Microorganisms detected using the Respiratory Pathogen ID/AMR Panel

Microorganism	MiSeq System	MiSeq i100 Plus System	Microorganism	MiSeq System	MiSeq i100 Plus System
Chlamydia pneumoniae	3	3	Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)	3	3
Human adenovirus B	3	3	*Bordetella parapertussis*	3	3
Human adenovirus C	3	3	*Bordetella pertussis*	3	3
Human metapneumovirus (HMPV)	3	3	Human coronavirus 229E (HCoV_229E)	3	3
Human parainfluenza virus 1 (HPIV-1)	3	3	Human coronavirus NL63 (HCoV_NL63)	3	3
Human parainfluenza virus 4 (HPIV-4)	3	3	Human coronavirus OC43 (HCoV_OC43)	3	3
Influenza A virus (H1N1)	3	3	Human parainfluenza virus 2 (HPIV-2)	3	3
Influenza A virus (H3N2)	3	3	Human parainfluenza virus 3 (HPIV-3)	3	3
Mycoplasmoides pneumoniae	3	3	Human respiratory syncytial virus A (HRSV-A)	3	3
Rhinovirus A (RV-A)	3	3	Influenza B virus (B/Victoria/2/87-like)	3	3
			Total detection	**60**	**60**

Table 7: Microorganisms detected using the Urinary Pathogen ID/AMR Panel

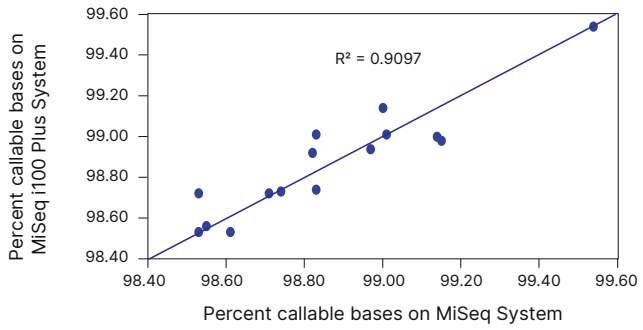
Microorganism	MiSeq System	MiSeq i100 Plus System
Cryptococcus neoformans	6	6
Enterococcus faecalis	6	6
Escherichia coli	6	6
Listeria monocytogenes	6	6
Pseudomonas aeruginosa	6	6
Salmonella enterica	6	6
Staphylococcus aureus	6	6
Total detections	**42**	**42**

Analysis of results from sequencing ATCC influenza virus samples with Illumina Microbial Amplicon Prep showed similar performance between the MiSeq i100 Series and the MiSeq System in detection of viral genomes (Table 8). The MiSeq i100 Plus System shows highly concordant results with the MiSeq System for percent callable bases, a metric that reports on sequencing accuracy across the genome, including challenging regions such as repetitive sequences, high GC content, and regions of low complexity (Figure 6).

Table 8: Viral genomes detected with Illumina Microbial Amplicon Prep

Microorganism	MiSeq System	MiSeq i100 Plus System
Influenza A (H1N1)	6	6
Influenza A (H3N2)	6	6
Influenza B	8	8
Total detections	**20**	**20**

Figure 6: Percent callability comparison across systems

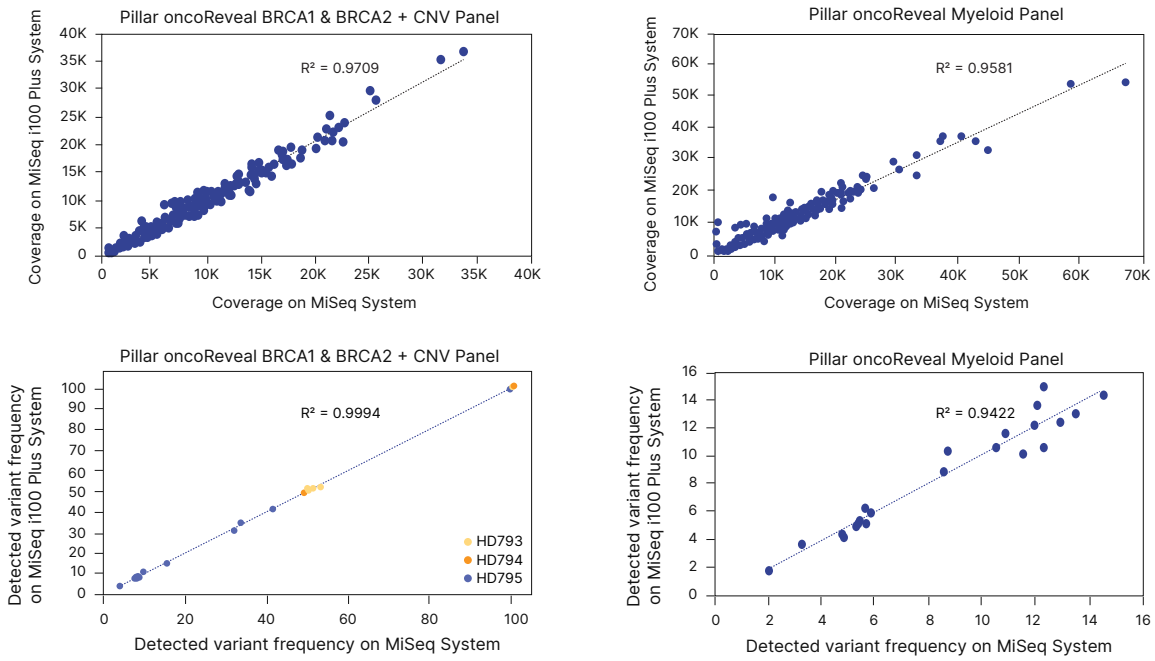


The MiSeq i100 Plus System shows highly concordant results with the MiSeq System for percent callability with values close to 100%, indicating exceptional performance in the ability to sequence viral genomes, including challenging regions.

Oncology panels

Analysis of results with Pillar oncoReveal Myeloid and BRCA1 & BRCA2 + CNV Panels showed similar amplicon coverage for both panels with the MiSeq i100 Plus and MiSeq Systems, indicating no differences in the amplification and performance of the panels (Figure 7). Both systems accurately called all expected variants in the reference samples analyzed (Figure 7). Additionally, analysis of results generated with the TruSight RNA Pan-Cancer Panel showed that both systems successfully detected known fusions in the reference sample HD834, except for fusions with expression levels below assay detection (Table 9). These data demonstrate that focused oncology panel results on the MiSeq i100 Plus System are equivalent to MiSeq System performance.

Figure 7: Comparable sequencing performance for oncology panels



The MiSeq i100 Plus System shows highly concordant results with the MiSeq System for sequencing coverage (top panels) and variant calling frequency (bottom panels) with the Pillar oncoReveal Myeloid Panel and the Pillar oncoReveal BRCA1 & BRCA2 + CNV Panel.

Table 9: Comparison of fusion detection with the TruSight RNA Pan-Cancer Panel across systems

Expected fusion	RNA expression level	Detected on MiSeq System	Detected on MiSeq i100 Plus System
TPM3-NTRK1	≥ 100 copies/ng	2/2	2/2
QKI-NTRK2	≥ 100 copies/ng	2/2	2/2
SLC34A2-ROS1	≥ 40 copies/ng	2/2	2/2
EML4-ALK	≥ 40 copies/ng	2/2	2/2
CCDC6-RET	≥ 4 copies/ng[a]	2/2	1/2
ETV6-NTRK3	≥ 4 copies/ng[a]	0/2	1/2
Sensitivity		**83%**	**83%**

a. Fusions with expression levels below assay detection are less reliably detected than fusions with higher expression.

Summary

The MiSeq i100 and MiSeq i100 Plus Sequencing Systems feature breakthrough advancements in system design, chemistry, and integrated analysis to deliver operational simplicity, high data accuracy, and exceptional speed. Data from key methods commonly run on the MiSeq System, including sWGS, 16S rRNA sequencing, and targeted sequencing panels across infectious disease and oncology applications were directly compared to data generated using the MiSeq i100 Plus System. Results show that performance on the MiSeq i100 Plus System meets or exceeds MiSeq System performance and supports more data-intensive applications with increased output and significantly shorter sequencing run times.

Learn more

MiSeq i100 Series

illumina®

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

M-GL-02246 v1.0