Performance and Compatibility Evaluation of Celemics NGS Panels with DNBSEQ-G400 Sequencing Platform

Introduction

The use of next-generation sequencing (NGS) technologies has become a more cost-effective screening method in clinical settings. With NGS, researchers and clinicians can make significant advances in oncology by performing whole-genome sequencing (WGS), whole-exome sequencing (WES), targeted sequencing, and more. NGS also offers the sensitivity to detect rare somatic variants, tumor subclones, and circulating DNA fragments.

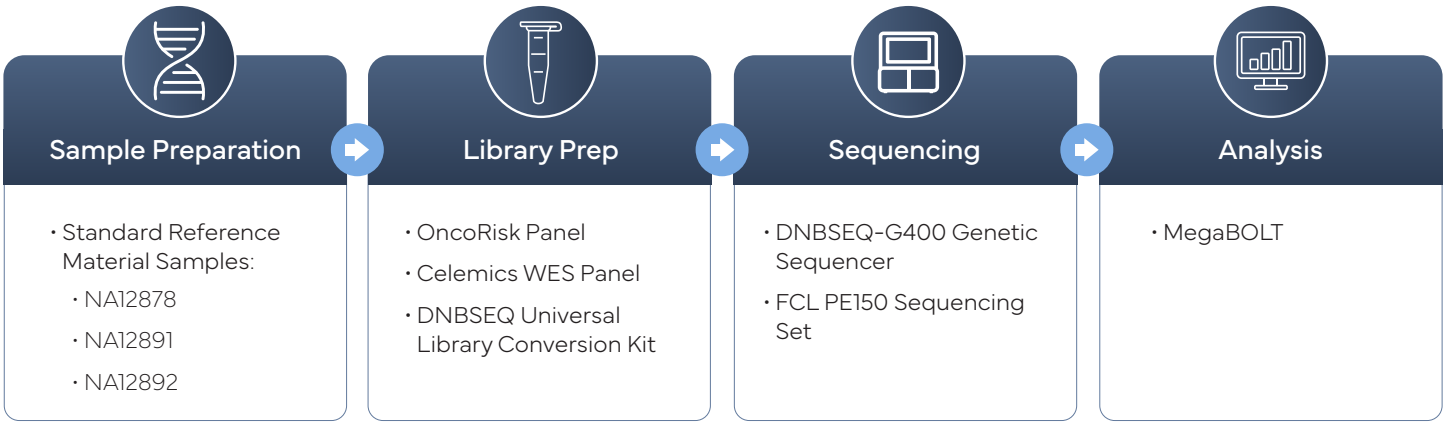
This study evaluates the performance of two Celemics library preparation kits and two target enrichment panels on the Complete Genomics DNBSEQ-G400 sequencing platform. Complete Genomics offers a full lineup of sequencers powered by its proprietary DNBSEQ™ technology. DNBSEQ utilizes Rolling Circle Replication (RCR) instead of traditional PCR amplification, resulting in highly accurate sequencing data with no clonal errors, index hopping, and low duplication rates.

The two panels evaluated were the WES Panel and the OncoRisk Panel. The WES panel is the most comprehensive exome panel on the market, with a target size of 37 Mb, and encompasses all exome and hard-to-capture regions. The OncoRisk Panel targets 31 well-known oncogenes, allowing cancer researchers to reduce sequencing costs compared to WGS or WES approaches.

Along with evaluating Celemics kits and panels, a comparative analysis of data from DNBSEQ-G400 and the Illumina NextSeq 500 sequencing platforms was conducted. Data from these two platforms were assessed for address detection sensitivity and compatibility with the two Celemics kits.

Highlights

- **The integration of Celemics library prep kit and target enrichment panels with Complete Genomics DNBSEQ-G400 platform offers a highly compatible solution and reproducible sequencing results, as well as high sensitivity in variant detection.**
- **Celemics WES and OncoRisk panels provide outstanding sensitivity and specificity, with the WES Panel providing comprehensive coverage of 37 Mb and the OncoRisk Panel targeting 31 key oncogenes for cost-effective, high-precision analysis.**
- **The DNBSEQ-G400 sequencer produces high-quality, reliable sequencing data and demonstrates better mapping rates, coverage, and on-target ratios for the Celemics panels than the Illumina NextSeq 500.**



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

© Copyright 2024 Complete Genomics. All rights reserved.

Methods

LIBRARY PREPARATION AND TARGET ENRICHMENT

For the assessment of the panels and the comparison between the two sequencing platforms, 200 ng of standard reference material samples, NA12878, NA12891, and NA12892 were used by Celemics to prepare libraries (Table 1). The capture probes for corresponding target regions totaling 96 Kb for oncogenes on the OncoRisk panel and 37 Mb for the WES panel. DNA input for the three reference samples was standardized at 200 ng for the library preparation process, which included the following steps: fragmentation, end repair, A-tailing, adapter ligation, and pre-PCR for the indexed NGS library. Each sample library underwent QC with Agilent ® TapeStation D1000 and High Sensitivity D1000 screen tape to check for appropriate library construction. Each sample library was prepared in duplicates and distinguished as L01 and L02 (for Library 1 and Library 2, respectively) to confirm reproducibility. The NGS-prepared DNA library and capture probes were hybridized in buffer to capture target regions of interest through the use of the Celemics target enrichment kit. After hybridization and washing process finished, the captured libraries were amplified by the post-PCR process.

Sample and Sequencing Details	
Sample Names	NA12878, NA12891, NA12892
Panels	Whole Exome Sequencing Panel OncoRisk Panel
Sequencing Kits	DNBSEQ-G400 High-Throughput Sequencing Rapid Sequencing Kit (FCL PE150) NextSeq 500/550 High Output Kit v2.5
Sequencing Devices	Complete Genomics DNBSEQ-G400 Illumina NextSeq 500
Sequencing Strategy	PE150 for both platforms

Table 1. Sample, kits, and sequencing details used in this evaluation.

SEQUENCING

The post-PCR products were prepared for sequencing on the two platforms. Samples for the Illumina NextSeq 500 were sequencing directly. The Complete Genomics samples underwent the additional step of converting them to DNA nanoballs (DNBs) using the DNBSEQ Universal Library Conversion Kit before sequencing on the DNBSEQ-G400.

DATA ANALYSIS

The data analyses were performed using MegaBOLT, a bioinformatics suite performing various analyses, such as alignment (invoking BWA), variant calling (invoking GATK), and quality control, for NGS sequencing data with high efficiency and accuracy. The variant validation analysis is uses RTG Tools and benchmark NA12878 SNP and INDEL sets. Human genome reference hg19 was used for the analysis.

Results

SEQUENCING

Sequencing metrics were compared across various categories of samples (Table 2; only the NA12878 sample is shown as representative data). The libraries' data volumes vary among the samples. Q30 of raw and clean reads is comparable for any library of any sample (~95%), whereas clean rates for OncoRisk are relatively lower than those of WES.

Panel	OncoRisk		WES	
Library No.	L01	L02	L01	L02
# of Raw Reads	16,112,328	17,951,376	67,818,040	73,274,128
Raw Data (Gb)	2.22	2.48	9.65	10.42
Rate of Clean Reads (%)	91.57	91.68	94.46	94.42
Q30 - raw (%)	95.76	95.96	94.68	94.97
Q30 - clean (%)	95.69	95.90	94.64	94.93

Table 2. Sequencing metrics of NA12878 samples on DNBSEQ-G400 sequencer.

ALIGNMENT

Alignment metrics for both the OncoRisk and WES panels were evaluated using NA12878 samples (Table 3), and metrics were compared between the DNBSEQ and Illumina sequencing platforms. Mapping rates, representing the percentage of mapped reads, exceeded 99% for all libraries in both panels. The mapping rates, coverage and on-target ratios for both OncoRisk and WES panels on the DNBSEQ-G400 sequencer were slightly higher compared to the Illumina NextSeq 500. However, overall metrics of both platforms were relatively similar, indicating the comparable sequencing performance for both instruments.

Panel	OncoRisk		WES	
Sequencer	**DNBSEQ**	**NextSeq**	**DNBSEQ**	**NextSeq**
Mapping Rate (%)	99.93	99.79	99.88	99.83
Duplication Rate (%)	4.46	3.71	1.56	1.68
Fold-80 Base Penalty	1.49	1.84	2.24	2.11
On-Target Rate (%)	72.59	69.19	71.59	68.65
Coverage ≥ 100X (%)	99.50	99.39	-	-
Coverage ≥ 20X (%)	-	-	91.66	90.50

Table 3. Alignment metrics of NA12878 samples for OncoRisk and WES panels.

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

© Copyright 2024 Complete Genomics. All rights reserved.

"Fold-80 base penalty" is the fold over-coverage needed to raise 80% of bases in covered targets to the mean coverage and is a measure of sequencing uniformity. The lower the value, the better the uniformity. For OncoRisk panel, the Fold-80 base penalty was calculated to be 1.49 on DNBSEQ-G400, lower than the 1.84 obtained from the Illumina NextSeq 500. Conversely, for the WES panel, the Fold-80 base penalty on Illumina's platform was lower compared to DNBSEQ-G400. Coverage plots demonstrate the uniformity of the sequencing (Figure 1).

GC bias plots show the performance of sequencing in different GC enrichment regions. As mentioned previously, duplicated libraries are labelled L01 and L02 to distinguish each panel. The plots for all four NA12878 samples illustrate that balanced GC regions (30-50%) are well represented (~1), higher GC regions (60-80%) are over-represented (~2) and extreme GC regions are under-represented (~0) (Figure 2). For NA12891 and NA12892 samples, the pattern observed in OncoRisk and WES libraries closely resembles those of NA12878 samples.

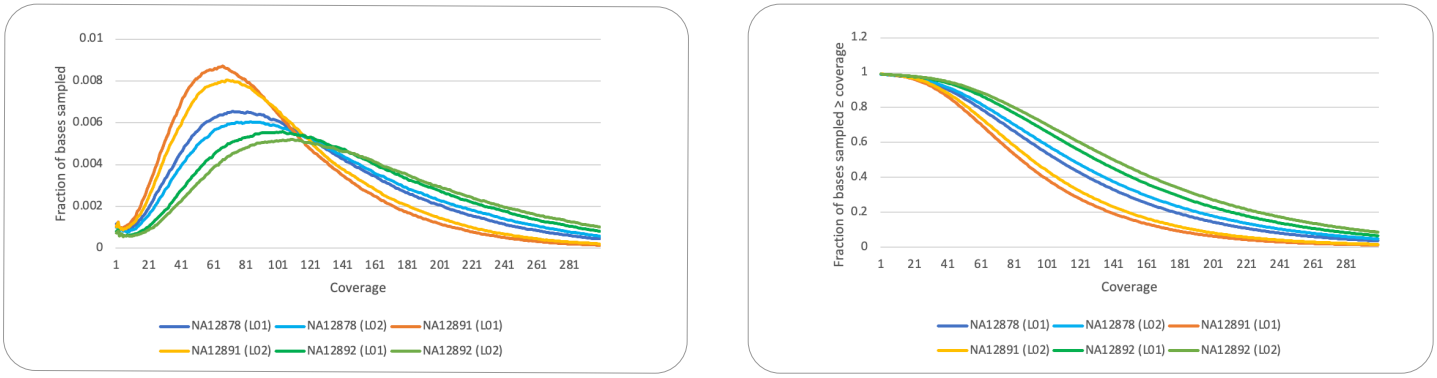


Figure 1. Coverage plots for WES samples on the DNBSEQ-G400 sequencer. Replicate libraries yielded similar coverage profiles, with a sharp peak indicating coverage uniformity. The mean coverage for each sample was ~100X with a high proportion of bases covered >30X.

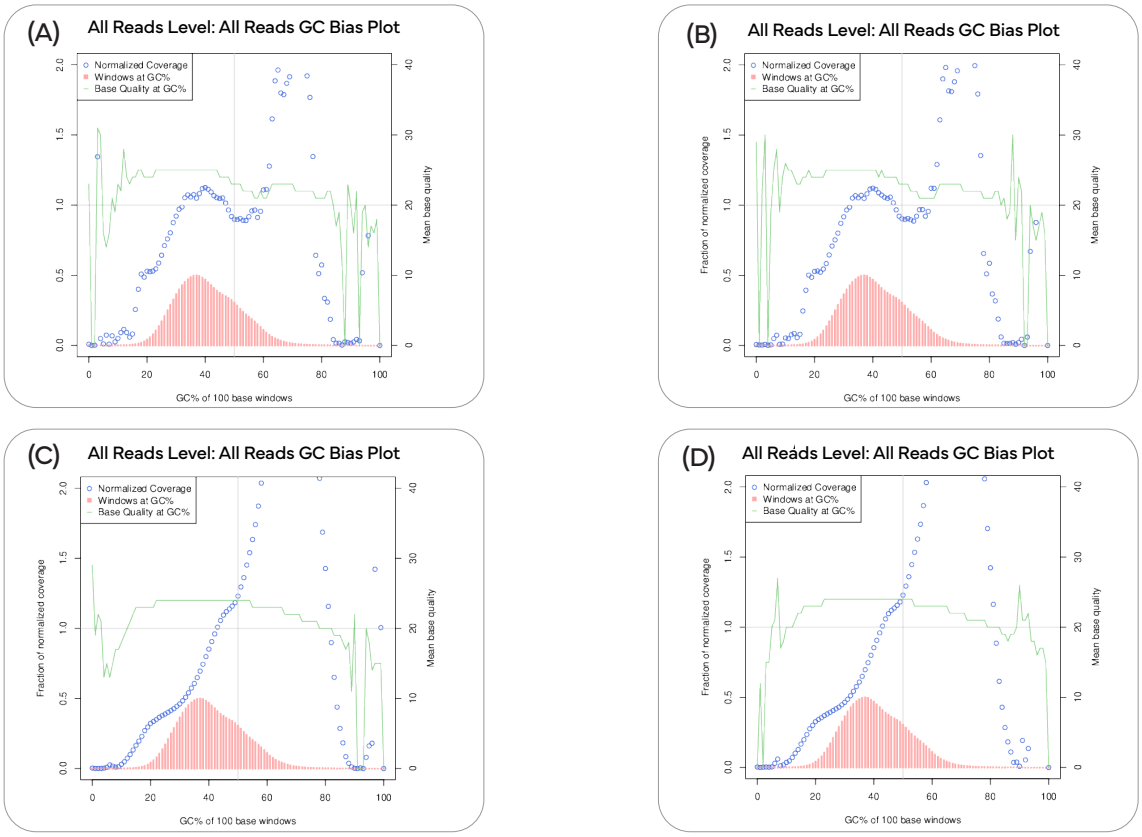


Figure 2. GC-bias plots for NA12878 samples: (A) OncoRisk L01 (B) OncoRisk L02 (C) WES L01 (D) WES L02 on DNBSEQ-G400. Red histograms represent the distribution of genomic GC content, calculated for the reference sequence in 100 bp bins. GC-bias was assessed by plotting normalized coverage for each bin (blue circles). If all sample-to-data processes were completely unbiased, all bins would be equally represented, i.e., the plot for each workflow would be a horizontal distribution centered on a normalized coverage of 1.

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

© Copyright 2024 Complete Genomics. All rights reserved.

PANEL SENSITIVITY ASSESSMENT

The sensitivity of variant calling was conducted by comparing detected variants with the list of known NA12878 standard variants for both platforms and analyzed using software "RTG Tools". Of the generated data set, the germline variants were selectively identified as this panel was designed to investigate germline mutations.

The number of germline variants detected and percent sensitivity using the OncoRisk Panel were compared between the DNBSEQ-G400 and NextSeq 500 platforms (Table 4). Of the germline variants successfully detected, we have observed the 100% concordance in results regardless of the sequencing platform. Moreover, the duplicated libraries sequenced for both DNBSEQ-G400 and NextSeq 500 showed absolute concordance, yielding 100% match in both groups, leading to 100% sensitivity for the OncoRisk, illustrating good performance of the panel and the instruments.

Library No.	Sequencer	No. of Germline Variants	Sensitivity (%)
OncoRisk (L01)	DNBSEQ-G400	52	100
	NextSeq 500	52	100
OncoRisk (L02)	DNBSEQ-G400	52	100
	NextSeq	52	100

Table 4. The numbers of germline variants and the sensitivity values for NA12878 from DNBSEQ-G400 and NextSeq 500 for the OncoRisk Panel.

The total number of single nucleotide variants (SNVs) detected for sample Library 01 and 02 with the WES panel were 29,097 and 30,450, respectively. Based on the number of SNVs detected, the sensitivity values were calculated as 98.27% and 98.42% for replicates L01 and L02, respectively (Table 6), demonstrating robust panel and sequencing performance.

Library No.	Sensitivity (%)
L01	98.27
L02	98.42

Table 5. The sensitivity for detection of SNVs called from the NA12878 samples using the WES panel sequenced on the DNBSEQ-G400 platform.

REPORTING

MegaBOLT generated reports for each sample analyzed. The report includes sequencing, alignment, and variant calling metrics and graphs (Figure 3).

QC Result

Sample	NA12891L01
Read_length	150:150
Read_raw	16,456,954
Read_clean	14,529,738
Read_clean_Rate	88.29%
Q20_raw	98.61%
Q20_1_raw	98.41%
Q20_2_raw	98.81%
Q30_raw	95.37%
Q30_1_raw	94.76%
Q30_2_raw	95.98%
GC_raw	41.19%
Q20_clean	98.62%
Q20_1_clean	98.30%
Q20_2_clean	98.94%
Q30_clean	95.33%
Q30_1_clean	94.41%
Q30_2_clean	96.25%
GC_clean	40.80%
AT_1_Separation	0.40%
AT_2_Separation	0.30%
GC_1_Separation	0.20%
GC_2_Separation	0.60%

Base Distribution

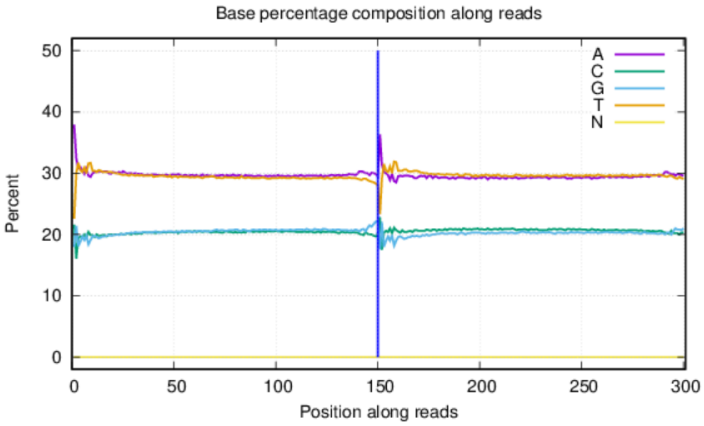


Figure 3. Part of an example of Complete Genomics MegaBOLT report.

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

© Copyright 2024 Complete Genomics. All rights reserved.

Summary

Celemics OncoRisk and WES panels were sequenced on Complete Genomics DNBSEQ-G400 and Illumina NextSeq 500 platform to evaluate panel performance and compatibility across different sequencing technologies. The NGS data sets were analyzed using the Complete Genomics MegaBOLT pipeline and cross-referenced with the Celemics pipeline for confirmation.

The DNBSEQ and NextSeq instruments produced comparable sequencing data, yielding similar performance metrics for both the OncoRisk and WES panels. These metrics demonstrated robust reproducibility among duplicated libraries. Furthermore, the variants detected with both OncoRisk and WES panels across the platforms were either identical or highly concordant. Taken together, the results and performance data generated in this study demonstrate that when coupled with the Celemics OncoRisk and WES panels, the DNBSEQ-G400 platform can produce sequencing data of equivalent, if not superior, quality compared to the NextSeq platform. Additionally, it highlights that Celemics' libraries and NGS panels exhibit high compatibility with various sequencing platforms, enabling the generation of robust sequencing results.

Ordering Information

Category	Product	Cat. No.	Vendor
Library Prep	Whole Exome Panel; Singleplex	WEXL.HM.ES16/96	Celemics
	Whole Exome Panel; Multiplex	WEXM.HM.EM2/4/8/16	Celemics
	OncoRisk Panel; Singleplex	ONCS.HM.ES16/96	Celemics
	OncoRisk Panel; Multiplex	ONCM.HM.EM2/4/8/16	Celemics

Category	Product	Cat. No.	Vendor
Library Prep	DNBSEQ Universal Library Conversion Kit	940-000963-00	Complete Genomics
Instruments	DNBSEQ-G400 Genetic Sequencer	900-000641-00	Complete Genomics
Sequencing Reagents	DNBSEQ-G400 High-Throughput Sequencing Set (FCL PE150)	940-000810-00	Complete Genomics
Software	MegaBOLT Bioinformatics Analysis Accelerator	900-000677-00	Complete Genomics

Contact Celemics, Inc. for more information

CELEMICS

Website: https://www.celemics.com
Tel: 951-805-8845
Email: sales-us@celemics.com

To learn more, visit completegenomics.com

Contact us:
info@completegenomics.com

Technical support:
United States: US-TechSupport@completegenomics.com
Canada: CA-TechSupport@completegenomics.com

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