Reliable NGS-based aneuploidy analysis in trophectoderm research samples with Embgenix PGT-A and DNBSEQ sequencing platforms

Highlights

- The Embgenix PGT-A Kit (RUO) is adaptable to multiple Complete Genomics DNBSEQ sequencing platforms using different sequencing library conversion kits
- Data produced using Embgenix PGT-A displayed consistently high quality, regardless of sequencing platform
- Copy number variation (CNV) analysis with data from Embgenix PGT-A showed reproducible accuracy across sequencing platforms

Introduction

Preimplantation Genetic Testing for Aneuploidy (PGT-A) is a process to identify potential chromosomal abnormalities in trophectoderm (TE) samples. Recently, researchers have started to use next-generation sequencing (NGS) as an effective, comprehensive approach to PGT-A. Embgenix PGT-A has been optimized to provide NGS-based analysis of whole chromosome, segmental, and mosaic aneuploidies in DNA from research samples (not for diagnostic use), but thus far, this library preparation kit has only been validated with Illumina® sequencing systems. As demand for genome profiling research expands, it is important to evaluate the compatibility of this kit with a broader range of sequencers.

This study investigates the compatibility of Embgenix PGT-A with DNBSEQ sequencing platforms—specifically, the Complete Genomics DNBSEQ-G400RS (DNBSEQ-G400), DNBSEQ-G99RS (DNBSEQ-G99), and DNBSEQ-T1+RS (DNBSEQ-T1+) sequencers. For this experiment, TE research samples were processed per the Embgenix PGT-A protocol. Libraries were generated with bead-based cleanup, normalization, and equimolar pooling, then subsequently sequenced on the the Illumina NextSeq® 550 or converted to DNBSEQ-compatible formats using one of two different methods: the DNBSEQ Universal Library Conversion Kit (Univ) or the DNBSEQ OneStep Library Conversion Kit (OS). Sequencing data from DNBSEQ platforms was directly compared with data acquired on the Nextseq 550.

Overall, data produced with Embgenix PGT-A using DNBSEQ platforms demonstrated exceptionally high concordance in CNV calling with data from the NextSeq 550. Concordance rates were 97.6% for the DNBSEQ-G400 (with the DNBSEQ OneStep Library Conversion Kit) and DNBSEQ-T1+ (with the DNBSEQ Universal Library Conversion Kit) and 100.0% for the DNBSEQ-G400 and DNBSEQ-G99 (with the DNBSEQ Universal Library Conversion Kit) as well as the G99 (with the DNBSEQ OneStep Library Conversion Kit). Additionally, base quality and GC content on all DNBSEQ platforms closely matched those of the NextSeq 550. These results confirm that Embgenix PGT-A can be seamlessly integrated into alternative sequencing workflows without compromising accuracy.

Methods

Samples

Samples for this study were derived from two primary sources: the control set consisted of genomic DNA (gDNA) from GM08331 cells, with each reaction containing 30 pg DNA (roughly the equivalent of five cells); the experimental set comprised of TE samples. In total, 43 samples were analyzed. A detailed breakdown of the sample types is provided in Table 1.

Sample type	Number of samples
TE	41 (one per TE sample)
gDNA from GM08331 (chr13 segmental aneuploid)	2

Table 1. Summary of sample types and number of samples. Chr: chromosome.

Sequencing

Sequencing libraries were prepared for each TE sample, as well as control gDNA, using the Embgenix PGT-A Kit (RUO) (Takara Bio, Cat. # 634760). Libraries were then converted to a DNBSEQ-compatible format using one of two methods:

- Univ: DNBSEQ Universal Library Conversion Kit (Complete Genomics, P/N 940-000963-00)
- OS: DNBSEQ OneStep Library Conversion Kit (Third Party) (Complete Genomics, P/N 940-001648-00); conversion performed by third party

For libraries converted with the DNBSEQ Universal Library Conversion Kit, sequencing was performed as follows:

For sequencing on the DNBSEQ-G99RS platform, 30 femtomoles (fmol) of single-stranded circular (ssCir) DNA was used as input for DNA nanoball (DNB) preparation and sequencing using the reagents provided in the DNBSEQ-G99RS High-throughput Sequencing Reagent Set V2.0 (App-D FCL PE150) (P/N 940-001871-00), according to the manufacturer's instructions.

For sequencing on the DNBSEQ-G400RS platform, 60 fmol of ssCir DNA was used for DNB preparation using the DNB buffer from the DNBSEQ High-throughput Sequencing Primer Kit (App-D) (Paired-End) (P/N 1000016952), and sequencing was performed with the DNBSEQ-G400RS High-throughput Sequencing Reagent Set (FCL PE150) (P/N 940-001801-00), according to the manufacturer's instructions.

For sequencing on the DNBSEQ-T1+ platform, 90 fmol of ssCir DNA was used for automated DNB generation and sequencing on the DNBSEQ-T1+, following the manufacturer's instructions. Sequencing (PE150+8+8) was performed using the DNBSEQ-T1+RS High-throughput Sequencing Reagent Set (FCL PE150) (P/N 940-002566-00), according to the manufacturer's instructions.

For libraries converted with the DNBSEQ OneStep Library Conversion Kit (Third Party), DNBs were directly loaded for sequencing on the DNBSEQ-G400 and DNBSEQ-G99 using the sequencing reagent sets above according to the manufacturer's instructions.

The NextSeq 550 served as the reference platform. Sequencing was run using mid output 150-cycle runs, with data merged across lanes and uniformly down-sampled to 1.5 x 10^6 reads per sample. The DNBSEQ platforms (Complete Genomics DNBSEQ-G400, DNBSEQ-G99, and DNBSEQ-T1+) all employed paired-end 150 bp runs with reads downsampled and trimmed to 2 x 76 bp to be consistent and compatible with the Embgenix Analysis Software (RUO).

Data analysis

Quality control (QC) was performed on sequencing data using FastQC and MultiQC, assessing metrics such as GC content and percentage of bases with a quality score ≥Q30.

For sequencing data analysis, Embgenix Analysis Software was used to determine the percentage of informative reads and to calculate the derivative log ratio spread (DLRS). CNV calls were generated automatically, and mosaic percentages were calculated by the software. Additional details on sequencing protocols and analytical workflows are summarized in Table 2.

Company	Sequencer	Sequencing kit	Conversion method	Data processing	Data QC and analysis
Illumina	NextSeq 550	Mid Output 150c	N/A	• Merged lanes • Downsampled to 1.5 x 10^6 total reads	FastQC and MultiQC
Complete Genomics	DNBSEQ-G400	FCS PE150	Univ	• Merged lanes • Downsampled to 1.5 x 10^6 total reads • Trimmed to 2 x 76 bp	Embgenix Analysis Software
	DNBSEQ-G99	PE150			
	DNBSEQ-G400	FCL PE150	OS		
	DNBSEQ-G99	PE150			
	DNBSEQ-T1+	PE150	Univ		

Table 2. Summary of materials and protocols used for sequencing and analysis.

Results

High-quality data generated from all sequencing conditions

For all combinations of sequencers and conversion methods, Embgenix PGT-A produced data with excellent quality. Base quality assessments indicated that the DNBSEQ-G400, DNBSEQ-G99, and DNBSEQ-T1+ platforms achieved higher percentages of bases with ≥Q30 compared with the NextSeq 550, while GC content remained consistent across all platforms. For all samples, % informative reads and DLRS were also assessed. Both library conversion methods (Univ and OS) were successful in producing high-quality libraries that yielded sequencing data with quality metrics comparable to those obtained from the NextSeq 550. Notably, data generated using the DNBSEQ-T1+ platform displayed the highest % bases ≥Q30 and % informative reads, as well as the lowest % GC content and DLRS, all of which suggest a superior signal-to-noise ratio. Quantitative results are detailed in Table 3.

Sequencer (with conversion method)	% bases ≥Q30	Average % GC content	% informative reads (average)	DLRS (average)
NextSeq 550	92.23%	42.19%	61.02%	0.0931
DNBSEQ-G400 (Univ)	98.44%	42.51%	60.59%	0.0937
DNBSEQ-G99 (Univ)	98.44%	42.43%	62.90%	0.0915
DNBSEQ-G400 (OS)	97.54%	42.53%	58.08%	0.0909
DNBSEQ-G99 (OS)	98.13%	42.49%	62.51%	0.0892
DNBSEQ-T1+ (Univ)	99.81%	40.65%	64.17%	0.0868

Table 3. Sequencing QC metrics for libraries across all sequencing conditions.

Before undergoing CNV analysis, sequencing data was strictly assessed for percentage of informative reads (i.e., reads that are useful for analysis) and DLRS (Figure 1). Data from all combinations of sequencers and conversion methods exceeded the minimum threshold for informative reads (>50%) and fell below the maximum threshold for DLRS (<0.23), confirming that the data was of sufficient quality for reliable downstream CNV analysis.

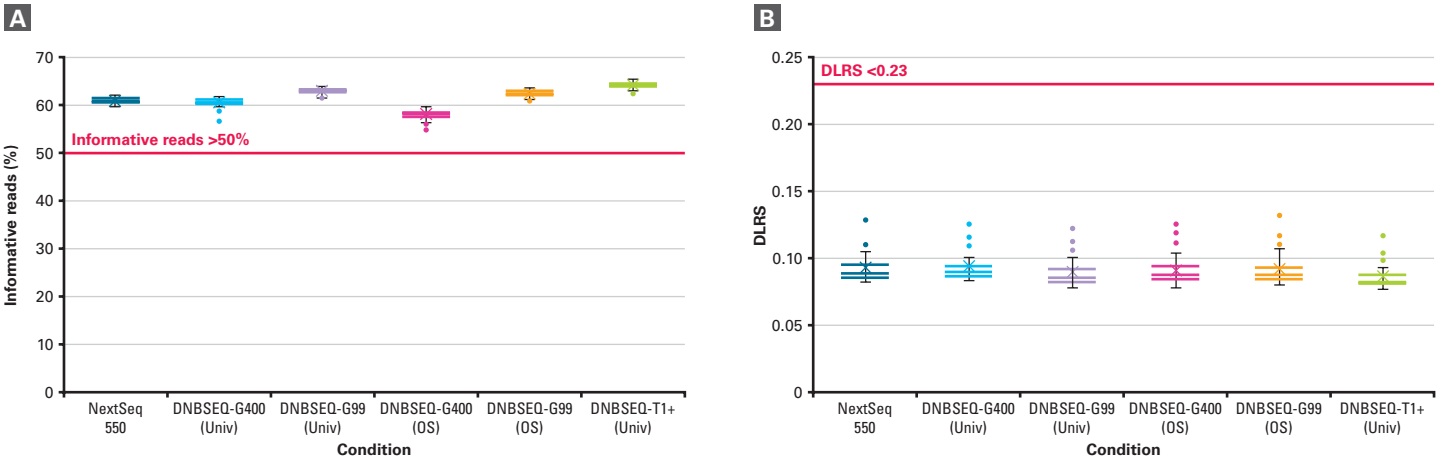

Figure 1. All combinations of sequencing platforms and conversion methods produced data of sufficient quality for CNV analysis. **Panel A.** Box and whisker plot showing % informative reads for data generated by the NextSeq 550, DNBSEQ-G400 (Univ or OS), DNBSEQ-G99 (Univ or OS), and DNBSEQ-T1+ (Univ) conditions. Red line indicates minimum threshold of % informative reads for high-quality data. **Panel B.** Box and whisker plot showing DLRS quantifications for data generated from each condition. Red line indicates maximum DLRS threshold for high-quality data.

Accurate CNV analysis with data from all sequencing conditions

Across all sequencing platforms and conversion methods, CNV analysis of control gDNA samples accurately identified both sample sex and the expected segmental aneuploidy on chromosome 13. In the TE sample cohort, concordance with the NextSeq 550 was assessed over a full spectrum of call types, including euploid, segmental aneuploid, low- or high-level mosaic, and aneuploid calls. Calls differing by 20% mosaicism or less—reflecting the contribution of a single cell in a five-cell sample—were classified as concordant, while any calls outside this margin were deemed discrepant. As an example of a concordant call, representative CNV plots below depict two CNVs in chromosomes 4 and 17 identified in a TE sample across all combinations of sequencing platforms and conversion methods (Figure 2).

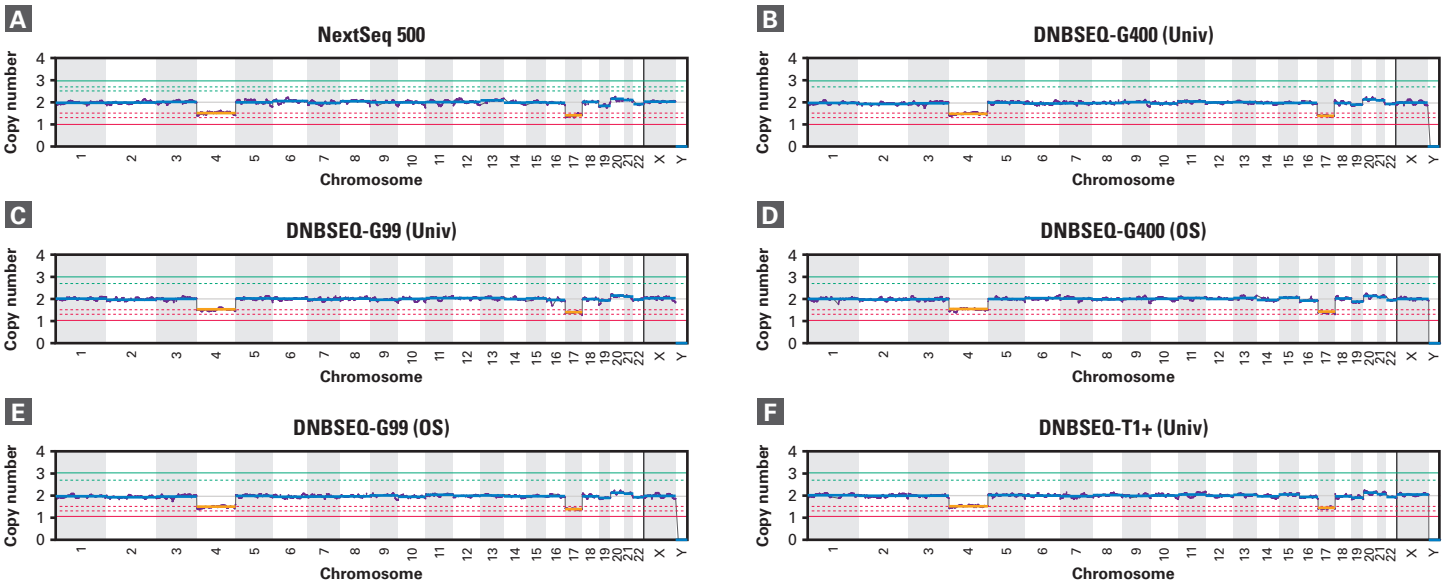

Figure 2. Example CNV plots show concordant identification of two CNVs in chromosomes 4 and 17 across different sequencers and conversion methods.

Comparing CNV analysis results from DNBSEQ platforms with the NextSeq 550 yielded high concordance rates (97.6–100%) for all sequencing conditions (Figure 3). For each of the two discrepant calls, the difference in mosaicism was minimal—under 22%, very close to the 20% agreement threshold (Table 4). Overall, the high concordance rates, which incorporate a variety of CNV call types, confirmed that all major aneuploidy categories were reliably detected by all combinations of sequencing platforms and conversion methods. These findings highlight the robustness and adaptability of Embgenix PGT-A for use with DNBSEQ sequencing platforms, providing laboratories the opportunity to expand their sequencing capabilities while maintaining accuracy.

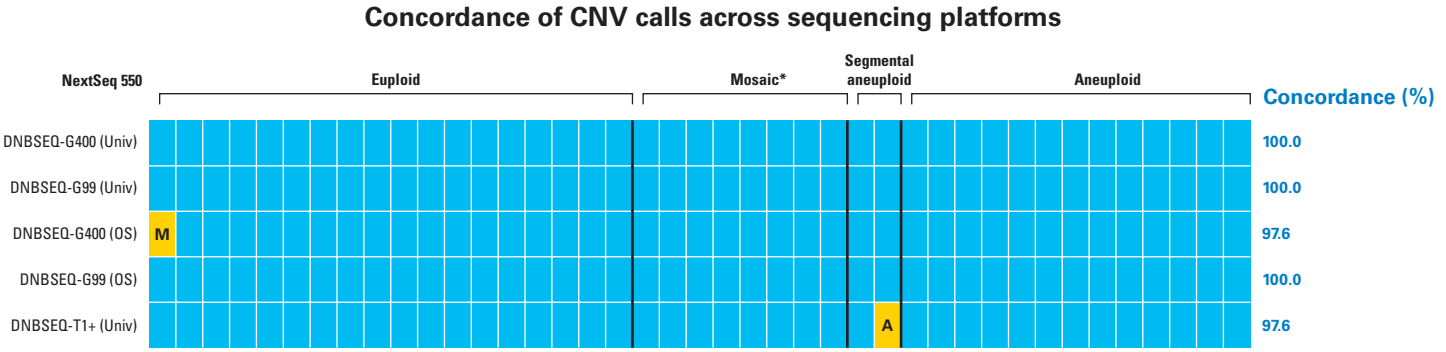

Figure 3. Chart displaying high concordance of CNV calls for 41 TE biopsy samples generated from each sequencing condition with those produced by the NextSeq 550. Concordance rates were calculated to incorporate CNV calls across all major aneuploidy categories. Each column represents one TE biopsy sample. Blue cells indicate CNV call agreement with the NextSeq 550, while yellow cells indicate disagreement. M: mosaic; A: aneuploid.

*Mosaic calls included low and high segmental mosaics and full chromosome mosaics.

Sample	Region (size)	Sequencer	Call	Difference in mosaicism (%)
1	Chr19 (58 Mb)	NextSeq 550	Euploid	20.7
		DNBSEQ-G400 (OS)	Mosaic segmental loss	
28	Chr1 (11 Mb)	NextSeq 550	Segmental gain	21.4
		DNBSEQ-T1+ (Univ)	Mosaic segmental gain	
	Chr22	NextSeq 550	Mosaic gain	21.5
		DNBSEQ-T1+ (Univ)	Gain	

Table 4. Details of discrepant calls between DNBSEQ sequencers and the NextSeq 550. Samples 1 and 28 displayed discrepant calls. For sample 28, the NextSeq 550 flagged the segmental gain as the primary call for chromosome 1 due to the system's prioritization of uniform aneuploidies over mosaics. For all discrepant calls, the differences in mosaicism were <22%, just above the threshold for concordance (20%).

Conclusion

This comprehensive analysis confirms that the Embgenix PGT-A Kit (RUO) is compatible with the Complete Genomics DNBSEQ-G400, DNBSEQ-G99, and DNBSEQ-T1+ sequencers when using either the DNBSEQ Universal Library Conversion Kit or the DNBSEQ OneStep Library Conversion Kit. For all combinations of sequencers and conversion methods, we observed high-quality QC metrics and reliable CNV detection, comparable to those achieved on the NextSeq 550. While laboratory-specific validation and verification are recommended, this data provides a strong technical foundation for research laboratories to integrate DNBSEQ sequencers and the Embgenix PGT-A Kit (RUO) into their PGT-A research workflows, thereby enhancing operational flexibility and cost efficiency without compromising performance.

Takara Bio USA, Inc.
United States/Canada: +1.800.662.2566 • Asia Pacific: +1.650.919.7300 • Europe: +33.(0)1.3904.6880 • Japan: +81.(0)77.565.6999

10.25 US (634188)