Automated generation of Twist Bioscience libraries for high-throughput sequencing with HotMPS chemistry on DNBSEQ platforms

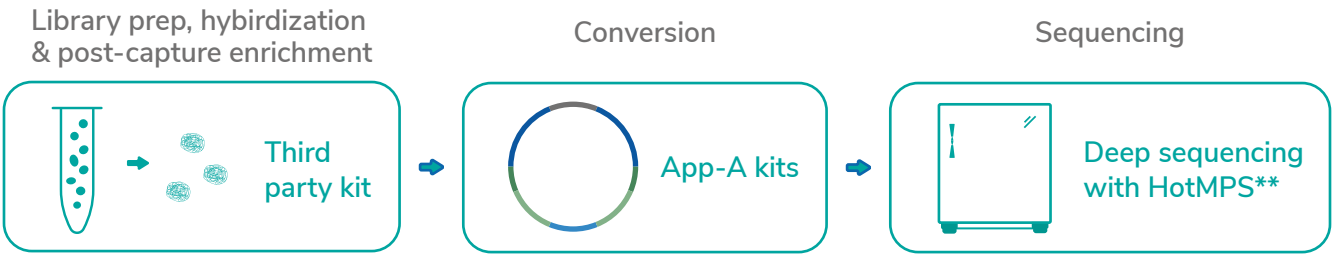
Highlights

- Automate Twist Bioscience library preparation, enrichment, and conversion with a liquid handling script that accommodates flexible throughput and greatly limits operator hands-on time.
- Easily convert banked and new Twist Bioscience libraries into high-fidelity DNA nanoball (DNB) libraries to take advantage of sequencing with exceptional read quality, mapping, and coverage uniformity.
- Establish high-throughput sequencing workflows that achieve consistently high recall and precision in SNP and Indel detection.

From linear library to DNB sequencing

Libraries for DNA and RNA sequencing are generally a pool of linear, double-stranded DNA fragments with adaptors designed for a specific sequencing platform. MGI's DNBSEQ™ technology, however, uses unique DNBs that offer the advantages of linear amplification, coverage uniformity, no amplified errors, and much lower index hopping. The MGIEasy Universal Library Conversion Kit (App-A) converts linear libraries into DNB libraries with a straight-forward, high-fidelity process that is easily automated. This data sheet summarizes details about the conversion of Twist Bioscience libraries and their performance in DNB sequencing.

Workflow overview



Linear DNA library prepared with the Twist Library Preparation EF Kit 2.0, custom panel hybridization and post-capture enrichment with Twist Bioscience Target Enrichment Standard Hybridization V1 on an external thermocycler. Most steps are automatable on the MGISP-960RS.

Library circularization and generation of DNB library with the MGIEasy Universal Library Conversion Kit (App-A).

High-throughput deep sequencing with HotMPS chemistry on the bench-top DNBSEQ-G400RS.

Experimental details

Library preparation using the Twist Bioscience Library Preparation Enzymatic Fragmentation Kit 2.0

The Twist Library Preparation EF Kit 2.0 (Twist Bioscience, Cat. No. 104207) generates end-repaired, dA-tailed, and adapter ligated templates. After enzymatic fragmentation to a defined size range, the library is constructed in a single-tube, high-efficiency reaction that amplifies libraries in a little over 3 hours and accommodates either full-length or universal adapters.

Library enrichment using the Twist Bioscience Target Enrichment Standard Hybridization V1

Library enrichment follows a 17-hour external hybridization to exon regions from amplified, indexed genomic DNA (gDNA) libraries generated with the TWIST Comprehensive Exome Panel, which targets 36.8 Mb of human protein-coding regions. The post-capture enrichment uses Twist Target Enrichment Standard Hybridization V1 to maximize on-targets reads. Automated, post-capture enrichment takes 3 hours and is followed by circularization, digestion, and DNB generation.

Starting material for each library preparation consisted of bottled genome NA12878 at a concentration of 1.35 ng/µl. The generated linear, double-stranded DNA fragments with Twist Universal Adapters and UDI primers designed for a third-party sequencing platform were collected into three pools for subsequent capture. Capture was performed for 2 enrichment reactions. The captured libraries, with concentrations ranging between 16.6 and 18.1 ng/µl, were then converted into DNBs using the MGIEasy Universal Library Conversion Kit (App-A). Sequencing was performed on the DNBSEQ-G400RS and data were evaluated with MegaBOLT's whole-exome sequencing analysis pipeline for germline data. Additional performance metrics were evaluated by analytics provider Platomics.[1]

Automated library preparation, enrichment, and conversion

The MGISP-960RS is a high-throughput sample preparation platform with flexible scripts to automate numerous applications. With integrated 96-channel pipette, systematic safety and contamination controls, and integrated PCR cycler, the MGISP-960RS is a versatile lab robot. The automation script to prepare Twist libraries performs all steps up to and after hybridization in approximately 10 hours. That includes template fragmentation, end repair and A-tailing (ERAT), adapter ligation, washing, capture, fragment circularization and digestion, and finally DNB generation. The library preparation script processes 12 samples and multiples thereof, up to 96. Capture following offline hybridization using the Twist Target Enrichment Standard Hybridization v1 can be done on 1 to 96 pools. Critically, this automated solution frees staff from lengthy and repetitive tasks. Hands-on time is reduced to a series of minimally disruptive intervals, each less than 1.25 hours (Figure 1). In addition, optimized planning of operator schedules is facilitated by the long walk away time and safe stopping points.

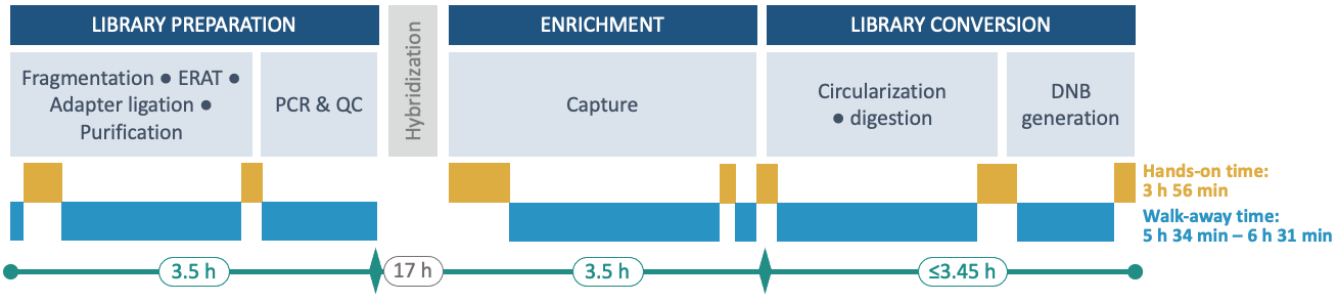


Figure 1: The automated library preparation, enrichment, and conversion workflow on MGISP-960RS limits hands-on time to short intervals between workflow steps, freeing up staff to perform other value-added tasks. Safe stopping points, where samples can be stored at –20°C for up to 24 h, are indicated by green diamonds. Note that hybridization is not included in the automation script.

Library conversion

The MGIEasy Universal Library Conversion Kit (App-A) Library Compatibility workflow produces DNB libraries from linear counterparts prepared with third-party kits (Figure 2). Linear, dsDNA fragments are heat-denatured and circularized. The resulting single-stranded circles serve as templates for rolling circle amplification (RCA) which extends each circle iteratively to create thousands of copies of the original DNA sequence. As a result, the DNB library has no GC bias, dropouts, or errors that accumulate during exponential amplification.

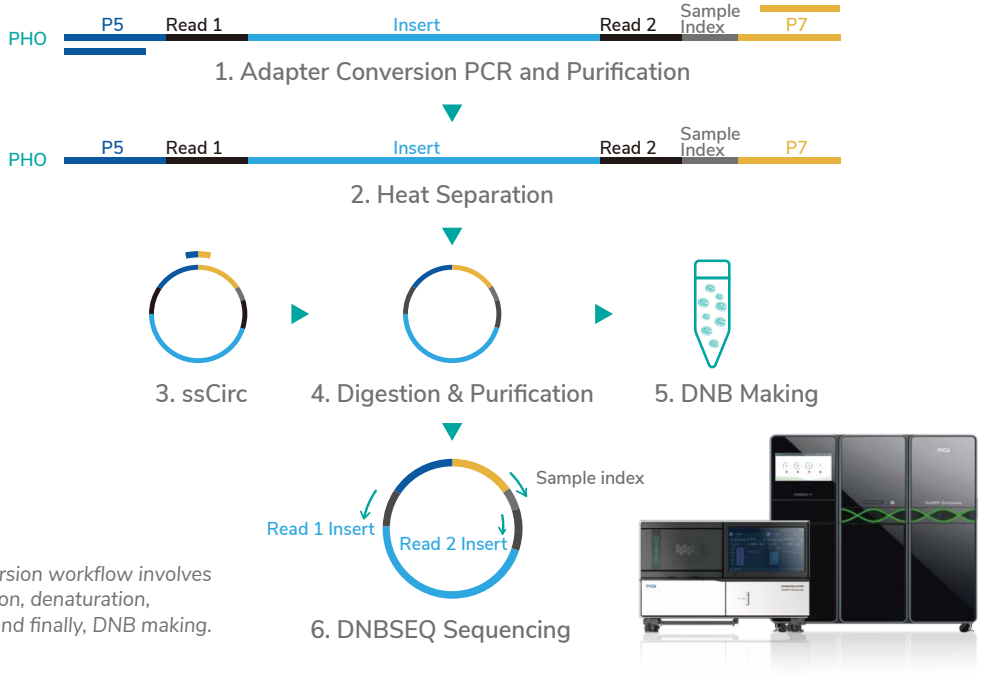
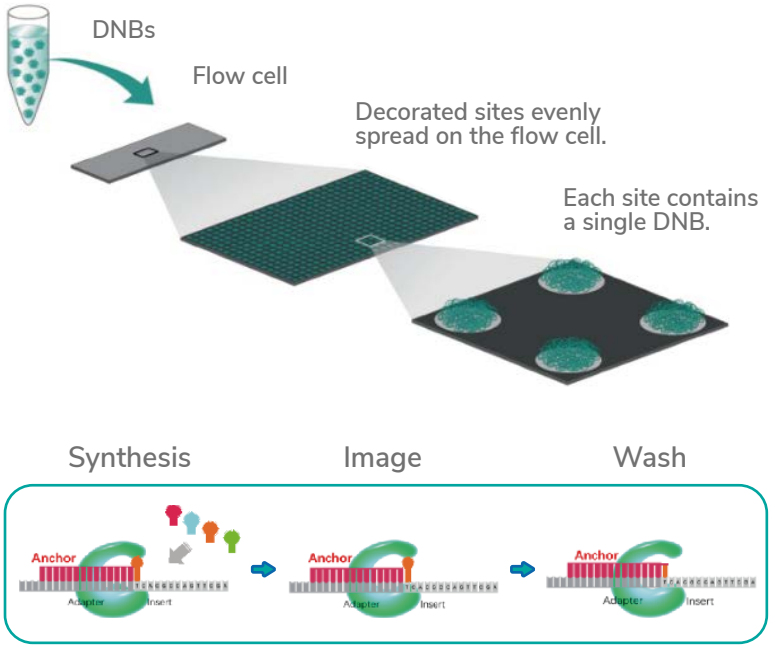


Figure 2: The automated library conversion workflow involves adapter conversion PCR and purification, denaturation, circularization, digestion, purification and finally, DNB making.

DNB Sequencing



DNB sequencing takes place on high-density pattern arrays that immobilize individual DNBs for highly parallelized tracking of dNTP incorporation during strand extension (Figure 3). Sequencing starts with the hybridization of DNBs to anchor spots. After primer hybridization, the flow cell is flushed with fluorescently labeled dNTP probes. Unbound probes are washed away, and bound probes are stimulated to fluoresce. High resolution imaging and proprietary algorithms transform signals into high-quality and highly accurate sequencing results.

Figure 3: DNB sequencing uses the combinatory probe anchor system (cPAS) on a patterned array flow cell to provide high sequencing accuracy with improved imaging and reduced index hopping.

Sequencing performance

DNB sequencing of converted Twist Bioscience libraries results in consistent high-quality mapping, uniformity, and clean read rates.

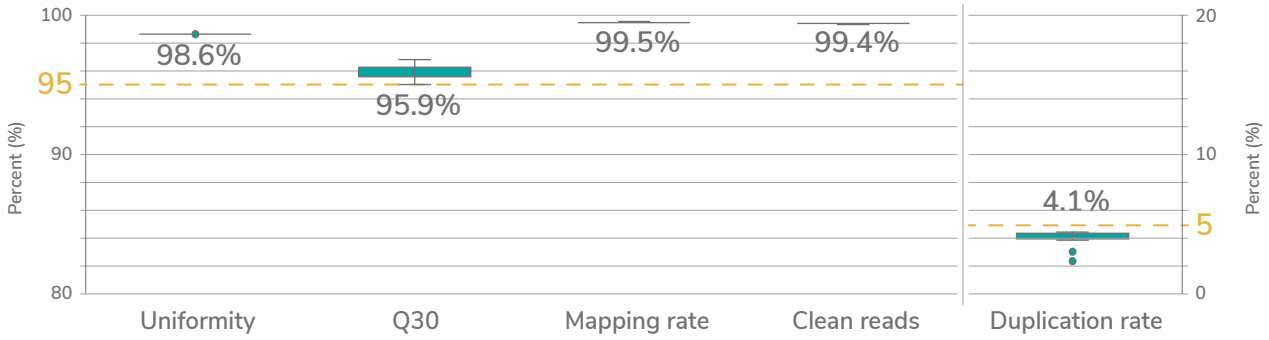
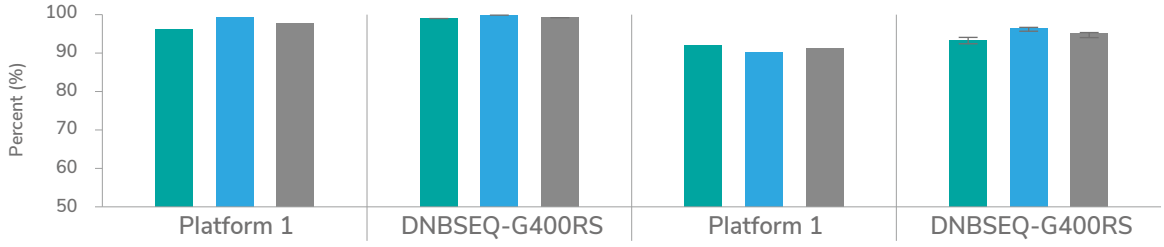


Figure 4. Performance data for 16 samples of a converted library pool sequenced on the DNBSEQ-G400RS. The percent reads with high quality scores and that were clean and mappable was over 95% in all samples. Duplication rate was below 5% and uniformity over 98%. Labeled are the median values.

Converted Twist Bioscience libraries show high recall and precision in variant detection (SNPs and Indels).



	SNPs		Indels	
Recall	96.26	99.13	92.18	93.42
Precision	99.34	99.76	90.27	96.45
F1 score	97.78	99.44	91.22	94.91
	Analyzed with MegaBOLT	Values provided by Platomics		

Figure 5: Performance of variant calling for single-nucleotide polymorphisms (SNPs) and insertion/deletions (Indels) of converted Twist Bioscience libraries. Converted libraries sequenced on the DNBSEQ-G400RS showed consistently high recall and precision with very low inter-sample variability (see error bars). Variant calling was performed with PlatoX LAB IVDD [CE IVD] and benchmarked against the GIAB cureated ground truth dataset for HG001 (analysis service and performance data provided by Platomics[1]).

Variant calling based on sequencing data of converted Twist Bioscience libraries is highly reproducible.

Reproducibility of variant detection was evaluated by Platomics[1] based on variant position.

	Reproducibility (%)
Position (all variant types)	99.99

Calculated as:

$$Reproducibility = \frac{P_B}{P}$$

Where P_B is the base positions that are concordant over all replicas in all runs, and P is all base positions.

Summary

Analyses of the pooled Twist Bioscience libraries converted from linear fragments to DNB libraries demonstrate the fidelity of the conversion and the suitability of the resulting libraries for high-throughput deep sequencing. By every measure of performance, sequencing results from the DNBSEQ-G400RS were of indisputable quality and a reliable basis for precise and reproducible variant calling. Furthermore, with a dedicated script to automate library preparation, enrichment, and conversion, the complete workflow is easily implemented within the context of existing laboratory routines. The precise operation of the MGISP-960RS ensures consistency from run to run - evidenced by very low inter-sample variability in performance metrics - as it minimizes handling errors and frees staff to focus on more value-added tasks.

References

1. *https://www.platomics.com/platox-lab-ivdd*

Ordering information

Product Name	Catalog No.	
Configuration1-MGISP-960 (CE, RUO)	900-000146-00	
Configuration7-MGISP-960 (CE, RUO)	900-000152-00	
Configuration2-MGISP-960 (CE, RUO)	900-000147-00	
Genetic Sequencer DNBSEQ-G400RS*	900-000493-00	
DNBSEQ-G400RS High-throughput Sequencing Set**	1000016950	
HotMPS High-throughput Sequencing Set (G400 HM FCL PE100)***	940-000091-00	
MGIEasy Universal Library Conversion Kit (App-A)	1000004155	DNBSEQ-G400/ DNBSEQ-T7
High-Throughput Pair-End Sequencing Primer Kit (App-A)	1000020832	
High-Throughput Barcode Primer 3 Reagent Kit (App-A)	1000014047	

**This sequencer is only available in selected countries, and its software has been specially configured to be used in conjunction with MGI's HotMPS sequencing reagents exclusively.*
***This sequencing reagent is only available in selected countries.*
****Unless otherwise informed, StandardMPS and CoolMPS sequencing reagents, and sequencers for use with such reagents are not available in Germany, USA, Spain, UK, Hong Kong, Sweden, Belgium, Italy, Finland, Czech Republic, Switzerland, and Portugal.*

Twist Bioscience HQ
681 Gateway Blvd
South San Francisco, CA 94080

MGI Tech Co., Ltd. | MGI-service@mgi-tech.com | en.mgi-tech.com
Building 11, Beishan Industrial Zone, Yantian District, Shenzhen, CHINA, 518083
Tel +86 4000 966 988