DATA SHEET

Encore™ NGS Library Systems *for* Ion Torrent™

Simple, fast, and affordable preparation of DNA libraries for a wide range of next generation sequencing applications

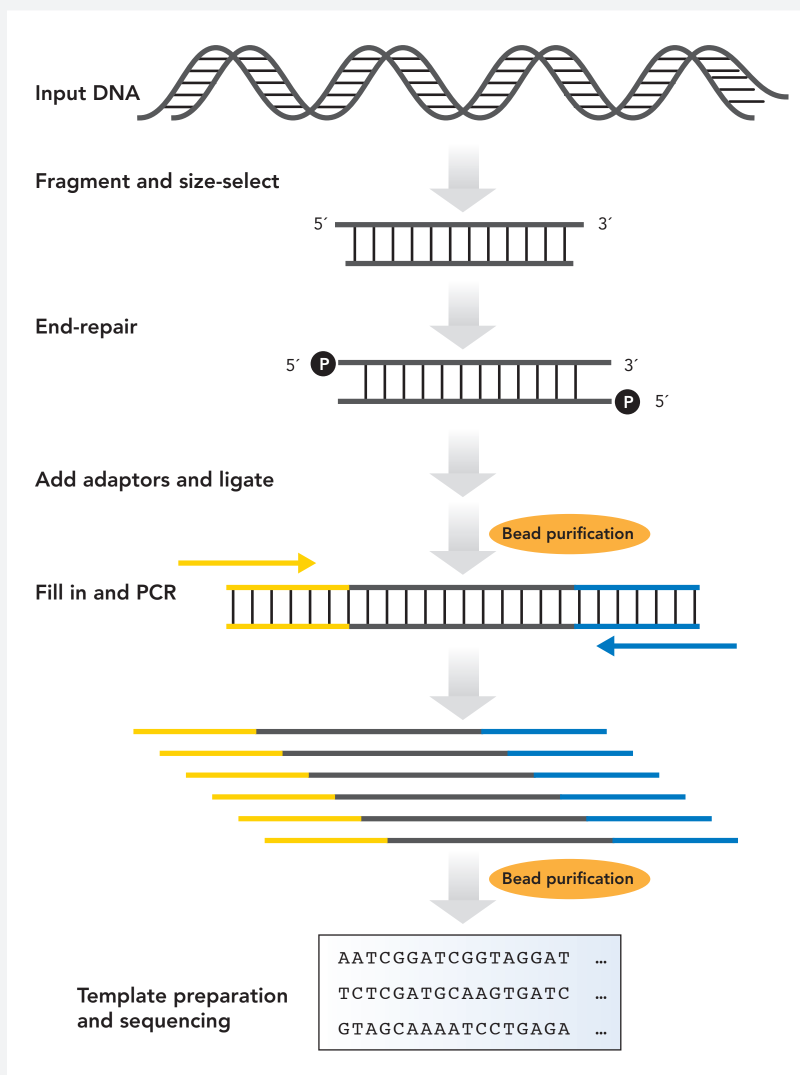
Highlights of the Encore NGS Library Systems *for* Ion Torrent

- **Simple, fast automatable workflow** — Library construction in four hours from as little as 100 ng DNA, with only two bead purification steps. Library size selection is performed using magnetic beads, eliminating the need for specialized equipment or gels.
- **A complete solution for a range of NGS applications** — All required components to make libraries for genomic DNA/exome sequencing, amplicon sequencing, RNA-Seq, Digital Gene Expression (DGE), ChIP-Seq, etc., for the Ion PGM™ System from Life Technologies™
- **Affordable and scalable** — Optional barcoding capability for multiplex sequencing to improve sample throughput and reduce sequencing costs

Introduction

Recent advances in Next-Generation Sequencing (NGS) technology have increased both the throughput and capacity of sequencing platforms, calling for increased efficiency in sample preparation. The Encore NGS Library Systems for Ion Torrent address this need by providing simple, rapid and affordable methods to construct DNA libraries for all major NGS applications.

FIGURE 1 Schematic of Encore NGS Library Systems for Ion Torrent workflow



As shown in **Figure 1**, the workflow consists of four main steps: (1) Fragmented double-stranded DNA is size selected using an innovative, automation-friendly magnetic bead-based procedure to isolate fragments with a median size of 150 base pairs (bp); (2) The ends of these fragments are repaired to generate blunt ends; (3) Adaptor molecules are ligated, placing specific adaptors on the 5′ and 3′ end of each fragment; and (4 Fragments with ligated adaptors are generated by a fill-in reaction and simultaneous PCR amplification to produce the final library with the optional incorporation of barcodes for multiplex sequencing.

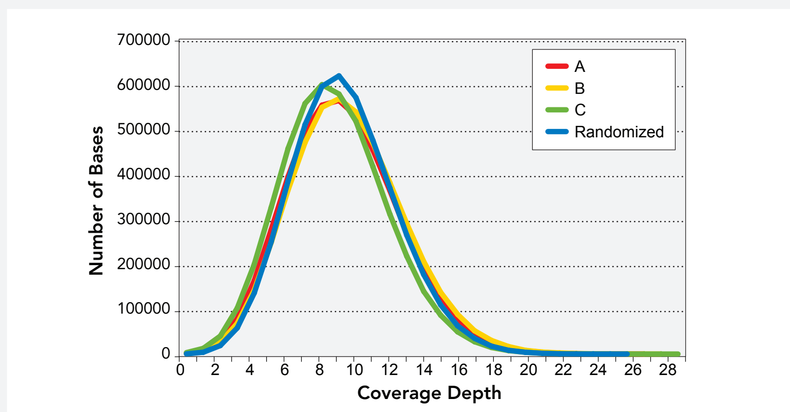
NuGEN® imagine more from less™

The entire workflow can be completed in four hours and yields DNA libraries ready for template preparation on the Ion PGM System. The Encore NGS Multiplex Library Systems for Ion Torrent use unique barcodes to facilitate sequencing of up to sixteen samples on a single chip, thereby dramatically reducing the per sample cost and time required to obtain sequence data.

Sequencing Libraries Prepared with *E. coli* Genomic DNA

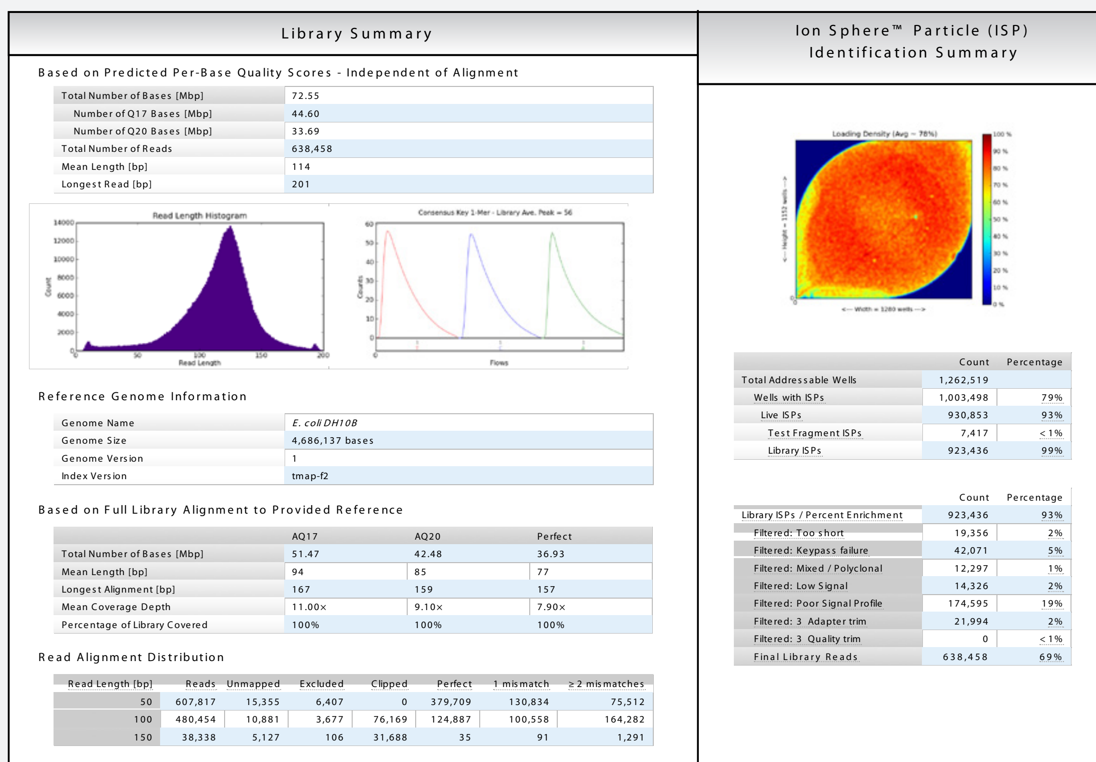
In order to evaluate the performance of the Encore NGS Library Systems for Ion Torrent, three independent sequencing libraries were constructed using 100 ng of *E. coli* genomic DNA and sequenced on the PGM. The distribution of reads from each sample was plotted to determine the depth of coverage across the *E. coli* genome, as well as to assess the reproducibility of library construction and sequencing. An average of 566,000 total reads was obtained for each sample and 430,000 of these were randomly

FIGURE 2 **Sequence coverage of *E. coli* genome**



Three independent sequencing libraries were constructed using 100 ng of *E. coli* genomic DNA using the Encore NGS Library System for Ion Torrent and sequenced on the PGM (A, B, C). The distribution of reads from each sample is plotted to determine the depth of coverage across the *E. coli* genome. The average number of total reads obtained for each sample was 566,000. 430,000 reads were randomly sub-sampled from each data set to generate the above plots. These experimental tracks are shown in red, gold, and green. The theoretical distribution based on randomly chosen sequences with lengths that emulate 430,000 reads from sample A is shown in blue.

FIGURE 3 **Key technical matrices of a typical run using libraries generated with the Encore NGS Library System *for* Ion Torrent**



Library Summary

Based on Predicted Per-Base Quality Scores - Independent of Alignment

Total Number of Bases [Mbp]	72.55
Number of Q17 Bases [Mbp]	44.60
Number of Q20 Bases [Mbp]	33.69
Total Number of Reads	638,458
Mean Length [bp]	114
Longest Read [bp]	201

Reference Genome Information

Genome Name	*E. coli DH10B*
Genome Size	4,686,137 bases
Genome Version	1
Index Version	tmap-f2

Based on Full Library Alignment to Provided Reference

	AQ17	AQ20	Perfect
Total Number of Bases [Mbp]	51.47	42.48	36.93
Mean Length [bp]	94	85	77
Longest Alignment [bp]	167	159	157
Mean Coverage Depth	11.00×	9.10×	7.90×
Percentage of Library Covered	100%	100%	100%

Read Alignment Distribution

Read Length [bp]	Reads	Unmapped	Excluded	Clipped	Perfect	1 mismatch	≥ 2 mismatches
50	607,817	15,355	6,407	0	379,709	130,834	75,512
100	480,454	10,881	3,677	76,169	124,887	100,558	164,282
150	38,338	5,127	106	31,688	35	91	1,291

Ion Sphere™ Particle (ISP) Identification Summary

	Count	Percentage
Total Addressable Wells	1,262,519	
Wells with ISPs	1,003,498	79%
Live ISPs	930,853	93%
Test Fragment ISPs	7,417	< 1%
Library ISPs	923,436	99%

	Count	Percentage
Library ISPs / Percent Enrichment	923,436	93%
Filtered: Too short	19,356	2%
Filtered: Keypass failure	42,071	5%
Filtered: Mixed / Polyclonal	12,297	1%
Filtered: Low Signal	14,326	2%
Filtered: Poor Signal Profile	174,595	19%
Filtered: 3 Adapter trim	21,994	2%
Filtered: 3 Quality trim	0	< 1%
Final Library Reads	638,458	69%

sub-sampled in line with the 4.7 million base pair *E. coli* reference genome.

As shown in **Figure 2**, the reads from each library (red, gold, and green trace) were mapped to the reference genome; the theoretical distribution based on randomly chosen sequences with lengths that emulate 430,000 reads from sample A is shown in blue. Mapping of these 430,000 reads provided an average coverage of 9X, with no positional bias or GC bias observed. These results demonstrate that DNA libraries constructed with the Encore NGS Library System for Ion Torrent provide sequence data with no bias in the positional mapping of reads.

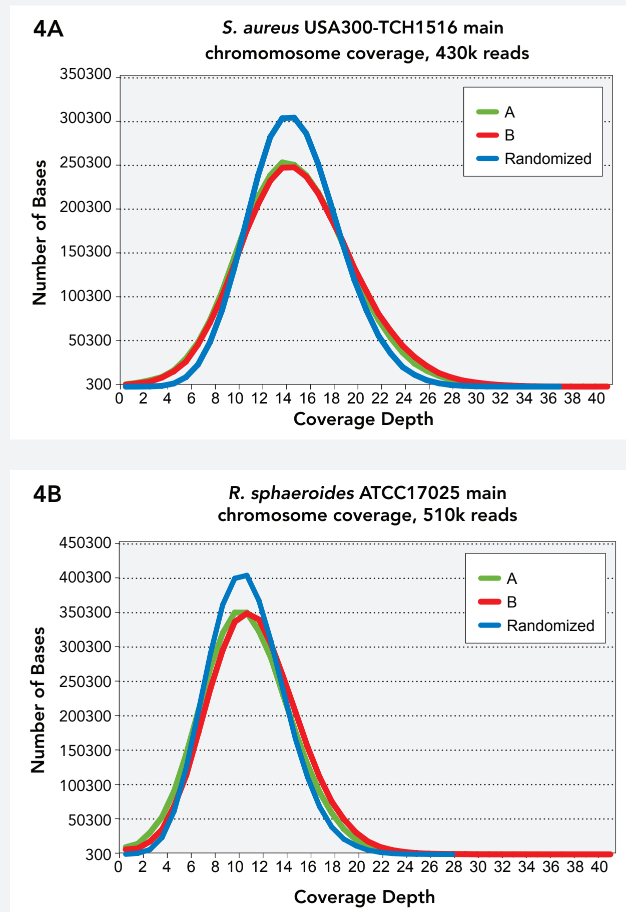
Figure 3 shows the key technical matrices of a typical run using libraries generated with the Encore NGS Library Systems for Ion Torrent and sequenced using the Ion 314 Chip.

Sequencing Libraries Prepared from Low and High GC Genomic DNA

Genomes with low and high GC contents can present a challenge for library construction and subsequent sequencing. Duplicate sequencing libraries were constructed using genomic DNA isolated from *Staphylococcus aureus* (33% GC) and *Rhodobacter sphaeroides* (69% GC) with the Encore NGS Library System for Ion Torrent and sequenced on the PGM.

Figure 4 shows that the actual sequence coverage of these two genomes closely resembles the theoretical coverage, indicating no positional bias or GC bias in the resulting data. **Table 1** shows that the GC contents and library coverage of the resulting sequences from *E. coli*, *S. aureus* and *R. sphaeroides* closely match the actual GC contents of the reference sequences.

FIGURE 4 **Sequence coverage of *Staphylococcus aureus* (33% GC) and *Rhodobacter sphaeroides* (69% GC).**



Sequence coverage of *Staphylococcus aureus* (33% GC) and *Rhodobacter sphaeroides* (69% GC). Duplicate sequencing libraries were constructed using 100 ng of genomic DNA isolated from *S. aureus* and *R. sphaeroides* and the Encore NGS Library System for Ion Torrent and sequenced on the PGM. An average of 542,000 and 555,000 total reads was obtained for *S. aureus* and *R. sphaeroides*; 430,000 and 510,000 of the reads were randomly sub-sampled from each data set to generate the plots, figure 4A and 4B. These experimental tracks are shown in red and green. The theoretical distribution based on randomly chosen sequences with lengths that emulate the experimental reads from sample A is shown in blue.

Unique Barcodes to Enable Sample Multiplexing

In order to further enhance the efficiency and cost-effectiveness of the PGM, the Encore NGS Multiplex Library Systems for Ion Torrent are available with an optional barcoding feature to increase the number of samples that can be sequenced on a chip. The Encore NGS Multiplex Systems for Ion Torrent are supplied with a set of sixteen barcoded adaptors to facilitate multiplex sequencing. The barcodes are intrinsic to the first sequencing read and are designed to provide unbiased representation across all nucleotide bases in the inserts. The barcode sequences are unambiguous, meaning that sequencing errors or deletion errors do not corrupt the code integrity.

As shown in **Figure 5**, use of the barcodes in an 8-plex sequencing experiment provides even read distribution, demonstrating unbiased representation from the barcoded libraries.

The Encore NGS Library Systems for Ion Torrent have been designed for seamless integration with the Ovation RNA-Seq System V2, the Ovation RNA-Seq FFPE System and the Ovation Prokaryotic RNA-Seq System to enable a complete end-to-end solution for transcriptome library construction using unfractionated total RNA samples.

Conclusions

The Encore NGS Library Systems for Ion Torrent offer a number of advantages for researchers engaged in Next Generation Sequencing:

- Simple, fast, automatable workflow — Library construction in as little as four hours, with only two bead purification steps
- A complete solution for a range of NGS applications such as genomic DNA sequencing, RNA-Seq, ChIP-Seq or DGE
- Affordable and scalable — Optional barcoding capability for multiplex sequencing up to 16 samples to improve sample throughput and reduce sequencing costs
- Seamless integration with NuGEN's Ovation RNA-Seq System

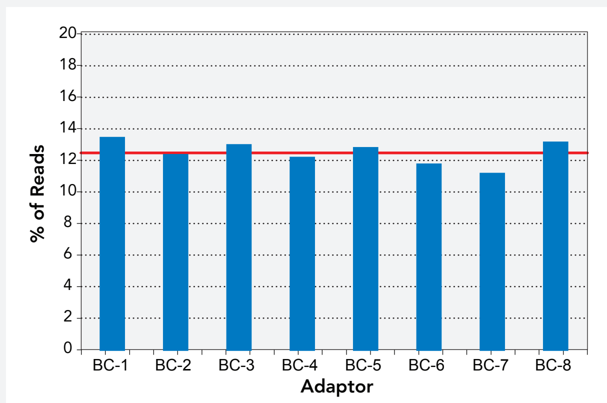
ORDERING INFORMATION

Part No.	Product Name
0306	Encore™ NGS Library System *for* Ion Torrent™
0307	Encore NGS Multiplex Library System I *for* Ion Torrent
0308	Encore NGS Multiplex Library System IB *for* Ion Torrent
7102	Ovation® RNA-Seq System V2
7150	Ovation RNA-Seq FFPE System
9030	Ovation Prokaryotic RNA-Seq System
Technical Documents	
Encore NGS Library Systems *for* Ion Torrent User Guide	

TABLE 1 **GC contents of the resulting sequences from *E. coli* (dh10b), *Staphylococcus aureus* and *Rhodobacter sphaeroides* as compared to the actual GC contents of the reference sequences.**

	Species	Expected GC Content	Mapped GC Content	% of Library Covered
Sample 1	*E. coli* dh10b	50.8%	51.2%	100%
Sample 2	*E. coli* dh10b	50.8%	51.2%	100%
Sample 3	*E. coli* dh10b	50.8%	51.1%	100%
Sample 4	*S. aureus*	32.8%	34.1%	100%
Sample 5	*S. aureus*	32.8%	34.1%	100%
Sample 6	*R. sphaeroides*	68.5%	67.1%	99%
Sample 7	*R. sphaeroides*	68.5%	67.8%	100%

FIGURE 5 **Read distribution of barcoded library reads for 8-plex sequencing**



Eight sequencing libraries were independently constructed with 100 ng *E. coli* genomic DNA using barcoded adaptors in the Encore NGS Multiplex Library System I for Ion Torrent. The libraries were then mixed based on the mass determined by the Agilent Bioanalyzer High Sensitivity DNA Chip. The results indicate an even distribution of reads derived from libraries containing each barcode with no biased presentation from any of the barcoded libraries. The theoretical distribution of barcoded reads for an 8-plex run is 12.5% as indicated by the red line.

NuGEN Technologies, Inc.

Headquarters USA
201 Industrial Road, Suite 310
San Carlos, CA 94070 USA
Toll Free Tel: 888.654.6544
Toll Free Fax: 888.296.6544
custserv@nugeninc.com
techserv@nugeninc.com

Europe
P.O. Box 149
6680 AC Bemmel
The Netherlands
Tel: +31-13-5780215
Fax: +31-13-5780216
europe@nugeninc.com

For our international distributors contact information, visit our website

www.nugeninc.com

©2011 NuGEN Technologies Inc. All rights reserved. The Ovation® and Applause™ families of products and methods are covered by U.S. Patent Nos. 6,692,918, 6,251,639, 6,946,251 and 7,354,717, and other issued and pending patents in the U.S. and other countries. NuGEN, the NuGEN logo, Ovation, SPIA, Ribo-SPIA, WT-Ovation, Encore, Applause, Prelude and Imagine More From Less are trademarks or registered trademarks of NuGEN Technologies, Inc. Other marks appearing in these materials are marks of their respective owners.

For research use only.