TWIST
BIOSCIENCE

Twist 96-Plex Ultra High-Throughput Rapid DNA Library Preparation for Next Generation Sequencing

Azeem Siddique,[1,3] Gaia Suckow,[1,3] Nils Homer,[2] Jorge Bahena,[1,3] Phillip Ordoukhanian,[1,3] Steve Head,[1,3] Keith Brown[3]
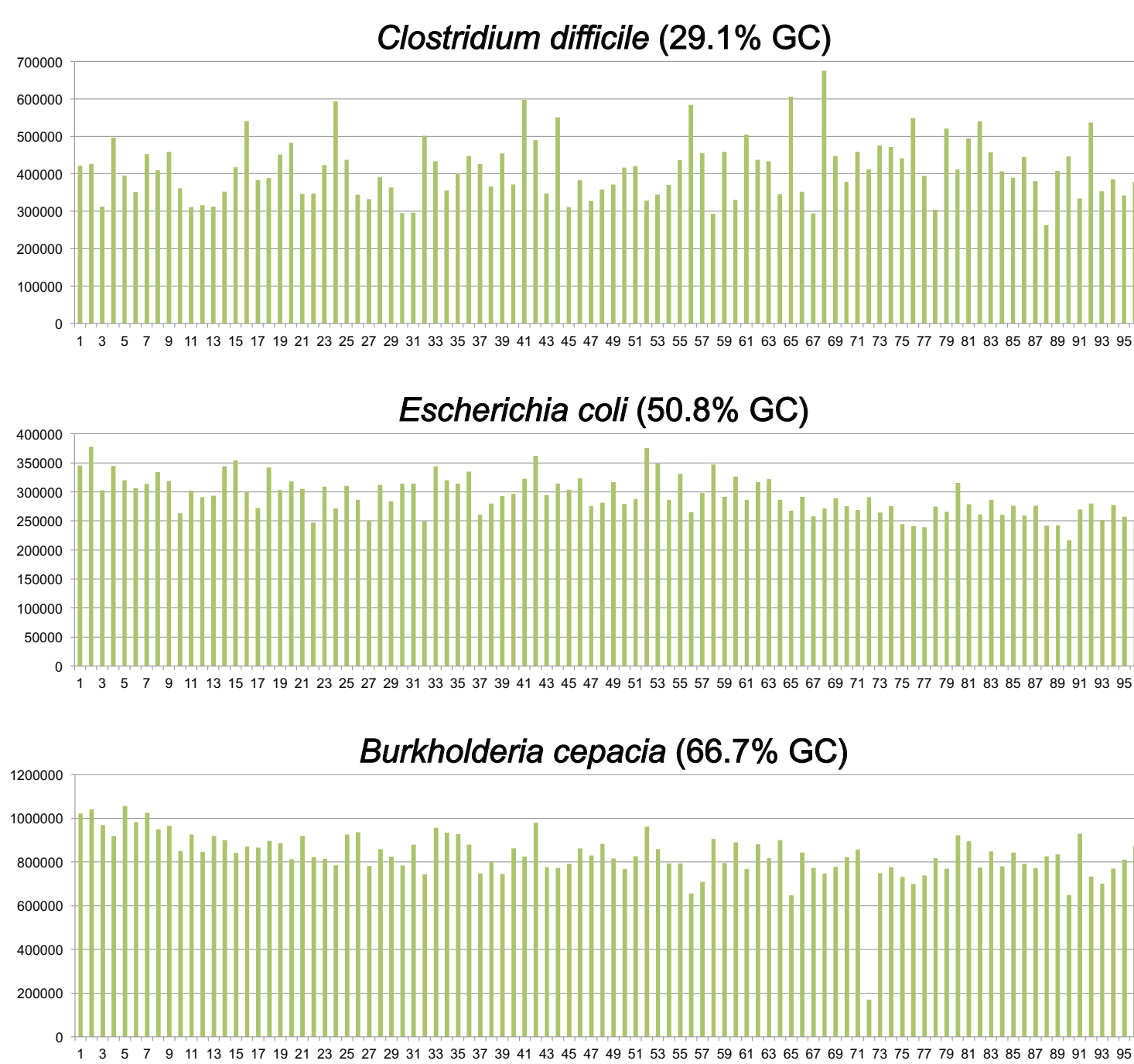
[1]The Scripps Research Institute, La Jolla, CA; [2]Fulcrum Genomics, Somerville, MA; [3]Twist Bioscience, South San Francisco, CA
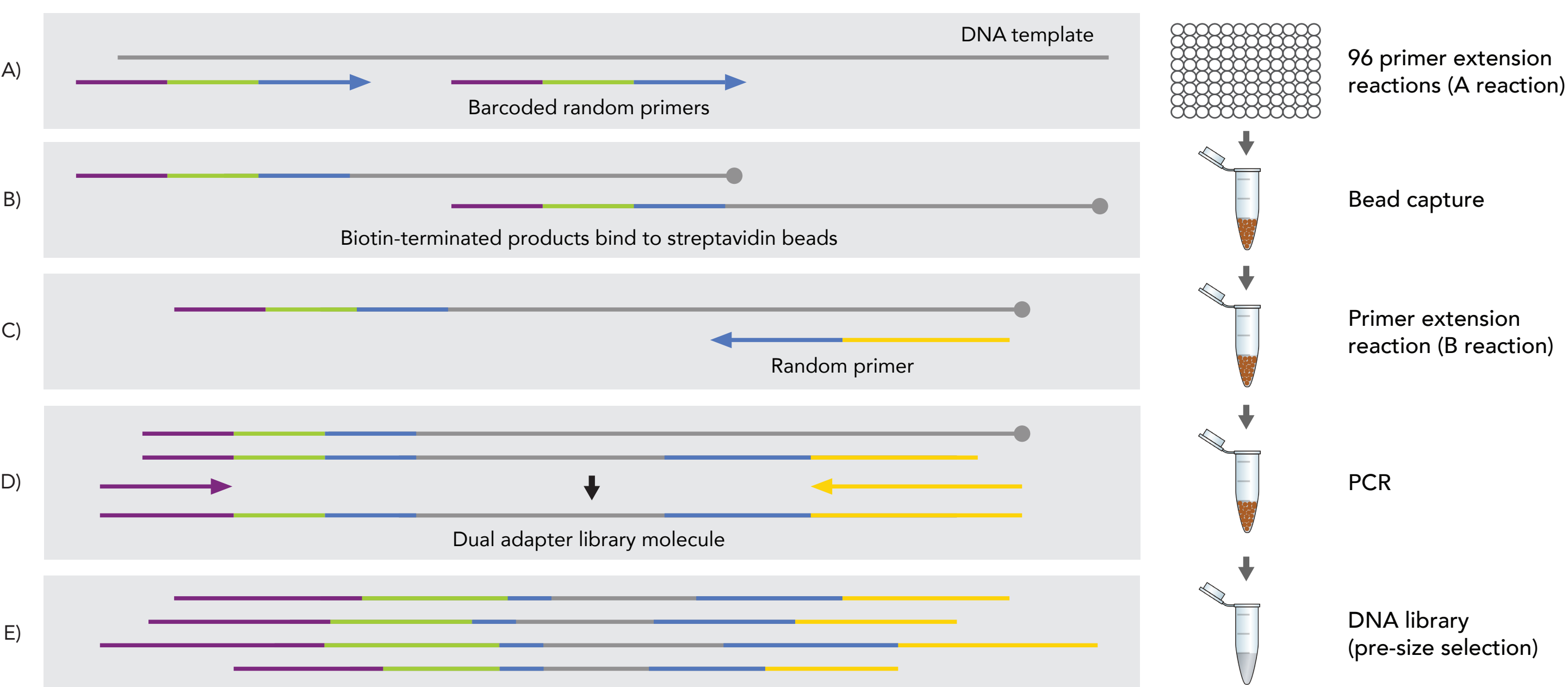
ABSTRACT

Whole Genome Shotgun Sequencing has become the tool of choice for microbial genome analysis. Rapidly declining costs of sequencing, data analysis, data storage and database access will continue to drive adoption. Library construction has not kept pace with these advancements, with costs of preparing a next generation sequencing (NGS) library often exceeding the cost of sequencing. Popular methods of library construction for NGS include fragmentation, end-repair and adapter ligation, and transposase-mediated adapter insertion. The Twist 96-Plex Library Preparation Kit is distinctly different in its approach because it relies on polymerasemediated primer extension for library preparation. The initial step of the prep, involving primer extension with barcoded random primers, is performed in 96-well plate. Each well of the plate contains primers with a unique barcode; consequently, the library generated from each well is uniquely identifiable and can be bioinformatically traced back to the original sample after sequencing. Following this step, the primer extension products are combined into one pool and all subsequent steps, including second strand synthesis and PCR, are performed with the single pool. The library prep is fast, easily automatable and can be tuned to genomes of high and low GC content. **With automation, 960 samples can be easily processed in a single day.** The technology will aid genetic research by helping to increase sample throughput and by reducing processing steps and operating costs. Presented here is Twist 96-Plex Library Preparation Kit sequencing data generated from multiple microbial genomes.

SCHEMATIC OF THE TWIST 96-PLEX LIBRARY PREPARATION KIT

(A) The DNA of interest is denatured and barcoded random primers with partial Illumina P5 adapter sequences (termed "Primer A") are used for primer extension with a DNA polymerase. Since each primer extension reaction occurs in one well of a 96-well plate and each well contains a uniquely barcoded primer, the library generated from each well is uniquely identifiable. The nucleotide mix contains a small fraction of dideoxynucleotides, which causes the primer extension reaction to self-terminate at lengths amenable to sequencing. **(B)** Because the dideoxynucleotides are covalently bound to biotin, the products of primer extension can be captured by streptavidin beads while other DNA fragments are washed away. **(C)** A second round of primer extension with a random primer with partial P7 adapter sequence (termed "Primer B") is performed on the beadcaptured DNA molecules to create the complementary strand. The newly created strand has Illumina adapter sequences at both ends of the molecule. **(D)** PCR is performed with full length Illumina adapters. **(E)** This generates a double-stranded DNA library, which must undergo a final size selection before it can be loaded on a sequencing instrument.
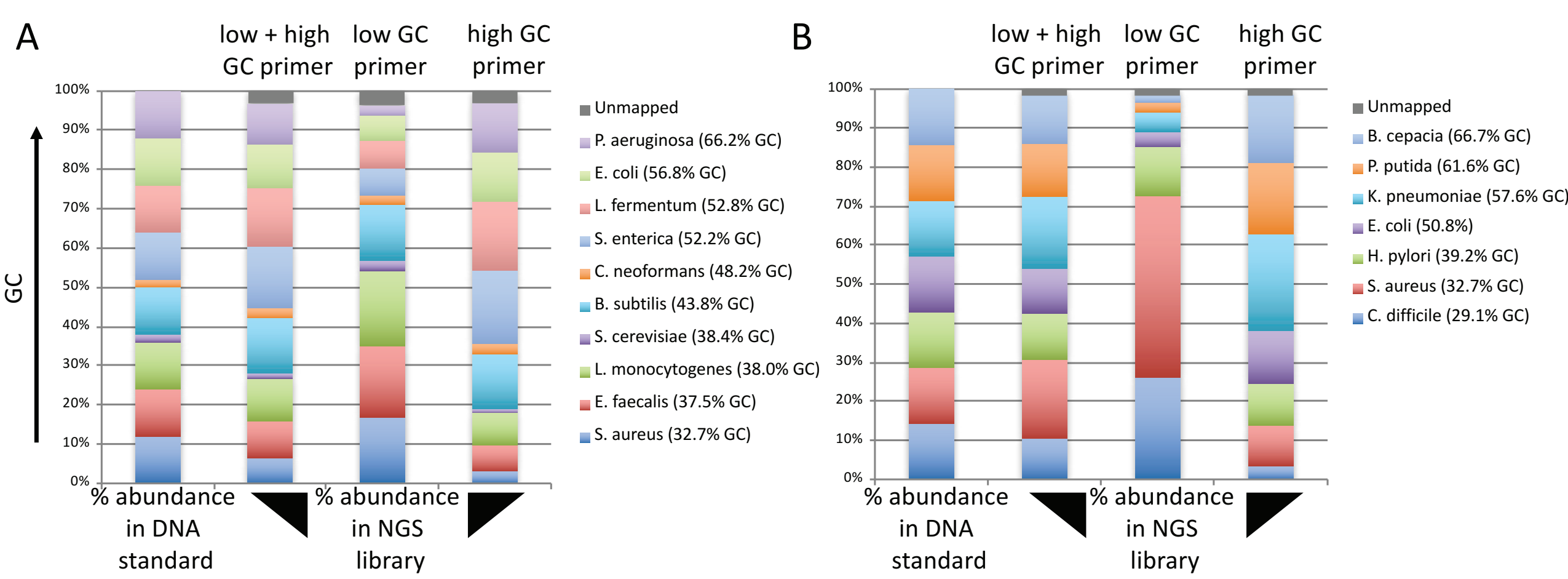
CONCLUSIONS

- Fundamentally different approach to library preparation allows for high-throughput processing of samples in a fast, convenient manner.
- NO fragmentation, NO end repair, NO ligation or associated clean-up steps are required. NO normalization of individual sample yields required.
- Tunable to samples of different GC content.
- Ideal for preparation of libraries from plasmids, synthetic constructs, small genomes and microbiome samples.
- Original content of complex multi-species samples is well-preserved in final sequencing data.

READ COUNT DATA FOR 96 REPLICATES OF THREE BACTERIAL SAMPLES (NO normalization of sample yields required)

50 ng of genomic DNA from a single organism was used as input for 96 primer extension reactions in a 96-well plate. Individually barcoded products of the primer extension reaction were combined for downstream library preparation steps. The final library was size selected and run on an Illumina NextSeq instrument for 2 x 150 nucleotide paired-end sequencing. Data was demultiplexed by index barcode (i.e., the barcode unique to each plate/organism) and by in-line barcode (i.e., sample barcode) to obtain read count data. The histograms depict sequencing read numbers for 96 replicate samples of genomic DNA from *Clostridium difficile*, *Escherichia coli* & *Burkholderia cepacia*, a low GC, medium GC and high GC organism, respectively.

SEQUENCING METRICS (96 technical replicates of each organism)

SAMPLES	SIZE OF GENOME (BP)	NUMBER OF READS	% READS ALIGNED	% DUPLICATION
C. difficile	4,290,252	412,658 +/- 80,701	94.56 +/- 0.34	3.31 +/- 0.36
E. coli	4,639,675	293,713 +/- 33,008	95.27 +/- 0.21	4.05 +/- 0.29
B. cepacia	8,605,950	833,066 +/- 109,518	96.38 +/- 0.14	5.41 +/- 0.43

SAMPLES	MEAN COVERAGE	% GENOME COVERAGE AT 1X	% GENOME COVERAGE AT 5X	% GENOME COVERAGE AT 10X
C. difficile	20.35 +/- 3.97	96.79 +/- 0.35	84.46 +/- 3.05	65.13 +/- 5.89
E. coli	13.55 +/- 1.50	98.24 +/- 0.34	83.54 +/- 3.44	56.25 +/- 5.96
B. cepacia	20.47 +/- 2.65	98.12 +/- 1.15	88.28 +/- 5.88	69.23 +/- 7.17

Key sequencing metrics for the RipTide libraries made from 96 replicates of *C. difficile*, *E. coli* and *B. cepacia* (as described in the legend of the previous figure) are shown in these two tables. Sequencing data was demultiplexed and analyzed for alignment rates, duplication rates and coverage statistics. All data shown is the mean and standard deviation of 96 demultiplexed replicates of each genome. **A total of 960 samples were sequenced on a single NextSeq 500 2 x 150 flow cell. Data for 288 out of 960 samples used for this poster presentation.**

MICROBIOME: READ ALIGNMENT DATA FOR SIX MIXED DNA SAMPLES

(A) 50 ng of a DNA standard representing a community of ten microbial organisms (from Zymo Research) was used as input in three Twist 96-Plex library preps, each performed with twelve replicate samples. **(B)** A simulated microbiome was created by mixing equal quantities of genomic DNA from seven different bacterial species. 50 ng of this DNA mixture was used for Twist 96-Plex Library Preparation, as in (A). Library preps in (A) and (B) were performed with (1) an A primer with low GC content in the 3' random sequence, (2) an A primer with high GC content in the random sequence, and (3) a 1:1 combination of the two primers. In both graphs, the left-most bar in the histogram represents the actual species composition of the DNA standard. Adjacent bars depict the species composition of the Twist 96-Plex libraries after sequencing and alignment of read data to the species in question. Data presented is an average of twelve replicates.