Simple and scalable genome analysis with TELL-Seq™: from haplotype phasing to *de novo* assembly in a tube

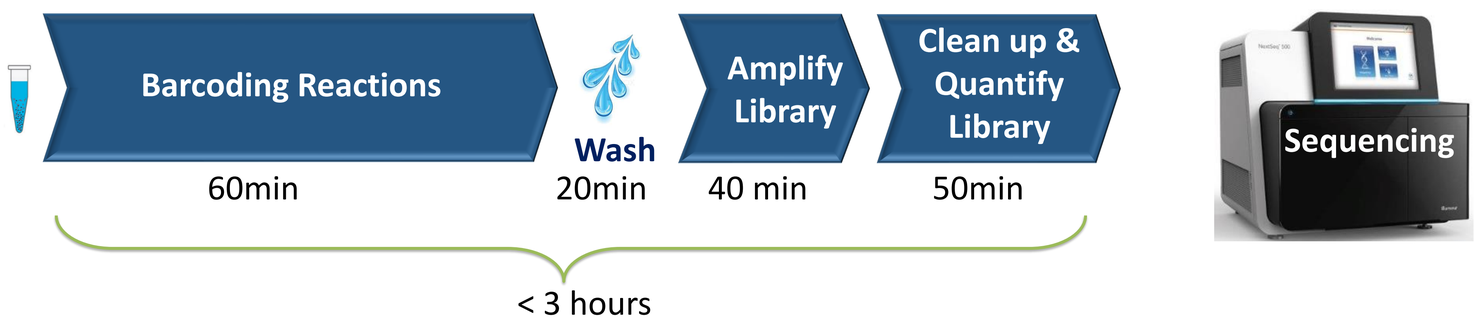
Tom Chen[1], Long Pham[1], Tsai-Chin Wu[1], Yu Xia[1], Gloria Mo[1], Tan Phan[2], Tuan Tran[2], Son Pham[2], Yong Wang[1], Ming Lei[1]
1. Universal Sequencing Technology Corporation, Carlsbad, CA 92011. 2. Bioturing Inc, San Diego, CA 92121
www.universalsequencing.com

INTRODUCTION

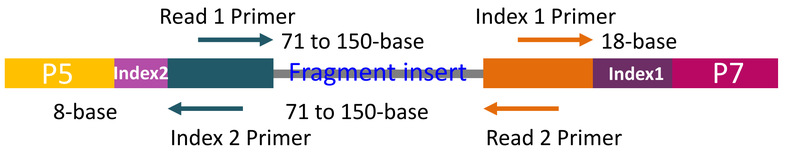
Transposase Enzyme Linked Long-read Sequencing (TELL-Seq™) is a simple and scalable NGS library technology to generate barcode linked-reads for genome scale haplotype phasing and/or *de novo* sequencing. In a PCR tube, under a standard NGS laboratory setting and without a need for any expensive protocol-specific instrument, TELL-Seq Whole Genome Sequencing (WGS) Library Prep Kit will generate an Illumina sequencing library in 3 hours. Multiple samples can be easily processed parallelly in a 96-well format when needed. The library protocol can be adjusted and used for various sizes of genomes ranging from bacteria to human. TELL-Seq library will replace both fragment library and mate-pair library altogether and become a new standard library method for WGS.

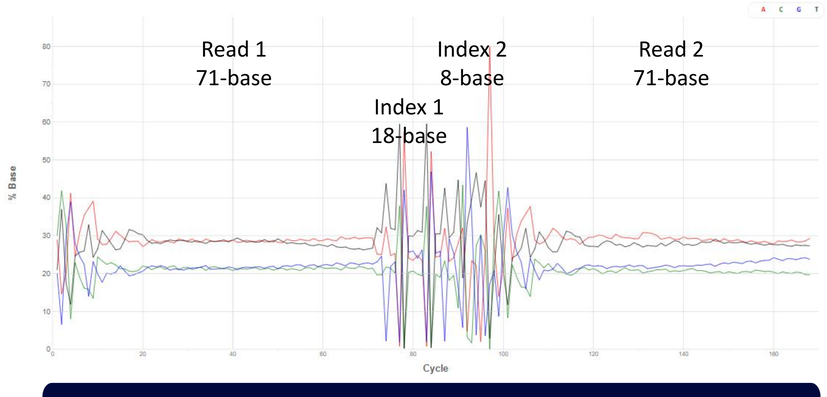
TELL-Seq Library Preparation and Sequencing Workflow





TELL-Seq Library Structure and Sequencing Scheme





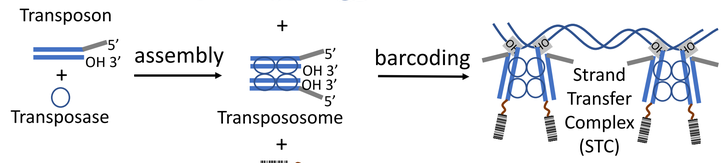
METHODS AND MATERIALS

Standard NGS Sequencing

TELL-Seq Technology

- Standard NGS laboratory setting without any protocol specific capital expenditure
- All barcoding reactions are done in a 0.2ml PCR tube
- 0.5ng – 5ng genomic DNA input based on the genome size
- Works for both small and large genomes ranging from bacteria to human genome
- Highly scalable workflow for preparation of multiple samples in parallel

TELL-Seq Barcoding Reactions

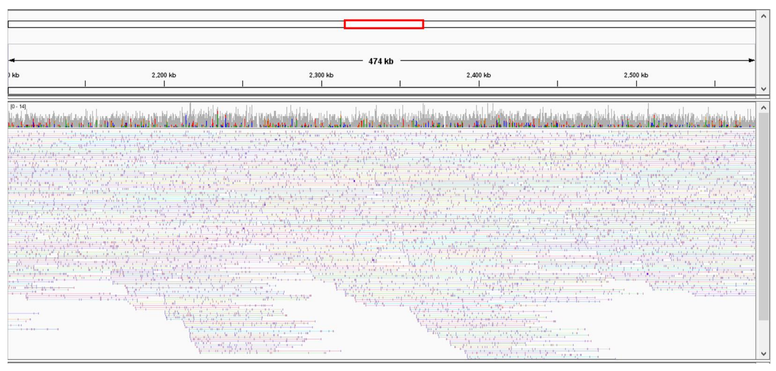


TELL-Seq WGS Library Prep kit includes

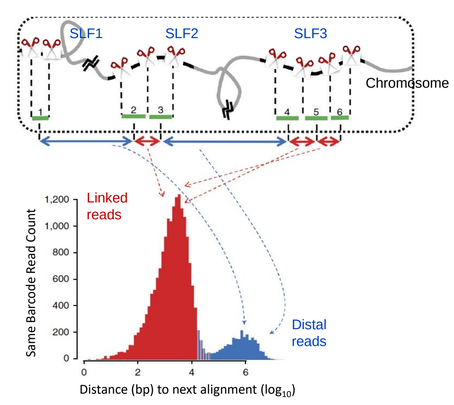
- TELL-Seq™ WGS Library Reagent Box 1
- TELL-Seq™ WGS Library Reagent Box 2
- TELL-Seq™ Library Index Primer Kit
- TELL-Seq™ Illumina Sequencing Primer Kit

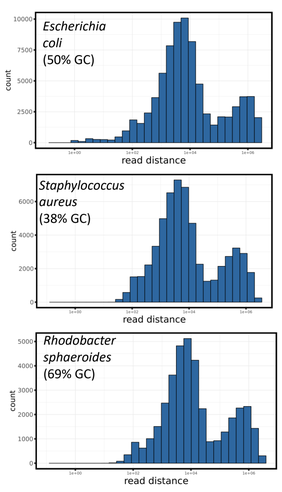
RESULTS

I. IGV Snapshot of TELL-Seq Linked Reads



II. TELL-Seq Results for Three Bacteria with Different GC Content





III. TELL-Seq Results Compared with Mate Pair Data

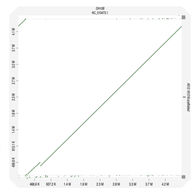
E. coli DH10B	Genomic coverage (%)	Average gap size (bp)	Number of gaps	Total gap length (bp)	Full genes†	Sample Processing
Illumina mate pair data 1*	96.72	3,576	43	153,750	4,181	Input DNA: 1-4μg Lib.Prep Time: 2 days
Illumina mate pair data 2*	96.59	3,191	50	159,572	4,164	
TELL-Seq data 1: 2x71bp	96.23	3,213	55	176,739	4158	Input DNA: 1ng Lib.Prep Time: 3 hrs
TELL-Seq data 2: 2x148bp	96.36	2,889	59	170,500	4166	

† The number of genes covered at 100%
* https://www.illumina.com/documents/products/appnotes/appnote-nextera-mate-pair-bacteria.pdf

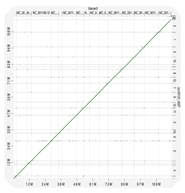
RESULTS

IV. *De Novo* Assembly Results Compared with Reference

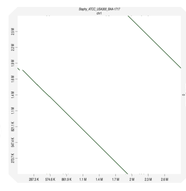
E. coli DH10B (4.6 Mb, 1 circular chromosome)



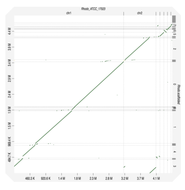
S. cerevisiae (12.2 Mb, 16 linear chromosomes)



S. aureus (2.9 Mb, 1 circular chromosome)

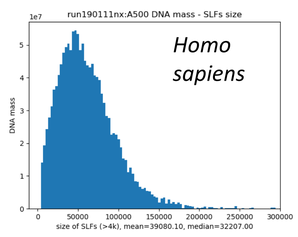


R. sphaeroides (4.6 Mb, 2 circular chromosomes, 5 plasmids)



V. Whole Genome Haplotype Phasing Results for NA12878

Method	TELL-Seq
Depth of genome coverage (Unique)	26X
Largest Phased Block	27.6 Mb
N50 Phased Block	4.6 Mb
Switch Error Rate	0.3%



CONCLUSION

TELL-Seq library technology enables short read 2nd generation sequencing systems to generate long read like sequencing data with innovative molecular barcoding library method. It is

- used for various sized genomes
- a single tube reaction without specialized instrument
- highly scalable for doing multiple samples in parallel
- one-stop solution for *de novo* sequencing and whole genome phasing
- automation friendly

TELL-Seq WGS library prep kit is available for early access soon. Visit www.universalsequencing.com for more information.

UNIVERSAL SEQUENCING
innovation for all

TELL-Seq™: Tell you the true genome today!