The Sequel II System

An introduction to Single Molecule Real-Time (SMRT) Sequencing and its applications

Deborah Moine – Scientist II, Field Applications Support

- SMRT Sequencing & Sequel II System
- Applications :
  - *de novo* assembly
  - Variant detection
  - Iso-Seq full-length cDNA sequencing
  - Microbe characterization
- Sequel II System updates

SMRT Sequencing Advantages

Long Reads
- Tens of kilobases
- Sequence from 500 bp to >50,000 inserts

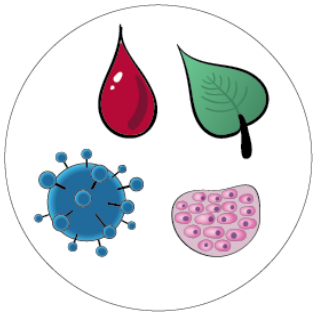
High Accuracy
- Free of systematic errors
- Achieves >99.999% (Q50)

Single-Molecule Resolution
- Sequence DNA or RNA
- Long reads with ≥Q20 (99%) single-molecule accuracy

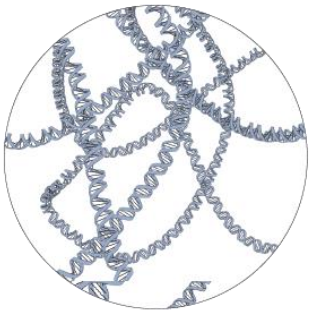
Uniform Coverage
- No DNA amplification
- Least GC content and sequence complexity bias

Simultaneous Epigenetic Detection
- Characterize epigenome
- No separate sample preparation required

FROM SAMPLE TO SMRT SEQUENCING

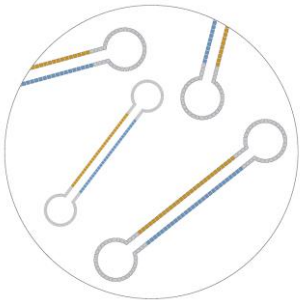


From viruses to vertebrates



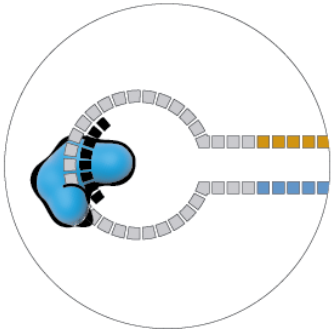
Isolate DNA or RNA

Ligate adapters +



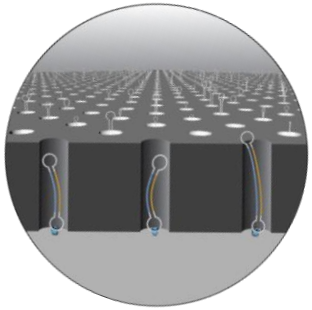
Generate SMRTbell libraries

\+ Primer & Polymerase



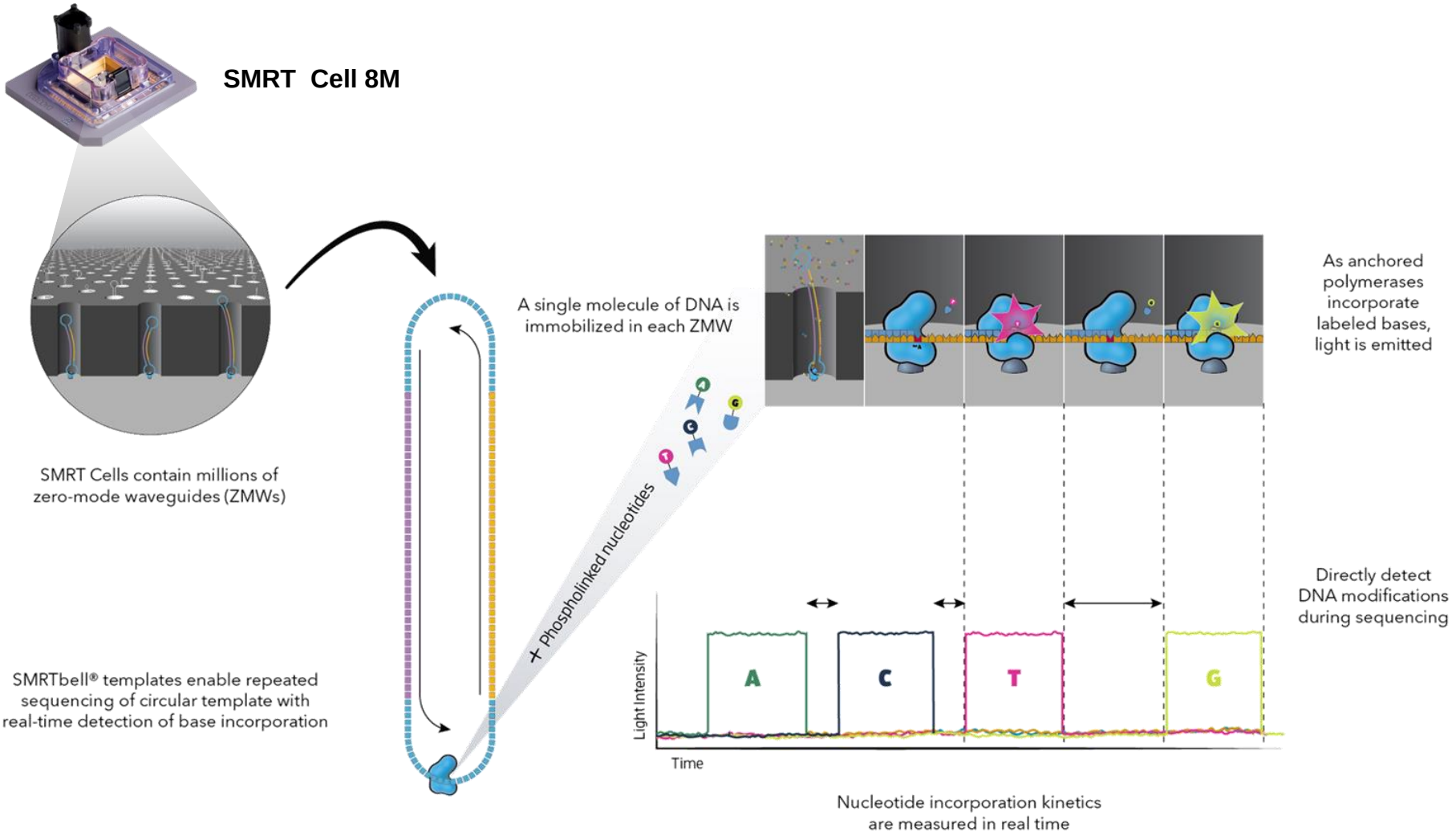
Prepare sequencing reaction

Use PacBio Sequel Systems to sequence genomes, transcriptomes, and epigenomes



SMRT Cells contain millions of zero-mode waveguides (ZMWs)

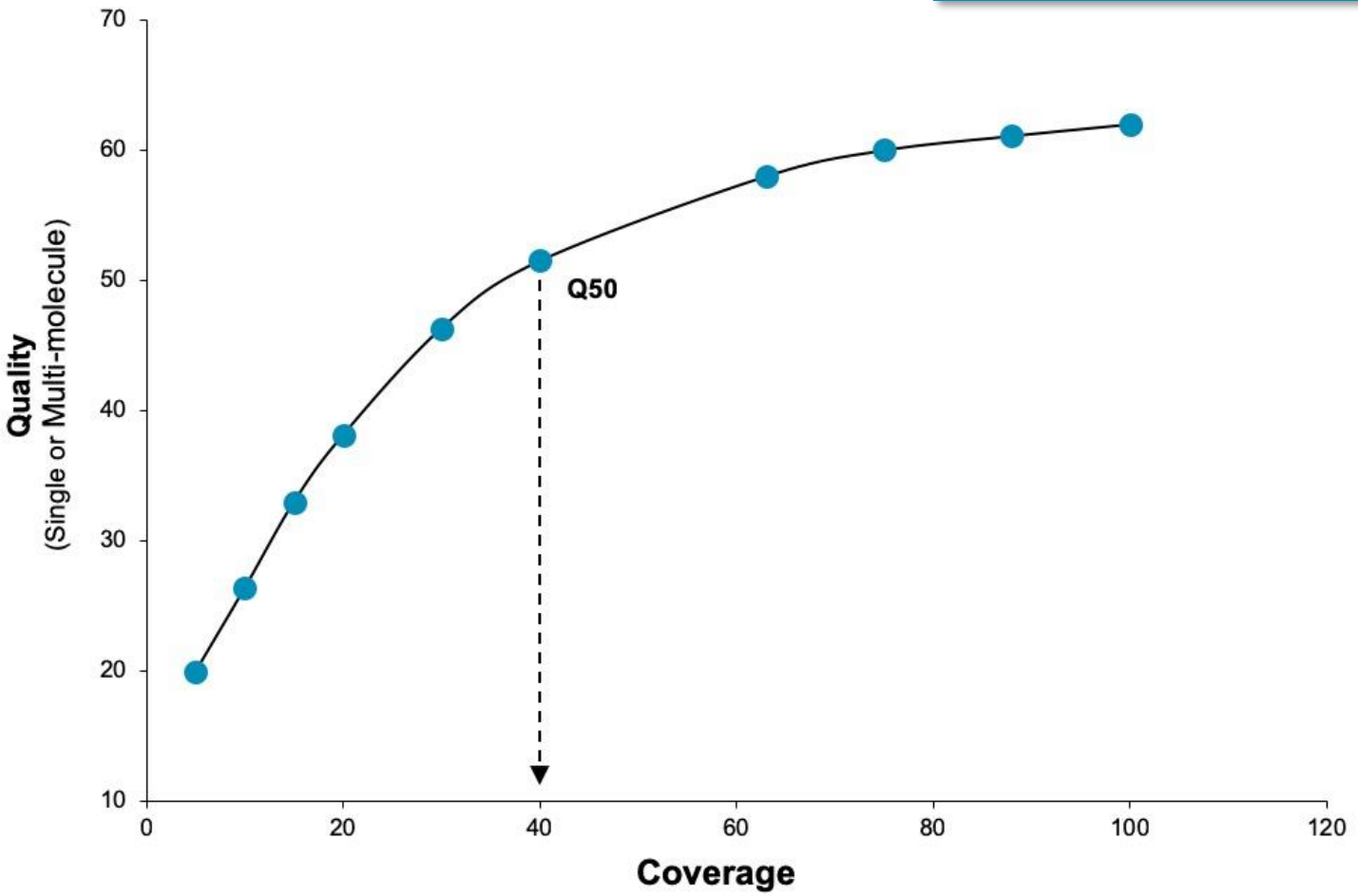
SINGLE MOLECULE, REAL-TIME (SMRT) SEQUENCING



SMRT Sequencing enables simultaneous collection of data from millions of wells using the natural process of DNA replication to sequencing long fragments of native DNA.

HIGH CONSENSUS ACCURACY

Achieves >99.999% (Q50)

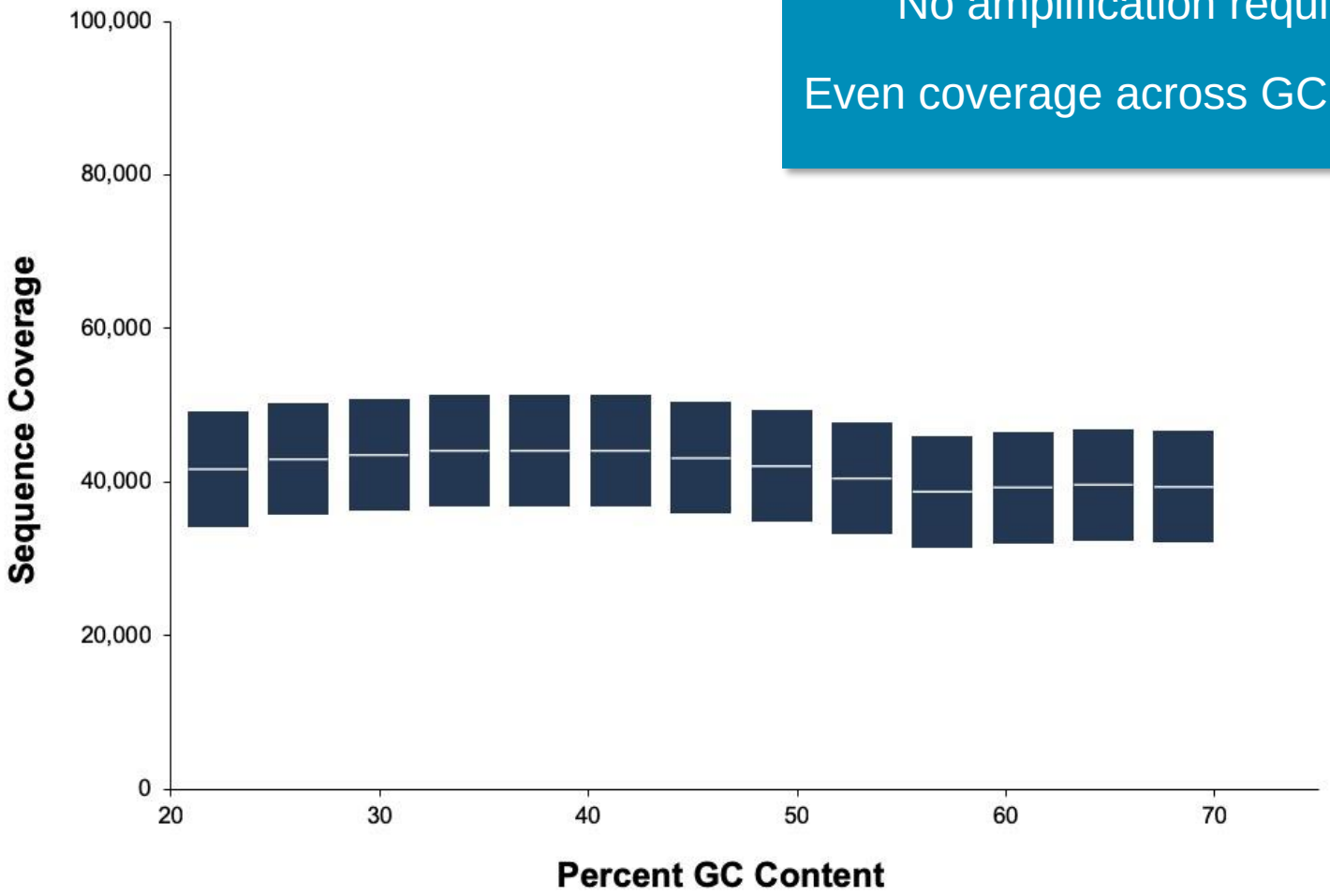


Consensus accuracy is a function of coverage and chemistry. The data above is based on a bacterial genome run on the Sequel II System (1.0 Chemistry, Sequel II System Software v7.0). Single-molecule accuracy has similar coverage requirements.

UNIFORM COVERAGE

No amplification required

Even coverage across GC content



Mean coverage per GC window across a human sample. Data generated with a 35 kb human library on a Sequel II System using 1.0 Chemistry and Sequel II System Software v7.0

PACBIO PRODUCT RELEASES OVER THE LAST EIGHT YEARS

May 2012
v1.3.1
- Epigenetic Modification

Nov 2012
v1.3.3
- XL Chemistry
- Stage Start

Aug 2012
v1.3.2
- MagBead Release

Feb 2012
v1.3.0
- C2 Launch

Jan 2013
v1.4
- HGAP/Quiver
- SMRT Cells v3

Oct 2013
v2.1
- P5-C3 release
- HGAP 2.0

Apr 2013
RS II Product Release (v2.0)
- 75K to 150K ZMW
- 2x Throughput
- P4 release

Mar 2014
v2.2
- Iso-Seq Method
- Long Amplicon Analysis

Oct 2014
v2.3
- P6-C4 release

Apr 2015
Barcode Kits

Oct 2015
Sequel System

Aug 2016
Sequel v3.1
- 1.2.1 Chemistry

Jan 2017
Sequel v4.0
- 2.0 Chemistry

Sept 2017
Sequel v5.0
- 2.1 Chemistry

Feb 2018
Sequel v5.1
- 2.1 Polymerase

Oct 2018
Sequel v6.0
- **3.0 Chemistry**
- **SMRT Cell 1M v3**

Apr 2019
Sequel II v7.0
- **1.0 Chemistry**
- **SMRT Cell 8M**

Increased throughput by over 10,000-fold

PacBio RS II System

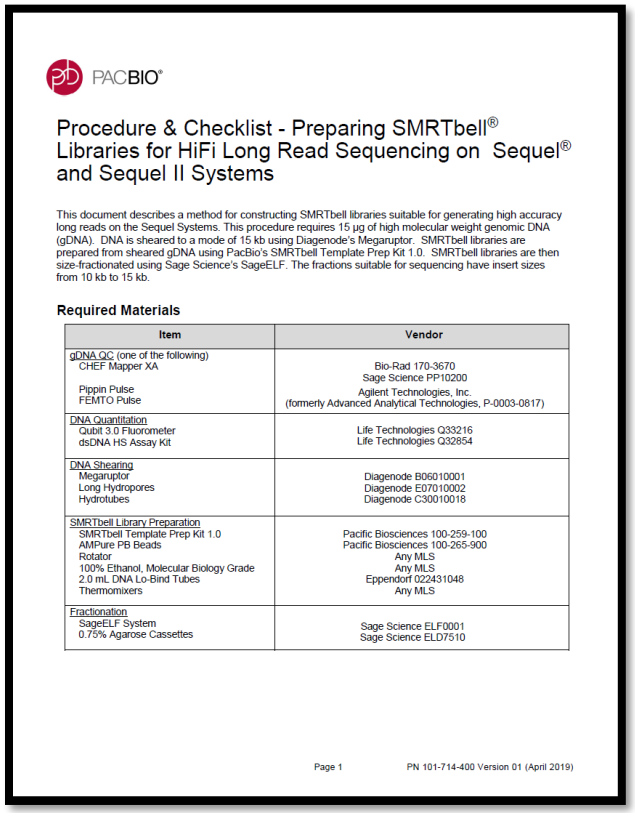
Sequel System

Sequel II System: Up to 160 Gb/Cell (gDNA)

SMRTBELL LIBRARY CONSTRUCTION

Existing Sequel library construction procedures are compatible with Sequel II System

- SMRTbell Template Prep Kit 1.0
- SMRTbell Express Template Prep Kit 2.0
  - Large insert gDNA library (≥10 kb)
  - Microbial multiplexing (10 kb)
  - Iso-Seq Express Method
  - Multiplexed amplicon sequencing
  - Low input DNA sequencing (150ng for a 300 Mb genome)
- Procedure & Checklist – Preparing SMRTbell Libraries for HiFi Long Read Sequencing



PACBIO®

Procedure & Checklist - Preparing SMRTbell® Libraries for HiFi Long Read Sequencing on Sequel® and Sequel II Systems

This document describes a method for constructing SMRTbell libraries suitable for generating high accuracy long reads on the Sequel Systems. This procedure requires 15 µg of high molecular weight genomic DNA (gDNA). DNA is sheared to a mode of 15 kb using Diagenode's Megaruptor. SMRTbell libraries are prepared from sheared gDNA using PacBio's SMRTbell Template Prep Kit 1.0. SMRTbell libraries are then size-fractionated using Sage Science's SageELF. The fractions suitable for sequencing have insert sizes from 10 kb to 15 kb.

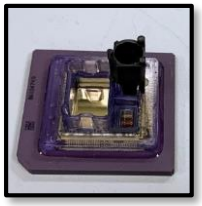
Required Materials

Item	Vendor
gDNA QC (one of the following) CHEF Mapper XA Pippin Pulse FEMTO Pulse	 Bio-Rad 170-3670 Sage Science PP10200 Agilent Technologies, Inc. (formerly Advanced Analytical Technologies, P-0003-0817)
DNA Quantitation Qubit 3.0 Fluorometer dsDNA HS Assay Kit	 Life Technologies Q33216 Life Technologies Q32854
DNA Shearing Megaruptor Long Hydropores Hydrotubes	 Diagenode B06010001 Diagenode E07010002 Diagenode C30010018
SMRTbell Library Preparation SMRTbell Template Prep Kit 1.0 AMPure PB Beads Rotator 100% Ethanol, Molecular Biology Grade 2.0 mL DNA Lo-Bind Tubes Thermomixers	 Pacific Biosciences 100-259-100 Pacific Biosciences 100-265-900 Any MLS Any MLS Eppendorf 022431048 Any MLS
Fractionation SageELF System 0.75% Agarose Cassettes	 Sage Science ELF0001 Sage Science ELD7510

Page 1 PN 101-714-400 Version 01 (April 2019)

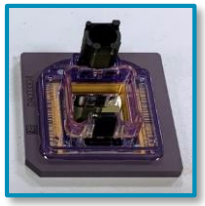
SEQUEL II SYSTEM – SEQUENCE WITH CONFIDENCE

Sequel System



1 million ZMWs
SMRT Cell 1M

Sequel II System



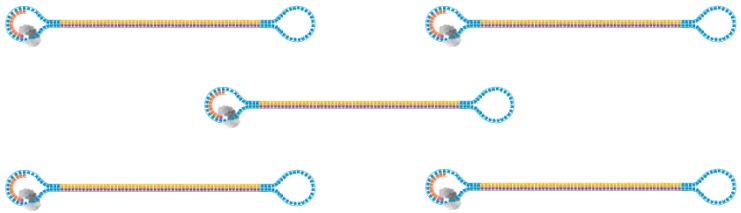
8 million ZMWs
SMRT Cell 8M

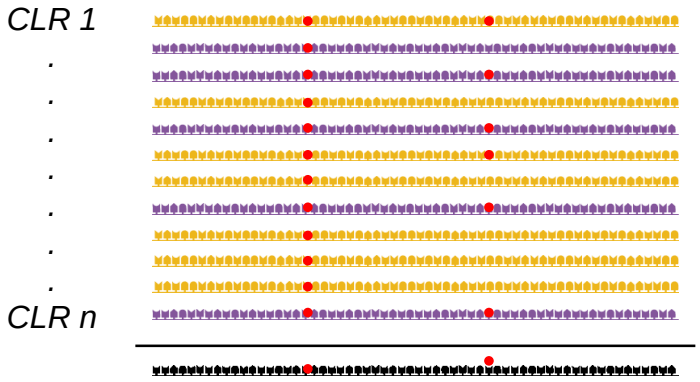
- ~8-fold increase in data yield
- Reduced project time
- Lower project cost
- Equivalent performance

Increased Throughput Capacity With the Sequel II System

TWO MODES OF SMRT SEQUENCING

- Continuous *Long Read* Sequencing (CLR)





Consensus sequence

Generate reads in ten's of kilobases

- Circular Consensus Sequencing (CCS)

HiFi Read

Generate high quality reads 1 – 20 kb

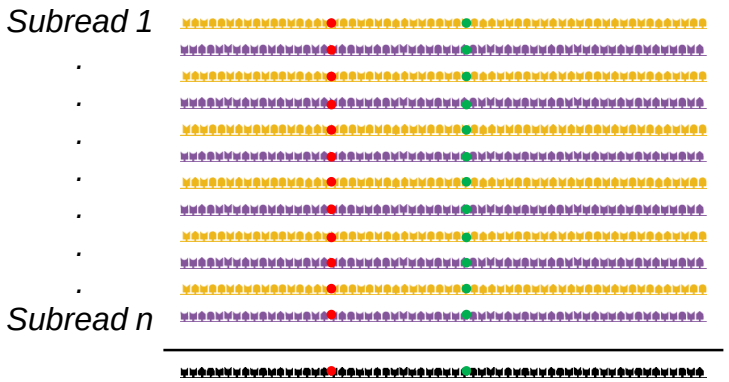
TWO MODES OF SMRT SEQUENCING

- Continuous *Long Read* Sequencing (CLR)

Consensus sequence

Generate reads in ten's of kilobases

- Circular Consensus Sequencing (CCS)

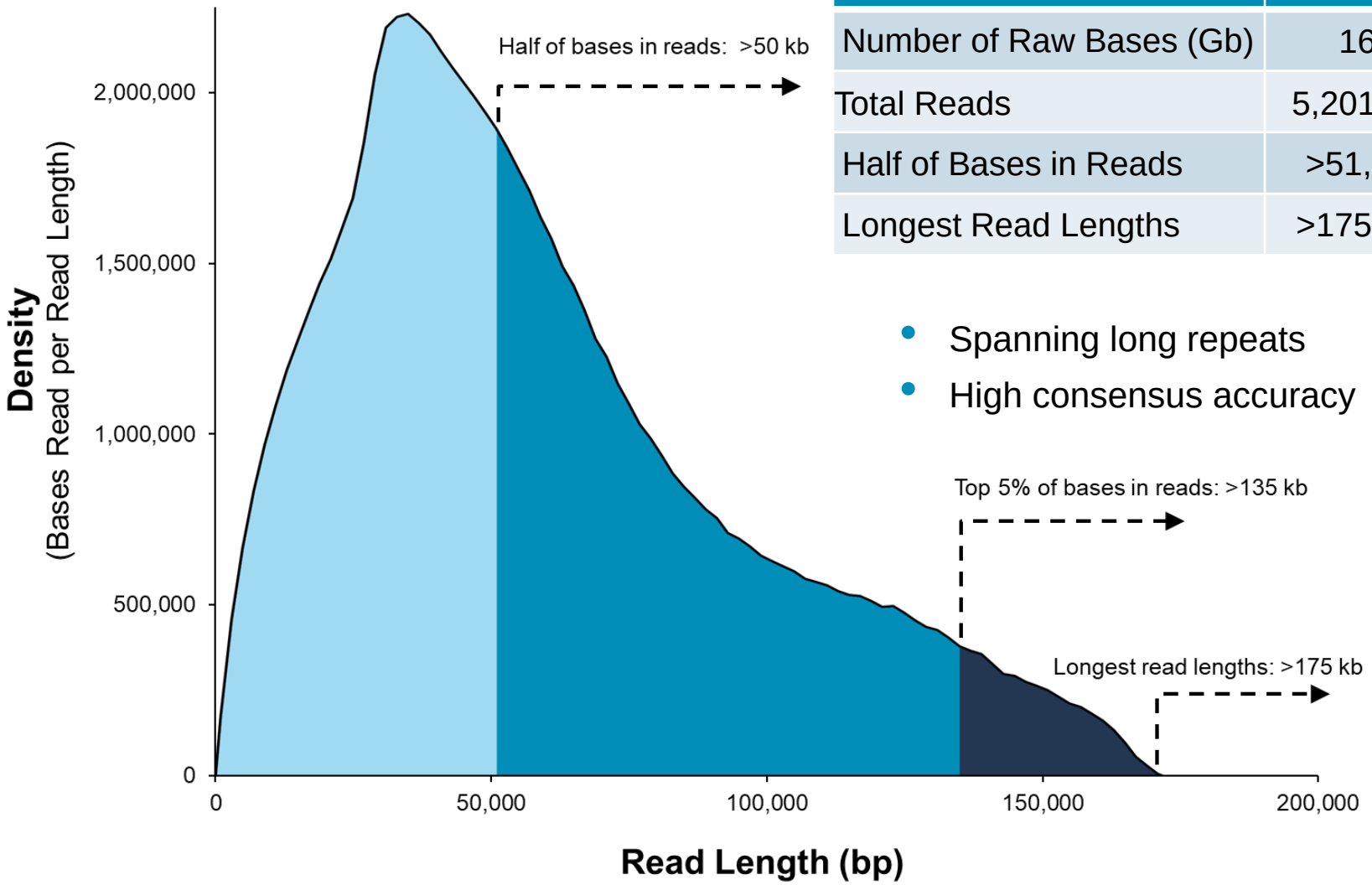


HiFi Read

Generate high quality reads 1 – 20 kb

LONG READ SEQUENCING

METRIC	
Number of Raw Bases (Gb)	166
Total Reads	5,201,973
Half of Bases in Reads	>51,863
Longest Read Lengths	>175,000

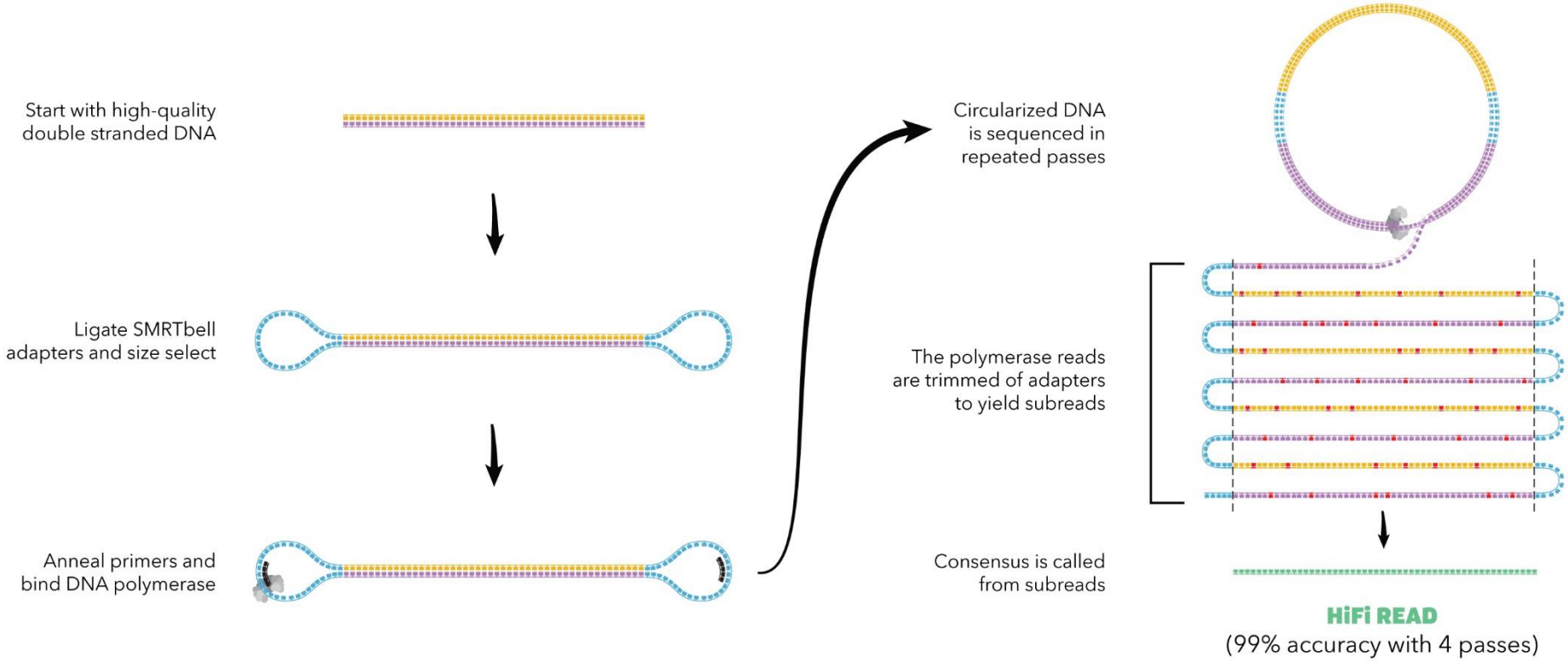


- Spanning long repeats
- High consensus accuracy

Data shown above from a 35 kb size-selected *E. coli* library using the SMRTbell Template Prep Kit on a Sequel II System (1.0 Chemistry, Sequel II System Software v7.0, 15-hour movie). Read lengths, reads/data per SMRT Cell 8M and other sequencing performance results vary based on sample quality/type and insert size.

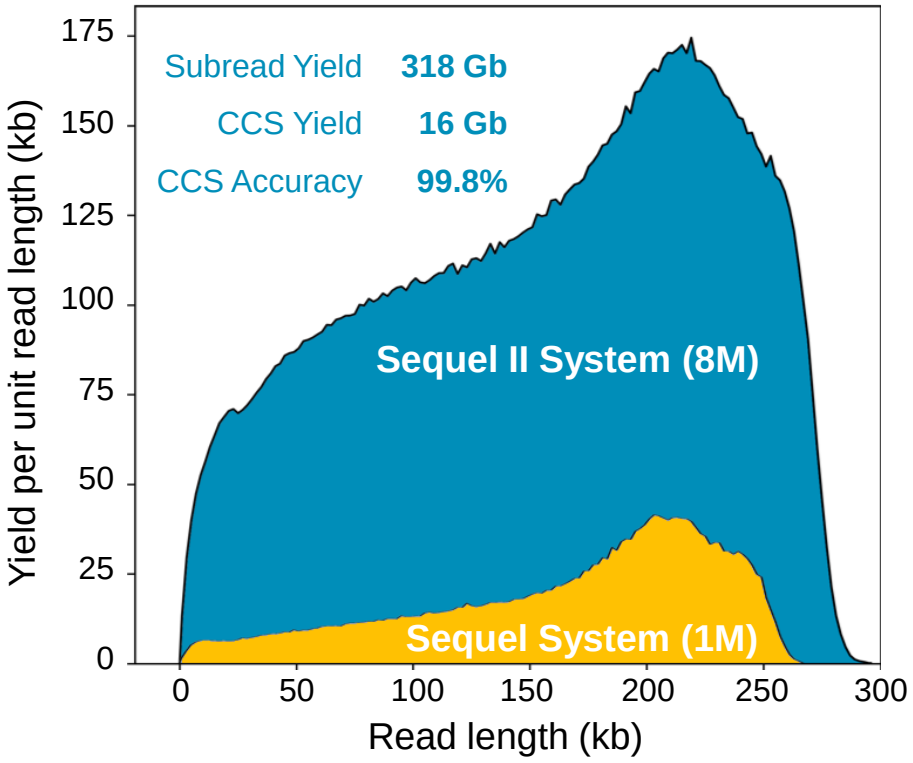
HiFi READS

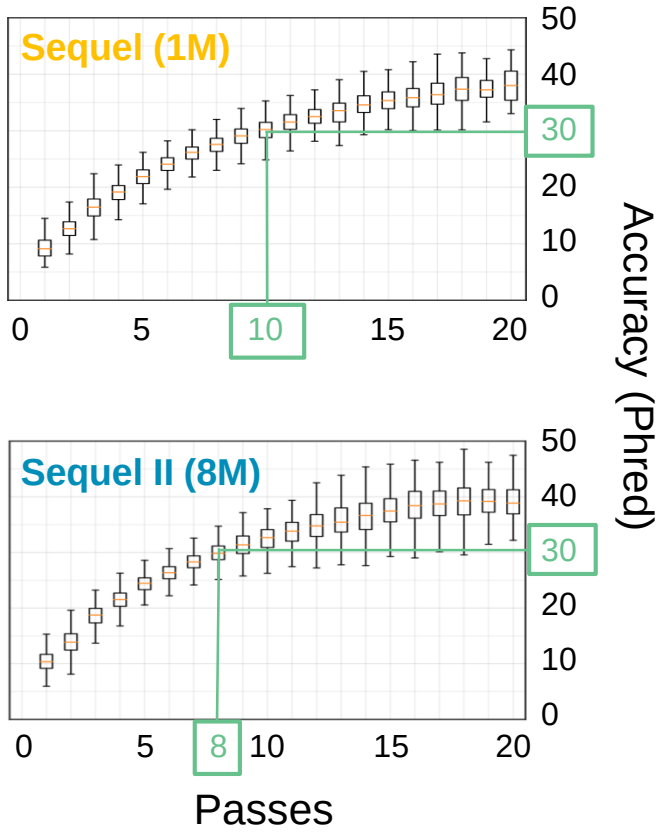
CIRCULAR CONSENSUS SEQUENCING: HIFI READ GENERATION



HIFI READ PERFORMANCE

- Increased Yield and Accuracy

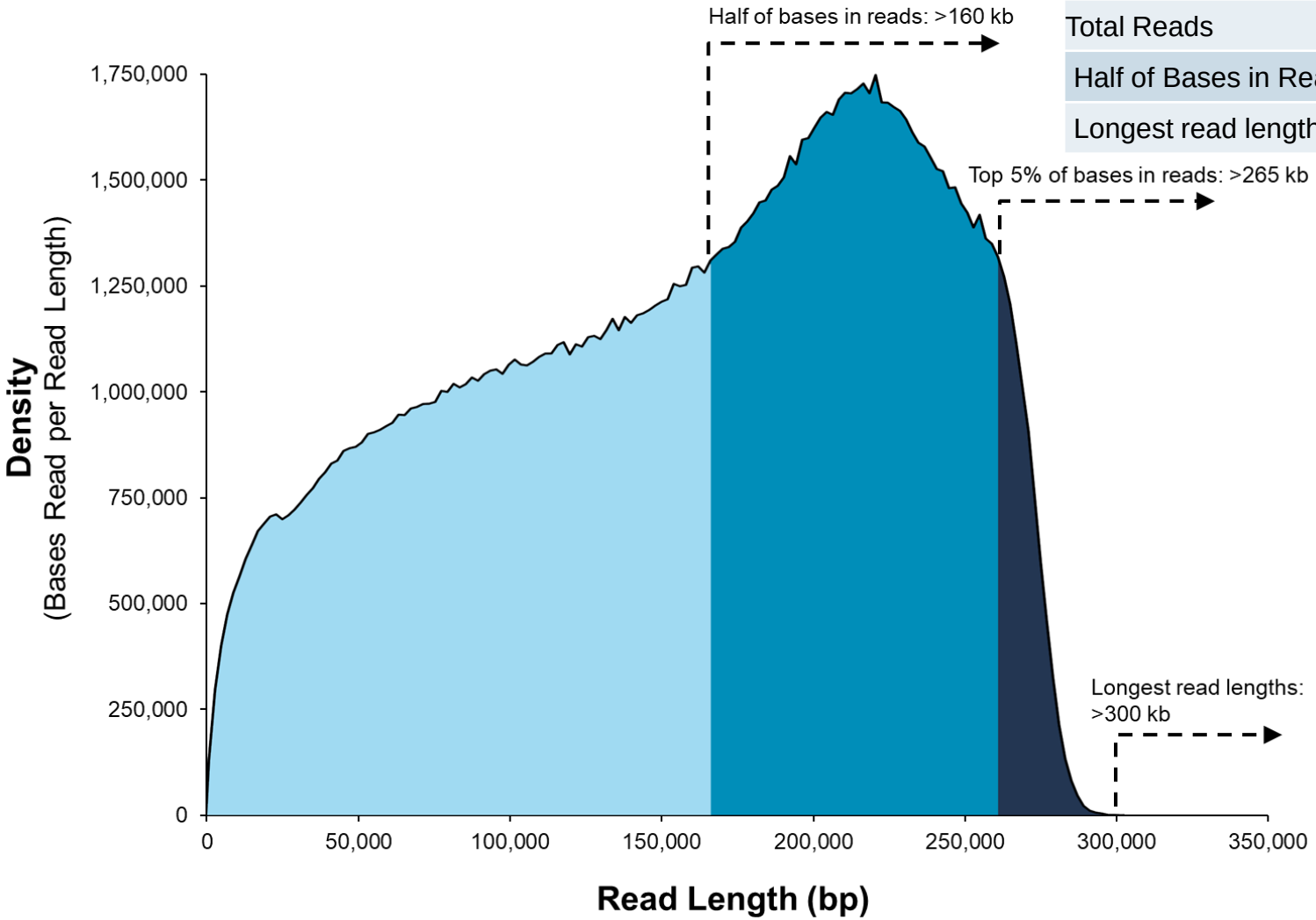




Data shown above from a 12 kb size-selected human library using the SMRTbell Template Prep Kit on a Sequel II System (1.0 Chemistry, Sequel II System Software v7.0, 30-hour movie). Read lengths, reads/data per SMRT Cell 8M and other sequencing performance results vary based on sample quality/type and insert size.

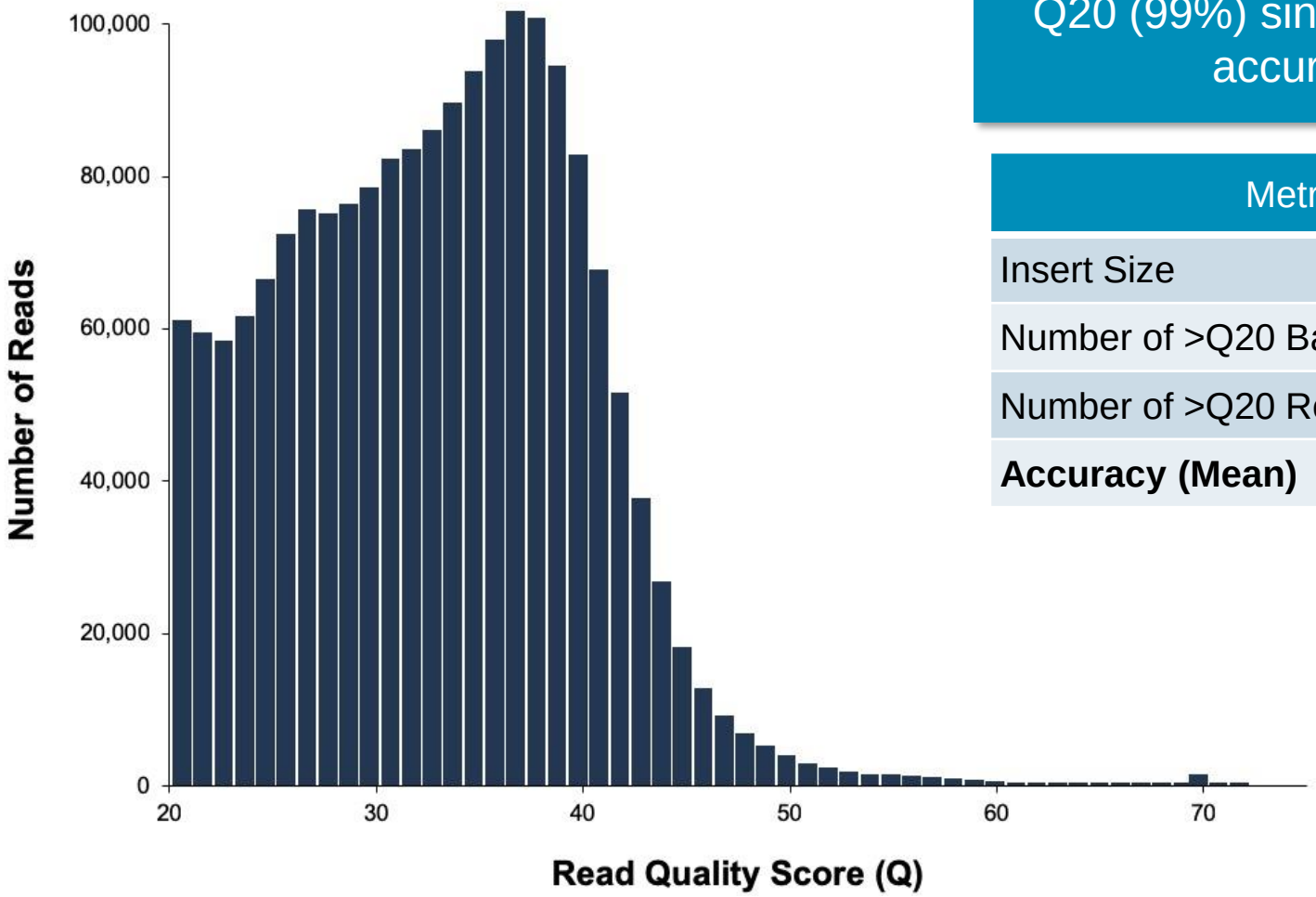
HIFI SEQUENCING PERFORMANCE

METRIC	
Number of Raw Bases (Gb)	320
Total Reads	4,053,000
Half of Bases in Reads	>166,571
Longest read lengths	>300,000



Data shown above from a 11 kb size-selected human library using the SMRTbell Template Prep Kit on a Sequel II System (1.0 Chemistry, Sequel II System Software v7.0, 30-hour movie). Read lengths, reads/data per SMRT Cell 8M and other sequencing performance results vary based on sample quality/type and insert size.

HIFI SEQUENCING PERFORMANCE



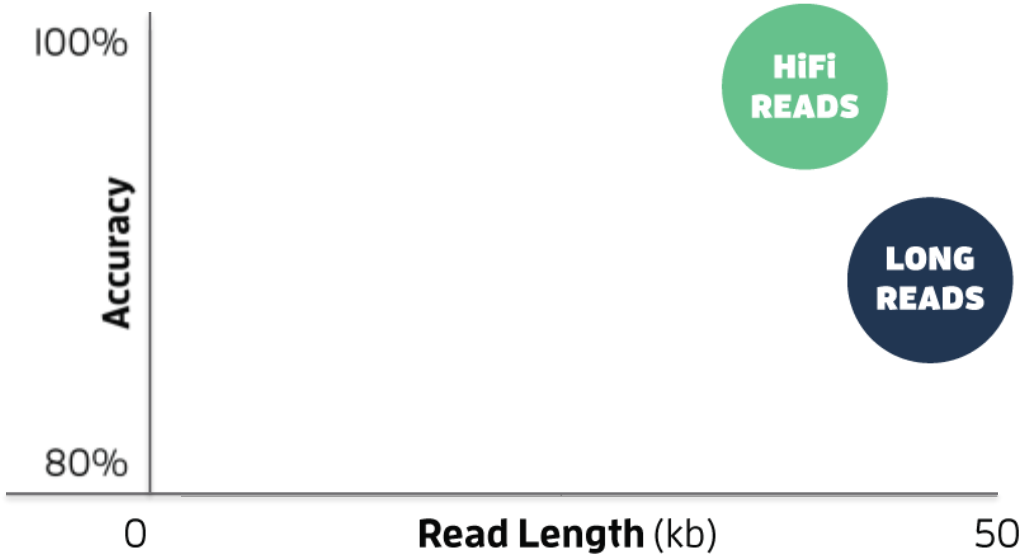
Q20 (99%) single-molecule accuracy

Metrics	
Insert Size	12 kb
Number of >Q20 Bases	21 Gb
Number of >Q20 Reads	1,855,642
Accuracy (Mean)	**99.8%**

Data shown above from a 11 kb size-selected human library using the SMRTbell Template Prep Kit on a Sequel II System (1.0 Chemistry, Sequel II System Software v7.0, 30-hour movie). Read lengths, reads/data per SMRT Cell 8M and other sequencing performance results vary based on sample quality/type and insert size.

THE SEQUEL II SYSTEM – SEQUENCE WITH CONFIDENCE

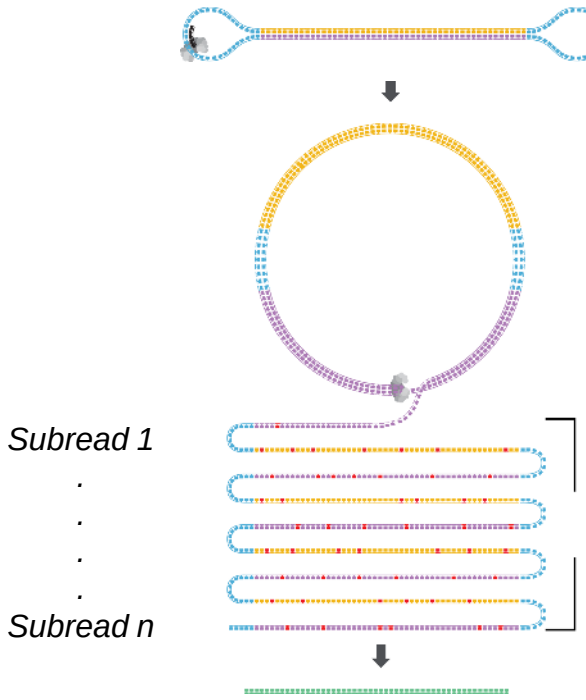
- ✓ Combines high accuracy and long read lengths
- ✓ Two sequencing modes for flexibility in your workflows



TWO MODES OF SMRT SEQUENCING

Circular Consensus Sequencing (CCS) Mode

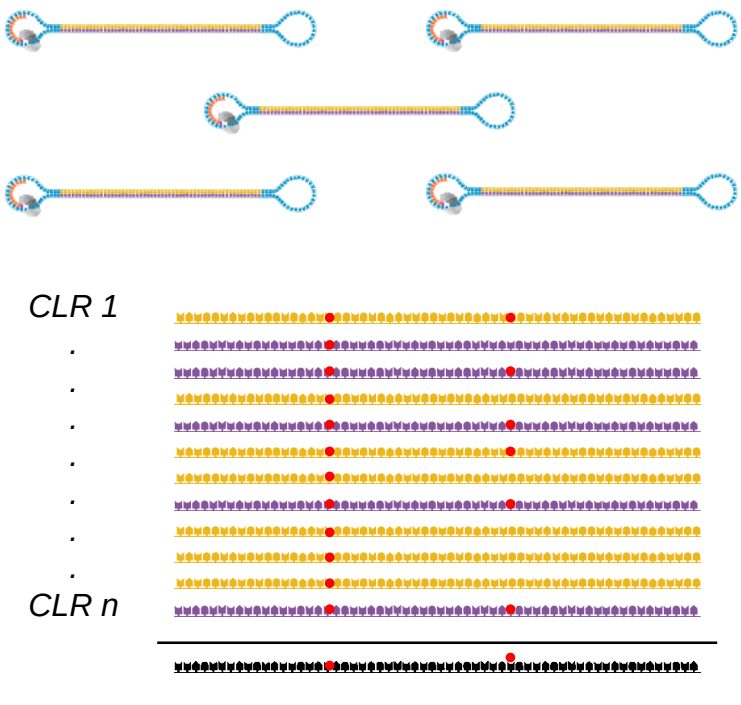
Inserts 10-20 kb



Single-molecule consensus sequence

Continuous Long Read (CLR) Sequencing Mode

Inserts >25 kb, up to 175 kb



Multi-molecule consensus sequence

WHOLE GENOME SEQUENCING (WGS) APPLICATIONS

HiFi READS

- Comprehensive detection of variants (SNVs, SVs, CNVs)
- High-quality, phased genome assembly

VARIANT DETECTION

ASSEMBLY

LONG READS

- Detection of structural variants (SVs)
- Assembly of very large genomes

STRUCTURAL VARIANT DETECTION

ASSEMBLY

WHAT CAN YOU DO WITH ONE SMRT CELL 8M?

HiFi Reads	Long Reads
Detect all variants – SNVs, indels, SVs, and CNVs – in a human genome (2 SMRT Cells 8M)	Detect structural and copy number variants in two human genomes
Assemble up to a 2 Gb genome	Assemble up to a 3 Gb genome

www.pacb.com/onesmrtcell

**Read lengths, reads/data per SMRT Cell 8M, and other sequencing performance results vary based on sample quality/type and insert size. Prices, listed in USD, are approximate and may vary by region. Pricing includes library and sequencing reagents run on a Sequel II System and does not include instrument amortization or other reagents.*

De novo Assembly

Using Whole Genome Sequencing on the Sequel II System

ANSWERING: WHAT'S THE SAME AND WHAT'S DIFFERENT?

High-quality genome assemblies are the most comprehensive way to assess genomic diversity among and between species.

Crop Improvements

Population or Disease-specific Studies

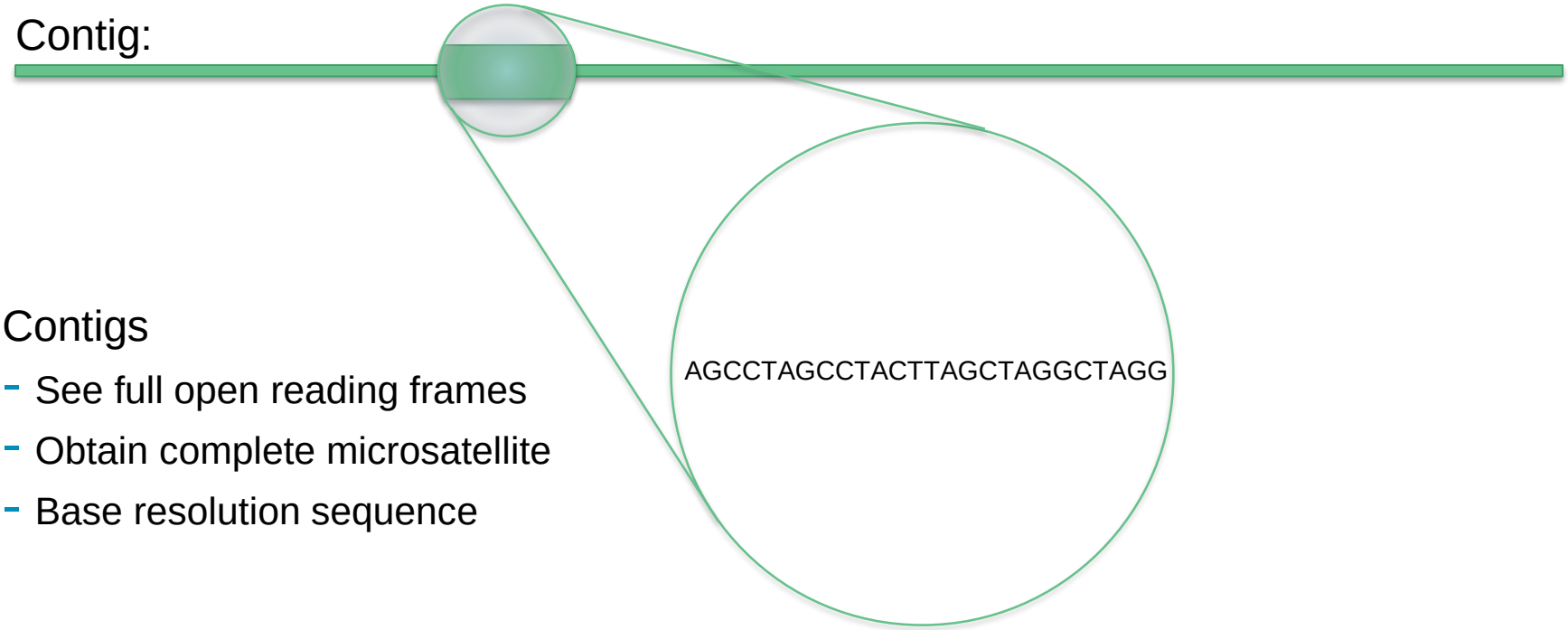
Animal Health & Breeding

Conservation

Disease Prevention

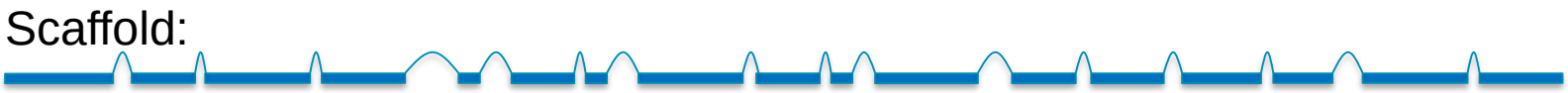
DRAFT VS HIGH-QUALITY GENOME ASSEMBLIES

You see more, and therefore have more opportunities to see differences, with a high-quality genome assembly.



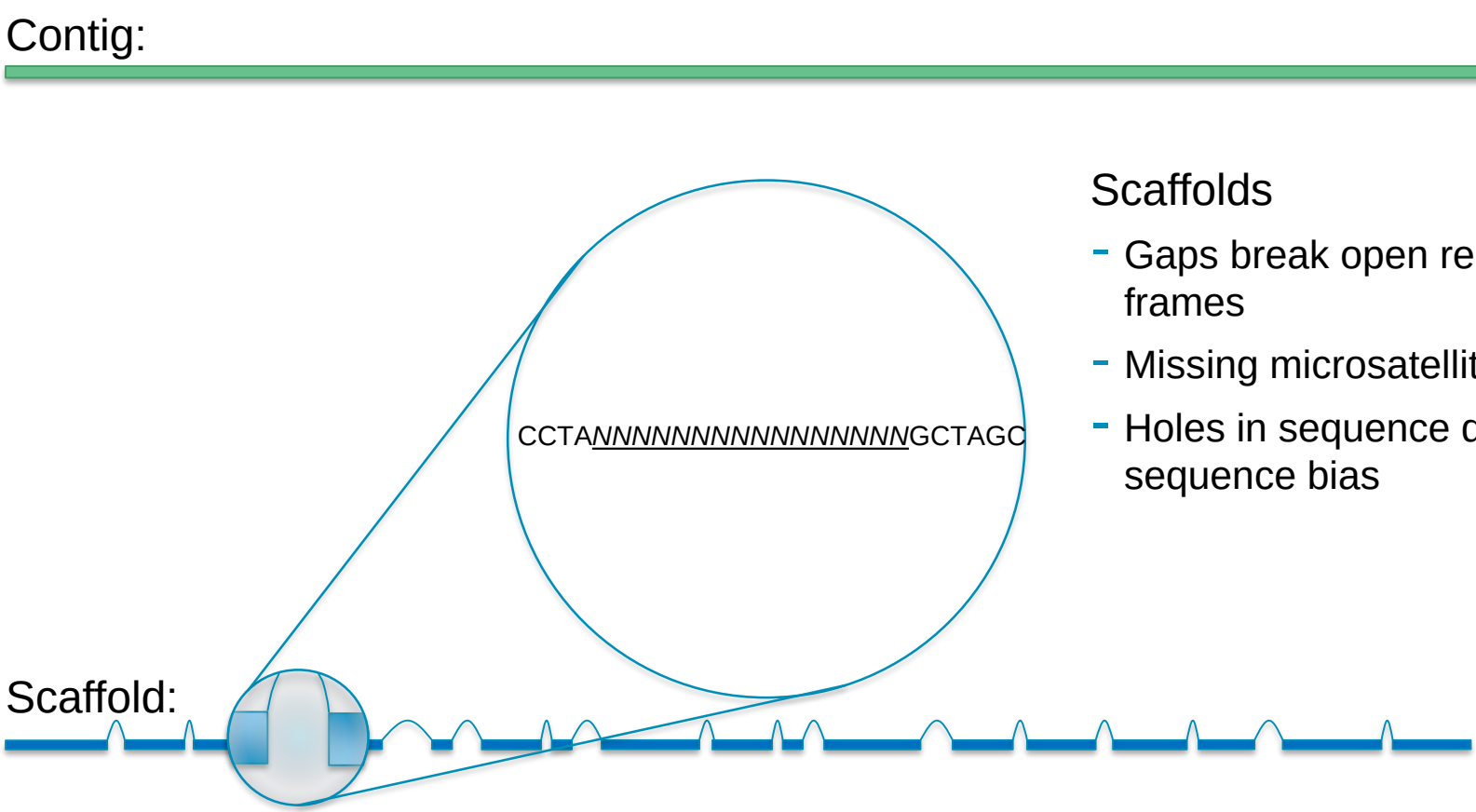
Contigs

- See full open reading frames
- Obtain complete microsatellite
- Base resolution sequence



DRAFT VS HIGH-QUALITY GENOME ASSEMBLIES

You see more, and therefore have more opportunities to see differences, with a high-quality genome assembly.



Scaffolds

- Gaps break open reading frames
- Missing microsatellites
- Holes in sequence due to sequence bias

UNIQUE CHALLENGES OF GENOME ASSEMBLY

Size and Complexity

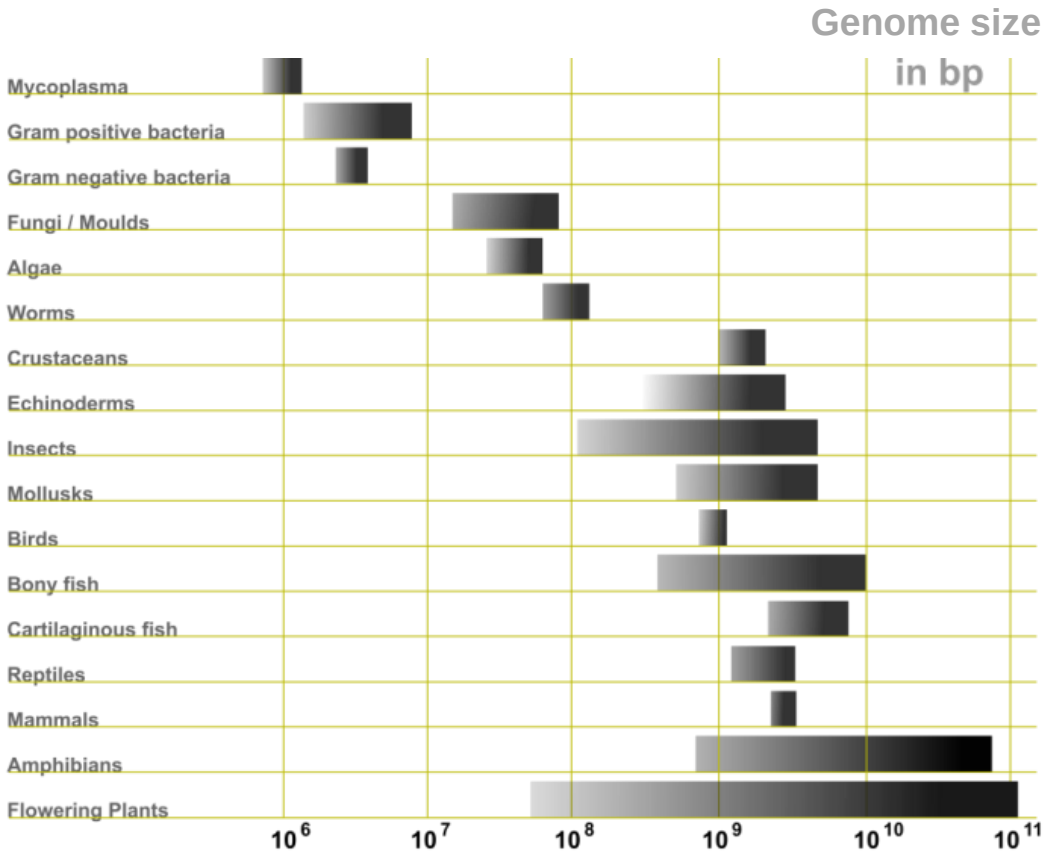
- Human genome has over 3 billion base pairs
- Plants often have even larger genomes

Extreme Repeat Content

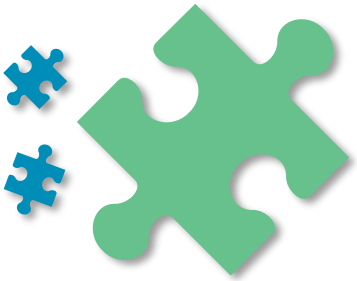
- Maize >60%
- Wheat >80%

Each Project is Unique

- Ranges in size, ploidy, and repeat content
- Custom strategy is commonly needed

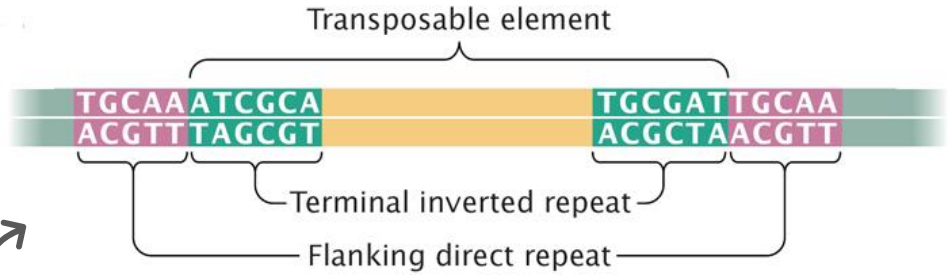


WHY DO LONG READS MATTER?



Bigger is better – when it comes to easy assembly

Phase haplotypes in outbred diploids and polyploids



Long, accurate reads span repetitive elements and allow assembly of repeats that are very similar

Even really big genomes

CHOOSE THE SEQUENCING MODE THAT'S RIGHT FOR YOUR PROJECT

- Highly accurate long reads with minimum accuracy of Q20 (99%)
- Small file sizes and fast analysis time
- Assemble up to a 2 Gb genome in a single SMRT Cell 8M*
- Run up to 200 samples (2 Gb) per year, per system*

- Longest reads, with half of data >50 kb and maximum read lengths up to 175 kb
- Short sequencing run times
- Assemble up to a 3 Gb genome in a single SMRT Cell 8M*
- Run up to 400 samples (3 Gb) per year, per system*

*Read lengths, reads/data per SMRT Cell 8M, and other sequencing performance results vary based on sample quality/type and insert size. Prices, listed in USD, are approximate and may vary by region. Pricing includes library and sequencing reagents run on a Sequel II System and does not include instrument amortization or other reagents.

WORKFLOWS: FROM DNA TO COMPREHENSIVE GENOME ASSEMBLIES

HiFi Reads

>15 µg of unamplified genomic DNA

Express Template Prep Kit 2.0

Size select 15-20 kb inserts

Use Circular Consensus Sequencing (CCS) mode

Generate HiFi reads with CCS analysis

Assemble and phase HiFi reads with FALCON and FALCON-Unzip3

LIBRARY PREP

SMRT SEQUENCING

DATA ANALYSIS

Long Reads

>5 µg of unamplified genomic DNA

Express Template Prep Kit 2.0

Size select >30 kb inserts

Use Continuous Long Read (CLR) sequencing mode

Assemble and phase long reads with PacBio analytical portfolio:

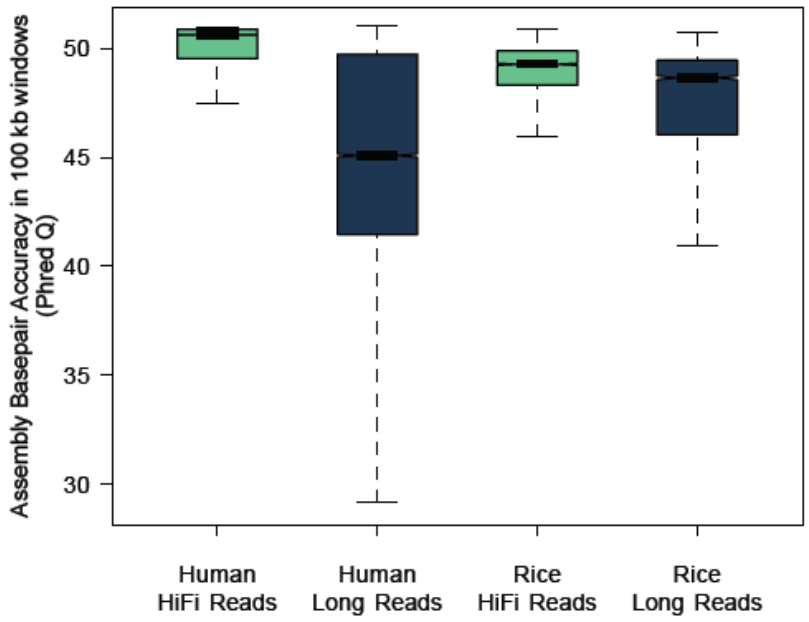
- -SMRT Analysis
- -FALCON and FALCON-Unzip
- -Bioinformatics service providers

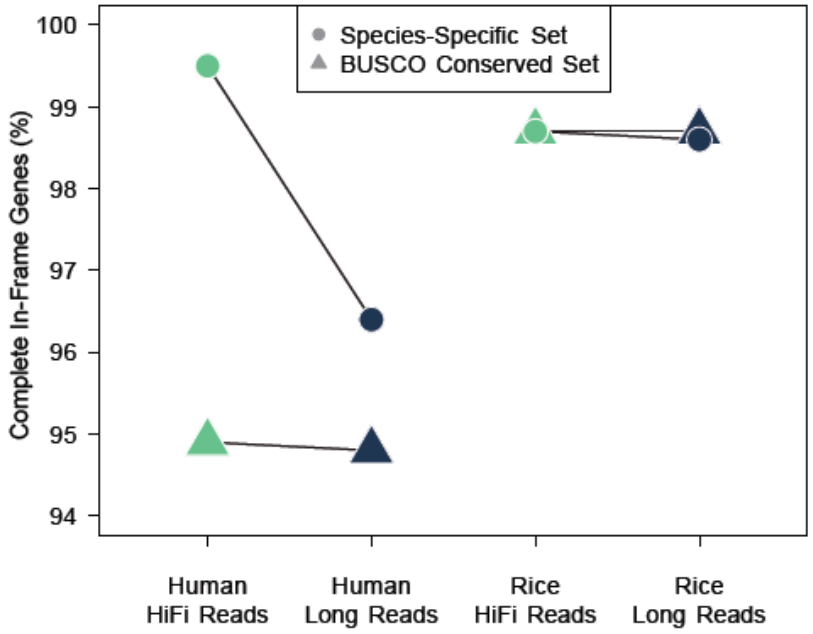
GENERATE CONTIGUOUS GENOME ASSEMBLIES

Dataset	Human		Rice	
Mode	Long Reads	**HiFi Reads**	Long Reads	**HiFi Reads**
Size Selection	BP >15 kb	**15 kb ELF**	BP >30 kb	**17 kb ELF**
Coverage	50-fold	**22-fold**	60-fold	**20-fold**
# of SMRT Cells 8M	2	**3**	1	**1**
Contig N50 (Mb)	12.6	**30.5**	11.2	**10.7**

Megabase size contig N50s

GENERATE COMPLETE AND ACCURATE GENOME ASSEMBLIES





Accuracies >Q40 (99.99%)

>94% of genes in frame

REDUCED ANALYSIS TIME WITH HIFI READS

HiFi READS

Compute times for *de novo* assembly of a human genome

Data Type		HiFi Reads	Long Reads
Input File Type		CCS.FASTQ.GZ	SUBREADS.BAM
Input File Size (GB)		48	323
Read Correction Method		CCS Analysis	Pre-assembly
Time to Results (Hours)	Read Correction	17.5	43.5
	Contig Assembly	13.7	18.9

Analyses run with PacBio recommended compute infrastructure

~31 hrs

~62 hrs

Fast analysis time

PACBIO®

Variant Detection

Using Whole Genome Sequencing on the Sequel II System

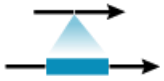
TYPES OF VARIANTS IN A GENOME

SMRT Sequencing provides comprehensive detection of all variant types.

Single Nucleotide Variant

Deletion

Insertion

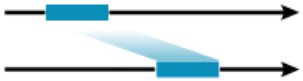


Tandem Duplication

Interspersed Duplication

Inversion

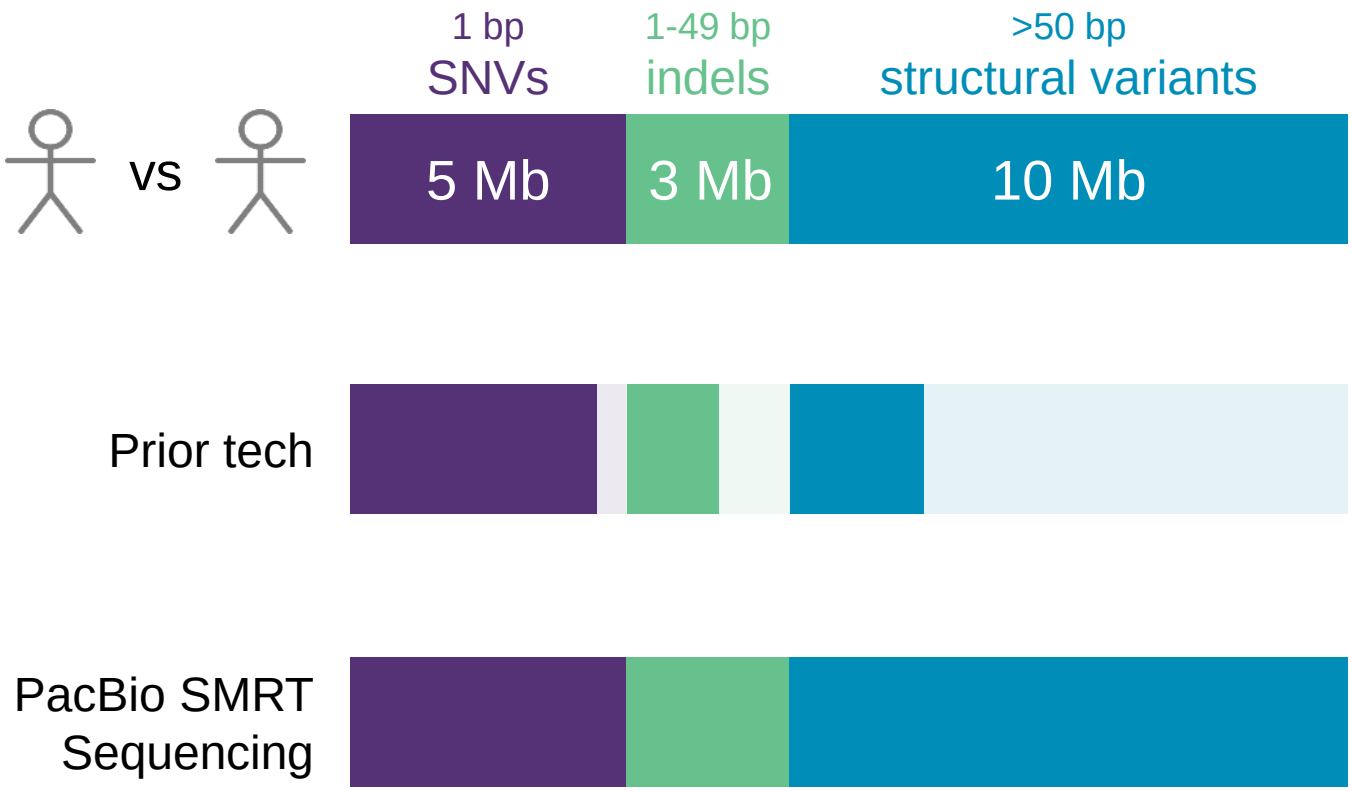
Translocation



Copy Number Variant

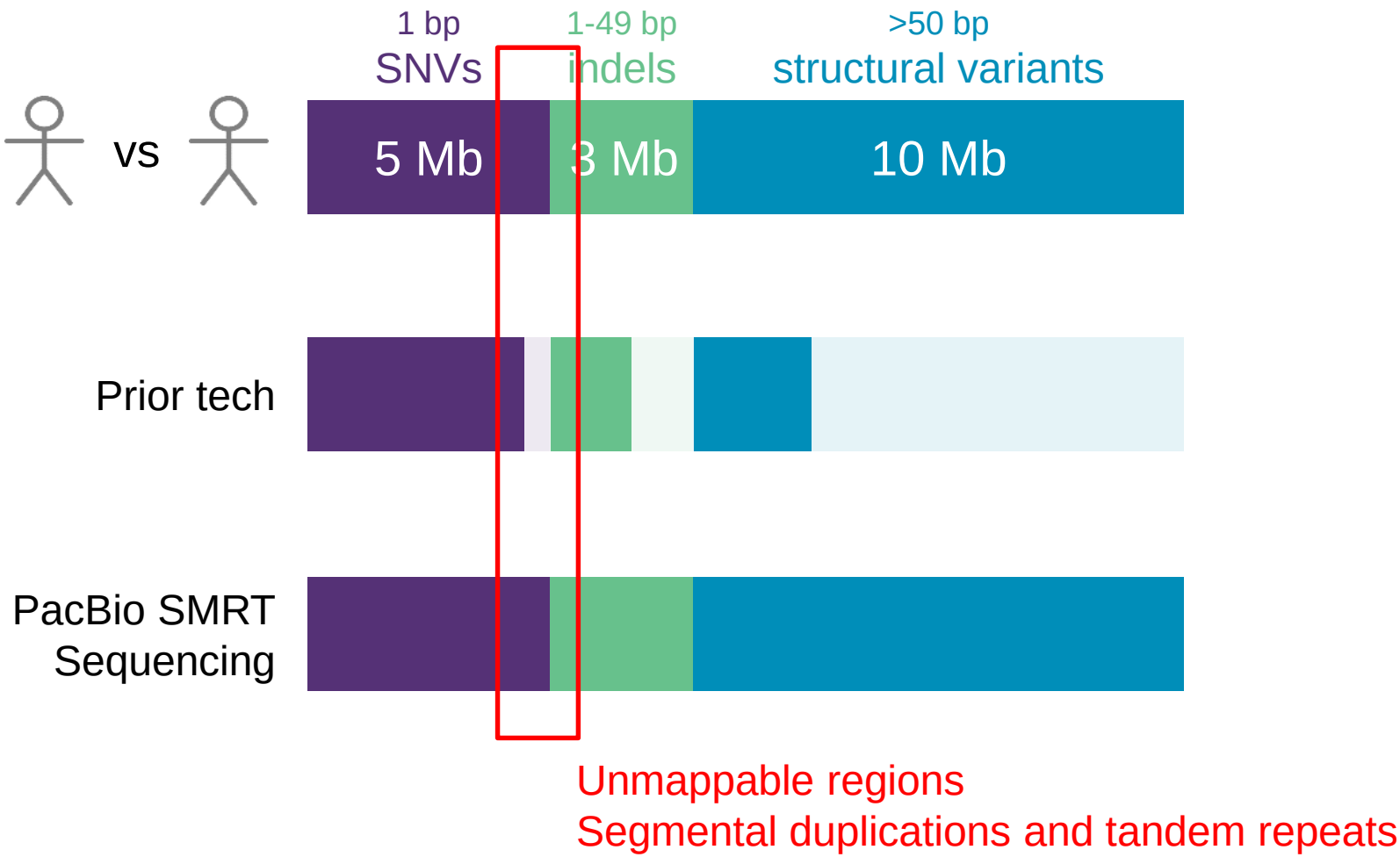
A COMPREHENSIVE VIEW OF THE GENOME

SMRT Sequencing provides a view into ***all*** variation between two humans.



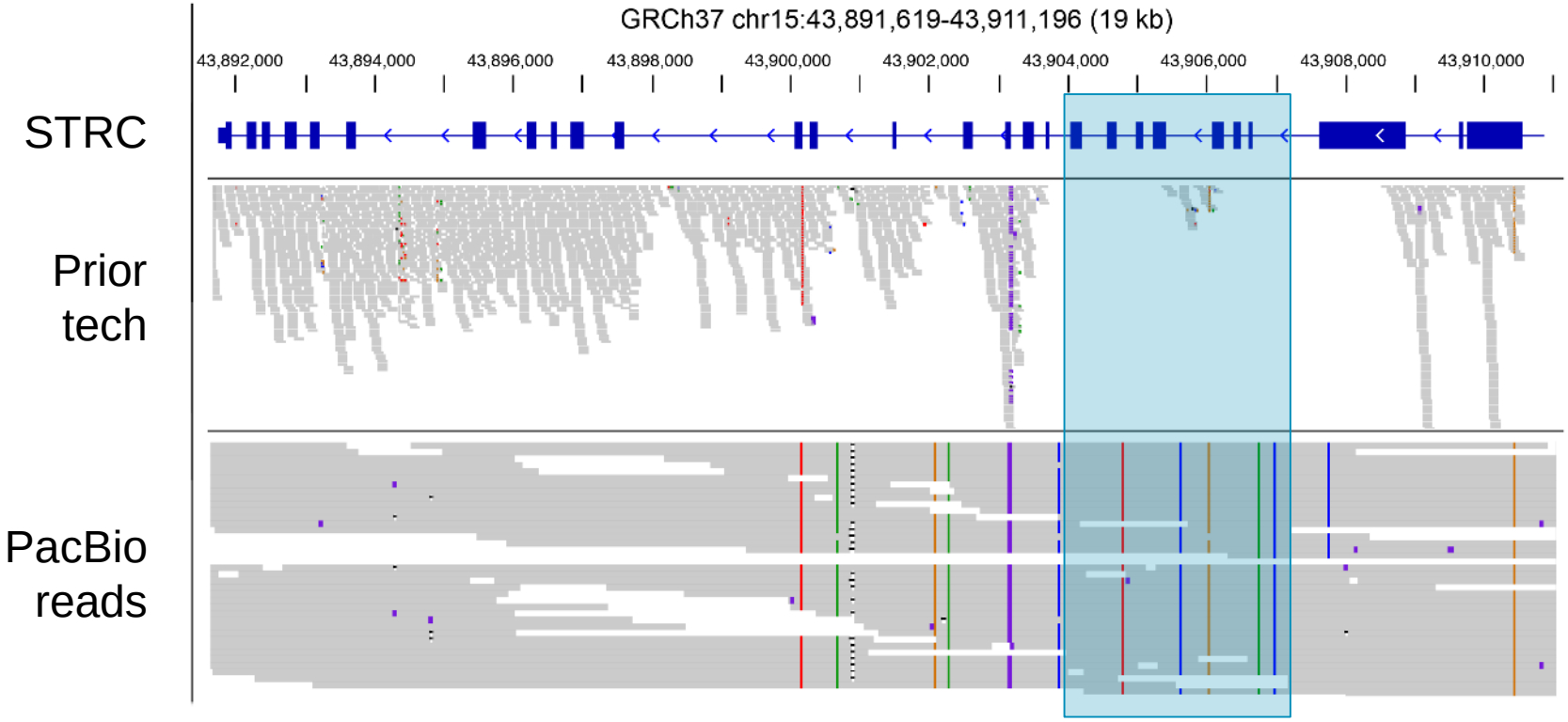
A COMPREHENSIVE VIEW OF THE GENOME

SMRT Sequencing provides a view into ***all*** variation between two humans.



A COMPREHENSIVE VIEW OF THE GENOME

SMRT Sequencing provides even coverage across difficult to sequence regions of the genome.

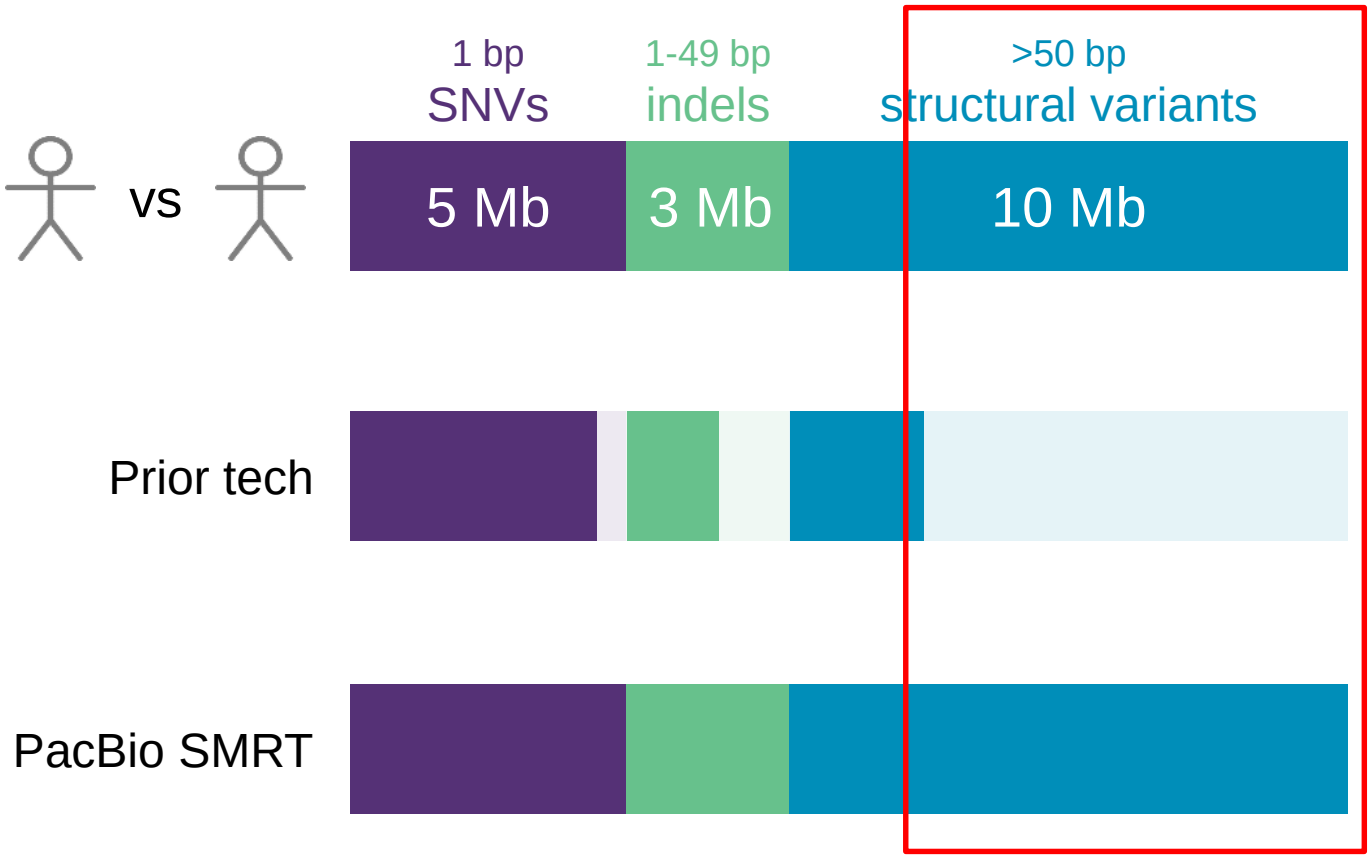


Almost no coverage with prior tech

PacBio reads sequence straight through and detect variants, some falling in coding regions

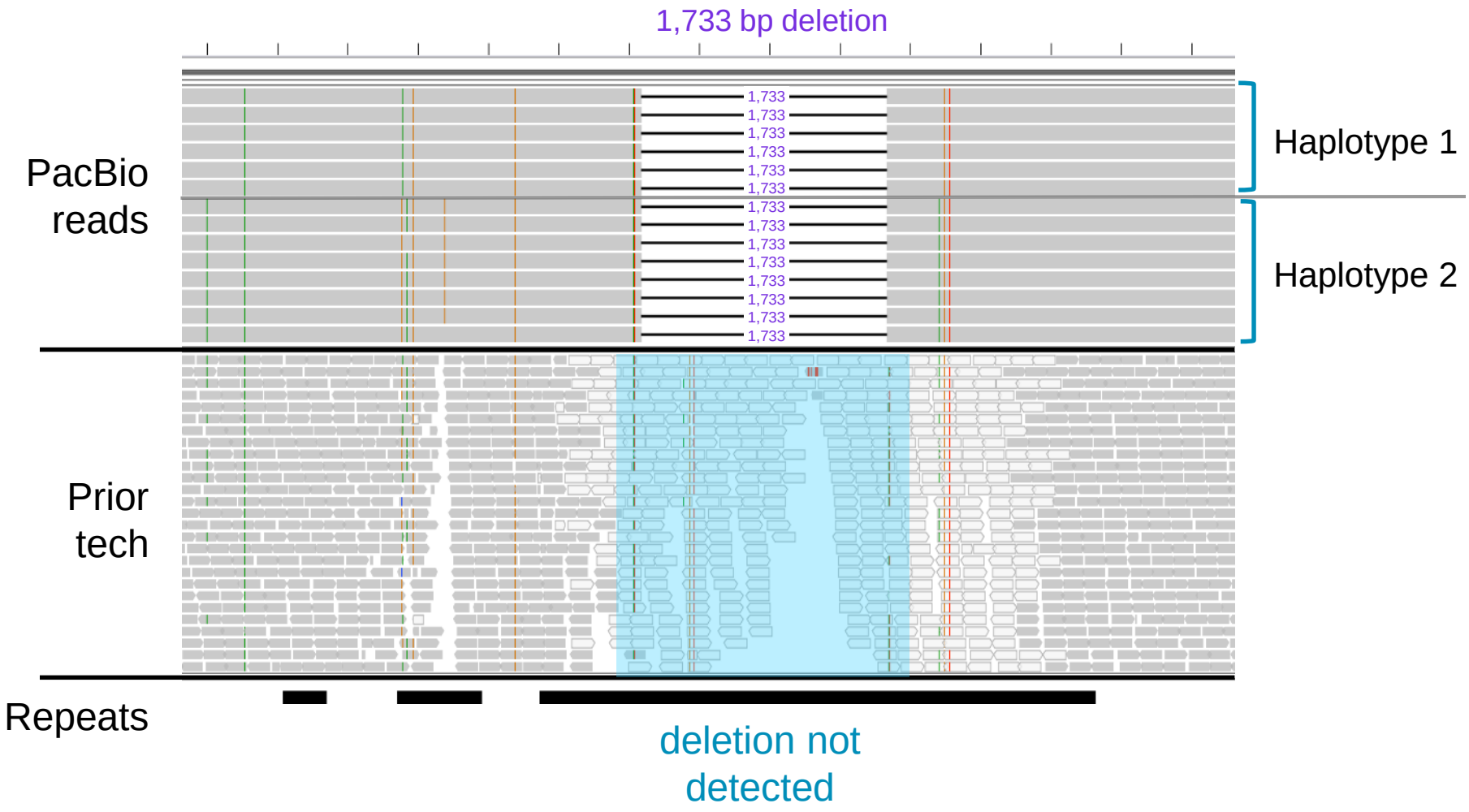
A COMPREHENSIVE VIEW OF THE GENOME

SMRT Sequencing provides a view into ***all*** variation between two humans.



A COMPREHENSIVE VIEW OF THE GENOME

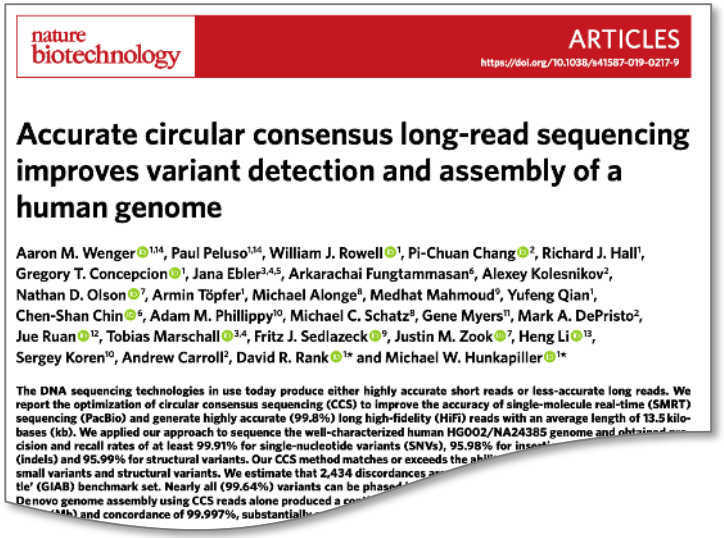
SMRT Sequencing provides long read lengths to span large structural variants.



COMPREHENSIVE VARIANT DETECTION WITH HiFi READS

The HiFi Difference

- High precision and recall for SNVs, indels, and SVs
- Detect 5% more variants in “medical exome”
- Detect structural variants
- Phase variants into haplotypes



nature biotechnology

ARTICLES

https://doi.org/10.1038/s41587-019-0217-9

Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome

Aaron M. Wenger[1,14], Paul Peluso[1,14], William J. Rowell[1], Pi-Chuan Chang[2], Richard J. Hall[1], Gregory T. Concepcion[1], Jana Ebler[3,4,5], Arkarachai Fungtammasan[6], Alexey Kolesnikov[2], Nathan D. Olson[7], Armin Töpfer[1], Michael Alonge[8], Medhat Mahmoud[9], Yufeng Qian[1], Chen-Shan Chin[6], Adam M. Phillippy[10], Michael C. Schatz[8], Gene Myers[11], Mark A. DePristo[2], Jue Ruan[12], Tobias Marschall[3,4], Fritz J. Sedlazeck[9], Justin M. Zook[7], Heng Li[13], Sergey Koren[10], Andrew Carroll[2], David R. Rank[1*] and Michael W. Hunkapiller[1*]

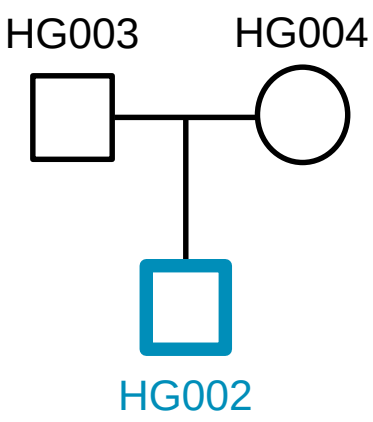
The DNA sequencing technologies in use today produce either highly accurate short reads or less-accurate long reads. We report the optimization of circular consensus sequencing (CCS) to improve the accuracy of single-molecule real-time (SMRT) sequencing (PacBio) and generate highly accurate (99.8%) long high-fidelity (HiFi) reads with an average length of 13.5 kilobases (kb). We applied our approach to sequence the well-characterized human HG002/NA24385 genome and obtained

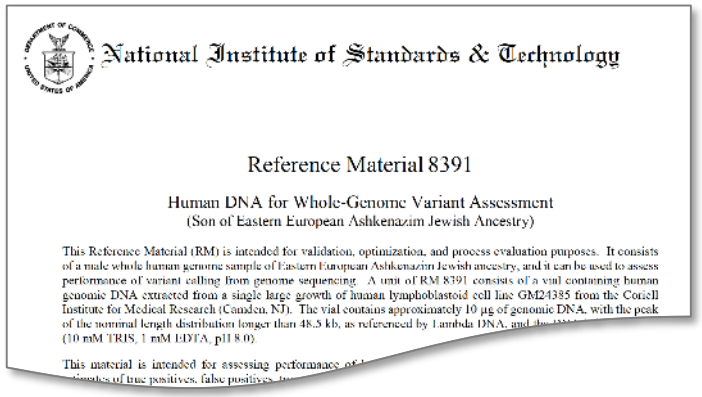
Wenger, A. et al., Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. 2019. *Nature Biotechnology*.

MEASURING PRECISION AND RECALL – GENOME IN A BOTTLE BENCHMARK

Reference materials

Well-characterized benchmark variants





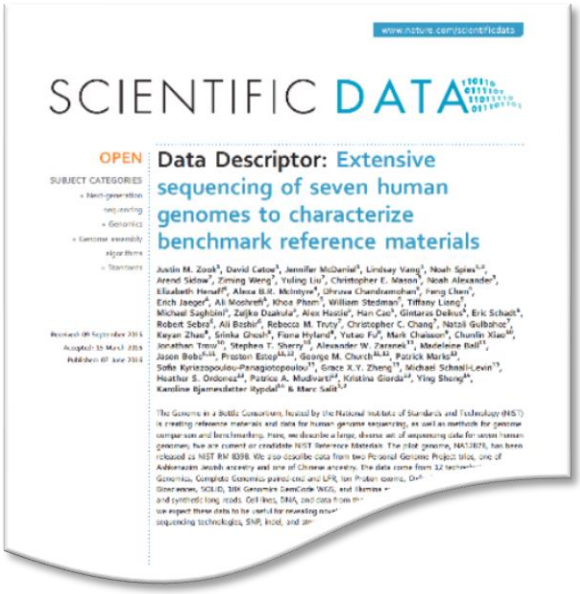
National Institute of Standards & Technology

Reference Material 8391

Human DNA for Whole-Genome Variant Assessment
(Son of Eastern European Ashkenazim Jewish Ancestry)

This Reference Material (RM) is intended for validation, optimization, and process evaluation purposes. It consists of a male whole human genome sample of Eastern European Ashkenazim Jewish ancestry, and it can be used to assess performance of variant calling from genome sequencing. A unit of RM 8391 consists of a vial containing human genomic DNA extracted from a single large growth of human lymphoblastoid cell line GM24385 from the Coriell Institute for Medical Research (Camden, NJ). The vial contains approximately 10 µg of genomic DNA, with the peak of the nominal length distribution longer than 48.5 kb, as referenced by Lambda DNA, and the [illegible] (10 mM TRIS, 1 mM EDTA, pH 8.0).

This material is intended for assessing performance of [illegible] samples of true positives, false positives, [illegible]



www.nature.com/scientificdata

SCIENTIFIC DATA

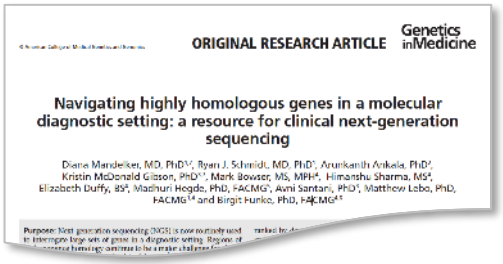
OPEN

SUBJECT CATEGORIES

Data Descriptor: Extensive sequencing of seven human genomes to characterize benchmark reference materials

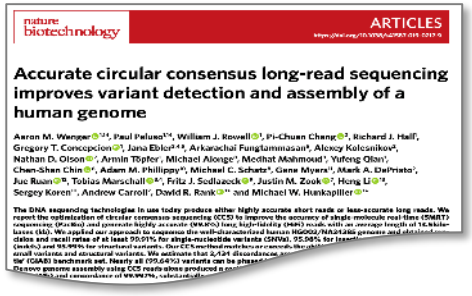
[NIST Genome in a Bottle Consortium]

MORE VARIANTS IN MEDICALLY-RELEVANT GENES


ORIGINAL RESEARCH ARTICLE Genetics inMedicine

Navigating highly homologous genes in a molecular diagnostic setting: a resource for clinical next-generation sequencing

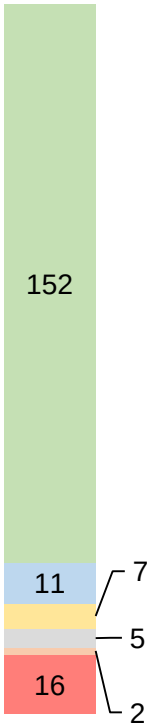
Diane Mandelker, MD, PhD[1], Ryan J. Schmidt, MD, PhD[1], Arunkanth Ankala, PhD[2], Kristin McDonald Gibson, PhD[2,3], Mark Bowser, MS, MPH[4], Himanshu Sharma, MS[4], Elizabeth Duffy, BS[4], Madhuri Hegde, PhD, FACMG[2], Avni Santani, PhD[1], Matthew Lebo, PhD, FACMG[1,4] and Birgit Funke, PhD, FACMG[4,5]


nature biotechnology ARTICLES

Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome

Aaron M. Wenger[1,14], Paul Peluso[1,14], William J. Rowell[1], Pi-Chuan Chang[1], Richard J. Hall[1], Gregory T. Concepcion[1], Jana Ebler[2,13], Arkarachai Fungtammasan[3], Alexey Kolesnikov[3], Nathan D. Olson[4], Armin Töpfer[1], Michael Alonge[5], Medhat Mahmoud[5], Yufeng Qian[5], Chen-Shan Chin[6], Adam M. Phillippy[6], Michael C. Schatz[5], Gene Myers[13], Mark A. DePristo[1], Jue Ruan[7], Tobias Marschall[2,8], Fritz J. Sedlazeck[9], Justin M. Zook[4], Heng Li[10,11], Sergey Koren[12], Andrew Carroll[1], David R. Rank[1,14] and Michael W. Hunkapiller[1,14]

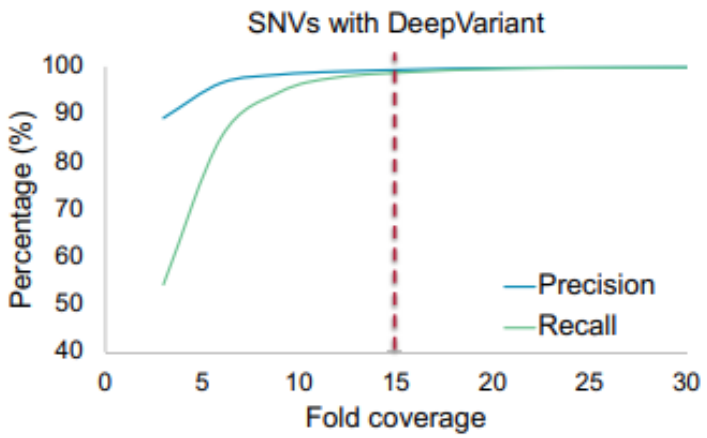
% problem exons resolved	Genes
100%	*ABCC6, ABCD1, ACAN, ACSM2B, AKR1C2, ALG1, ANKRD11, BCR, CATSPER2, CD177, CEL, CES1, CFH, CFHR1, CFHR3, CFHR4, CGB, CHEK2, CISD2, CLCNKA, CLCNKB, CORO1A, COX10, CRYBB2, CSH1, CYP11B1, CYP11B2, CYP21A2, CYP2A6, CYP2D6, CYP2F1, CYP4A22, DDX11, DHRS4L1, DIS3L2, DND1, DPY19L2, DUOX2, ESRRA, F8, FAM120A, FAM205A, FANCD2, FCGR1A, FCGR2A, FCGR3A, FCGR3B, FLG, FLNC, FOXD4, FOXO3, FUT3, GBA, GFRA2, GON4L, GRM5, GSTM1, GYPA, GYPB, GYPE, HBA1, HBA2, HBG1, HBG2, HP, HS6ST1, IDS, IFT122, IKBKG, IL9R, KIR2DL1, KIR2DL3, KMT2C, KRT17, KRT6A, KRT6B, KRT6C, KRT81, KRT86, LEFTY2, LPA, MST1, MUC5B, MYH6, MYH7, NEB, NLGN4X, NLGN4Y, NOS2, NOTCH2, NXF5, OPN1LW, OR2T5, OR51A2, PCDH11X, PCDHB4, PGAM1, PHC1, PIK3CA, PKD1, PLA2G10, PLEKHM1, PLG, PMS2, PRB1, PRDM9, PROS1, RAB40AL, RALGAPA1, RANBP2, RHCE, RHD, RHPN2, ROCK1, SAA1, SDHA, SDHC, SFTPA1, SFTPA2, SIGLEC14, SLC6A8, SMG1, SPATA31C1, SPTLC1, SRGAP2, SSX7, STAT5B, STK19, STRC, SULT1A1, SUZ12, TBX20, TCEB3C, TLR1, TLR6, TMEM231, TNXB, TRIOBP, TRPA1, TTN, TUBA1A, TUBB2B, UGT1A5, UGT2B15, UGT2B17, UNC93B1, VCY, VWF, WDR72, ZNF419, ZNF592, ZNF674*
[75%, 100%)	*ANAPC1, C4A, C4B, CHRNA7, CR1, DUX4, FCGR2B, HYDIN, OTOA, PDPK1, TMLHE*
[50%, 75%)	*ADAMTSL2, CDY2A, DAZ1, GTF2I, NAIP, OCLN, RPS17*
[25%, 50%)	*DAZ2, DAZ3, KIR3DL1, OPN1MW, PPIP5K1*
(0%, 25%)	*NCF1, RBMY1A1*
0%	*BPY2, CCL3L1, CCL4L1, CDY1, CFC1, CFC1B, GTF2IRD2, HSFY1, MRC1, OR4F5, PRY, PRY2, SMN1, SMN2, TSPY1, XKRY*

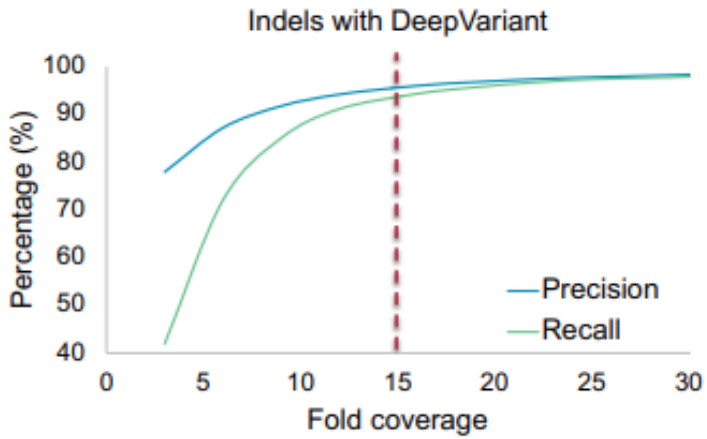


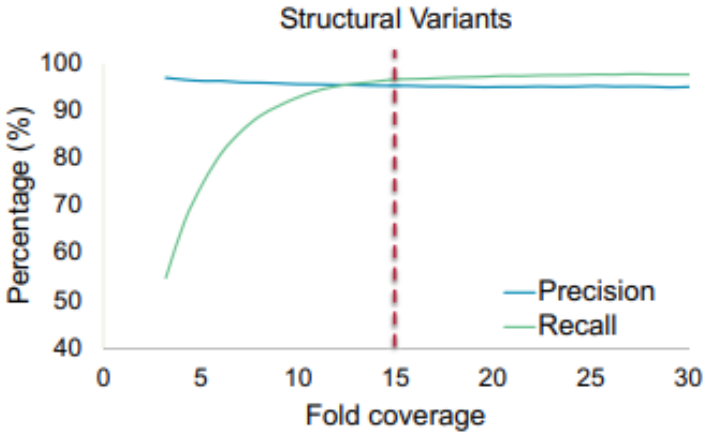
Mandelker, D. et al., Navigating highly homologous genes in a molecular diagnostic setting: a resource for clinical next-generation sequencing. 2016. *Genetics in Medicine* 18, 1282-1289

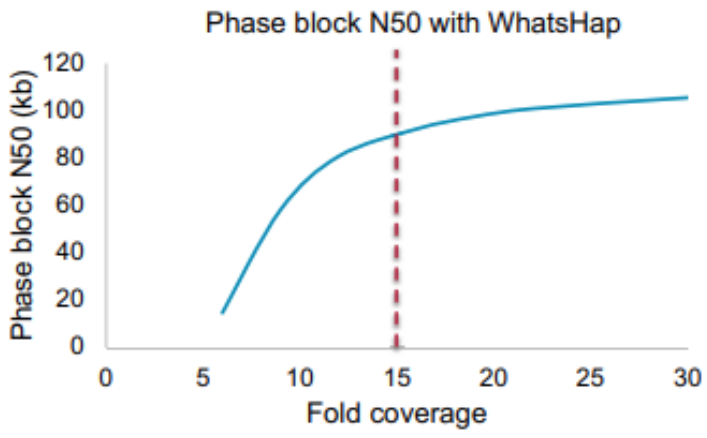
Wenger, A. et al., Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. 2019. *Nature Biotechnology.*

15-FOLD HiFi READ COVERAGE RECOMMENDATION









Wenger, A. et al., Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. 2019. *Nature Biotechnology.*

WORKFLOW: FROM DNA TO COMPREHENSIVE VARIANT DETECTION WITH HiFi READS

HiFi READS

LIBRARY PREP

3-6 µg DNA input per sample

Prepare 1 SMRTbell library
per sample with Express Template Prep Kit 2.0
(15-20 kb size fraction)

SMRT SEQUENCING

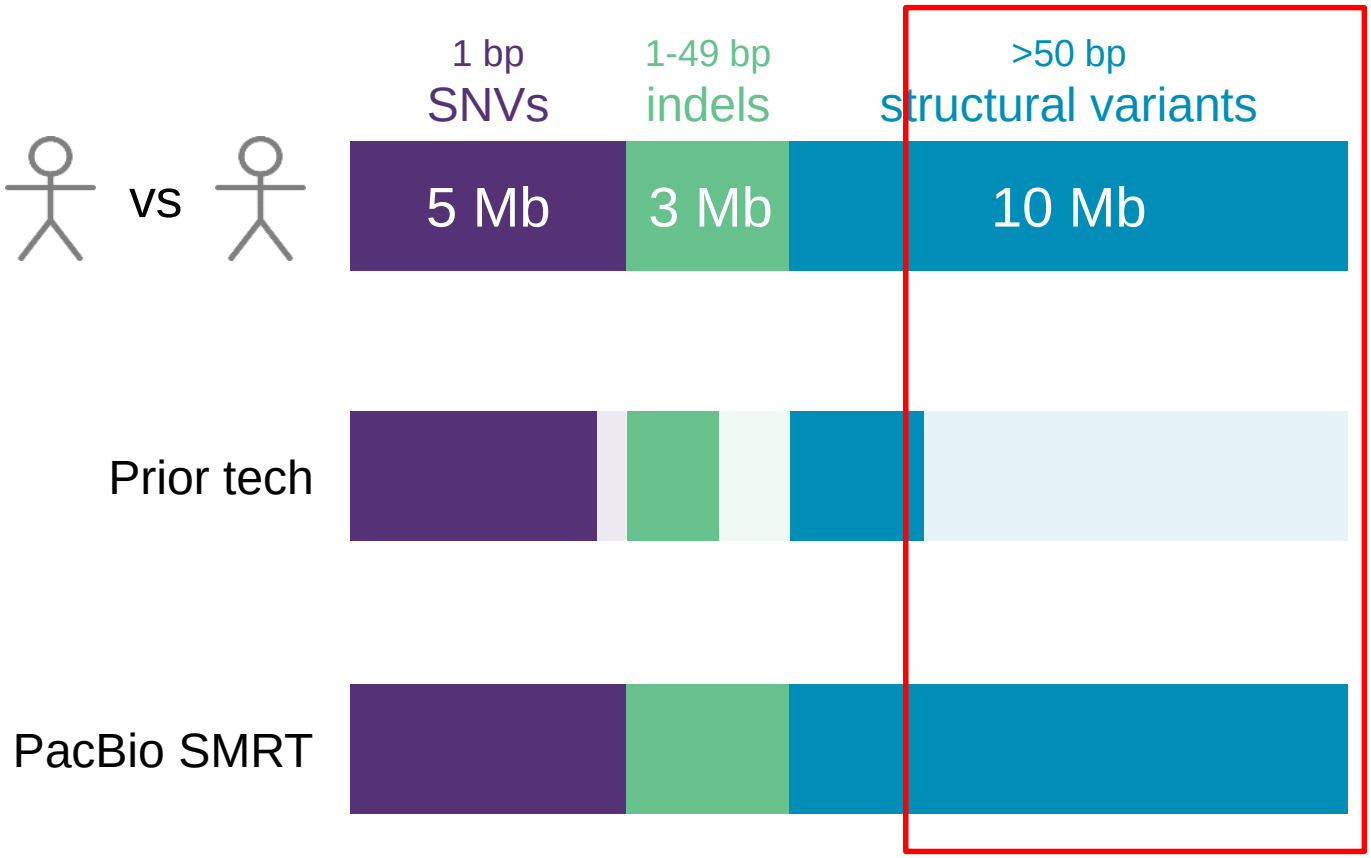
Sequence 2 SMRT Cells 8M per sample for 30 hrs (in CCS mode)

DATA ANALYSIS

Call small variants with DeepVariant or GATK
Call structural variants with pbsv

STRUCTURAL VARIANT DETECTION IN A SINGLE SMRT CELL 8M

SMRT Sequencing detects novel SVs of all types and lengths with base pair resolution of breakpoints.



STRUCTURAL VARIANT DETECTION WITH LONG READS

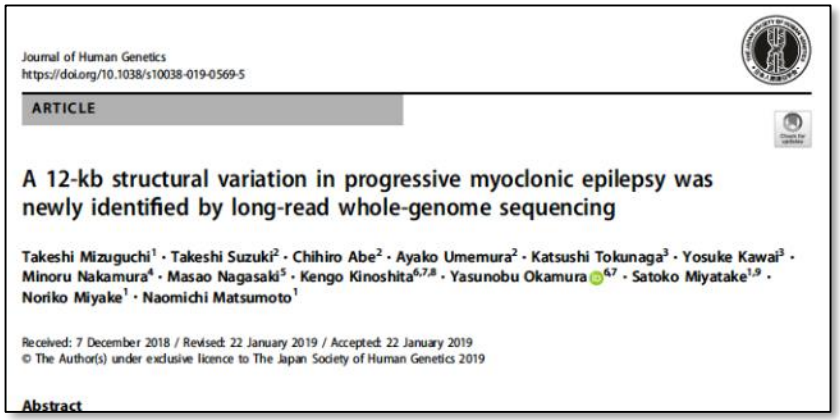
The SMRT Sequencing Difference

High precision and recall for detection of:

- Insertions and deletions
- Inversions
- Translocations
- Copy-number variants

Experiment type	SMRT Sequencing	+ Short-read Sequencing
Disease Research	10-fold CLR coverage to detect SVs (multiplex 2 samples per SMRT Cell 8M)	30-fold coverage to detect SNVs and indels
Population Genetics	5- to 10-fold CLR coverage per sample to discover SVs (multiplex 2-3 samples per SMRT Cell 8M)	30-fold coverage per sample in a large cohort to genotype discovered SVs

CASE STUDIES



Journal of Human Genetics
https://doi.org/10.1038/s10038-019-0569-5

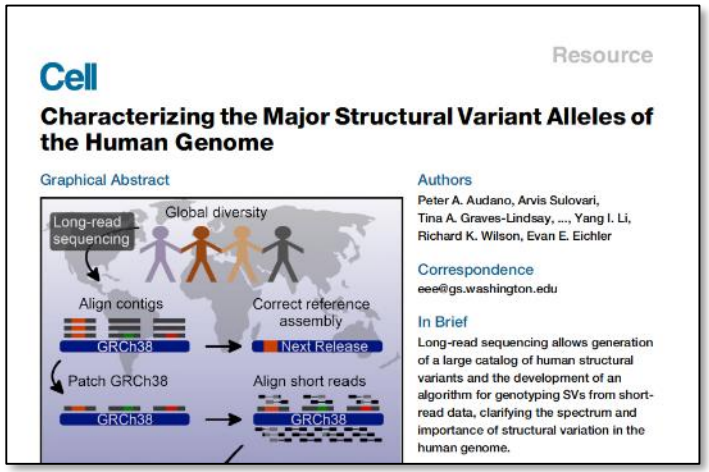
ARTICLE

A 12-kb structural variation in progressive myoclonic epilepsy was newly identified by long-read whole-genome sequencing

Takeshi Mizuguchi[1] · Takeshi Suzuki[2] · Chihiro Abe[2] · Ayako Umemura[2] · Katsushi Tokunaga[3] · Yosuke Kawai[3] · Minoru Nakamura[4] · Masao Nagasaki[5] · Kengo Kinoshita[6,7,8] · Yasunobu Okamura[4,7] · Satoko Miyatake[1,9] · Noriko Miyake[1] · Naomichi Matsumoto[1]

Received: 7 December 2018 / Revised: 22 January 2019 / Accepted: 22 January 2019
© The Author(s) under exclusive licence to The Japan Society of Human Genetics 2019

Abstract



Resource

Cell

Characterizing the Major Structural Variant Alleles of the Human Genome

Graphical Abstract

Authors

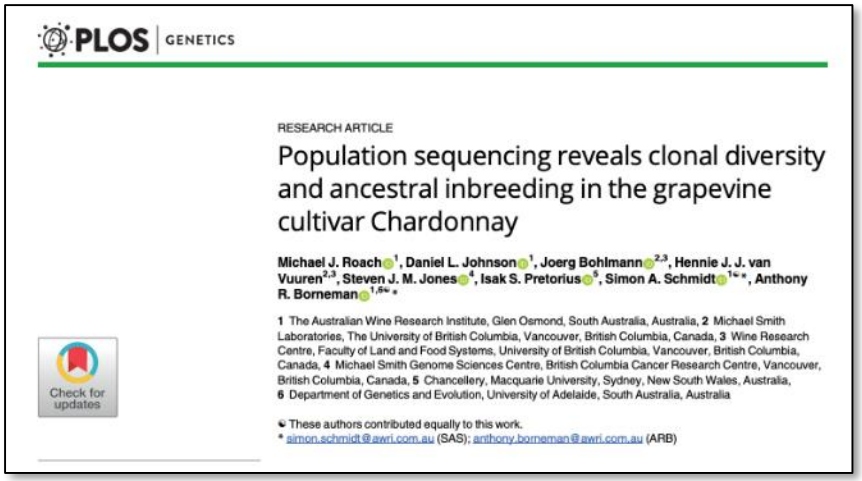
Peter A. Audano, Arvis Sulovari, Tina A. Graves-Lindsay, ..., Yang I. Li, Richard K. Wilson, Evan E. Eichler

Correspondence

eee@gs.washington.edu

In Brief

Long-read sequencing allows generation of a large catalog of human structural variants and the development of an algorithm for genotyping SVs from short-read data, clarifying the spectrum and importance of structural variation in the human genome.



PLOS GENETICS

RESEARCH ARTICLE

Population sequencing reveals clonal diversity and ancestral inbreeding in the grapevine cultivar Chardonnay

Michael J. Roach[1], Daniel L. Johnson[1], Joerg Bohlmann[2,3], Hennie J. J. van Vuuren[2,3], Steven J. M. Jones[4], Isak S. Pretorius[5], Simon A. Schmidt[1☯*], Anthony R. Borneman[1,6☯*]

1 The Australian Wine Research Institute, Glen Osmond, South Australia, Australia, 2 Michael Smith Laboratories, The University of British Columbia, Vancouver, British Columbia, Canada, 3 Wine Research Centre, Faculty of Land and Food Systems, University of British Columbia, Vancouver, British Columbia, Canada, 4 Michael Smith Genome Sciences Centre, British Columbia Cancer Research Centre, Vancouver, British Columbia, Canada, 5 Chancellery, Macquarie University, Sydney, New South Wales, Australia, 6 Department of Genetics and Evolution, University of Adelaide, South Australia, Australia

☯ These authors contributed equally to this work.
* simon.schmidt@awri.com.au (SAS); anthony.borneman@awri.com.au (ARB)

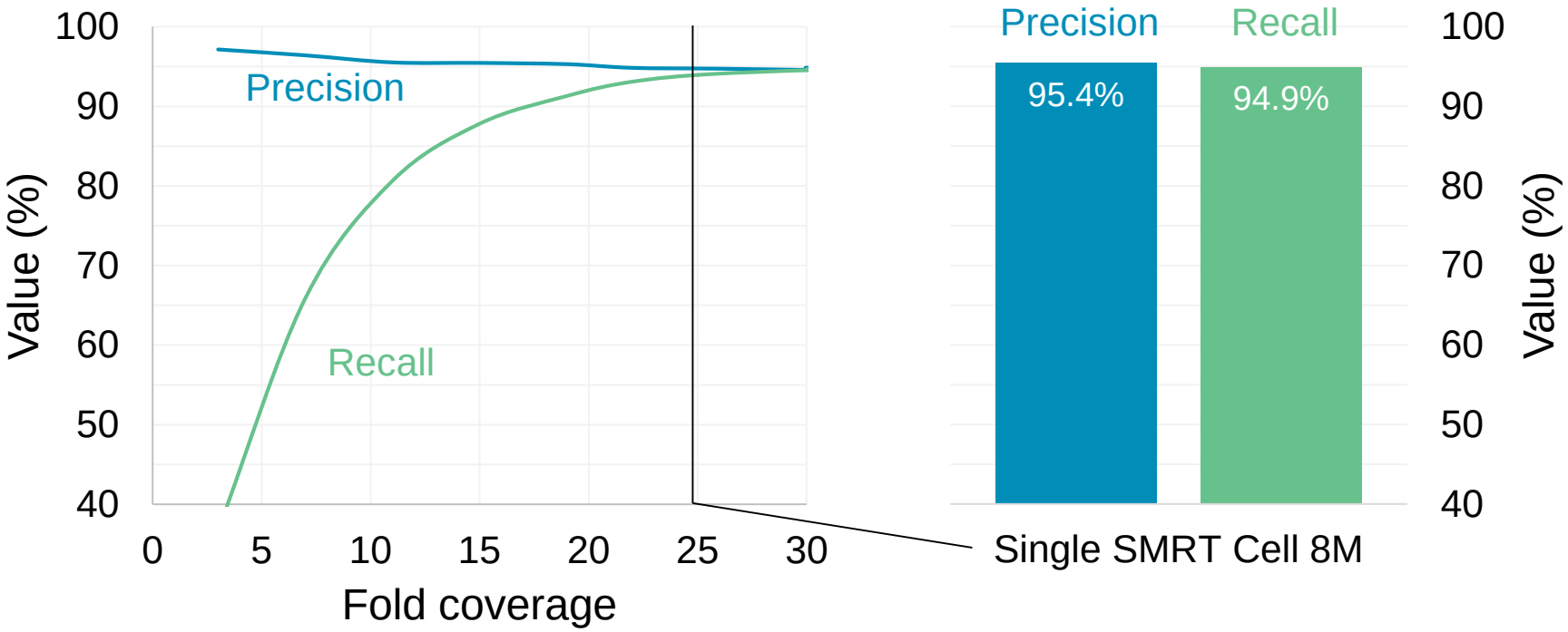
Check for updates

Mizuguchi T, et al. A 12-kb structural variation in progressive myoclonic epilepsy was newly identified by long-read whole-genome sequencing. 2019. *J Human Genetics*, 64, 359-368.

Audano P, et al. Characterizing the Major Structural Variant Alleles of the Human Genome. 2019. *Cell*, 176 (3), 663-675.

Roach, M.J., et al. Population sequencing reveals clonal diversity and ancestral inbreeding in the grapevine cultivar Chardonnay. 2018. *PLOS Genetics* 14.

RECALL VS COVERAGE FOR SV DETECTION



Precision and recall for a single SMRT Cell 8M is for a human-sized genome. For a genome >3 Gb, more SMRT Cells may be required.

WORKFLOW: FROM DNA TO SENSITIVE STRUCTURAL VARIANT DETECTION

LONG READS

LIBRARY PREP

Prepare 3-6 µg DNA (>20 kb) input per sample

Prepare 1 SMRTbell library per sample using barcoded adapters and Express Template Prep Kit 2.0
(size select >15 kb)

SMRT SEQUENCING

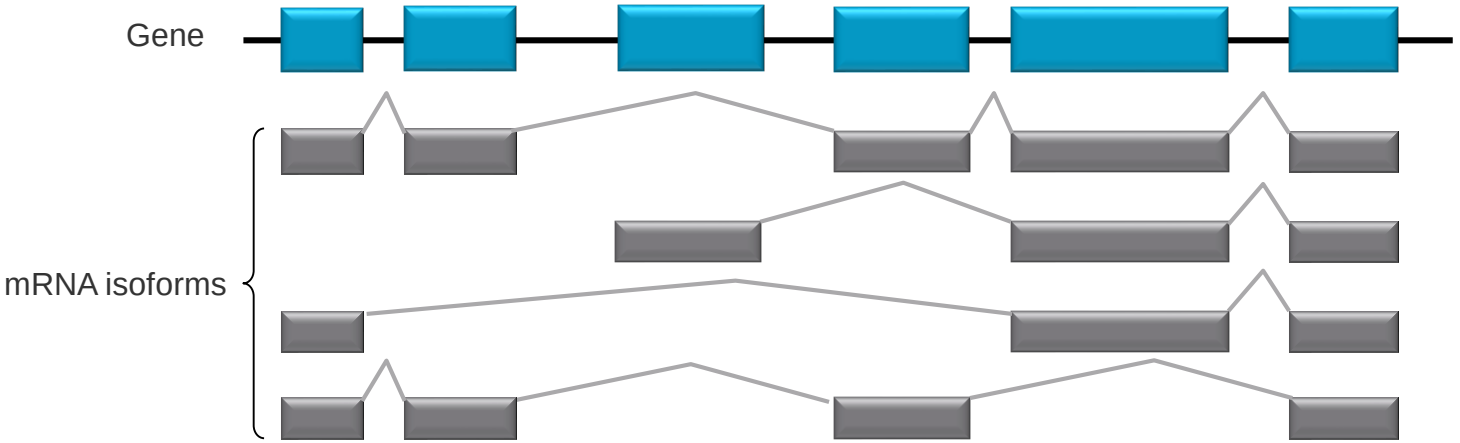
Multiplex 2 samples per 1 SMRT Cell 8M, sequencing for 15 hrs (in CLR mode)

DATA ANALYSIS

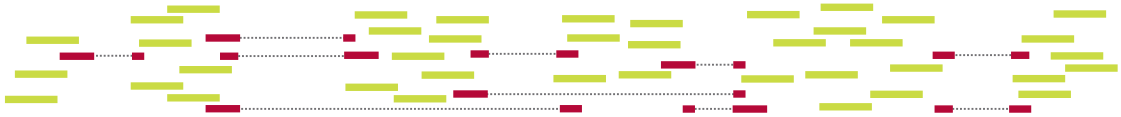
Call structural variants jointly with pbsv

RNA Sequencing with the Iso-Seq Method for Transcriptome Analysis

DETERMINATION OF TRANSCRIPT ISOFORMS



Other technologies:



Insufficient Connectivity
Splice Isoform Uncertainty

Reads spanning splice junctions

PacBio's Iso-Seq solution:

Full-length cDNA Sequence Reads
Splice Isoform Certainty – No Assembly Required

ISO-SEQ METHOD: FULL-LENGTH RNA SEQUENCING

Sequence full-length cDNA sequences – from 5’ end to the poly-A tail – without the need for transcript reconstruction.

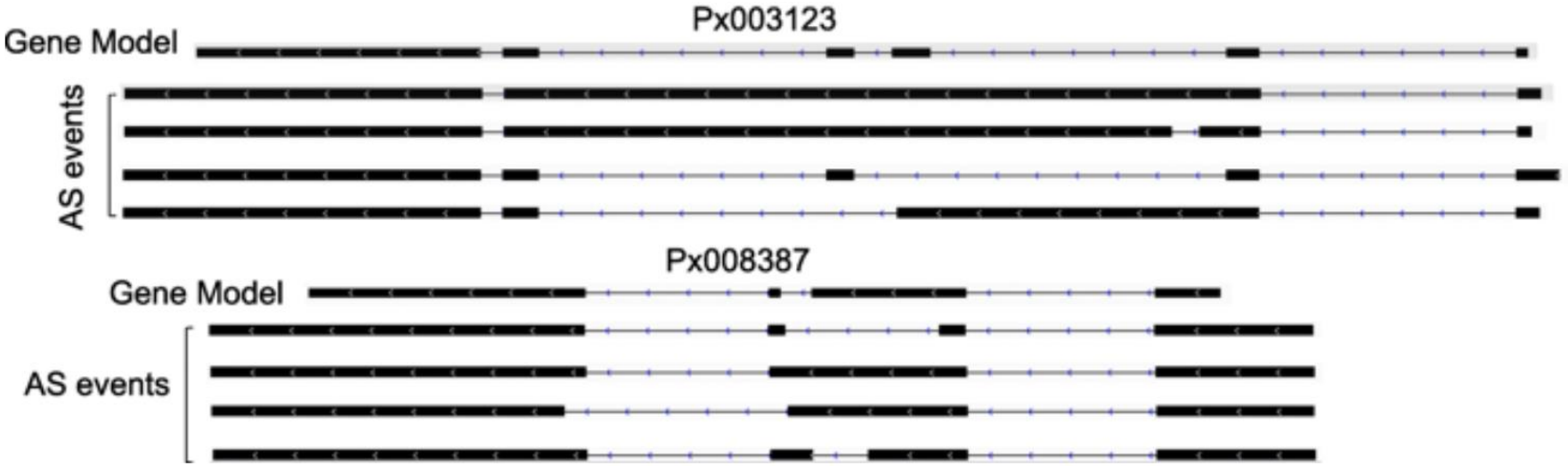
The PacBio Iso-Seq method allows you to:

- Profile whole transcriptomes exhaustively at the isoform level
- Discover novel genes and isoforms
- Identify exon skipping events and alternative 5' / 3' sites
- Characterize function without reference genomes
- Combine with and complement RNA-seq to quantify at isoform-level

www.pacb.com/applications/rna-sequencing

IMPROVED GENOME ANNOTATION WITH ISO-SEQ

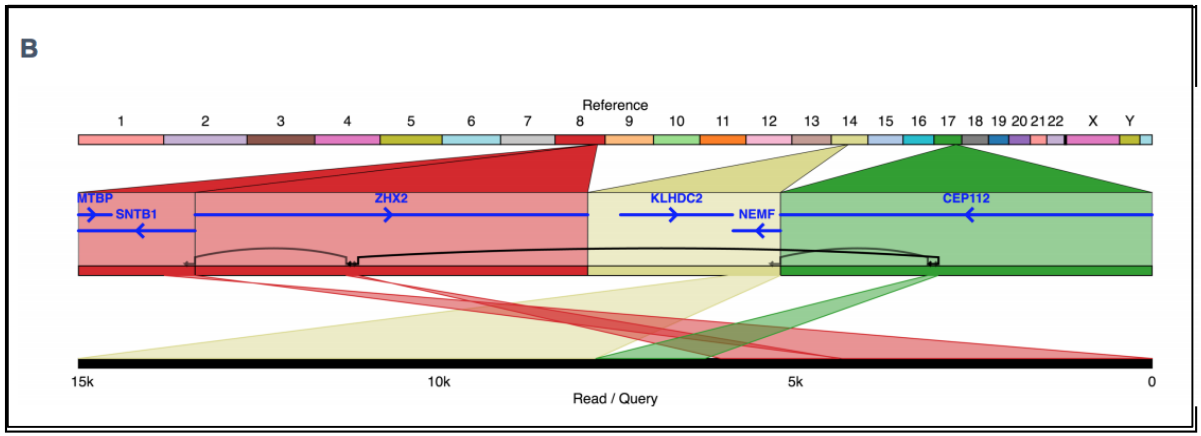
Plutella xylostella
Diamondback moth



Alternative splicing contributes to a **much more diverse transcript set** compared to original gene model

Zhao, Q. et al., 2019. Genome-wide profiling of the alternative splicing provides insights into development in Plutella xylostella, *BMC Genomics*, 20:463.

SEQUENCING THE CANCER GENOME & TRANSCRIPTOME



Supplementary Figure 18. Ribbon plot of "3-hop" KLHDC2-SNTB1 gene fusion captured by long reads. This is a "3-hop" gene fusion in SK-BR-3 created by a series of three variants (A). These variants are captured together in several individual SMRT sequencing reads, one of which is shown in (B).

In total, Iso-Seq identified **15 fusion genes** with genomic evidence

Further reading: Medium blogpost

Nattestad, M. *et al.* Complex rearrangements and oncogene amplifications revealed by long-read DNA and RNA sequencing of a breast cancer cell line. *Genome Res.* 1–19 (2018).

WORKFLOW: THE ISO-SEQ METHOD FOR TRANSCRIPTOME CHARACTERIZATION

LIBRARY PREP

300 ng of total RNA input per sample

Prepare 1 SMRTbell library per sample with Express Template Prep Kit 2.0

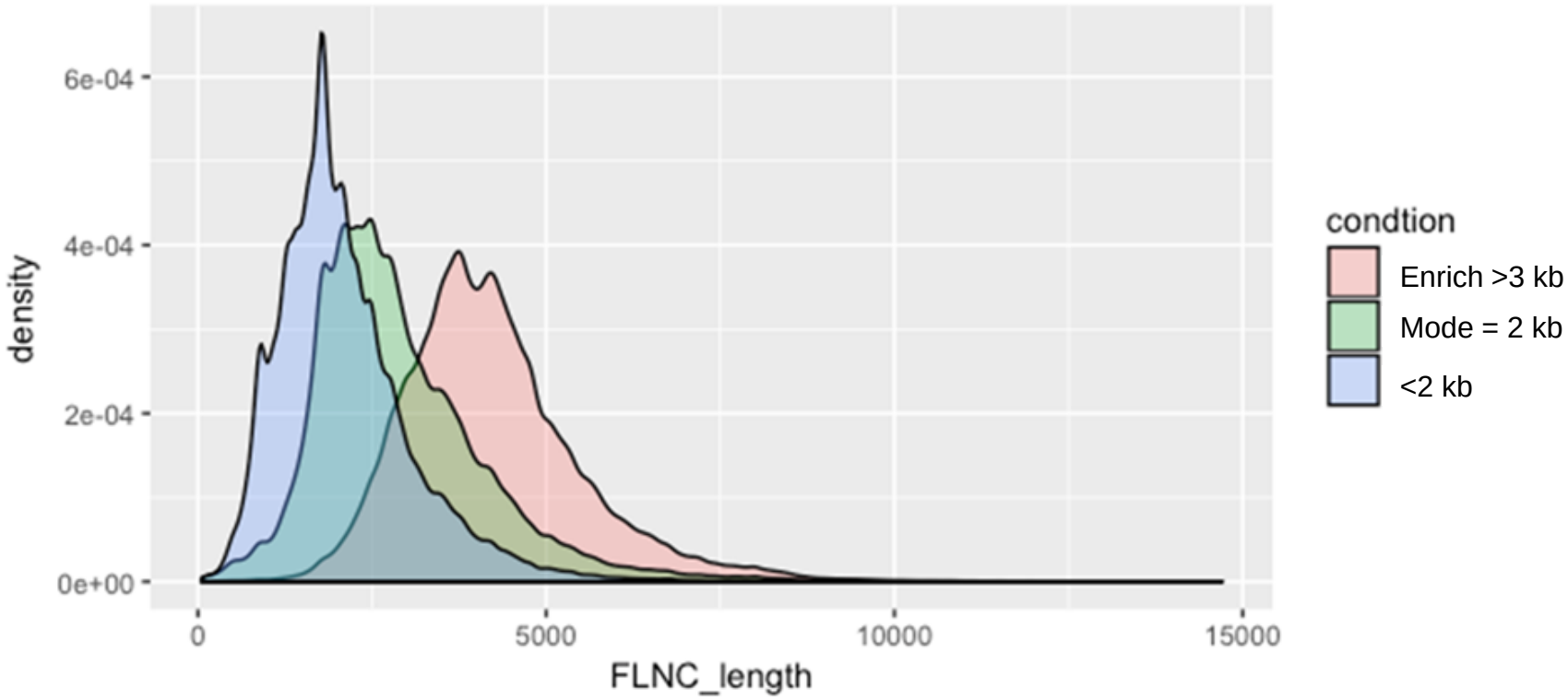
SMRT SEQUENCING

Use the Sequel II System to generate up to 4 million full-length, non-concatemer (FLNC) reads per SMRT Cell 8M

DATA ANALYSIS

Use the Iso-Seq analysis in SMRT Link to output high-quality, full-length transcript sequences

PRONEX BEAD PURIFICATION ENABLES MODULATION OF FULL-LENGTH TRANSCRIPT SIZE-DISTRIBUTION



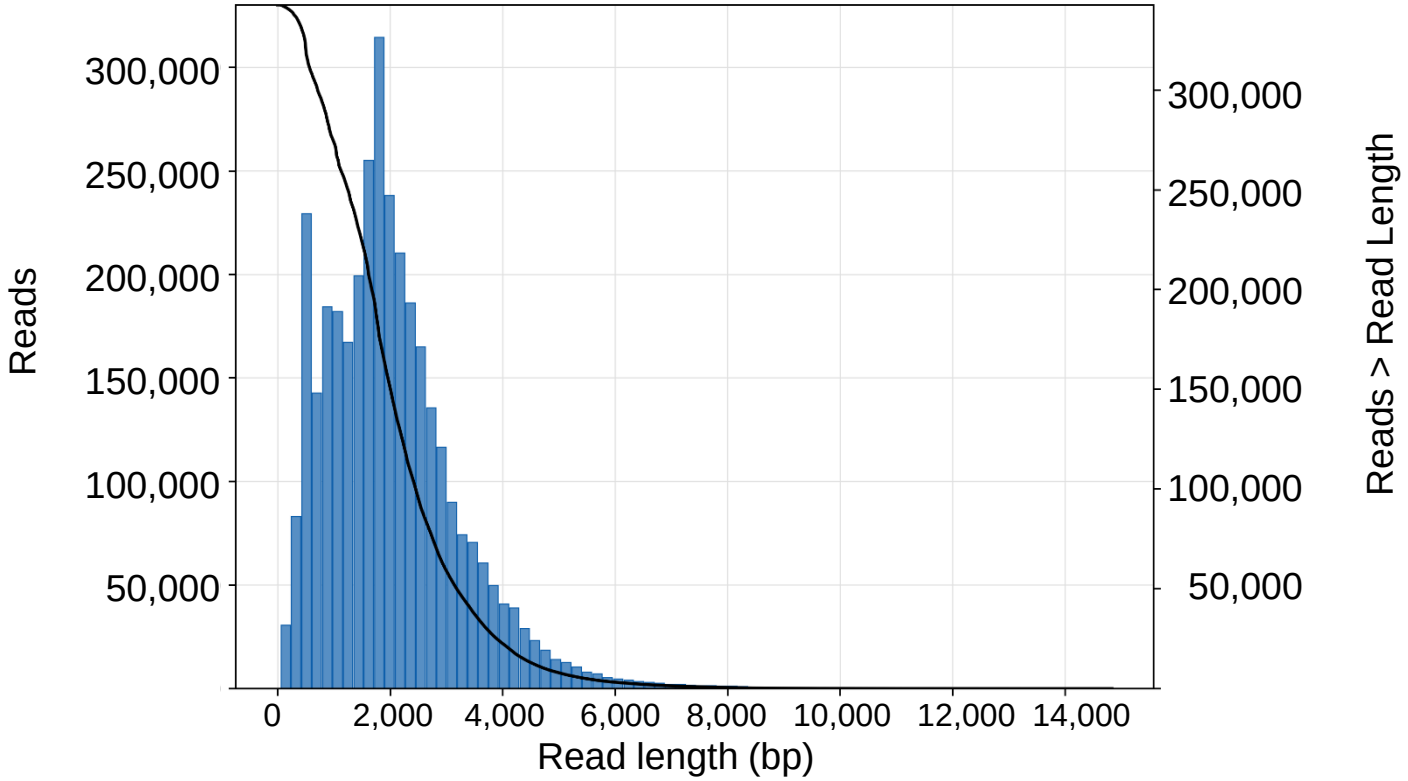
ISO-SEQ EXPRESS RESULTS ACROSS VARIOUS SAMPLES (SEQUEL II SYSTEM)

Sample #	Sample Description	Protocol	# FLNC Reads	% FLNC Reads
1	UHR	Standard	3,466,513	85%
2	Mouse Liver	Standard	3,431,638	87%
3	MCF7	Standard	3,531,419	84%
4	Brain	Standard	2,943,148	86%
5	Alz Brain Tissue	Standard	3,142,634	83%
6	Heart	Standard	2,753,509	87%
7	Liver	long	3,542,983	85%
8	ColT Cell Line	short	2,852,434	84%

ISO-SEQ PERFORMANCE

- Iso-Seq Universal Human Control

Metric	
Number of Raw Bases (Gb)	255
Total Reads	4,460,955
Full Length Non-chimeric Reads	3,436,022
CCS Passes (Mean)	8



Data shown above from a Universal Human Reference RNA (human) and Lexogen SIRV spike-in controls. The library was constructed using the SMRTbell Template Prep Kit on a Sequel II System (1.0 Chemistry, Sequel II System Software v7.0, 20-hour movie). Read lengths, reads/data per SMRT Cell 8M and other sequencing performance results vary based on sample quality/type and insert size.

PACBIO®

Sequencing for High-Resolution Metagenomics and Closed Microbial Genomes

PACBIO PROVIDES HIGHER-RESOLUTION VIEWS OF BOTH METAGENOME MEMBERS AND FUNCTIONS

Full-length 16S provide species and strain-level information

- Highly accurate, full-length 16S sequences distinguishes keystone or critical *species* from genus-level noise
- Follow our 16S amplification and sequencing protocol or use an all-in-one kit from our partner Shoreline Biome

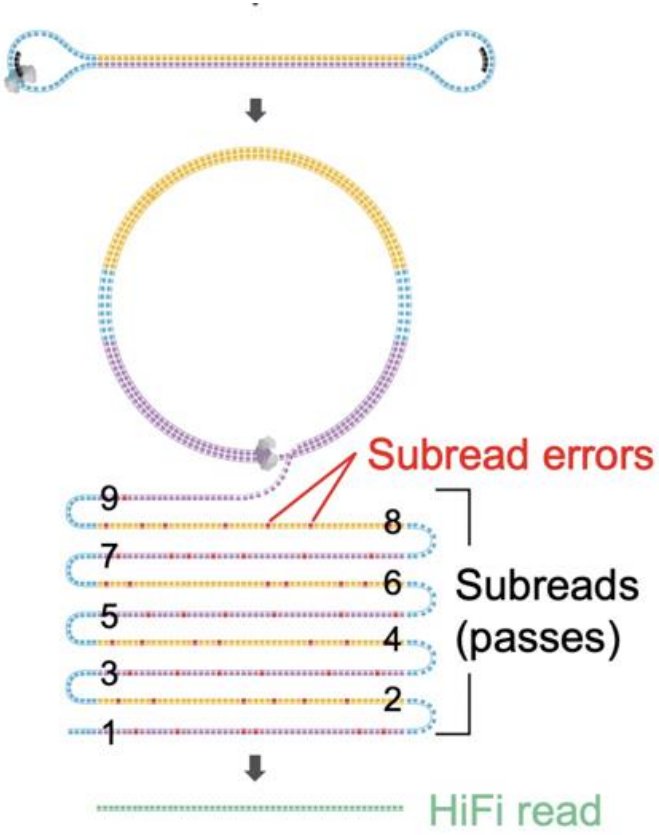
HiFi Shotgun profiling reveals intact genes and operons without assembly

- Move beyond correlation and reveal community functions with highly accurate, 10 kb HiFi reads
- Obtain high-confidence information even for low-abundance sequences seen only once

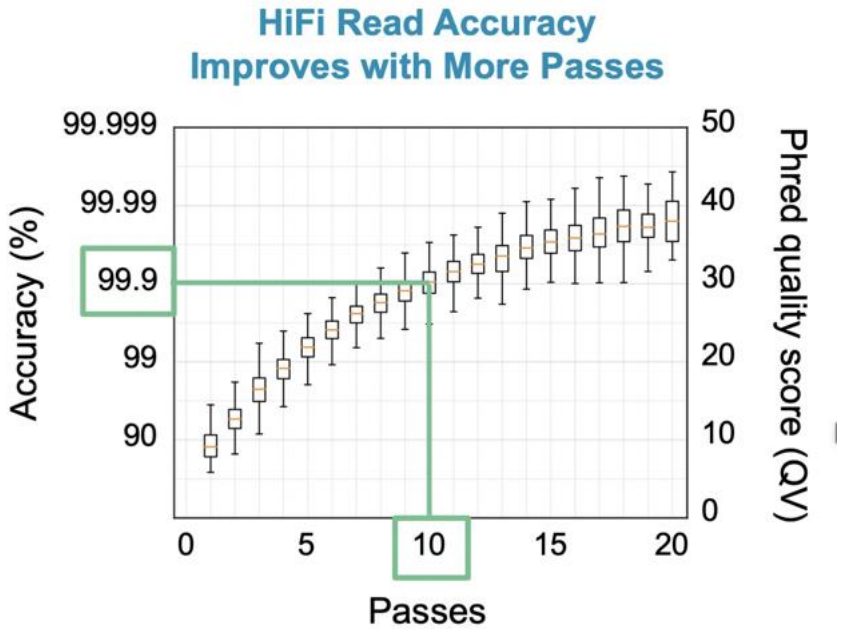
Metagenome assembly from long reads generates new references

- Resolve genomes of microbes that can’t be easily cultured
- Leverage epigenomic data to cluster contigs and plasmids from the same strain

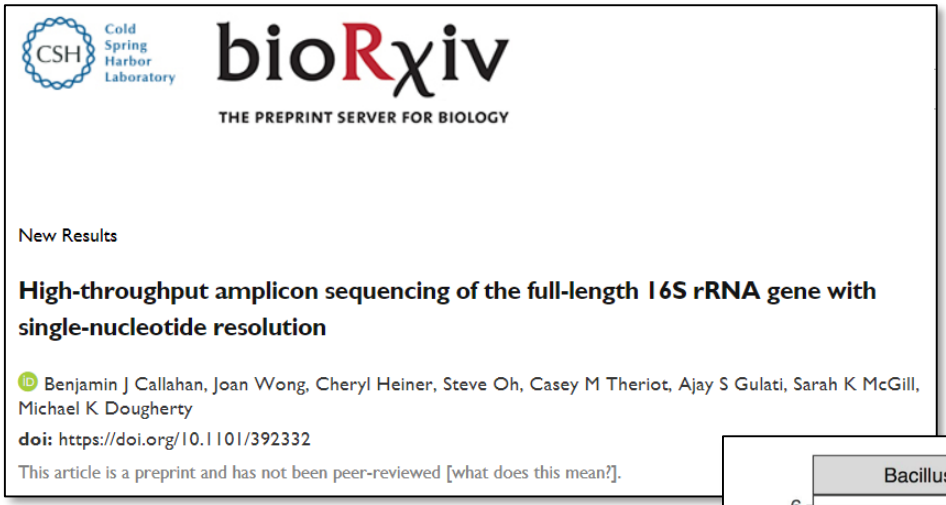
ONLY PACBIO CAN DELIVER SINGLE MOLECULE LONG READS WITH HIGH ACCURACY



- With 100 kb average read lengths, both 16S and shotgun libraries can be sequenced at very high accuracy on the Sequel II system
- **16S**: Up to 2 M Q30 full-length 16S sequences
- **Shotgun**: Up to 1.7 M Q20 reads from 10 kb libraries



FULL-LENGTH 16S SEQUENCING



Cold Spring Harbor Laboratory

bioRxiv

THE PREPRINT SERVER FOR BIOLOGY

New Results

High-throughput amplicon sequencing of the full-length 16S rRNA gene with single-nucleotide resolution

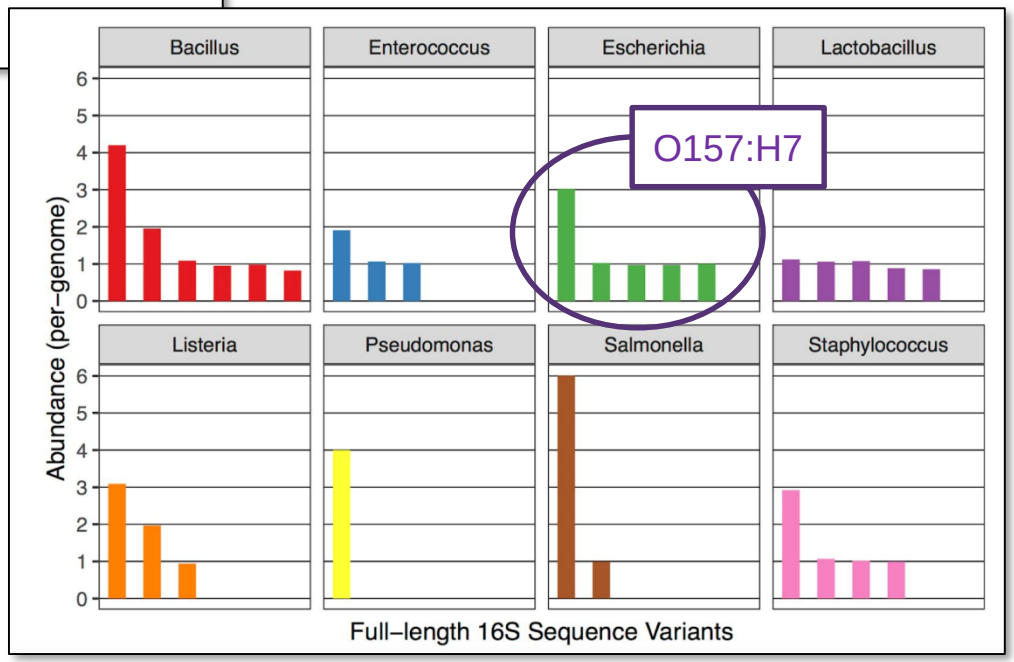
Benjamin J Callahan, Joan Wong, Cheryl Heiner, Steve Oh, Casey M Theriot, Ajay S Gulati, Sarah K McGill, Michael K Dougherty

doi: https://doi.org/10.1101/392332

This article is a preprint and has not been peer-reviewed [what does this mean?].

"The high resolution and accuracy we are reporting derives in part from the exceptional and not-entirely-appreciated accuracy of PacBio CCS sequencing."

- Nearly every bacteria has multiple copies of the 16S housekeeping gene, but in many cases they are not perfect duplicates
- PacBio CCS produced multiple distinct 16S sequences per bacterial genome, and they appear in integer ratios that reflected their copy number in each genome



Callahan, BJ et. al. (2018) High-throughput amplicon sequencing of the full-length 16S rRNA gene with single-nucleotide resolution. bioRxiv doi: http://dx.doi.org/10.1101/39233.

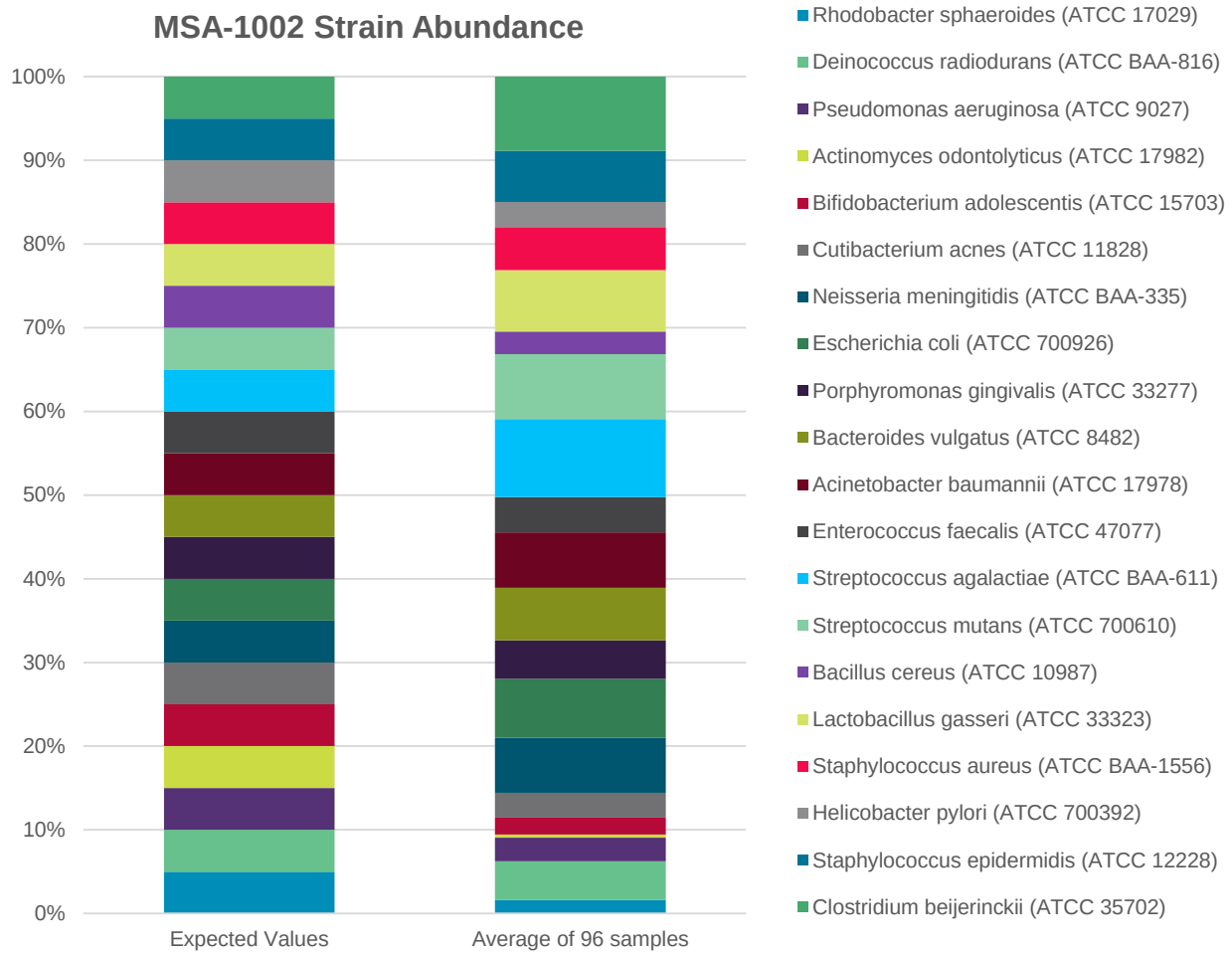
FULL-LENGTH 16S TYPICAL PERFORMANCE ON THE SEQUEL II SYSTEM

16S	>Q20 BC reads	>Q20 Read Quality	BC Samples / SMRT Cell	Avg reads / BC
MSA-1002	2,426,218	Q42	96	35,000
MSA-1003 (1a)	1,743,260	Q34	48	36,317
MSA-1003 (2a)	1,738,543	Q34	96	18,109

Full-length 16S sequencing and shotgun profiling data were collected for mock communities

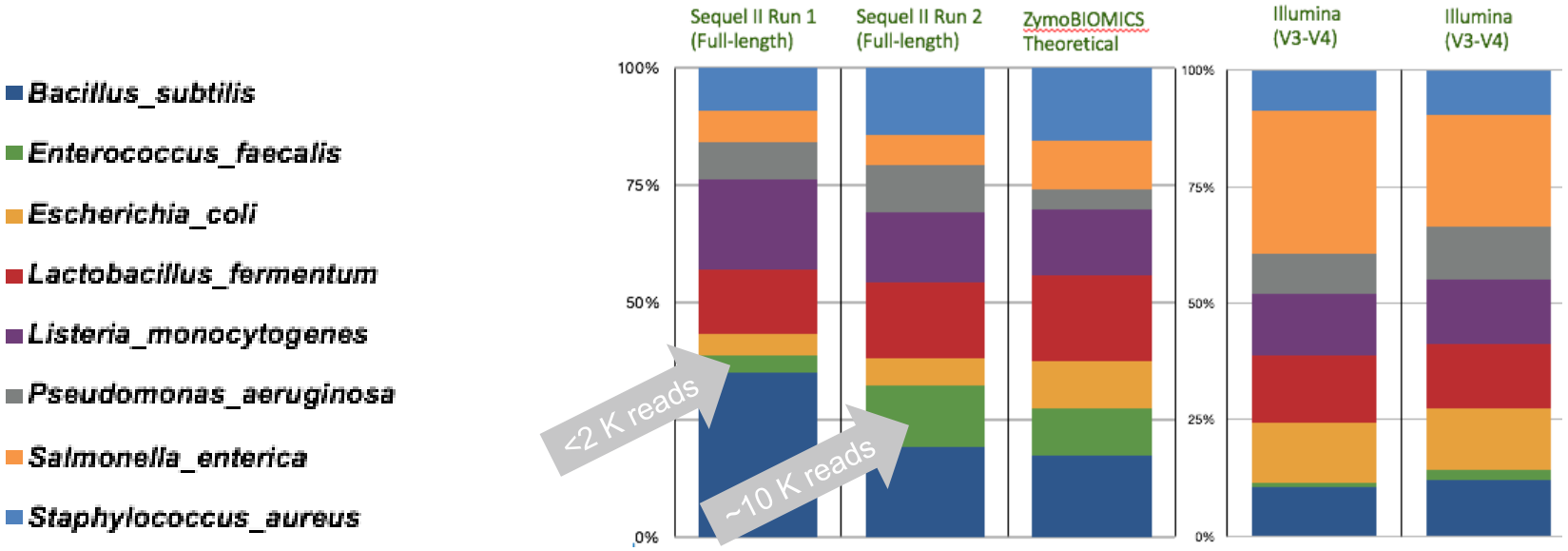
- ATCC 20 strain staggered (MSA-1002)
- ATCC 20 strain even (MSA-1003)

PACBIO FULL-LENGTH 16S SEQUENCING FAITHFULLY RECAPITULATES A MOCK COMMUNITY SAMPLE



V1-V9 amplicons were sequenced on a single SMRT Cell 8M at 96-plex

16S: UNBIASED TAXONOMIC REPRESENTATION COMPARING TO ZYMOBIOMICS MICROBIAL COMMUNITY STANDARD



- Full-length 16S sequencing on the Sequel II System allows you to more faithfully characterize the composition of your metagenome
- With our internal protocol, 48-plex runs reliably generate at least 10 K reads / barcode. At 96-plex, ~5% of barcodes drop below this floor.

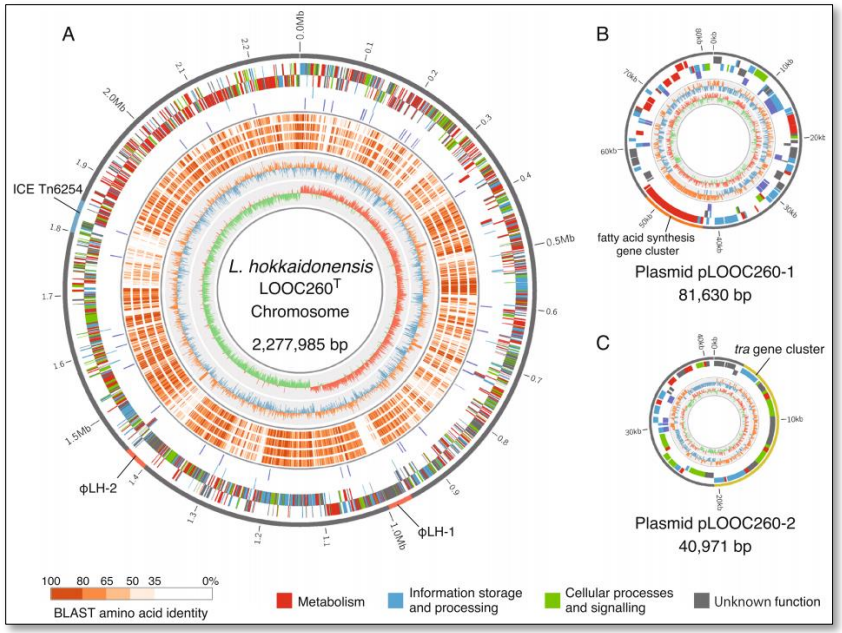
UNIVERSITY of MARYLAND SCHOOL OF MEDICINE INSTITUTE FOR GENOME SCIENCES

LONG READS + HIGH ACCURACY MEANS GENE DISCOVERY CAN BE DONE *DIRECTLY* ON HIFI READS, WITHOUT ASSEMBLY

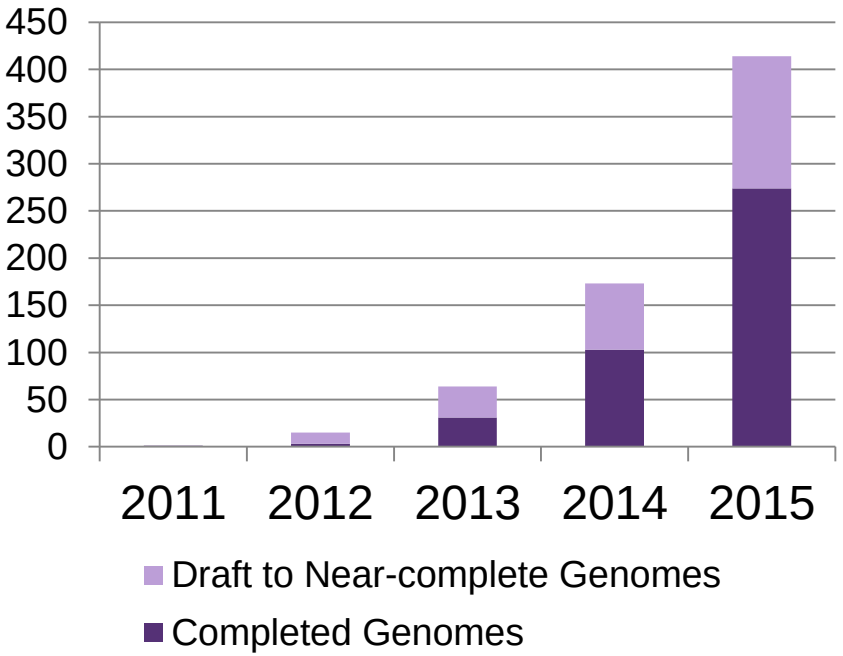
Sample	# HiFi Reads	# Amino acid sequences	Genes / read	Mean protein size (amino acids)
Human fecal 1	2,802,471	16,429,903	5.9	320.2
Human fecal 2	1,646,208	11,993,089	7.3	325.1
Human fecal 3	1,593,641	13,054,811	8.2	321.7

- FragGeneScan was used to predict proteins directly on unassembled HiFi reads
- Error-free genes can be found even from species with too little coverage for assembly
- High accuracy means existing NGS tools and pipelines can be used without modification

PACBIO IS THE NEW GOLD STANDARD FOR BACTERIAL GENOMICS

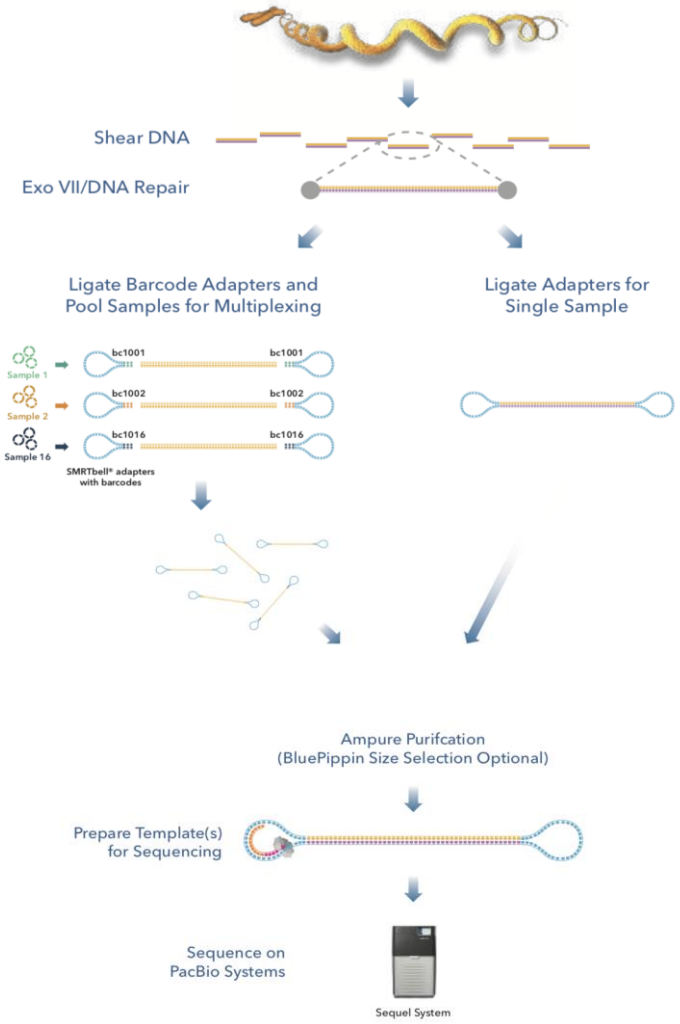


Growth in published microbial genomes completed with SMRT Sequencing



Tanizawa et al. (2015) Complete genome sequence and analysis of Lactobacillus hokkaidonensis LOOC260T, a psychrotrophic lactic acid bacterium isolated from silage. BMC Genomics 16: 240.

AFFORDABLY CHARACTERIZE COMPLETE MICROBIAL GENOMES ON THE SEQUEL PLATFORM



- Multiplex up to 30 Mb of microbial genomes / 16 samples on one SMRT Cell 1M
- Adjust planned multiplexing depth to balance cost constraints with your requirements for genome completeness
- Use our Microbial Multiplexing Calculator to simplify equimolar pooling
- Assemble most bacterial chromosomes into 5 contigs or fewer

ACHIEVE HIGHER COST EFFICIENCY AFTER BUILDING INITIAL EXPERIENCE WITH 30 MB POOLED LIBRARIES

Barcode ID	Sample ID	Pre-Assembly Yield (%)	GC Content (%)	Genome Size (bp)	Contigs (#)	Concordance w. Reference (QV)
BC1002	E. coli control 1	77	50.08	4,642,523	1	55.21
BC1015	E. coli control 2	77.9	50.08	4,642,523	1	56.25
BC1004	B. sub control 1	76.4	43.94	4,045,593	1	51.16
BC1016	B. sub control 2	74.9	43.94	4,045,593	1	51.45
BC1009	E. coli 1	77.4	50.08	4,642,523	1	54.63
BC1018	E. coli 2	77.7	50.08	4,642,523	1	54.91
BC1012	S. sonnei 1	76.4	51.03	4,813,450	1	59.83
BC1020	S. sonnei 2	78.1	51.03	4,813,450	1	54.27
BC1014	L. monocytogenes 1	71.8	37.94	3,032,269	1	53.06
BC1022	L. monocytogenes 2	73.2	37.94	3,032,269	1	49.63

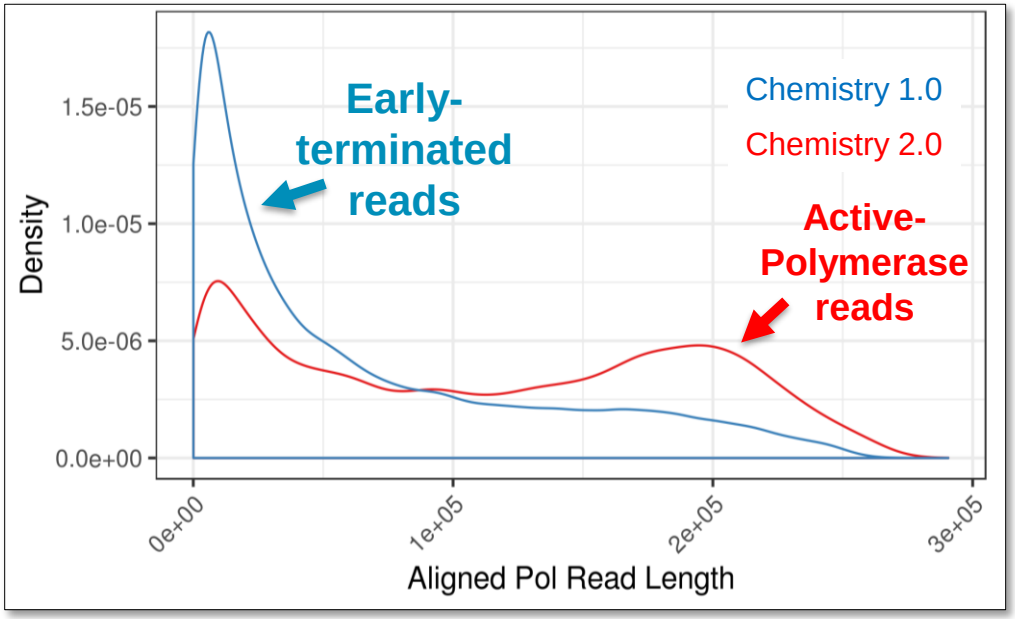
- 10-plex run of 42.4 MB total microbial genomes shown
- Combined use of Barcoded Adapter Kit 8A & Barcoded Adapter Kit 8B
- Less than 5 contig assemblies for main chromosomal genomes can be achieved with Advanced HGAP parameters
- QV50 is roughly 99.999% base accuracy

PACBIO®

Sequel II System Updates

SEQUEL II CHEMISTRY 2.0 (EARLY ACCESS (EA) INITIATED END OF JUNE)

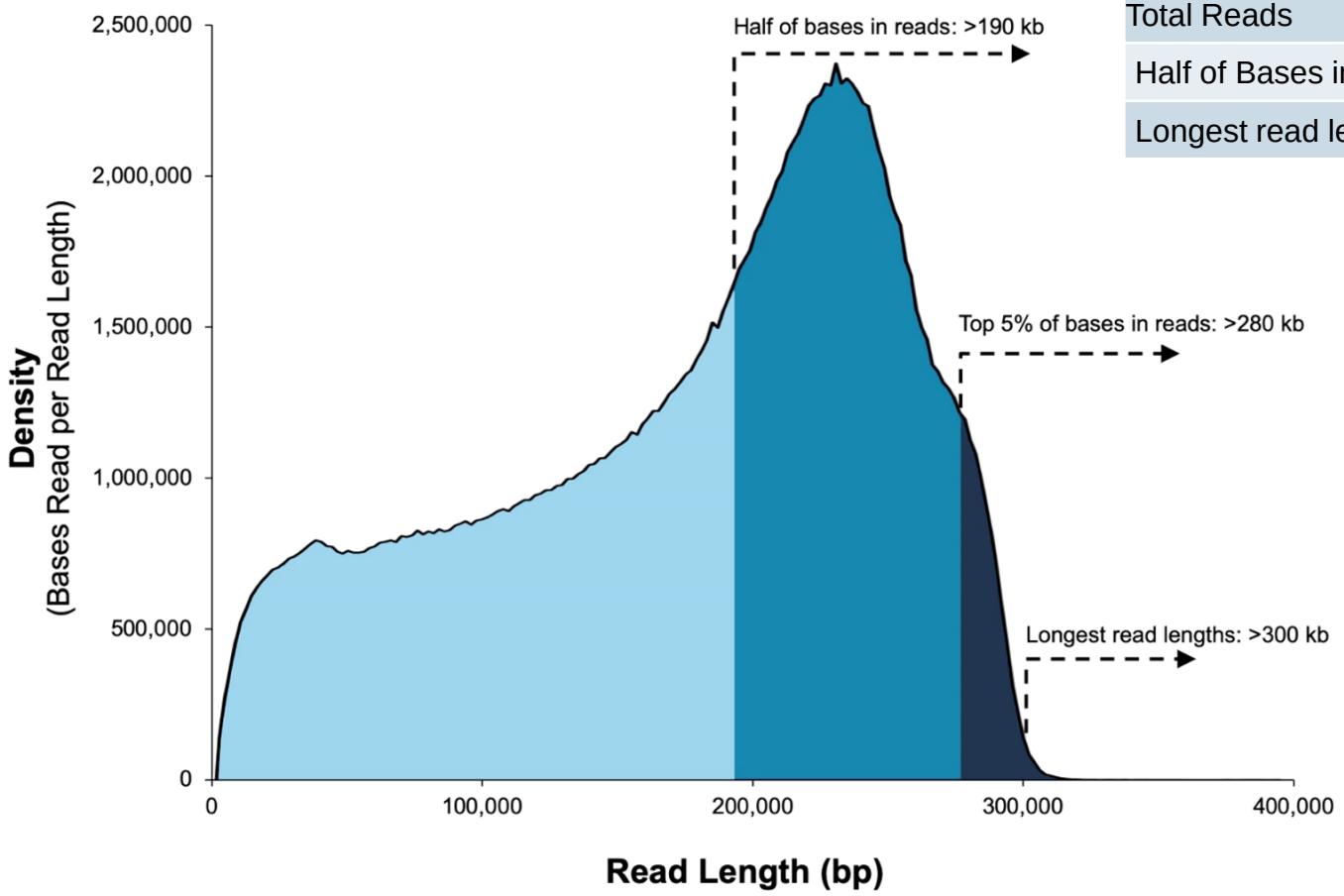
- Higher yield of long HiFi reads due to greater efficiency of getting reads into rolling circle synthesis and continuing to the end of the movie
- Translates to ~25-50% average RL increase and ~40% higher yield of reads > Q20 and Q30
- Also allows increase in insert size from ~11-13 to 15-20+ kb, depending on application



Site	Library Name	Sample Biology	Insert Size (Kb)	[Loading]	Yield (Gb)	Pol RL	Subread	P0	P1	Base Rate	≥Q20 Yield (Gb)	≥ Q20 Reads (count)	≥ Q20 Read Length (mean, bp)	≥ Q20 Read Quality (median)	Percent Reads ≥ Q20	UMY (Gb)
Site 3 *	GM19734	Human	16.1	50pM	384.9	98259	11162	49%	49%	1.97	**51.2**	4,643,311	11029	Q36		
	GM19734	Human	16.1	60pM	437.3	96301	11595	41%	57%	1.95						
	GM19030	Human	22.8	60pM	380.8	93318	10222	47%	51%	1.97	**47.1**	4629362	10170	Q36		
	GM19030	Human	22.8	60pM	414.7	91150	10376	40%	57%	1.94						
	HG03736	Human	19.1	60pM	388.3	87645	12765	42%	55%	1.90	**52.2**	3,916,231	13332	Q33		
	HG03736	Human	19.1	60pM	367.8	85863	12833	44%	53%	1.86						

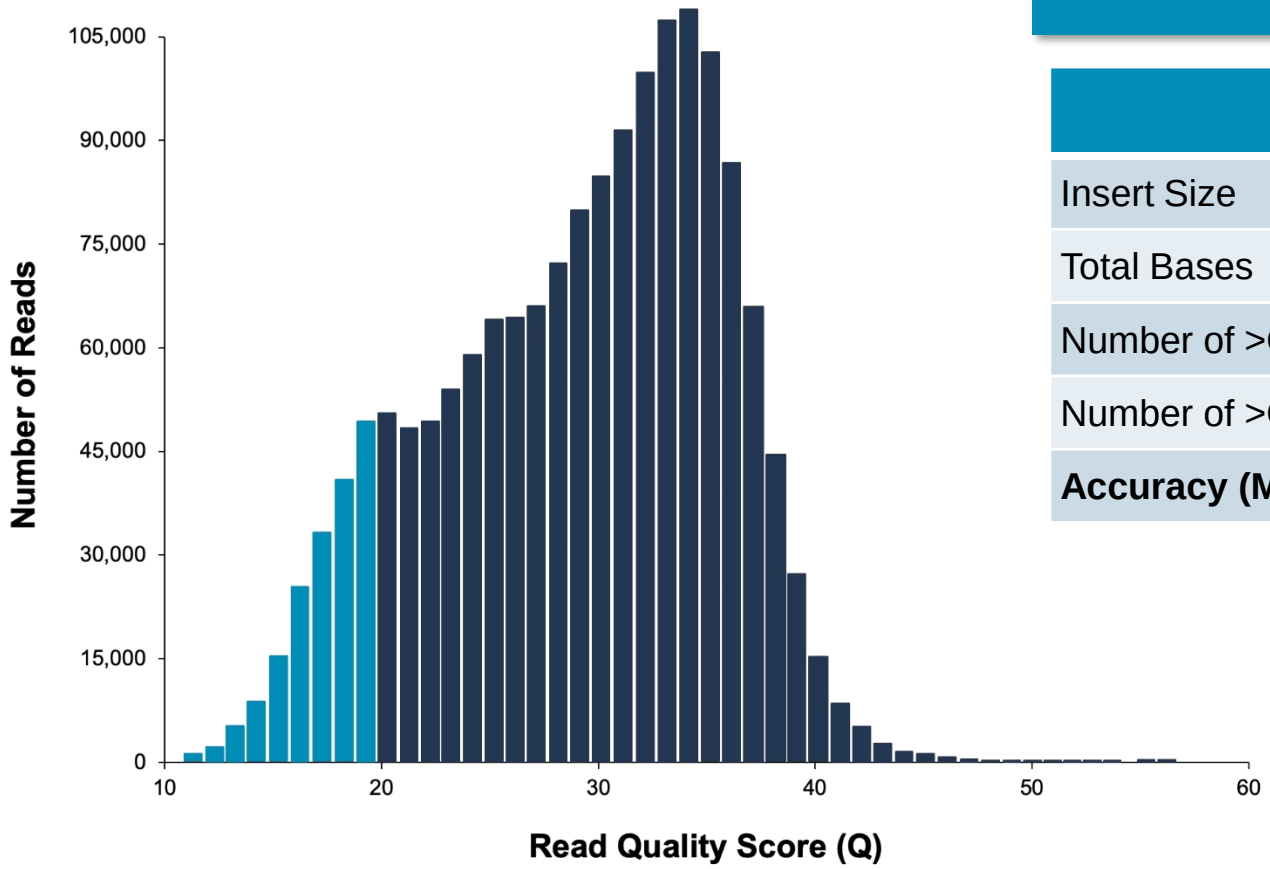
HIFI SEQUENCING PERFORMANCE

METRIC	
Insert Size	20 kb
Number of Raw Bases (Gb)	377
Total Reads	4,266,403
Half of Bases in Reads	>193,403
Longest read lengths	>300,000



Data shown above from a 20 kb size-selected human library using the SMRTbell Template Prep Kit on a Sequel II System (2.0 Chemistry, Sequel II System Software v8.0, 30-hour movie). Read lengths, reads/data per SMRT Cell 8M and other sequencing performance results vary based on sample quality/type and insert size.

HIFI SEQUENCING PERFORMANCE



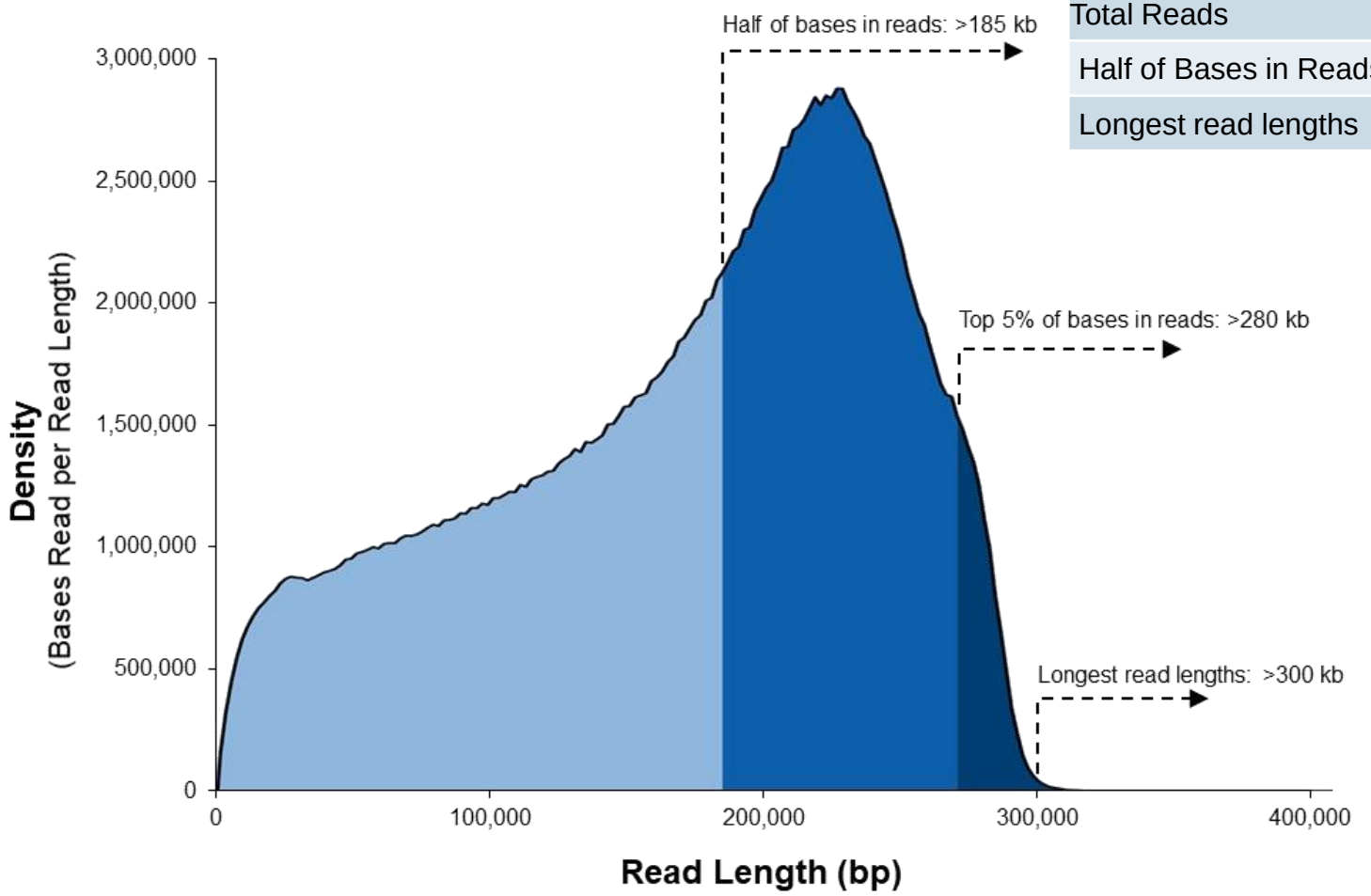
Q20 (99%) single-molecule accuracy

Metrics	
Insert Size	20 kb
Total Bases	30 Gb
Number of >Q20 Bases	26 Gb
Number of >Q20 Reads	1,423,277
Accuracy (Mean)	**99.92%**

Data shown above from a 20 kb size-selected human library using the SMRTbell Template Prep Kit on a Sequel II System (2.0 Chemistry, Sequel II System Software v8.0, 30-hour movie). Read lengths, reads/data per SMRT Cell 8M and other sequencing performance results vary based on sample quality/type and insert size.

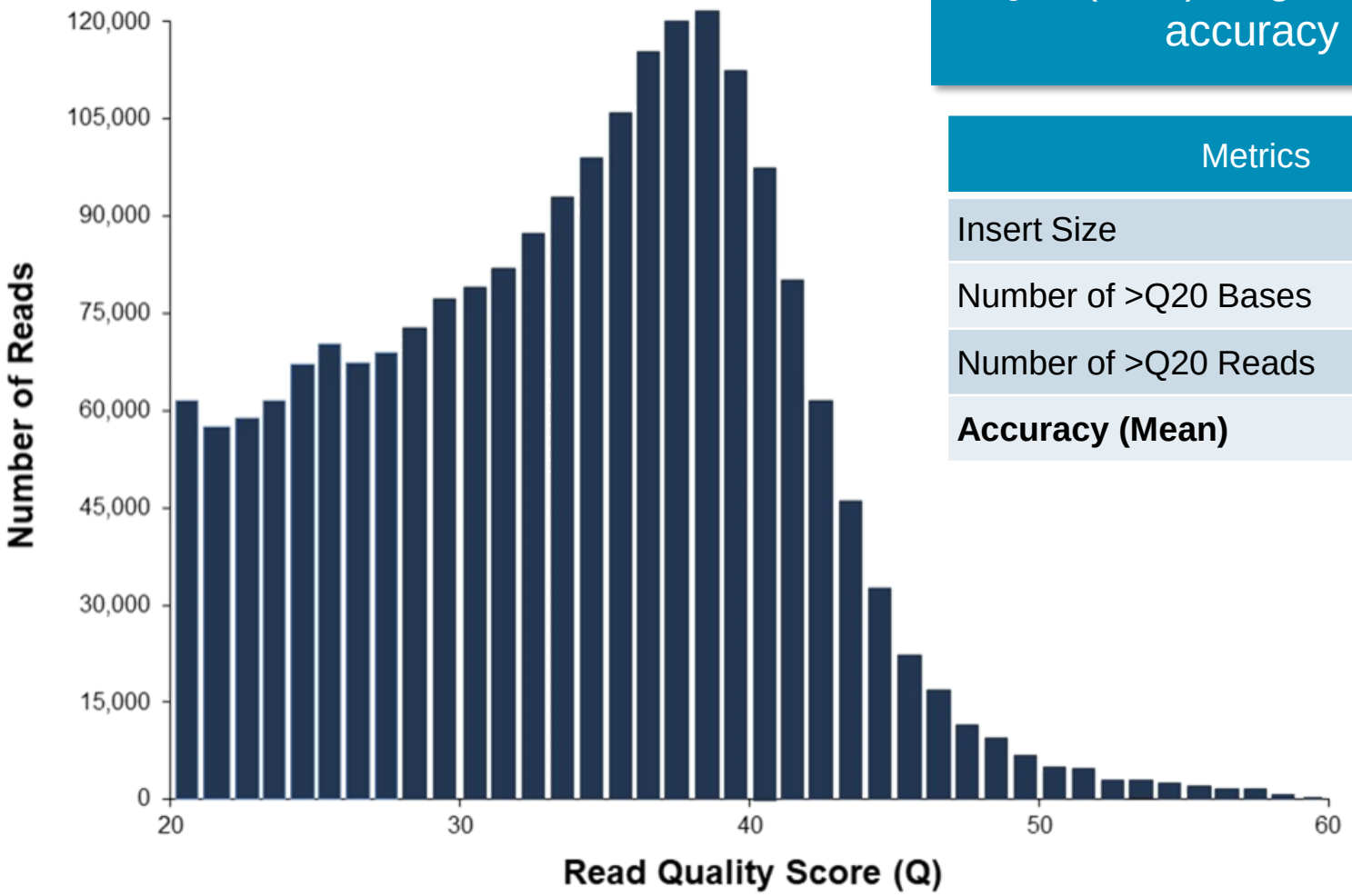
HIFI SEQUENCING PERFORMANCE

METRIC	
Insert Size	15 kb
Number of Raw Bases (Gb)	449
Total Reads	4,976,882
Half of Bases in Reads	>186,554
Longest read lengths	>300,000



Data shown above from a 15 kb size-selected human library using the SMRTbell Template Prep Kit on a Sequel II System (2.0 Chemistry, Sequel II System Software v8.0, 30-hour movie). Read lengths, reads/data per SMRT Cell 8M and other sequencing performance results vary based on sample quality/type and insert size.

HIFI SEQUENCING PERFORMANCE

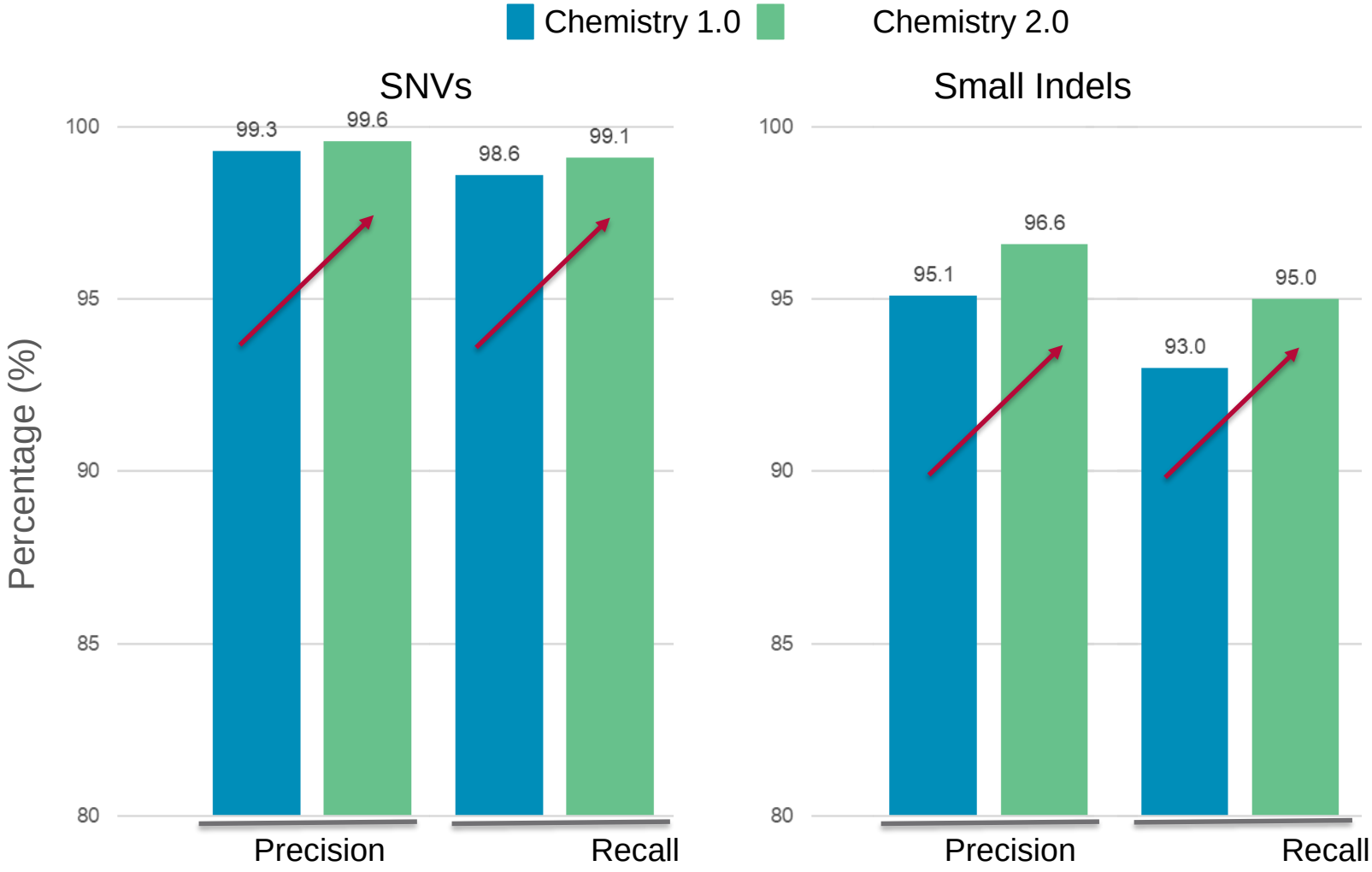


Q20 (99%) single-molecule accuracy

Metrics	
Insert Size	15 kb
Number of >Q20 Bases	27.4 Gb
Number of >Q20 Reads	2,099,017
Accuracy (Mean)	**99.96%**

Data shown above from a 15 kb size-selected human library using the SMRTbell Template Prep Kit on a Sequel II System (2.0 Chemistry, Sequel II System Software v8.0, 30-hour movie). Read lengths, reads/data per SMRT Cell 8M and other sequencing performance results vary based on sample quality/type and insert size.

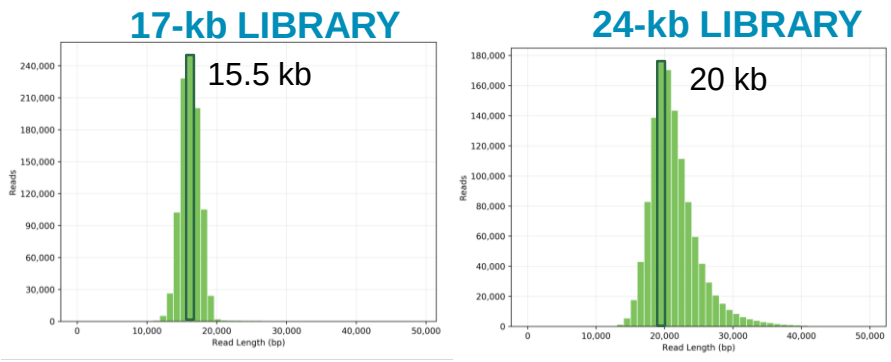
USING HIFI READS FOR VARIANT DISCOVERY

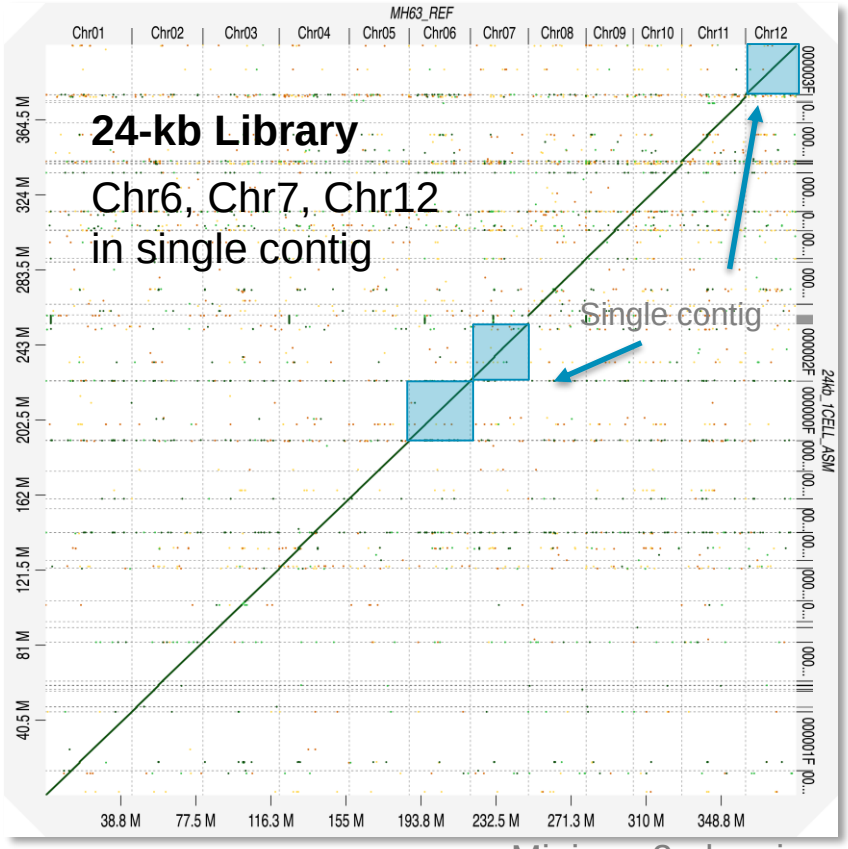


Variant calls from ~15-fold HiFi read coverage of a human genome (HG002) were measured against the Genome in a Bottle small variant benchmark (v3.3.2) for SNVs and indels using Deep Variant. Libraries were generated using an 11 kb insert for Chemistry 1.0 and 15 kb insert for Chemistry 2.0.

DE NOVO ASSEMBLY WITH HIFI READS – RICE GENOME

Library Size	17 kb	24 kb
Raw Yield	218 Gb	366 Gb
1 SMRT Cell Yield	15 Gb	25 Gb
Median Read Length	16 kb	21 kb
Coverage (400Mb)	38-fold	63-fold
FALCON Asm Length	403 Mb	405 Mb
N Contigs	209	211
Contig N50	14 Mb	20 Mb
N Chrom in 1 Contig	0	3





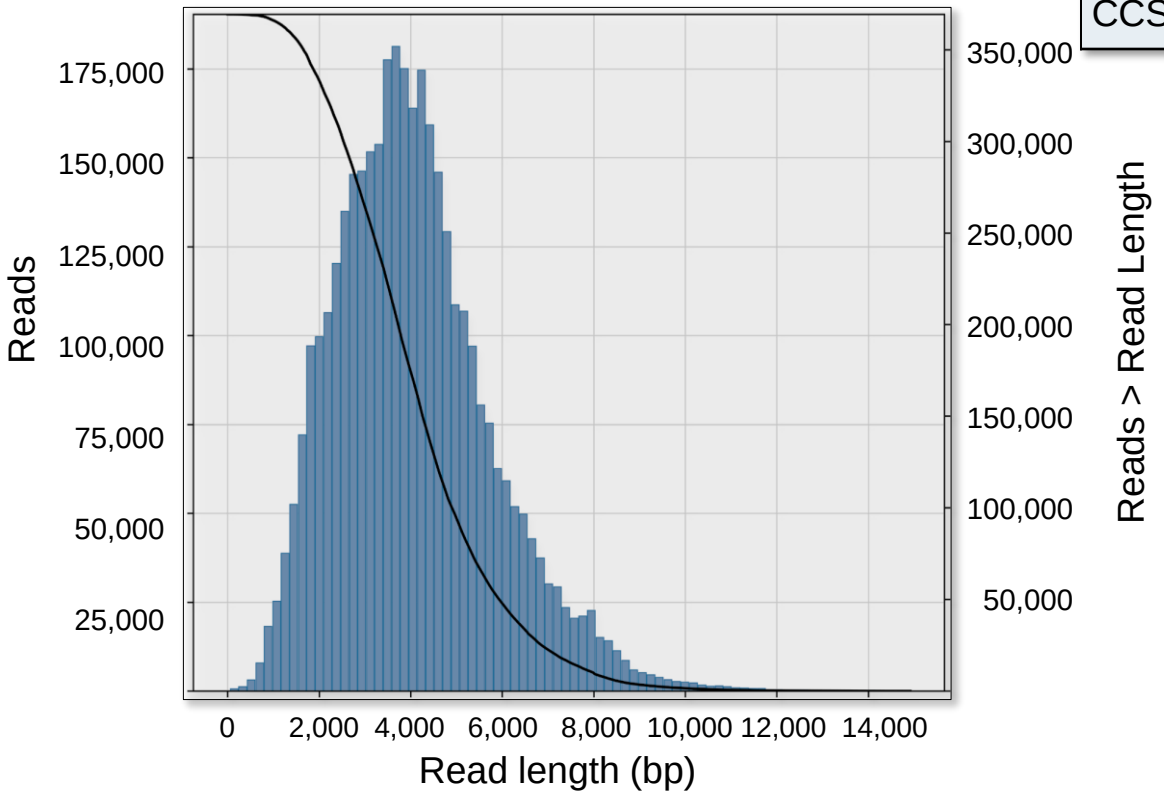
Minimap2, dgenies

Data shown above from a 17 kb & 24 kb size-selected rice library using the SMRTbell Template Prep Kit on a Sequel II System (2.0 Chemistry, Sequel II System Software v8.0, 30-hour movie). Falcon assembly was performed post CCS analysis. Read lengths, reads/data per SMRT Cell 8M and other sequencing performance results vary based on sample quality/type and insert size.

ISO-SEQ PERFORMANCE

Metric	
Number of Raw Bases (Gb)	502
Total Reads	5,433,706
Full Length Non-chimeric Reads	3,693,801 (86%)
CCS Passes (Mean)	8

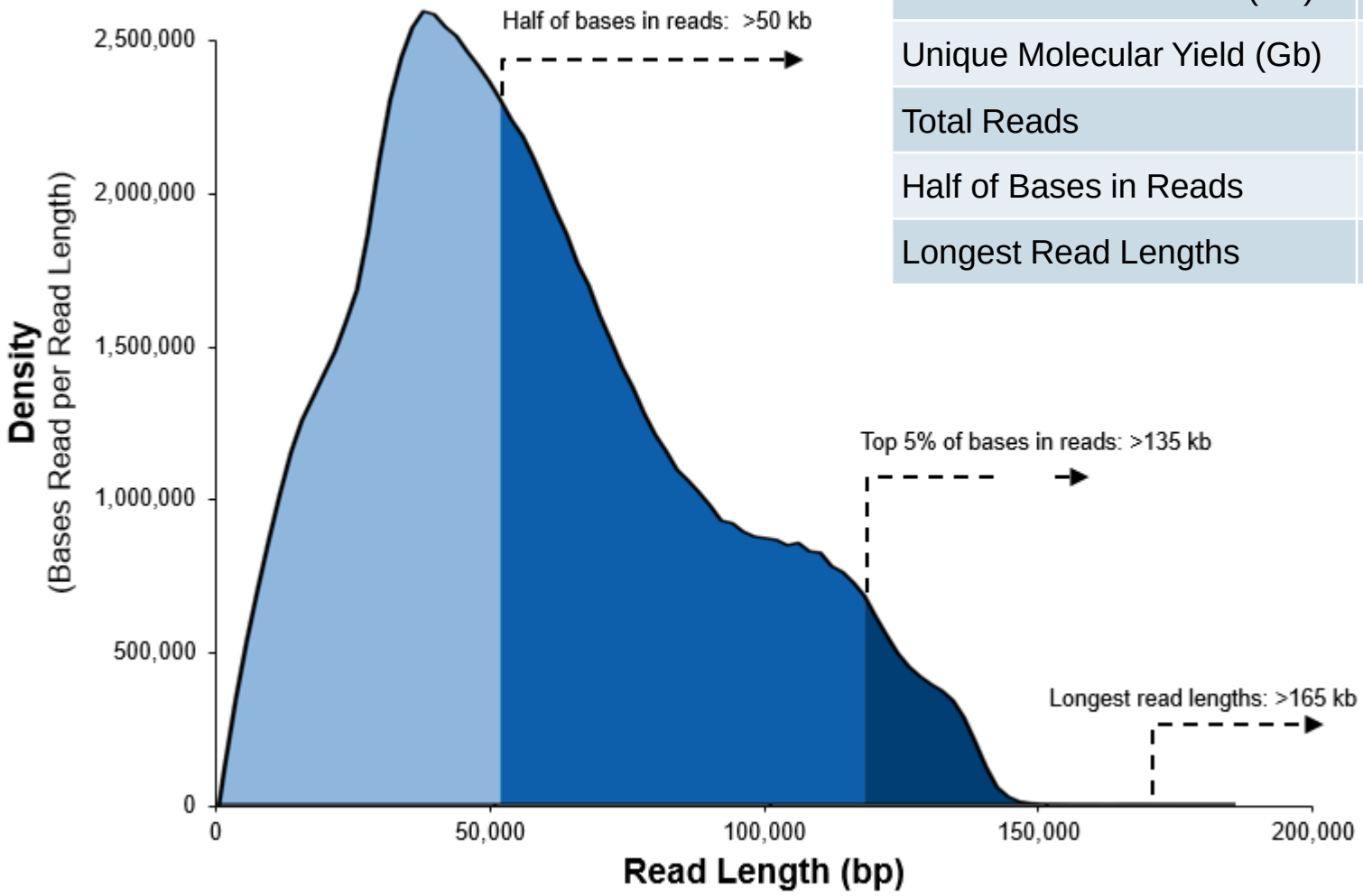
- Iso-Seq: Universal Human Reference RNA



Data shown above from a Universal Human Reference RNA (human) and Lexogen SIRV spike-in controls. The library was constructed using the Iso-Seq Express workflow including the SMRTbell Express Template Prep Kit 2.0 on a Sequel II System (Sequel II Sequencing kit 2.0, Sequel II Binding Kit 2.0, Sequel II System Software v8.0, 24-hour movie). Read lengths, reads/data per SMRT Cell 8M and other sequencing performance results vary based on sample quality/type and insert size.

LONG READ SEQUENCING

METRIC	
Number of Raw Bases (Gb)	182
Unique Molecular Yield (Gb)	161
Total Reads	5,210,363
Half of Bases in Reads	>52,456
Longest Read Lengths	>165,000



Data shown above from a 35 kb size-selected *E. coli* library using the SMRTbell Template Prep Kit on a Sequel II System (2.0 Chemistry, Sequel II System Software v8.0, 15-hour movie). Read lengths, reads/data per SMRT Cell 8M and other sequencing performance results vary based on sample quality/type and insert size.

PACBIO®

SMRT Link v8.0 Preview

SMRT LINK V8.0 – SUMMARY OVERVIEW OF MAJOR IMPROVEMENTS

New Workflow Engine - Cromwell

- Scalable, widely adopted in scientific environments, developed by Broad

Further Enhancements for CCS Analysis

- Faster CCS analysis (4-6 times faster time-to-results)

Analysis Applications

- **NEW**: Microbial Assembly application
- *de novo* assembly for CCS data – Falcon Unzip (Bioconda)

UI and Usability improvements

- New style GUI – same functionality
- Support for IGV visualization
- Analysis restart from the point of failure

ANALYSIS APPLICATIONS

Microbial Assembly – NEW application

- *de novo* assembly for small genomes between 1.9-10 Mb
- Circularization
- Plasmid assembly (2 – 220 kb)

Iso-Seq Analysis

- Analysis from CCS data

de novo assembly with HiFi reads - Bioconda

- Direct CCS (FASTA+FASTQ) input to FALCON and FALCON-Unzip
- Accurate phasing - good phased contig QV approaching Q50
- Faster read tracking and polishing using CCS reads

COMPATIBILITY AND SUPPORT

SMRT Link v8.0 supports Sequel and Sequel II Systems only

- Use SMRT Link v7.0 or earlier for PacBio RS II data

Data Visualization

- IGV compatibility - URL links to alignment results (BAM and BAI)
- SMRT View is no longer included with SMRT Link

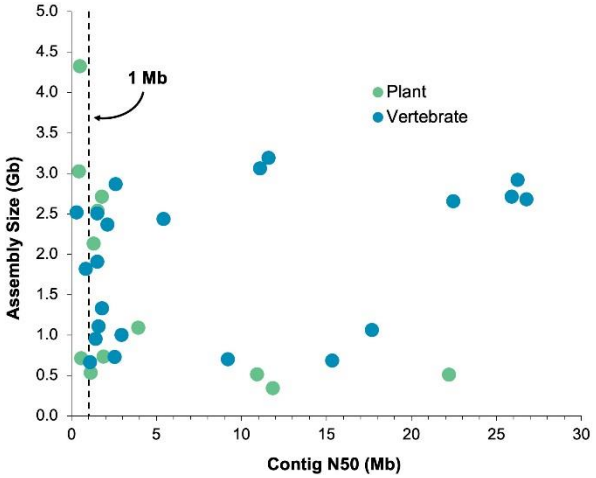
SMRT Link v8.0 migration

- This is a major release – more efforts for customers to upgrade
  - Relatively minimal impact on UI users, cmd and APIs users more impacted
- **Sept.11, 2019** - **SMRT Link v8.0 notification email - details on upcoming changes**

WHAT COULD YOU DO WITH A SEQUEL II SYSTEM TODAY?

With a single SMRT Cell 8M:

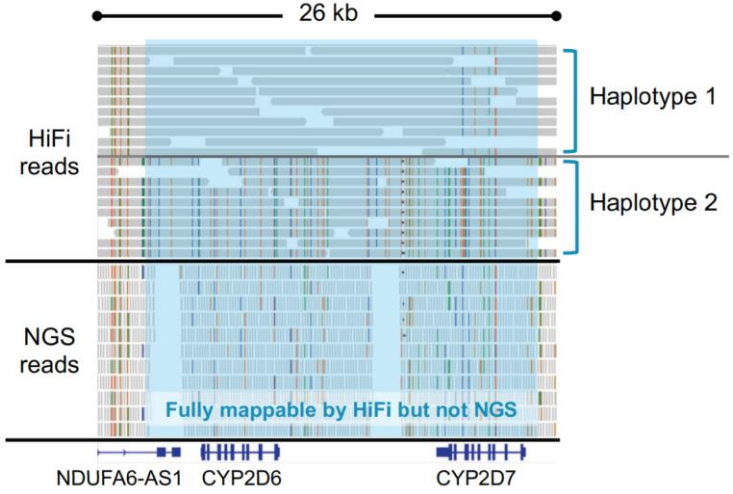
- Generate a 2 Gb genome assembly
- Call structural variants in a human genome
- Sequence a whole transcriptome
- Determine the composition of a >90 gut microbiome samples



Plant and vertebrate genome assemblies generated with SMRT Sequencing data from the genome database on NCBI showing PacBio assemblies readily achieve contig N50s ≥1 Mb.

With 2-3 SMRT Cells 8M:

- Detect all variants in a human genome
- Phase a diploid assembly of a human genome



Coverage of the Cytochrome P450 2D6 (CYP2D6) and CYP2D7 genes with HiFi reads and NGS reads visualized in IGV. CYP2D6 is responsible for the metabolism and elimination of approximately 25% of clinically used drugs.

www.pacb.com