

REVEALING SMRT™ BIOLOGY

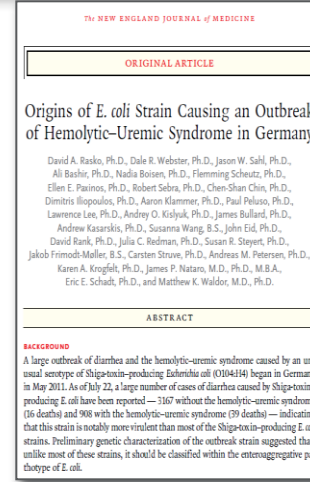
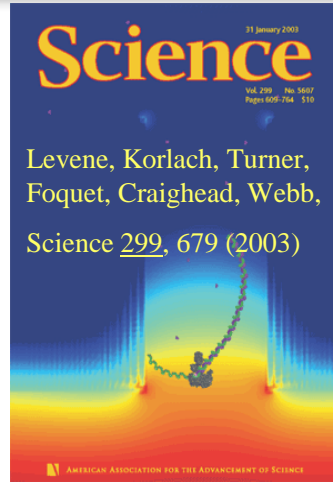
PacBio RS II

技术简介及原理介绍

FU Wei

Senior Application Specialist

History of Pacific Biosciences



2011 Max Delbrück
德国分子医学中心 (MDC)
Korf & 默克尔总理

Menlo Park, CA, USA

Cornell University
Steve Turner &
Jonas Korf

2003 First Report
First
system
sold
2009

2011 IF>50

Company
IPO

Company
founded
2000

Proof of
concept
2007

Venture
funded
2004

Company
"launch"
2008

LPR
2010

FCR
2011

14 Years, \$370 Million, 400 people, customers>70

PACIFIC BIOSCIENCES™ CONFIDENTIAL

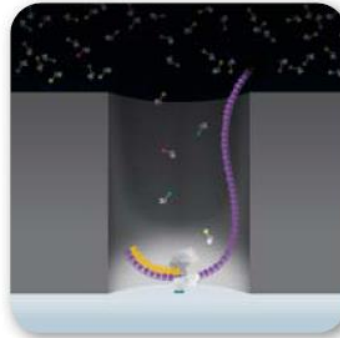
PacBio 公司被美国麻省理工学院(MIT)的《Technology Review》杂志评为2010年度全球50家最具创新力的企业之一

Single Molecule, Real-Time (SMRT®) DNA Sequencing

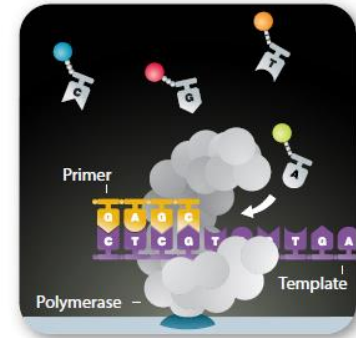
SMRT® Cells



Zero-Mode Waveguides



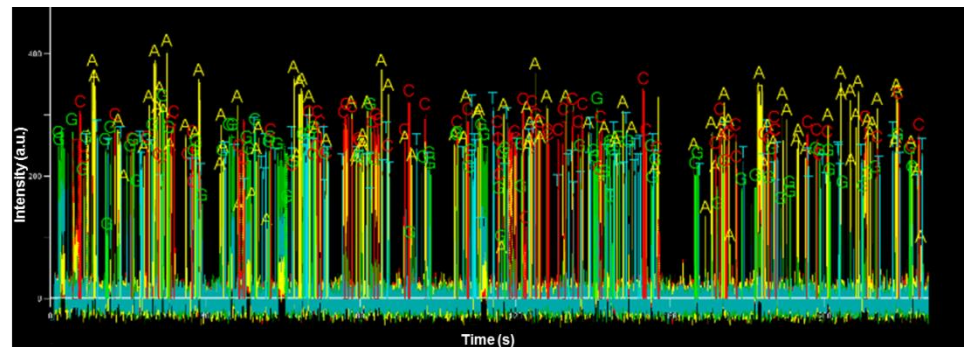
Phospholinked Nucleotides



PacBio® RS II



Trace



PACIFIC BIOSCIENCES™ CONFIDENTIAL

(Science 299,682)

PacBio RS——革命性的第三代测序平台

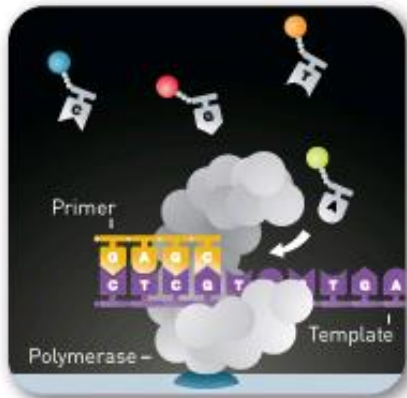


SMRT™ Cell

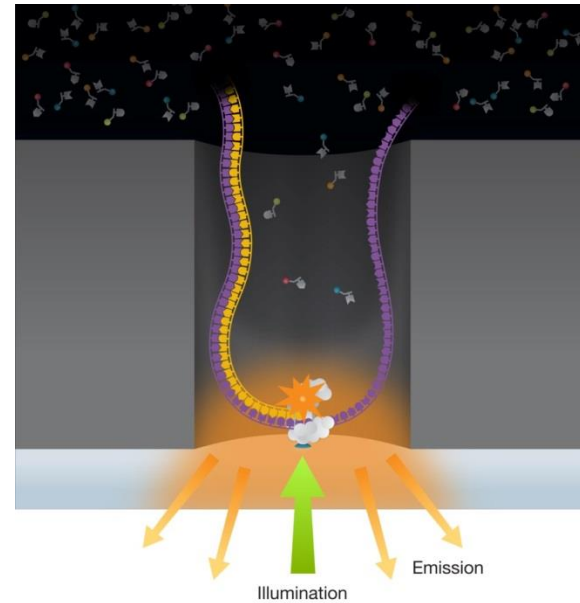
Single Molecule, Real-Time (SMRT®) DNA Sequencing



SMRT® Cells

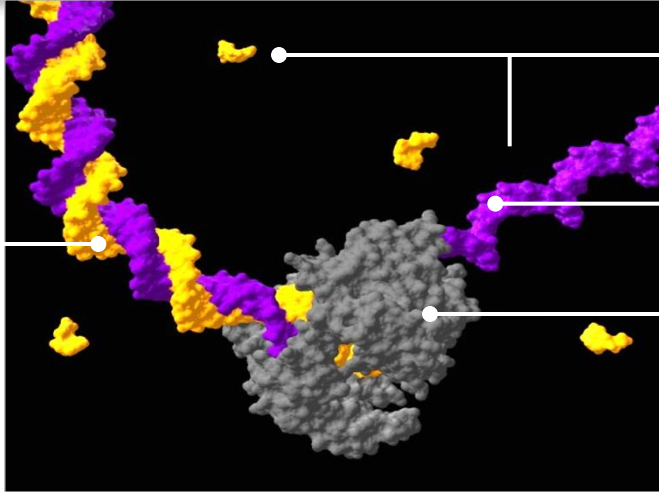


Phospholinked Nucleotides



Science, Vol 299, Jan 31 2003, pp682-686
J. Appl. Phys. 103, 034301 (2008)

DNA Polymerase



Primer

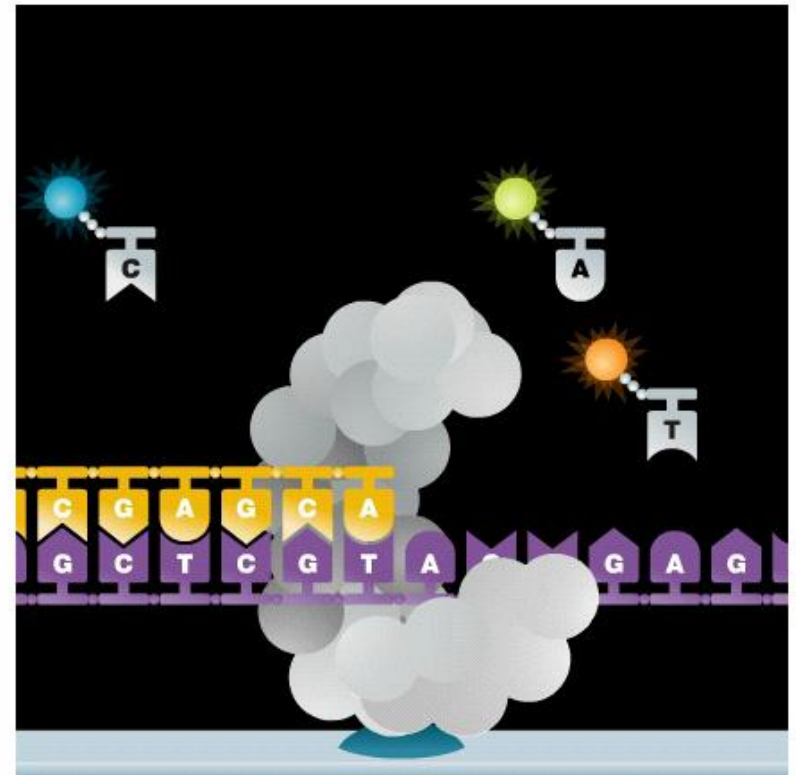
Nucleotides

Template

DNA polymerase

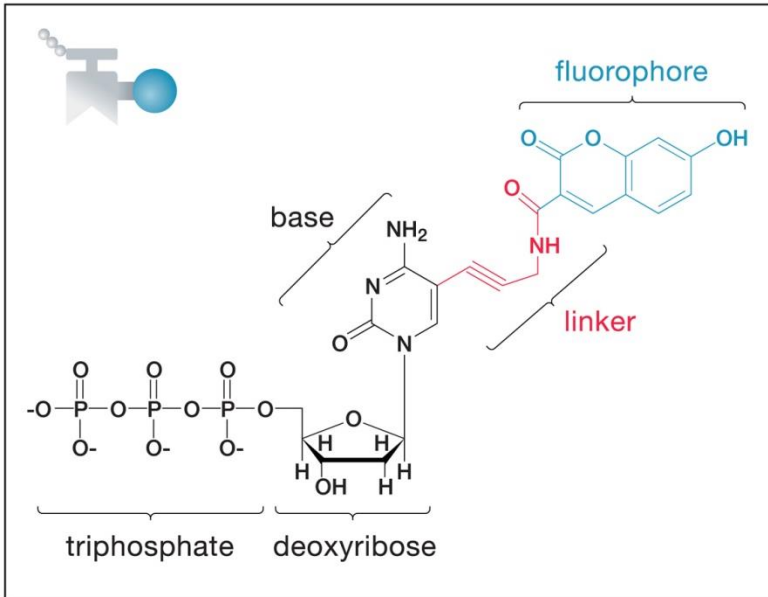
DNA聚合酶是序列合成的引擎

- Fast: 750 bp/s
- Processive
- Frugal
- Faithful: 1 in 10^5
- Small: 15 nm



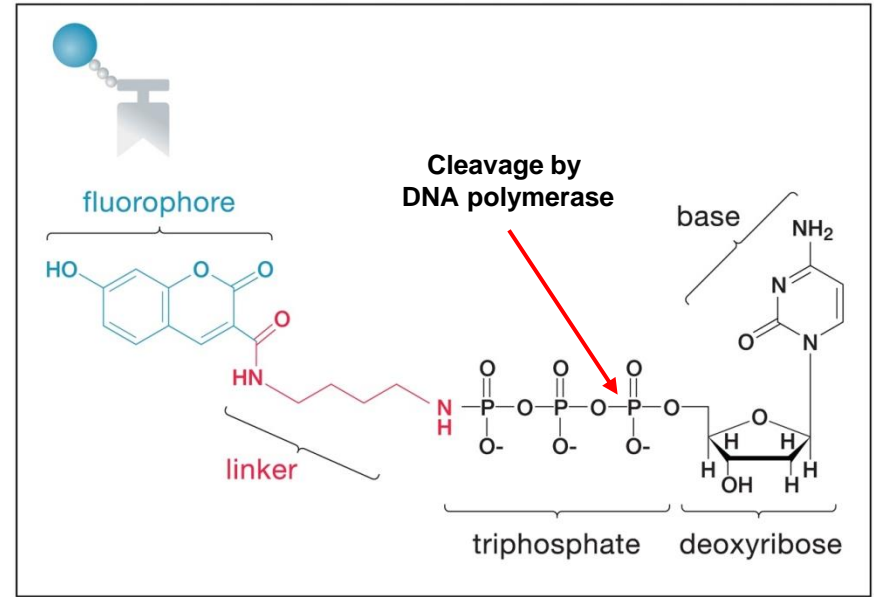
新型核苷酸标记方法，可维持DNA聚合酶的高活性

Base-linked



- Fluorophore stays in DNA
- Inhibits enzyme
- Creates background light

Phospho-linked



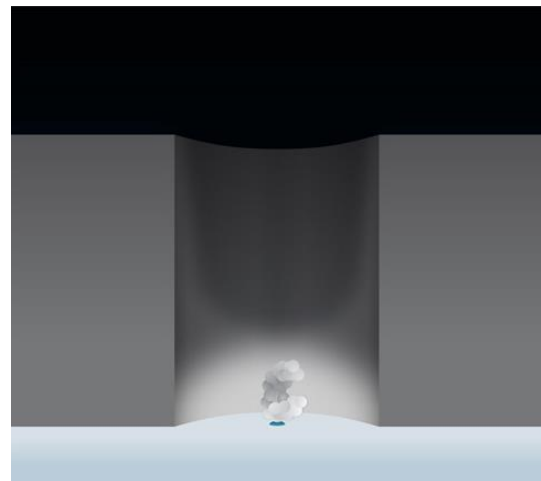
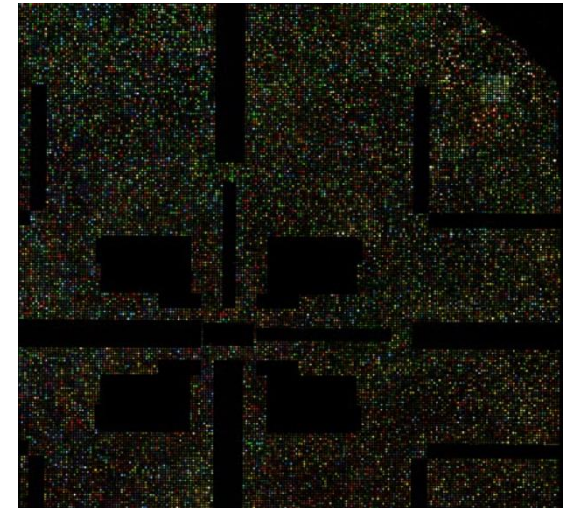
- Fluorophore clipped off by polymerase
- DNA synthesized is natural
- No clogging, background decay

SMRT Cell

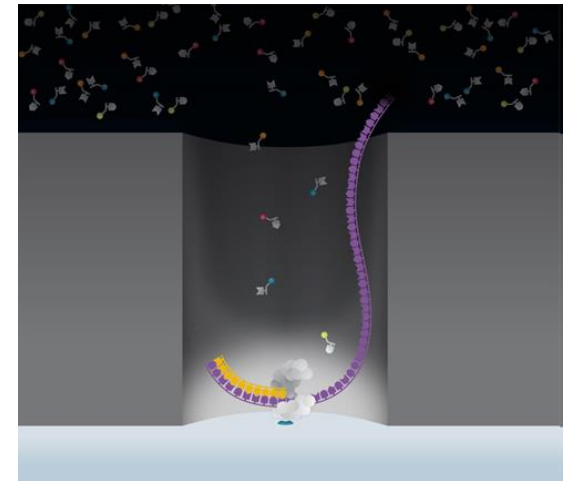
- 150,000 ZMWs (zero-mode waveguide, 零模波导孔) in each SMRT Cell
- approx. 33% of ZMWs have one and only one polymerase (>50,000 ZMWs give valid data)
- Each ZMW provides a window that enables the real-time observation of a single molecule of DNA polymerase
 - 在基质表面锚定单个DNA聚合酶分子以使DNA合成



SMRT cell

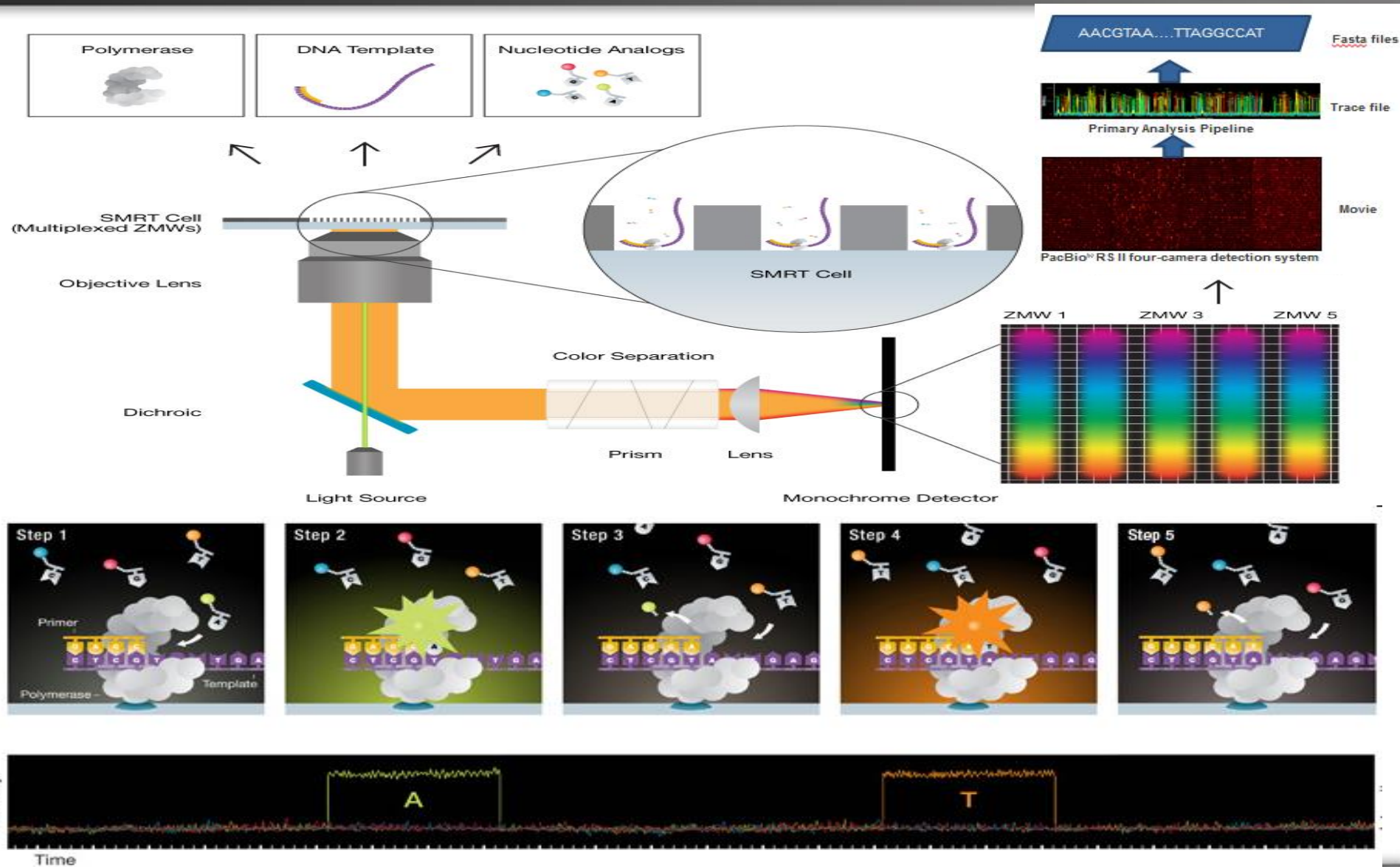


ZMW with DNA polymerase



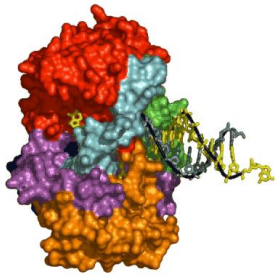
ZMW with DNA polymerase and phospholinked nucleotides
(Science 299,682)

在单个ZMW中，DNA合成期间可检测到单个核苷酸的掺入

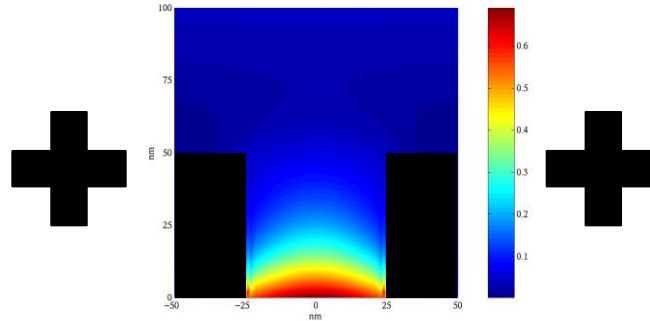


Single-Molecule, Real-Time DNA Sequencing Is:

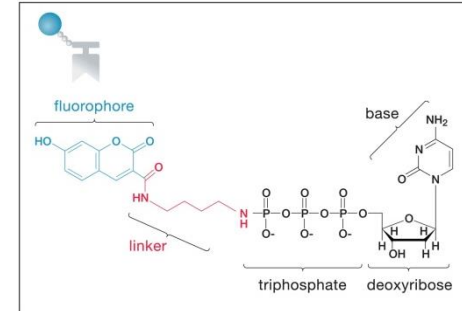
DNA Polymerase



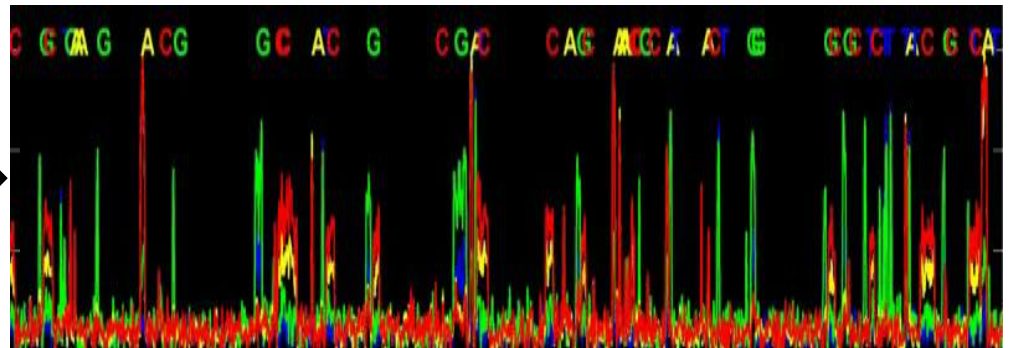
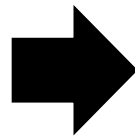
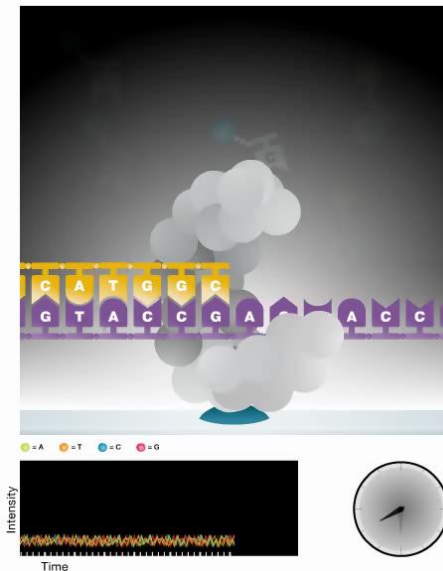
ZMW Confinement



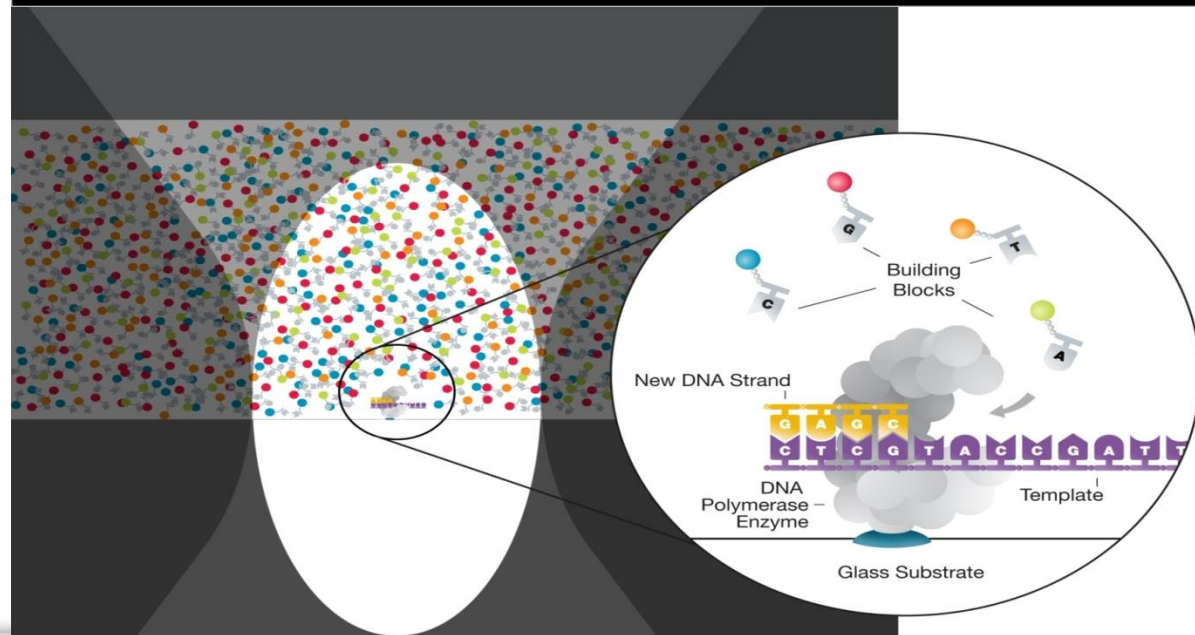
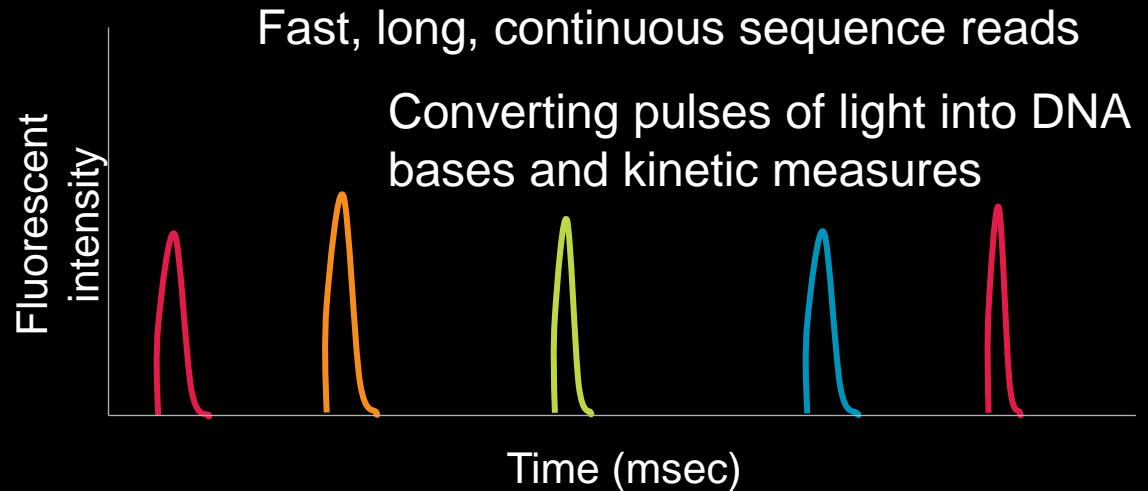
荧光基团标记在磷酸链上



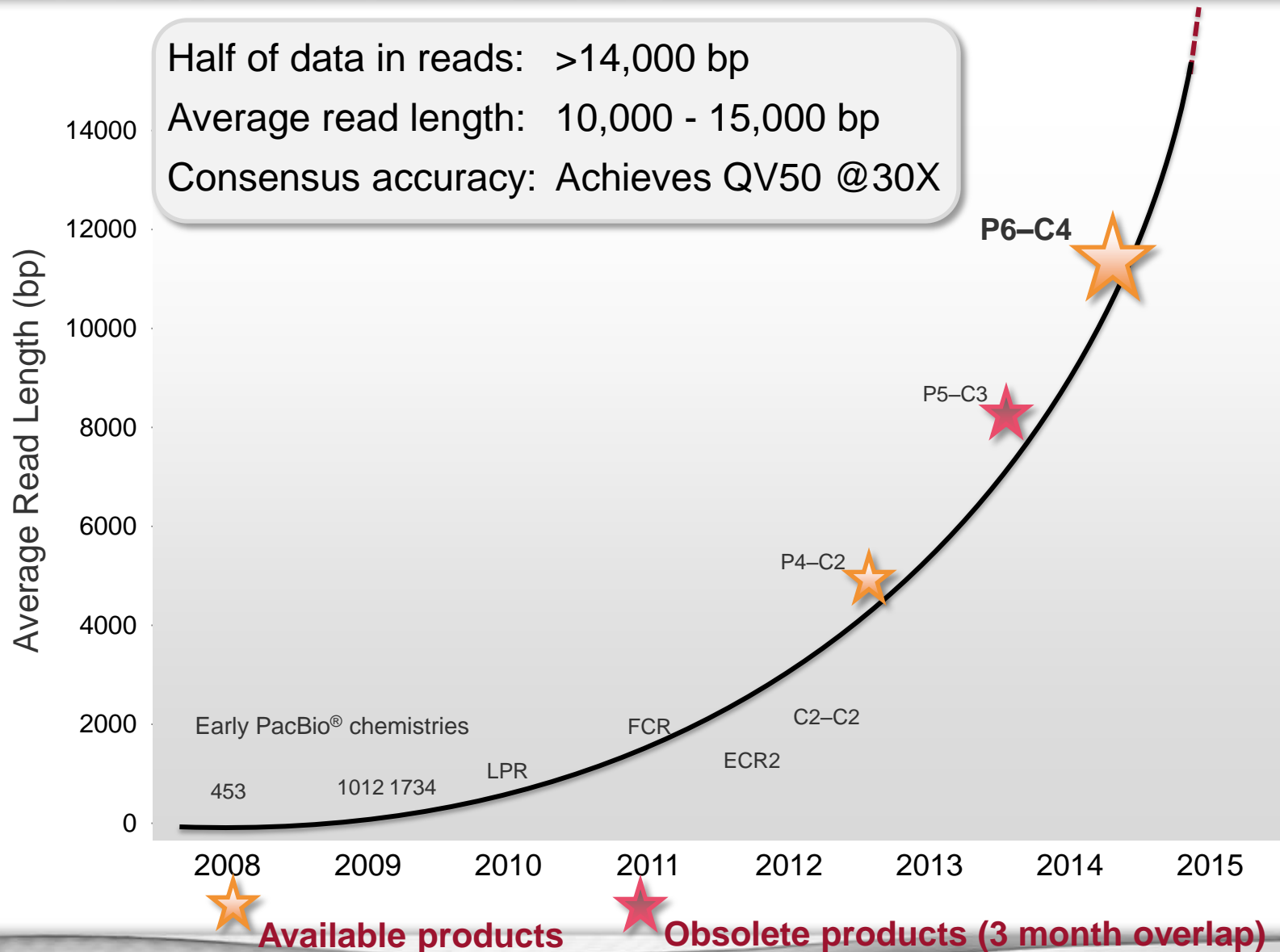
=



Single Molecule Real Time

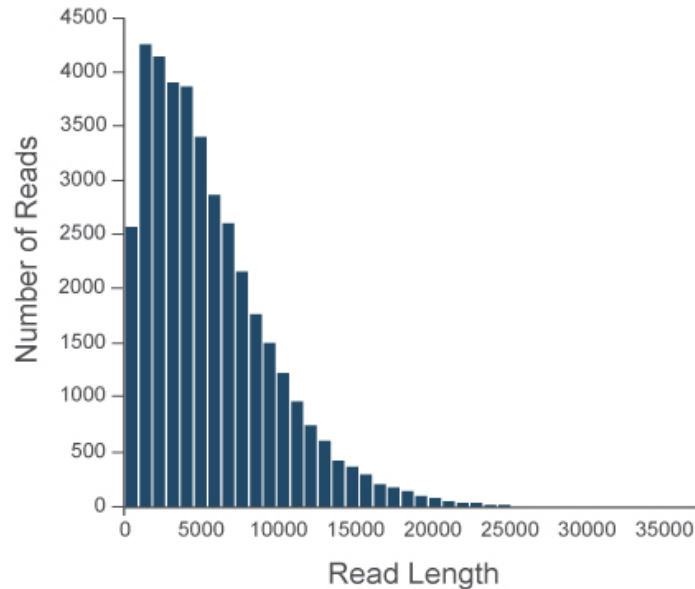


NEW! P6-C4 Chemistry



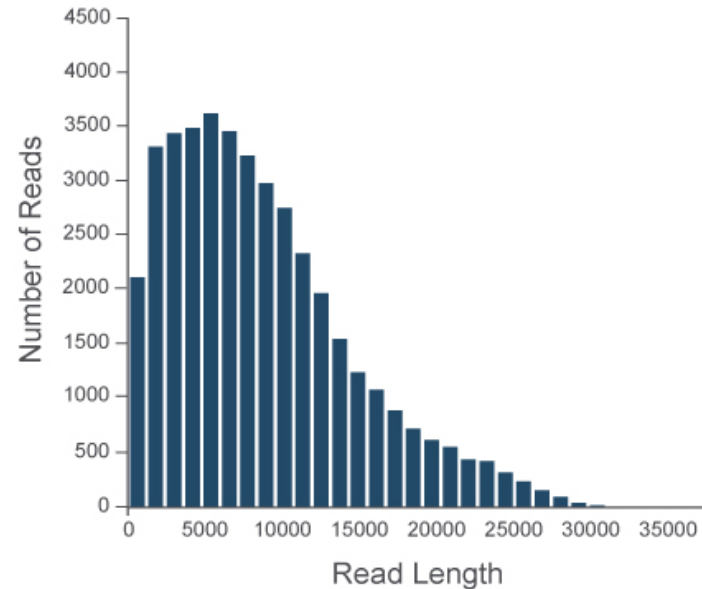
Typical PacBio® RS II Performance

P4-C2 Chemistry



Average: ~ 5.5 kb
Maximum: > 24 kb
Top 5% of reads: > 11 kb
Half of data in reads: > 8 kb
Data per SMRT® Cell: ~ 275 Mb

P5-C3 Chemistry

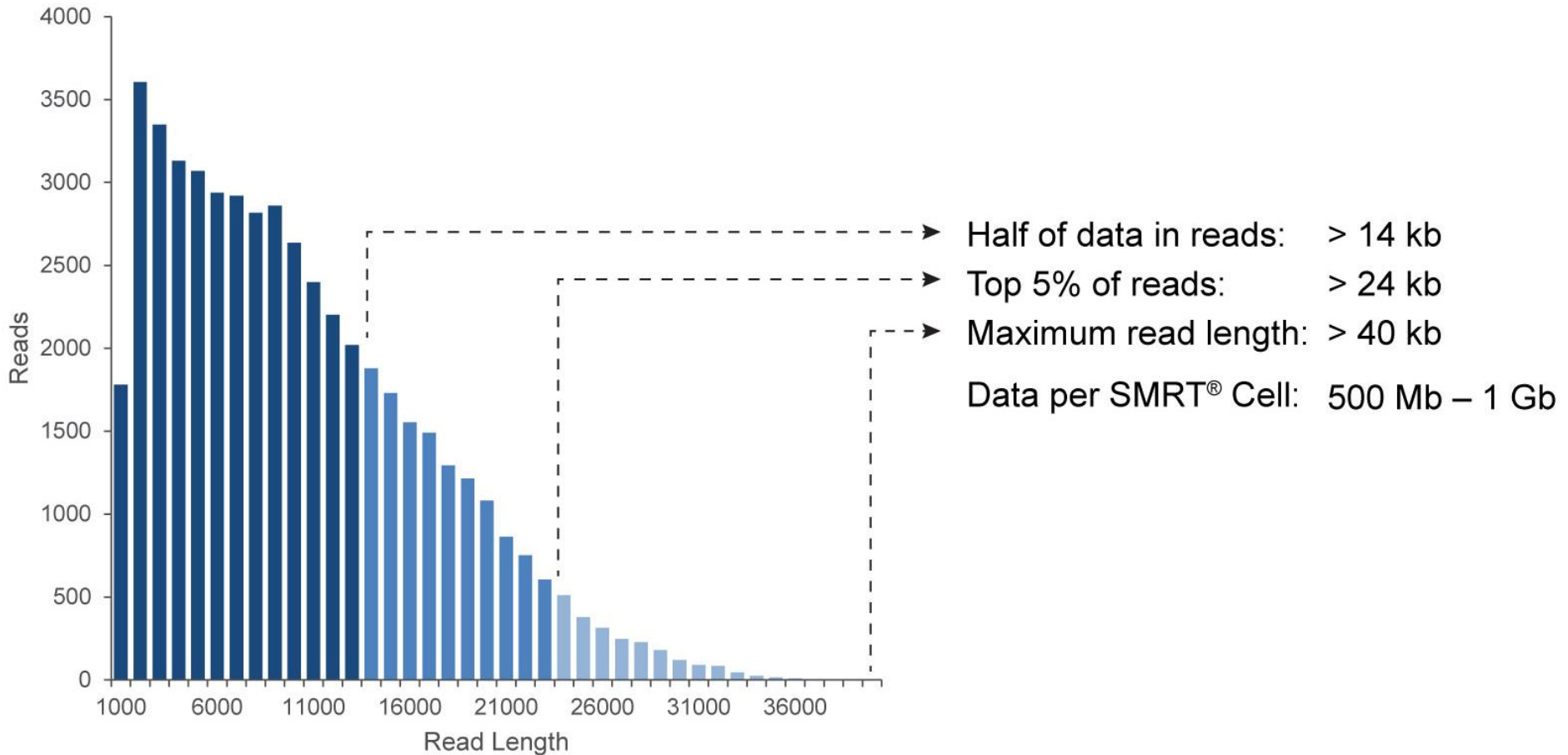


Average: ~ 8.5 kb
Maximum: > 30 kb
Top 5% of reads: > 18 kb
Half of data in reads: > 10 kb
Data per SMRT® Cell: ~ 375 Mb

Based on data from a 20 kb size-selected *E. coli* library using a 180-minute movie.

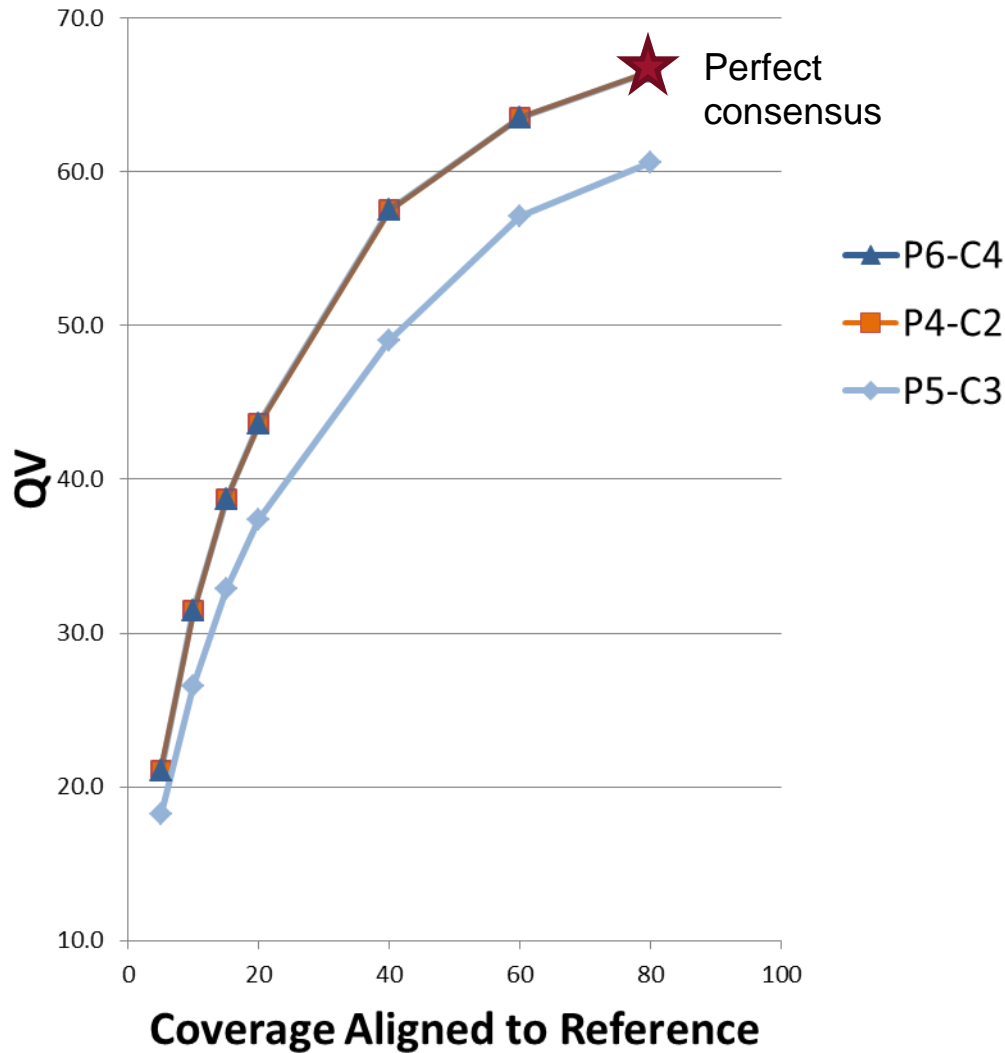
Each SMRT Cell yields ~ 50,000 reads.

P6-C4: Read Length Performance



P6-C4, 4-hr movie, 20-kb BluePippin™ size-selected *E. coli* library (1 SMRT Cell)

Consensus Accuracy Performance Comparison



~QV 50 consensus accuracy coverage:

- P6-C4 30x ± 10
- P4-C2 30x ± 10
- P5-C3 45x ± 10

- 20 kb *E. coli* library
- Resequencing analysis with SMRT® Analysis v2.2

Products and Workflow

Library Preparation



Template Prep Kit
Polymerase Binding Kit
MagBead Kit
AMPure® PB Kit

**No amplification
required**

Instrument Run



PacBio® RS II
RS Remote
RS Touch
RS Dashboard
SMRT® Cells
DNA Sequencing Kit

**Sequencing time 30 to
180 min per SMRT Cell**

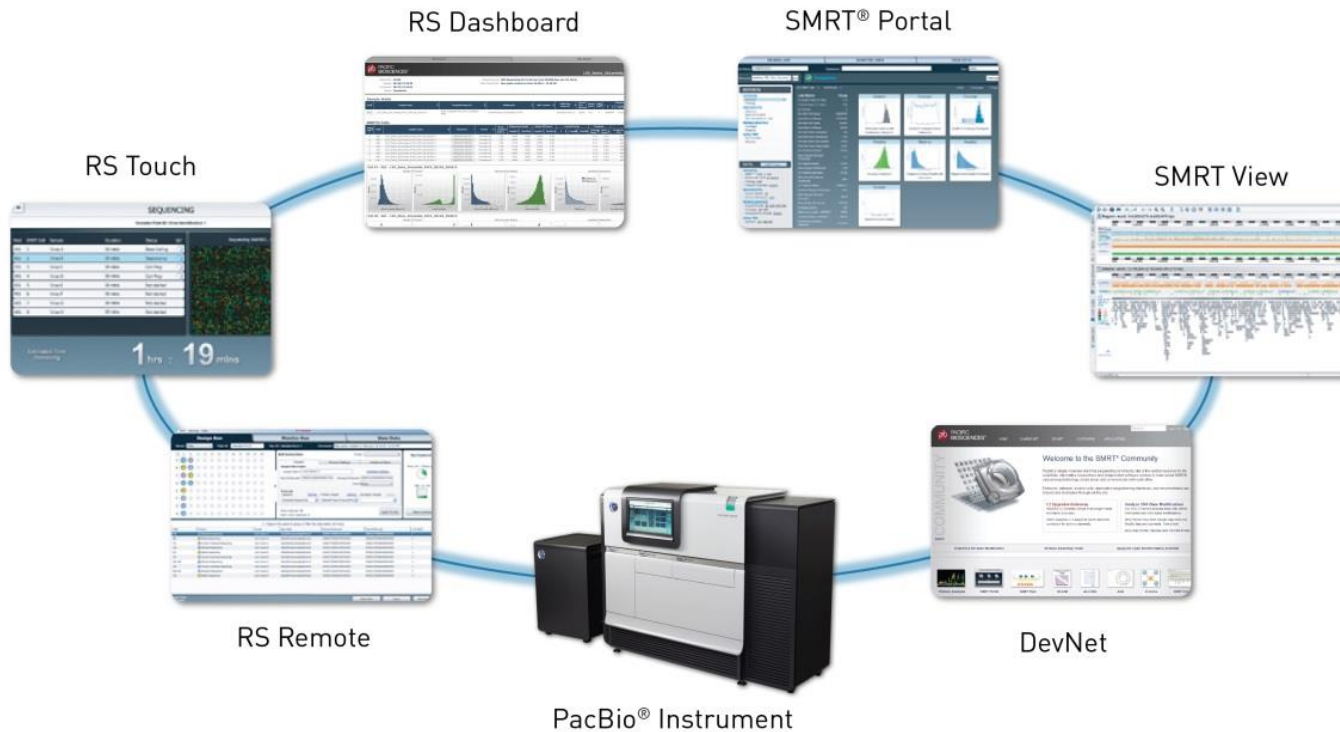
Data Analysis



SMRT Analysis
SMRT Portal
SMRT View

**Open source, open
standards**

Integrated End-to-End Solutions



Easy, user-friendly, web-based solutions

Streamlined data analysis and viewing

Support for novice and expert users

Applications

De Novo Assembly



Very long reads

Targeted Sequencing



Single Molecule Accuracy

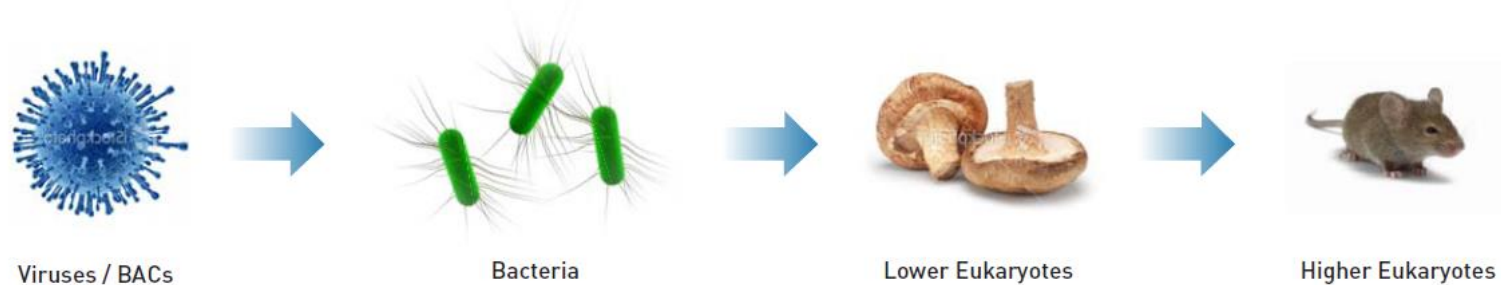
Base Modification Detection



Kinetic Information

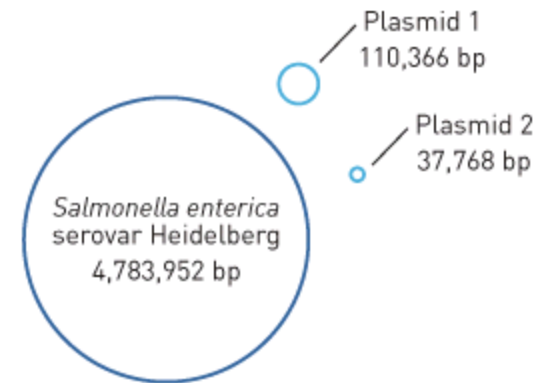
De Novo Assembly: Reduce Ambiguities

Complete microbial genomes and improve assemblies of larger organisms



Read lengths up to 30 kb, unbiased genome coverage, and high accuracy

- Resolve mobile elements and structural-variation events
- Generate complete, accurate and contiguous genome assemblies
- Annotate more genes



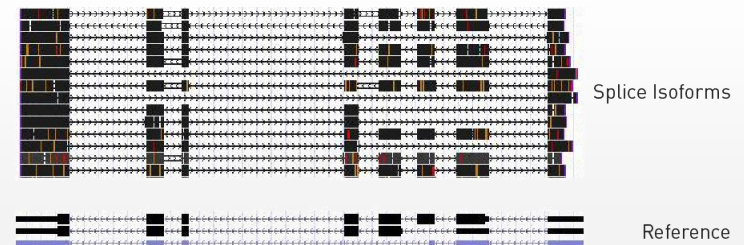
De novo, finished assembly of *Salmonella enterica* subsp. *enterica* serovar Heidelberg, with predicted accuracy of >99.999% (QV50)

Targeted Sequencing: High-Resolution Insights

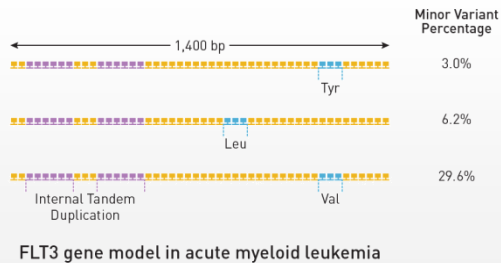
Exquisite sensitivity and specificity to fully characterize genetic complexity

- Multi-kilobase reads
- Achieves 99.999% consensus accuracy
- Linear variant detection to <0.1% frequency
- Access to the entire genome

Iso-Seq™ Full-Length Transcript Sequencing

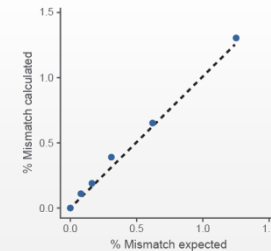


Compound Mutations and Haplotype Phasing

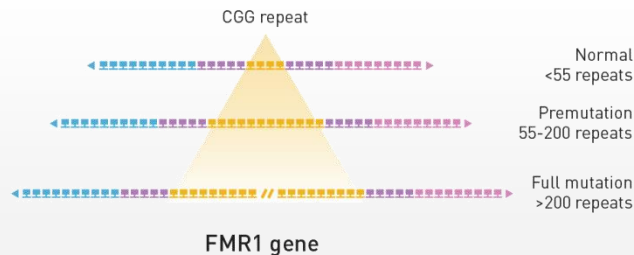


Minor-variant Detection

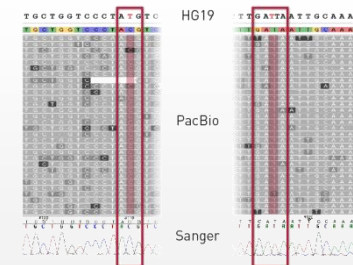
SNP254 C>A



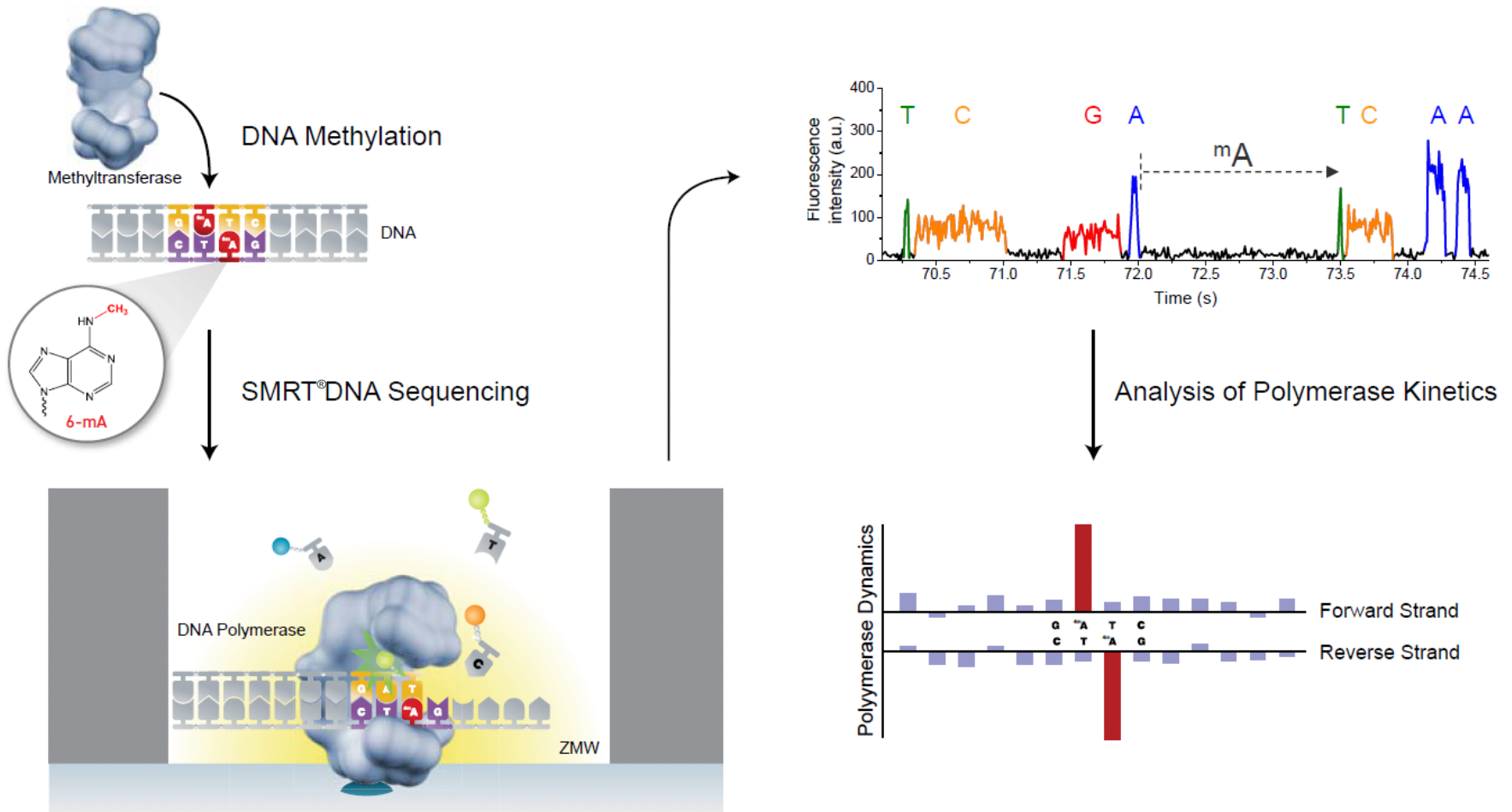
Repeat Expansions



SNP Detection and Validation



Base Modification: Discover the Epigenome



Detect base modifications using the kinetics of the polymerization reaction during normal sequencing

Read Customer Publications



Key Sequencing Characteristics

1. Contiguity

- Sequence reads >10,000 bases
- Some reads >30 kb

2. Accuracy

- Achieves >99.999% (QV50)
- Lack of systematic sequencing errors

3. Uniformity

- Lack of GC content or sequence complexity bias

4. Originality

- No DNA amplification
- Epigenome characterization

ARTICLES

Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data

Chen-Shan Chin¹, David H Alexander¹, Patrick Marks¹, Aaron A Klammer¹, James Drake¹, Cheryl Heiner¹, et al & Jonas Korlach¹

Ross et al. *Genome Biology* 2013, 14:R51
<http://genomebiology.com/2013/14/5/R51>

Genome Biology

RESEARCH Open Access

Characterizing and measuring bias in sequence data

Michael G Ross¹, Carsten Russ, Maura Costello, Andrew Hollinger, Niall J Lennon, Ryan Hegarty, Chad Nusbaum and David B Jaffe

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in Microbiology

ELSEVIER

Entering the era of bacterial epigenomics with single molecule real time DNA sequencing

Brigid M Davis, Michael C Chao and Matthew K Waldor

Roberts et al. *Genome Biology* 2013, 14:405
<http://genomebiology.com/2013/14/6/405>

Genome Biology

CORRESPONDENCE

The advantages of SMRT sequencing

Richard J Roberts^{1*}, Mauricio O Carneiro² and Michael C Schatz²

Abstract
Of the current next-generation sequencing technologies, SMRT sequencing is sometimes overlooked. However, attributes such as long reads, modified base detection and high accuracy make SMRT a useful technology and an ideal approach to the complete sequencing of small genomes.

Pacific Biosciences' single molecule, real-time sequencing technology, SMRT, is one of several next-generation sequencing technologies that are currently in use. In the

Now a new technology, SMRT sequencing from Pacific Biosciences [1], has been developed that not only produces considerably longer and highly accurate DNA sequences from individual unamplified molecules, but can also show where methylated bases occur [2] (and thereby provide functional information about the DNA methyltransferases encoded by the genome).

SMRT sequencing is a sequencing-by-synthesis technology based on real-time imaging of fluorescently tagged nucleotides as they are synthesized along individual DNA template molecules. Because the technology uses a DNA polymerase to drive the reaction, and because it images single molecules, there is no degradation

[1]. Derivatives of methylated deoxythymine and the deoxycytosine and 5-carboxyl cytosine regulate gene expression, a different mechanism(s) [4]. Methylation is also well-established in bacteria, including *Escherichia coli*, in which methylation (Dam and CcrM, respectively) is involved in chromosome replication, DNA repair (reviewed in [5,6]). However, DNA methylation in prokaryotes is more than in eukaryotes [7], and the consequences of DNA modification are being investigated for most of the

sequencing platforms in the last decade of whole genome sequencing. Meanwhile, large scale analyses are lagged far behind, resulting in

Summary-----PacBio 3rd Generation Solution

Novel System Architecture with Years of Headroom

Speed

Single Molecule Real Time Technology

Granularity

Built-in Flexibility

- Multiple sequencing protocols enabled the same sample preparation
 - Reagent Configuration
- Cost-effective

Long Readlength

- Simplified assembly and mapping
 - Enables access to novel, medically important portions of genome
 - Capture last 5% of human genome
- Ease of Use

- Time to result – sample prep to sequence in <1 day

- Sequences 1-3 bases per second

PACBIO RS

- Polymerases
- Data
- Not
- Preparation



The world's first 3rd generation DNA Sequencer

- Simplified sample prep
- Up to 24 hours of unattended operation



PACIFIC
BIOSCIENCES®