Illumina Sequencing Systems – Power, Simplicity, Versatility.

Jeremy Preston, Ph.D.
Director, Product Marketing
Systems & Consumables
San Diego

Illumina Sequencing

It keeps getting easier – do more with less!

MiSeq

iSeq 2

HiSeq 2500

Circa Sep 2012

Single Instrument Workflow for MiSeq and HiSeq

iNTEGRATION

Streamlined end-to-end solution

ign illumina genome network

Services

FastTrack SERVICES

Sample Prep

Sequencing

Analysis

BaseSpace

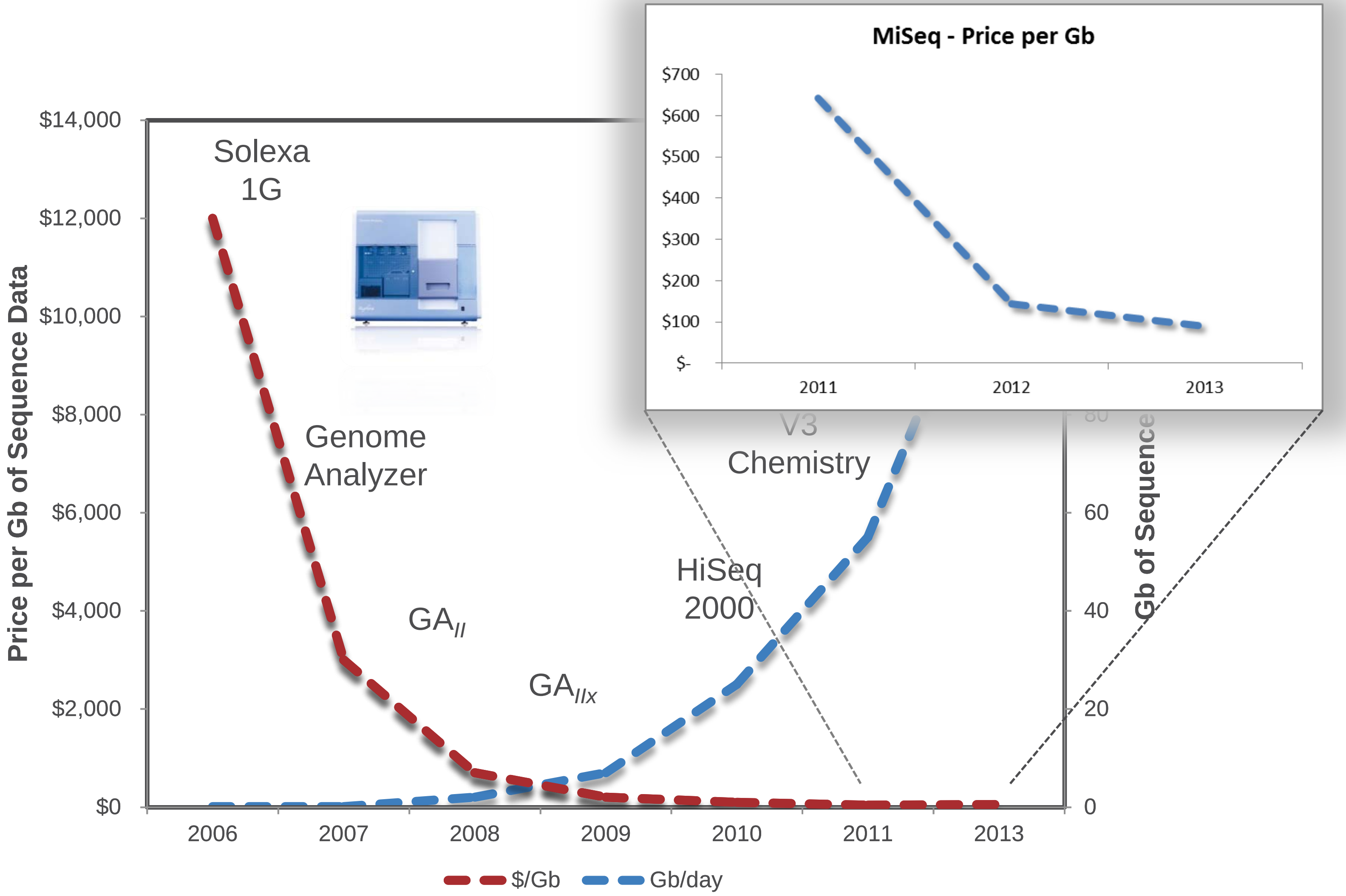
Suite of DNA, RNA & Targeted Solutions

Industry's leading NGS instruments

Storage, Processing, Analysis & Collaboration

The Pace of Change in Sequencing

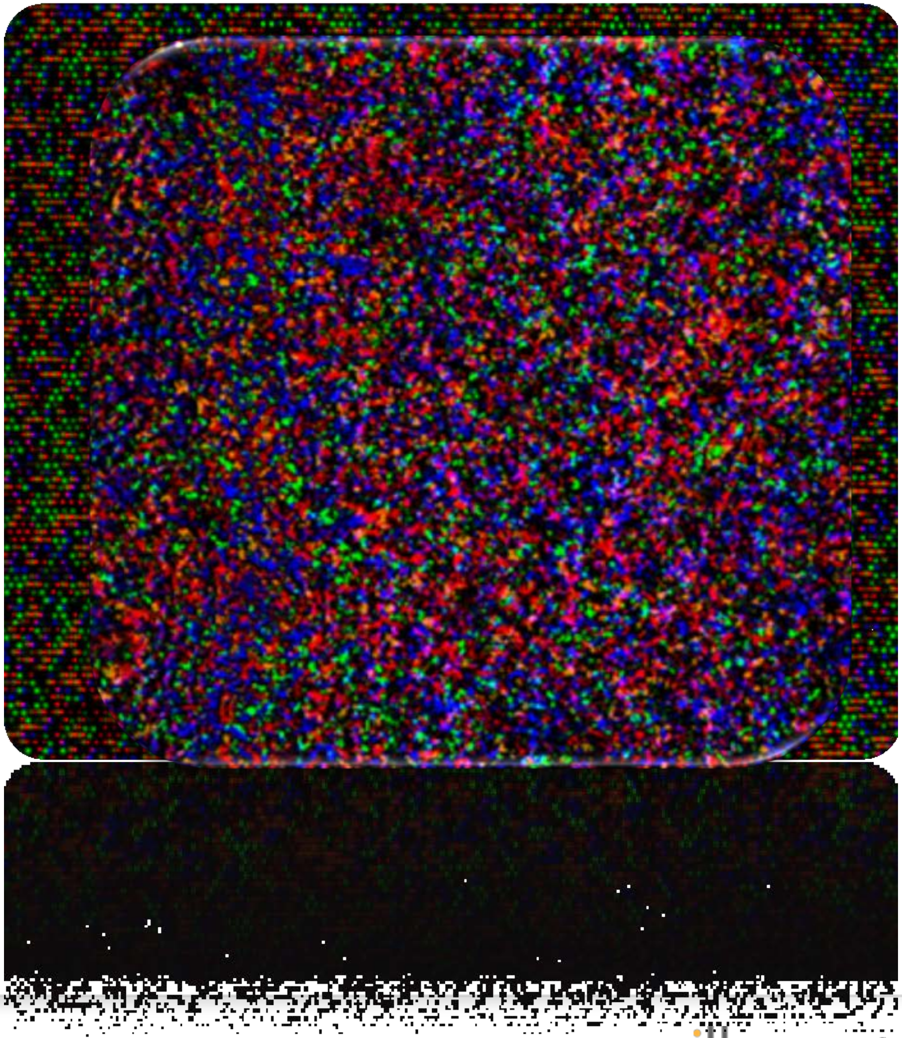
Making sequencing cheaper and faster



Optimizing Cluster Size and Spacing

Random cluster generation to patterned flow cell

- New patterned substrate technology
  - Transforms random cluster generation into densely packed ordered arrays
- Universal technology that can benefit multiple Illumina platforms
- Increased performance
  - Potential increase in reads
  - Faster cluster detection timing



Ordered Array

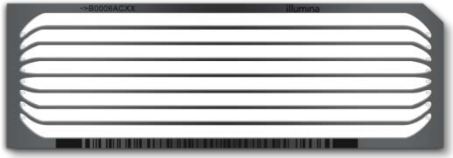
Random Cluster Generation

HiSeq 2500 Sequencing System

Speed and versatility

HIGH OUTPUT

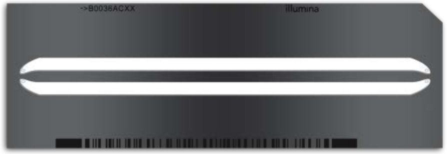
50 – 600Gb
2 – 11 days
2 x 100bp max



Larger projects, fewer runs

RAPID RUN

10 – 180Gb
7 – 40 hours
2 x 150bp max



Smaller projects, quick results

HiSeq 2500 WGS – from DNA to Disease in ~50hrs

Rapid sequencing and annotation of neonatal intensive care patients

Neonate at 5 mo - developmental regression & seizures

Samples prepared, sequenced and analyzed (HiSeq 2500)

Identified novel variant in gene linked to copper metabolism

Confirmed diagnosis of Menkes disease

Top 10 Medical Breakthroughs

7. Speeding DNA-Based Diagnosis for Newborns

34 of 61
VIEW ALL

By Alice Park | Dec. 04, 2012 | Add a Comment

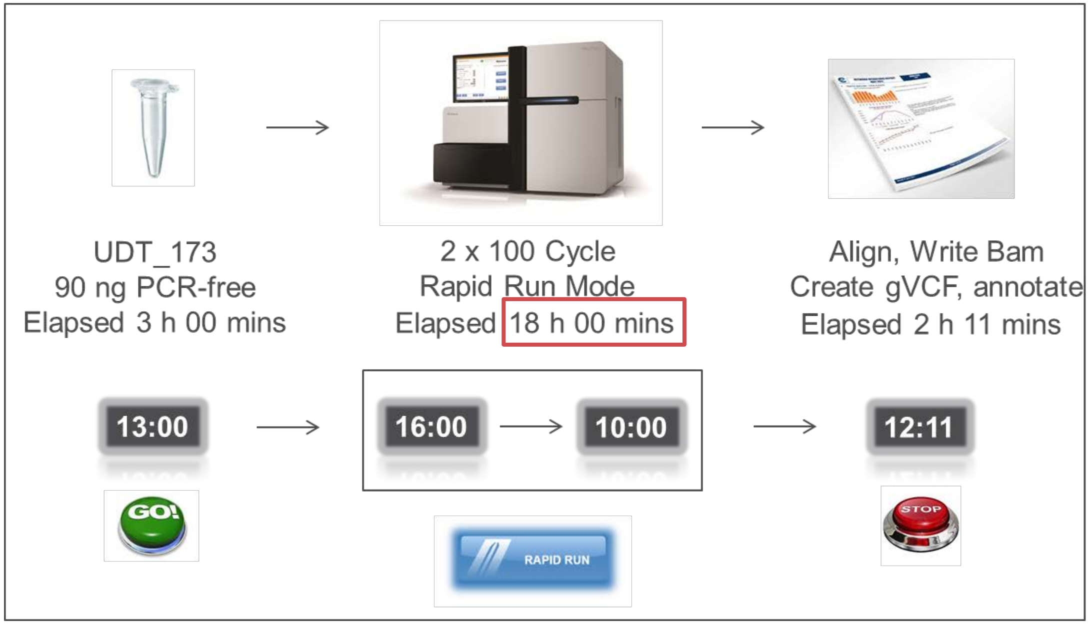
Fifty hours. That's how long it now takes to decode and interpret a newborn baby's genome — an undertaking that used to take weeks or even months. And those two days can mean the difference between life and death for a critically ill infant. The speedier genomic analysis is possible thanks to advances in sequencing technology as well as innovative software that links the 3,500 known genetic defects to their childhood diseases, allowing doctors to quickly decide on the right treatment that could save a baby's life. About 30% of babies admitted to the neonatal intensive care unit each year have inherited a genetic disease, and sequencing their genomes may become a critical part of improving their care in coming years — the sooner the better.

DAVID AARON TROY / GETTY IMAGES

Courtesy of Stephen Kingsmore, Children's Mercy Hospital, Kansas City

Sample to Variant Call in Less than 24 Hours!

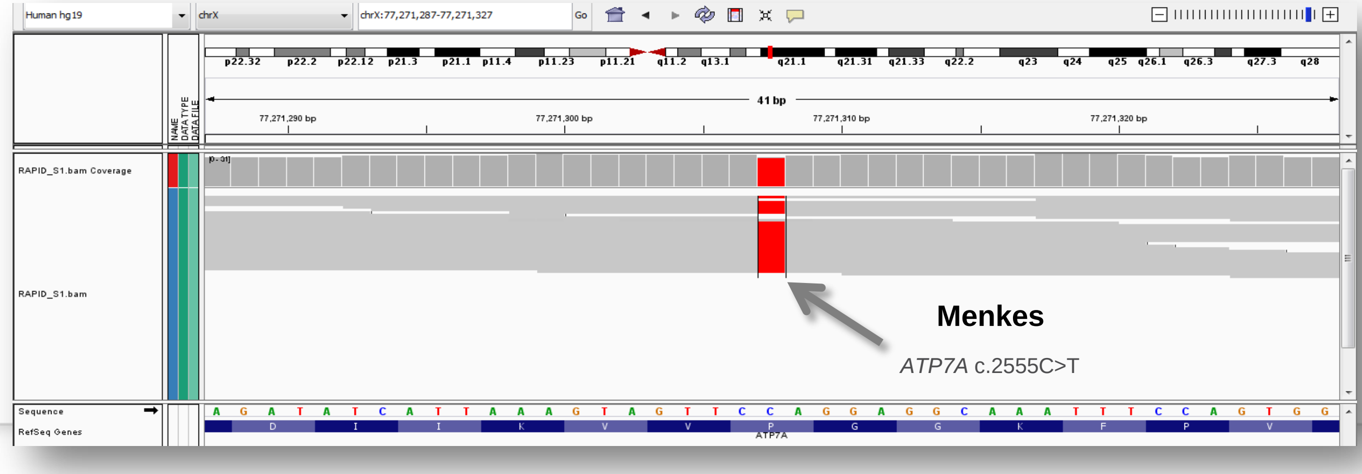
Blinded re-run of Menkes disease sample using Illumina pipeline



UDT_173
90 ng PCR-free
Elapsed 3 h 00 mins

2 x 100 Cycle
Rapid Run Mode
Elapsed 18 h 00 mins

Align, Write Bam
Create gVCF, annotate
Elapsed 2 h 11 mins



Menkes

ATP7A c.2555C>T

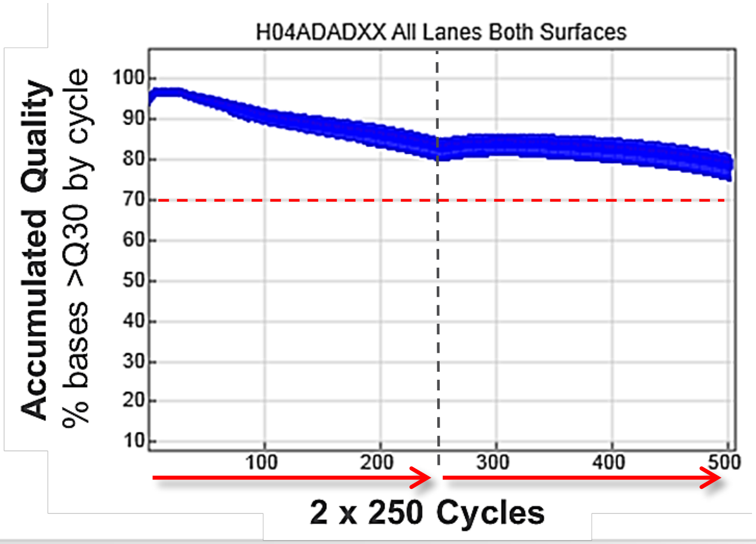
HiSeq 2500 – Extending Performance in 2013

Commercial path to 2x250bp reads in rapid run mode

- 2x250bp in rapid mode
- Up to 300Gb in ~60 hours
  - ~150Gb per flow cell
- Increased application breadth
  - Transcript isoforms/gene fusions
  - Metagenomics & complex de novo assembly
  - Resolve difficult regions of the genome
- Improved accuracy for all read lengths

New: reagents & software

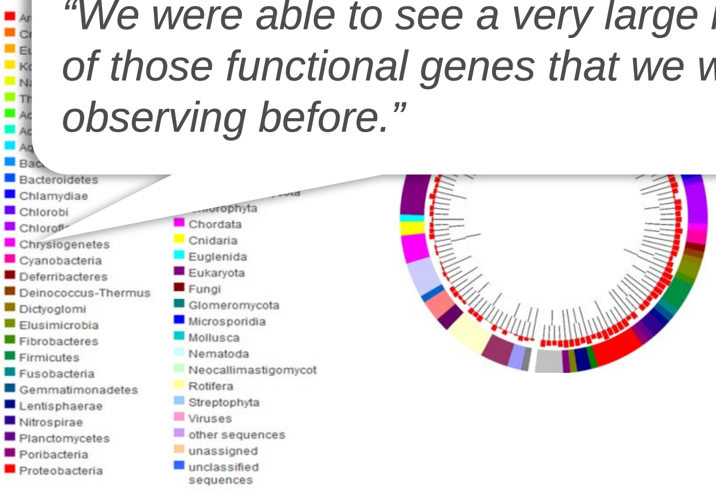
Sample	Yield per Flow Cell	Depth (x)	% >Q30	Reads per Flow Cell
HCC2218 Normal	129Gb	38	78.6	258M
HCC2218 Tumor	156Gb	45	76.3	311M
Current R&D performance	147Gb	42	87.5	294M



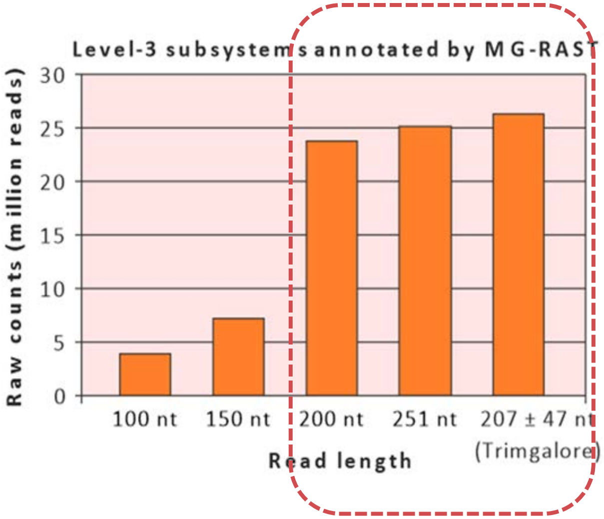
The Longer Read Advantage

Synthetic gut Metagenomics

- *In vitro* model of large intestine inoculated with human gut microbiota
- 2x250 bp on HiSeq 2500
- Examine bacterial community & assess metabolic pathways (MG-RAST)
- Reads mapping to known bacterial genes **highly dependent on read length**
- **Better resolution** of closely related sequences with longer reads (species/strain)



> *"We were able to see a very large number of those functional genes that we were not observing before."*

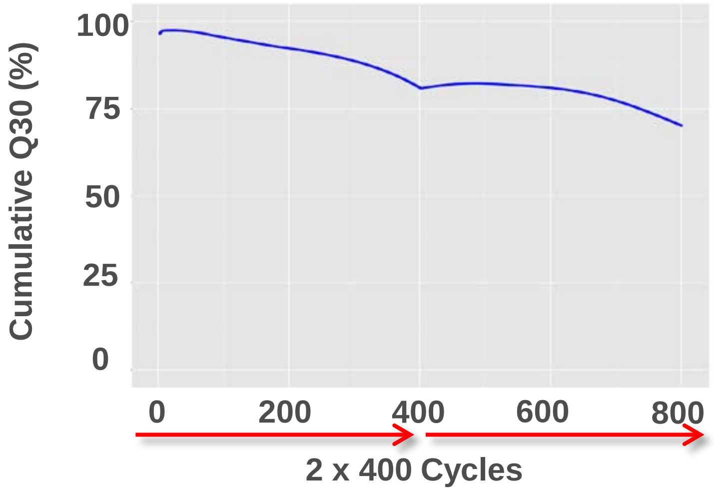


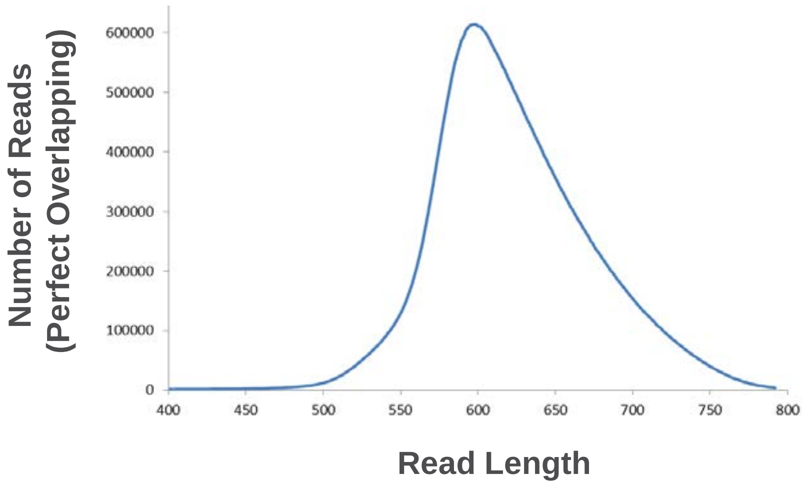
Reads mapping to genes	
100nt	3,963,158
150nt	7,170,312
200nt	23,763,325
251nt	25,160,962

Sample and analysis courtesy of Bas Dutilh, Radboud University Medical Centre Nijmegen, Holland

HiSeq 2500 – R&D Demonstrated Scalability

Overlapping 2 x 400 bp Reads (In Research)





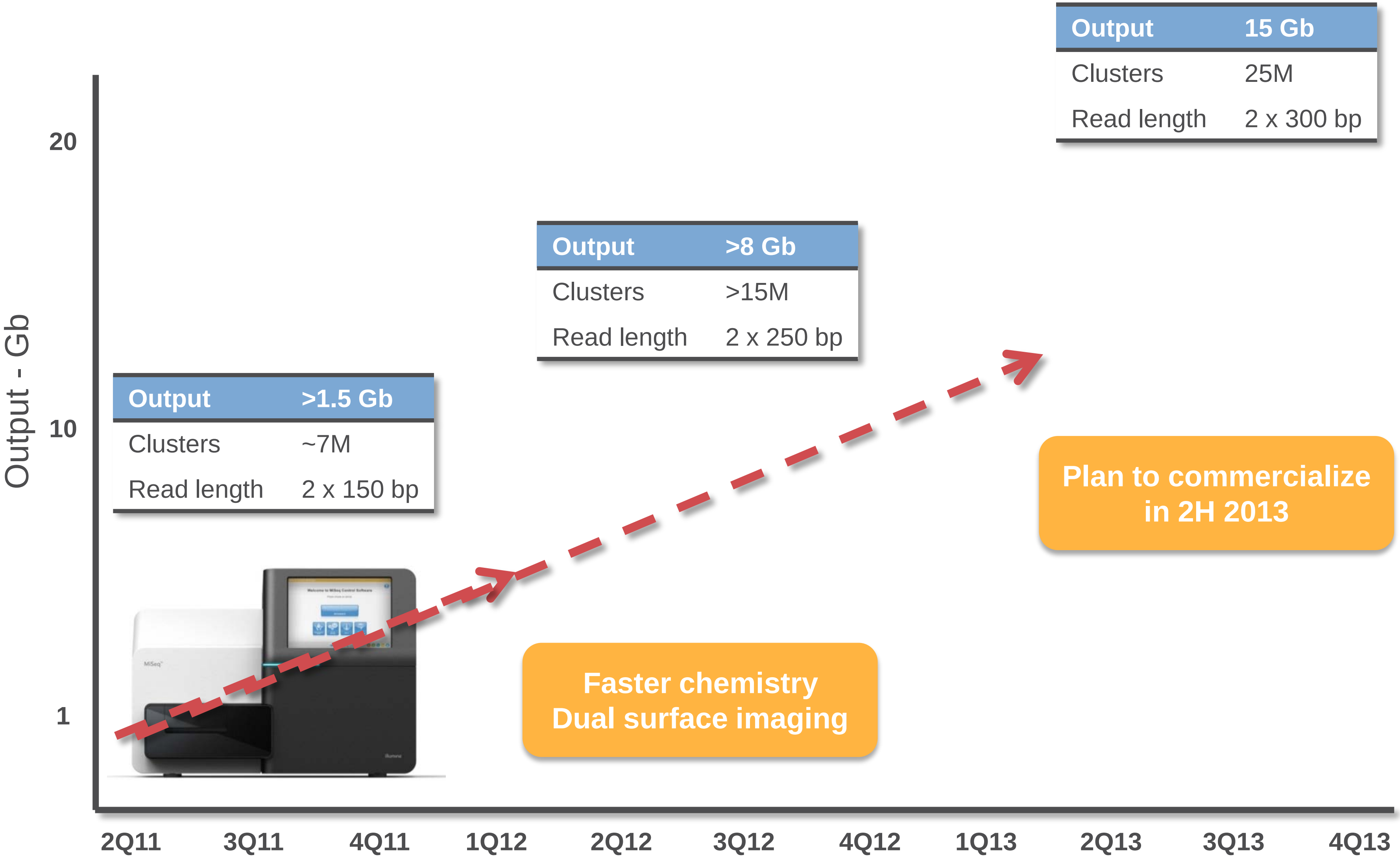
Run metrics	
Read length	2 x 400
Output (single flow cell)	>180Gb
Quality	70% Q30 (All bases)

Read Length	# of reads
>= 500	64.8M
>= 600	42.5M
>= 700	5.3M
>= 750	739K
>= 775	128K

- Human NA12882 (CEPH Child) library – 650 bp median insert size

MiSeq – Continuous Performance Improvements

Path towards 15Gb per run; enabling broader range of applications



Output	>1.5 Gb
Clusters	~7M
Read length	2 x 150 bp

Output	>8 Gb
Clusters	>15M
Read length	2 x 250 bp

Output	15 Gb
Clusters	25M
Read length	2 x 300 bp

Faster chemistry
Dual surface imaging

Plan to commercialize in 2H 2013

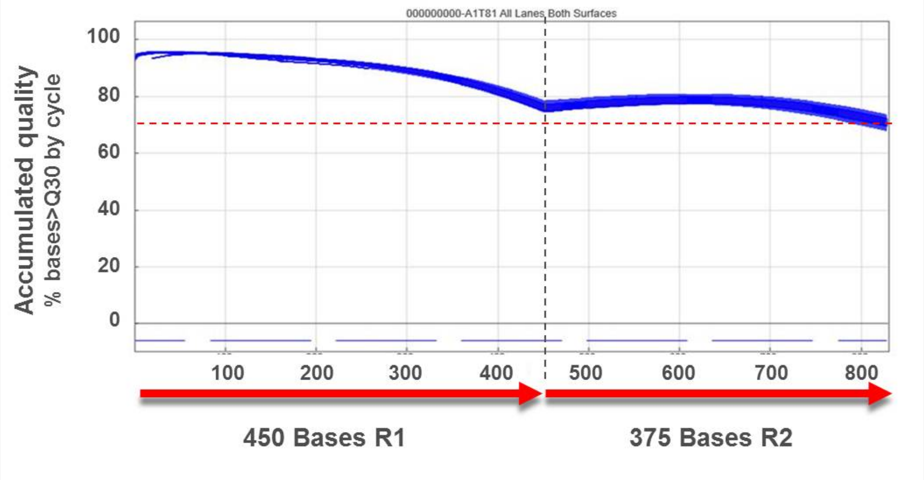
MiSeq – R&D Demonstrated Scalability

Early Development data

Output	15.2 Gb
Clusters	25.1M
Read length	2 x 300
Quality	90% ≥Q30
MMR	0.17% (R1), 0.37% (R2)
Run time	72 hours

New: 150 & 600 cycle PE kits at launch

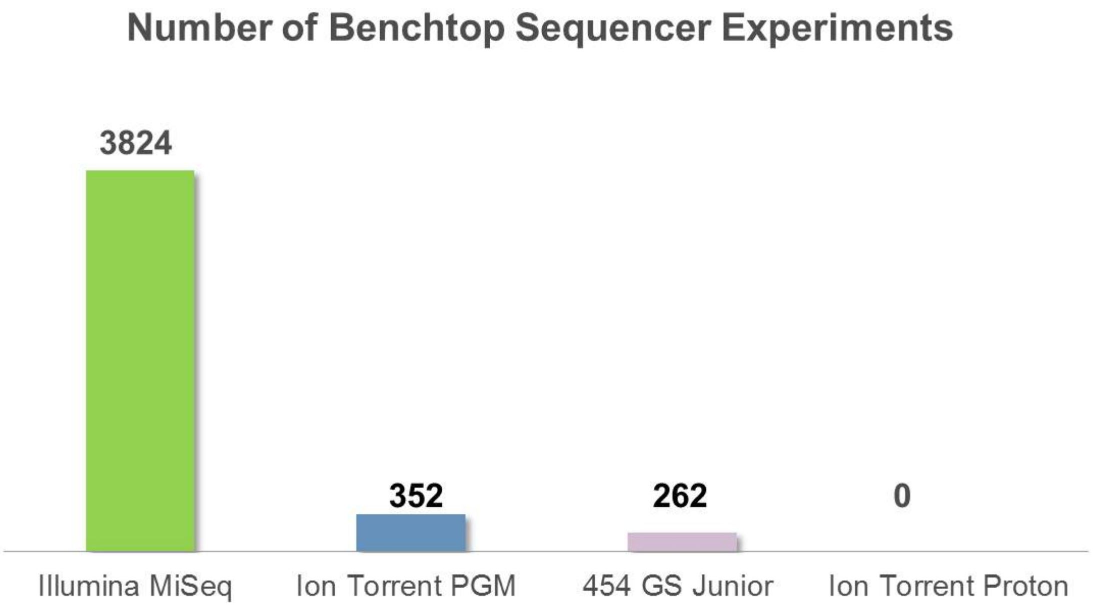
“Research” *longer read runs in excess of 20Gb*

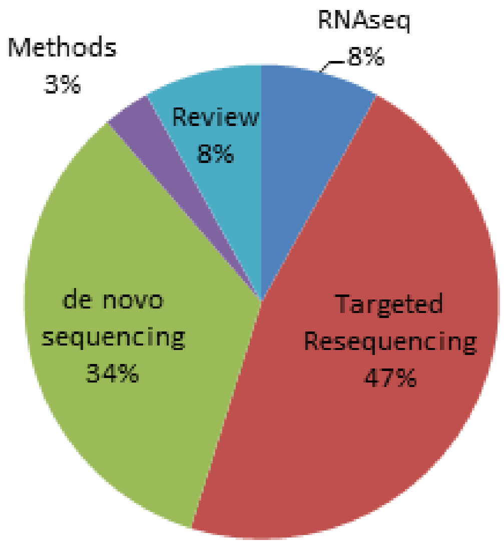


Output	22 Gb
Clusters	26.5M
Read length	450 (R1), 375 (R2)
Quality	71% ≥Q30
MMR	1.7% (R1), 2.4% (R2)

MiSeq – Greater than 85% of Desktop Data

Analysis of data submissions to the NCBI Sequence Read Archive (SRA)





**n = 80+ publications*

NIH's primary archive of NGS data and part of the international partnership of archives at the NCBI, the European Bioinformatics Institute and the DNA Database of Japan.

MiSeq – Transforming Clinical Microbiology

Rapid bacterial genome sequencing

Whole-genome sequ
meticillin-resistant

imon R Harris[*], Edward J P Cartwright[*], M Est
ichael A Quail, Stephen D Bentley, Julian Park

JCM
Journals.ASM.org

Rapid Wh
Tuberculo

Downloaded from bmjopen.bmj.com on August

Access

pen
accessible medical research

A pilot study of
sequencing of St
Clostridium difficile
detection and surve

David W Eyre,[1,2] Tanya Golubchik,[2,3] N
Paolo Piazza,[4] Elizabeth M Batty,[2,3] Can
Xavier Didelot,[2,3] Lily O'Connor,[1,5] Roch
Angela M Kearns,[6] Angela Shaw,[7] John
Tim E A Peto,[1,2,5] A Sarah Walker,[1,2,10]

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Rapid Whole-Genome Sequencing for Investigation of a Neonatal MRSA Outbreak

Claudio U. Köser, B.A., Matthew T.G. Holden, Ph.D., Matthew J. Ellington, D.Phil.,
Edward J.P. Cartwright, M.B., B.S., Nicholas M. Brown, M.D.,
Amanda L. Ogilvy-Stuart, F.R.C.P., Li Yang Hsu, M.R.C.P.,
Claire Chewapreecha, B.A., Nicholas J. Croucher, M.A.,
Simon R. Harris, Ph.D., Mandy Sanders, B.Sc., Mark C. Enright, Ph.D.,
Gordon Dougan, Ph.D., Stephen D. Bentley, Ph.D., Julian Parkhill, Ph.D.,
Louise J. Fraser, Ph.D., Jason R. Betley, Ph.D., Ole B. Schulz-Trieglaff, Ph.D.,
Geoffrey P. Smith, Ph.D., and Sharon J. Peacock, Ph.D., F.R.C.P.

REVIEWS

QUENCING

microbiology
e sequencing

lson[2,4], *Tim E. A. Peto*[3,4] *and*

PLoS

OPEN ACCESS Freely available online

Opinion

Routine Use of Microbial Whole Genome Sequenci
Diagnostic and Public Health Microbiology

Claudio U. Köser[1,2*], Matthew J. Ellington[2], Edward J. P. Cartwright[1,2], Stephen H. Gillespie[3],
Nicholas M. Brown[2], Mark Farrington[2], Matthew T. G. Holden[4], Gordon Dougan[4], Stephen D.
Julian Parkhill[4], Sharon J. Peacock[1,2,4,5]

illumina

MiSeq – Adopted by Leading Public Health Organizations

Identification of food-borne pathogen outbreaks

Cost effective, public health genomic epidemiology

Novel Avian-Origin Influenza A (H7N9) Virus Genome Sequencing at China CDC Using MiSeq

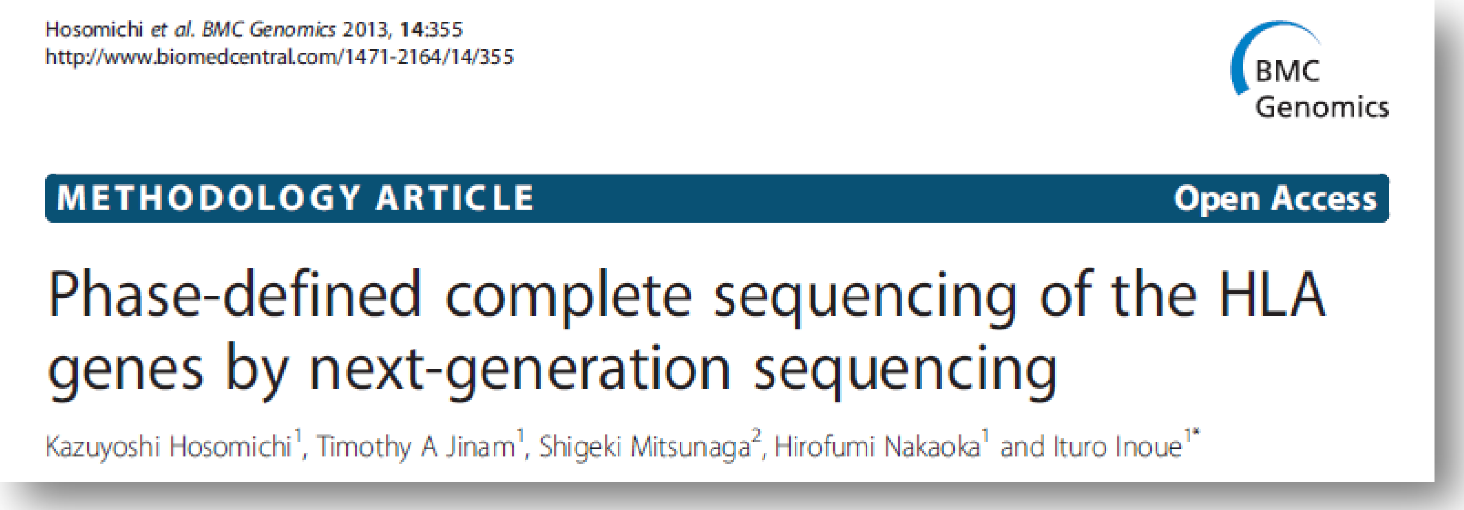
The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Preliminary Report: Epidemiology of the Avian Influenza A (H7N9) Outbreak in China

Qun Li, M.D., Lei Zhou, M.D., Minghao Zhou, Ph.D., Zhiping Chen, M.D., Furong Li, M.D., Huanyu Wu, M.D., Nijuan Xiang, M.D., Enfu Chen, M.P.H., Fenyang Tang, M.D., Dayan Wang, M.D., Ling Meng, M.D., Zhiheng Hong, M.D., Wenxiao Tu, M.D., Yang Cao, M.D., Leilei Li, Ph.D., Fan Ding, M.D., Bo Liu, M.D., Mei Wang, M.D., Rongheng Xie, M.D., Rongbao Gao, M.D., Xiaodan Li, M.D., Tian Bai, M.D., Shumei Zou, M.D., Jun He, M.D., Jiayu Hu, M.D., Yangting Xu, M.D., Chengliang Chai, M.D., Shiwen Wang, M.D., Yongjun Gao, M.D., Lianmei Jin, M.D., Yanping Zhang, M.D., Huiming Luo, M.D., Hongjie Yu, M.D., M.P.H., Lidong Gao, M.D., Xinghuo Pang, M.D., Guohua Liu, M.D., Yuelong Shu, Ph.D., Weizhong Yang, M.D., Timothy M. Uyeki, M.D., M.P.H., M.P.P., Yu Wang, M.D., Fan Wu, M.D., and Zijian Feng, M.D., M.P.H.

MiSeq – Phased HLA sequencing



Hosomichi *et al. BMC Genomics* 2013, **14**:355
http://www.biomedcentral.com/1471-2164/14/355

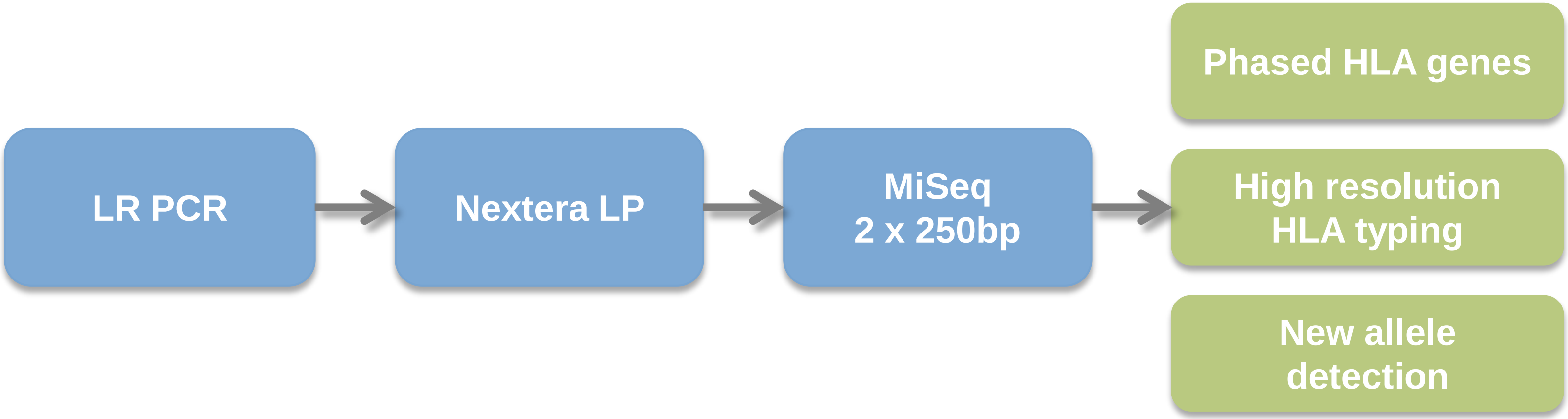
BMC Genomics

METHODOLOGY ARTICLE **Open Access**

Phase-defined complete sequencing of the HLA genes by next-generation sequencing

Kazuyoshi Hosomichi[1], Timothy A Jinam[1], Shigeki Mitsunaga[2], Hirofumi Nakaoka[1] and Ituro Inoue[1*]

Division of Human Genetics, National Institute of Genetics, Shizuoka, Japan



LR PCR

Nextera LP

MiSeq
2 x 250bp

Phased HLA genes

High resolution HLA typing

New allele detection

Illumina's Suite of Sample Prep Solutions

DNA Sample Prep

- Nextera & Nextera XT
- Nextera Mate Pair
- TruSeq DNA PCR-free
- TruSeq Nano DNA

Targeted Resequencing

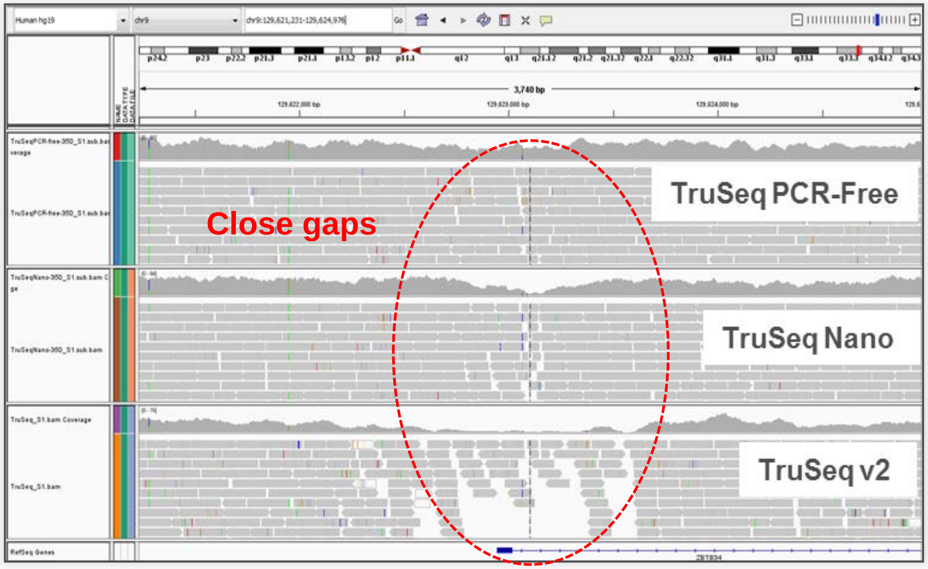
- TruSeq Custom Amplicon
- Nextera Rapid Capture Custom & Exome
- TruSight Panels

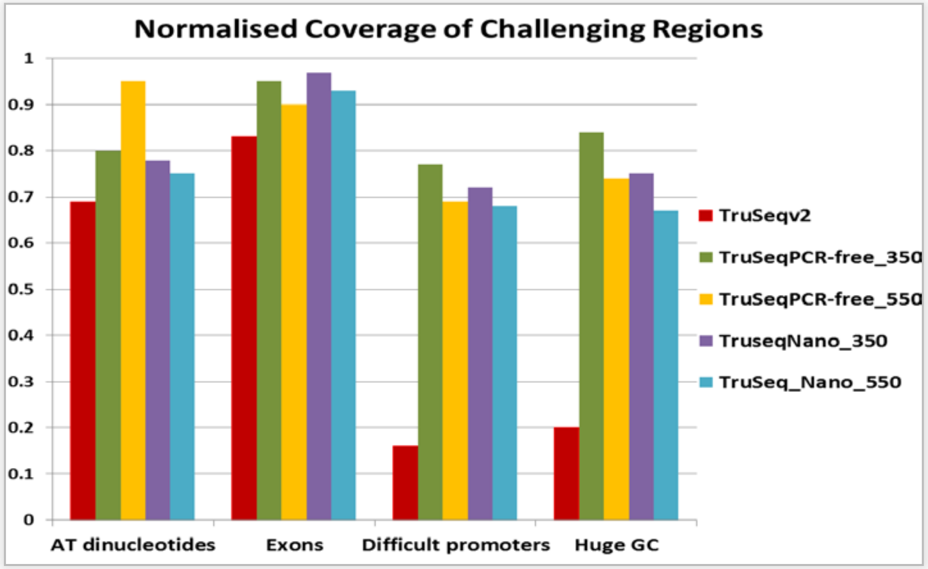
RNA & Regulation Sample Prep

- TruSeq Small RNA
- TruSeq RNA
- TruSeq Stranded RNA (FFPE)
- TruSeq ChIP
- TruSeq Targeted RNA Expression

New: TruSeq PCR Free and Nano

Improved coverage of challenging genomic regions





- TruSeq PCR Free – 1ug input
  - Fast, no PCR duplicates
- TruSeq Nano – 100 ng input
  - Low input, precious samples

- ✓ High library diversity
- ✓ Excellent coverage uniformity
- ✓ All-inclusive kit w/gel-free size-selection
- ✓ 24 & 96 sample options

Nextera Rapid Capture Exome

Fastest exome sequencing - sample to data in less than 2.5 days!

- **Fast assay**
  - Only 1.5 days to prep 96 exomes
  - 12-plex pre-enrichment pooling
- **Lowest DNA input**
  - Only 50ng
- **Efficient**
  - 37Mb exonic content requires only 4Gb total sequence*
- *Also available:*
  - *Nextera Rapid Capture Expanded Exome*
  - *Nextera Rapid Capture Custom Enrichment*

1 exome
in <13 hours

50+ or 100+ exomes
in ~8 days

10 or 20 exomes
in ~20 hours

Low throughput &
fast turnaround

✓ **Flexibility to run on HiSeq & MiSeq**

*Recommended 2 x 75bp reads

TruSeq Targeted RNA

Rapid & economical RNA profiling

- Accurate targeting for Human, Mouse & Rat transcriptomes
  - Alternative isoforms
  - Individual exons and splice junctions
  - cSNP detection for allele specific expression
  - Non-coding RNA transcripts
- Add custom content to Fixed Panels

✓ Rapid – sample to answer in 1.5 days; <4hrs HOT

✓ Low RNA input – 50 ng or less

✓ Low price per sample; 48-384 spls/run

✓ Validation of over 10,000 assay designs

✓ Custom and pre-validated fixed panels available

Pre-Validated Fixed Panels	
Immune Response	Cardiotox
Lung Cancer	Apoptosis
Breast Cancer	Neuro Panel
Stem Cell	Prostate Cancer
P53 Pathway	Wnt Pathway
Cytochrome P450	Cell Cycle
NFκB Pathway	Hedgehog

TruSeq Targeted RNA

Rapid RNA profiling

Single MiSeq run

**Equivalent to 15,000 qPCR reactions
(40 x 384 well plates)**

=

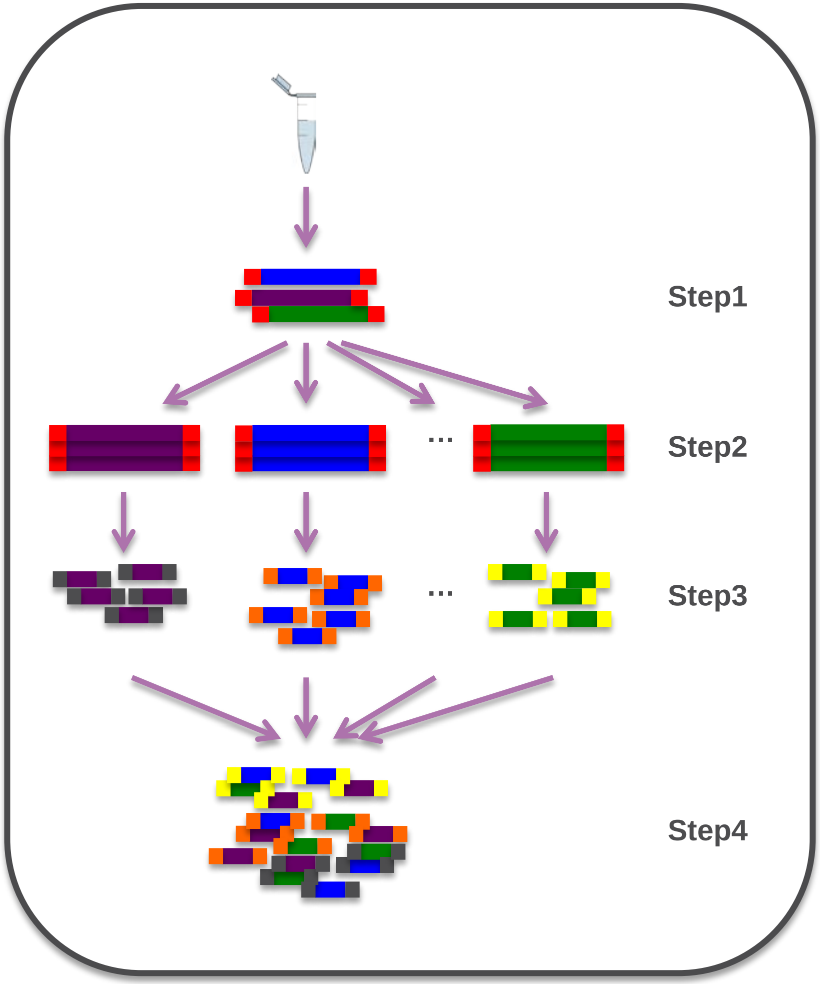
Moleculo Technology Enables Synthetic Long Reads

Up to 10Kb from Illumina short reads

- ✓ Synthetic long reads 8 – 10kb
- ✓ Enables fully phased genomes
- ✓ Accurate de novo assembly of large, complex genomes

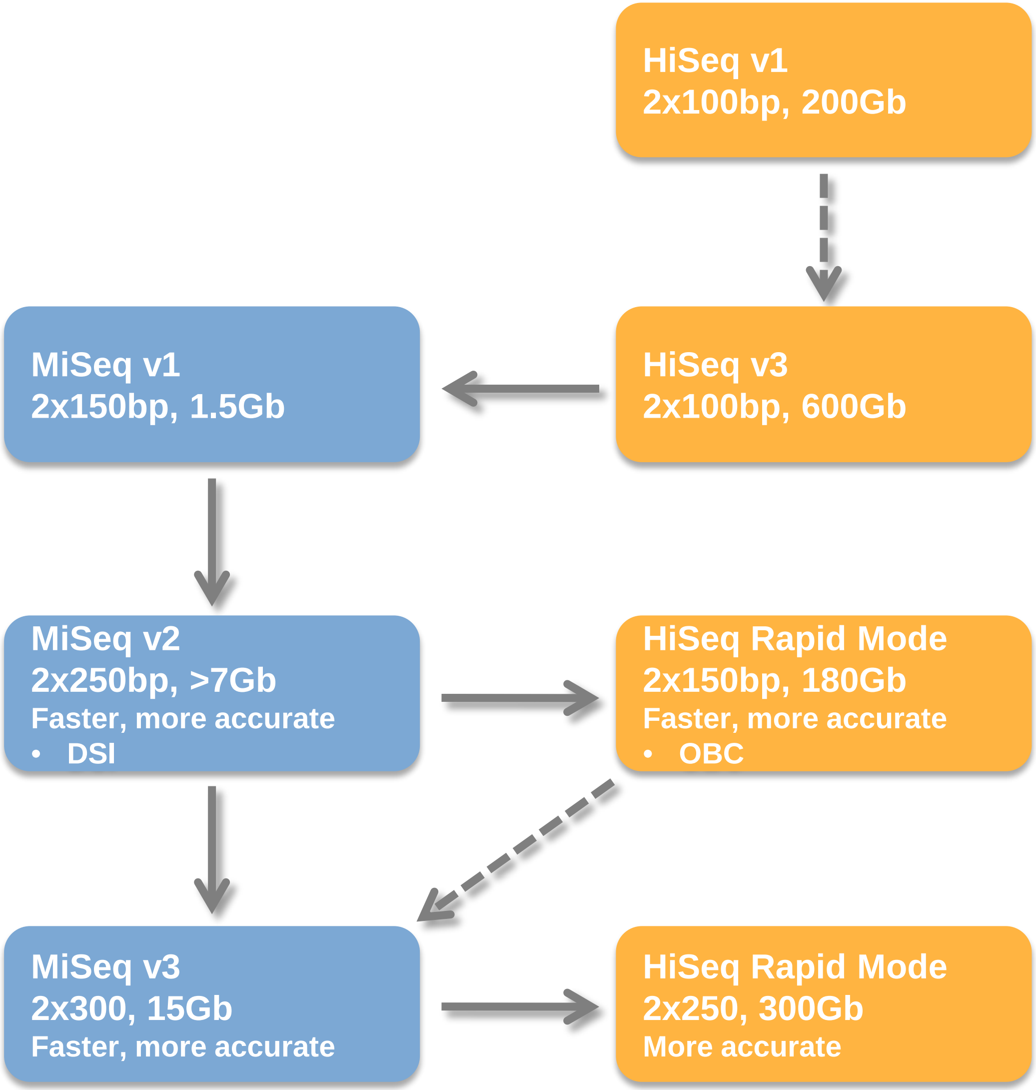
Available:

- ✓ Illumina services 2H13
- ✓ Kit format early 2014



Platform Synergy

Leveraging synergy for fast development



HiSeq v1
2x100bp, 200Gb

HiSeq v3
2x100bp, 600Gb

MiSeq v1
2x150bp, 1.5Gb

MiSeq v2
2x250bp, >7Gb
Faster, more accurate
- **DSI**

HiSeq Rapid Mode
2x150bp, 180Gb
Faster, more accurate
- **OBC**

MiSeq v3
2x300, 15Gb
Faster, more accurate

HiSeq Rapid Mode
2x250, 300Gb
More accurate

Thank You!

Jeremy Preston
jpreston@illumina.com