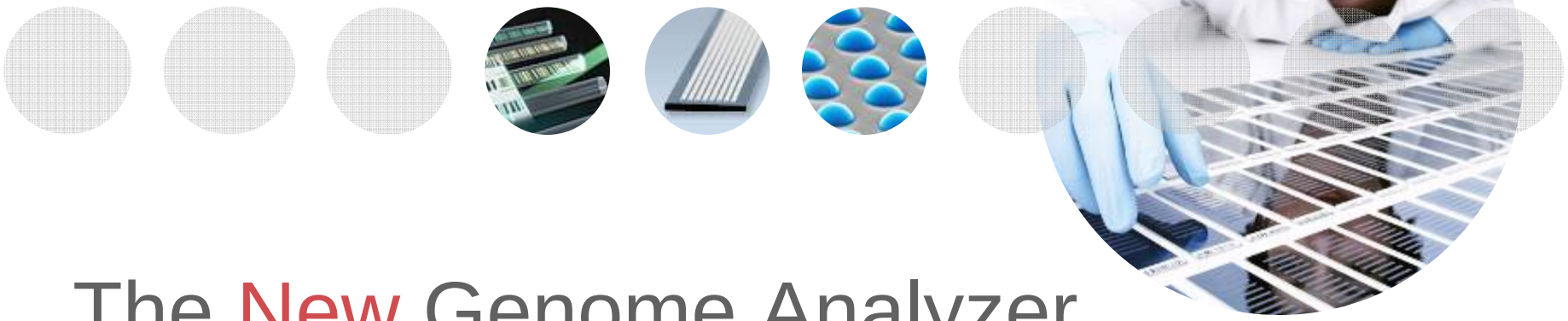
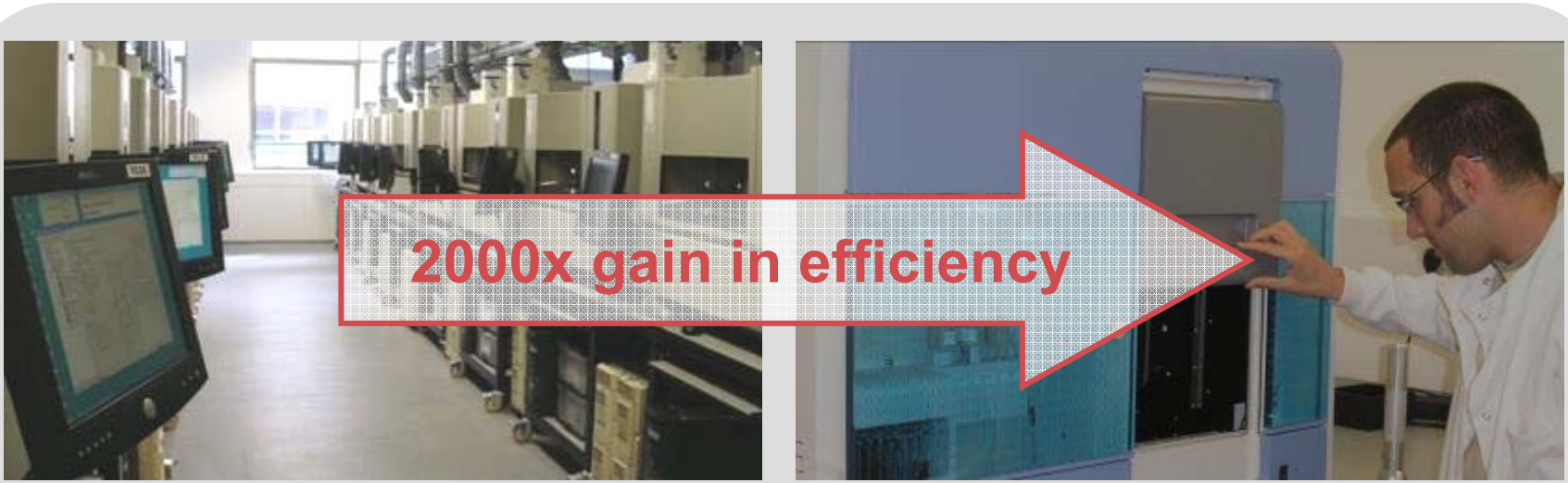


The New Genome Analyzer*IIx*

Delivering more data, faster, and easier than ever before

Jeremy Preston, PhD
Marketing Manager, Sequencing

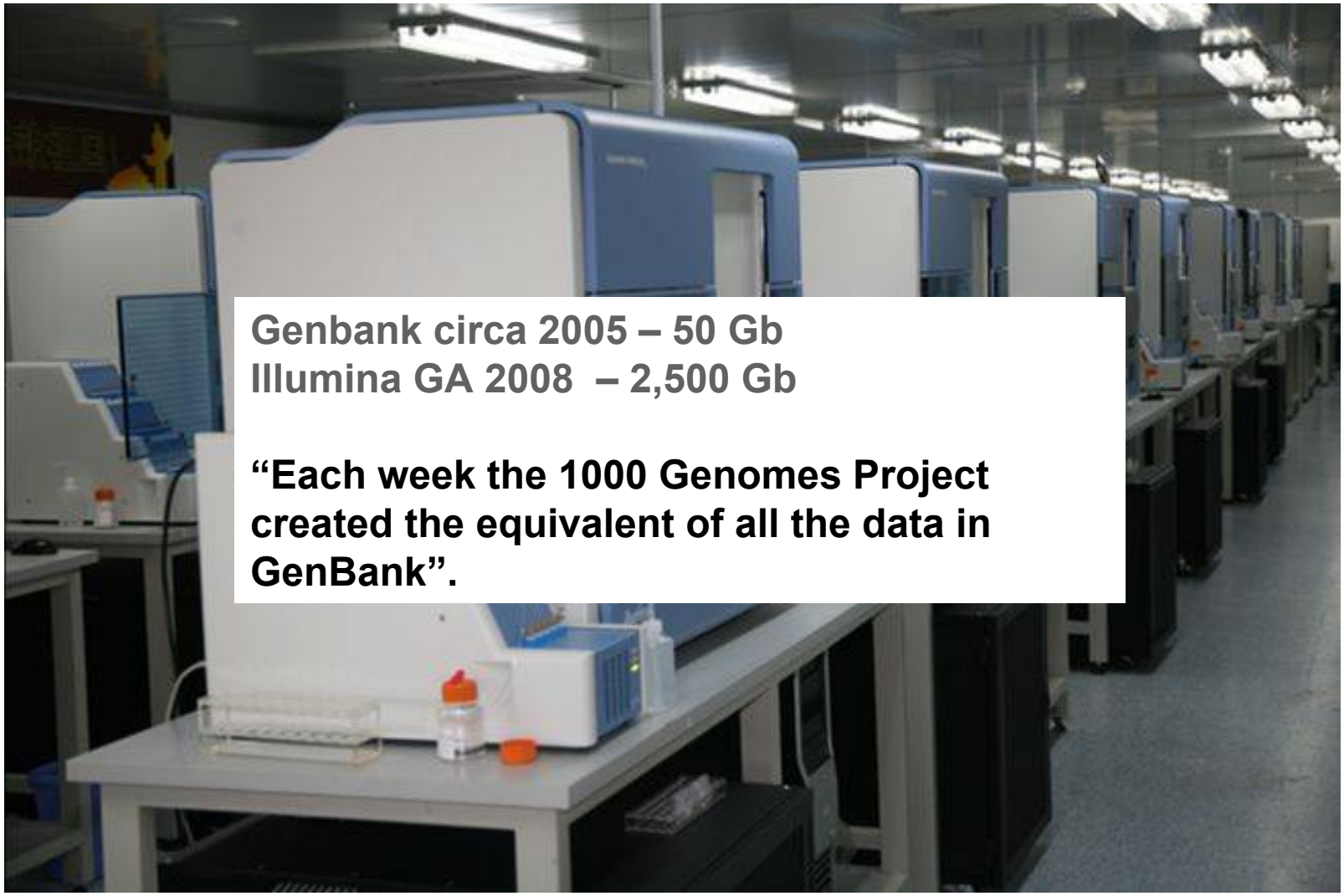
Illumina Genome Analyzer: a Paradigm Shift



Then…	Now…
Library prep robots, clones, preps etc	1 lab bench
~100 sequencers	1 GA
Dozens of lab staff	1 guy (with sideburns)
1,200,000 bases/day/instrument	2,500,000,000 bases/day/instrument
$1-2M for 1Gb raw data	$400 for 1Gb raw data

illumina®

Illumina Genome Analyzer: A paradigm shift



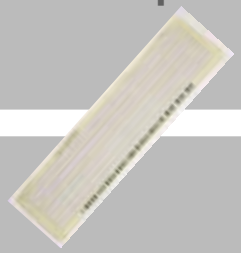
Simplest Sequencing Process

1 *Library prep (~ 6 hrs)*

Fragment DNA
↓
Repair ends / Add A overhang
↓
Ligate adapters
↓
Select ligated DNA

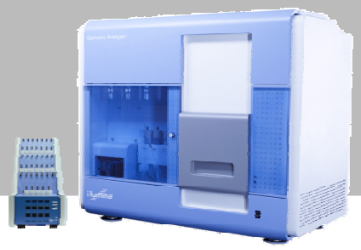
2 *Automated Cluster Generation (~ 5 hrs)*

Up to 96 samples



Hybridize to flow cell
↓
Extend hybridized oligos
↓
Perform bridge amplification

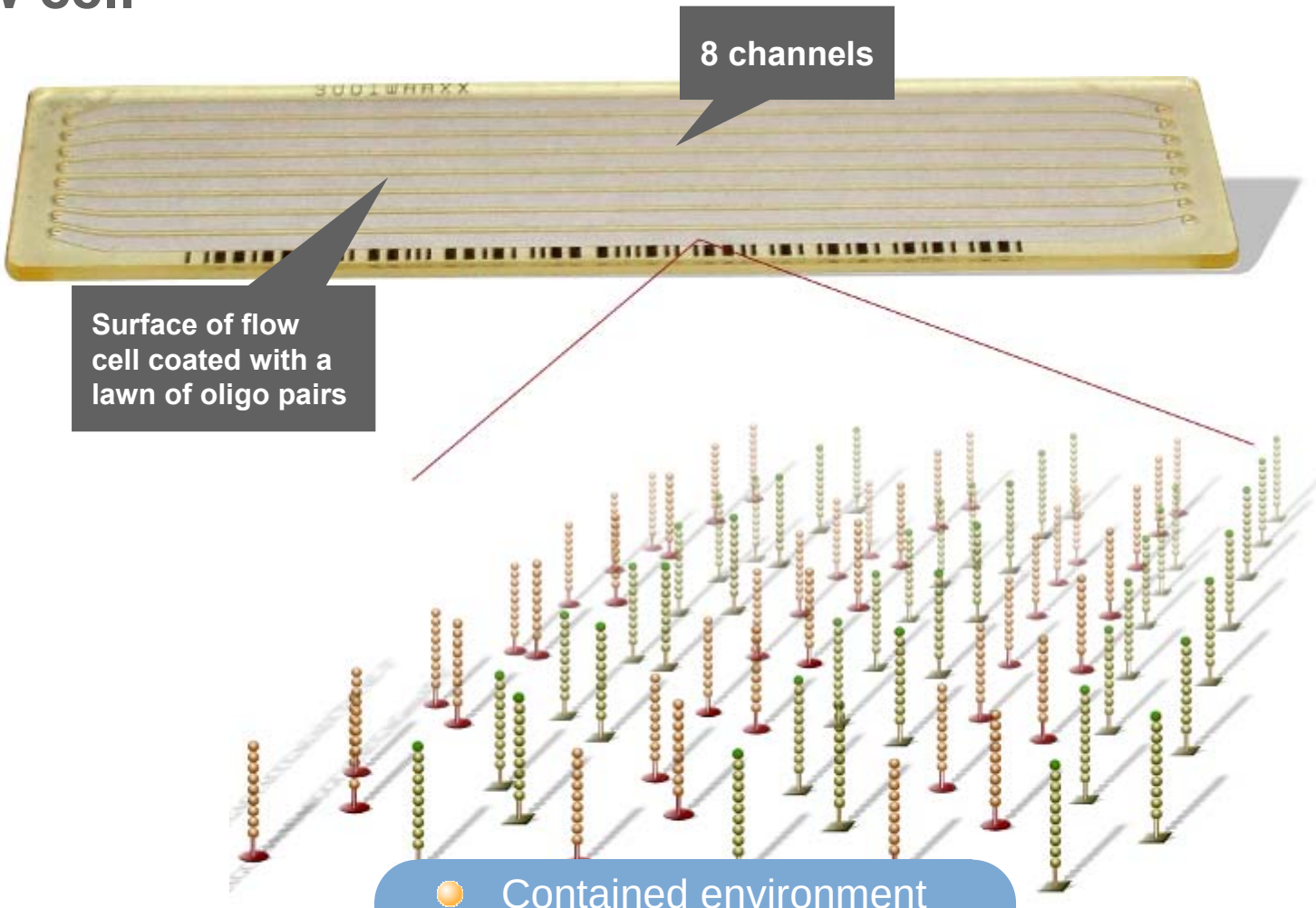
3 *Sequencing (~ 2-9 days*)*



Up to 96 samples

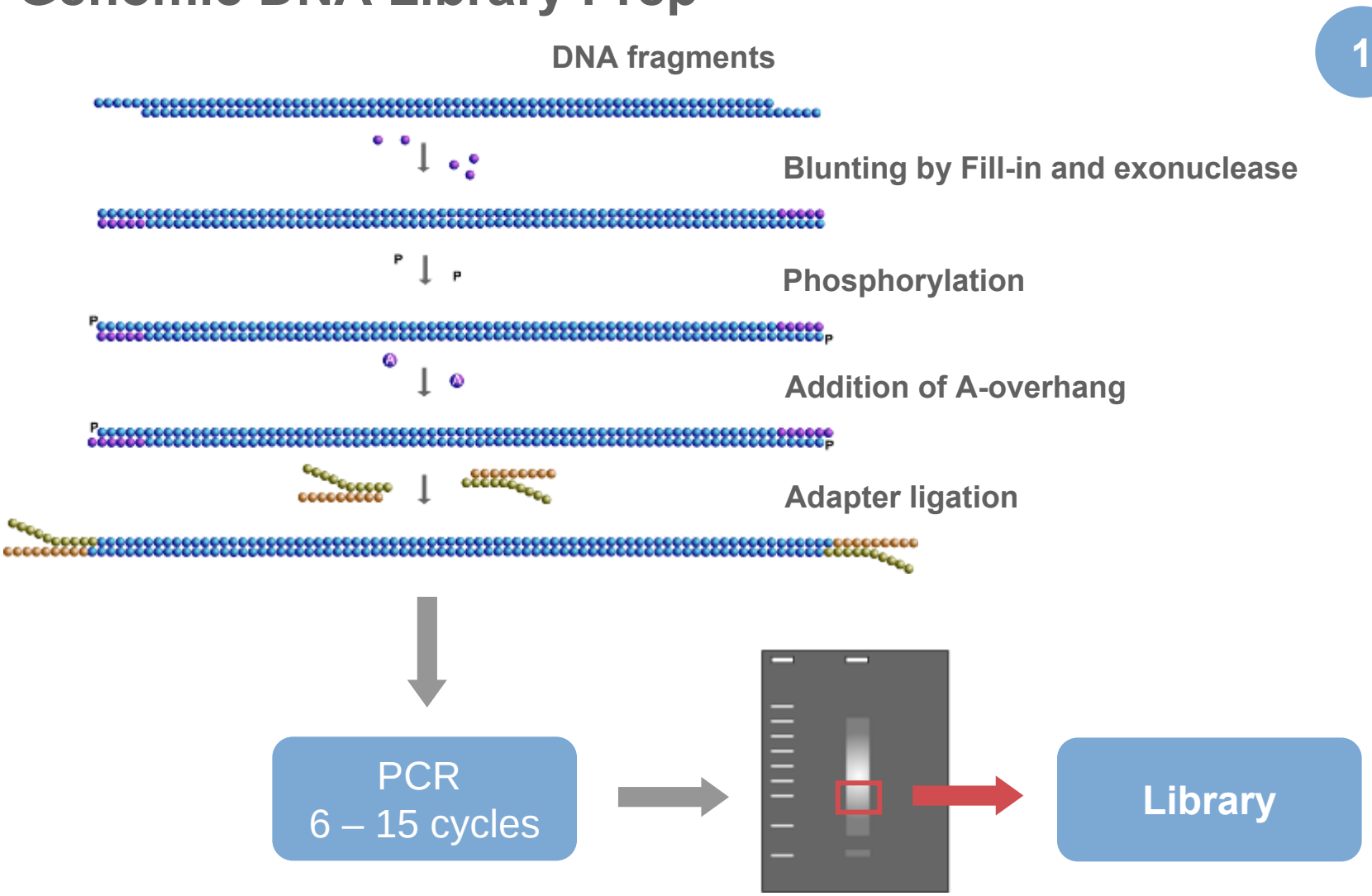
Perform sequencing on forward strand
↓
Re-generate reverse strand
↓
Perform sequencing on reverse strand

Flow cell



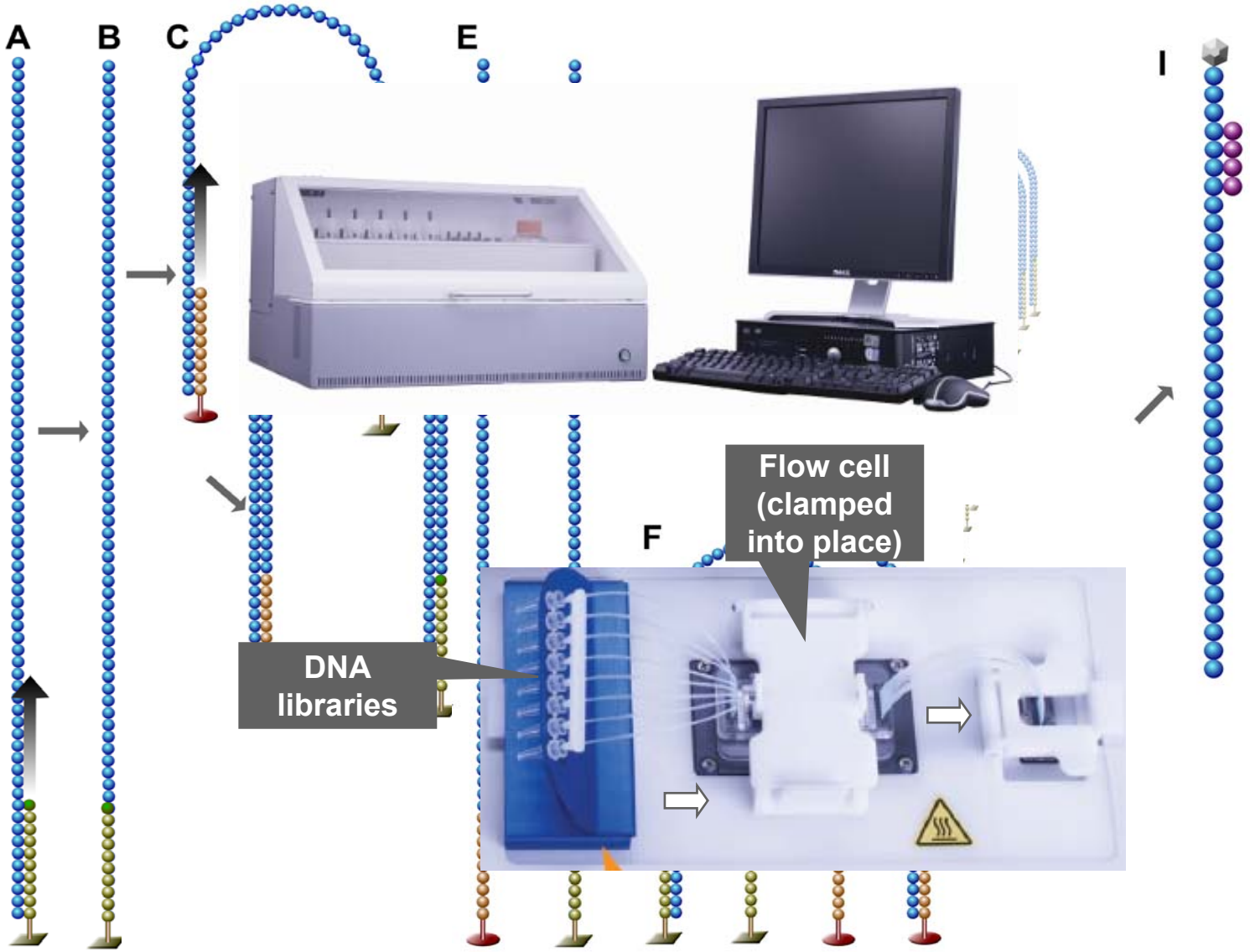
- Contained environment
- No need for clean rooms
- Sequencing performed inside the flow cell

Genomic DNA Library Prep



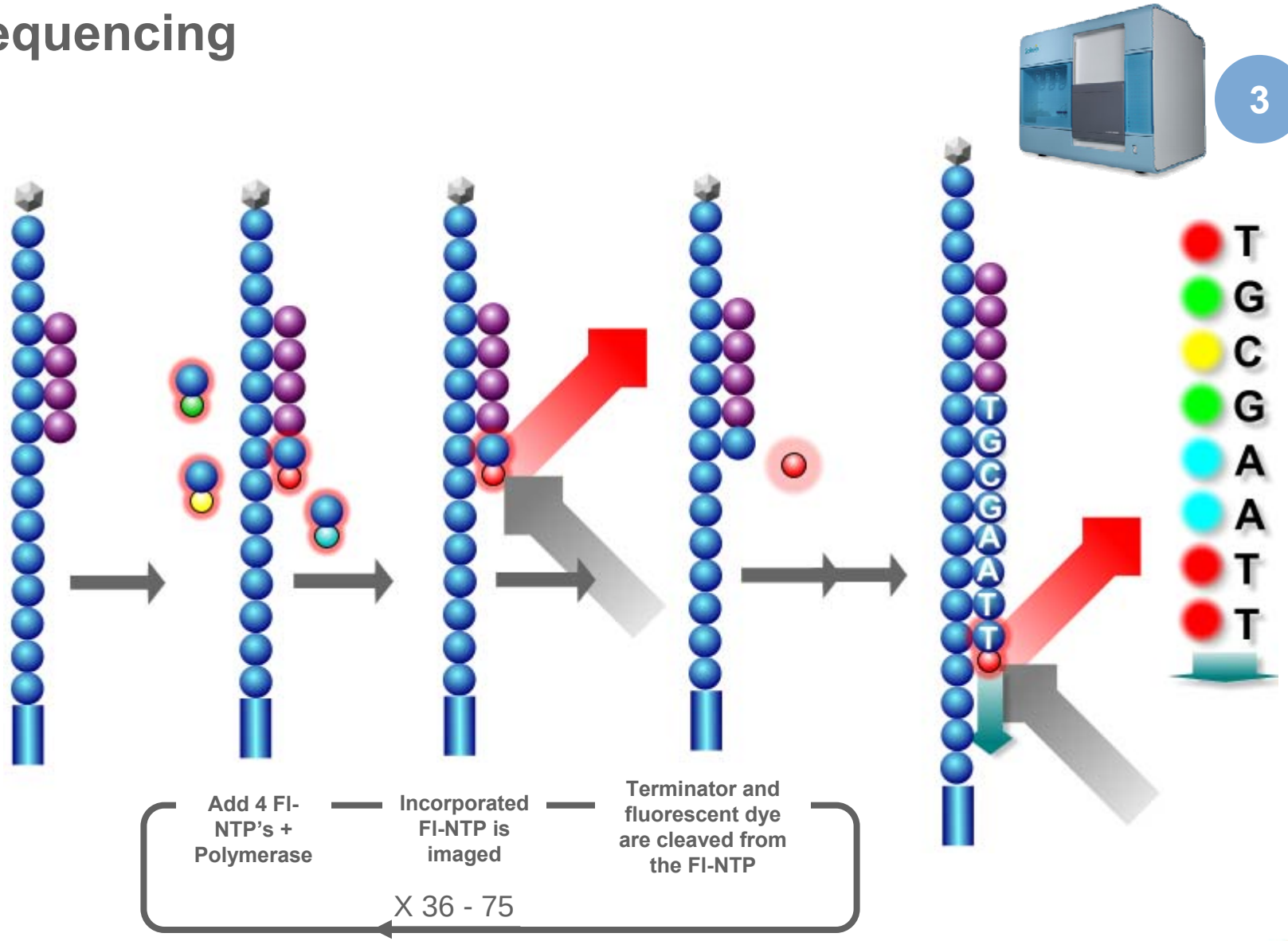
Cluster Generation

2



illumina

Sequencing


3
T
G
C
G
A
A
T
T
TGCGAATT
Add 4 Fl-NTP's + Polymerase
Incorporated Fl-NTP is imaged
Terminator and fluorescent dye are cleaved from the Fl-NTP
X 36 - 75

Sequencing with Paired Ends

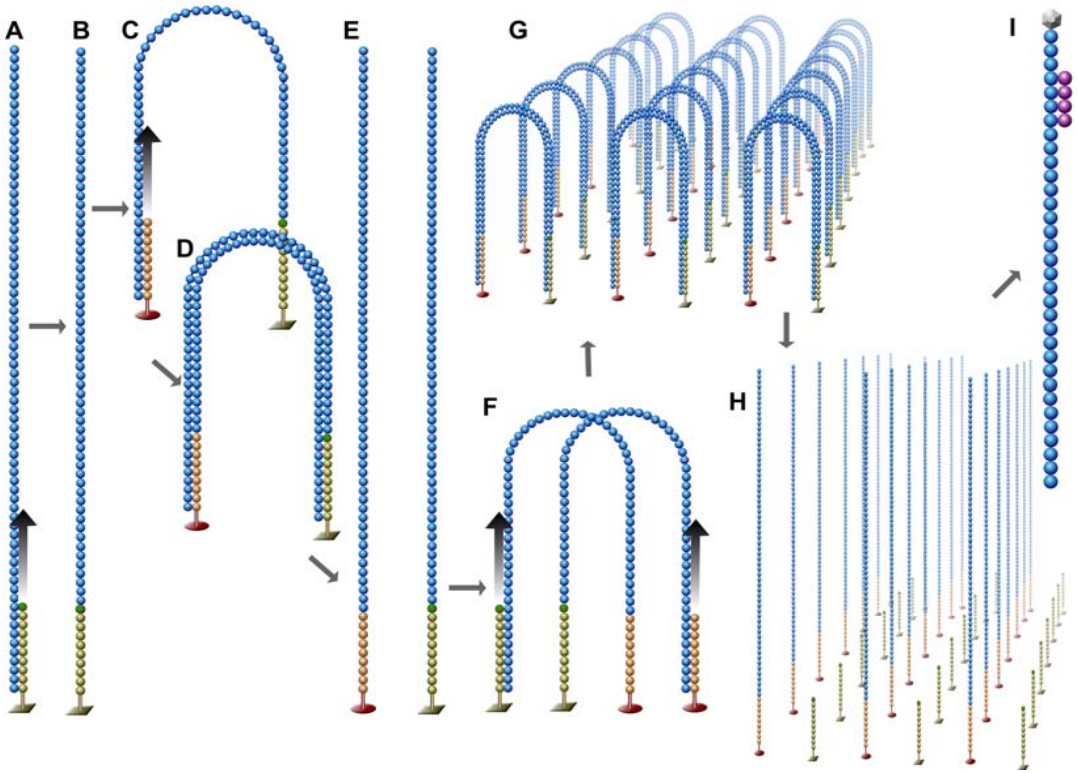
Reference	This is really the best way to do sequencing
Single-reads	This is
……	is really
……	really the
……	the best
……	sequencing
Paired-reads	This is (------26 characters-------) sequencing

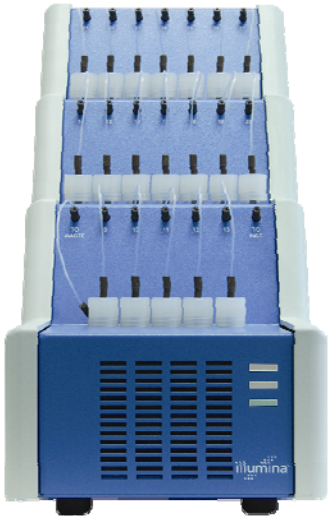
Assembly becomes easier!!

illumina

Paired-End Sequencing

- Sequenced strand is stripped off
- Unblock the 3'ends of templates and lawn primers
- Regenerate clusters and cleave forward strand → sequence reverse strand





Broadest range of applications

Optimized, streamlined and easy-to-use reagent solutions

Sample Prep

Whole genome
- Resequencing
- De-novo
- Targeted

Transcriptome
- RNA-Seq
- DGE
- Small RNA
- miRNA

Regulation
- Methylation
- ChIP-Seq

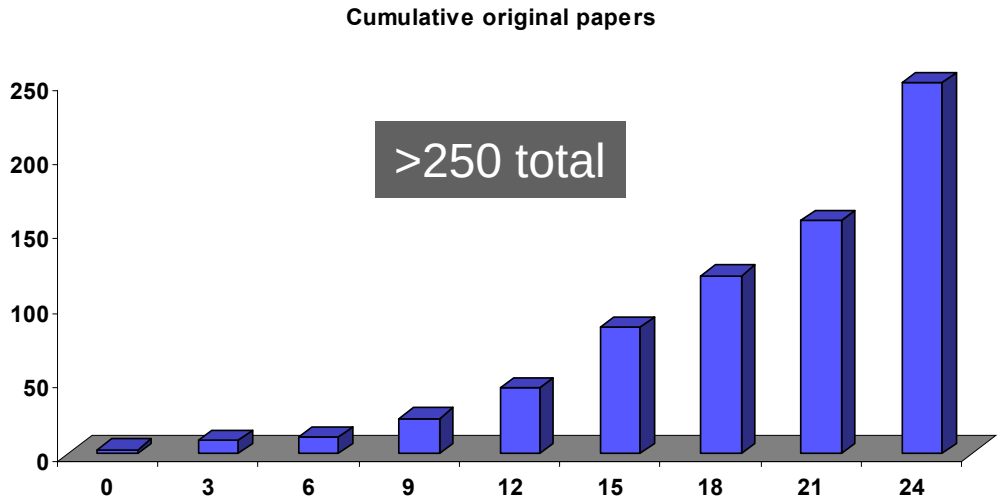
Automated Cluster Generation

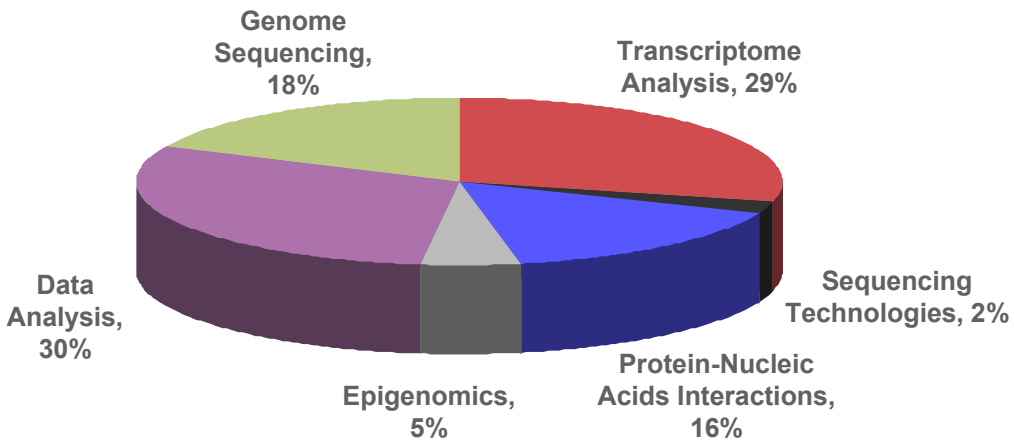
Sequencing

illumina

GA Applications Published to February 2009

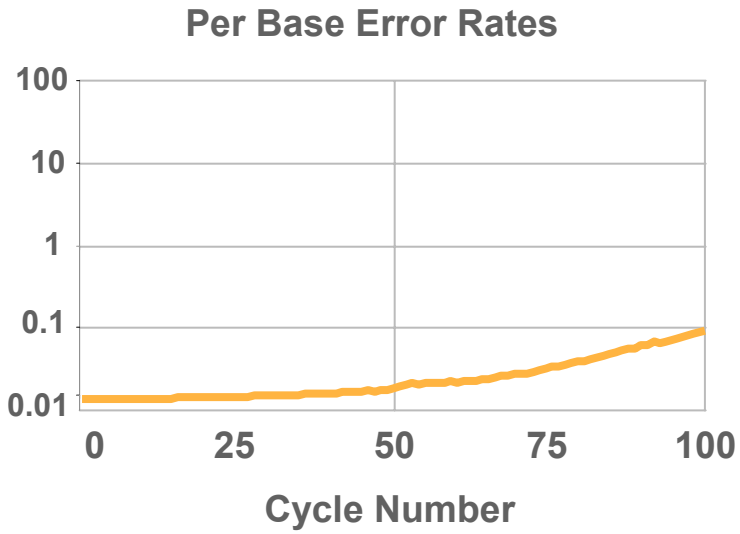
nature	42
Science	14





de-novo Sequencing

READ LENGTH†	RUN TIME (DAYS)	PER BASE READ ACCURACY	% PERFECT READS
1 × 35 bp	~ 2.5	≥ 99%	≥ 90%
2 × 35 bp	~ 5	≥ 99%	≥ 90%
2 × 50 bp	~ 6.5	> 98.5%	≥ 80%
2 × 75 bp	~ 9.5	≥ 98.5%	≥ 70%



Genome	Researchers
Apple Scab *(Venturia inaequalis)*	U Western Cape, SA
Buchnera aphidicola	U Arizona
Human	BGI
Giant Panda	BGI
Eight Pine Plastomes	USDA Forest Service
Pseudomonas syringae	Sainsbury Lab

illumina®

Complex Genome Sequencing

De-novo Sequencing - Giant Panda

Beijing Genomics Institute

- 2 x 75 bp reads
- 50x coverage of the 3Gb genome
- N50 contig size of ~300 kb

Ancient DNA - Neandertal

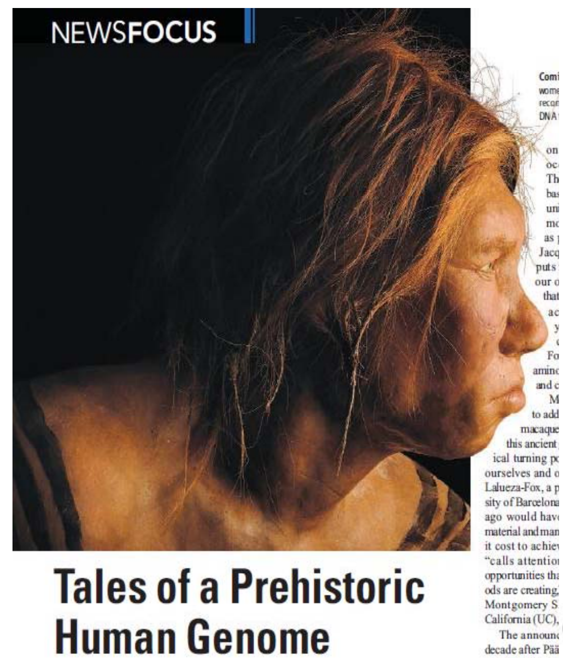
(*Svante Pääbo, MPI*)

- Illumina and the Roche/454 platforms
- 1.5x coverage
- Plans for 15- to 20x coverage

“Taste good, sequence it”
“Look cute, sequence it.”

Wine grape, the Tree of Life sequencing projects and the Giant Panda Project

“data suggest that the human and Neandertal lineages began to diverge some 800,000 years ago”


NEWSFOCUS

Tales of a Prehistoric Human Genome

Cancer Genome Sequencing

The First Complete Cancer Genome

DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome

Timothy J. Ley[1,2,3,4]*, Elaine R. Mardis[2,3]*, Li Ding[2,3], Bob Fulton[3], Michael D. McLellan[3], Ken Chen[3], David Dooling[3], Brian H. Dunford-Shore[3], Sean McGrath[3], Matthew Hickenbotham[3], Lisa Cook[3], Rachel Abbott[3], David E. Larson[3], Dan C. Koboldt[3], Craig Pohl[3], Scott Smith[3], Amy Hawkins[3], Scott Abbott[3], Devin Locke[3], LaDeana W. Hillier[3,8], Tracie Miner[3], Lucinda Fulton[3], Vincent Magrini[2,3], Todd Wylie[3], Jarret Glasscock[3], Joshua Conyers[3], Nathan Sander[3], Xiaoqi Shi[3], John R. Osborne[3], Patrick Minx[3], David Gordon[8], Asif Chinwalla[3], Yu Zhao[1], Rhonda E. Ries[1], Jacqueline E. Payton[5], Peter Westervelt[1,4], Michael H. Tomasson[1,4], Mark Watson[3,4,5], Jack Baty[6], Jennifer Ivanovich[4,7], Sharon Heath[1,4], William D. Shannon[1,4], Rakesh Nagarajan[4,5], Matthew J. Walter[1,4], Daniel C. Link[1,4], Timothy A. Graubert[1,4], John F. DiPersio[1,4] & Richard K. Wilson[2,3,4]

Nature 2008. 456:66-72

nature

- Sequenced leukemia genome vs. matched normal (skin)
- Low sample input - primary tumor could be studied rather than cell lines
- 8 new mutations discovered in AML (coding genes)
  - Out of millions of total SNPs

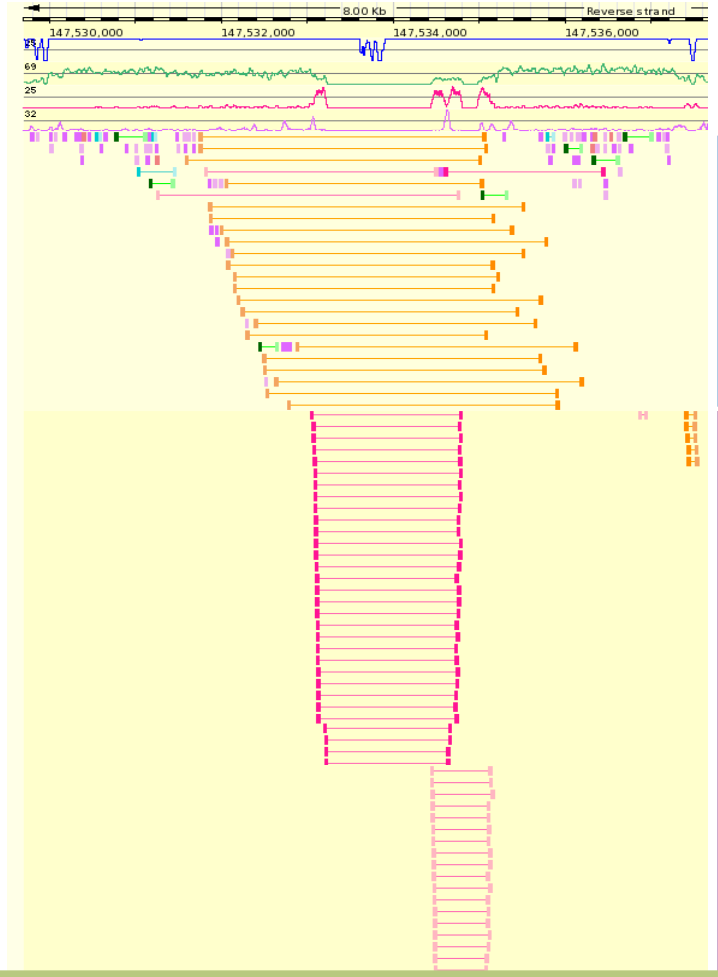
"Most of these genes would not have been candidates for directed re-sequencing on the basis of our current understanding of cancer"

300 cancer genomes in next 12 months

illumina

Short and Long Insert Paired-End Reads

The ultimate combination for detecting structural variation



Long Inserts | **Deletion?**

Up 5 kb (10kb EA)

Short Inserts | **Inversions?**

200bp – 500bp
100ng gDNA

Genomes are complex and dynamic – you need a flexible set of tools

The Genome Analyzer$_{IIx}$ and Software Advancements

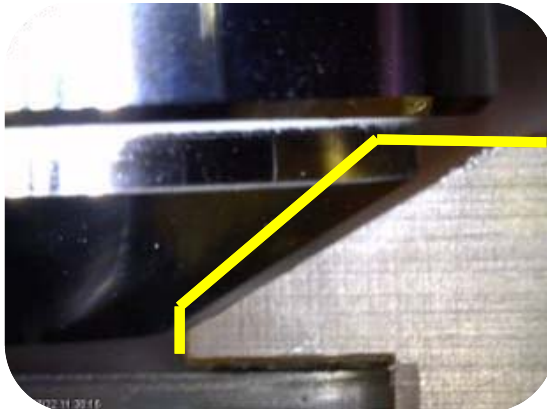
65% Increase in data output & simplified computing

- 20-25 Gb of high quality data / run
- 2.5 Gb / day
- >300M reads per paired-end run
- 2 x 75bp supported read length
- Raw Accuracy: ≥ 98.5%

Genome Analyzer$_{IIx}$

Increased output and ease-of-use for long reads

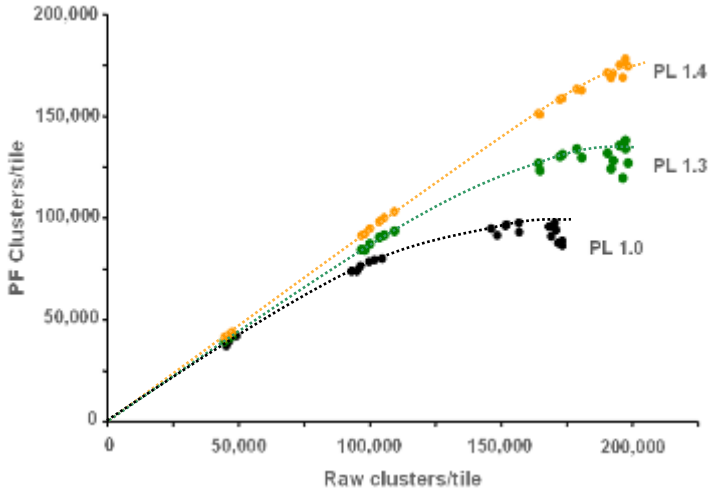
Improved manifold design increases reads per flow cell by 20%



Larger reagent chiller enables long read sequencing runs (100+bp)

New Software Delivers Up to 40% More Data Per Run

Increases yield, improves accuracy & lowers error rates



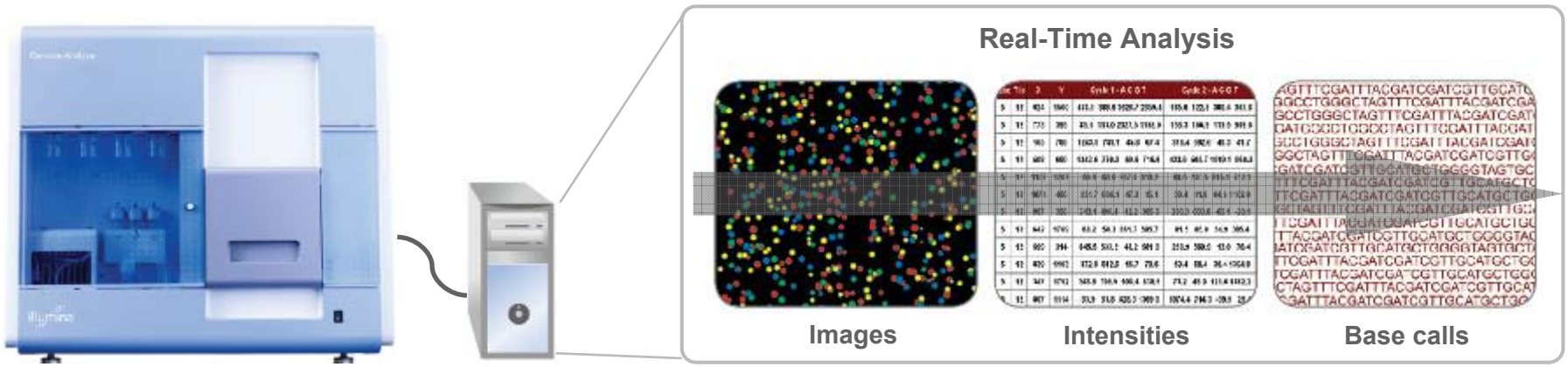
illumina®

Increased output with reduced computing infrastructure

More gigabases of data for fewer gigabytes of computing power!

Sequencing Control Software v2.4

- Includes new Real Time Analysis (RTA) feature
- Image extraction and real time base calling on instrument computer
- Shorter time to results
  - Performed simultaneously with sequencing
  - Eliminates need to transfer images and intensities across network
  - Base calls and quality scores within hours of end of run



Targeted Resequencing

- Bottleneck in NGS workflow
  - Ultimate resolution for mutation discovery is sequencing
  - Cost of whole genome sequencing is prohibitive for most researchers
  - Specific regions of interest are preferred – candidate regions

"Ability to fully leverage the power of NGS is crippled by the lack of corresponding 'front end' targeting technologies"
Porreca et al., 2007. Nature Methods

illumina®

What is Targeted Resequencing?

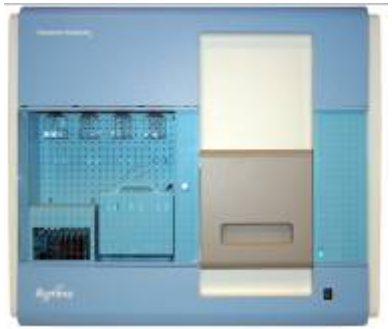
- What is targeted resequencing?
  - Target Enrichment, Sequence Capture, Genome Partitioning, Genomic Capture, Target Capture, Targeted Pullout
- Focus on subset of the genome
  - Remaining genomic material discarded
- Find exact disease causing genes and mutations
  - Gene families (eg. kinases), human exome, gene pathways, any candidate region, GWAS follow-up
- **Understand complex human traits**
  - Discover rare variants
  - SNPs, small insertions and deletions

illumina®

Agilent SureSelect Targeted Enrichment System

- Co-marketing agreement between Illumina and Agilent
  - Licensed capture protocol developed by the Broad Institute*
- Customizable, any regions of interest
- Optimized and validated on GA
- Uniform coverage, high reproducibility and specificity, little or no bias
- Low DNA input
- 75bp+ long reads provide even coverage across captured regions

Agilent Technologies
SureSelect™
Target Enrichment System



Nature Biotechnology **27**, 182 - 189 (2009)
Published online: 1 February 2009 | doi:10.1038/nbt.1523

Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing

Andreas Gnirke[1], Alexandre Melnikov[1], Jared Maguire[1], Peter Rogov[1], Emily M LeProust[2], William Brockman[1,5], Timothy Fennell[1], Georgia Giannoukos[1], Sheila Fisher[1], Carsten Russ[1], Stacey Gabriel[1], David B Jaffe[1], Eric S Lander[1,3,4] & Chad Nusbaum[1]

illumina

Delivering on Roadmap Milestones

15x increase in 2008

4-5x increase in 2009