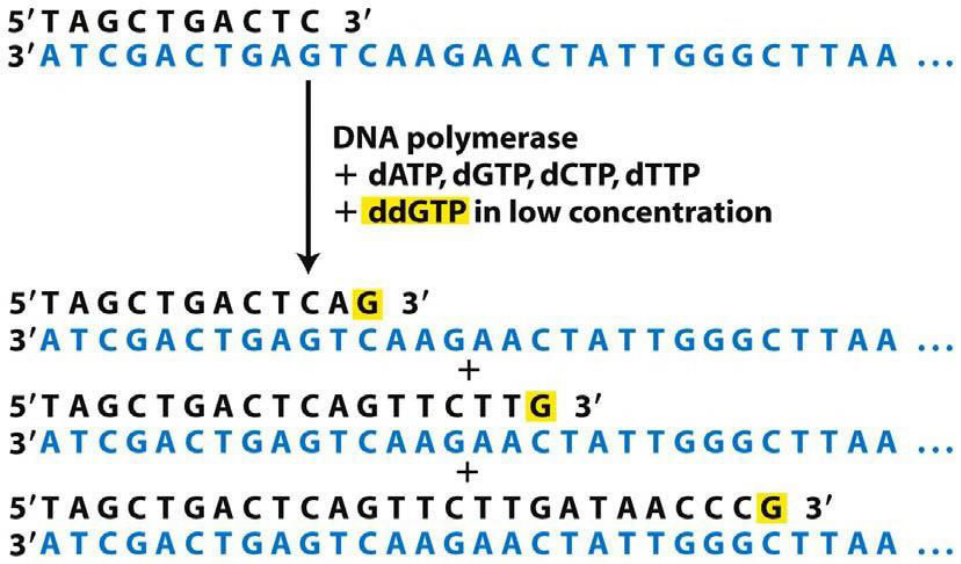
RAMA UNIVERSITY

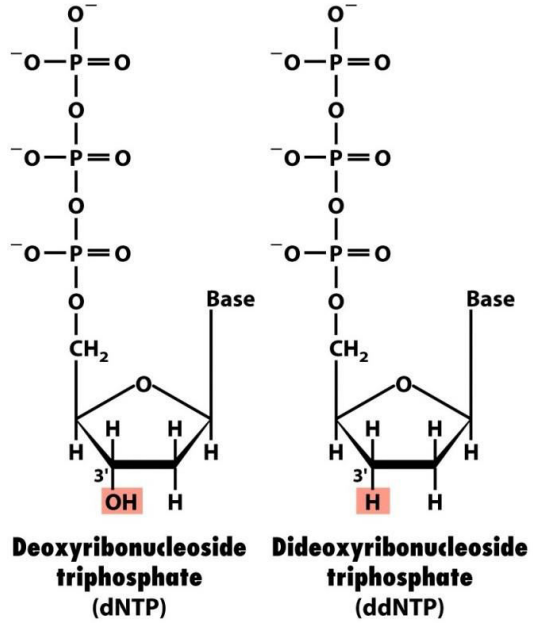
www.ramauniversity.ac.in

FACULTY OF ENGINEERING &TECHNOLOGY
DEPARTMENT OF BIOTECHNOLOGY

Dr. Simranjit Singh
Assistant Professor
Department of Biotechnology
Rama University, Kanpur

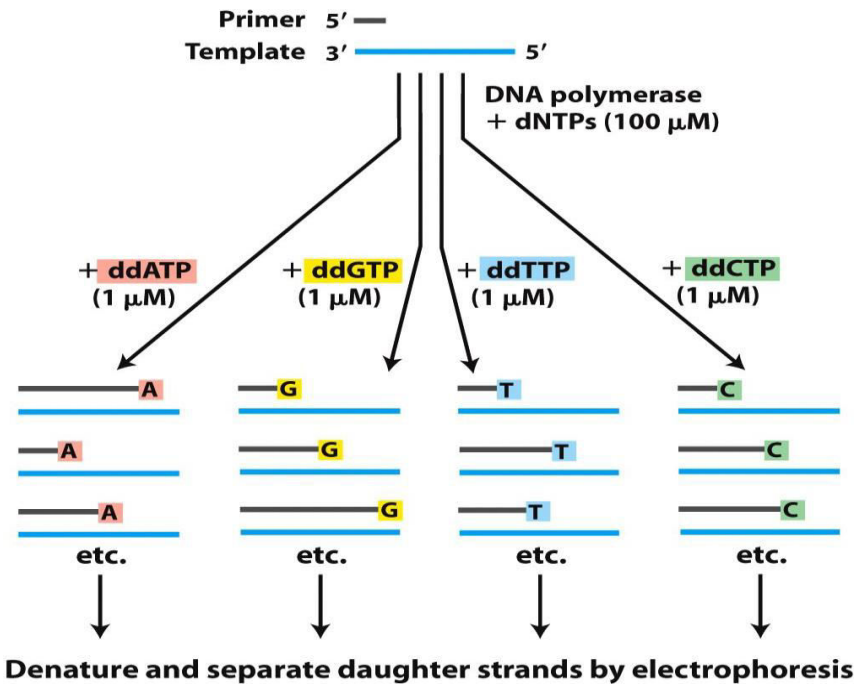
DNA sequencing: the Sanger (di-deoxy) method

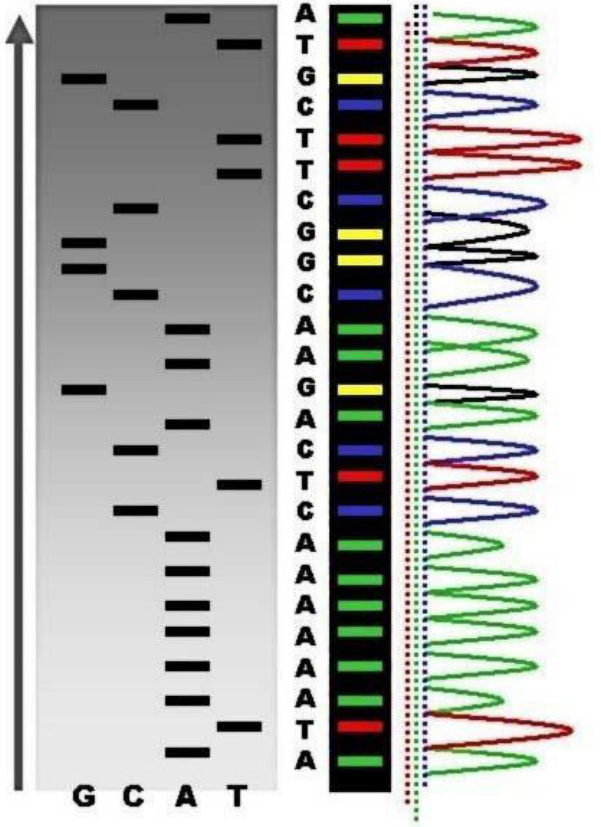




Sequencing

DNA sequencing: the Sanger (dideoxy) method





DNA sequencing: the Sanger (dideoxy) method

Next generation sequencing

Technologies/Platforms:

- Roche/454 FLX: 2004
- Illumina Solexa Genome Analyzer: 2006
- Applied Biosystems SOLiD System: 2007
- Helicos Heliscope: 2010
- Pacific Biosciencies SMRT: 2010
- LifeTechnologies Ion Torrent: 2011

Parameters:

- Cost (device, cost/Mb)
- Read length
- Speed
- Accuracy
- Preperation time
- Manipulation steps (amplification needed)

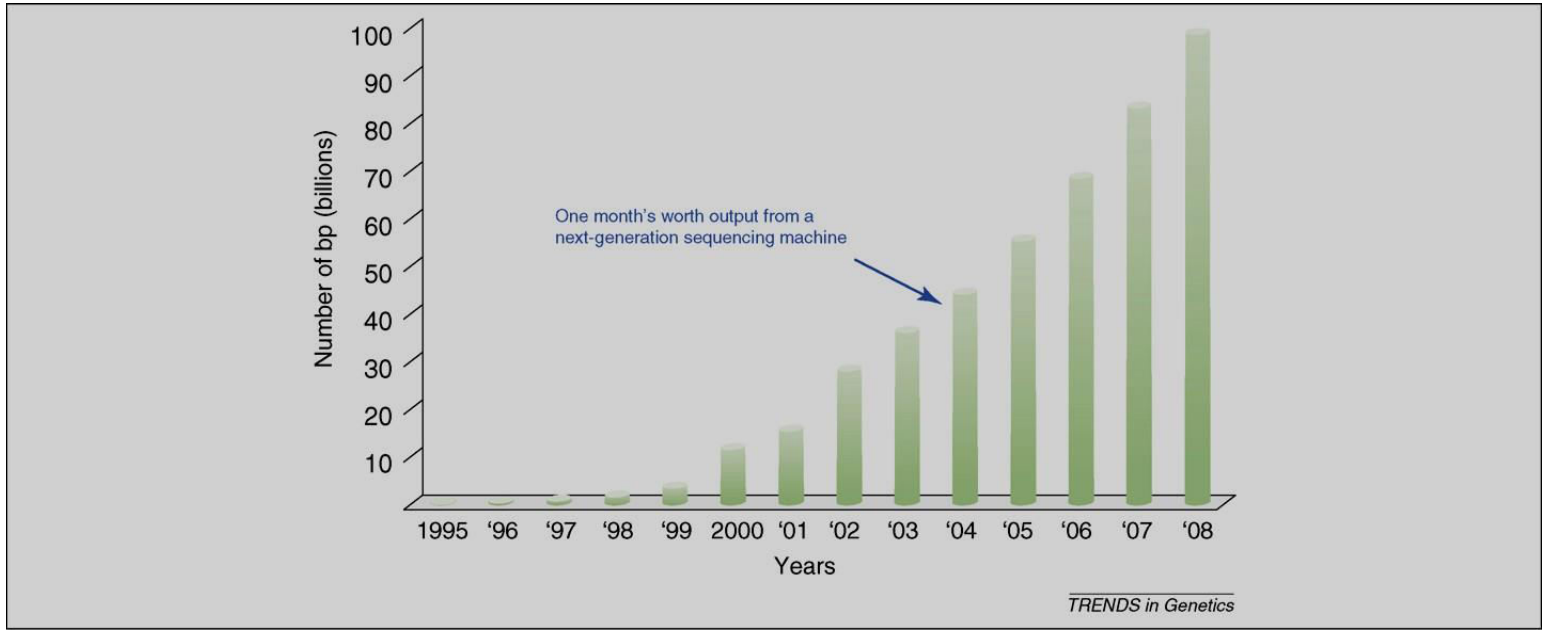


Table 1 | **Comparison of next-generation sequencing platforms**

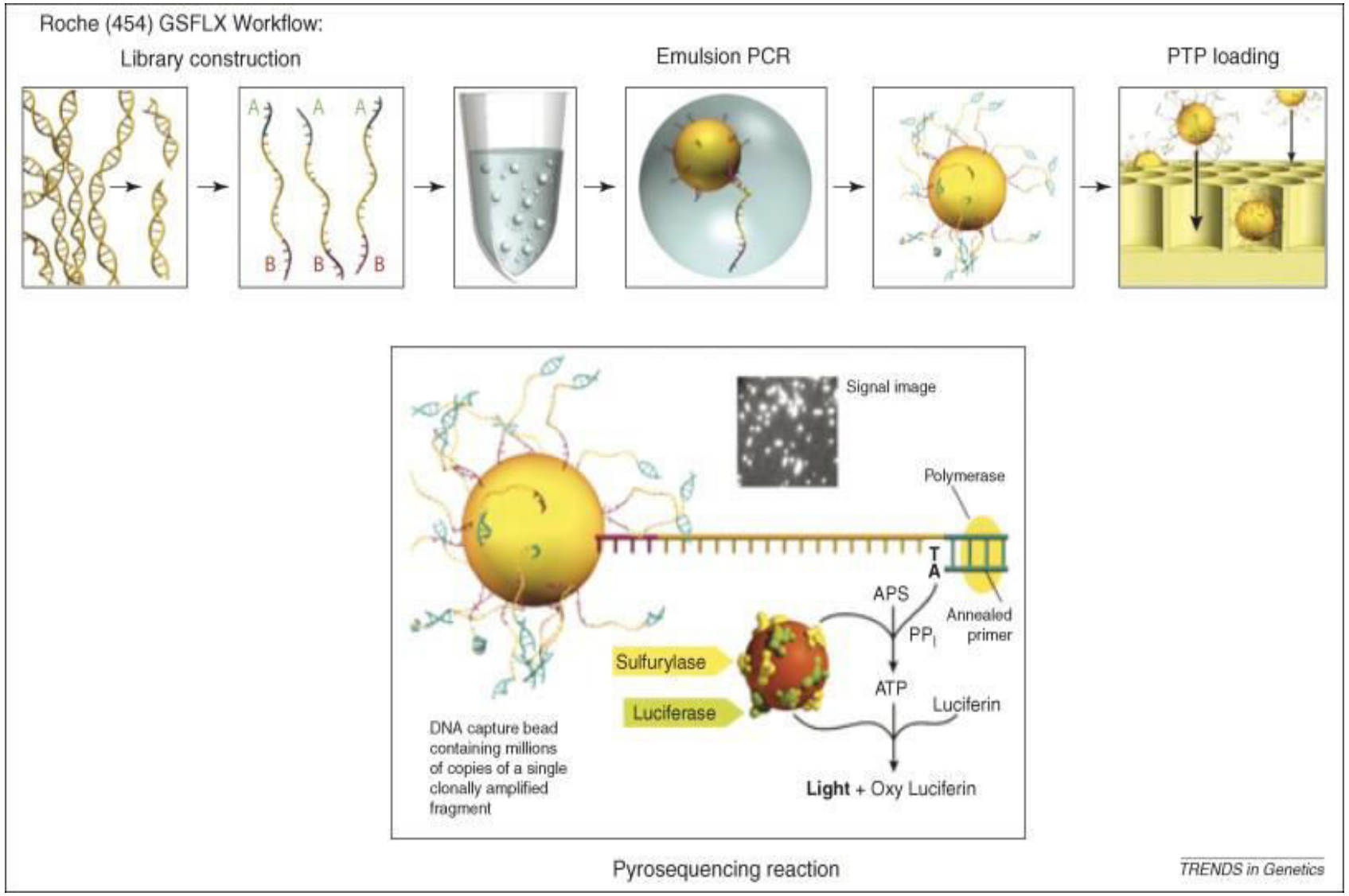
Platform	Library/ template preparation	NGS chemistry	Read length (bases)	Run time (days)	Gb per run	Machine cost (US$)	Pros	Cons	Biological applications	Refs
Roche/454's GS FLX Titanium	Frag, MP/ emPCR	PS	330*	0.35	0.45	500,000	Longer reads improve mapping in repetitive regions; fast run times	High reagent cost; high error rates in homo-polymer repeats	Bacterial and insect genome *de novo* assemblies; medium scale (<3 Mb) exome capture; 16S in metagenomics	D. Muzny, pers. comm.
Illumina/ Solexa's GA$_{II}$	Frag, MP/ solid-phase	RTs	75 or 100	4‡, 9§	18†, 35§	540,000	Currently the most widely used platform in the field	Low multiplexing capability of samples	Variant discovery by whole-genome resequencing or whole-exome capture; gene discovery in metagenomics	D. Muzny, pers. comm.
Life/APG's SOLiD 3	Frag, MP/ emPCR	Cleavable probe SBL	50	7‡, 14§	30†, 50§	595,000	Two-base encoding provides inherent error correction	Long run times	Variant discovery by whole-genome resequencing or whole-exome capture; gene discovery in metagenomics	D. Muzny, pers. comm.
Polonator G.007	MP only/ emPCR	Non-cleavable probe SBL	26	5§	12§	170,000	Least expensive platform; open source to adapt alternative NGS chemistries	Users are required to maintain and quality control reagents; shortest NGS read lengths	Bacterial genome resequencing for variant discovery	J. Edwards, pers. comm.
Helicos BioSciences HeliScope	Frag, MP/ single molecule	RTs	32*	8‡	37†	999,000	Non-bias representation of templates for genome and seq-based applications	High error rates compared with other reversible terminator chemistries	Seq-based methods	91
Pacific Biosciences (target release: 2010)	Frag only/ single molecule	Real-time	964*	N/A	N/A	N/A	Has the greatest potential for reads exceeding 1 kb	Highest error rates compared with other NGS chemistries	Full-length transcriptome sequencing; complements other resequencing efforts in discovering large structural variants and haplotype blocks	S. Turner, pers. comm.

*Average read-lengths. ‡Fragment run. §Mate-pair run. Frag, fragment; GA, Genome Analyzer; GS, Genome Sequencer; MP, mate-pair; N/A, not available; NGS, next-generation sequencing; PS, pyrosequencing; RT, reversible terminator; SBL, sequencing by ligation; SOLiD, support oligonucleotide ligation detection.

Next generation sequencing

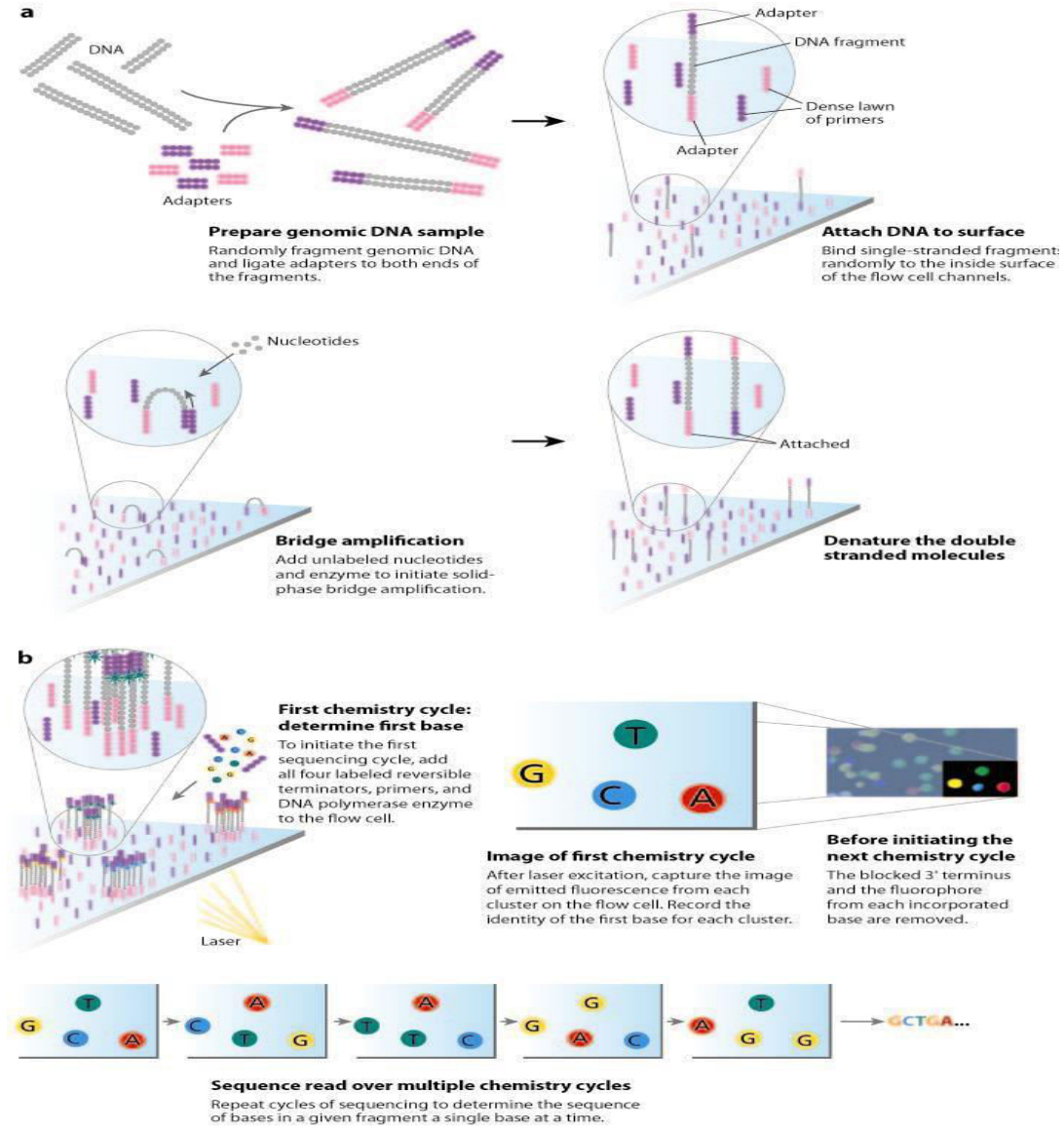
	Roche (454)	Illumina	SOLiD
Chemistry	Pyrosequencing	Polymerase-based	Ligation-based
Amplification	Emulsion PCR	Bridge Amp	Emulsion PCR
Paired ends/sep	Yes/3kb	Yes/200 bp	Yes/3 kb
Mb/run	100 Mb	1300 Mb	3000 Mb
Time/run	7 h	4 days	5 days
Read length	250 bp	32-40 bp	35 bp
Cost per run (total)	$8439	$8950	$17447
Cost per Mb	$84.39	$5.97	$5.81

Roche (454) Workflow



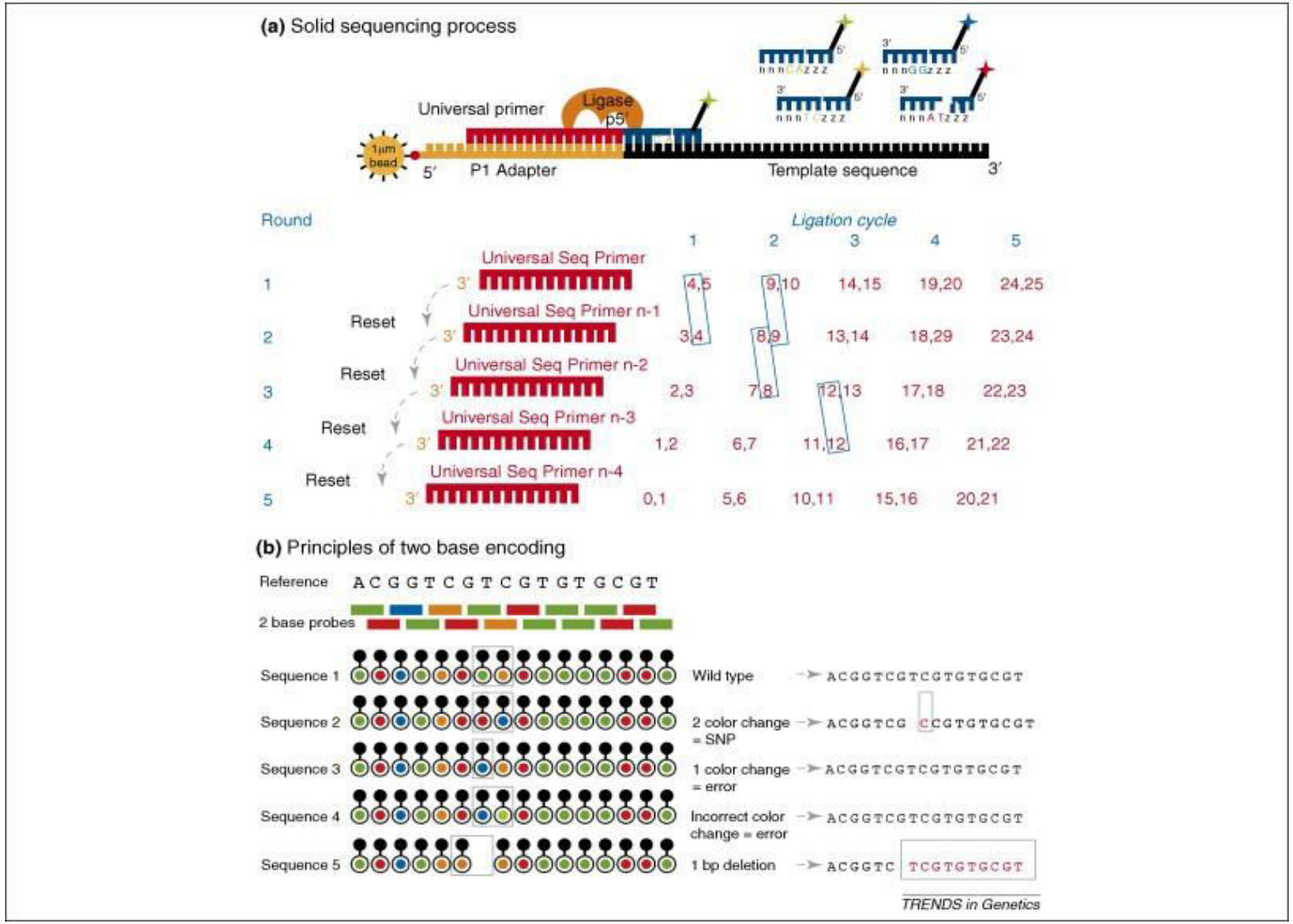
Sequencing

Illumina (Solexa) Workflow



Mardis ER. 2008.
Annu. Rev. Genomics Hum. Genet. 9:387–402

ABI SOLiD Workflow



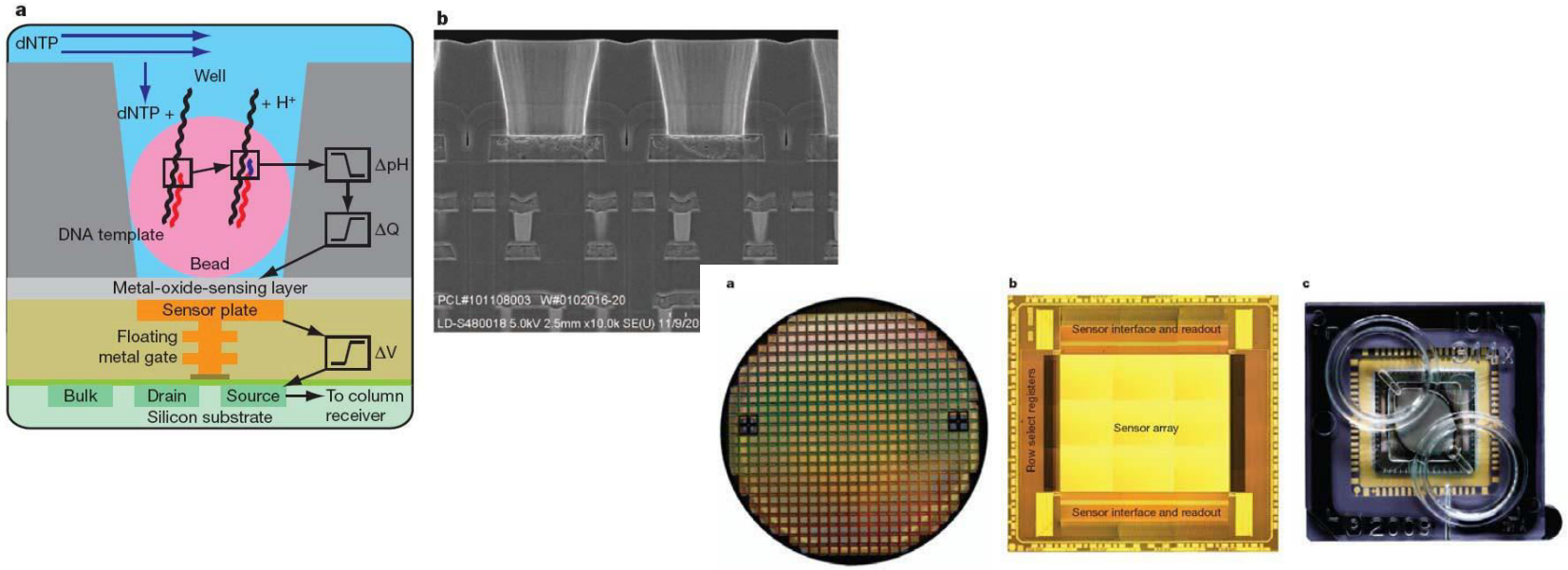
Sequencing

Ion Torrent

348 | NATURE | VOL 475 | 21 JULY 2011

An integrated semiconductor device enabling non-optical genome sequencing

Jonathan M. Rothberg[1], Wolfgang Hinz[1], Todd M. Rearick[1], Jonathan Schultz[1], William Mileski[1], Mel Davey[1], John H. Leamon[1], Kim Johnson[1], Mark J. Milgrew[1], Matthew Edwards[1], Jeremy Hoon[1], Jan F. Simons[1], David Marran[1], Jason W. Myers[1], John F. Davidson[1], Annika Branting[1], John R. Nobile[1], Bernard P. Puc[1], David Light[1], Travis A. Clark[1], Martin Huber[1], Jeffrey T. Branciforte[1], Isaac B. Stoner[1], Simon E. Cawley[1], Michael Lyons[1], Yutao Fu[1], Nils Homer[1], Marina Sedova[1], Xin Miao[1], Brian Reed[1], Jeffrey Sabina[1], Erika Feierstein[1], Michelle Schorn[1], Mohammad Alanjary[1], Eileen Dimalanta[1], Devin Dressman[1], Rachel Kasinskas[1], Tanya Sokolsky[1], Jacqueline A. Fidanza[1], Eugeni Namsaraev[1], Kevin J. McKernan[1], Alan Williams[1], G. Thomas Roth[1] & James Bustillo[1]



Applications

- Genomes
- Re-sequencing Human Exons (Microarray capture/amplification)
- small (including mi-RNA) and long RNA profiling (including splicing)
- ChIP-Seq:
  - Transcription Factors
  - Histone Modifications
  - Effector Proteins
- DNA Methylation
- Polysomal RNA
- Origins of Replication/Replicating DNA
- Whole Genome Association (rare, high impact SNPs)
- Copy Number/Structural Variation in DNA
- ChIA-PET: Transcription Factor Looping Interactions
- The $1000 genome

Current bottle neck: Data management!!!

The New York Times **DNA Sequencing Caught in Deluge of Data (Published: November 30, 2011)**