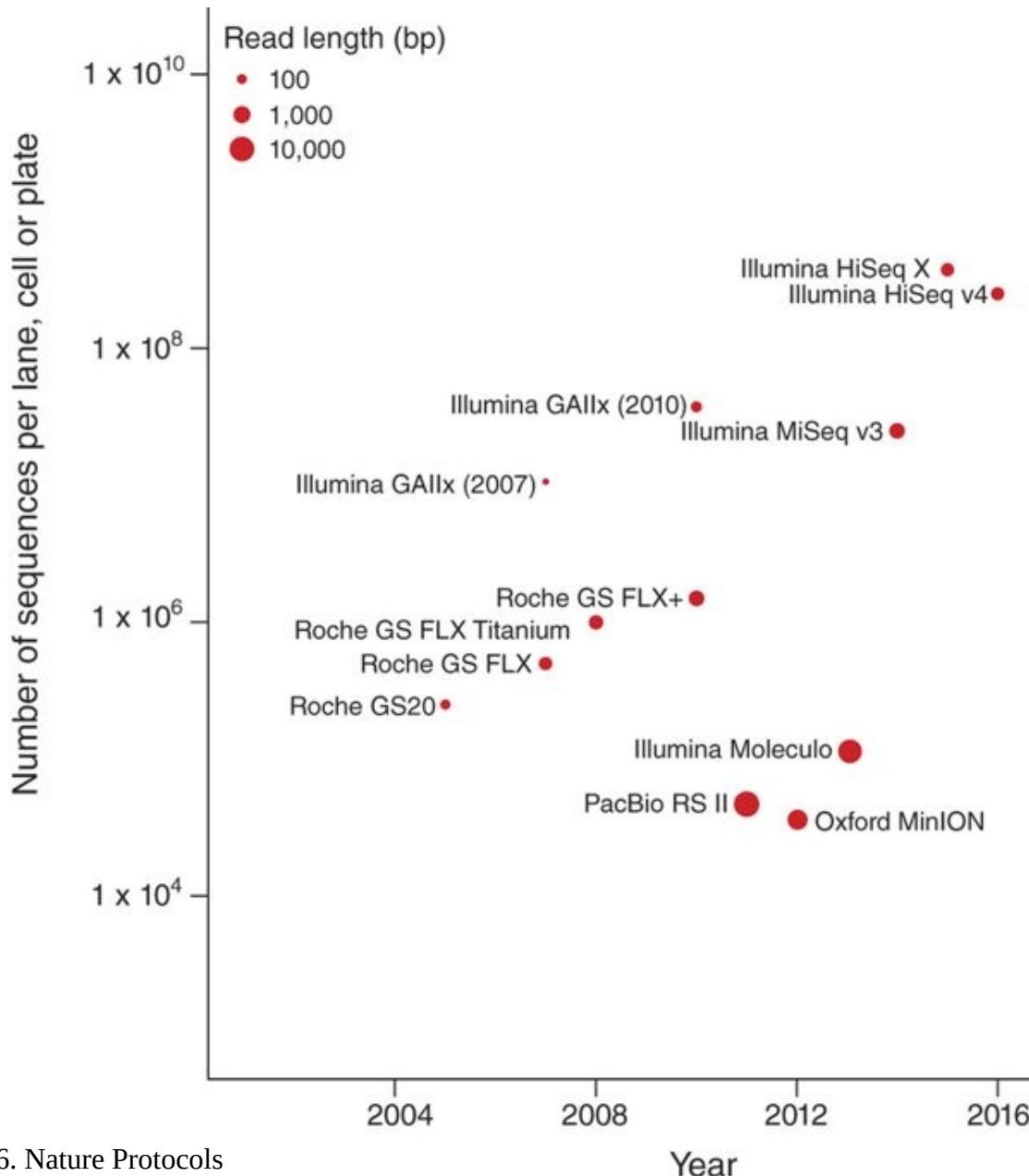


Dr. Richard Allen White III
Washington State University
FABI Workshop - Oct 26, 2018
Day 5

Learning Objectives

1. Sequencing technologies past, present and future
2. Metagenomic binning review
3. Metatranscriptomics sample prep
4. Metatranscriptomics analysis reads based
5. Metatranscriptomics analysis assembly based

Sequencing technologies

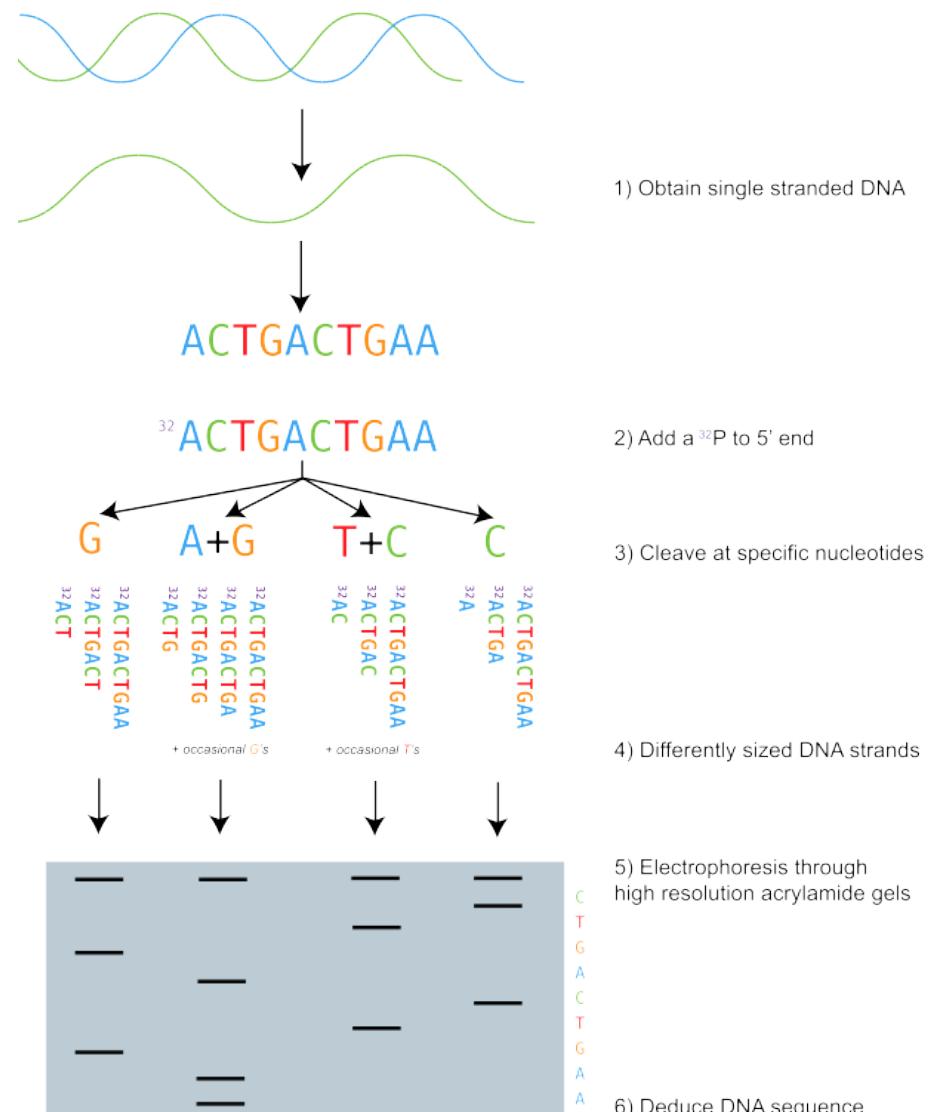


Sequencing platform - maxam gilbert

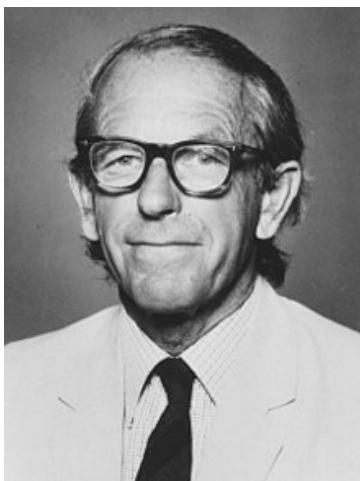


Walter Gilbert
1932 - Current

1976-1977

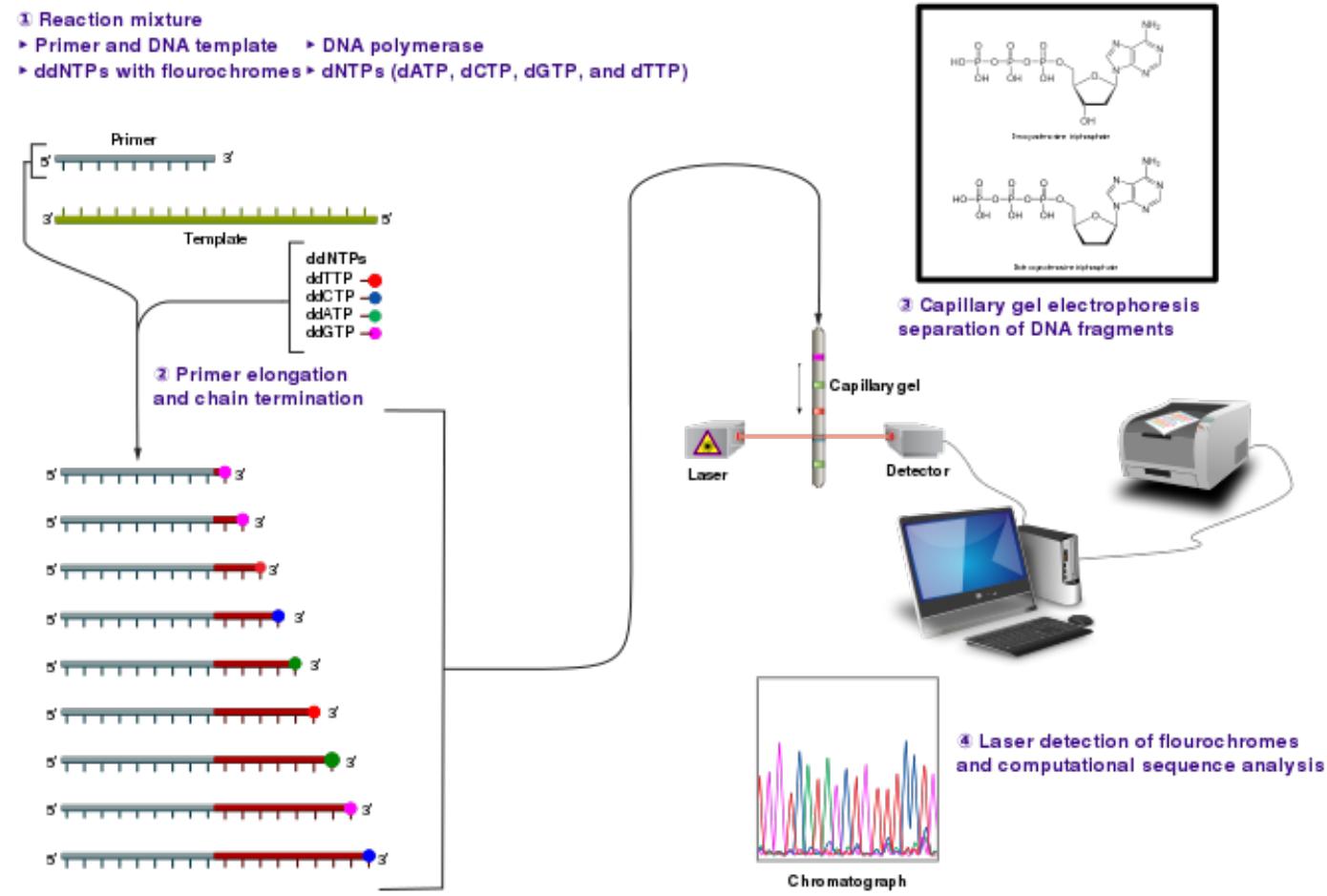
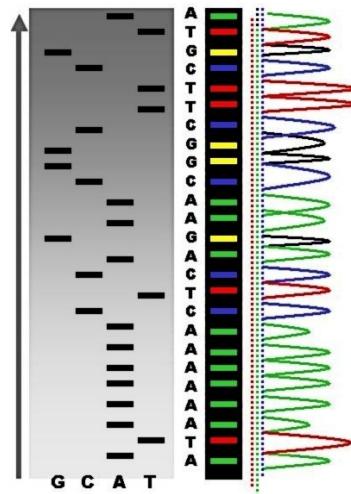


Sequencing platform - Sanger method dideoxy-chain termination

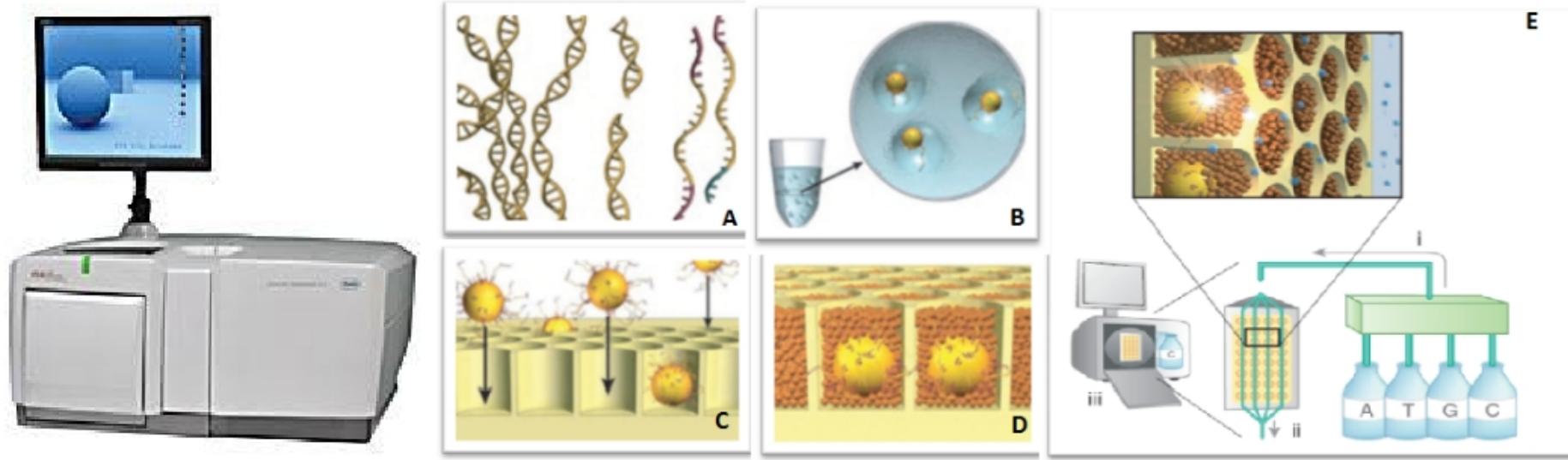


Fred Sanger
1932 - 2013

1976-1977



Sequencing platform - pyrosequencing Roche 454



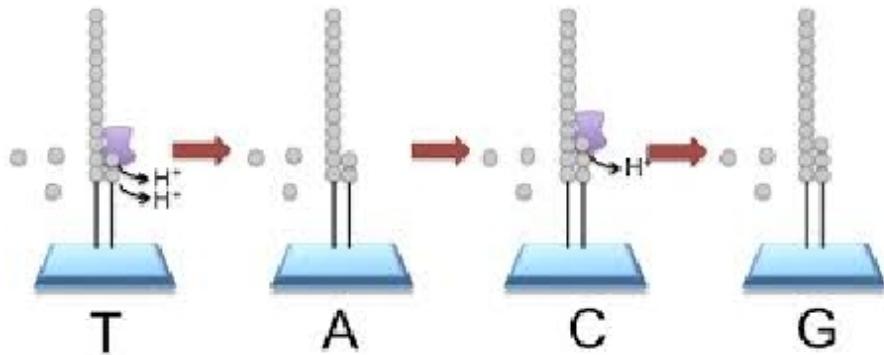
PRO

- First NGS sequencer
- Shorter than sanger
- Up to 1 M reads

CON

- Few reads ~ 1 M
- High Insertion/Deletion (INDEL) errors (homopolymers)
- NOW EXTINCT 2016

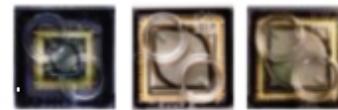
Sequencing platform - pH detection base Life Tech Ion Torrent/Ion Proton



Ion PGM™ Sequencer



Ion Proton™ Sequencer



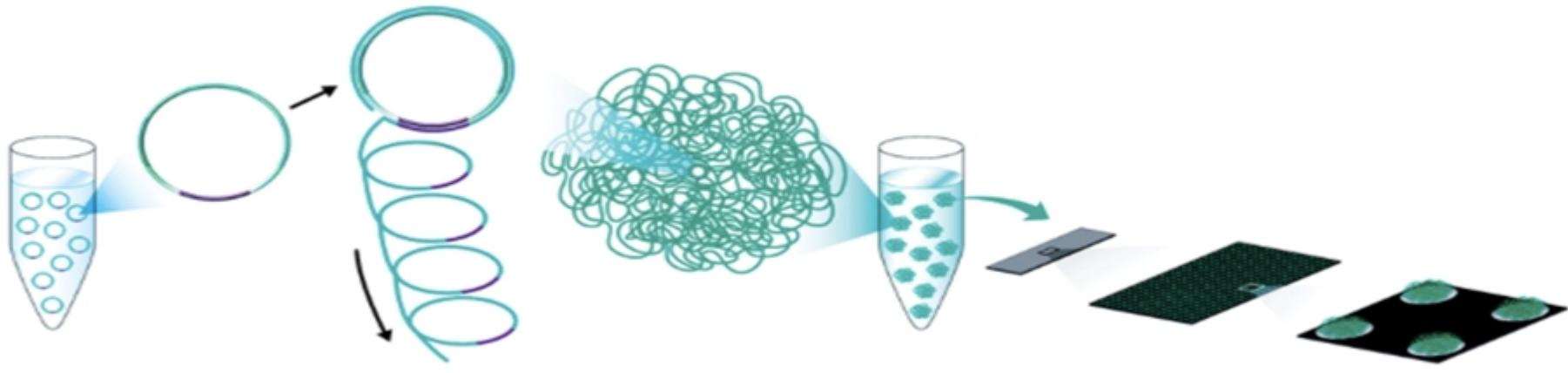
PRO

- NGS sequencer
- Shorter than sanger
- Up to 1 M reads

CON

- Few reads ~ 1 M
- High Insertion/Deletion (INDEL) errors (homopolymers)
- NOW EXTINCT ?

Sequencing platform – Complete genomics (MGI-BGI) Nanoballs technology



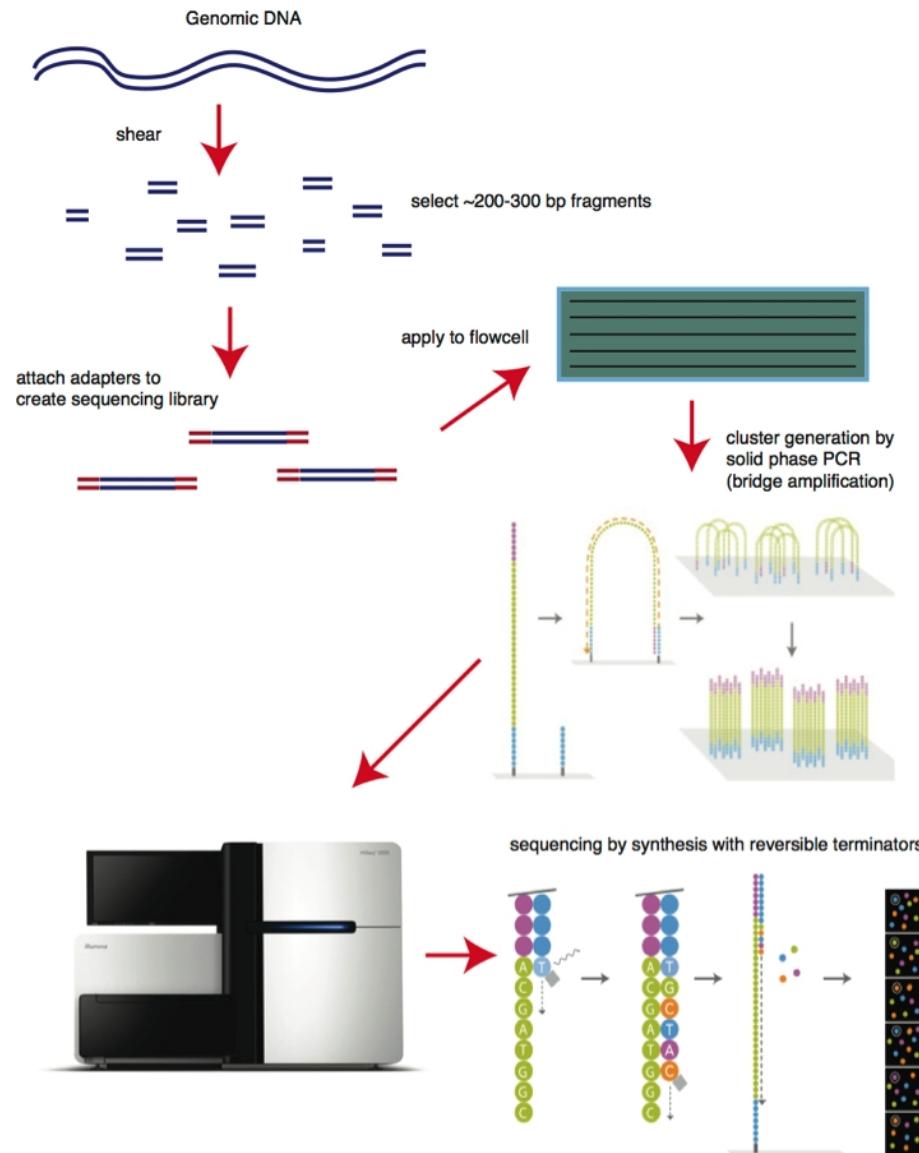
Pro

eradication of optical duplicates
intense signal

Con

really short reads
hard time with repeats

Sequencing platform - Reversible dideoxy-chain termination Illumina



PRO

- NGS sequencer
- Very short (~100 bp, 250 bp)
- Up to 100 B reads

CON

- Short reads
- PCR based
- EXTINCT WHEN ?

Sequencing platform - Reversible dideoxy-chain termination

Illumina hybrid subassembly - Moleculo



RESEARCH ARTICLE



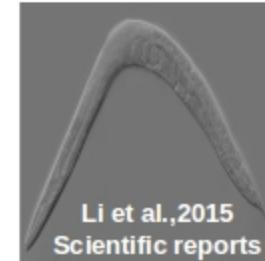
The genome sequence of the colonial chordate, *Botryllus schlosseri*

Ayelet Voskoboinik^{1,2*}, Norma F Neff^{2†}, Debashis Sahoo^{1†}, Aaron M Newman^{1†}, Dmitry Pushkarev^{3†}, Winston Koh^{3†}, Benedetto Passarelli³, H Christina Fan³, Gary L Mantas³, Karla J Palmeri^{1,3}, Katherine J Ishizuka^{1,3}, Carmela Gissi⁴, Francesca Griggio⁴, Rachel Ben-Shlomo⁵, Daniel M Corey⁵, Lolita Penland⁵, Richard A White III⁶, Irving L Weissman^{1,2,6*}, Stephen R Quake^{1,2}

¹Department of Pathology, Institute for Stem Cell Biology and Regenerative Medicine, Stanford University, Stanford, United States; ²Hopkins Marine Station, Stanford University, Pacific Grove, United States; ³Departments of Applied Physics and Bioengineering, Howard Hughes Medical Institute, Stanford University, Stanford, United States; ⁴Dipartimento di Bioscienze, Università degli Studi di Milano, Milano, Italy; ⁵Department of Biology, University of Haifa-Oranim, Tivon, Israel; ⁶Ludwig Center for Cancer Stem Cell Research and Medicine, Stanford University School of Medicine, Stanford, United States



Li et al., 2014
Plos One



Li et al., 2015
Scientific reports

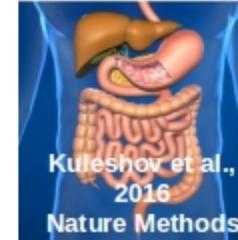


Barrio et al., 2016
eLife

Voskoboinik et al., 2013



Sharon et al., 2015
Genome Research



Kuleshov et al., 2016
Nature Methods

Why not soil?

Sequencing platform - Reversible dideoxy-chain termination

Illumina hybrid subassembly - Moleculo

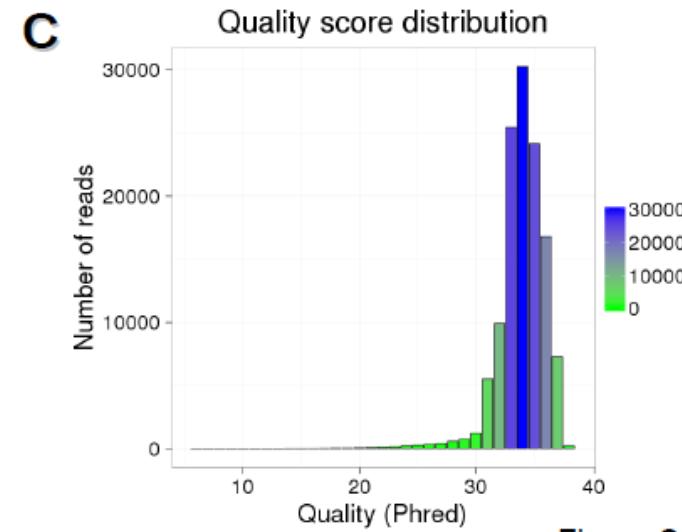
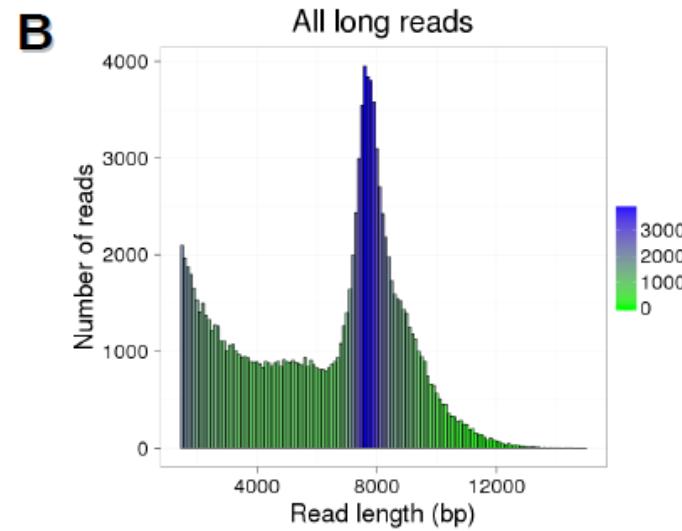
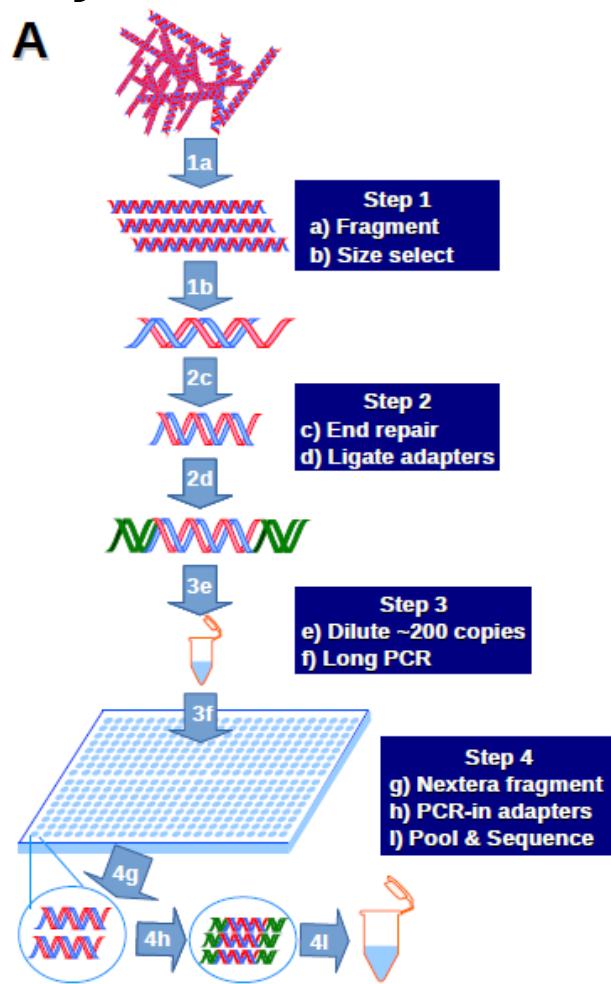


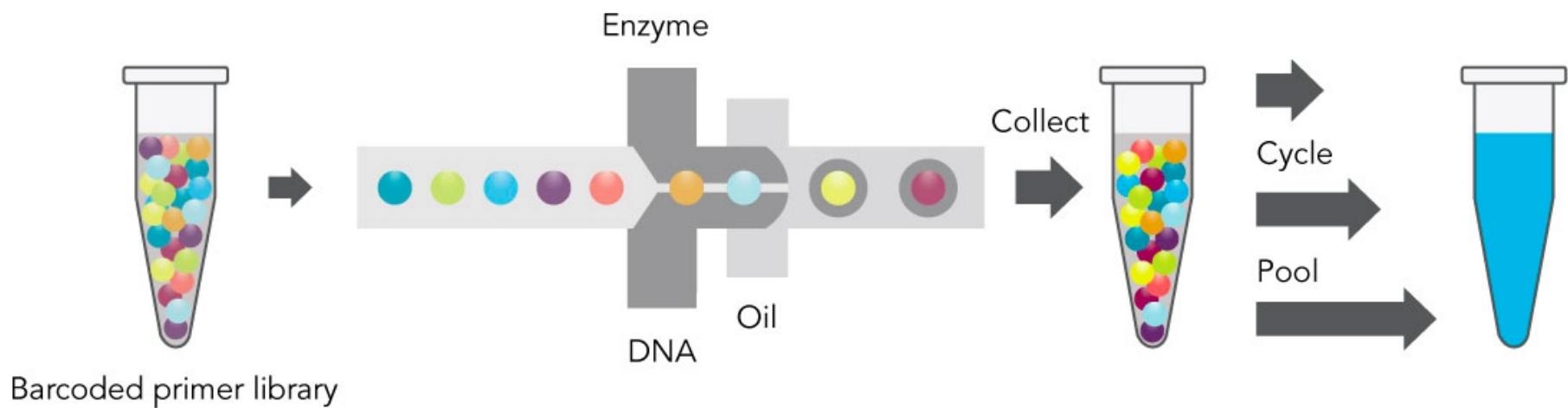
Figure S4

White III et al., 2016. *Msystems*

Link to article: <https://msystems.asm.org/content/1/3/e00045-16>

Sequencing platform - Reversible dideoxy-chain termination

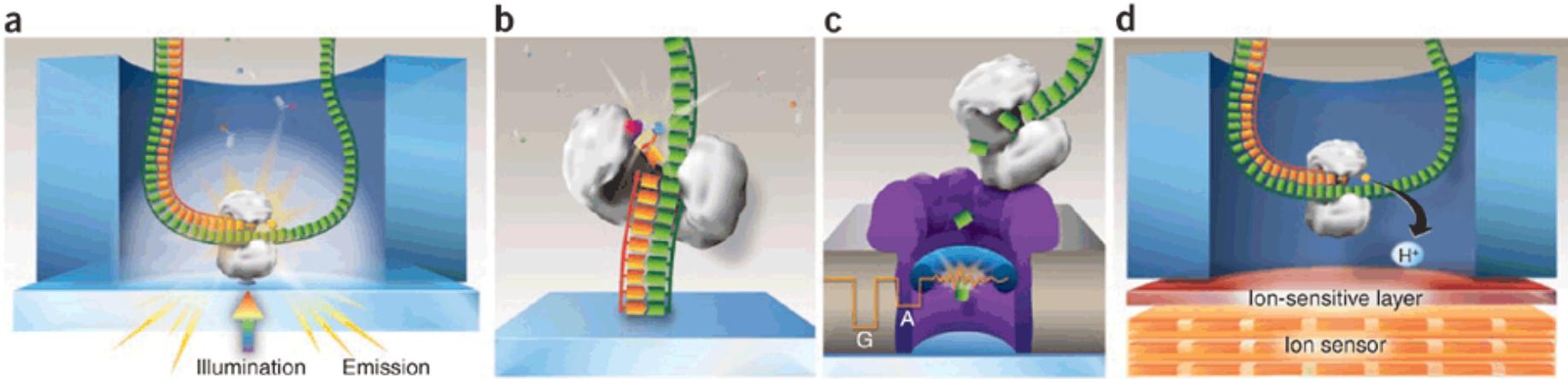
10X technology - Molecule in a droplet



- PRO**
- NGS sequencing library tech (~1 ng)
 - Long >100k partition
 - Phasing genomes

- CON**
- Expensive (\$5 run/\$75k instrument)
 - PCR based

Sequencing platform - Single molecule Pacific Biosciences (PacBio) zero-mode waveguides



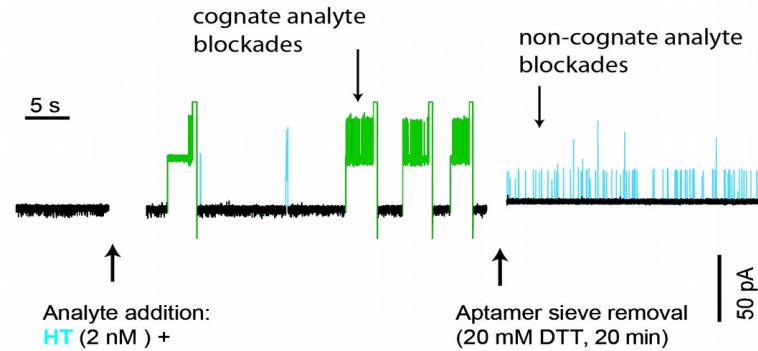
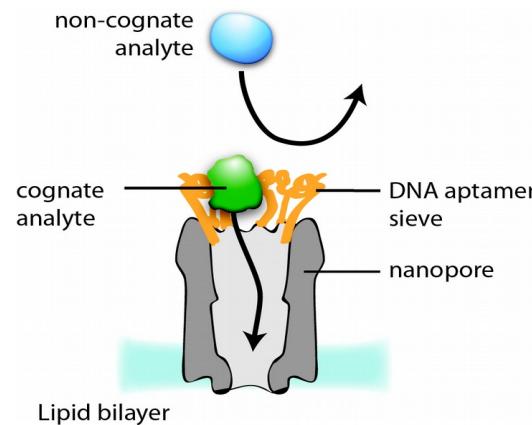
PRO

- Long Reads (>20 kb)
- Improved assembly/scaffold size
- Detects novel and native DNA modifications

CON

- Prior to 2013 long reads (<Q20) very low quality (<Q25)
- High Insertion/Deletion (INDEL) errors
- Higher cost than Illumina

Sequencing platform - Single molecule Oxford nanopore sequencing



PRO

- Long Reads (>20 kb)
- Improved assembly/scaffold size
- Detects novel and native DNA modifications

CON

- Prior to 2013 long reads (<Q20) very low quality (<Q25)
- High Insertion/Deletion (INDEL) errors
- Higher cost than Illumina

Sequencing platform - Single molecule Oxford nanopore sequencing



MinION
10 to 20 Gb ~6 h
Weights 100 g
PC/Laptop portable
~\$1k



GridION
100 GB per run
5 minIONs
\$50k



PromethION
48 MinIONs
Each cell 3,000 nanopores
Total 144,000 nanopores
\$150k



SmidgION

Sequencing cost by platform

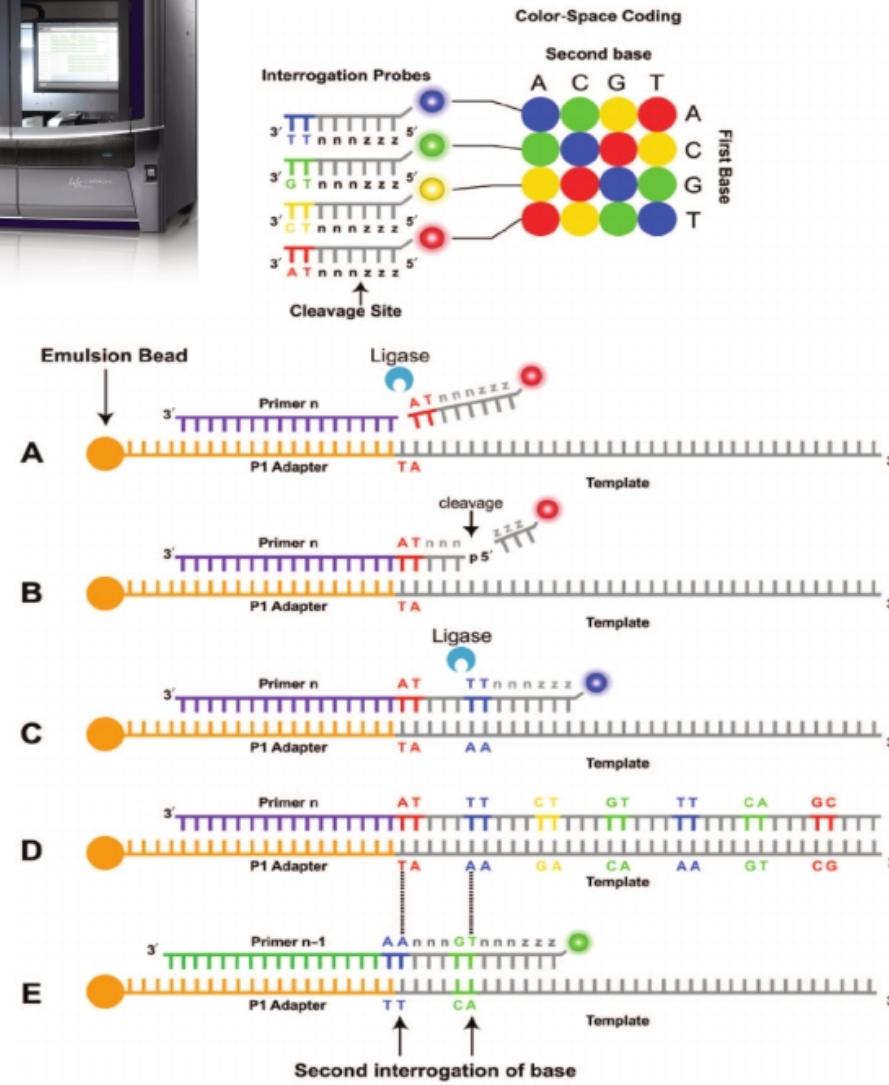
	Date	Time	Read length	Quality	Amount	MB cost
454 FLX+	2005 (2011)	24 h	700 bp	> Q20	1 Gb	\$10.00
Ion torrent	2011 (2014)	2 h	400 bp	> Q20	1 Gb	\$1.00
Illumina MiSeq	2011 (2014)	27 h	2x300 bp	> Q30	15 Gb	\$0.15
Illumina HiSeq 2500	2010 (2014)	1 - 10 days	2x250 bp	> Q30	3000 GB	\$0.05
Illumina Moleculo	2012 (2014)	2 days	5 - 15 kb	> Q30	2 Gb	\$0.35
Illumina Xten	2014	14 days	2x100 bp	> Q30	1.8 Tb	\$0.001
PacBio	2011 (2013)	30 m - 4 h	10 - 40 kb	> Q20	500 Mb	\$0.13

Extinct Sequencing platforms



Sequencing platform - Sequencing by ligation

Life Tech SOiLD platform



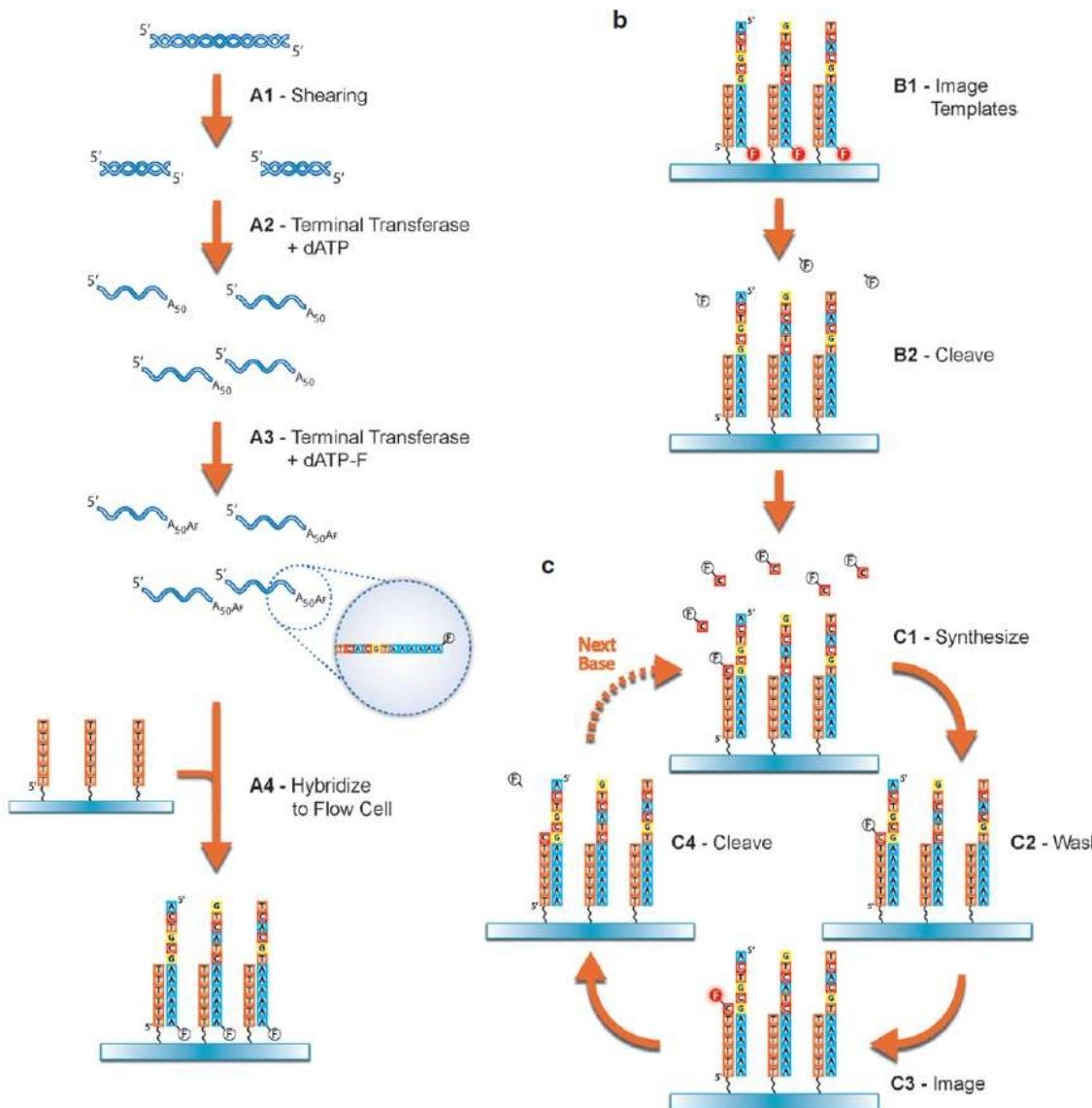
PRO

- NGS sequencer
- Very short (<100 bp)
- Up to 10 B reads

CON

- Short reads
- emPCR based
- NOW EXTINCT 2017

Sequencing platform - single molecule fluorescent sequencing Helicos Biosciences heliscope



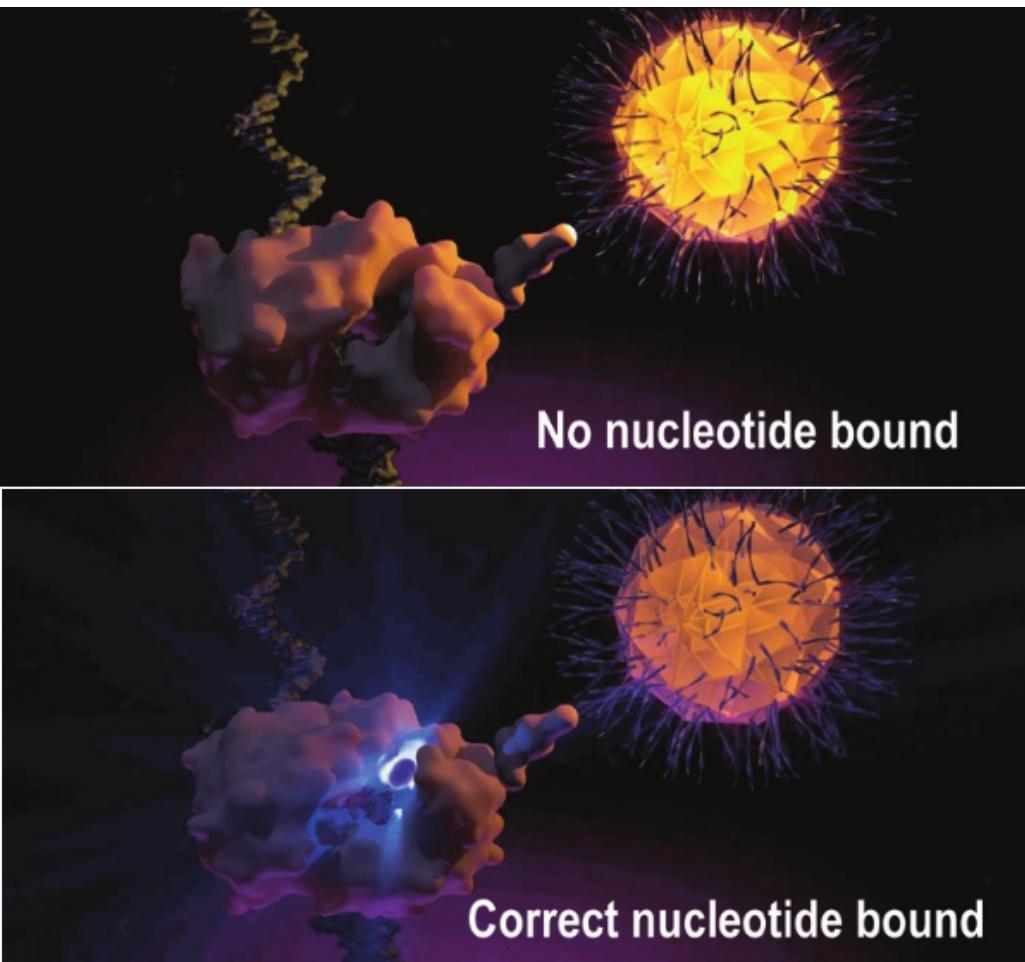
PRO

- NGS sequencer
- Very short (<50 bp)
- Up to 1 B reads
- Direct RNA (polyA oligos)

CON

- Short reads
- expensive
- NOW EXTINCT 2012

Sequencing platform - single molecule FRET sequencing Life Tech “Project Starlight”

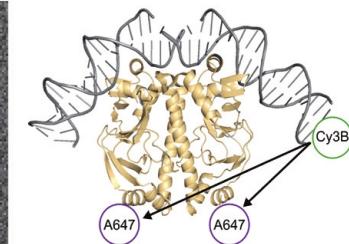
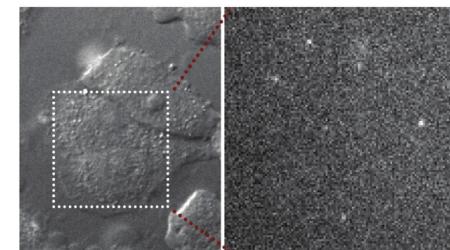


PRO

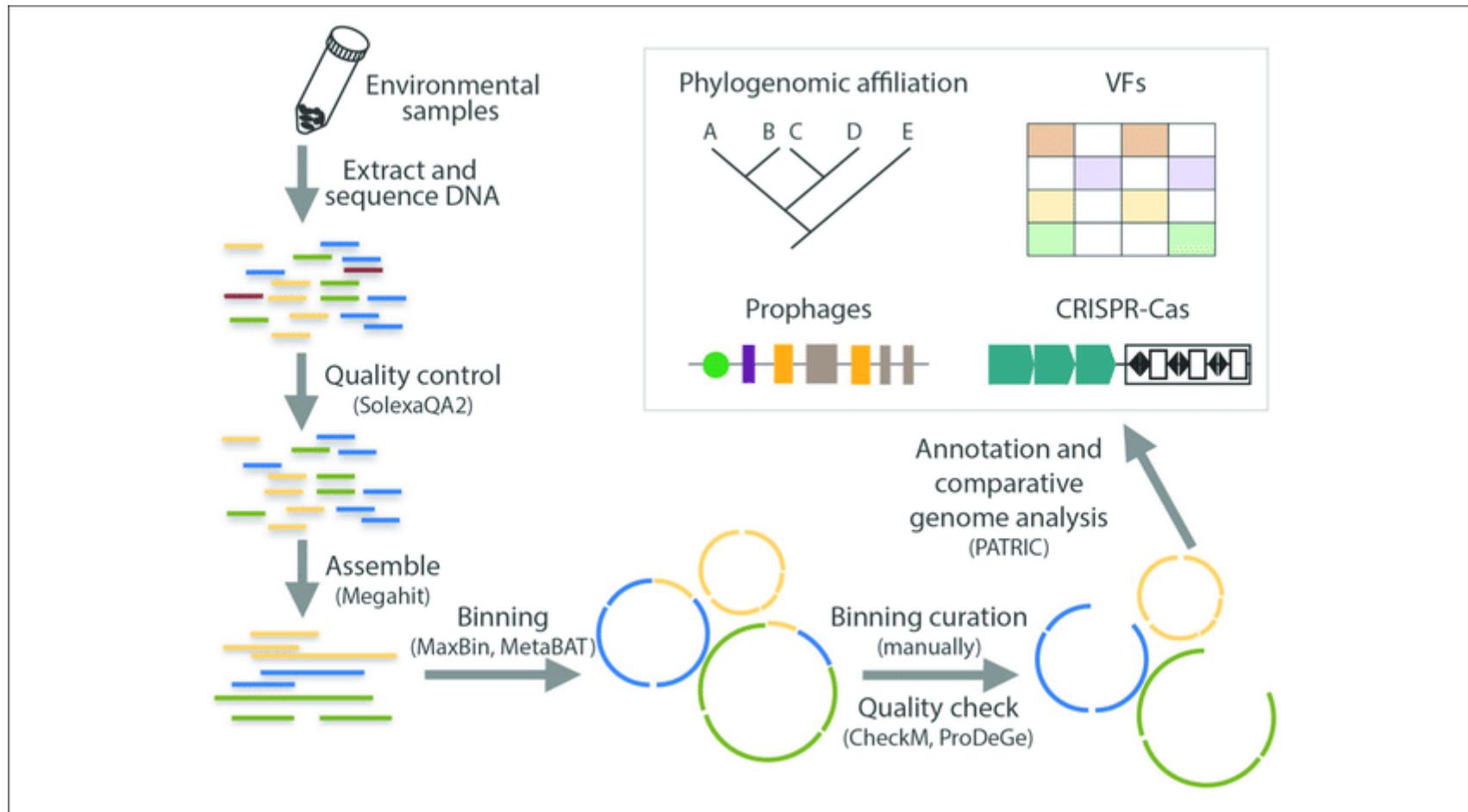
- NGS sequencer
- Uses FRET
- Truly single molecule

CON

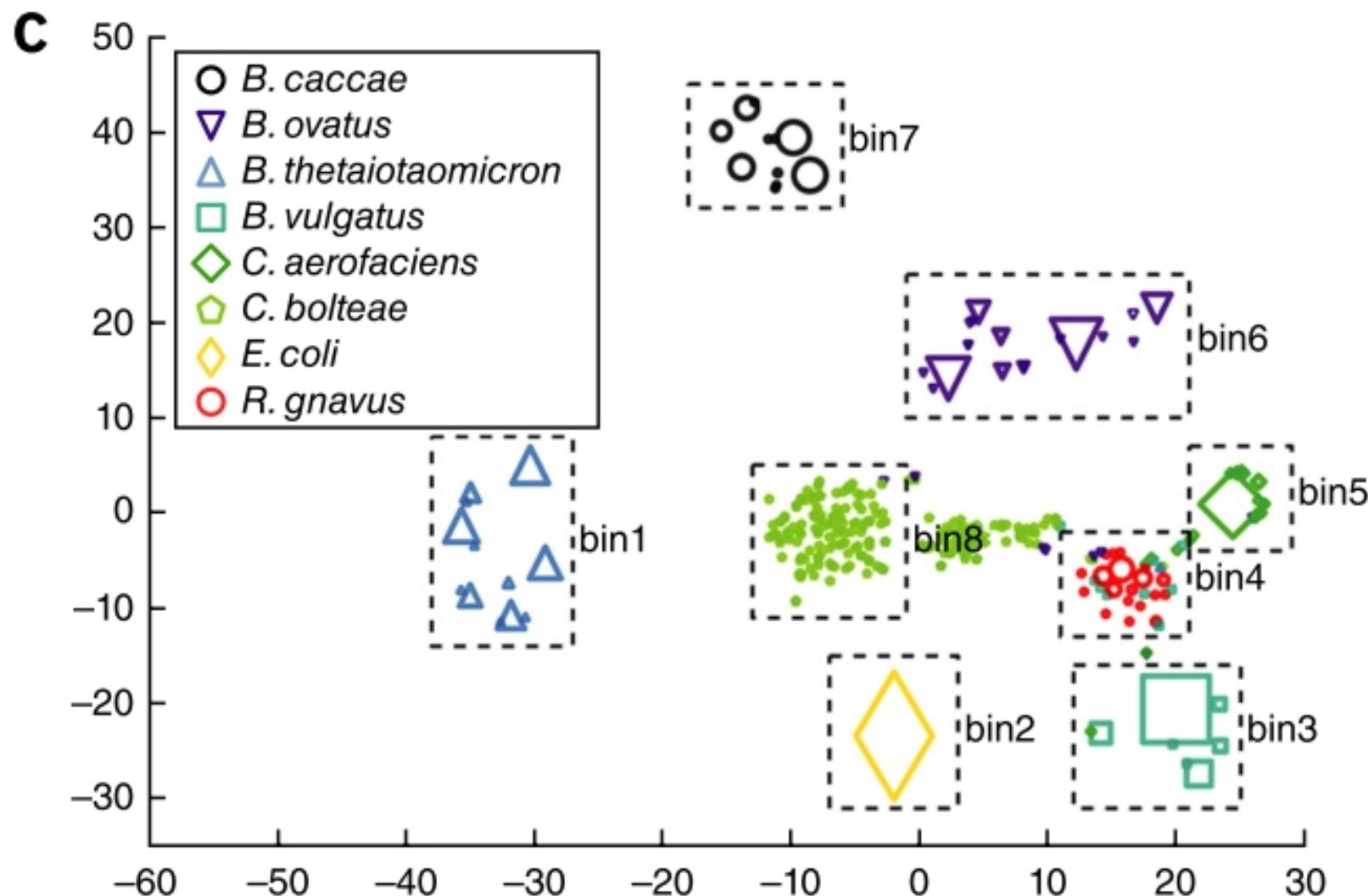
- Can it sequence reads?
- NOW EXTINCT 2010



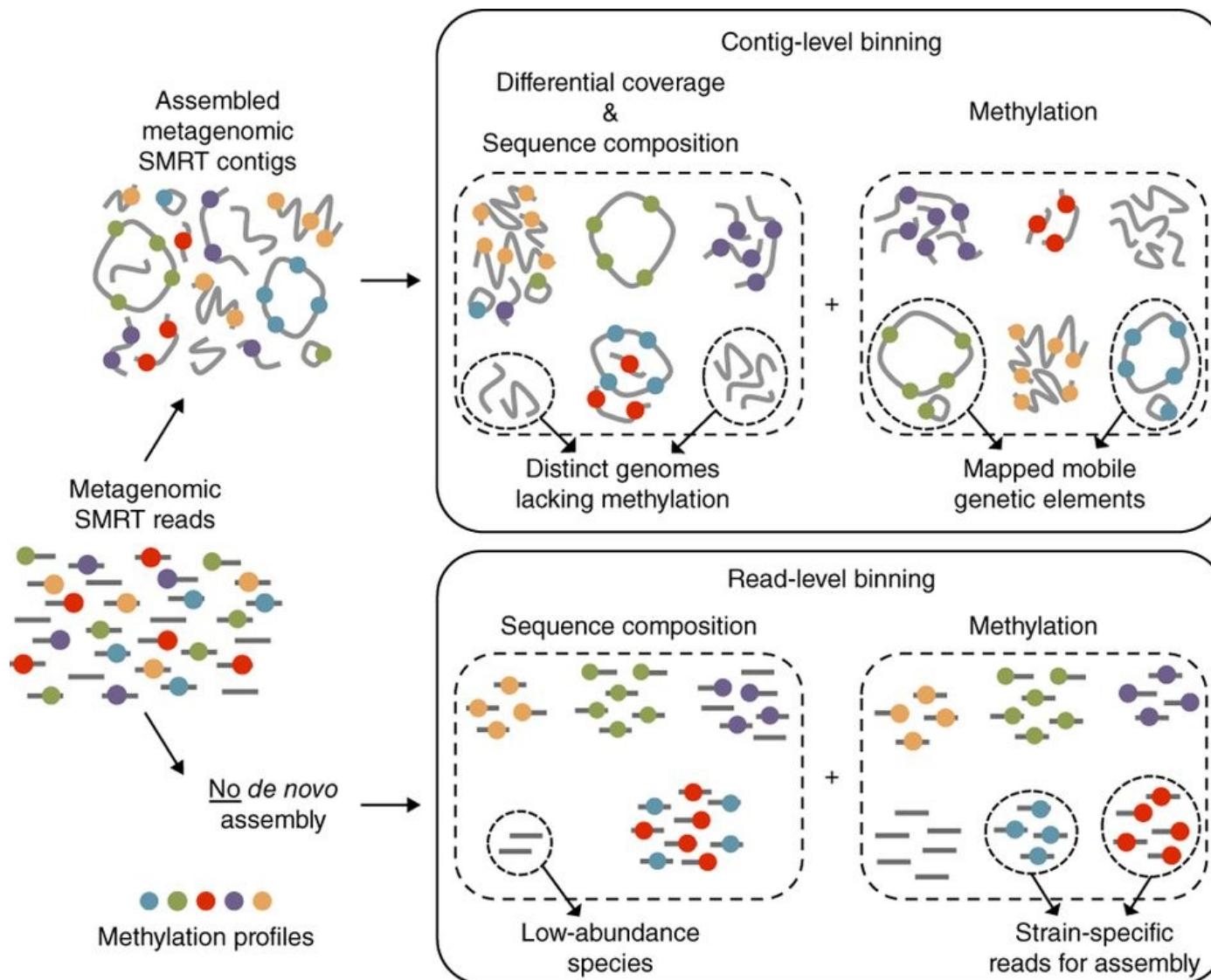
Metagenomic binning (MAGs)



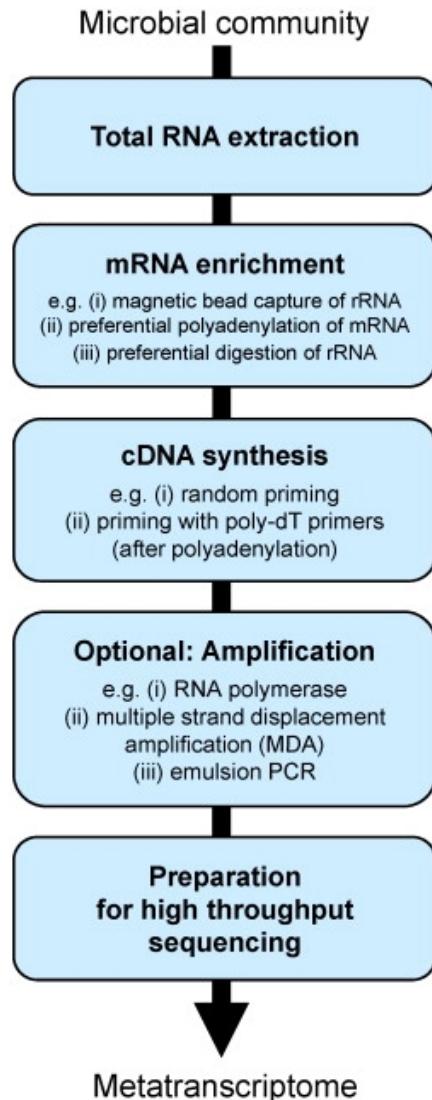
Metagenomic binning (MAGs)



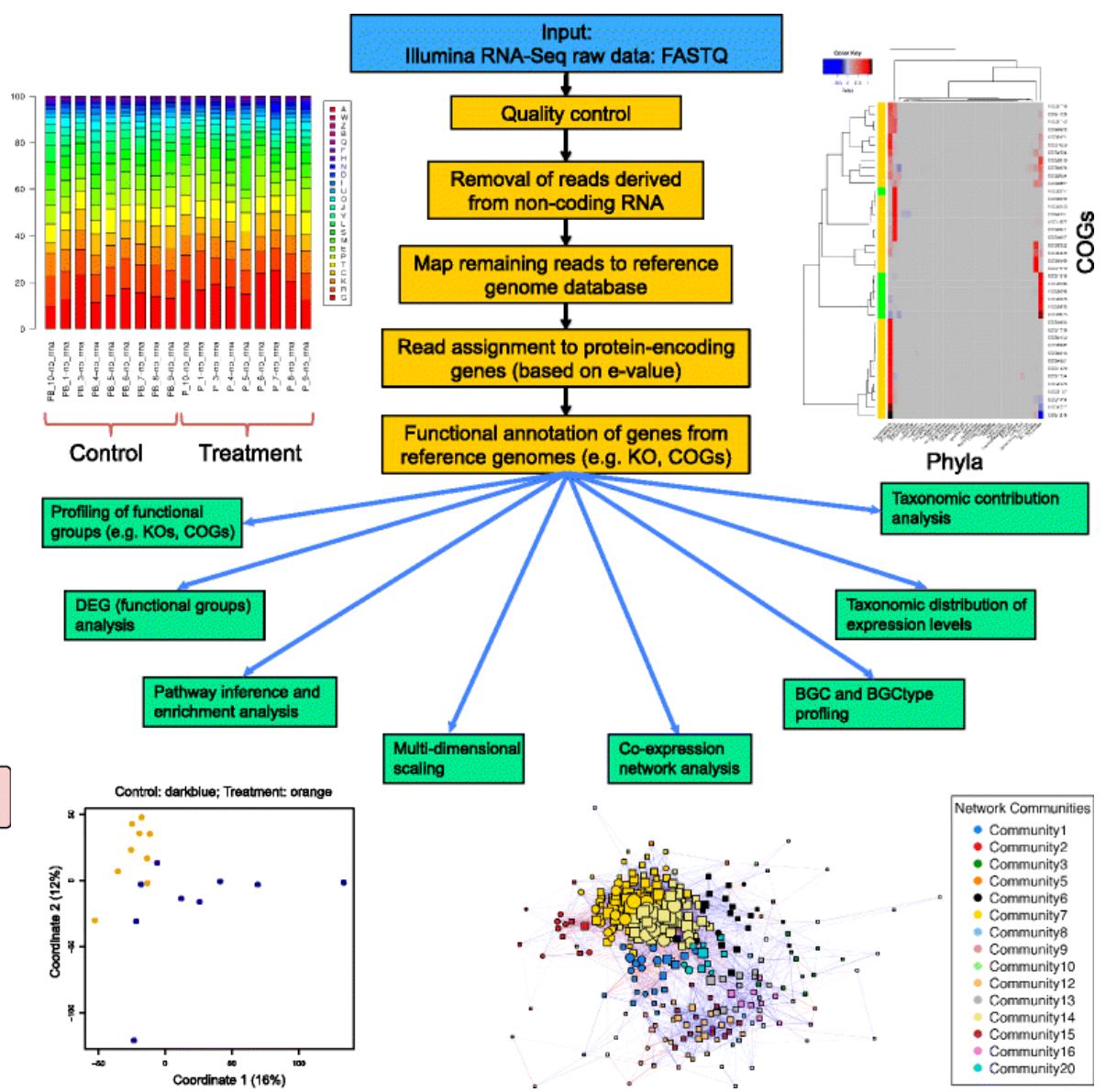
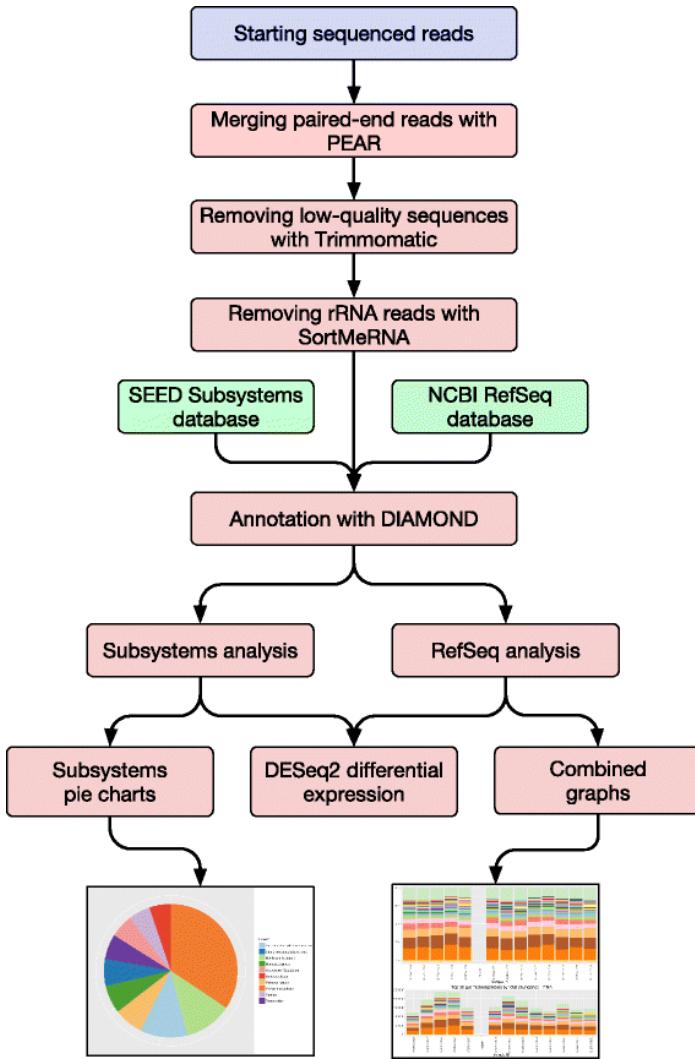
Metagenomic binning (MAGs) – methylation use



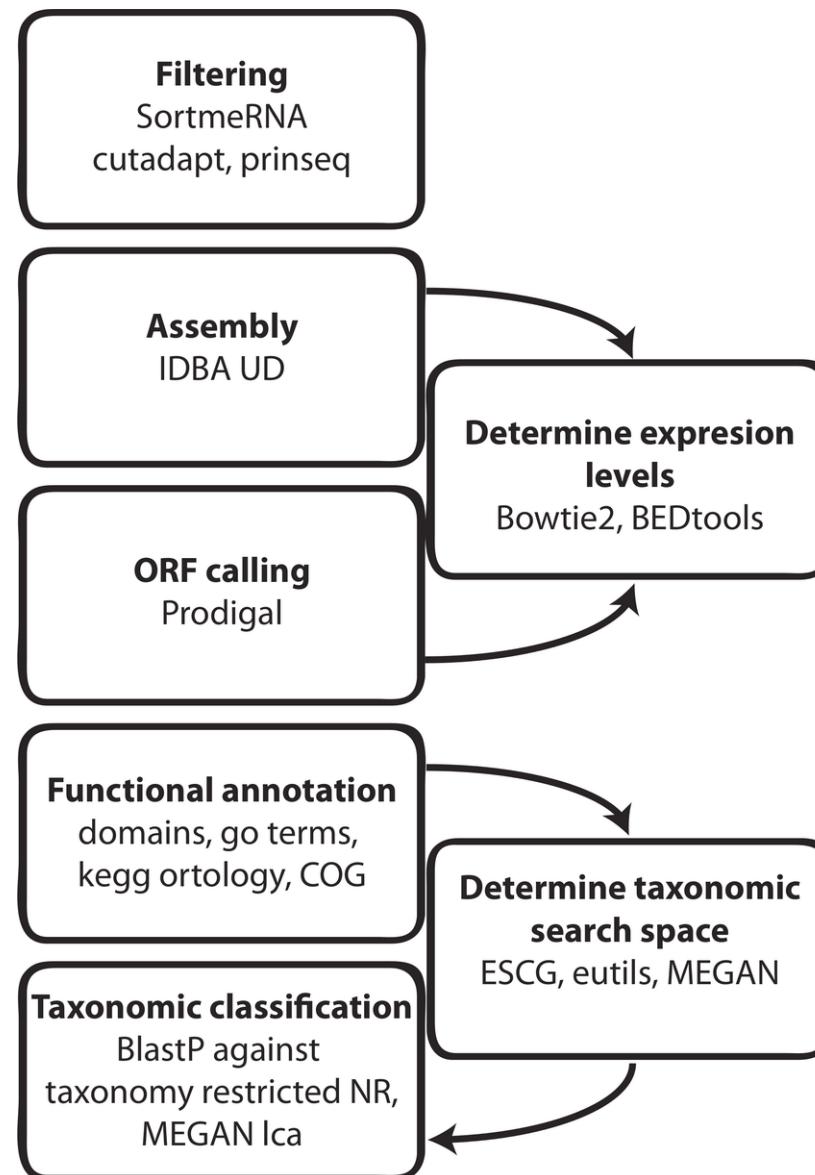
Metatranscriptomics – sample prep



Metatranscriptomics analysis - Read based



Metatranscriptomics analysis – assembly based



Metatranscriptomics analysis – hybrid based

