Emerging NGS Platforms

Ultima Genomics

The UG100

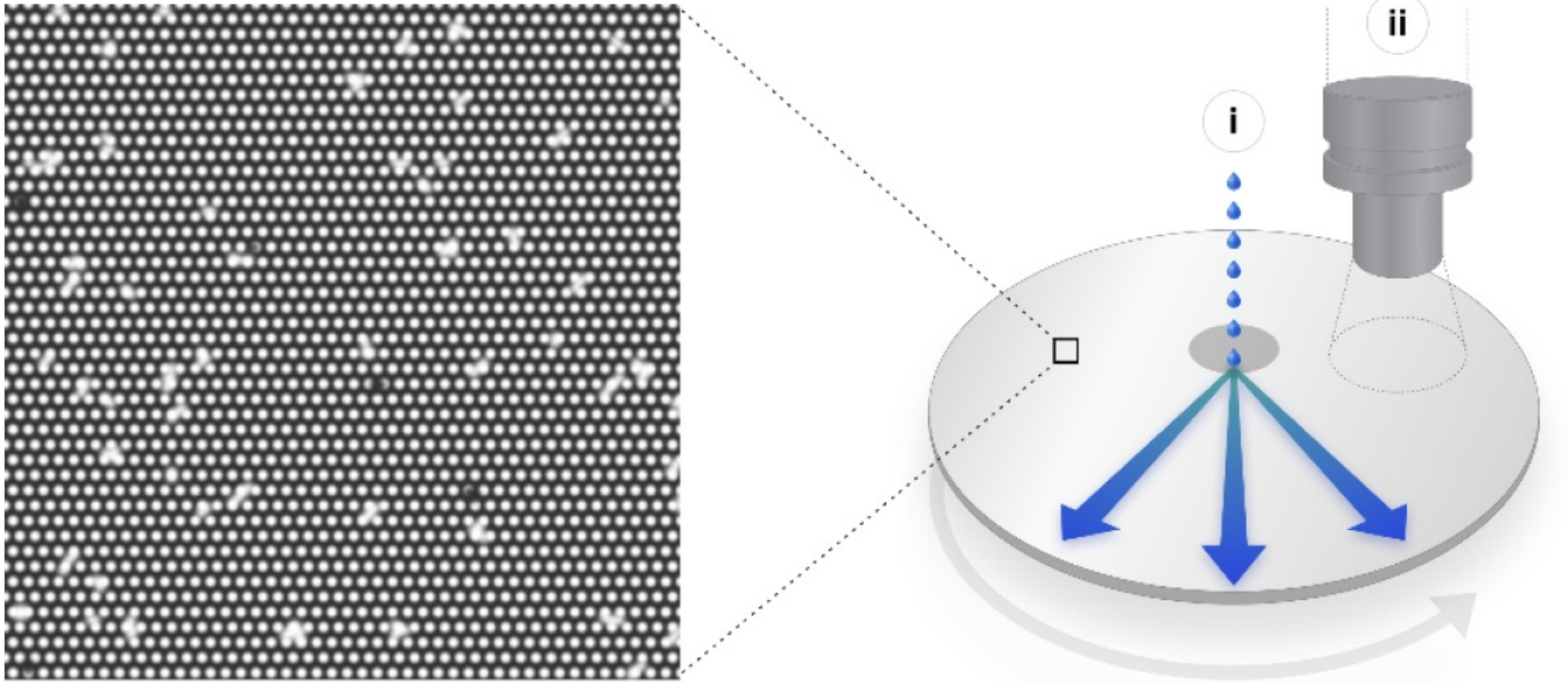
Footprint is the “Ultima”te of its kind

- Three instruments total
  - Clustering
  - Sequencing
  - Computational

Ultima Genomics

Three main innovative components: (1) *open fluidics and optics system*

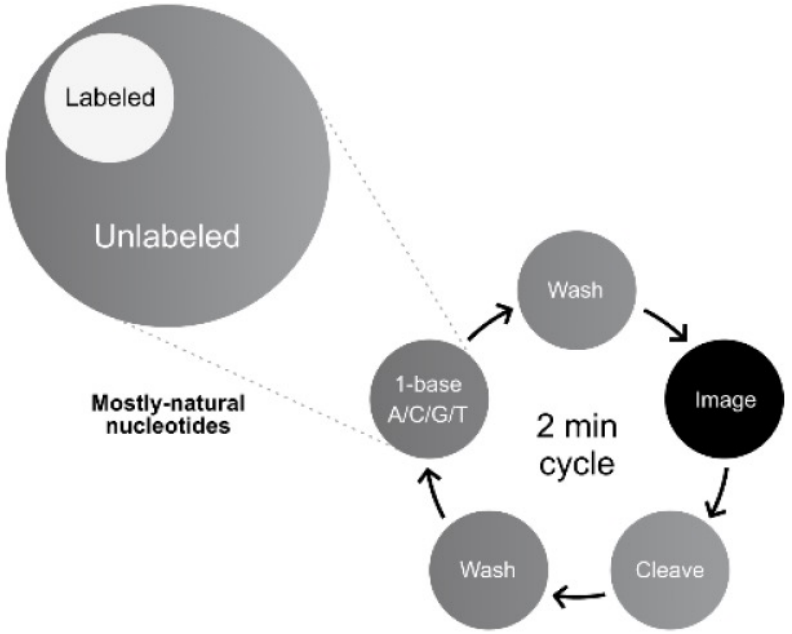
- Circular silicon wafer as an “open flow-cell”
- Patterned – dense array of electrostatic landing pads to bind sequencing beads
  - ~10 billion clonally amplified beads
- Spin-dispense system

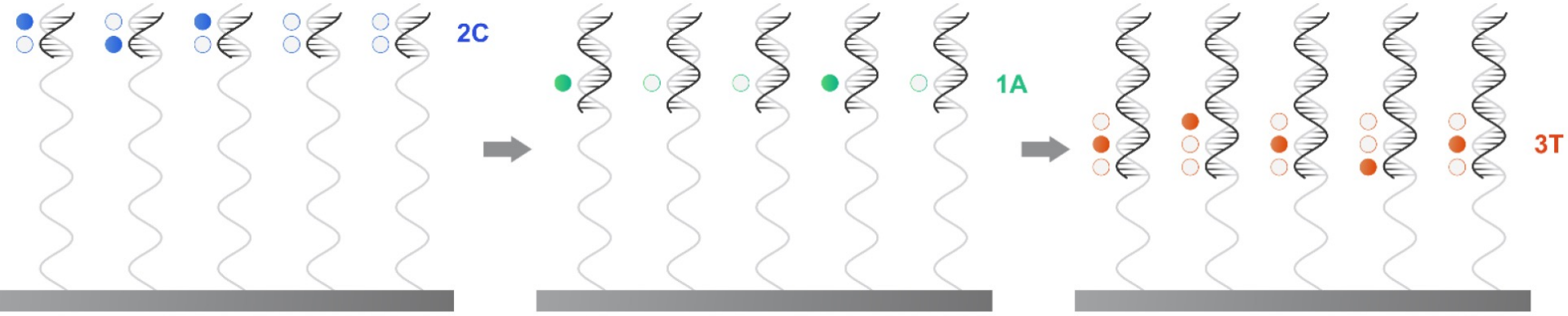


Ultima Genomics

Three main innovative components: (2) *mostly natural sequencing chemistry*

- Sequencing-by-synthesis (mnSBS) uses a mixture of native dNTPs and one-at-a-time fluorescently labeled dNTP
- Polymerase extends 0, 1 or several bases depending on respective homopolymer
- Detected signal proportional to length of homopolymer





Ultima Genomics

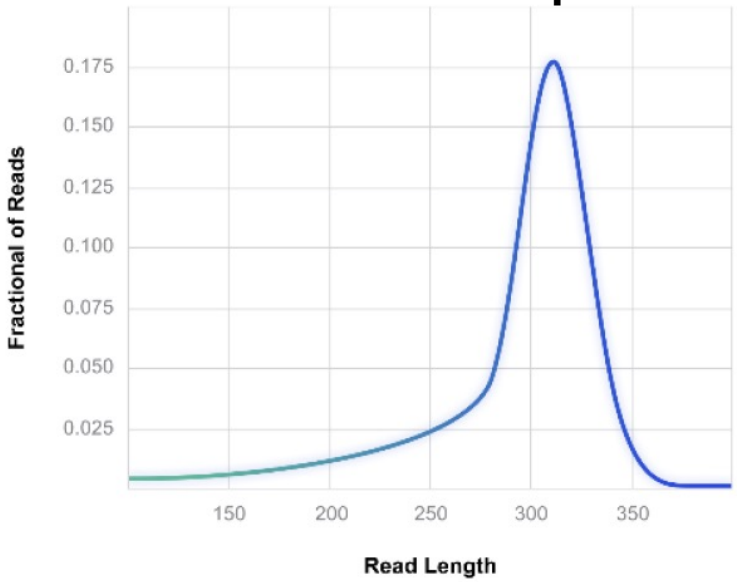
Three main innovative components: (3) *neural network-enabled base-calling*

- Machine learning and convolutional neural network (CNN) to convert raw signals to sequence reads; run-specific calibration process; CRAM file

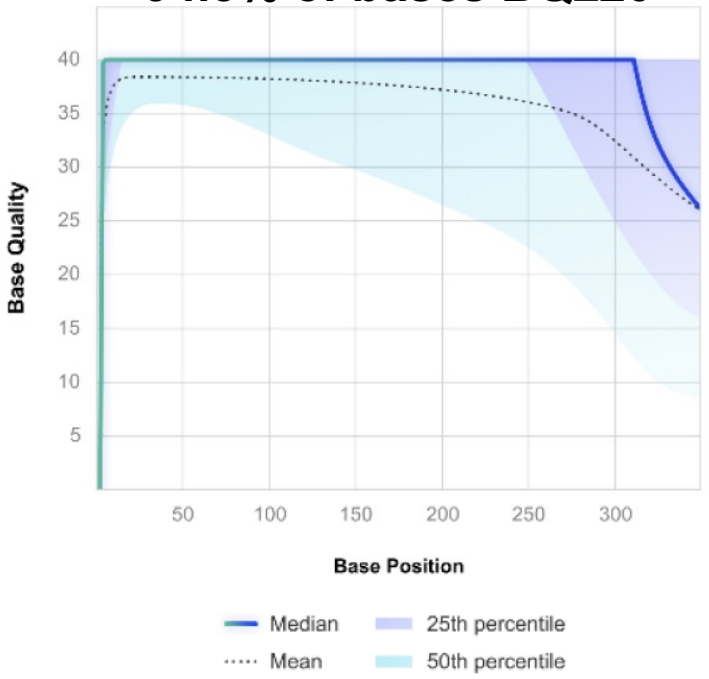
HG001-HG007 Generated Data

- 444 cycles and a 20 hr run time

Mode 310 bp

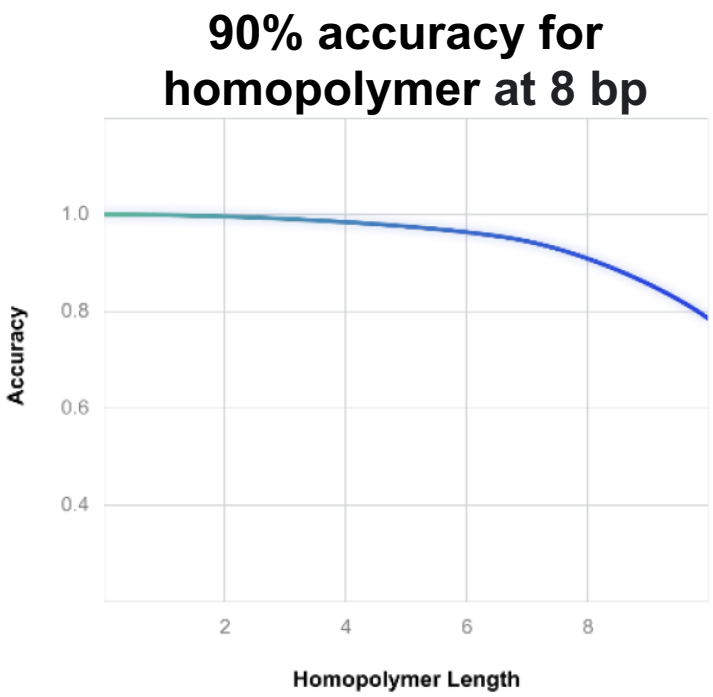


94.5% of bases BQ≥20



Ultima Genomics

Three main innovative components: (3) *neural network-enabled base-calling*



	HG001	HG002	HG003	HG004	HG005	HG006	HG007	GIAB Mean
Mean coverage	40.05	39.58	39.67	38.56	39.16	39.53	38.32	39.27
% >20X	95.9%	95.7%	95.9%	96.6%	96.8%	96.1%	96.7%	96.2%
% Duplication	4.1%	6.6%	6.2%	7.3%	7.6%	6.1%	8.2%	6.6%
F90*	1.483	1.466	1.469	1.377	1.398	1.464	1.368	1.432
F95*	1.821	1.885	1.803	1.677	1.632	1.797	1.666	1.754
PF reads (M)**	469	435	436	4338	430	443	433	440
% reads aligned	99.80%	99.96%	99.96%	99.97%	99.96%	99.95%	99.97%	99.94%
Mean read length	264.8	284.9	284.4	286.7	288.6	280.3	286.4	282.3
Median read length	291	302	302	303	304	301	302	300.7
Modal read length	309	311	311	311	311	310	311	310.6
% chimeras	2.3%	1.1%	1.2%	1.3%	1.4%	1.5%	1.3%	1.5%
Raw Indel error	0.27%	0.30%	0.29%	0.28%	0.29%	0.38%	0.28%	0.30%
HQ Mismatch error†	0.07%	0.08%	0.07%	0.07%	0.07%	0.10%	0.07%	0.07%
% BQ20 bases	95.5%	94.8%	95.0%	95.2%	95.0%	93.6%	95.2%	94.9%
% BQ30 bases	87.3%	86.4%	86.8%	87.6%	87.1%	84.6%	87.7%	86.8%
Ti/Tv ratio Exome	2.97	2.89	2.95	2.90	2.98	2.93	2.97	2.94
Ti/Tv ratio‡	2.09	2.09	2.09	2.09	2.09	2.09	2.09	2.09
SNP recall‡	99.7%	99.6%	99.6%	99.7%	99.7%	99.6%	99.7%	99.7%
SNP precision‡	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%	99.6%
SNP F1‡	99.7%	99.6%	99.6%	99.7%	99.7%	99.6%	99.7%	99.6%
Indel recall‡	96.7%	96.4%	96.6%	95.4%	96.0%	95.9%	96.0%	96.1%
Indel precision‡	97.0%	96.8%	97.1%	96.4%	97.0%	96.2%	96.7%	96.7%
Indel F1‡	96.9%	96.6%	96.8%	95.9%	96.5%	96.1%	96.3%	96.4%

Table 1: Performance metrics for Genome in a Bottle (GIAB) reference genomes HG001-7, and average performance metrics for 224 additional 1000 Genomes (1000G) reference genomes.

* F90/95: Ratio of coverage between the median and the 10th or 5th percentile lowest coverage, respectively.
** PF: Pass-filter reads. All other metrics were calculated over these reads.
† HQ Mismatch error rate was corrected for germline SNPs and alignment errors (see Methods section).
‡ Variant calling metrics were measured on GIAB HCR excluding long homopolymers and repetitive regions (UG-HCR, see Methods).

GIAB Mean
SNP: 99.7% recall; 99.6% precision
Indel: 96.1% recall; 96.7% precision

Ultima Genomics in Summary

- Technology: mostly natural sequencing-by-synthesis (mnSBS)
- Output: 10 billion clonally amplified beads; 2 Wafers at a time
- Runtime: 20 hr
- Length of reads: 310 bp mode size (GIAB)
- Accuracy:
  - SNP: 99.7% recall; 99.6% precision
  - Indel: 96.1% recall; 96.7% precision
- Instrument cost: Unknown (comparable to Illumina NovaSeq 6000)
- Instrument size (relative): Several NovaSeqs worth
- Partnerships/acquisitions:
  - AI partnerships with Google DeepVariant, NVIDIA and Senteon
  - Exact Sciences, Regeneron Pharmaceuticals, NYGC and Broad Institute: beta testing

Element Biosciences – AVITI Benchtop Sequencer

Highlights

- **Two** independent, random-access flow cells (~240Gb/800 million paired-end reads output per flow cell)
- **FASTQ** file output (conversion performed by AVITI Operating Software Bases2Fastq workflow)
- Two library workflows (**Adept** – conversion of existing libraries ; **Elevate** – library preparation)
- **Tunable optical throughput** (full scan, ½ scan, ¼ scan etc..) to select desired TAT/read depth

Element Biosciences - AVITI

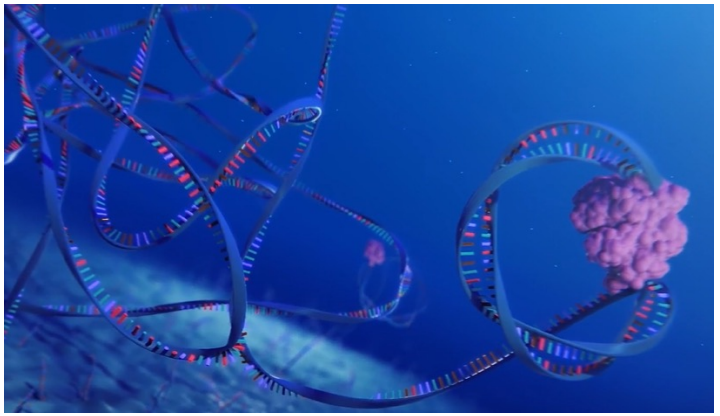
Circularization

Off-instrument library circularized prior to sequencing

"*Working to move onto the flow cell*"

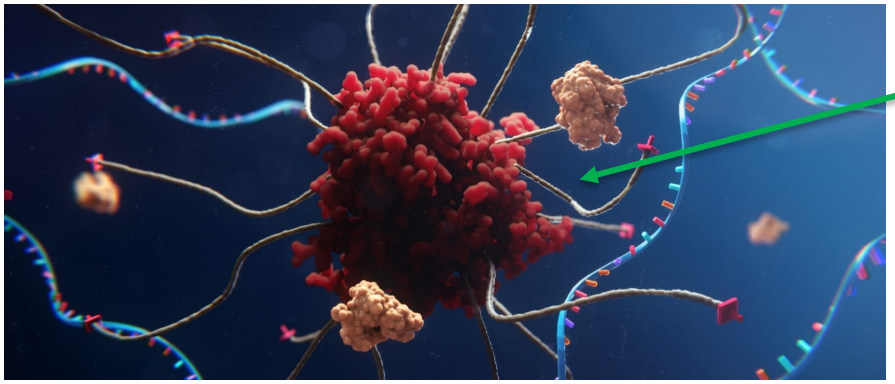
Pollination

Rolling circular amplification on flow cell surface creates clonal copies of the library molecule or *"polony"*



Sequencing

Instead of binding an individual labeled nucleotide at each location of the cluster, AVITI uses single fluor with many octopus-like tentacle arms (Avidite)



Avidite nucleotide substrate

Each arm, single nt type

One avidite = multiple binding sites within the polony = increased binding avidity

Element Biosciences - Resources

Element Biosciences - AVITI

Sequencing metrics at 800M Reads PF/flow cell; 1600M Reads PF combined when running both flow cells.

	Read Length	Output (Gb) 1 Flow Cell	Output (Gb) 2 Flow Cells	Run Time	Data Quality
300-cycle Sequencing Kit	2x150	240	480	48hrs	%Q30 > 90
	2x100	160	320	35hrs	
150-cycle Sequencing Kit	2x75	120	240	29hrs	
	2x50	80	160	23hrs	
	2x25	40	80	17hrs	

Tunable run time

Sequencing turnaround times tunable to shallower read depths for time saving options; common paired end read length examples below.

Read Throughput[1]	2x25	2x50	2x75	2x100	2x150
800M (full scan)	17hrs	23hrs	29hrs	35hrs	48hrs
600M (3/4 scan)	16hrs	21hrs	27hrs	32hrs	43hrs
400M (1/2 scan)	15hrs	20hrs	24hrs	29hrs	39hrs
200M (1/4 scan)	14hrs	18hrs	22hrs	26hrs	34hrs
100M (1/8 scan)	13hrs	17hrs	21hrs	24hrs	31hrs

[1] Paired-end reads

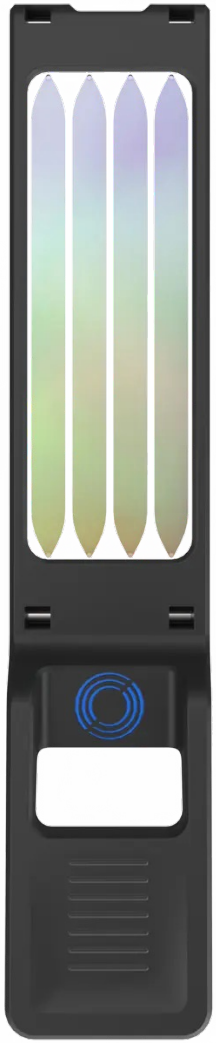
Element Bioscience (AVITI) in Summary

- Technology: **Sequencing-by-synthesis**
- Output:
  - **2 x 150 240Gb / 800M read pairs per flow cell**
  - **2 x 150 260-300Gb / >800M read pairs per flow cell (Broad Institute beta test)**
  - **2 x 75 120Gb / 800M read pairs per flow cell**
  - **2 x 75 130-150Gb / >800M read pairs per flow cell (Broad Institute beta test)**
- Runtime: **2 x 150bp (full-scan) - 48 hours**
- Accuracy:
  - **SNP: 99.8% recall / 99.2% precision**
  - **Indel: 99.7% recall / 99.2% precision**
  - **>90% Q30 (2 x 150, PCR-free libraries)**
  - **>97% Q30 (Broad Institute beta test, 2 x 150, PCR-free libraries, GIAB)**
  - **>87% Q40 (Broad Institute beta test, 2 x 150, PCR-free libraries, GIAB)**
- Instrument cost: **$289K**
  - **300-cycle sequencing cartridge: $1,680**
  - **150-cycle sequencing cartridge: $1,080 (Q4 2022)**
  - **Price per Gb: $5 - $7 USD**
- Instrument size (relative): **Large benchtop (29" x 36" x 29")**
- Partnerships: Fabric Genomics/Senteon/Jumpcode/Dovetail/NEB/QIAGEN/Agilent

Singular Genomics G4 Platform

G4 Sequencing Platform
Bench Top

Singular Genomics



Sequencing

- Rapid SBS chemistry – 4 color
- On board cluster generation
- Engineered polymerase

Flow Cell Design

- 4 lanes per flow cell
  - Lane segregation
- 4 flow cells per sequencing run

		F2 Kits	F3 Kits*
	NUMBER OF READS (CLUSTERS)	**150–165 M**	**300–330 M**
Sequencing Output (Base calls)	1 x 50 bp (50 cycles)		15–17 Gb
	2 x 50 bp (100 cycles)	15–17 Gb	30–33 Gb
	2 x 100 bp (200 cycles)	30–33 Gb	60–66 Gb
	2 x 150 bp (300 cycles)	45–50 Gb	90–100 Gb
Run Time	1 x 50 bp (50 cycles)		6-8 hrs
	2 x 50 bp (100 cycles)	8-10 hrs	8-10 hrs
	2 x 100 bp (200 cycles)	12–15 hrs	12–15 hrs
	2 x 150 bp (300 cycles)	16–19 hrs	16–19 hrs

*F3 planned for next release

Singular Genomics Performance

METRIC	Flow Cell 1	Flow Cell 2	Flow Cell 3	Flow Cell 4
Configuration	2x150	2x150	2x150	2x150
Paired-Reads (M)	168	169	169	186
Output (Gb)	51	51	51	56
% Bases ≥ Q30 R1	90	90	89	88
% Bases ≥ Q30 R2	91	91	90	90

Two Separate Reactions of NA12878 WGS

METRIC	20x Target Coverage	30x Target Coverage	40x Target Coverage
%PF Reads Aligned	99.99	99.16	98.37
Duplication Rate (%)	17.7	16.3	15.5
Median Insert Size (bp)	251	249	248
Mean Coverage (X)	22.2	33.6	45.9
%Bases >=10x Coverage	98.49	99.44	99.69
SNP Precision	99.30	99.62	99.71
SNP Sensitivity	98.94	99.14	99.17
SNP F1-Score	99.12	99.38	99.44
Indel (<50bp) Precision	95.02	96.49	97.13
Indel (<50bp) Sensitivity	93.51	95.49	96.4
Indel F1-Score	94.26	95.99	96.77
Total SNPs	3738914	3741923	3744535
Het:Hom Ratio	1.49	1.48	1.46
Ti:Tv Ratio	2	1.99	1.99

40x WGS Coverage

SNP:

99.17% Sensitivity

99.71% Precision

Indel:

96.4% Sensitivity

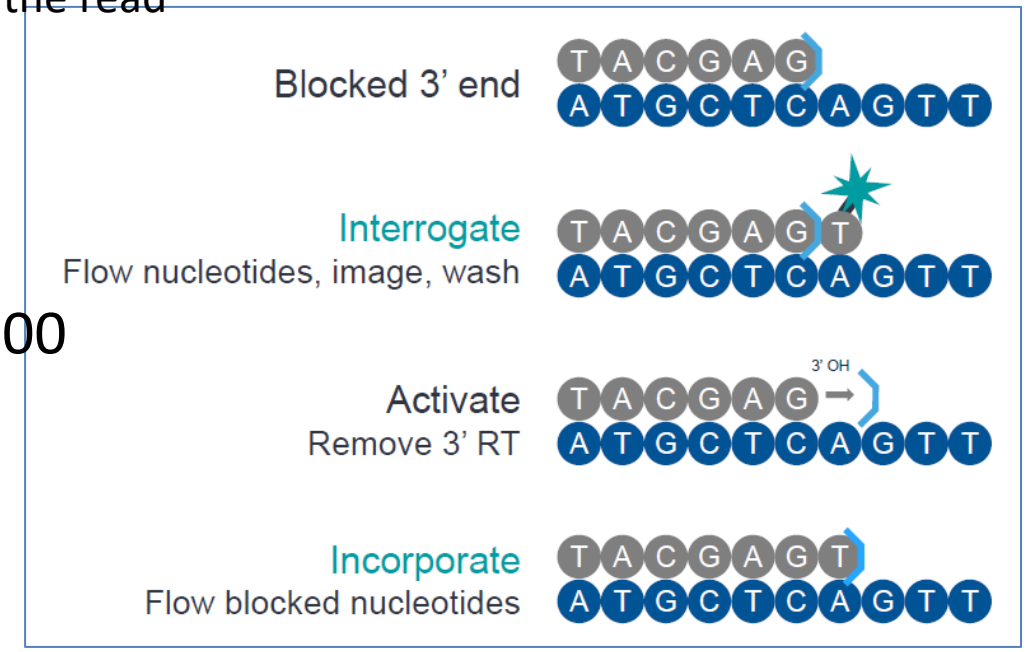
97.13% Precision

Singular Genomics: Summary

- Technology: Rapid SBS – 4 Color Chemistry
- Output: 150 – 165 million clusters (F2 flow cell)/four flow cells per instrument
- Runtime (2 x 150 bp): 16 – 19 hr
- Length of reads: 2 x 150 bp
- Accuracy:
  - SNP: 99.17% Sensitivity; 99.71% Precision
  - Indel: 96.4% Sensitivity; 97.13% Precision
- Instrument cost: $350,000
- Instrument size (relative): Bench Top; Slightly bigger than a MiSeq
- Partnerships/acquisitions:
  - Dovetail Genomics, Lexogen, NEB, Twist Bioscience, Watchmaker Genomics, Broad Institute's Terra Platform

PacBio: Onso

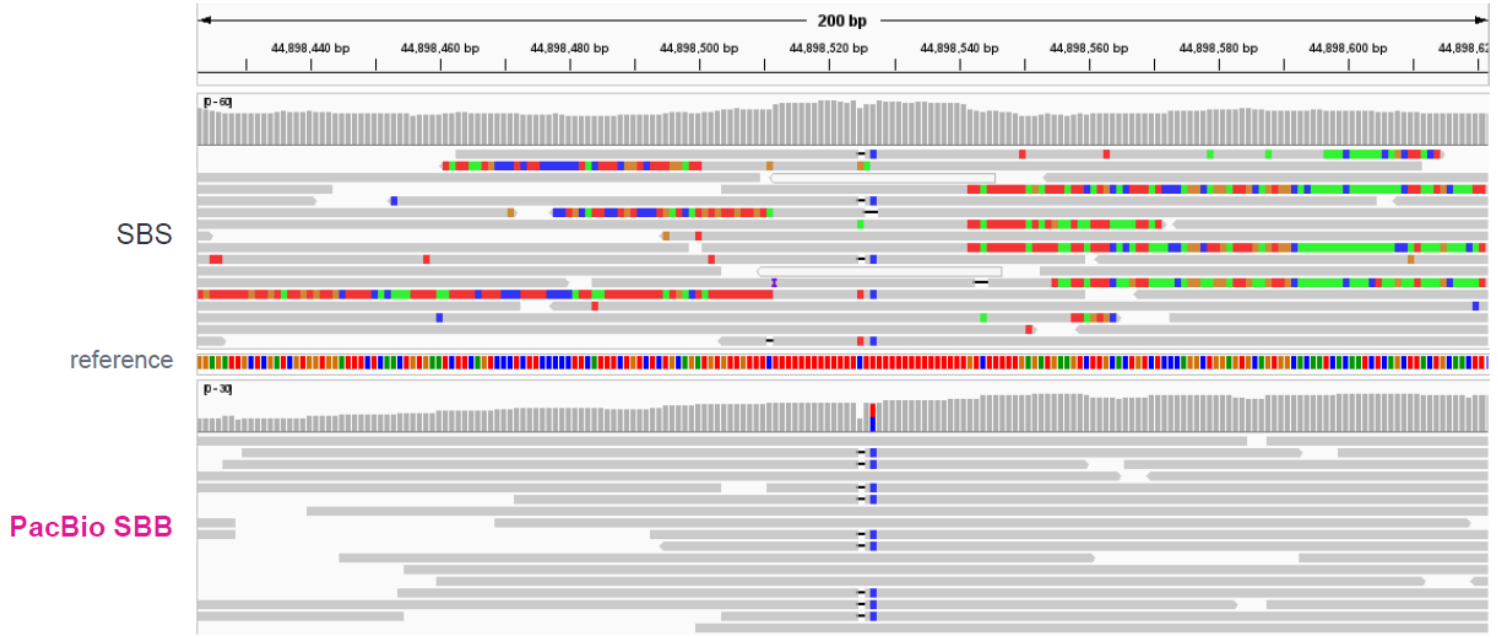
- Sequencing By Binding (SBB) Technology: interrogation followed by incorporation
  - Better for homopolymer stretches
  - Chemistry doesn't die at the end of the read
- Output: Up to 500M reads
  - Demultiplexed FASTQ
- Runtime: <24 hours
- Length of reads: 2x150, or 1x200
  - Illumina libraries
- Accuracy:
  - Google DeepVariant calling
  - 99.25% Indels
  - 99.7% SNVs



PacBio: Sequencing by Binding (SBB)

TOMM40; chr19:44,898,425-44,898,624

NovaSeq 2x151 vs PacBio SBB 1x200



No high-quality reads

No variants called

Two het variants called

Full resolution & phasing:

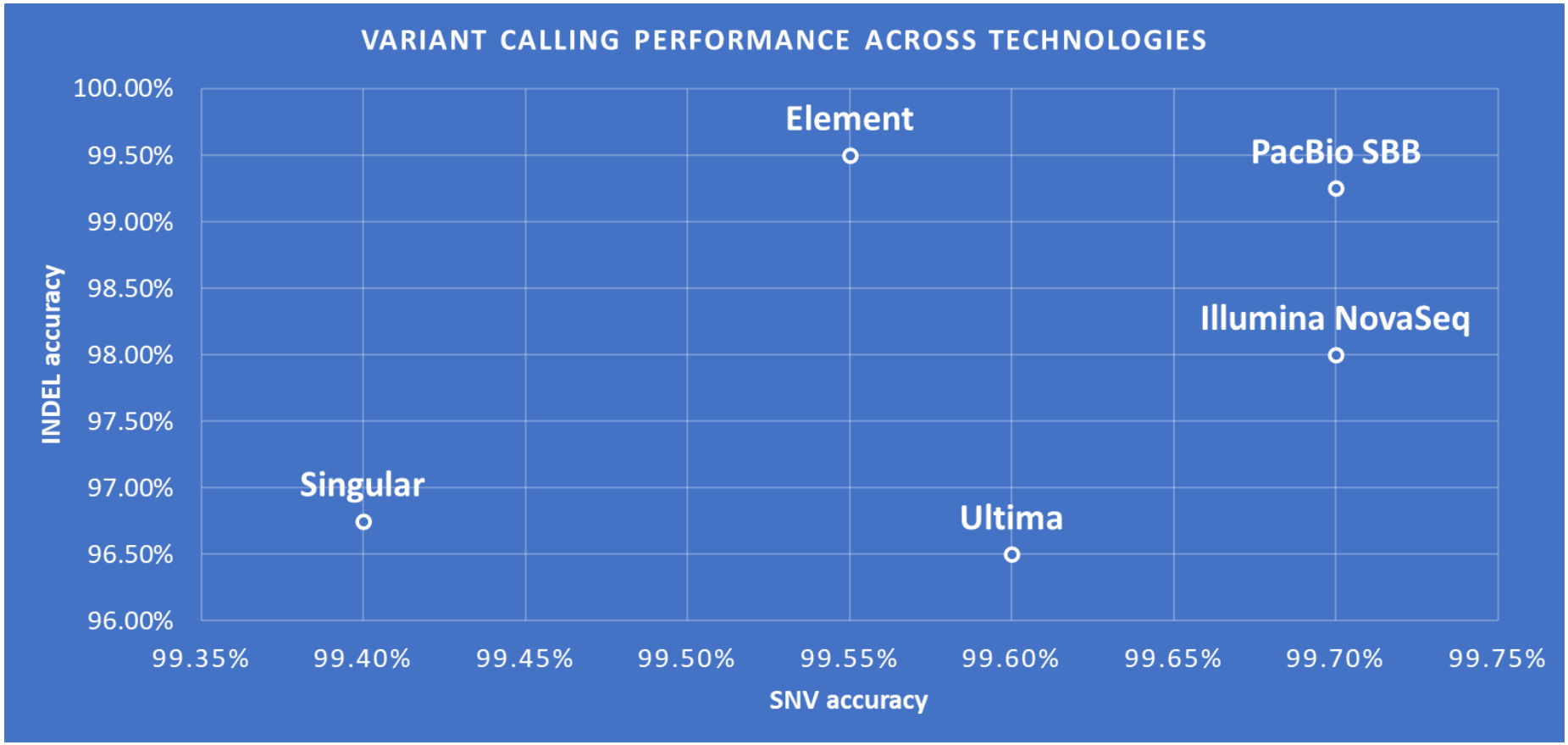
TTTTTTTTTTTTTCTTTTTTTTTTTTTTTT

TTTTTTTTTTTTT-TCTTTTTTTTTTTTTT

PacBio: Onso

- Instrument cost: $259,000
- Instrument size: Benchtop
- Partnerships/acquisitions: Omniome
- Commercial availability: First half of 2023
- Other details
  - Jonas Korlach's presentation: IGM Tech Dev\AGBT2022\PacBio
  - Summary of technology: PacBio Doubles Down on Accuracy by Acquiring Omniome

PacBio SBB Performance Comparison



Metrics from PacBio's AGBT 2022 presentation

Illumina NovaSeq X Series

Key features:

- Shipping starts 1Q 2023 – X Plus price is 1.25 million (NovaSeq X is $985,000)
- NovaSeqX is single flowcell; XPlus is two flowcells
  - 8 lanes; segregation possible
- New optics, new chemistry, new flowcell design, and software updates
  - 50x more stable chemistry – ambient temp for reagents
  - 90% reduction in packaging
  - 2.5x faster base calling
  - Greater accuracy – 50% reduction in sequencing error rate and phasing
  - 5x reduced data footprint (lossless genomic data compression)
  - 320% higher density on flowcell – 10 billion clusters – 6 TB (2x150) in 24 hrs for NovaSeq X - $200 genome (including clustering, sequencing, primary and secondary analysis)
  - Demultiplexing while sequencing
  - The NovaSeq X series achieves ultra-high-resolution imaging through higher numerical aperture, a custom CMOS sensor, and two-channel SBS with blue-green optics

Chemistry “X”: XLEAP-SBS chemistry

- 2024 availability for NextSeq1000/2000 – no compatibility for NovaSeq6000
- Used within NovaSeq X

Illumina Complete Long-Read

- 1Q 2023 – WGS protocol w/ 6-7 kb reads on avg with up to 30 kb
- NextSeq, NovaSeq6000 and NovaSeq X compatible

Illumina NovaSeq X Series

	Output per flow cell run			Reads passing filter per flow cell	Instrument run time		
Platform/Flow Cell	**2 x 50 bp**	**2 x 100 bp**	**2 x 150 bp**	**Paired-End Reads**	**2 x 50 bp**	**2 x 100 bp**	**2 x 150 bp**
Illumina NovaSeq SP	65 – 80 Gb	134 – 167 Gb	200 – 250 Gb	1.3 – 1.6 billion	~ 13 hr	~ 19 hr	~ 25 hr
Illumina NovaSeq S1	134 – 167 Gb	266 – 333 Gb	400 – 500 Gb	2.6 – 3.2 billion	~ 13 hr	~ 19 hr	~ 25 hr
Illumina NovaSeq S2	333 – 417 Gb	667 – 833 Gb	1000 – 1250 Gb	6.6 – 8.2 billion	~ 16 hr	~ 25 hr	~ 36 hr
Illumina NovaSeq S4	N/A	1600 – 2000 Gb	2400 – 3000 Gb	16 – 20 billion	N/A	~ 36 hr	~ 44 hr
Illumina NovaSeq X 1.5B	165 Gb	330 Gb	500 Gb	3.2 billion	~ 13 hr	~ 18 hr	~ 21 hr
Illumina NovaSeq X 10B	1000 Gb	2000 Gb	3000 Gb	20 billion	~ 18 hr	~ 22 hr	~ 24 hr
Illumina NovaSeq X 25B	N/A	N/A	8000 Gb	52 billion	N/A	N/A	~ 48 hr

Short-Read Sequencer Summary

Platform	Reads x run max: (M)	Read length max: (paired-end*, Half of data in reads**)	Run time max: (d)	Yield low: (Gb)	Yield high: (Gb)	Rate max: (Gb/d)	Accuracy
ElemBio AVITI 2022 2fcell	1600	PE150*	2	40	800	400	90% > Q30 80% > Q40 PCR-free
ElemBio AVITI 2022 2fcell 2x75bp (Q4 2022)	1600	PE75*	1	120	240	240	
Singular Genomics G4 F2 4fcell Standard	600-650	PE150*	0.79	15	200	253	
Singular Genomics G4 F2 4fcell Max Read	4000	50	1	180	200	200	
Singular Genomics G4 F3 4fcell (late 2022)	1200-1320	PE150*	0.79	15	400	505	75-90% > Q30
Singular Genomics Systems PX (2023)	NA	PE150*	NA	NA	NA	NA	
PACB SBB (2023H1)	400-500	PE150*	1	120	150	150	90% > Q40
Ultima Genomics UG 100 (4g4-flow runs)	8000-10000	300	0.83	625	10000	NA	85% > Q30
ILMN iSeq 100 1fcell	4	150*	0.79	1.2	1.2	0.79	80% > Q30
ILMN MiniSeq 1fcell	25	150*	1	1.65	7.5	7.5	80% > Q30
ILMN NextSeq 1000 P1/P2 1fcell	400	PE150*	2	30	120	60	85% > Q30
ILMN NextSeq 2000 P3 1fcell	1200	PE150*	2	60	360	150	85% > Q30
ILMN NovaSeq S1 2fcells	3200	150*	1.04	134	1000	600	85% > Q30
ILMN NovaSeq S2 2fcells	6600	150*	1.5	333	2000	1333	85% > Q30
ILMN NovaSeq S4 v1.5 2fcells	20000	150*	1.83	280	6000	3600	85% > Q30

Short-Read Sequencer Summary

Instrument	Chemistry	Max Output	Runtime	Length of reads	SNP recall	SNP precision	Indel recall	Indel precision
Ultima UG100	mnSBS	10B clonally amplified beads	20 hr	310 bp (GIAB)	99.7%	99.6%	96.1%	96.7%
Element AVITI	SBS	2 x 150 240 Gb / 800M read pairs	48 hr	150 bp	99.8%	99.2%	99.7%	99.2%
Singular G4	SBS	2 x 150 50 Gb / 165M clusters	16 hr	150 bp	99.17%	99.71%	96.4%	97.13%
PacBio SBB Instrument	SBB	2 x 150	< 24 hr	150 bp	99.7%		99.25%	
Illumina NextSeq 2000	SBS	2 x 300 360 Gb 2.4B PE reads		300 bp	98%		98%	