
SINGULAR
GENOMICS
JP Morgan
40th Annual Healthcare
Conference 2022
Drew Spaventa | Chairman and CEO

FORWARD-LOOKING STATEMENTS

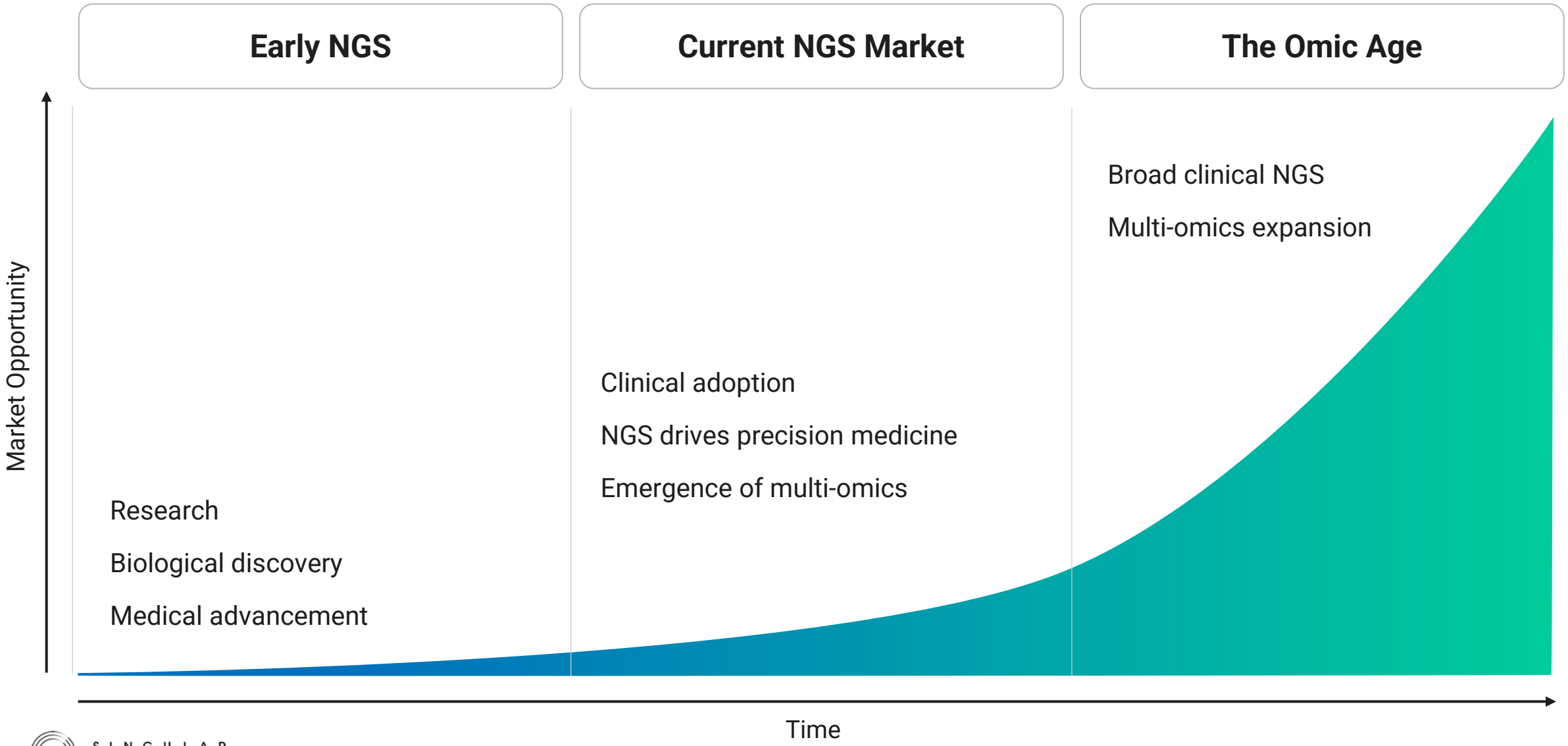
All statements in this presentation and the associated discussion that are not statements of historical facts constitute forward-looking statements within the meaning of the Private Securities Litigation Reform Act of 1995. Forward-looking statements include, but are not limited to, statements regarding: (i) our ability to successfully complete the development of our G4 and PX Integrated Solutions; (ii) our ability to meet our commercial launch and product delivery timelines and objectives; and (iii) our ability to achieve customer and scientific acceptance for our G4 and PX Integrated Solutions. Any such forward-looking statements are based on our management's current expectations and are subject to a number of risks and uncertainties that could cause our actual future results to differ materially from our management's current expectations or those implied by the forward-looking statements. These risks and uncertainties include, but are not limited to: (i) we have incurred significant losses since inception, we expect to incur significant losses in the future and we may not be able to generate sufficient revenue to achieve and maintain profitability; (ii) we have no history commercializing our products or technology; (iii) the life sciences technology market is highly competitive, and if we fail to compete effectively, our business and operating results will suffer; (iv) if we are sued for infringing, misappropriating or otherwise violating intellectual property rights of third parties, this litigation could be costly and time consuming and could prevent or delay us from developing or commercializing our product candidates; (v) if our products fail to achieve early customer and scientific acceptance, we may not be able to achieve broader market acceptance for our products, and our revenues and prospects may be harmed; and (vi) the COVID-19 pandemic and efforts to reduce its spread have adversely impacted, and may materially and adversely impact our business and operations in the future. These and other risk factors that may affect our future results of operations are identified and described in more detail in our filings with the SEC, including our Quarterly Report on Form 10-Q for period ended September 30, 2021, filed with the SEC on November 9, 2021. Accordingly, you should not rely upon forward-looking statements as predictions of future events or our future performance. We disclaim any intention or obligation to revise or update any forward-looking statements, whether as a result of new information, future events, or otherwise.

This presentation also contains estimates and other statistical data made by independent parties and by us relating to market size and growth and other data about our industry. This data involves a number of assumptions and limitations, and you are cautioned not to give undue weight to such estimates.

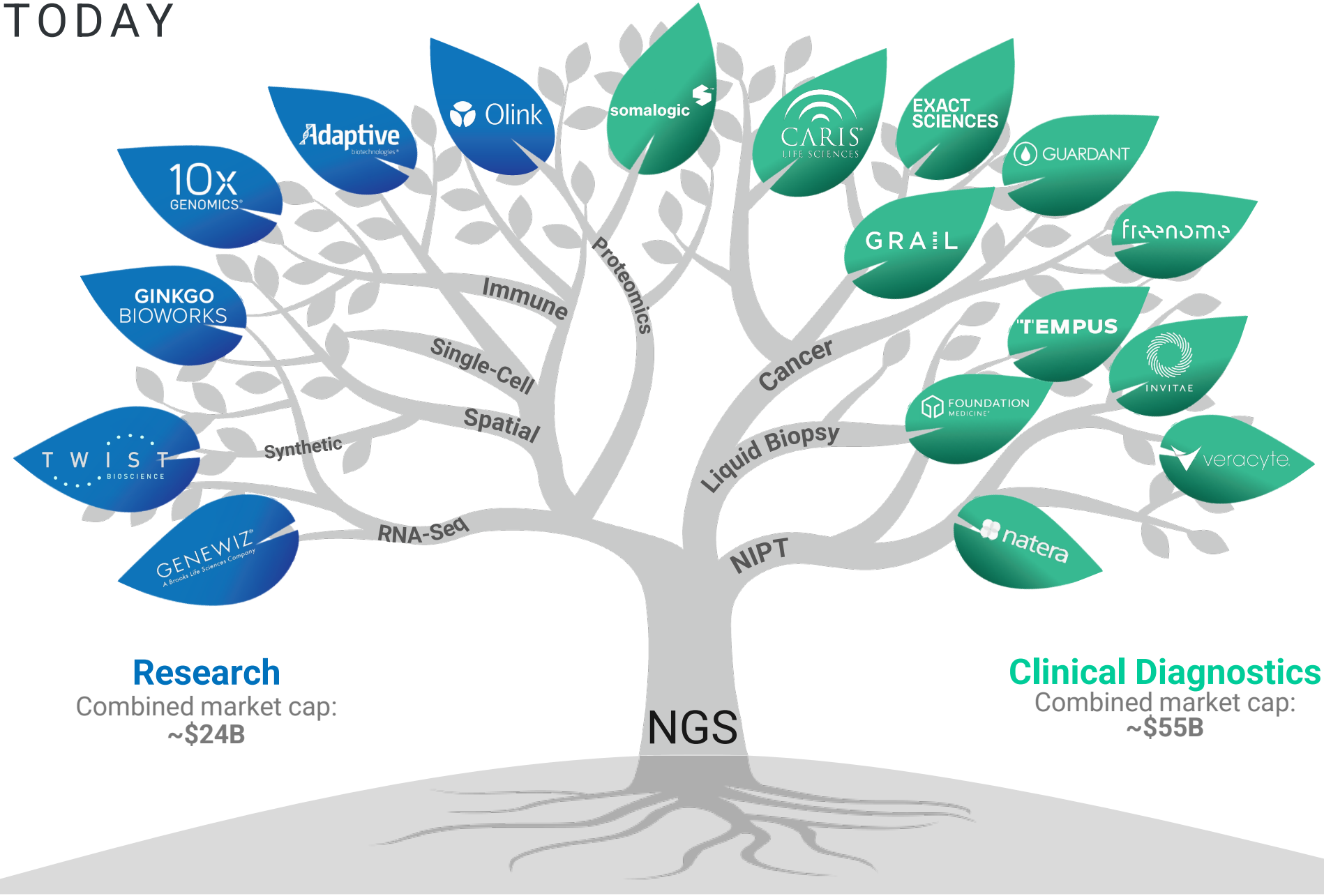
A NEW ERA OF NGS HAS ARRIVED

Setting a new benchmark in power, speed, flexibility and accuracy

WE HAVE ENTERED THE OMIC AGE
VALUE CREATION ON THE BACK OF NGS



NGS TODAY



Research
Combined market cap:
~$24B

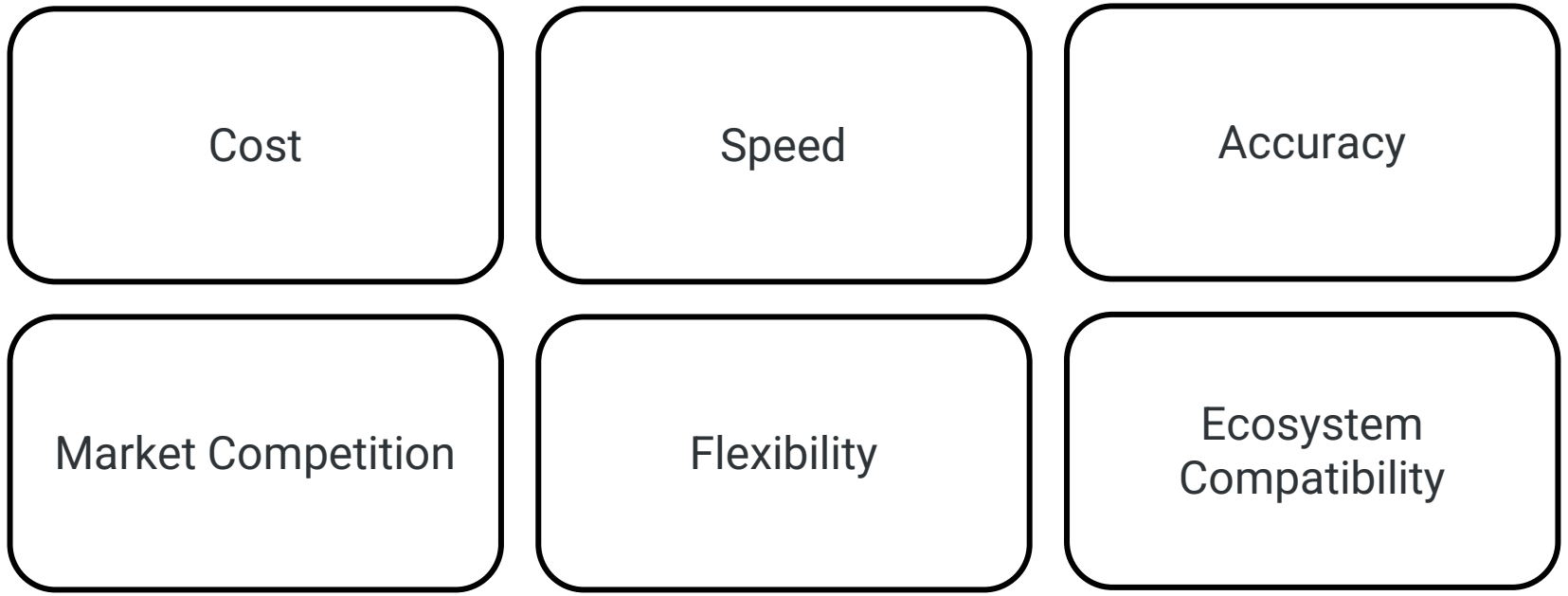
Clinical Diagnostics
Combined market cap:
~$55B

SINGULAR GENOMICS

* Combined market caps based on recent public information and company estimates

VOICE OF CUSTOMER

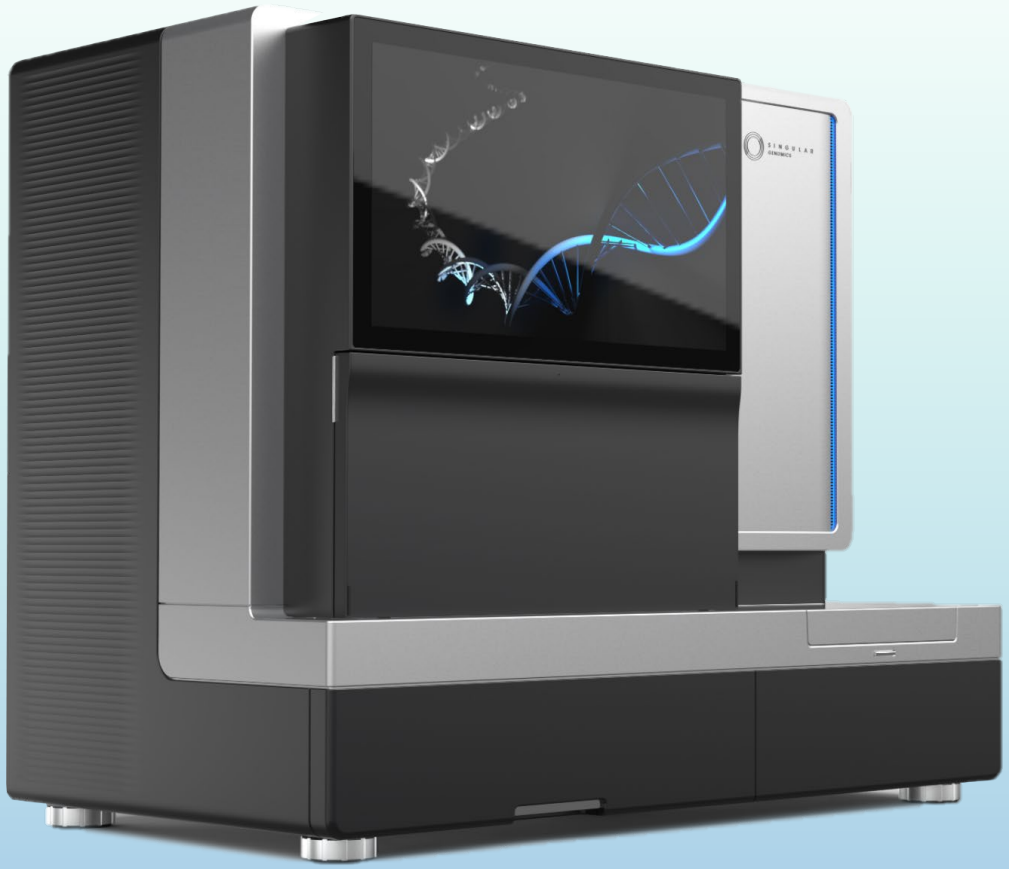
WE LISTENED TO WHAT THEY'RE ASKING FOR



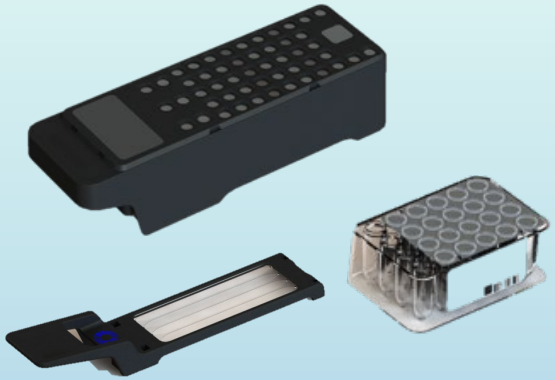
SINGULAR IS ADDRESSING ALL OF THESE

Cost effective ■ Alternate supplier ■ Fast chemistry ■ Run versatility ■ Industry standard accuracy ■ Ecosystem enabled

LAUNCHING THE G4



Fully Integrated
Consumable Kits



Shipping Q2 2022

DISCOVER G4

Power
15—400 Gb output range

More data per day than any other benchtop sequencer

Speed
6—19 hour run time

Industry leading run times

Flexibility
1—4 flow cell
16 lanes

Unparalleled operational efficiency

Accuracy
75–90% bases ≥ Q30

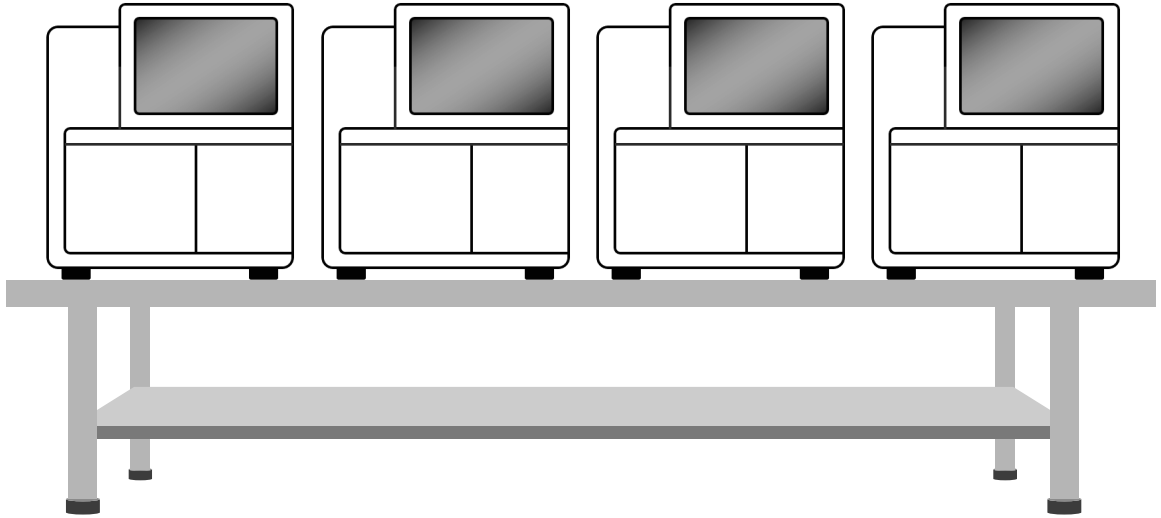
State-of-the-art industry standard

SINGULAR GENOMICS

* All presentation specs are targets

POWER
EQUIVALENT TO 3-4 LEADING BENCHTOP INSTRUMENTS

≥



15—400 Gb run modes

A single G4 will process 4 whole human genomes in 16 hours

SINGULAR GENOMICS

* System comparison in data output per hour

SPEED

ENGINEERED FOR SUB 3-MINUTE CYCLE TIMES

Fast SBS chemistry built from the ground up

Advanced high-resolution optical system

Rapid fluidics

Run times 6 –19 hours

FLEXIBILITY
OPERATIONAL AND COST-EFFECTIVE

Scale experiments up or down

Daily runs

4 independent flow cells

16 independent lanes (16 individual samples)

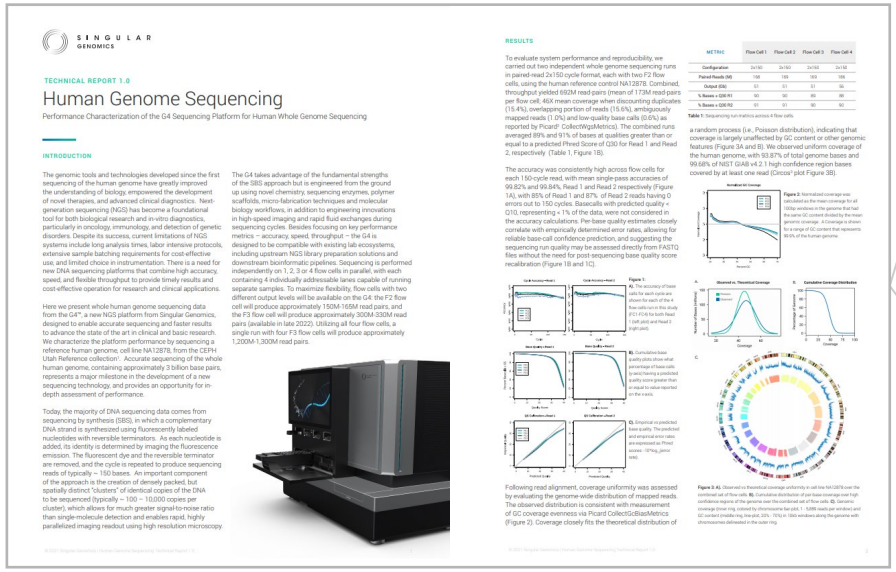
ACCURACY

MATCHING INDUSTRY LEADING SPECIFICATIONS

Proprietary 4-color SBS chemistry

Novel method of paired read sequencing

75–90% bases ≥ Q30 across all kits



SINGULAR GENOMICS

TECHNICAL REPORT 1.0

Human Genome Sequencing

Performance Characterization of the G4 Sequencing Platform for Human Whole Genome Sequencing

INTRODUCTION

The genomic tools and technologies developed since the first sequencing of the human genome have greatly improved the understanding of biology, empowered the development of novel therapies, and advanced clinical diagnostics. Next-generation sequencing (NGS) has become a foundational tool for both biological research and in-vitro diagnostics, particularly in oncology, immunology, and detection of genetic disorders. Despite its success, current limitations of NGS systems include long analysis times, labor intensive protocols, extensive sample batching requirements for cost-effective use, and limited choice in instrumentation. There is a need for new DNA sequencing platforms that combine high accuracy, speed, and flexible throughput to provide timely results and cost-effective operation for research and clinical applications.

Here we present whole human genome sequencing data from the G4™, a new NGS platform from Singular Genomics, designed to enable accurate sequencing and faster results to advance the state of the art in clinical and basic research. We characterize the platform performance by sequencing a reference human genome, cell line NA12878, from the CEPH Utah Reference collection[1]. Accurate sequencing of the whole human genome, containing approximately 3 billion base pairs, represents a major milestone in the development of a new sequencing technology, and provides an opportunity for in-depth assessment of performance.

Today, the majority of DNA sequencing data comes from sequencing by synthesis (SBS), in which a complementary DNA strand is synthesized using fluorescently labeled nucleotides with reversible terminators. As each nucleotide is added, its identity is determined by imaging the fluorescence emission. The fluorescent dye and the reversible terminator are removed, and the cycle is repeated to produce sequencing reads of typically ~ 150 bases. An important component of the approach is the creation of densely packed, but spatially distinct "clusters" of identical copies of the DNA to be sequenced (typically ~ 100 – 10,000 copies per cluster), which allows for much greater signal-to-noise ratio than single-molecule detection and enables rapid, highly parallelized imaging readout using high resolution microscopy.

The G4 takes advantage of the fundamental strengths of the SBS approach but is engineered from the ground up using novel chemistry, sequencing enzymes, polymer scaffolds, micro-fabrication techniques and molecular biology workflows, in addition to engineering innovations in high-speed imaging and rapid fluid exchanges during sequencing cycles. Besides focusing on key performance metrics – accuracy, speed, throughput – the G4 is designed to be compatible with existing lab ecosystems, including upstream NGS library preparation solutions and downstream bioinformatic pipelines. Sequencing is performed independently on 1, 2, 3 or 4 flow cells in parallel, with each containing 4 individually addressable lanes capable of running separate samples. To maximize flexibility, flow cells with two different output levels will be available on the G4: the F2 flow cell will produce approximately 150M-165M read pairs, and the F3 flow cell will produce approximately 300M-330M read pairs (available in late 2022). Utilizing all four flow cells, a single run with four F3 flow cells will produce approximately 1,200M-1,300M read pairs.

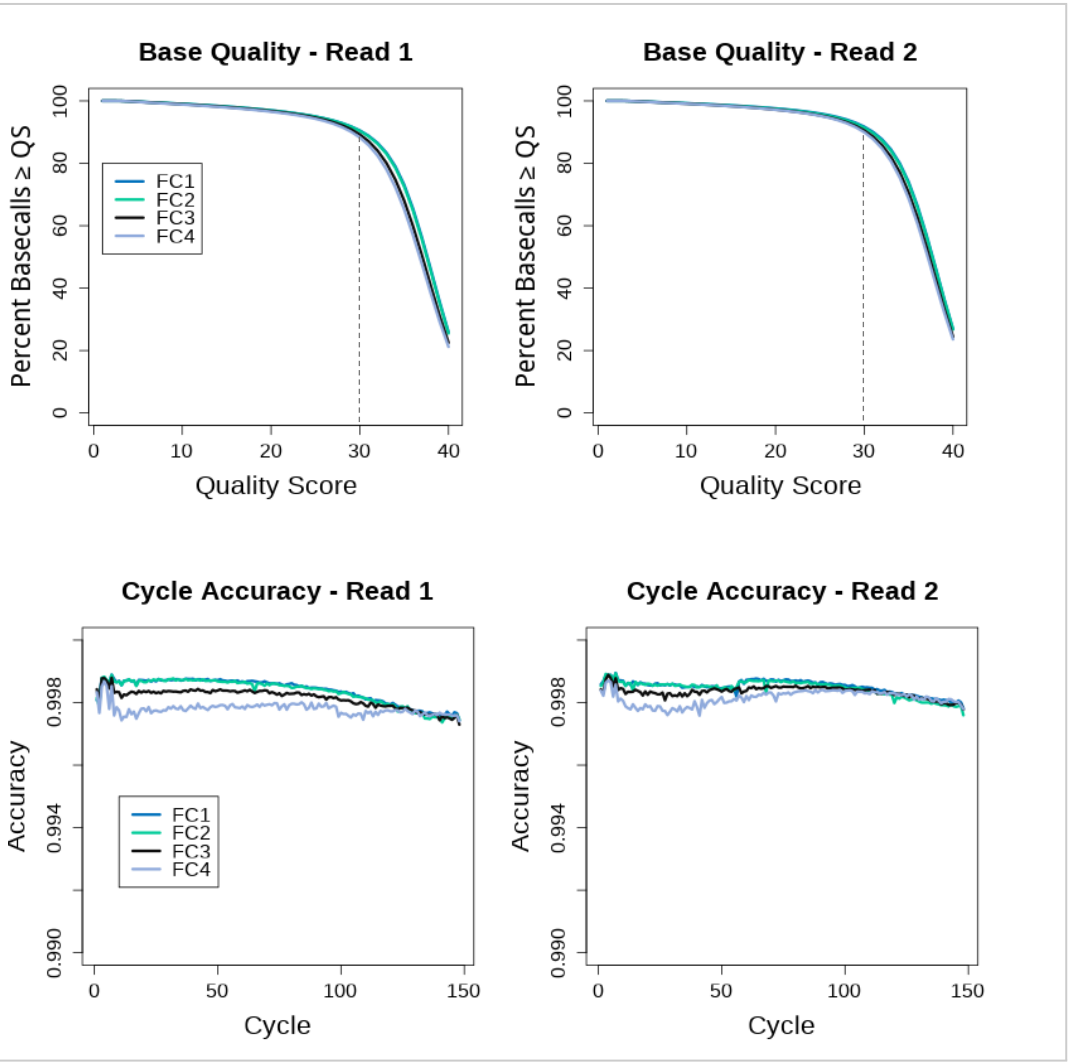
RESULTS

To evaluate system performance and reproducibility, we carried out two independent whole genome sequencing runs in paired-read 2x150 cycle format, each with two F2 flow cells, using the human reference control NA12878. Combined, throughput yielded 692M read-pairs (mean of 173M read-pairs per flow cell; 46X mean coverage when discounting duplicates (15.4%), overlapping portion of reads (15.6%), ambiguously mapped reads (1.0%) and low-quality base calls (0.6%) as reported by Picard® CollectWgsMetrics). The combined runs averaged 89% and 91% of bases at qualities greater than or equal to a predicted Phred Score of Q30 for Read 1 and Read 2, respectively. (Table 1, Figure 1B).

The accuracy was consistently high across flow cells for each 150-cycle read, with mean single-pass accuracies of 99.82% and 99.84%, Read 1 and Read 2 respectively (Figure 1A), with 85% of Read 1 and 87% of Read 2 reads having 0 errors out to 150 cycles. Basecalls with predicted quality = Q10, representing ~ 1% of the data, were not considered in the accuracy calculations. Per-base quality estimates closely correlate with empirically determined error rates, allowing for reliable base-call confidence prediction, and suggesting the sequencing run quality may be assessed directly from FASTQ files without the need for post-sequencing base quality score recalibration (Figure 1B and 1C).

Following read alignment, coverage uniformity was assessed by evaluating the genome-wide distribution of mapped reads. The observed distribution is consistent with measurement of GC coverage evenness via Picard CollectGcBiasMetrics (Figure 2). Coverage closely fits the theoretical distribution of a random process (i.e., Poisson distribution), indicating that coverage is largely unaffected by GC content or other genomic features (Figure 3A and B). We observed uniform coverage of the human genome, with 93.87% of total genome bases and 99.68% of NIST GIAB v4.2.1 high confidence region bases covered by at least one read (Circos® plot Figure 3B).

Enlarged images from technical report



SEAMLESS INTEGRATION INTO EXISTING WORKFLOWS

PARTNERING WITH LEADING PROVIDERS

Prepare

Simple run planning and library loading

TWIST BIOSCIENCE

NEW ENGLAND BioLabs Inc.

LEXOGEN

WATCHMAKER GENOMICS

Dovetail GENOMICS

Sequence

Integrated clustering and sequencing

Analyze

Rapid and accurate analysis

Terra

BROAD INSTITUTE

Microsoft

verily

aws

Other Commercial Informatic Platforms

G4 KIT PERFORMANCE SPECIFICATIONS

READS AND OUTPUT PER FLOW CELL

		F2 FLOW CELL	F3 FLOW CELL
	Number of Reads (clusters)	150—165M	300—330M
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
Quality	75—90% bases ≥ Q30 across all kits		
Accuracy	99.6—99.9% across all kits		

APPLICATIONS

COVERING A WIDE RANGE OF CUSTOMER NEEDS

	F2 KITS (150M READS)				F3 KITS (300M READS)			
	Run Hours	Samples / Lane	Samples / FC	Samples / Run	Run Hours	Samples / Lane	Samples / FC	Samples / Run
RNA Gene Expression [1] (2x50 bp, 10M reads)	8—10	3.75	15	60	6–8	7.5	30	120
Single Cell RNA-Seq (130 cycles, 7,.500 cells/sample and 20,000 reads/cell)	8—10	0.25	1	4	8–10	0.50	2	8
Total RNA-Seq (2x100 bp, 50M reads)	12—15	0.75	3	12	12—15	1.5	6	24
Exome (2x100 bp, 35Mb at 100x coverage)	12—15	1.25	5	20	12—15	2.75	11	44
Target Enrichment (2x150 bp, 800Kb at 4,000x coverage)	16–19	1.5	6	24	16—19	3	12	48
Human Whole Genome [2] (2x150bp, 3Gb at 30x coverage)	–	–	–	–	16—19	0.25	1	4

(1) F2 uses 100 cycle kit; 50 cycle available only on F3 and shows run hours for a 1x50 bp run

(2) F3 kits only for Human Whole Genome

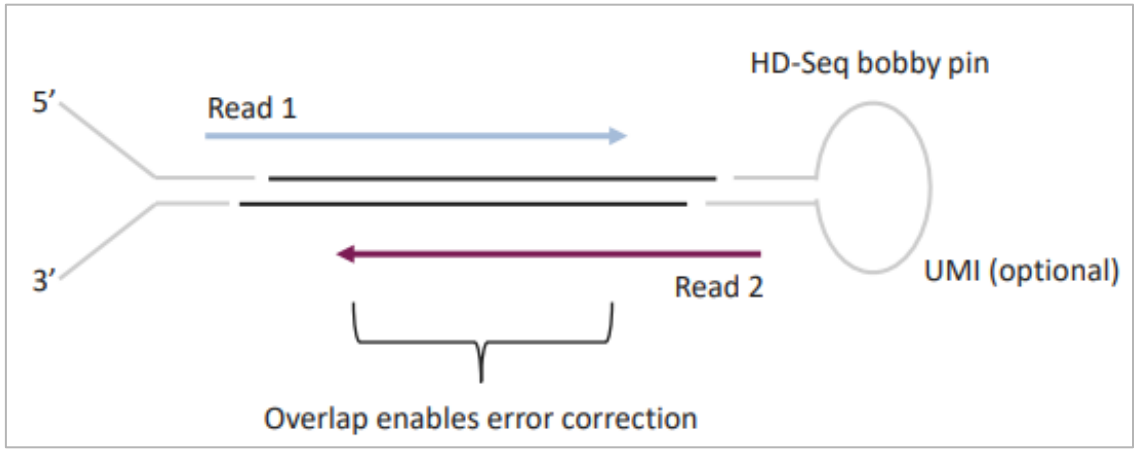
SPECIALIZED APPLICATIONS

NOVEL KITS ENGINEERED TO DRIVE DISCOVERY; EXPECTED LATE 2022

HD-Seq

Q50 accuracy for rare variant detection

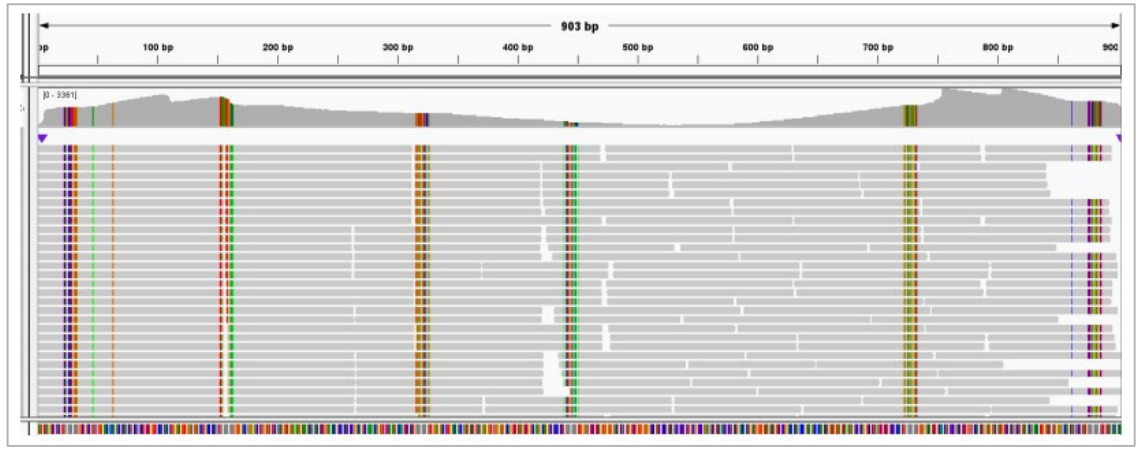
Applications in oncology



XR-Seq

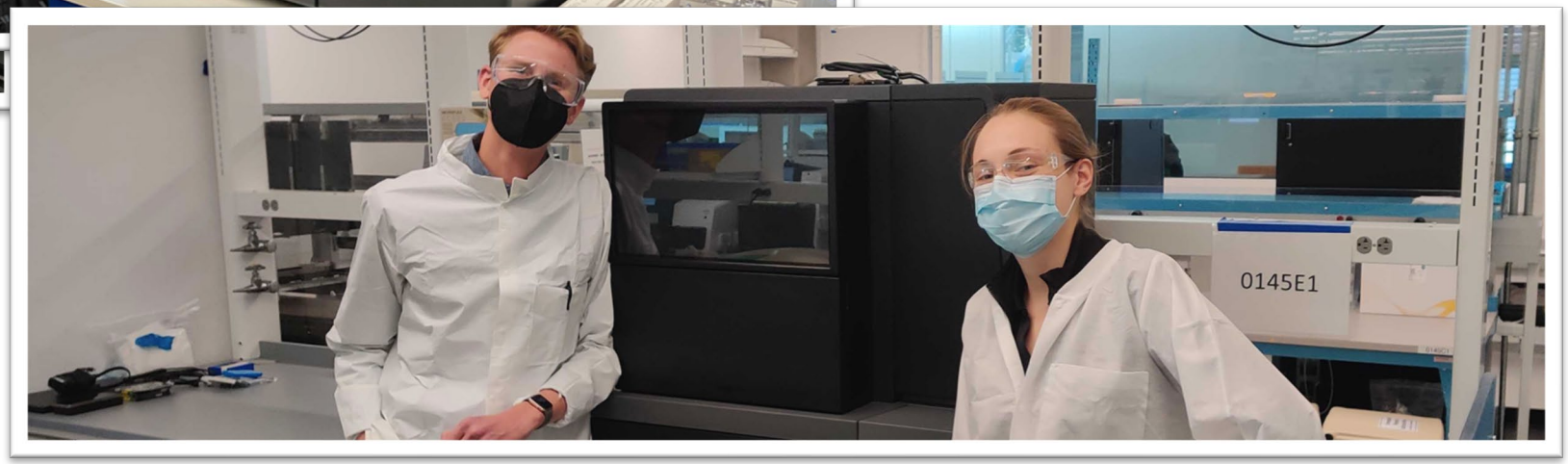
Targeted extended-range sequencing 500—3,000 bp

Applications in immunology and bacterial genomics



EARLY ACCESS PARTNERS

SUCCESSFUL EXECUTION OF FIRST PLACEMENTS



SINGULAR GENOMICS

EARLY ACCESS PARTNERS
ACHIEVING TARGET PERFORMANCE

	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6
Lab Type	Academic core	Commercial clinical	Commercial clinical	Government core	Commercial CRO	Commercial clinical
Application	Spatial transcriptomics	Targeted sequencing	Liquid biopsy	Microbial genome sequencing	Liquid biopsy	TBD
On site status	Complete ☑	Complete ☑	Complete ☑	Complete ☑	In process	Pending
Reads per flowcell	136M	>150M	167M	169M	TBD	TBD
Accuracy	99.6%--99.7% (>75% of bases ≥Q30)	99.6%--99.8% (>75% of bases ≥Q30)	99.7%--99.8% (>80% of bases ≥Q30)	99.7%--99.9% (>80% of bases ≥Q30)	TBD	TBD

TARGET CUSTOMER PROFILES
DELIVERING ON A BROAD AND DIVERSE OPPORTUNITY

Strong value proposition for customer segments

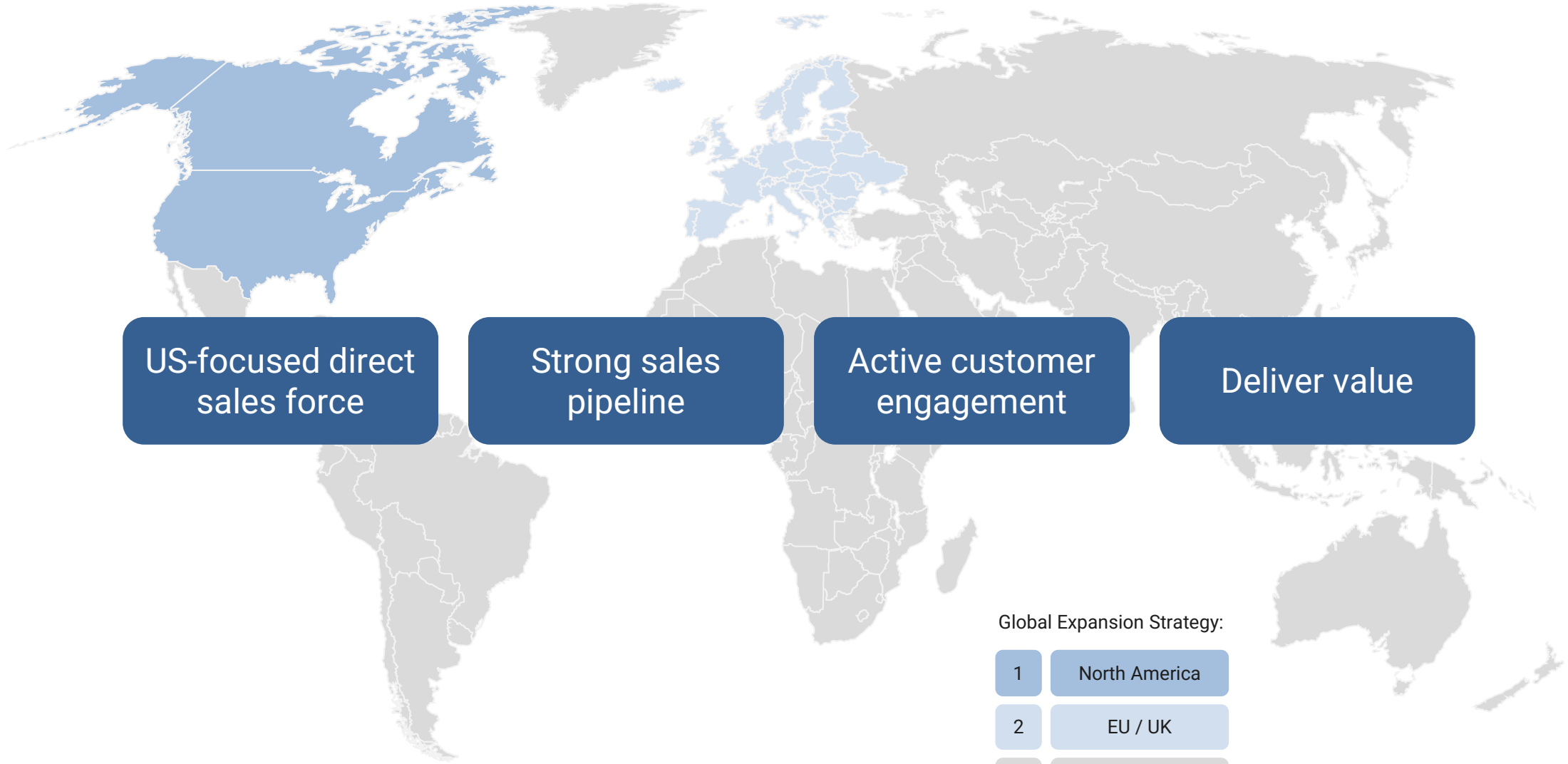
Academic Core	RNA, single cell, targeted panels, exomes, WGS
Clinical *	Targeted panels, RNA, exomes, rapid WGS
Emerging Growth	RNA, single cell, targeted panels, spatial

SINGULAR GENOMICS

* Initial products are RUO

POISED FOR STRONG COMMERCIAL ADOPTION

FOCUSED ROLLOUT AND GROWTH STRATEGY



US-focused direct sales force

Strong sales pipeline

Active customer engagement

Deliver value

Global Expansion Strategy:

1	North America
2	EU / UK
3	Rest of World

SINGULAR GENOMICS

MULTI-OMICS REIMAGINED

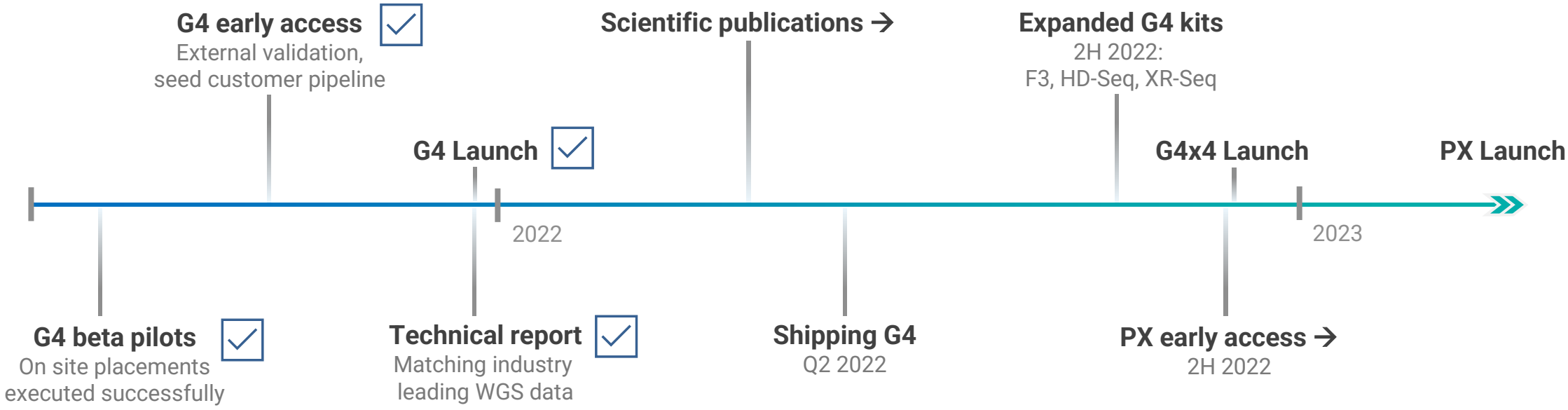
Harnessing the power of sequencing in single cell analysis and spatial profiling

We are busy in the lab…

PX launching 2023

TIMELINE

PATTERN OF EXECUTION, AND KEY UPCOMING MILESTONES



SINGULAR IS UNIQUELY POSITIONED FOR GROWTH

STRONG TAIL WINDS

Market	Customers desire competition
Technology	G4 offers superior benchtop performance
Clinical	Addressing high growth clinical needs
Multi-omics	Leveraging sequencing to address the next frontier

SINGULAR GENOMICS