The power of sensitivity and coverage at scale

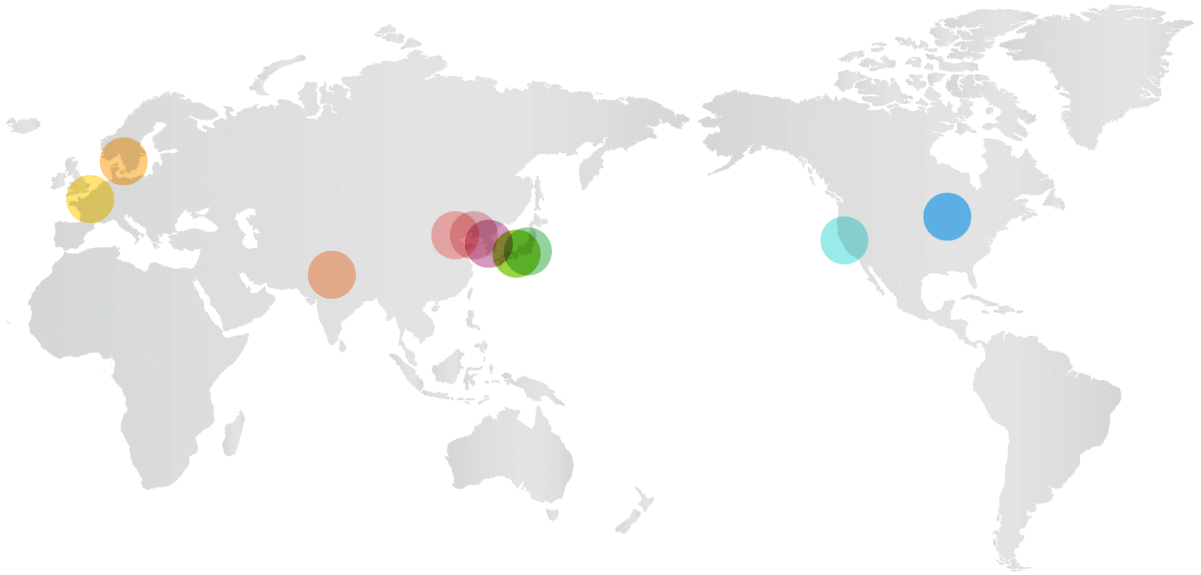
Bryan Bell, PhD, associate director, market strategy
Xuan Li, PhD, senior staff engineer, platform integration
Peng Xu, PhD, staff scientist

that's GOOD science!®

Agenda

- Introducing the Takara Bio group
- Innovations in single-cell NGS
- Unmet needs in biomarker discovery
- New high-throughput single-cell whole-genome amplification (WGA)
- New high-throughput single-cell total RNA-seq

Takara Bio group worldwide locations



>1,500 employees worldwide

Takara Bio Inc.

Corporate headquarters: Kusatsu, Shiga, Japan

Worldwide affiliates

- Takara Bio Europe SF (Gothenburg)
- Takara Bio Europe S.A.S. (Paris)
- Takara Bio USA, Inc. (San Jose)
- Takara Biomedical Technology Co. Ltd. (Beijing)
- Takara Biotechnology Co. Ltd. (Dalian)
- Takara Korea Biomedical Inc. (Seoul)
- DSS Takara Bio India Pvt. Ltd. (New Delhi)

Additional facilities

- Center for Gene and Cell Processing (Kusatsu)
- Stem cell research & manufacturing (Gothenburg)
- Satellite offices (Madison, Tokyo)

About Takara Bio USA, Inc.

1984

Established as Clontech Laboratories, Inc.

30+ years of experience supporting the life sciences' research community

2005

Acquired as a wholly owned subsidiary of Takara Bio Inc.

50+ years of experience providing a diverse array of biotechnology products and services

2017

Acquired Rubicon Genomics Inc. and WaferGen Bio-systems, Inc.

Providing unmatched sensitivity for demanding sequencing applications and complete workflows for automated sample processing and analysis

2021

Headquarters relocated to San Jose, CA

Innovating single-cell NGS technologies for 15 years

Plate-based single-cell RNA-seq

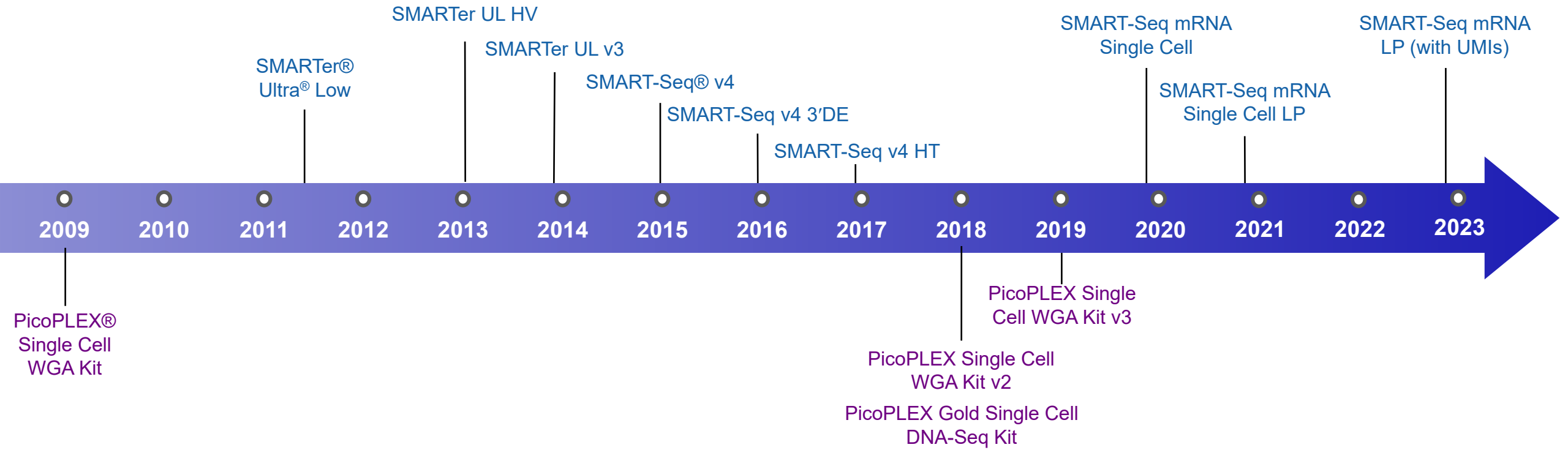


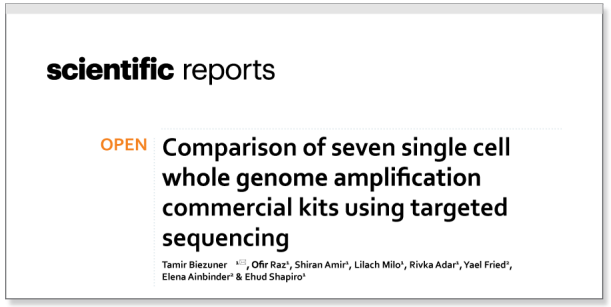
Plate-based single-cell DNA-seq

TaKaRa

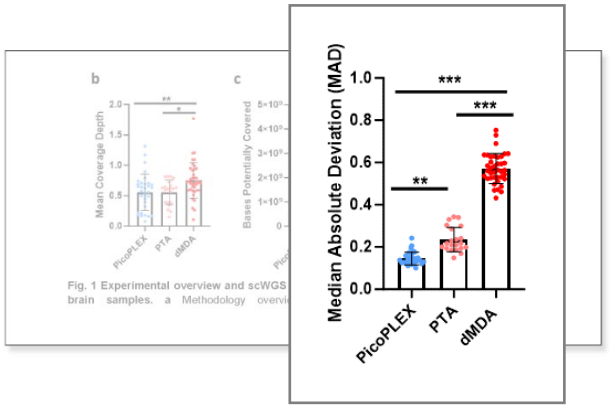
Achieve single-cell WGA sensitivity, uniformity, and reproducibility

Unlock the genome with PicoPLEX technology

- Achieve high reproducibility and accuracy
- Outperform the leading dMDA and PTA technologies for improved CNV detection
- Attain highly uniform coverage with low allele drop-in and drop-out rates



"[PicoPLEX] was the most reliable kit, showing reproducible results for all cells, both in the coverage perspective and both in reproducibility perspective, with low variance for all analyzed cells."



"We noted clear differences across methods, with PicoPLEX performing best…"

Harness the power of sensitivity and full gene body coverage for single-cell RNA-seq (scRNA-seq)

Prestigious institutions have profiled the full transcriptome with SMART-Seq technology

- Unparalleled sensitivity and reproducibility
- Robust, full-length chemistry
- More than differential gene expression

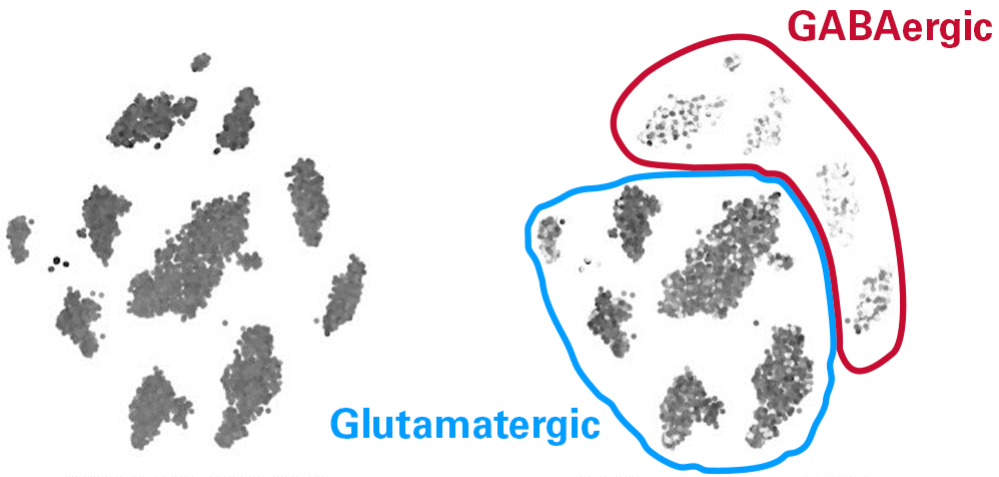
PRESS RELEASE

Takara Bio USA, Inc. elevates the sensitivity of scRNA-seq in its continuing mission to support the efforts of single-cell researcher

DATE: January 26, 2021 AUTHOR: Takara Bio USA, Inc. CATEGORIES: DNA clean-up | Press release

Mountain View, CA—January 26, 2021—Takara Bio USA, Inc. (TBUSA), a wholly owned subsidiary of Takara Bio Inc., is proud to announce that researchers in the laboratory of Stephen Quake, Professor at Stanford University and a pioneer in single-cell genomics applications, have pushed the sensitivity of gene detection to new heights using the recently released SMART-Seq Single Cell Kit from TBUSA.

"We consistently find that the Takara Bio SMART-Seq Single Cell kit demonstrates greater sensitivity than our Smart-seq2 protocol. On mammalian cells from tissue culture we found the Takara Bio kit on average detected ~7500 genes, compared to ~5000 genes found using the Smart-seq2 protocol, while with toxoplasma parasites we detected ~2300 with Takara's kit compared to 866 with Smart-seq2," said Yuan Xue, a graduate researcher in the Quake Lab.



Nondifferential gene: ***Unc5c***

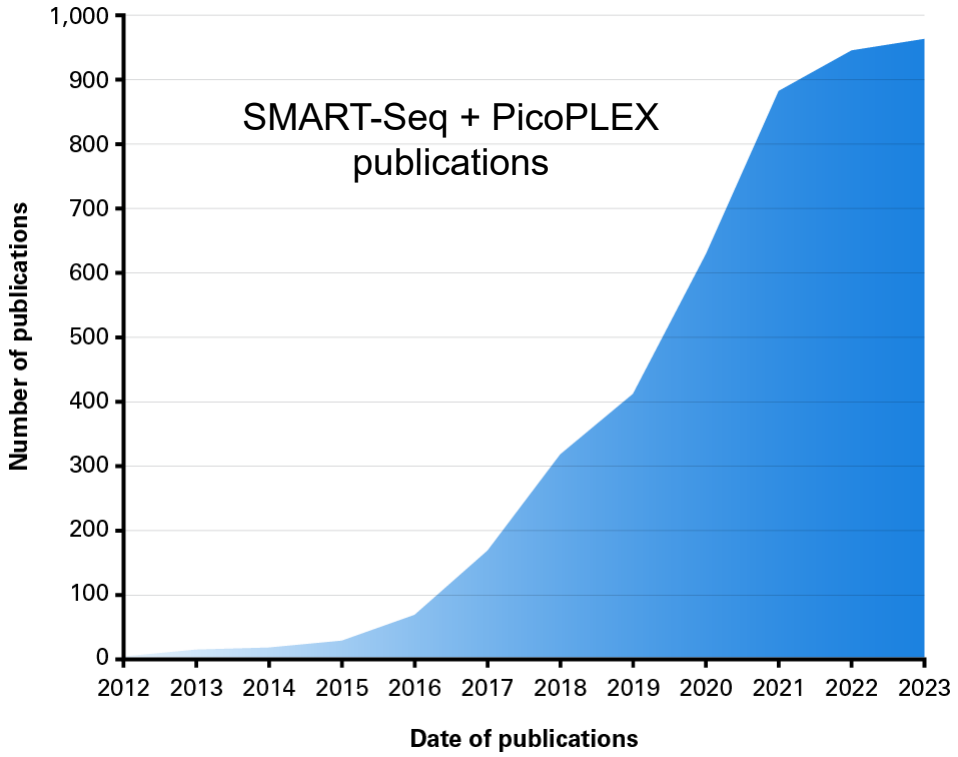
Differential isoform: ***Unc5c-208***

Employing SMART-Seq chemistry, the **Allen Institute for Brain Science** published a preprint paper that was later published in Nature.

Figure adapted from "Isoform cell type specificity in the mouse primary motor cortex" (Booeshaghi et al. 2020, *bioRxiv*) under a CC BY 4.0 license.

The desire for scale has grown

Takara Bio single-cell solutions maintain sensitivity and coverage but lack high scale



Advances in scale of scRNA-seq over the years sacrifice sensitivity and detection of biomarkers

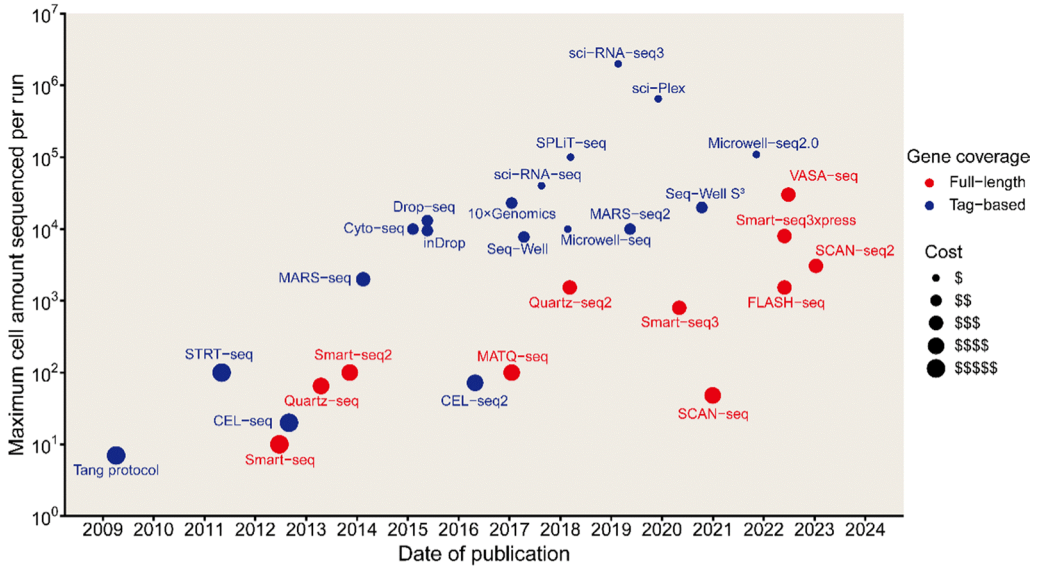
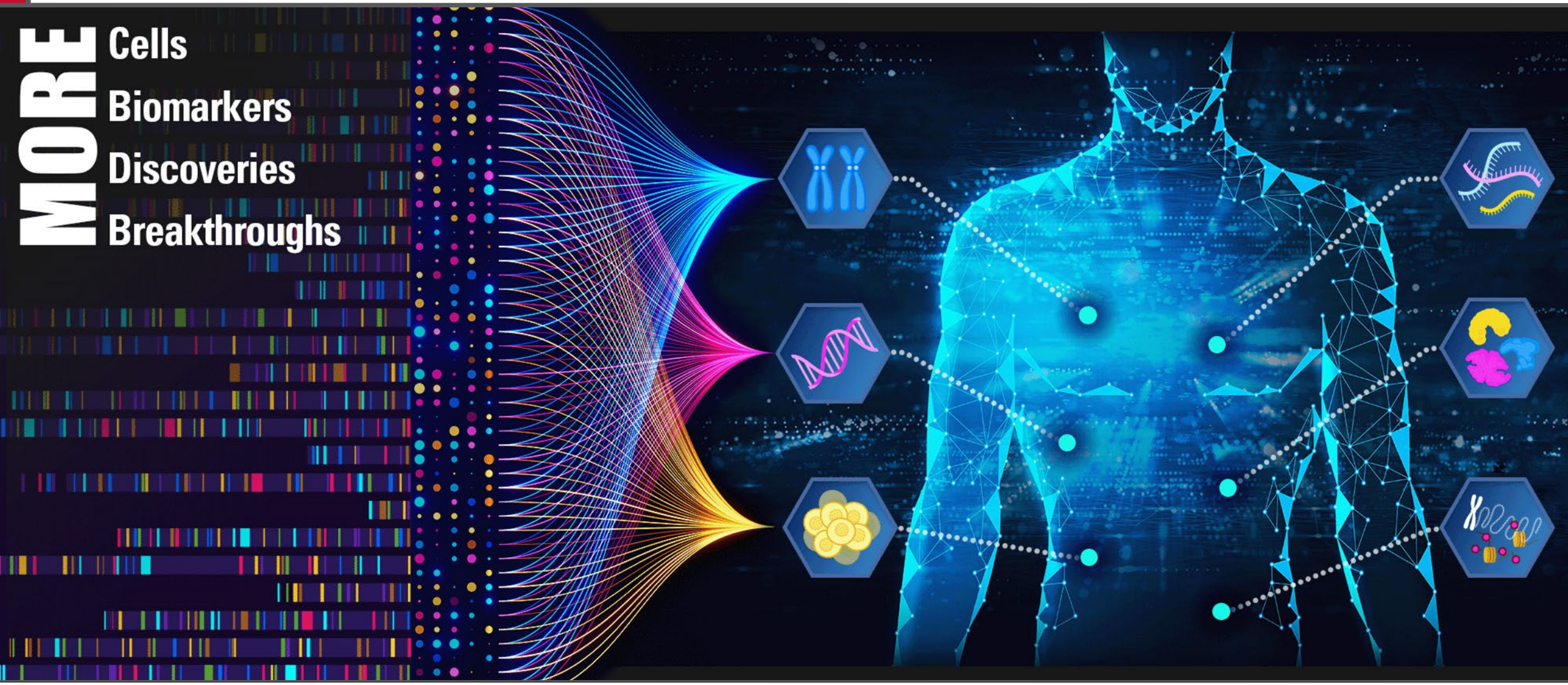


Figure adapted from "Advances in single-cell RNA sequencing and its applications in cancer research" (Huang et al. 2023, *J. Hematol. Oncol.)* under a CC BY 4.0 license.

Next-generation single-cell biomarker discovery, scaled

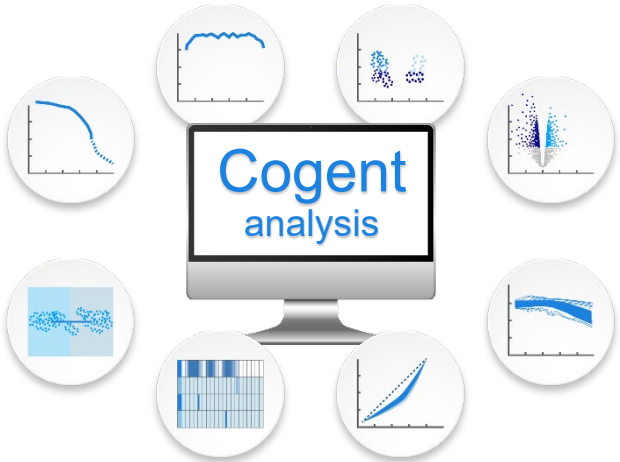


TaKaRa

Next-generation single-cell biomarker discovery, scaled

Shasta™ Single-Cell System

Cogent™ NGS Analysis Pipeline and Discovery Software



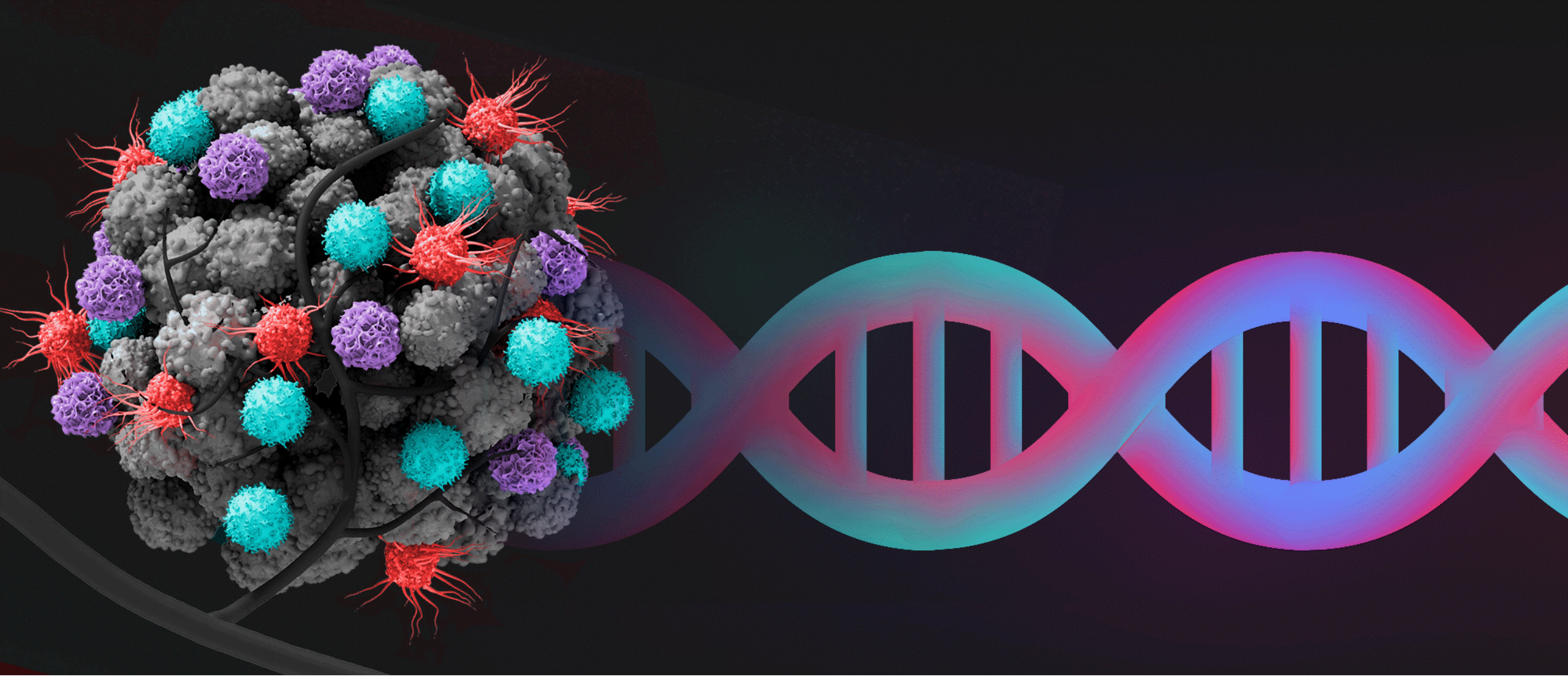
Shasta Whole-Genome Amplification Kit

- Analyze over 1,500 single cells per run
- Profile copy number variation (CNV) data, including chromosomal aneuploidies, and single-nucleotide variation (SNV) data
- Resolve tumor heterogeneity and track clonal evolution with user-friendly bioinformatics tools

Shasta Total RNA-Seq Kit

- Analyze full-length transcriptomes of up to 100,000 single cells per run with outstanding sensitivity
- Uncover multiple RNA biotypes
- Detect splicing isoforms and gene fusions
- Use bioinformatics tools to decode expression patterns of protein-coding and noncoding genes

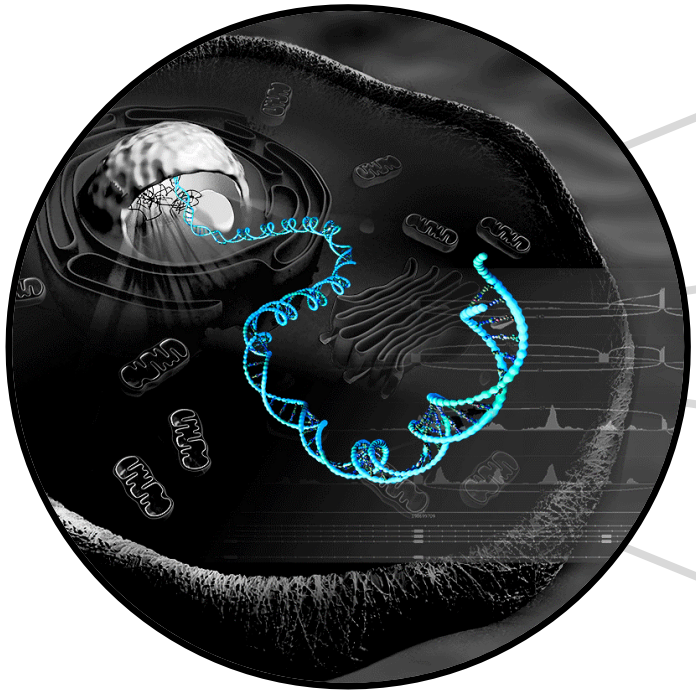
TaKaRa

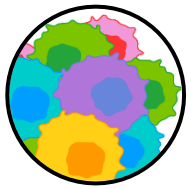


Shasta Whole-Genome Amplification Kit

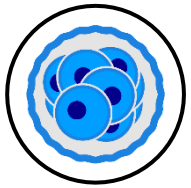
TaKaRa

The power of single-cell WGA

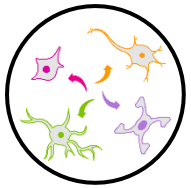




Tumor evolution and heterogeneity



Developmental biology



Neuronal diversity

Biomarker discovery

Copy number variation (CNV)
Single nucleotide variation (SNV)
Structural variation (SV)

TaKaRa

Timeline of key WGA technology developments

Current shortcomings of WGA

- Plate-seq approaches
  - Lack throughput
  - Require deep (expensive) sequencing
  - Lack automated solutions
- Droplet and combinatorial approaches
  - Have lower data quality and limited resolution
  - Require targeted sequencing (not WGA) for commercial options

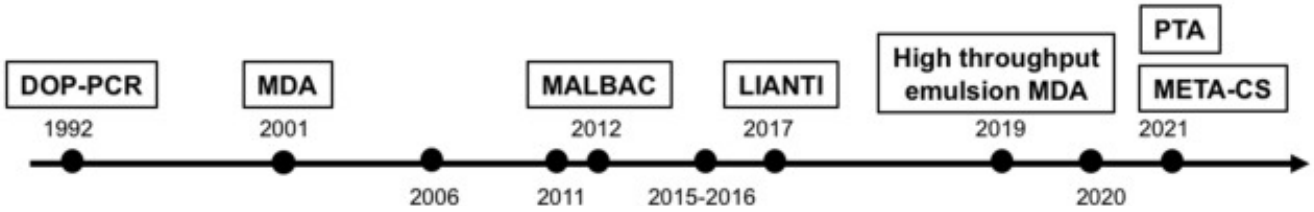
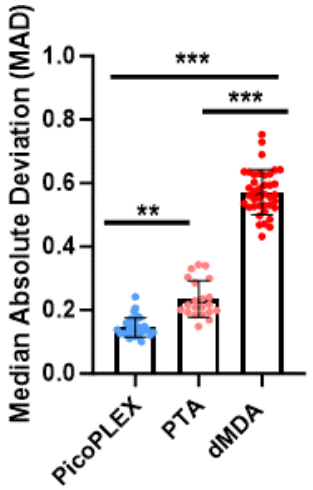


Figure adapted from “Single-Circulating Tumor Cell Whole Genome Amplification to Unravel Cancer Heterogeneity and Actionable Biomarkers” (Khan et al. 2022, *Int. J. Mol. Sci.)* under a CC BY 4.0 license.

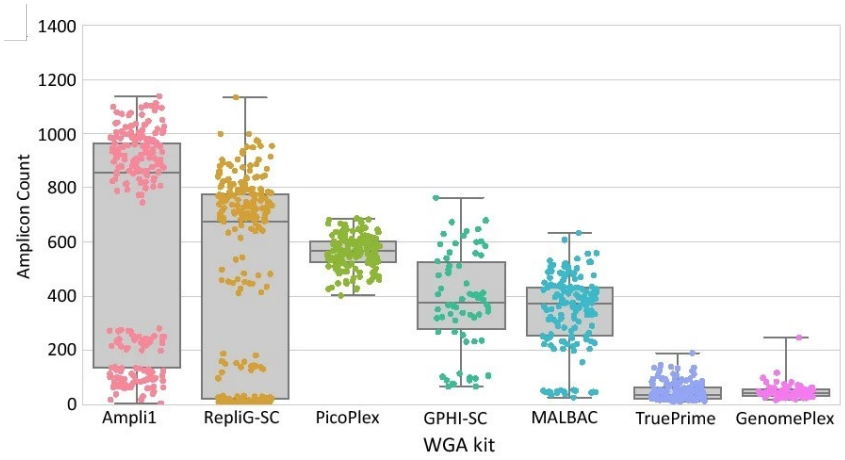
Sensitive, uniform, and reproducible WGA by PicoPLEX technology

Most even



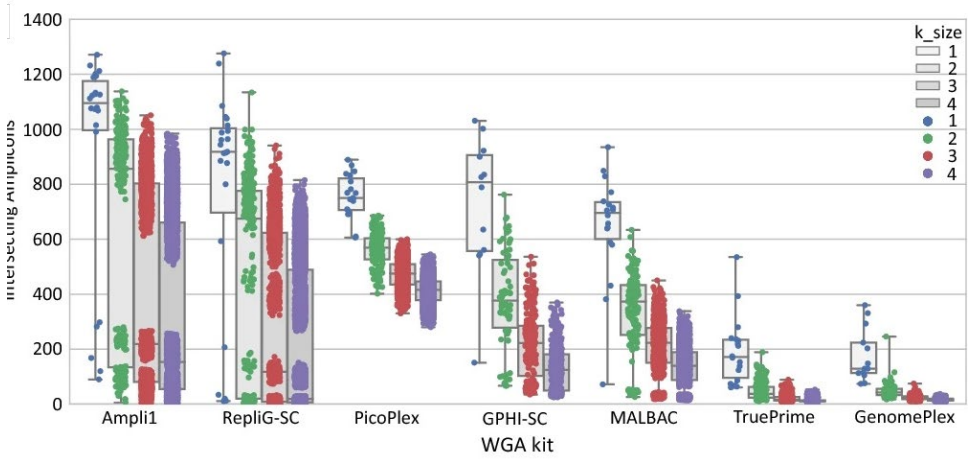
PTA provided the broadest amplification, but PicoPLEX application provided the most even amplification (2023)

Most reliable



scWGA genome coverage analysis: PicoPLEX kit was the most reliable, with the tightest interquartile region (IQR) of all kits, and no failed cells.

Most reproducible



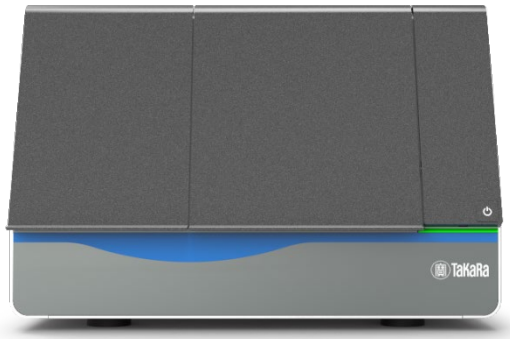
scWGA reproducibility analysis: PicoPLEX application demonstrated high reproducibility for all cells

Figure adapted from "Single-cell somatic copy number variants in brain using different amplification methods and reference genomes" (Kalef-Ezra. et al. 2023, *bioRxiv)* under CC BY 4.0 license.

Figures adapted from "Comparison of seven single cell whole genome amplification commercial kits using targeted sequencing" (Biezuner et al. 2021, *Sci. Rep.)* under a CC BY 4.0 license.

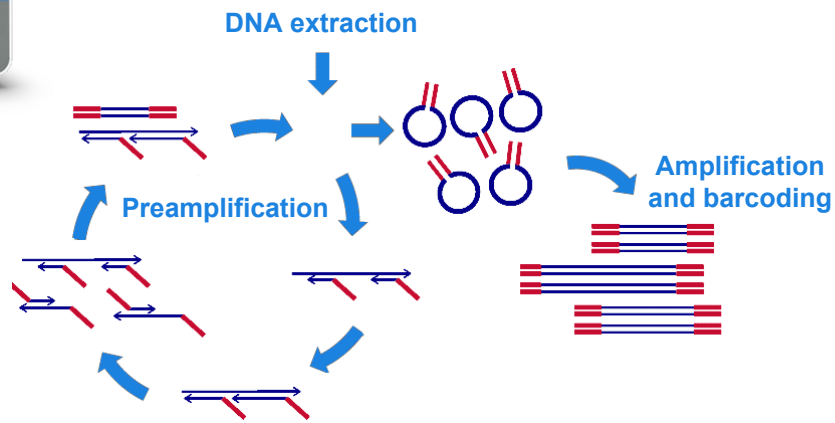
Automated, high-throughput solution for single-cell WGA

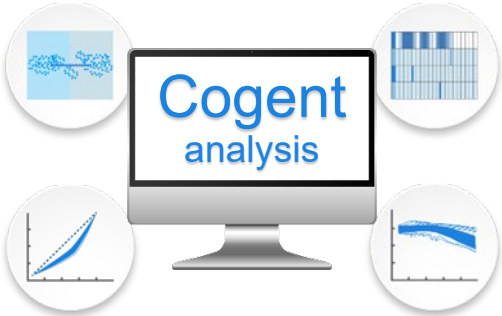
Shasta WGA Kit



- ✓ High-throughput WGA
  Process up to 1,500 cells per run
- ✓ Lower sequencing cost
  Analyze CNV, SNV, and SV at low depth
- ✓ Automated workflow on the Shasta instrument
  Obtain library in one day
- ✓ Leading chemistry for uniformity and reproducibility
  Take advantage of PicoPLEX WGA chemistry
- ✓ End-to-end solution
  Use free Cogent bioinformatics tools

PicoPLEX WGA



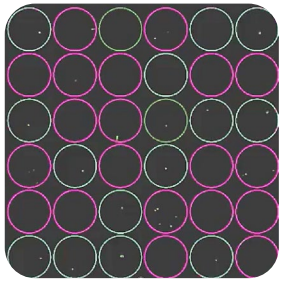


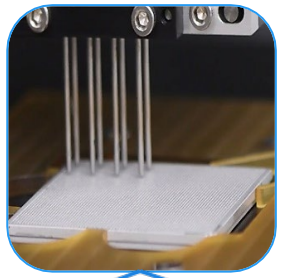
TaKaRa

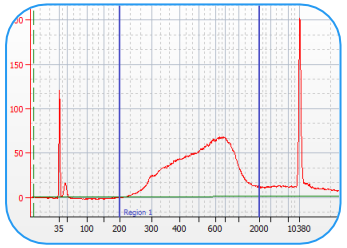
Shasta WGA workflow

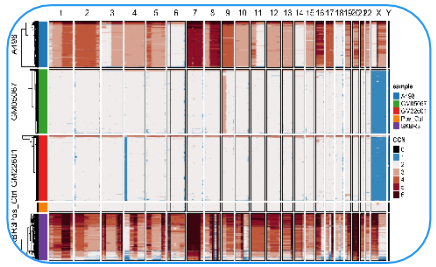
DAY 1: Laborious pipetting replaced by automatic dispensing

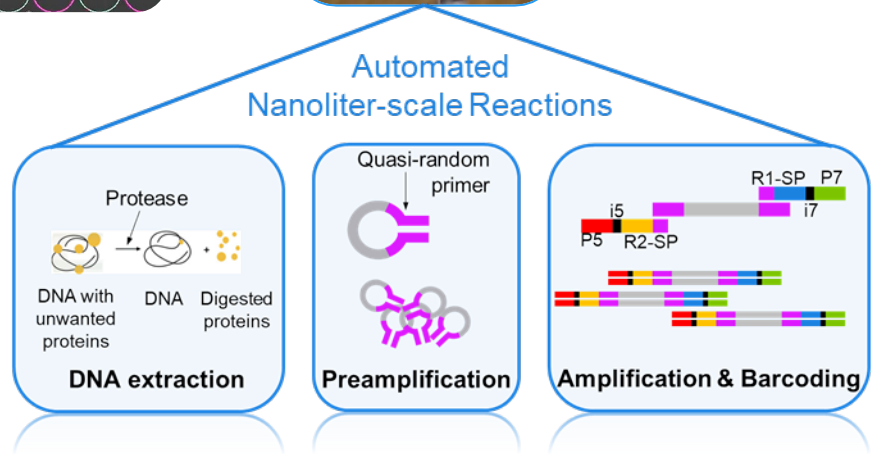
Stained cell suspension → Cell dispense & selection → Library generation → Cleanup & quantification → Sequencing → Analysis & visualization





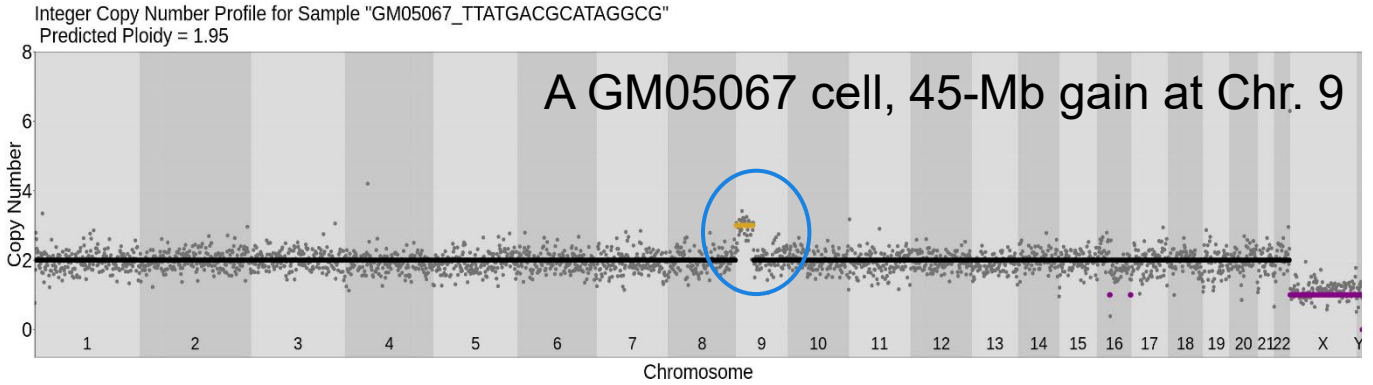


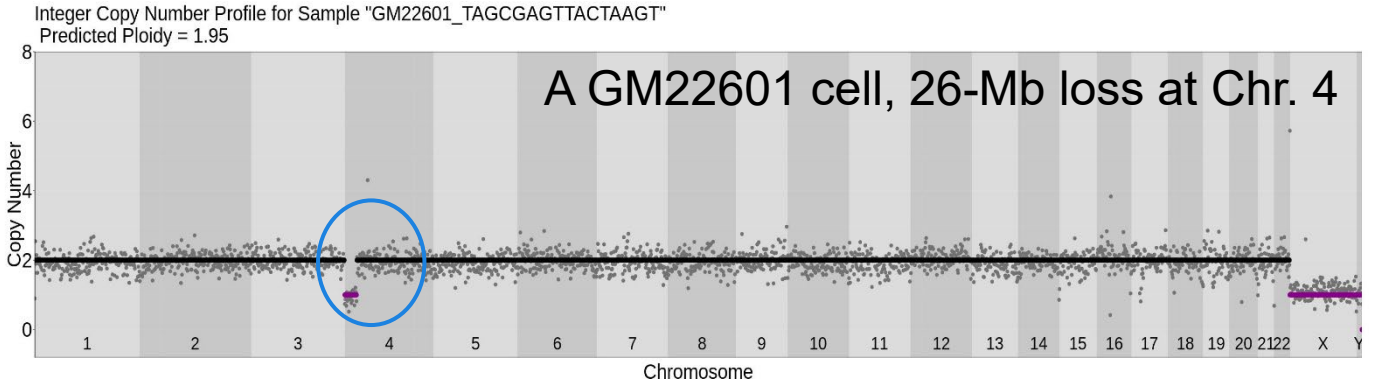




Created with BioRender.com.

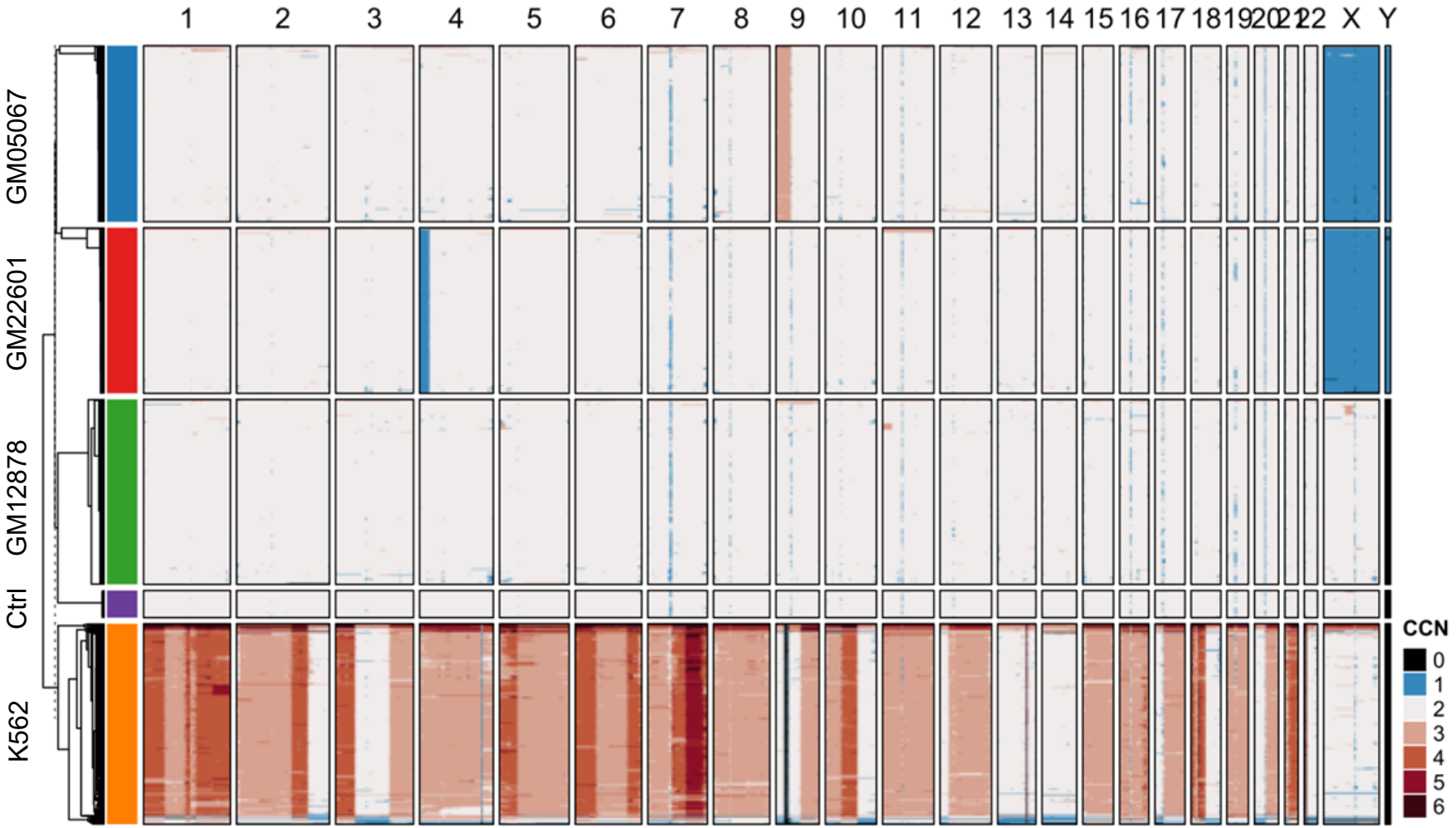
Get single-cell-level CNV with shallow sequencing





1 x 10^6 reads/cell, 2 x 75 bp, 1 Mb average bin size

Copy number profiles of 1,124 single cells



Four cell lines

- GM05067
- GM22601
- GM12878
- K562

250,000 reads/cell

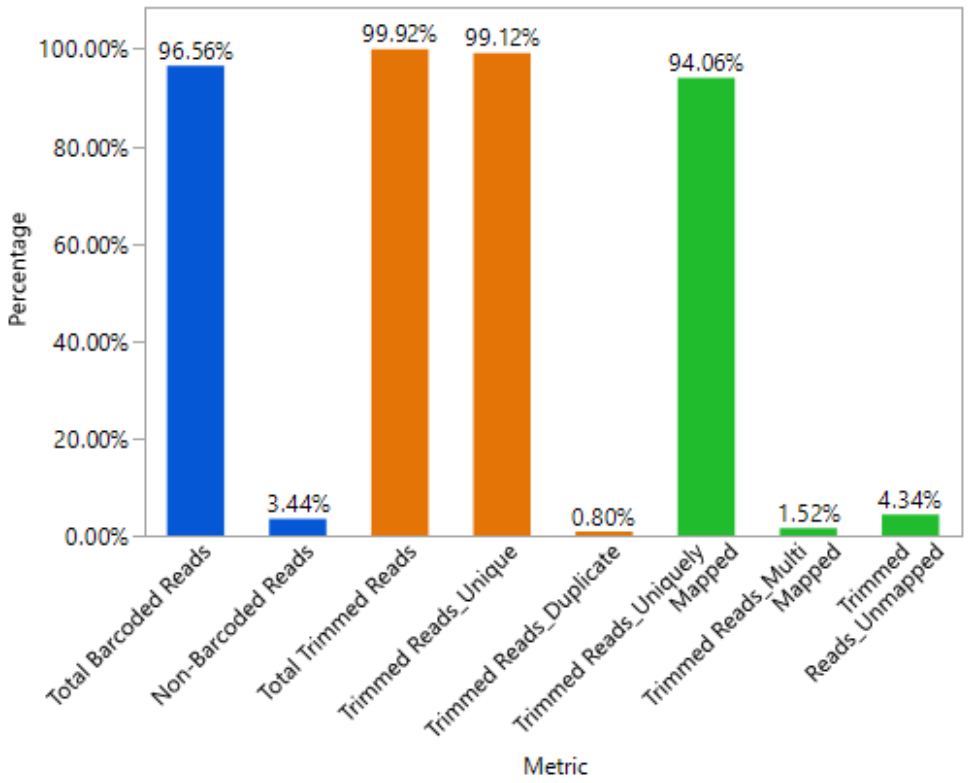
2 x 75 bp

1 Mb average bin size

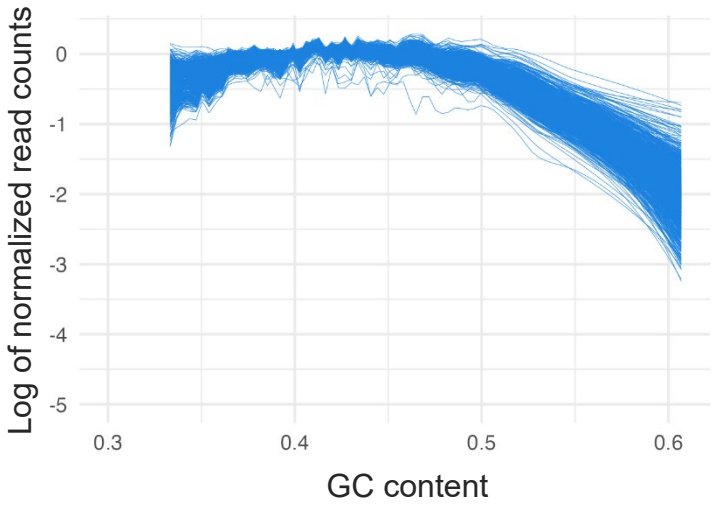
TaKaRa

Obtain good data quality with our nanoliter-scale chemistry

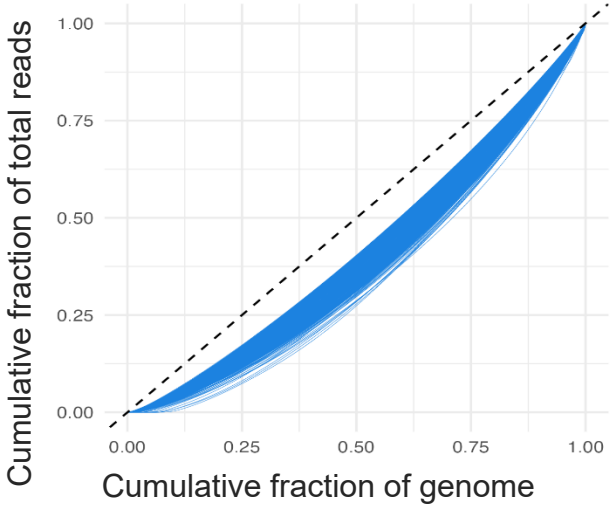
QC metrics



GC content vs. bin counts



Lorenz curve for coverage uniformity

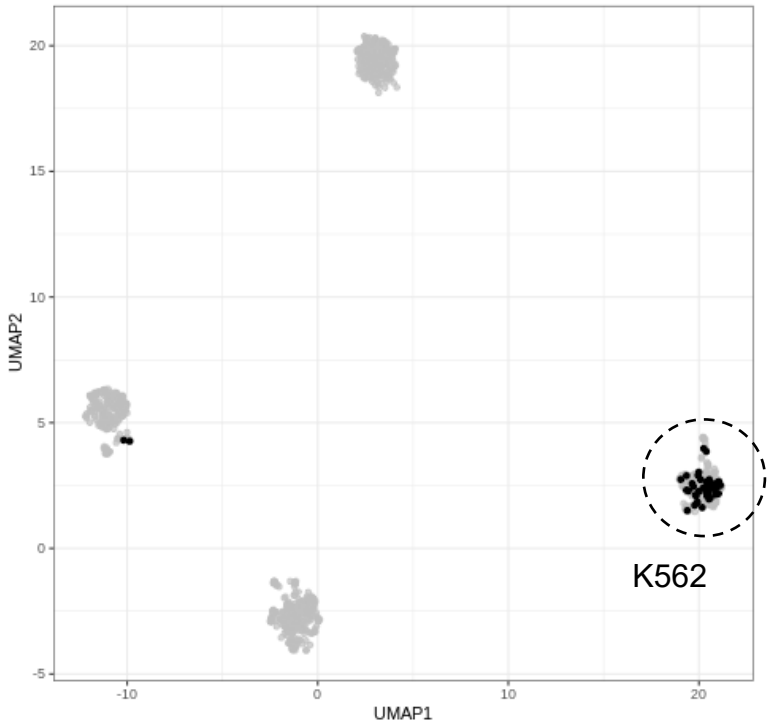


Pair sensitivity and sequencing affordability

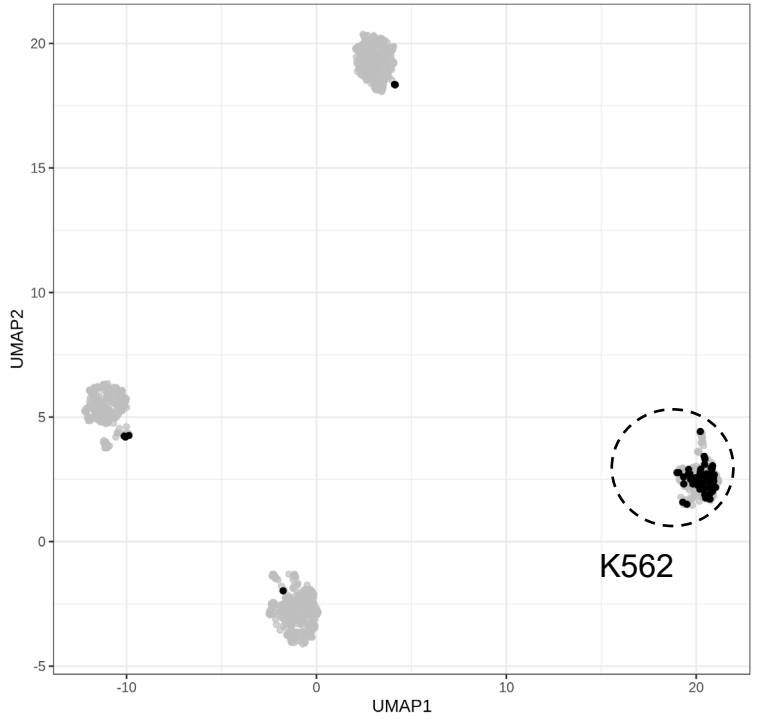
Pseudo-bulk SNV analysis for cell clusters

- Clustered single cells by their CNV profiles (~100 million reads)
- Performed pseudo-bulk SNV analysis (Monopogen [Dou et al. 2023])
- Called germline variants for each cluster and putative somatic variants for each single cell

Chr. 7: 55034534: A:C
A variant within the *EGFR* gene



Chr. 4: 105200804: G:A
A variant within the *TET2* gene

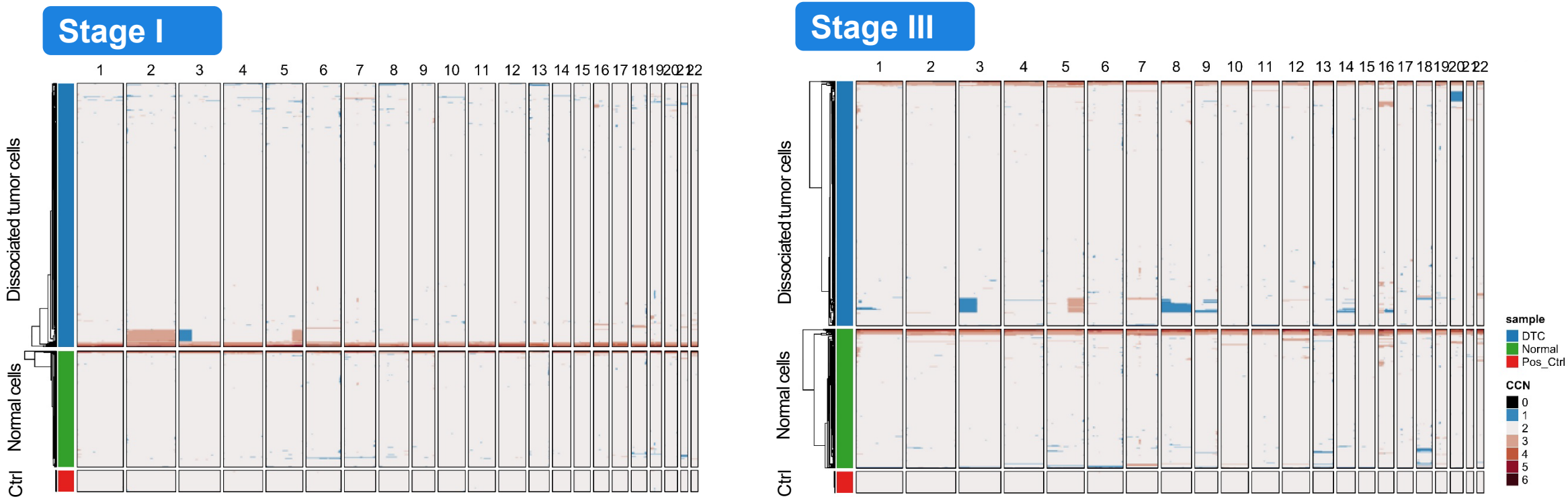


Dou, et al. “Single-nucleotide variant calling in single-cell sequencing data with Monopogen.” *Nat. Biotechnol.* (2023)

Characterize tumor heterogeneity with Shasta WGA

Find CNV events in small subclones amongst a heterogeneous tumor sample

- Dissociated cells from tumor tissues of Stage I (815 cells) and Stage III (858 cells) clear-cell renal-cell carcinoma (ccRCC) and adjacent normal tissue
- ~370,000 reads/cell, 2 x 75 bp



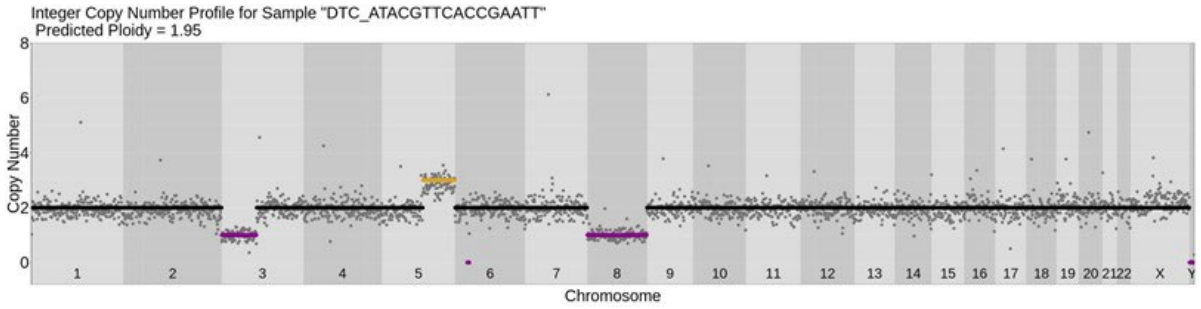
Detect crucial CNV events associated with tumor progression or better prognosis

−3p: a cytogenic hallmark of ccRCC, encompassing four commonly mutated genes: *VHL*, *PBRM1*, *BAP1*, and *SETD2* (Creighton et al. 2013).

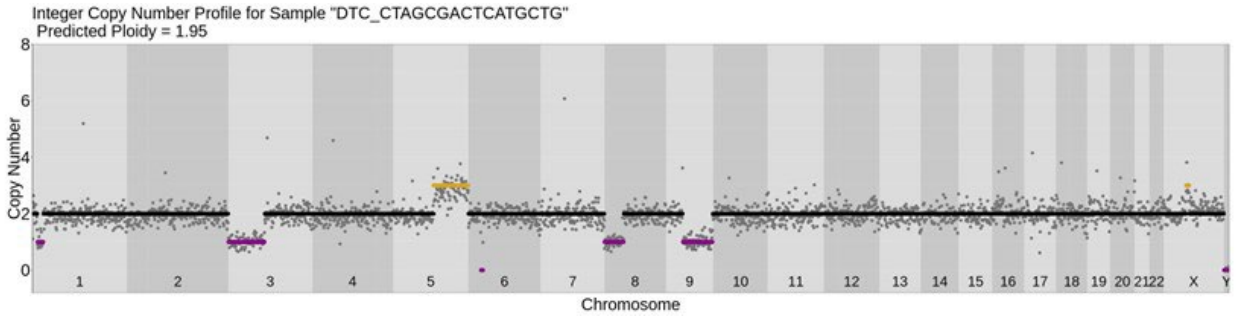
+5q: associated with better patient survival (Creighton et al. 2013).

Partial or complete loss of chr. 8: associated with *TCEB1* mutations (Sato et al. 2013).

Stage III tumor cell 1



Stage III tumor cell 2

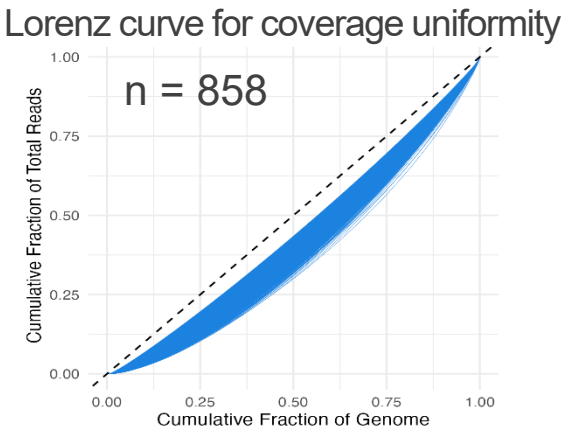


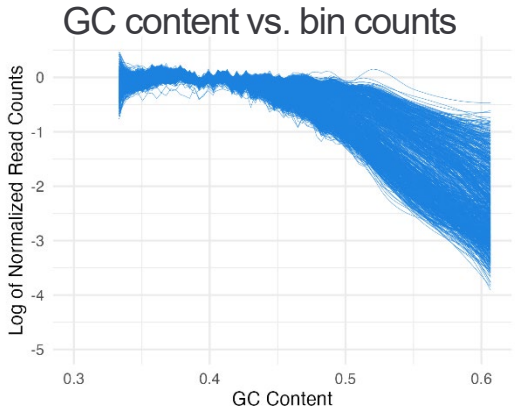
Creighton et al. "Comprehensive molecular characterization of clear cell renal cell carcinoma." *Nature* **499**, 43–49 (2013).

Sato et al. “Integrated molecular analysis of clear-cell renal cell carcinoma.” *Nature Genetics* **45**, 860–867 (2013).

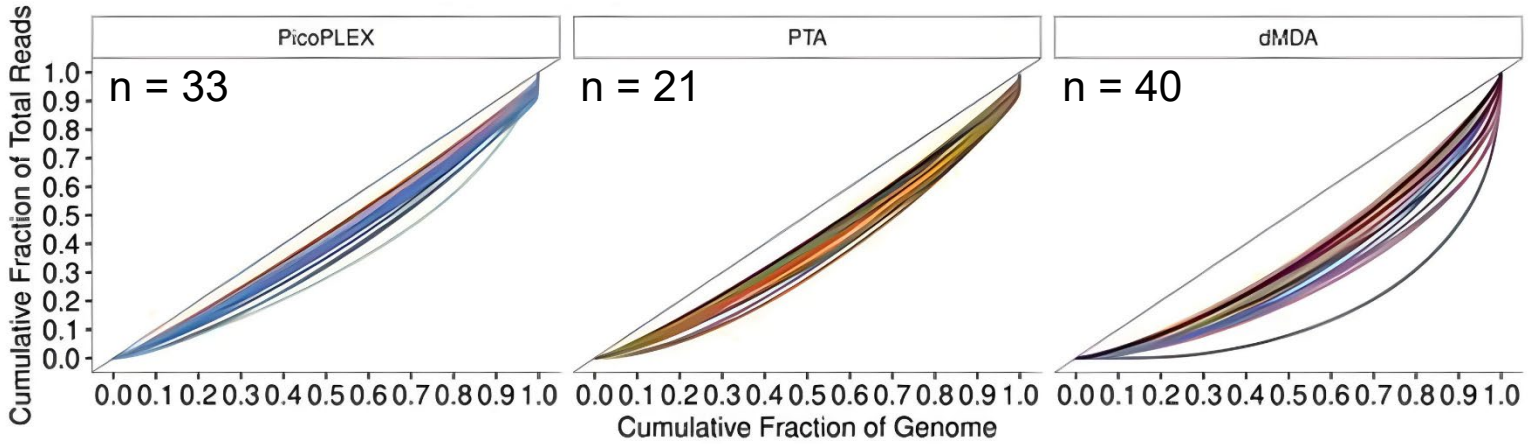
Compare performance with other WGA chemistries

Shasta WGA kit





PicoPLEX WGA kit | **Competitor R kit** | **Competitor X kit**



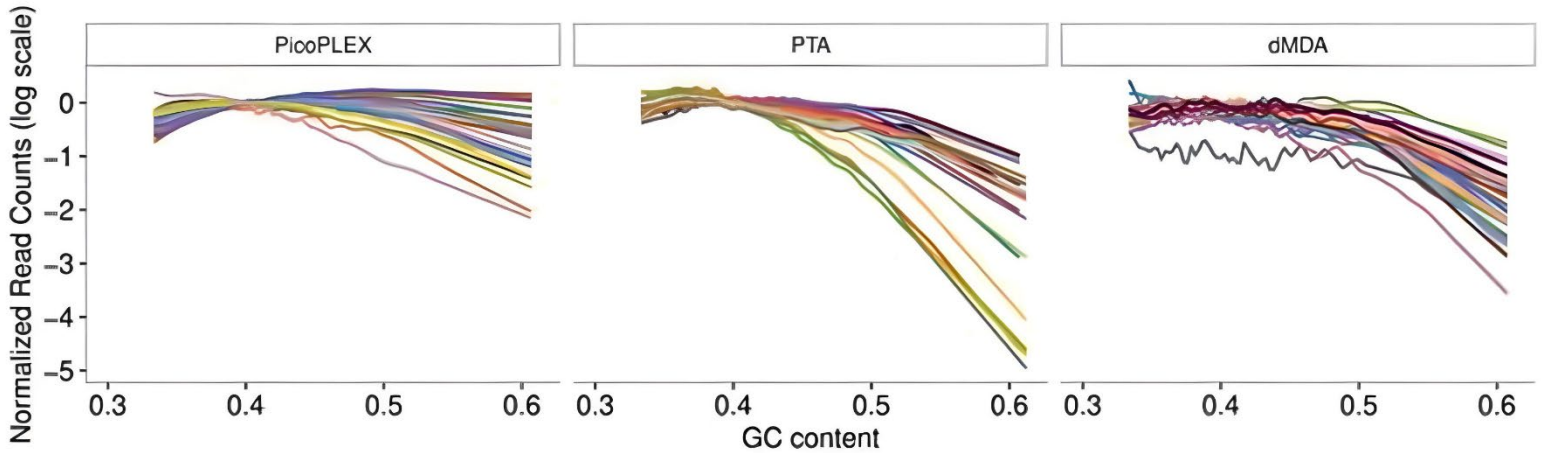
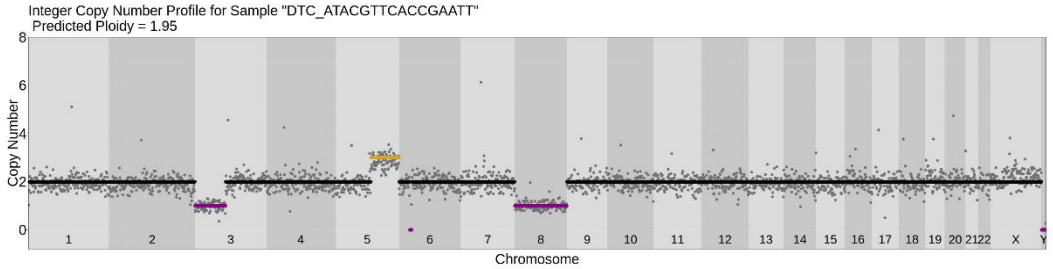


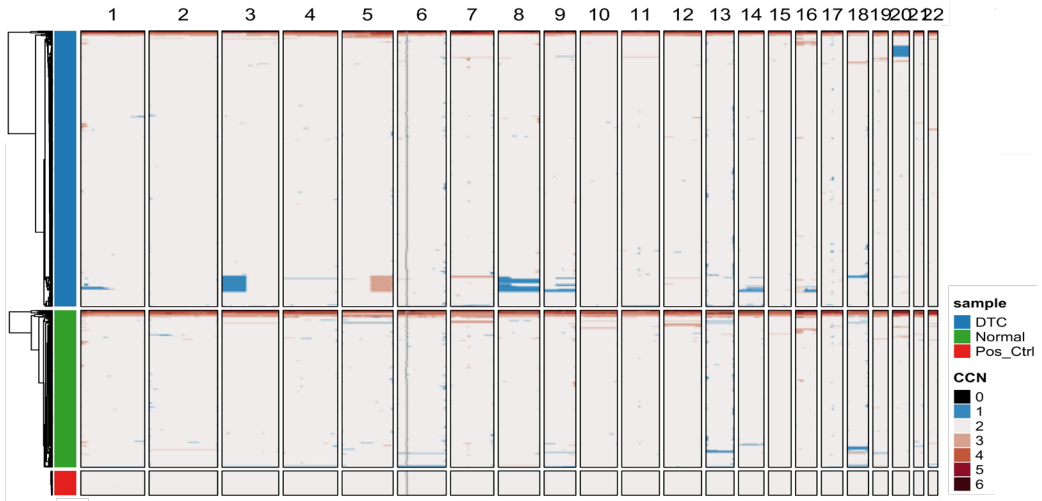
Figure adapted from "Single-cell somatic copy number variants in brain using different amplification methods and reference genomes" (Kalef-Ezra et al. 2023, *bioRxiv*) under a CC BY 4.0 license.

Summary: first commercial solution for scaled WGA

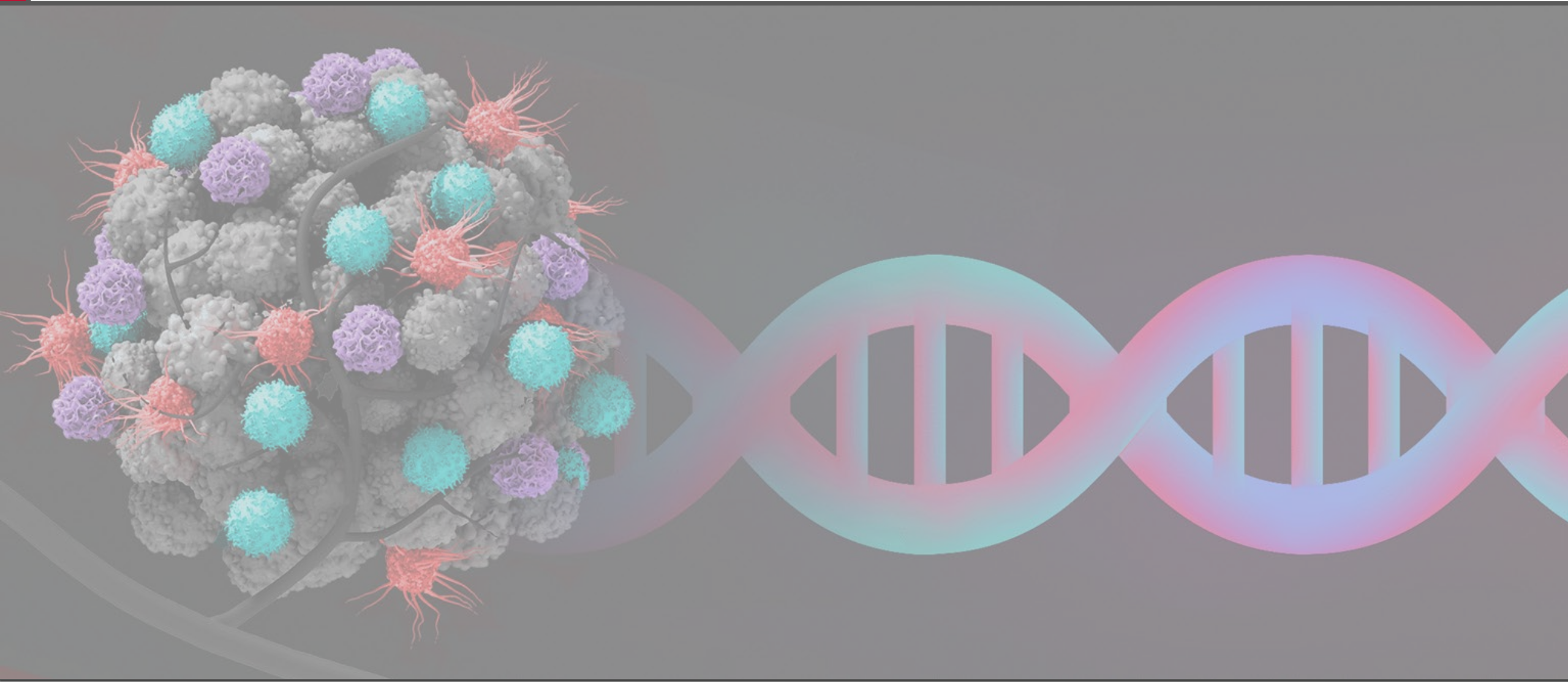
Shasta Whole-Genome Amplification

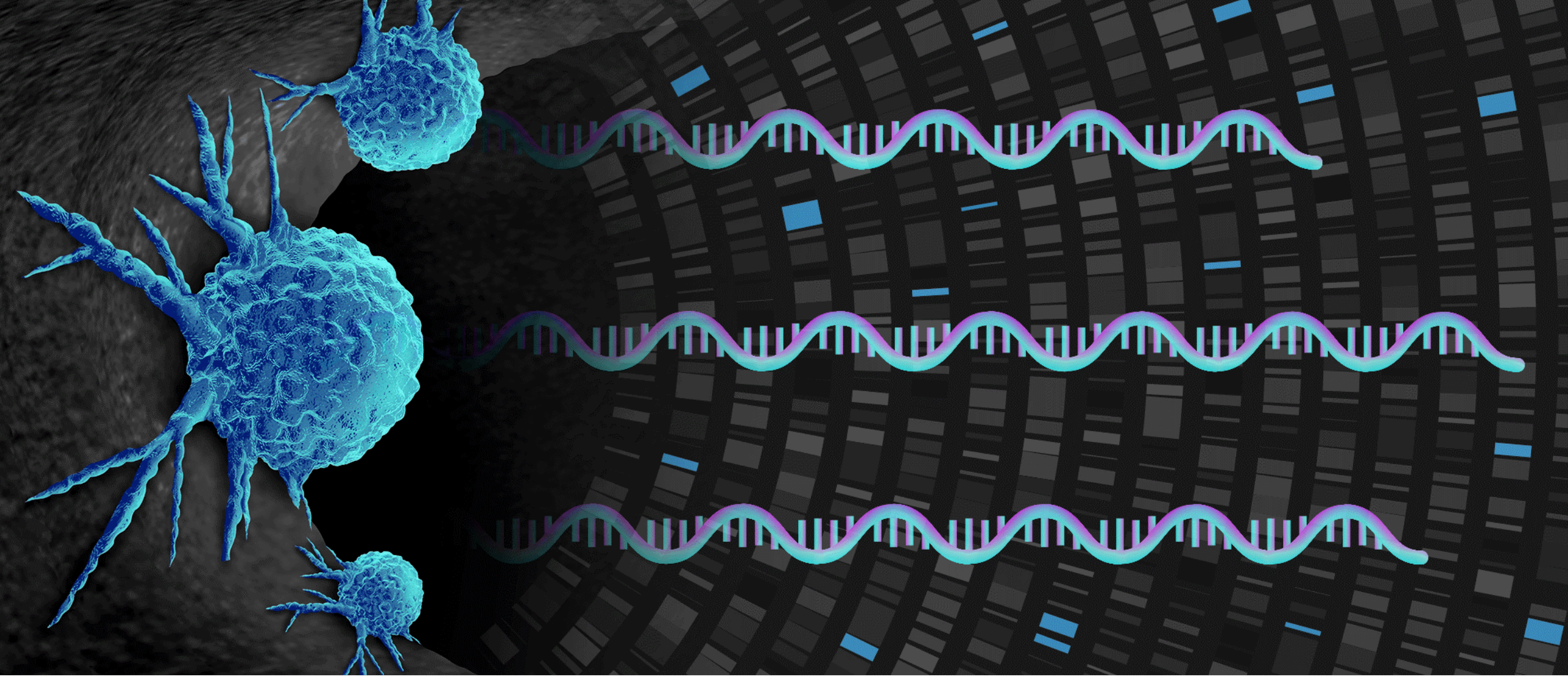
- Up to 1,500 single-cell WGA libraries
- CNV/SNV at a lower sequencing cost
- Fully automated workflow
- High uniquely mapped reads, low duplicates, and good coverage uniformity
- End-to-end solution, including free Cogent bioinformatics tools





Questions



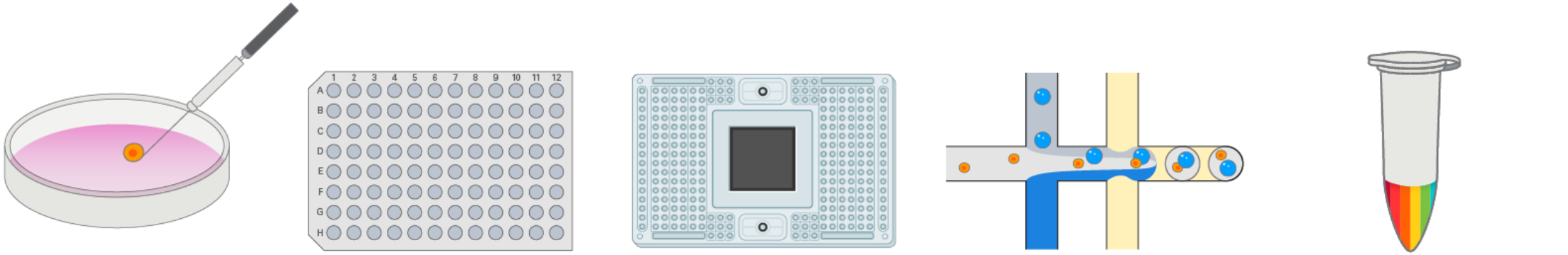


Shasta Total RNA-Seq Kit

TaKaRa

Why do we need high-throughput scRNA-seq?

The advancement of scRNA-seq technology coupled with decreasing sequencing expenses offers an exceptional opportunity to explore the transcriptomes of millions of individual cells.



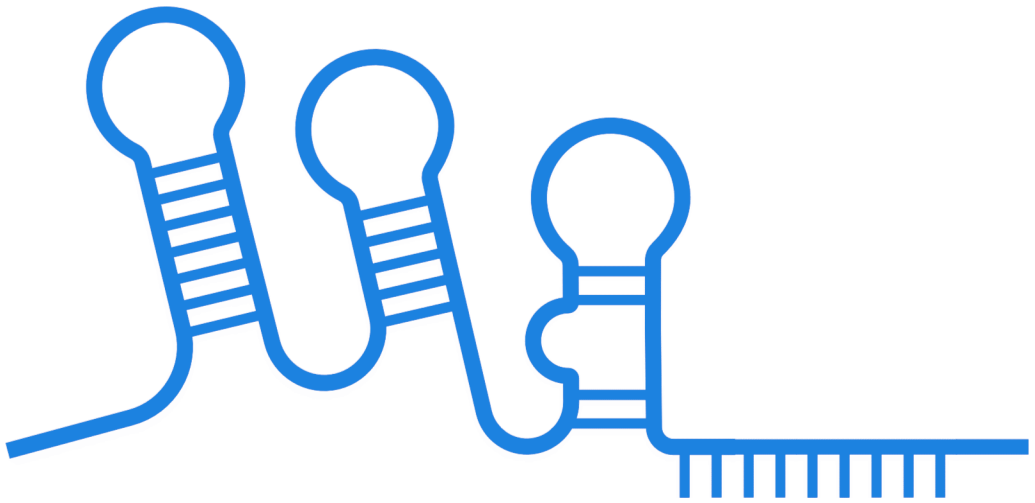
Micro-manipulation **Plate** **Microwell or nanowell** **Droplet** **Combinatorial indexing**

1–1,000s scale
Full coverage, high sensitivity, low throughput

10,000s–1,000,000s scale
End counting, low sensitivity, mRNA only

Created with BioRender.com.

Advancement of biological studies will benefit from incorporation of single-cell lncRNA analyses



Created with BioRender.com.

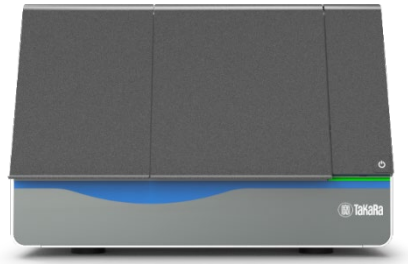
Long noncoding RNAs (lncRNAs)

- Newly defined transcripts that make up most of the transcriptome
- Tissue- and cell-specific regulators of multiple important cellular processes
- Functional analysis of the majority of lncRNAs is still insufficient

Automated, high-throughput solution for single-cell total RNA-seq

Shasta Total RNA-Seq Kit

- ✓ High throughput and sensitivity
  - *Process up to 100,000 cells/experiment*
  - *Process up to 96 samples/experiment*
  - *Obtain full gene body coverage*
- ✓ Automated two-day workflow
  - *Perform two rounds of barcoding*
  - *Reduce labor and human error*
  - *Lower reagent cost*
- ✓ Complete solution
  - *Use free Cogent bioinformatics tools*
  - *Perform both protein-coding and long noncoding RNA analysis*

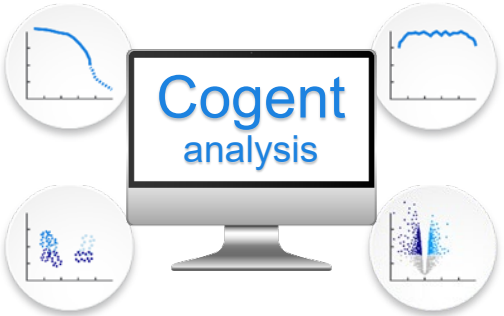


Unique indexing strategy

SMART-Seq for full gene body coverage

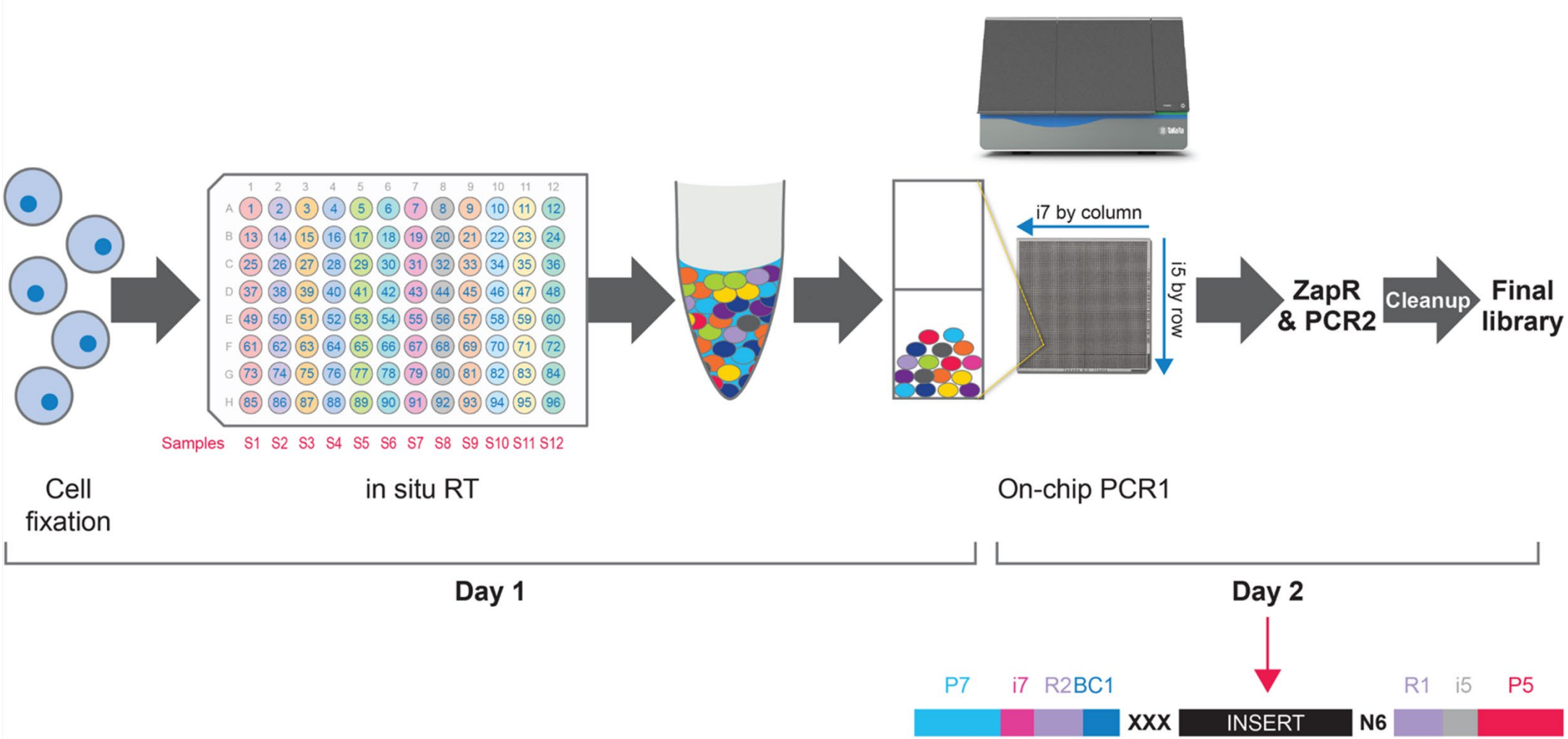
Shasta single-cell automation system

Cogent NGS pipeline



TaKaRa

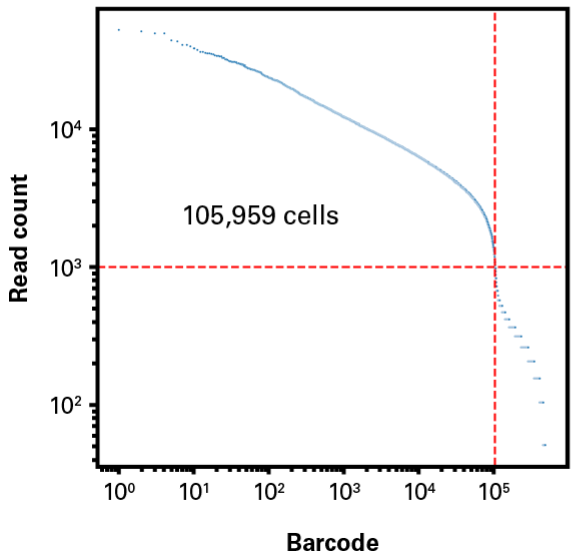
Shasta Total RNA-Seq workflow



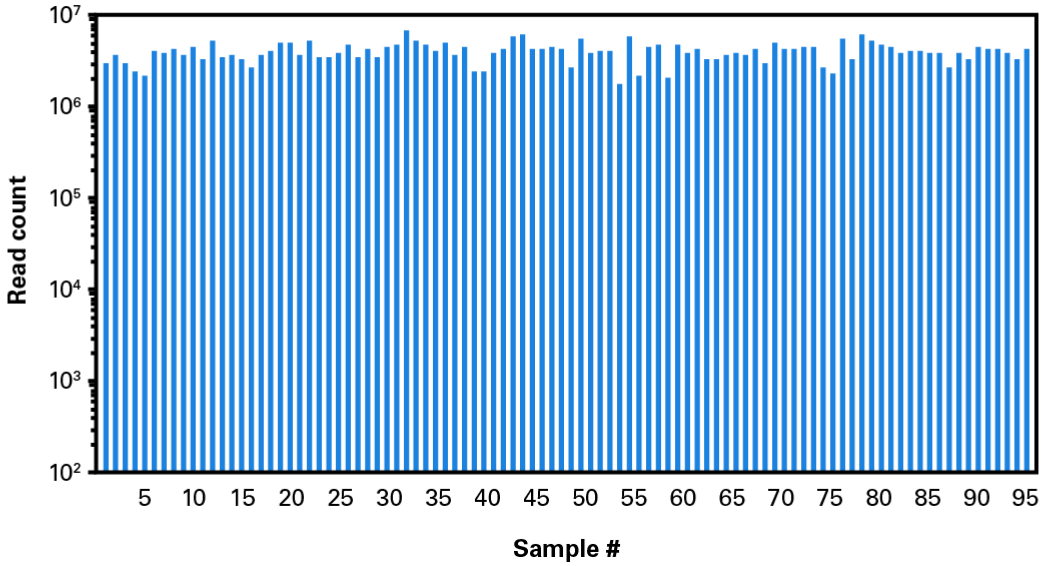
Enable high-throughput total scRNA-seq

- Cell throughput: ~100,000 cells per experiment
- Sample throughput: up to 96 samples per experiment
- Full 5'–3' gene body coverage

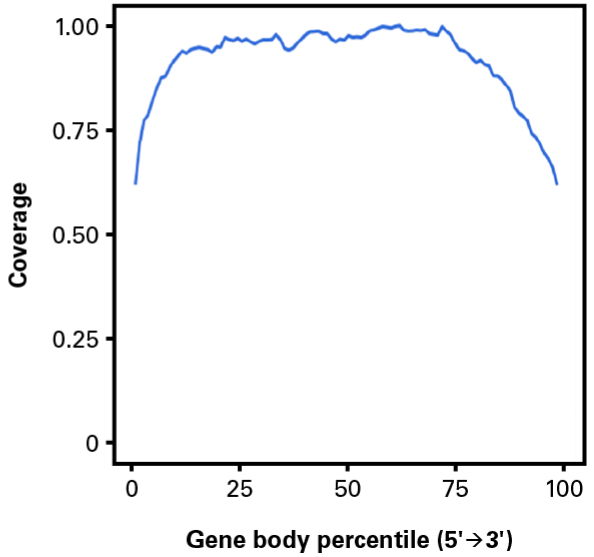
Cell throughput



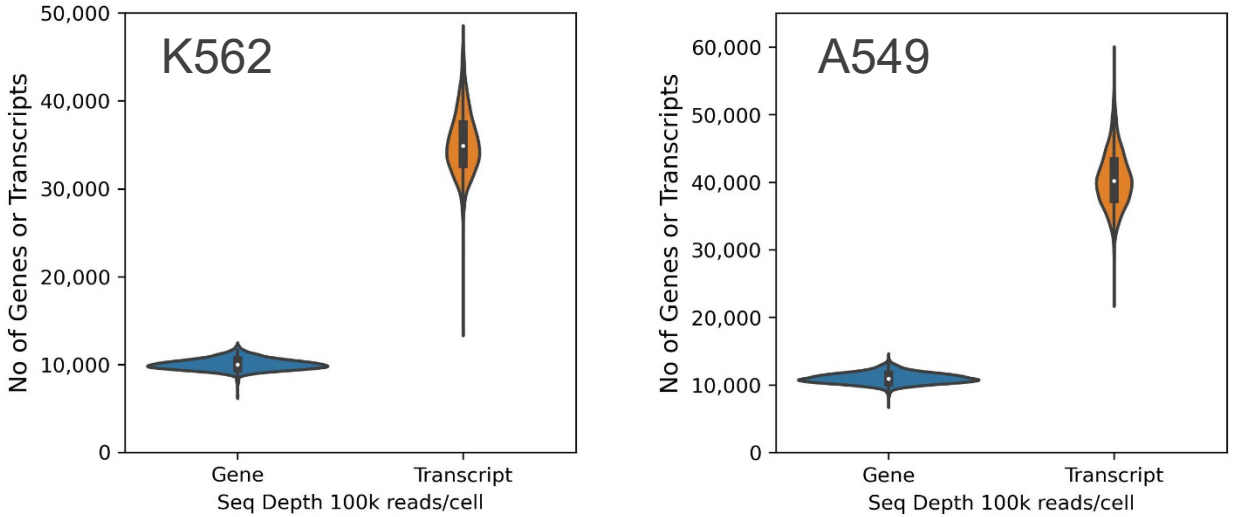
Sample throughput

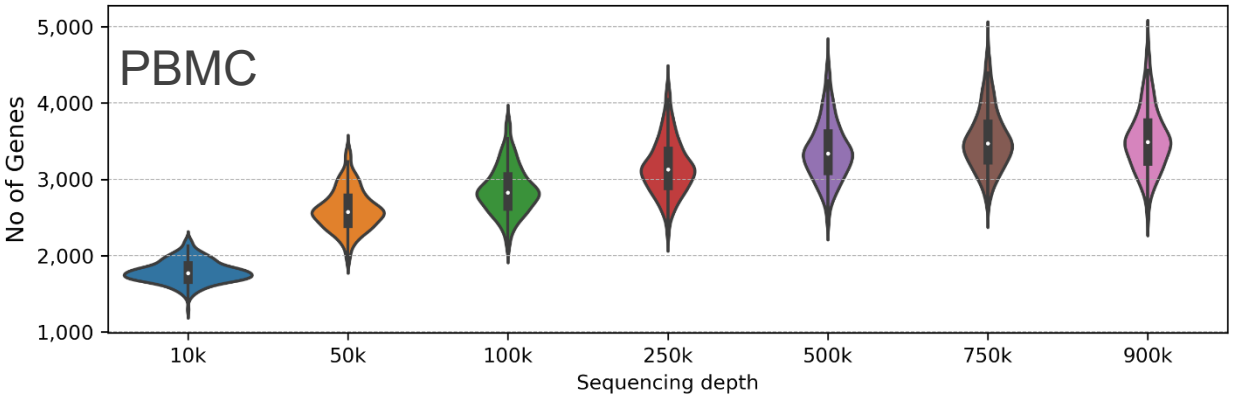


Gene body coverage

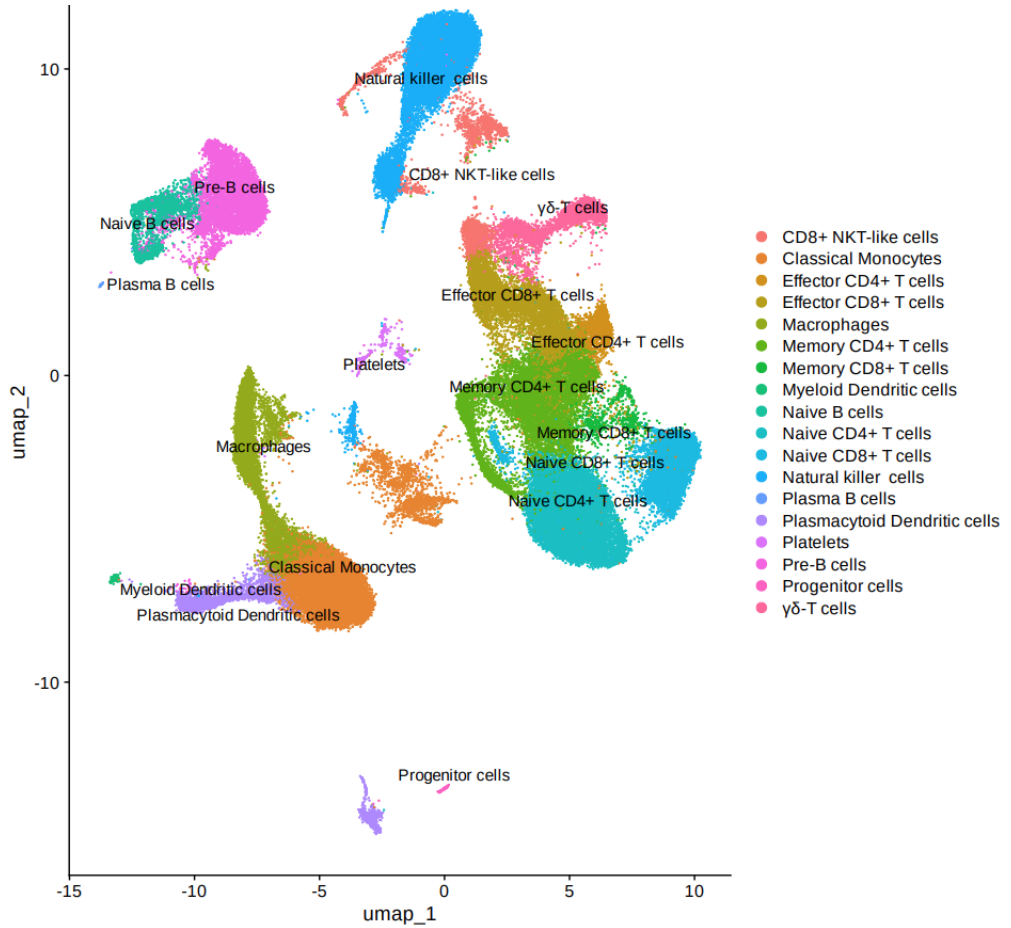


Achieve high sensitivity across various cell types and detect more genes with greater sequencing depth





80,000 PBMCs



Case study: discovering biomarkers regulated by p53 and epitherapy treatment

Four cell samples from Dr. Ting Wang's lab at Washington University in St. Louis School of Medicine

WT

P53-WT

WT_T

P53-WT + epitherapy treatment

KO

P53-KO

KO_T

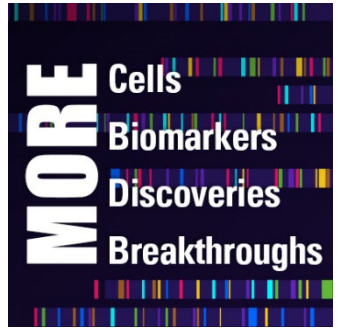
P53-KO + epitherapy treatment

Shasta Total RNA-Seq Kit

Cogent NGS Analysis Pipeline

- ✓ Standard analysis
- ✓ lncRNA analysis

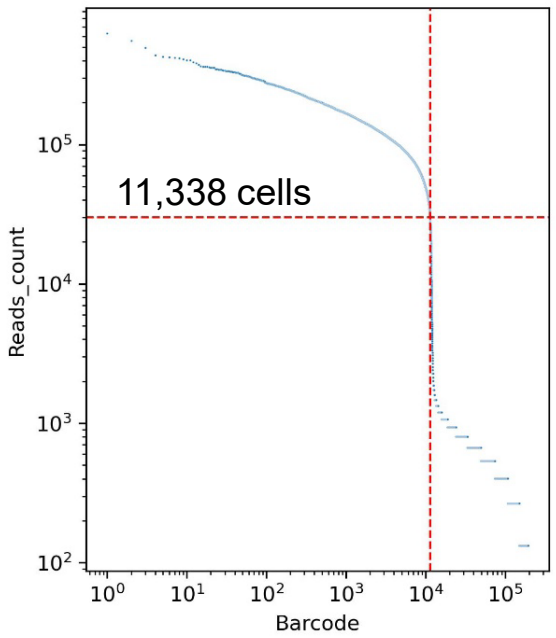
Biomarker discovery

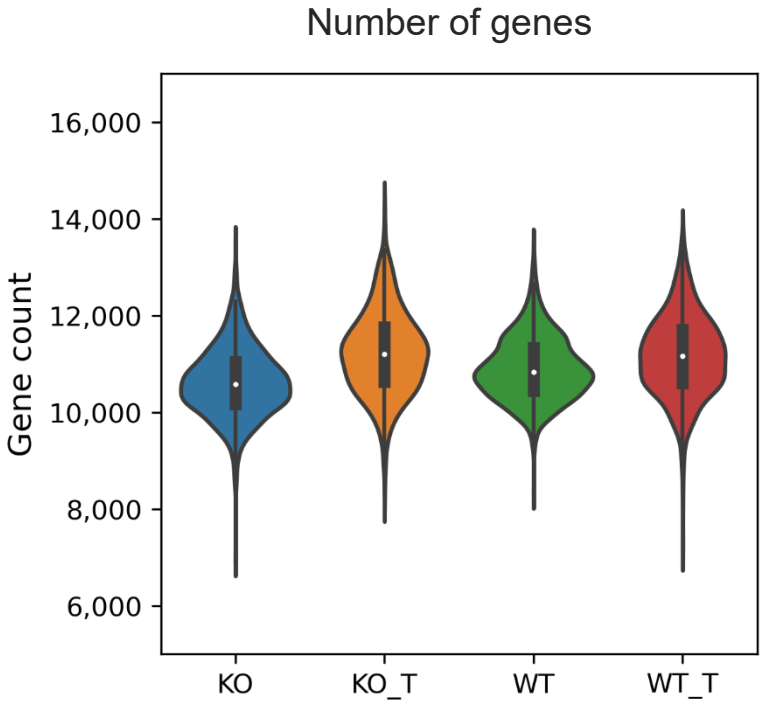


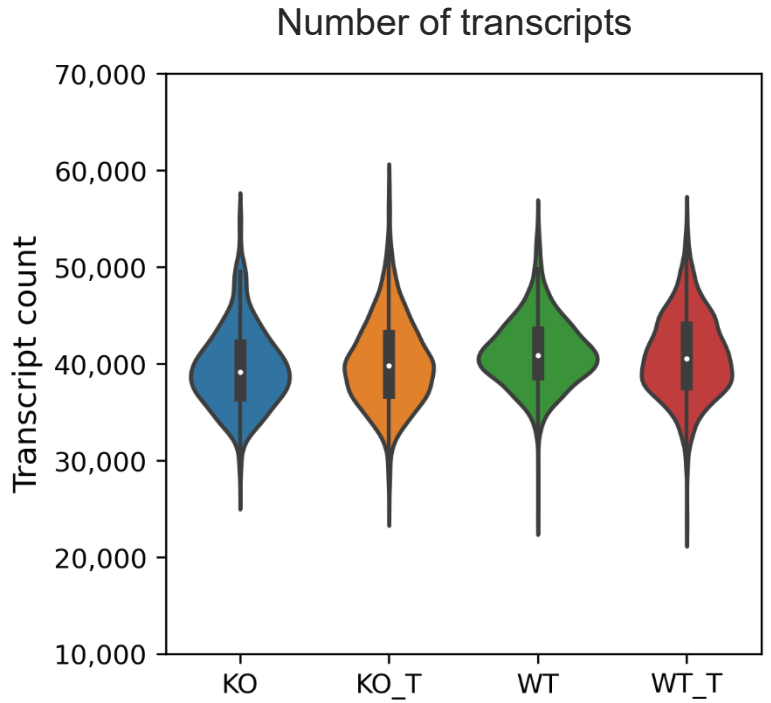
Created with BioRender.com.

Achieved outstanding sensitivity for both genes and transcripts

- ✓ Analyzed *~11,000 cells*
- ✓ Detected *~10,000 genes and ~40,000 transcripts per sample at a sequencing depth of 100,000 reads/cell*

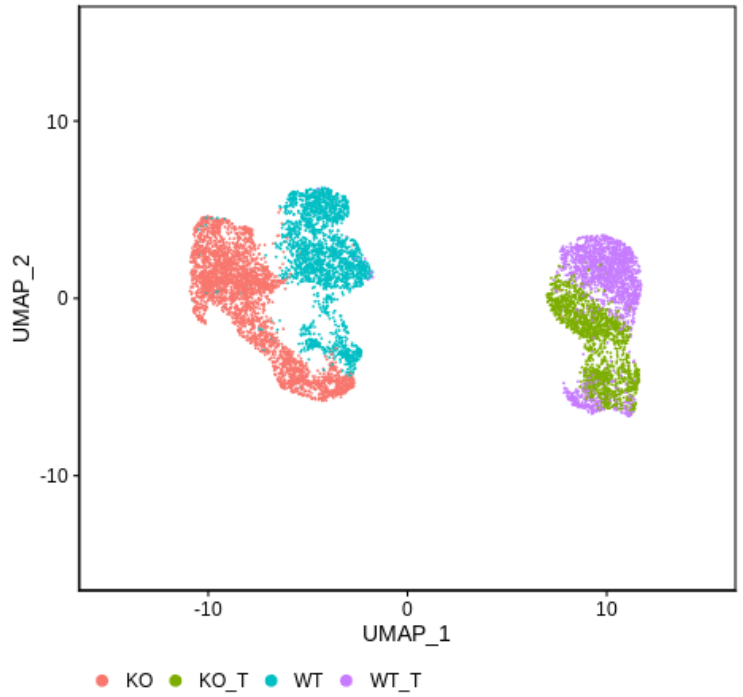


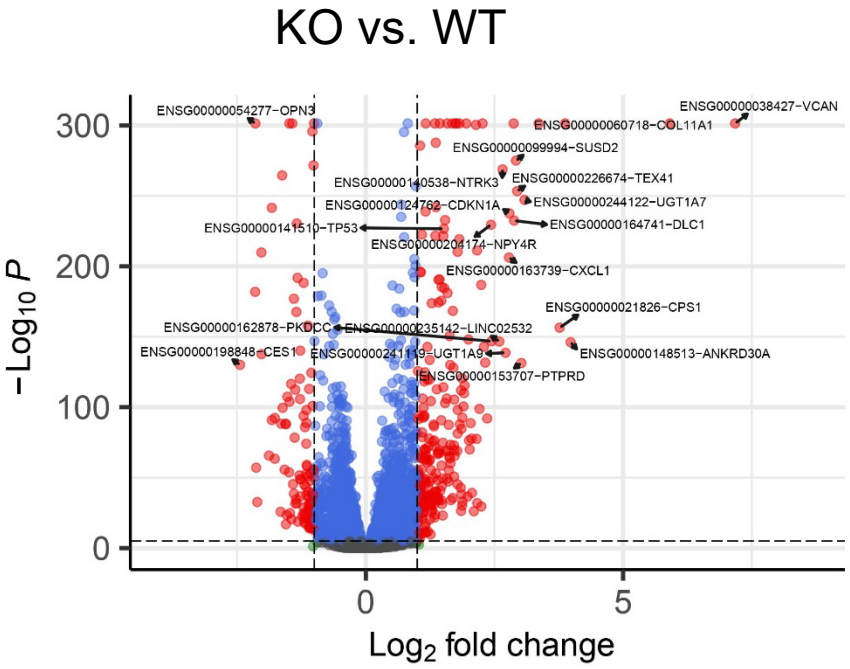




Biomarkers discovered using the standard Cogent pipeline

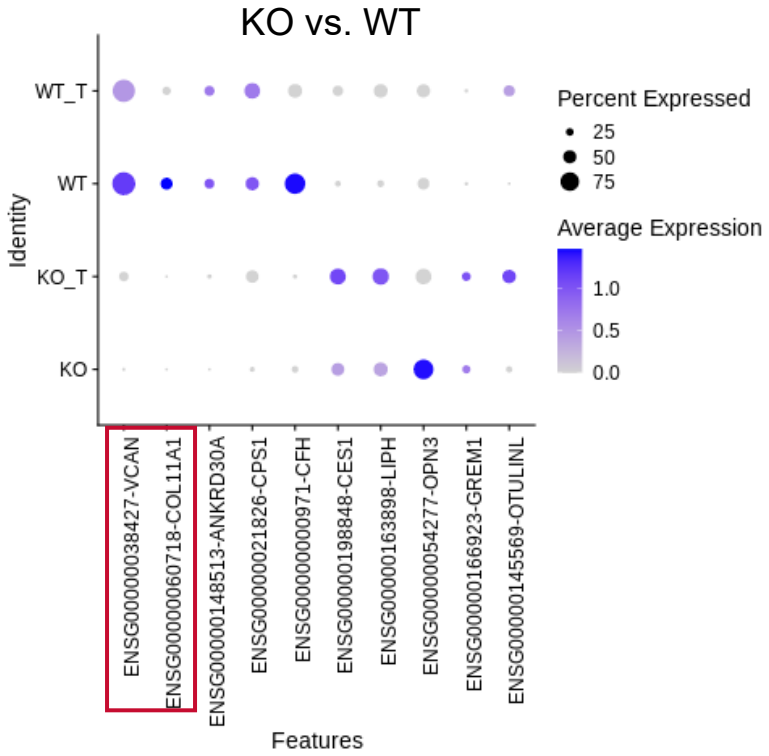
- UMAP separated the four samples using the standard (protein-coding) Cogent NGS Analysis Pipeline
- Differentially expressed biomarker genes from the four samples were identified

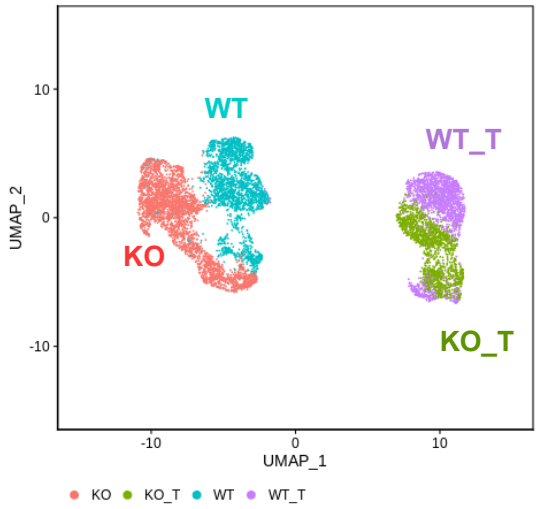


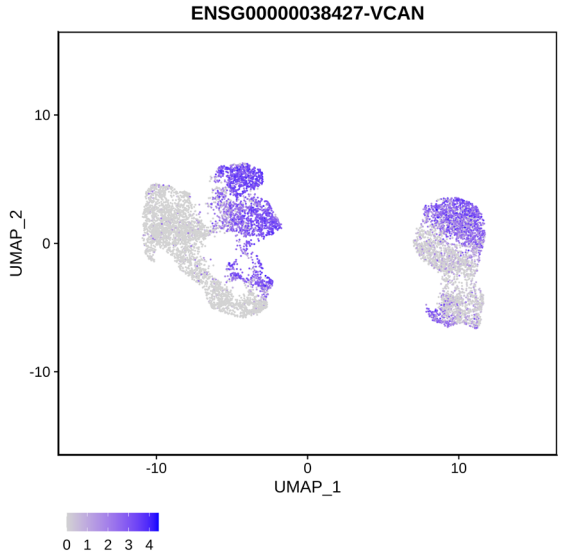


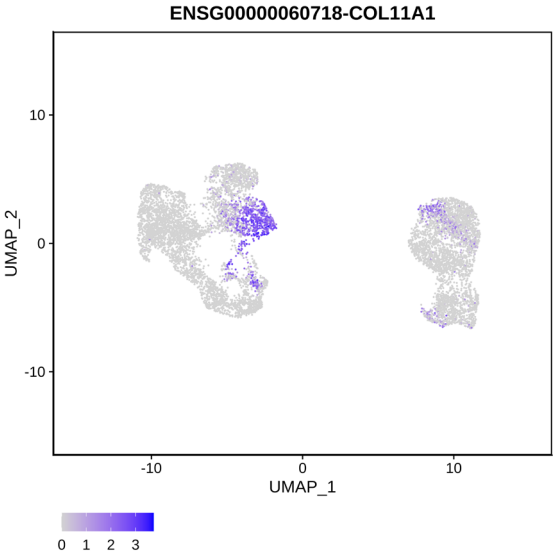
Examples of protein-coding genes identified with the Cogent pipeline

- Many of the identified differentially expressed genes are potential biomarkers regulated by p53 and epitherapy treatment
- *VCAN* and *COL11A1* have been associated with many cancers



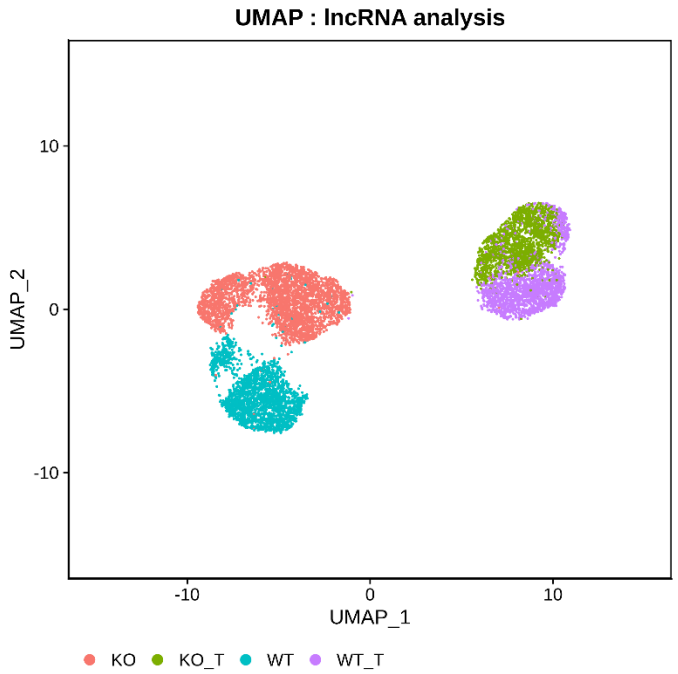


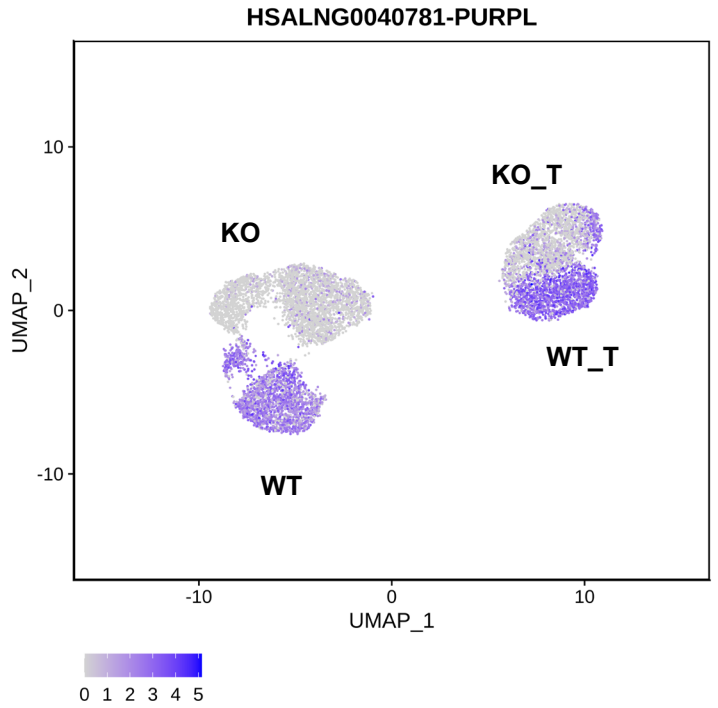


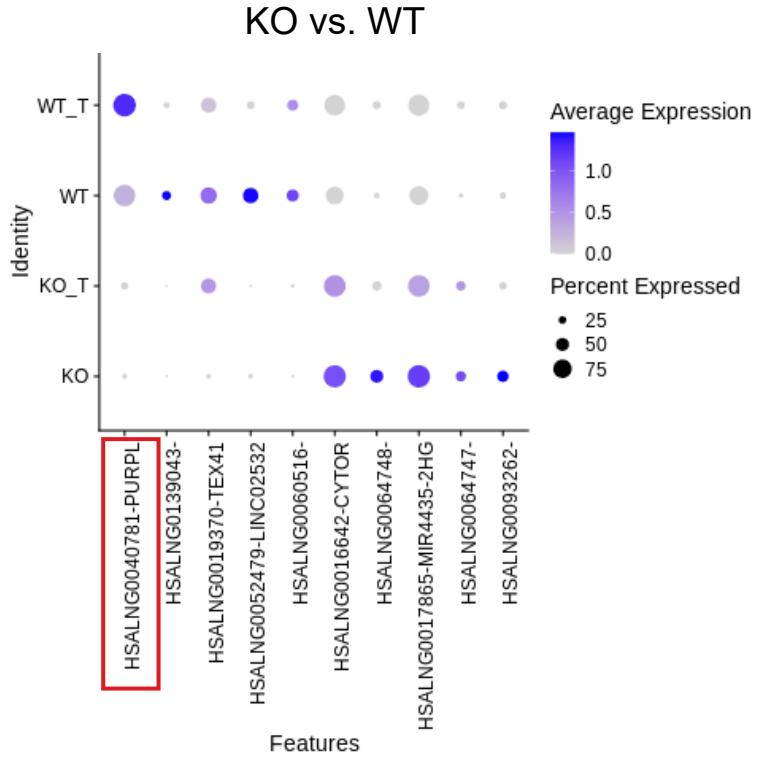


Example of noncoding gene identified with the Cogent lncRNA pipeline

p53 upregulated regulator of P53 levels (*PURPL*) has been shown to be a biomarker involved in many cancers (Li et al. 2017; Schmitt et al. 2016)





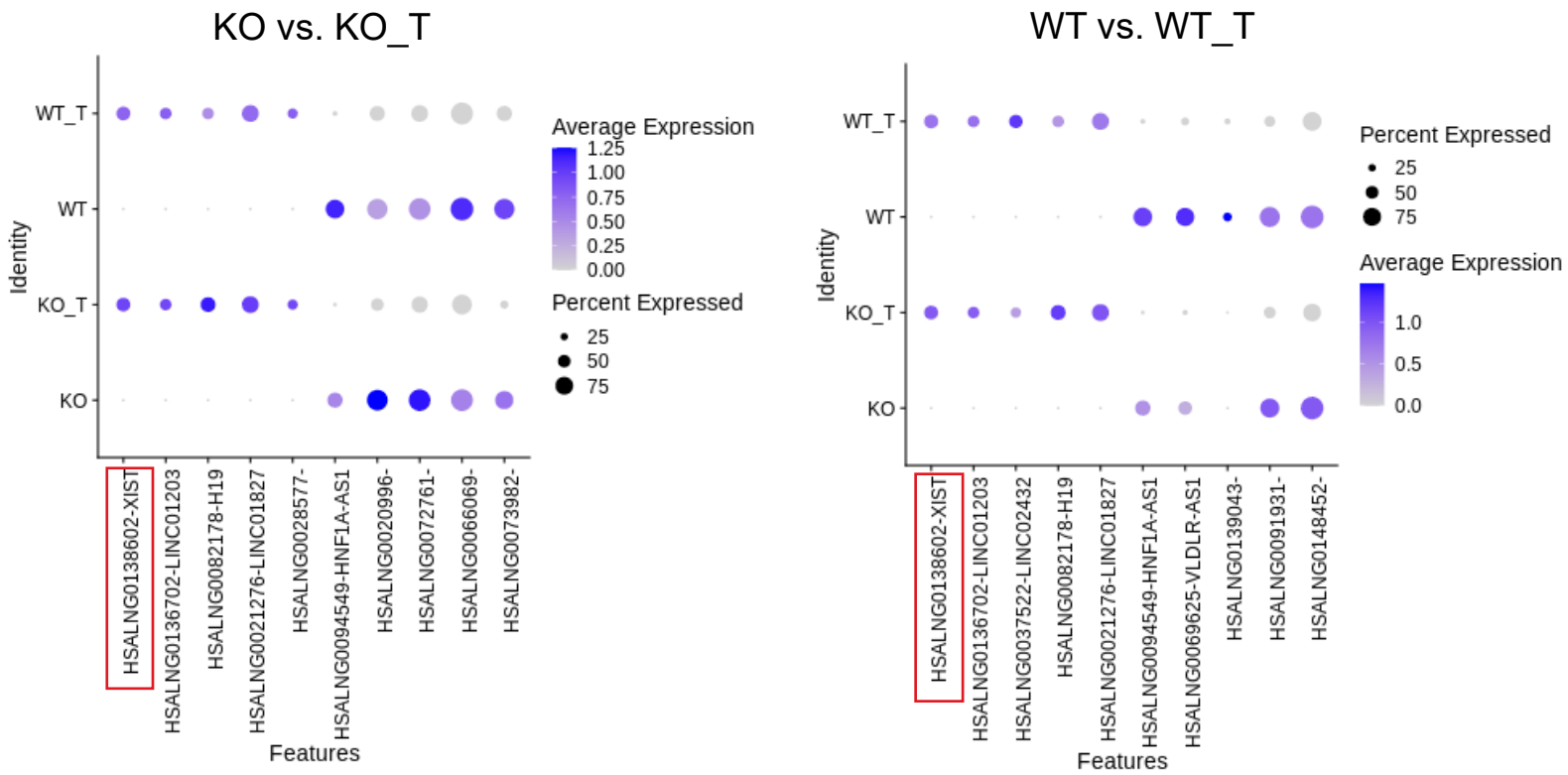


Li et al. “Long Noncoding RNA PURPL Suppresses Basal p53 Levels and Promotes Tumorigenicity in Colorectal Cancer.” *Cell Rep.* **20,** 2408–2423 (2017)

Schmitt et al. “An inducible long noncoding RNA amplifies DNA damage signaling.” *Nat. Genet.* **48,** 1370–1376 (2016).

Example of noncoding gene identified with the Cogent lncRNA pipeline

XIST was identified as a biomarker and has been shown to be involved in many cancers (Han et al. 2022; Richart et al. 2022)



Han et al. “Pan-cancer analysis of LncRNA XIST and its potential mechanisms in human cancers.” *Heliyon* **8,** e10786 (2022)

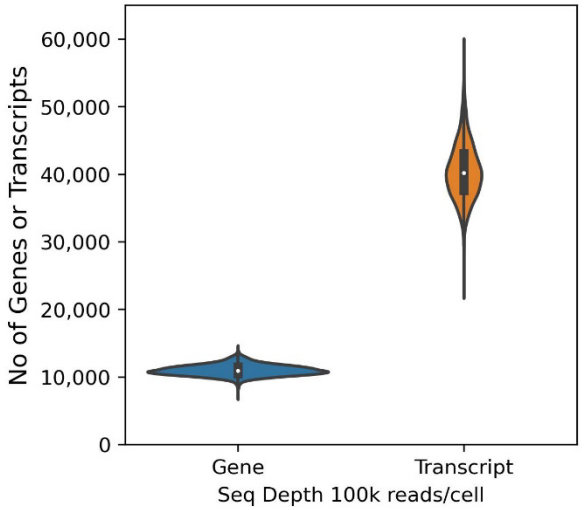
Richart et al. “XIST loss impairs mammary stem cell differentiation and increases tumorigenicity through Mediator hyperactivation.” *Cell* **185,** 2164–2183 (2022).

TaKaRa

Summary: first commercial automated solution for scaled total RNA-seq

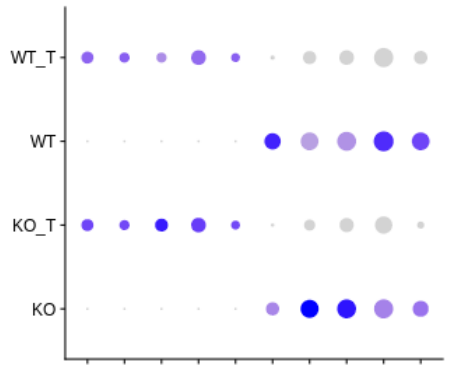
High throughput

High sensitivity



Shasta automation

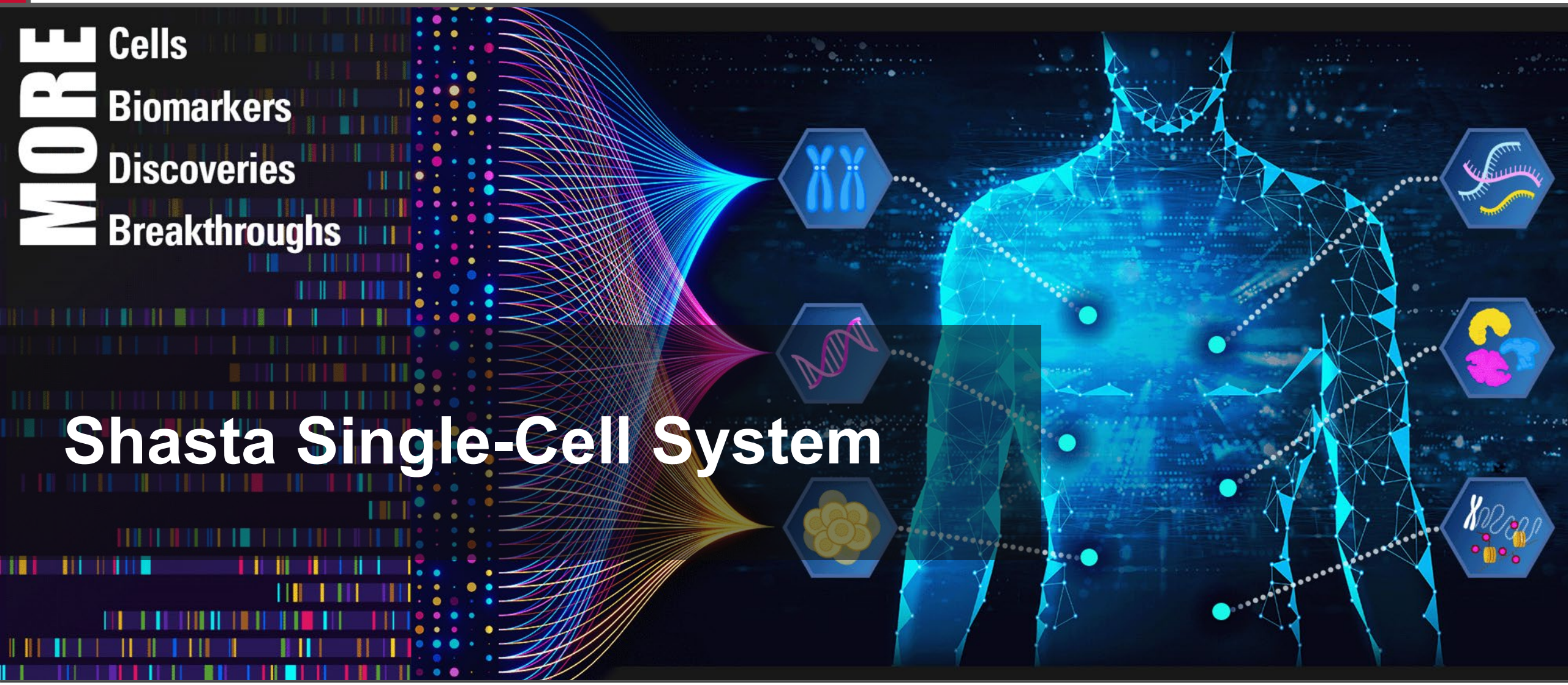
Cogent NGS pipeline



- ✓ Up to 100,000 cells
- ✓ Up to 96 samples
- ✓ Coding and noncoding information
- ✓ Full-length gene body coverage
- ✓ Reduced labor and human error
- ✓ Lower reagent cost
- ✓ Free-to-use analysis tools

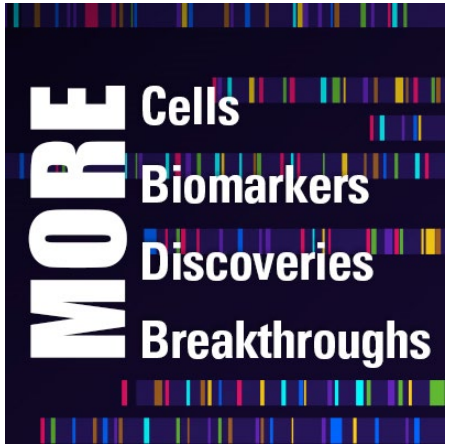
TaKaRa

Generating meaningful biological discoveries



TaKaRa

Generating meaningful biological discoveries



The Shasta single-cell solution

- Three novel, validated applications:

  Shasta Whole-Genome Amplification Kit
  Resolve tumor heterogeneity and track clonal evolution

  Shasta Total RNA-Seq Kit
  Go beyond mRNA to discover multiple RNA biotypes

  Shasta mRNA-Seq Kit
  Detect low-expressed biomarkers

- Free, easy-to-use Cogent bioinformatics software
- Intuitive user interface and easy maintenance

TaKaRa

that's
GOOD
science!®