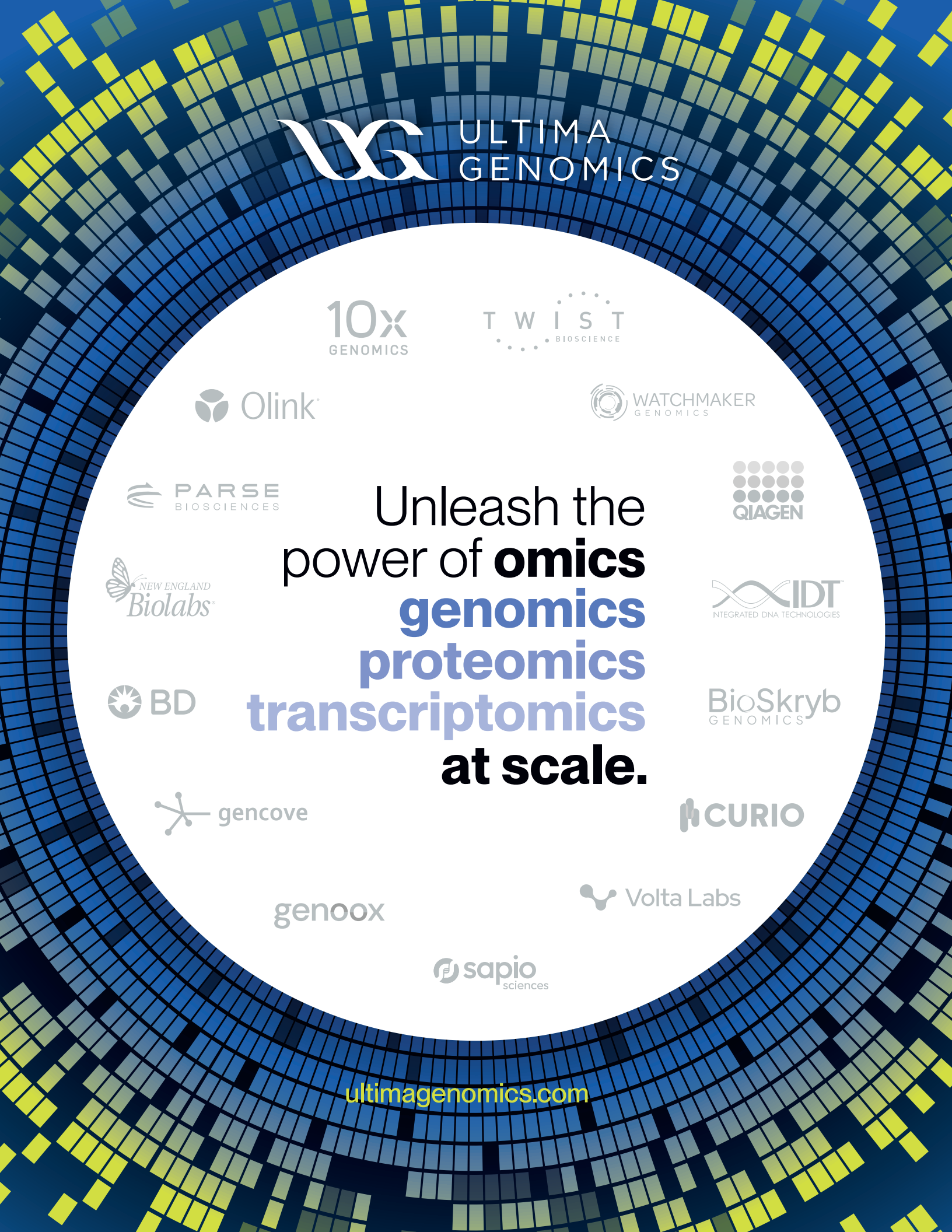
AGBT2026™
GENERAL MEETING

February 23–26, 2026 • Signia by Hilton Orlando Bonnet Creek • Orlando, FL


ULTIMA GENOMICS
10x GENOMICS
TWIST BIOSCIENCE
Olink
WATCHMAKER GENOMICS
PARSE BIOSCIENCES
QIAGEN
Unleash the power of omics
genomics
proteomics
transcriptomics
at scale.
NEW ENGLAND Biolabs
IDT INTEGRATED DNA TECHNOLOGIES
BD
BioSkryb GENOMICS
gencove
CURIO
genoox
Volta Labs
sapio sciences
ultimagenomics.com

AGBT 2026™
GENERAL MEETING

Signia by Hilton Orlando Bonnet Creek

FEBRUARY 23-26, 2026

Welcome to the 2026 Advances in Genome Biology and Technology General Meeting.

We are delighted to welcome you to Orlando for the 2026 AGBT General Meeting and to continue AGBT's longstanding role in advancing collaboration, innovation, and progress across the genomics community.

Over the coming days, the meeting will highlight breakthroughs in DNA sequencing technologies, emerging experimental and analytical approaches, and their applications across research, medicine, and industry. The General Meeting remains a distinctive forum for sharing early insights, fostering interdisciplinary exchange, and accelerating discovery, bringing together scientists, technologists, and leaders from around the world.

As part of its broader mission, AGBT convenes three high-impact meetings and networking events each year that span genomic technology, agriculture, and precision health. We invite you to join us at our upcoming meetings:

AGBT Agriculture
April 12-15, 2026
Phoenix, Arizona
Registration now open

AGBT Precision Health
September 14-16, 2026
San Diego, California

Thank you for being part of the AGBT community and for contributing to the shared pursuit of discovery and excellence. We hope you find this year's General Meeting both inspiring and impactful.

AGBT is dedicated to advancing research, promoting education, and expanding commerce in genome science and technology.

Wendy Vlieks
President, AGBT

AGBT 2026™

AGRICULTURE

APRIL 12-15, 2026

SHERATON GRAND AT WILD HORSE PASS | PHOENIX, AZ

FOR MORE INFORMATION, INCLUDING REGISTRATION AND ABSTRACT SUBMISSION DETAILS, SCAN HERE:

www.agbt.org

Organizing Committee

We thank Co-Chairs Eric Green, Len Pennacchio, Beth Shapiro, and the Scientific Organizing Committee for dedicating countless hours to shaping a rigorous, engaging AGBT 2026 General Meeting.

Federica Di Palma
University of British Columbia

Kim Doheny
Johns Hopkins University

Stacey Gabriel
Broad Institute of MIT and Harvard

Eric Green
Former Director, National Human Genome Research Institute (NHGRI)

Elinor Karlsson
UMass Chan Medical School

John McPherson
University of California, Davis

Stephen Montgomery
Stanford University

Len Pennacchio
Lawrence Berkeley National Laboratory; Joint Genome Institute

Michael Schatz
Johns Hopkins University

Beth Shapiro
University of California, Santa Cruz

Emma Teeling
University College Dublin

Michael Zody
New York Genome Center

Sponsors

We gratefully acknowledge the generous support provided by the following companies

GOLD

illumina®

SILVER

ULTIMA GENOMICS

BRUKER

Roche

BRONZE

New England Biolabs

Complete Genomics

IDT Integrated DNA Technologies

Watchmaker Genomics

Oxford Nanopore Technologies

10x Genomics

Element Biosciences

BD

Codetta Bio

CONTRIBUTING

biomatter

canal biosciences

cellanome

CHRYSALIS

ELLIS BIO

GeneDx

Genome Research

LatchBio

MilliporeSigma

Nabsys

Novogene

PacBio

PARSE BIOSCIENCES

QIAGEN

revvity

SINGULAR GENOMICS

Stellaromics

TWIST BIOSCIENCE

vizgen

Volta Labs

General Information

Hospitality Desk Hours

- Sunday, Feb. 22: Noon-8 p.m.
- Monday, Feb. 23: 8 a.m.-8 p.m.
- Tuesday, Feb. 24: 7:30 a.m.-9:30 p.m.
- Wednesday, Feb. 25: 7:30 a.m.-8 p.m.
- Thursday, Feb. 26: 7:30 a.m.-6 p.m.

Name Badges

Badges are required for conference security and networking. All attendees must:

- Wear badges at all times, visible on the provided lanyard, with names facing outward.
- Keep badges with you throughout the meeting.
- Present badges for entry to all conference entrances and events.

Important: Badges are nontransferable. Proper badge display helps ensure attendee safety and facilitates meaningful connections.

If you forget your badge, you will be asked to return to your hotel room to retrieve it.

Transportation

Airport Shuttle Service

AGBT is pleased to offer complimentary shuttle service between Orlando International Airport (MCO) and the Signia by Hilton Orlando Bonnet Creek / Waldorf Astoria Orlando (21 miles; approximately 30 minutes).

Shuttle service is available only to registered attendees. Unfortunately, we cannot accommodate additional family members.

Arrival Shuttles (every 30 minutes)

Pick-up location: Baggage claim area

- Sunday, Feb. 22: 1-8 p.m.
- Monday, Feb. 23: 7 a.m.-4 p.m.

Departure Shuttles (every 30 minutes)

Pick-up locations: Signia by Hilton Group Transportation entrance; Waldorf Astoria main entrance
Friday, Feb. 27: 4 a.m.-2 p.m.

General Information

Arrival Instructions

After landing at MCO, take the monorail to the main terminal and proceed to baggage claim for your airline (Terminal A or B), even if you did not check bags. E2 staff wearing lime green polo shirts will meet you at the bottom of the escalator and escort you to the shuttle.

If you arrive in Terminal C, go to the baggage claim carousel and look for AGBT signage.

For assistance onsite, call or text **Kristi Mazon** (407) 632-8246 or **Christina Geraghty** (407) 912-5116.

Alternative Transportation Options

- **Taxi:** Approximately $50-85 each way
- **Rideshare (Uber/Lyft):** Approximately $35-70 each way
- **Rental car:** Available from multiple agencies at the airport
- **Orlando International Airport:** Ground Transfers

Social Events

- **Welcome Reception:** Monday, Feb. 23, 7-10 p.m., Promenade and Signature Island
- **Women's Networking Event:** Tuesday, Feb. 24, 5:30-7:30 p.m., Waterside Lanai
- **Sponsor's Social:** Tuesday, Feb. 24, 9:45 p.m., Sponsor Promenade, Signia Executive Suites and Presidential Suites
- **AGBT Dinner:** Wednesday, Feb. 25, 6:10-7:25 p.m., Waterside Ballroom
- **Farewell Dinner:** Thursday, Feb. 26, 6-11 p.m., 19th Green

Conference Technology

Official AGBT Mobile App

The AGBT mobile app includes the agenda, speaker details, maps, abstracts and other essential resources.

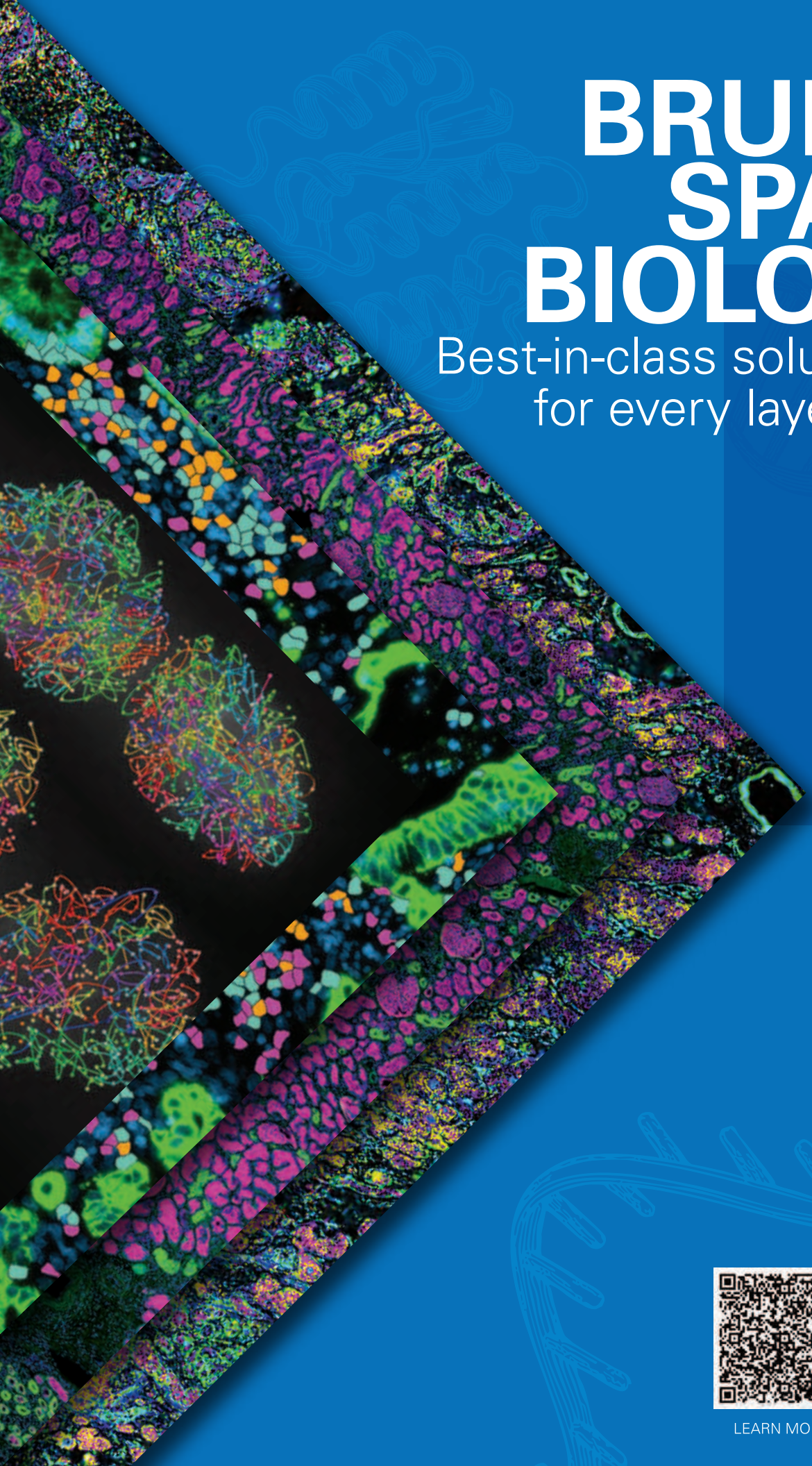
Social Media Engagement

Share your experience throughout the meeting:

- **Primary Conference Hashtag:** #AGBTGM26

Tech Support

- **AGBT Hospitality Desk:** Bonnet Creek North Foyer
- **Wi-Fi Network:** AGBTGM26
- **Wi-Fi Password:** AGBTGM26



BRUKER SPATIAL BIOLOGY

Best-in-class solutions for every layer of biology

LEARN MORE!

BRUKER

Agenda

Sunday, February 22

12:00 p.m. – 8:00 p.m.	**Meeting Registration / Hospitality Desk** Bonnet Creek North Foyer
1:00 p.m. – 8:00 p.m.	**Transfers from Orlando International Airport (MCO)**

Monday, February 23

7:00 a.m. – 4:00 p m.	**Transfers from Orlando International Airport (MCO)**
8:00 a.m. – 8:00 p.m.	**Meeting Registration / Hospitality Desk** Bonnet Creek North Foyer

Scientific Program

5:00 p.m. – 5:15 p.m.	**Opening Remarks** Len Pennacchio, Meeting Co-Chair, Lawrence Berkeley National Laboratory, The Genome Partnership Board Member **AGBT Distinguished Service Award Presentation** Nathan Lakey, Orion Genomics, The Genome Partnership Board Member Leisa Zigman, President Emerita and Board Member, The Genome Partnership **Recipient:** Eric Green, Former Director, National Human Genome Research Institute; Illumina
Plenary Session:	**Genomics** (Len Pennacchio, Meeting Co-Chair, Lawrence Berkeley National Laboratory) Floridian Ballroom
5:15 p.m. – 5:45 p.m.	**William Pao, Revelio Therapeutics, Inc.** “Decoding nature: how curiosity-driven research leads to breakthrough medicines”
5:45 p.m. – 6:15 p.m.	**Scott Edwards, Harvard University** “Finding the genes behind flightlessness in birds”
6:15 p.m. – 6:45 p.m.	**Emma Teeling, University College Dublin** “Bats: New models of extended healthspan and disease resistance”
7:00 p.m. – 10:00 p.m.	**Welcome Dinner – *Moonlight on the Creek*** **Sponsored by Twist Bioscience** Promenade and Signature Island

Agenda

Tuesday, February 24

7:30 a.m. – 9:00 a.m.	**Breakfast Sponsored by Ellis Bio** Waterside Lanai and Patio
7:30 a.m. – 9:30 p.m.	**Meeting Registration / Hospitality Desk** Bonnet Creek North Foyer
Plenary Session:	**AI in Genomics** (Mike Schatz, Johns Hopkins University, Session Chair) Floridian Ballroom
9:00 a.m. – 9:30 a.m.	**Barbara Engelhardt, Stanford University** “Machine learning methods for live-cell imaging data to design immunotherapies”
9:30 a.m. – 10:00 a.m.	**Jack Gilbert, University of California, San Diego** “The human microbiome in precision medicine”
10:00 a.m. – 10:30 a.m.	**Coffee Break Sponsored by Novogene** Sponsor Promenade
10:30 a.m. – 11:00 a.m.	**Marylyn Ritchie, Medical University of South Carolina** "AI-powered translational bioinformatics: integrating health records and multiomics”
11:00 a.m. – 11:30 a.m.	**Hilary Finucane, Harvard Medical School, Broad Institute of MIT and Harvard** “Using predicted protein structures to interpret disease genetics”
11:30 a.m. – 11:50 a.m.	**Poster Flash Talks I (2 minutes each)**
12:00 p.m. – 1:30 p.m.	**Gold Sponsor – Illumina – Workshop and Lunch** Jacob Thaysen, PhD, CEO, Steve Barnard, PhD, Chief Technology Officer, Head of Research & Product Development, and Cande Rogert, PhD, VP, Global Head of Advanced Sciences “Illumina insight ecosystem: the path from innovation to impact” Bonnet Creek Ballroom VII-XII
12:00 p.m. – 1:30 p.m.	**AGBT Lunch** Waterside Boulevard
1:30 p.m. – 3:30 p.m.	**Poster Session with Coffee and Dessert** Bonnet Creek Ballroom I-VI
Plenary Session:	**Biology** (Beth Shapiro, University of California, Santa Cruz; Meeting Co-Chair and Session Chair) Floridian Ballroom
3:30 p.m. – 4:00 p.m.	**Tom Gilbert, The Globe Institute, University of Copenhagen; NTNU University Museum** “Domestication hologenomics – how did our ancestors tame the wolf?”

Agenda

Tuesday, February 24

4:00 p.m. – 4:30 p.m.	**Hernán Morales, University of Copenhagen** "Irreversible loss? The genomic legacy of population collapse and the challenge of genetic diversity restoration"
4:30 p.m. – 5:00 p.m.	**Adam Phillippy, Center for Genomics and Data Science Research, National Human Genome Research Institute, National Institutes of Health** "Complete, telomere-to-telomere genomes as a foundation for predictive genomics"
5:00 p.m. – 5:20 p.m.	**Kiran Garimella, Broad Institute of MIT and Harvard** (Abstract-Selected Talk) "Scaling long-read genomics for population health: 14,000 diverse genomes from the All of Us Research Program"
5:30 p.m. – 7:30 p.m.	**Dinner on Your Own**
5:30 p.m. – 7:30 p.m.	**Women's Networking Event** Waterside Lanai
Concurrent Session:	**Abstract Selected Talks**
7:30 p.m. – 9:45 p.m.	**Cancer** (Federica Di-Palma, Genome British Columbia, Session Chair) Floridian Ballroom
7:30 p.m. – 7:45 p.m.	**Hanlee Ji, Stanford University** "Determining the function for cancer mutation of unknown significance with single-cell, genome-engineering screens"
7:45 p.m. – 8:00 p.m.	**Nicholas Navin, MD Anderson Cancer Center** "Reprogramming of the stromal and immune microenvironment in preinvasive breast cancer"
8:00 p.m. – 8:15 p.m.	**Courtney Hall, Johns Hopkins University** "Resolving tumor heterogeneity with targeted single-cell nanopore sequencing"
8:15 p.m. – 8:30 p.m.	**Jiwoon Park, Weill Cornell Medicine** "SPOT-Met: learning metastatic tropism from 1,000 spatially profiled colorectal tumors"
8:30 p.m. – 8:45 p.m.	**Peter Park, Harvard University Medical School** "Seeing where GRCh38/CHM13 miss: personalized diploid references uncover hidden somatic structural variation in cancer"
8:45 p.m. – 9:00 p.m.	**Rendong Yang, Northwestern University** "3′UTR exitron splicing reshapes the post-transcriptional regulatory landscape in cancer"
9:00 p.m. – 9:15 p.m.	**Mark Cowley, Children's Cancer Institute** "Advancing precision diagnosis in high-risk paediatric oncology through integrated whole-genome, RNA and methylome analysis"

Agenda

Tuesday, February 24

9:15 p.m. – 9:30 p.m.	**Arutha Kulasinghe, Frazer Institute, University of Queensland** "Spatial single-cell resolved multiomic profiling of the lung cancer tumour microenvironment, one-cell, one-niche and one-neighbourhood at a time"
9:30 p.m. – 9:45 p.m.	**Jude Dunne, Bruker Spatial Biology** "In-situ 3D spatial single-cell profiling of genome architecture in Estrogen receptor-positive (ER+) breast cancer"
Concurrent Session:	**Abstract Selected Talks**
7:30 p.m. – 9:45 p.m.	**Computational Biology** (Mike Zody, New York Genome Center, Session Chair) Bonnet Creek Ballroom VII-XII
7:30 p.m. – 7:45 p.m.	**Can Cenik, University of Texas** "Predicting the translation efficiency of messenger RNA in mammalian cells"
7:45 p.m. – 8:00 p.m.	**Lulu Shang, MD Anderson Cancer Center** "Statistical Identification of cell type-environment interaction-driven gene expression in spatial transcriptomics"
8:00 p.m. – 8:15 p.m.	**Yasin Uzun, Pennsylvania State University** "A benchmarking framework for single-cell multi-omics-based gene regulatory network inference"
8:15 p.m. – 8:30 p.m.	**Linghua Wang, University of Texas** "AI-enabled framework for tertiary lymphoid structure detection, phenotyping, and patient stratification"
8:30 p.m. – 8:45 p.m.	**Hao Feng, University of Texas Health Science Center at Houston** "Mist: a hierarchical Bayesian framework for detecting differential DNA methylation dynamics in single-cell data"
8:45 p.m. – 9:00 p.m.	**Hyun Min Kang, University of Michigan** "CartoScope: a harmonized, molecular-resolution web platform for spatial omics visualization, analysis, and sharing"
9:00 p.m. – 9:15 p.m.	**Dawei Li, Texas Tech University** "From integration to prediction: computational tools for virome and transposable element analysis"
9:15 p.m. – 9:30 p.m.	**Ioannis Mouratidis, University of Texas** "Fundamental limitations of large language models for genomic sequence generation"
9:30 p.m. – 9:45 p.m.	**Nicolas Fernandez, Broad Institute of MIT and Harvard** "Celldega: toolkit for analysis and visualization of spatial and single-cell data"
Starting at 9:45 p.m.	**Sponsor Social** Sponsor Promenade, Signia Suites

Agenda

Wednesday, February 25

7:30 a.m. – 9:00 a.m.	**Breakfast sponsored by Volta Labs** Waterside Lanai and Patio
7:30 a.m. – 8:00 p.m.	**Meeting Registration / Hospitality Desk** Bonnet Creek North Foyer
Plenary Session:	**Genetics** (Federica Di-Palma, Genome British Columbia, Session Chair) Floridian Ballroom
9:00 a.m. – 9:30 a.m.	**Elaine Mardis, The Steve and Cindy Rasmussen Institute for Genomic Medicine** "Precision pediatric medicine: a game-changer"
9:30 a.m. – 10:00 a.m.	**Megan Dennis, University of California, Davis** "Querying complex genomic regions important in human disease & evolution"
10:00 a.m. – 10:30 a.m.	**Coffee Break Sponsored by BD Biosciences** Sponsor Promenade
10:30 a.m. – 11:00 a.m.	**Kathryn Paige Harden, University of Texas** "Disease or moral failing? Rethinking what psychiatric genetics leaves out—and why"
11:00 a.m. – 11:20 a.m.	**Chandrima Bhattacharya, Weill Cornell Medicine (Late-Breaking Abstract Selected Talk)** "A global multi-omics atlas of microbial biosynthetic gene clusters"
11:20 a.m. – 11:50 a.m.	**Next Gen Leadership Awards** (Elinor Karlsson, University of Massachusetts Medical School; Beth Shapiro, University of California, Santa Cruz, Meeting Co-Chair) Floridian Ballroom
12:00 p.m. – 2:15 p.m.	**Silver Sponsor Workshops and Lunch** Bonnet Creek Ballroom VII-XII
12:05 p.m. – 1:05 p.m.	**Silver 1 Sponsor – Ultima Genomics – Workshop and Lunch**
1:10 p.m. – 2:10 p.m.	**Silver 2 Sponsor – Bruker Spatial Biology – Workshop and Lunch** Mark R. Munch, PhD, President of Bruker Nano Group, Jude Dunne, PhD, Vice President, New Product Development, Assays and Assay Systems, Bruker Spatial Genomics, Joseph Beechem, PhD, Chief Scientific Officer, Bruker Spatial Biology, Oliver Braubach, PhD, Director of Research & Development, Bruker Spatial Biology, Nigel Jamieson, PhD, Professor of Surgery, University of Glasgow "Expanding the frontiers of spatial biology: best-in-class solutions for every layer of biology"

Agenda

Wednesday, February 25

12:00 p.m. – 2:10 p.m.	**AGBT Lunch** Waterside Boulevard
2:15 p.m. – 4:44 p.m.	**Bronze Sponsor Workshops – (John McPherson, University of California, Davis, Session Chair)** Floridian Ballroom
2:20 p.m. – 2:40 p.m.	**Bronze 1 Sponsor – New England Biolabs, Inc.** Bernard Lam, Ph.D, Associate Director, Translational Genomics Laboratory, Ontario Institute for Cancer Research “Cross-platform evaluation of whole methylome sequencing in research and clinical specimens”
2:40 p.m. – 3:00 p.m.	**Bronze 2 Sponsor – Complete Genomics** Yuxuan Wang, MD, PhD, Assistant Professor of Oncology, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine “Clinical Oncology Applications of Circulating Tumor DNA Sequencing”
3:00 p.m. – 3:20 p.m.	**Bronze 3 Sponsor – Integrated DNA Technologies, Inc.** Ajay Gannerkote, IDT President and Steven Henck, Chief Technology Officer “Accelerating enhanced genomics solutions”
3:20 p.m. – 3:35 p.m.	**Bronze 4 Sponsor – Watchmaker Genomics** Travis Sanders, Senior Research Scientist “From bottlenecks to breakthroughs: unlocking high-quality WGS at scale”
3:35 p.m. – 3:50 p.m.	**Bronze 5 Sponsor – Oxford Nanopore Technologies** Marilyn Li, Director of Cancer Genomic Diagnostics, Children’s Hospital of Philadelphia and Lakmal Jayasinghe, Chief Scientific Officer, Oxford Nanopore Technologies “A new era of biological discovery”
3:50 p.m. – 4:05 p.m.	**Bronze 6 Sponsor – 10x Genomics** Jason Kim, Senior Manager, Science & Technology “Building impactful AI models for real-world results”
4:05 p.m. – 4:20 p.m.	**Bronze 7 Sponsor – Element Biosciences** Abbey Cutchin, Director Product Management and Rahul Lodhavia, Sr. Product Manager "Relentless innovation where it counts: from high-throughput genomics to direct tissue sequencing"
4:20 p.m. – 4:32 p.m.	**Bronze 8 Sponsor – BD Biosciences** Aruna Ayer, PhD, VP, Multiomics, Innovation & Scientific Affairs and Luciano Martelotto, PhD, Director of Global Market Development, Single Cell “High-confidence multiomics: unlocking deeper biological insights with BD Rhapsody”

Agenda

Wednesday, February 25

4:32 p.m. – 4:44 p.m.	**Bronze 9 Sponsor – Codetta Bio** William Henley, Co-Founder, Vice President of Research & Innovation "Redefining multi-omics with integrated protein and nucleic acid analysis in one harmonized workflow on Codetta Bio's Concerto system"
4:45 p.m. – 6:10 p.m.	**Poster Session and Wine & Cheese Reception Sponsored by Roche** Bonnet Creek Ballroom I-VI
6:10 p.m. – 7:25 p.m.	**AGBT Dinner** Waterside Ballroom
Concurrent Session:	**Abstract Selected Talks**
7:30 p.m. – 9:45 p.m.	**Technology (Stacey Gabriel, Broad Institute of MIT and Harvard, Session Chair)** Floridian Ballroom
7:30 p.m. – 7:45 p.m.	**Joshua Weinstein, University of Chicago** "Volumetric DNA microscopy: spatial transcriptomics in intact tissue"
7:45 p.m. – 8:00 p.m.	**Darren Segale, Illumina** "Spatial transcriptomic gene-expression profiles in breast cancer progression"
8:00 p.m. – 8:15 p.m.	**Rui Ye, MD Anderson Cancer Center** "Decoding genotypes and phenotypes during esophageal cancer progression with single-cell multi-omics"
8:15 p.m. – 8:30 p.m.	**Wei-Yu Chi, Weill Cornell Medicine** "Single-cell mapping of regulatory DNA: protein interactions"
8:30 p.m. – 8:45 p.m.	**Kanishk Asthana, University of California, San Diego** "ChronoSeq: minute-scale automated scRNA-seq for real-time profiling of perturbations"
8:45 p.m. – 9:00 p.m.	**Gary Schroth, Cellanome** "Multi-modal platform for studying cell-cell interactions and dynamics of neuron and glial cells with single-cell specificity"
9:00 p.m. – 9:15 p.m.	**John Doench, Broad Institute of MIT and Harvard** "High-resolution functional genomics: Integrating CRISPR perturbations and single-cell RNA-Seq"
9:15 p.m. – 9:30 p.m.	**Chao Yan, New York Genome Center** "SPACE-Tag Enables Spatial Chromatin Profiling With CUT&Tag"
9:30 p.m. – 9:45 p.m.	**Yutaka Suzuki, University of Tokyo** "Sequence-based spatial analysis of cancers: Trial use of Roche SBX sequencer for the VISIUM HD libraries"

Agenda

Wednesday, February 25

Concurrent Sessions:	Abstract Selected Talks
7:30 p.m. – 9:45 p.m.	**Biology (Emma Teeling, University College Dublin, Session Chair)** Bonnet Creek Ballroom VII-XII
7:30 p.m. – 7:45 p.m.	**Nigel Jamieson, University of Glasgow** “First application of true spatial multiomic integration on a single slide of pancreatic cancer using high plex protein and whole transcriptome”
7:45 p.m. – 8:00 p.m.	**Christopher Mason, Weill Cornell Medicine** “Embryonic program reactivation during tumorigenesis discovered with AI-driven subcellular mapping of the human body with SAHA data”
8:00 p.m. – 8:15 p.m.	**Miranda Orr, Washington University School of Medicine** “Subcellular spatial multiomics enables multimodal, multiscale 3D reconstruction of neuropathology in the human Alzheimer’s hippocampus”
8:15 p.m. – 8:30 p.m.	**Billy Lau, Stanford University School of Medicine** “Joint single-molecule cfDNA and cfRNA nanopore sequencing enables integrated multiomic cancer detection and monitoring”
8:30 p.m. – 8:45 p.m.	**Andrew Mungall, Canada’s Michael Smith Genome Sciences Centre** “Rapid long-read sequencing for genetic diagnosis in the NICU: Early results from a multi-family cohort”
8:45 p.m. – 9:00 p.m.	**Mamoru Kato, National Cancer Center Japan** “Simulator of cancer-cell evolution for personalized drug administration in cancer genome medicine”
9:00 p.m. – 9:15 p.m.	**Aparna Ganapathy, Strand Life Sciences Private Limited** “Whole exome sequencing in idiopathic female infertility identifies novel candidate targets for non-hormonal contraceptive discovery”
9:15 p.m. – 9:30 p.m.	**Winston Timp, Johns Hopkins University** “Targeting Long-read RNA sequencing to resolve splicing and modifications”
9:30 p.m. – 9:45 p.m.	**Brenda Murdoch, University of Idaho** “Three-dimensional chromosome architecture in developing mammalian tissues”

Agenda

Thursday, February 26

7:30 a.m. – 9:00 a.m.	**Breakfast Sponsored by Latch Bio** Waterside Lanai and Patio
7:30 a.m. – 6:00 p.m.	**Hospitality Desk** Bonnet Creek North Foyer
Plenary Session:	**Evolutionary Genomics** (Elinor Karlsson, University of Massachusetts Medical School, Session Chair) Floridian Ballroom
9:00 a.m. – 9:30 a.m.	**Katie Hinde, Arizona State University** “Celebrating the tree of life and the science of science engagement”
9:30 a.m. – 10:00 a.m.	**Chris Jiggins, University of Cambridge** “Genomes and the network of life”
10:00 a.m. – 10:30 a.m.	**Coffee Break Sponsored by Latch Bio** Sponsor Promenade
10:30 a.m. – 11:00 a.m.	**Dan Rokhsar, University of California, Berkeley** “You can see a lot just by looking: deeply conserved synteny and the evolution of animals”
11:00 a.m. – 11:30 a.m.	**Tomas Marques-Bonet, University Pompeu Fabra** “Primate genomes as a model for biomedicine, evolution, and conservation”
11:30 a.m. – 12:00 p.m.	**Poster Flash Talks II (2 min. each)**
12:00 p.m. – 1:30 p.m.	**AGBT Lunch** Waterside Boulevard
12:15 p.m. – 1:15 p.m.	**Silver 3 Sponsor – Roche Sequencing Solutions, Inc. – Workshop and Lunch** Mitu Chaudhary, International Business Leader, Sequencing Systems, Roche Diagnostics Solutions, Johanna Whitacre, Network Lead, Sequencing Research & Development, Roche Diagnostics Solutions, Gustav Karlberg, Lifecycle Leader, Sequencing Systems, Roche Diagnostics Solutions, and Edwin Cuppen, Scientific Director, Hartwig Medical Foundation; Mark Kokoris, Head of SBX Technology “The Power of One" and “Cancer Whole Genome Sequencing” Bonnet Creek Ballroom VII-XII
Plenary Session:	**Technology** (Stephen Montgomery, Stanford University, Session Chair) Floridian Ballroom

Agenda

Thursday, February 26

1:30 p.m. – 2:00 p.m.	**Neville Sanjana, New York Genome Center** “Massively-parallel approaches to understand the human noncoding genome and transcriptome”
2:00 p.m. – 2:30 p.m.	**Job Dekker, University of Massachusetts and Howard Hughes Medical Institute** "Mitotic inheritance of chromosome folding programs"
2:30 p.m. - 3:00 p.m.	**Elinor Karlsson, Broad Institute of Harvard and MIT**
3:00 p.m. – 3:20 p.m.	**Jon Bezney, Stanford University** (Abstract-Selected Talk) “CRISPR depletion of whole-blood RNA-seq for enhanced transcriptional phenotyping of rare Mendelian disorders”
3:20 p.m. – 3:40 p.m.	**Yasuhiro Murakawa, Kyoto University** (Late-Breaking Abstract Selected Talk) “A new flow cell–based spatial transcriptomics technology for transcription start site profiling at submicron resolution”
3:40 p.m. – 4:00 p.m.	**James Occean, Stanford University** (Abstract-Selected Talk) “Pan-cancer whole-genome profiling of DNA methylation in primary human tumors”
4:00 p.m. – 4:10 p.m.	**Closing Comments** Beth Shapiro, University of California, Santa Cruz, Meeting Co-Chair Len Pennacchio, Lawrence Berkeley National Labs, Meeting Co-Chair
6:00 p.m. – 11:00 p.m.	**Final Night Dinner – Cirque Soirée Sponsored by New England Biolabs, Inc.** 19th Green

Friday, February 27

4:00 a.m. – 2:00 p.m.	**Transfers to Orlando International Airport (MCO)**

AXELIOS

Roche

A new beginning starts here

Roche Sequencing Solutions welcomes you to the world of AXELIOS 1, the platform powered by SBX that heralds a new era in genomics.

Explore the new “next” in NGS – join us at the **Jade Lexington Suite**

The AXELIOS 1 sequencing platform and sequencing by expansion (SBX) technology are in development and not commercially available. The content of this material reflects current study results and/or design goals. The AXELIOS 1 sequencing platform based on SBX technology will be launched for Research Use Only. Not for use in diagnostic procedures. AXELIOS is a trademark of Roche.

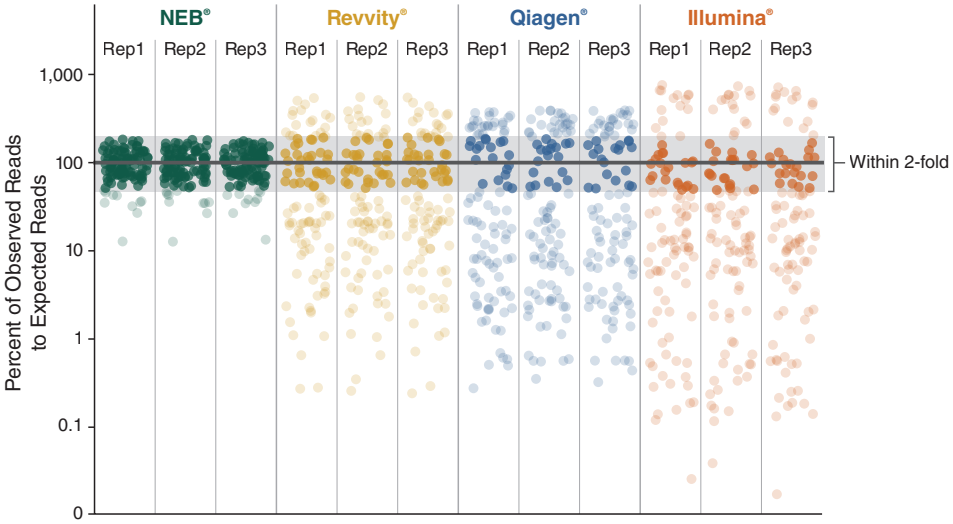
Abstract Table of Contents

The Peacock Mantis Shrimp (Odontodactylus scyllarus) is very small but can unfurl its appendages and become quite powerful.

Successfully see more small RNAs with the NEBNext® Low-bias Small RNA Library Prep Kit

Bias in small RNA library prep can seriously impact the accuracy of your data, skewing the representation of rare RNA species. The NEBNext Low-bias Small RNA Library Prep Kit minimizes biased representation of small RNAs, more accurately reflecting the number and proportion of unique small RNAs present. NEB's next-generation kit utilizes a streamlined workflow for the capture of small RNAs (<120 nt) with a 5′ phosphate and 3′ hydroxyl group, and can be used to identify microRNAs (miRNAs), transfer RNAs (tRNAs) and tRNA-derived fragments, small nucleolar RNAs (snoRNAs), piwi-interacting RNAs (piRNAs), and plant miRNAs. Designed for the preparation of Illumina®-compatible small RNA sequencing libraries, this latest NEBNext kit is kind of a big deal!

NEBNext Low-bias Small RNA Library Prep Kit produces libraries with the lowest bias.



NEBNext Low-bias Small RNA libraries were made using a mix of 100 synthetic control miRNAs, including five that had 3′ 2′-O-methyl ends. Libraries were prepared from 0.3 ng of the synthetic miRNA mix to compare library-generated bias between the NEBNext Low-bias Small RNA Library Prep Kit and small RNA kits from Revvity® (NEXTFLEX Small RNA Sequencing Kit V4), Qiagen® (QIAseq® miRNA Library Kit) and Illumina (TruSeq® Small RNA Library Prep). Libraries were sequenced on an Illumina NextSeq 500 (1 x 56 bases), and expected reads were calculated from total reads mapped to the synthetic controls, divided by the total number of control sequences (black line at 100%). Percent of observed reads to expected reads was calculated for each control sequence and plotted across replicates.

Visit New England Biolabs in **Sarasota** to learn more.

NEW ENGLAND Biolabs®

ABSTRACT CATEGORY 01

Cancer Omics

Molecular classification of cancers by transcriptome profiling of from FFPE specimens

Cancer Omics

POSTER 101

Hiroyuki Aburatani[1], Shuichi Tsutsumi[1], Kenji Tatsuno[1], Motohiro Kato[2], Kayoko Tao[3]

1. The University of Tokyo, Research Center for Advanced Science and Technology, Tokyo, Japan
2. The University of Tokyo, Graduate School of Medicine, Tokyo, Japan
3. National Cancer Center Hospital, Department of Pediatrics, Tokyo, Japan

The TOP2 hybrid capture-enrichment system was utilized to analyze 737 DNA-panel genes and 455 RNA-panel genes (Kohsaka, 2019). In the JCCG-TOP2 study conducted between January 2022 and February 2023, a total of 210 pediatric cancer cases were enrolled from 50 institutions nationwide in collaboration with the Japan Children's Cancer Group (Tao, in revision). Among the 205 evaluable samples, six failed to pass quality control due to insufficient quantities of extracted DNA and RNA (n=1) or inadequate pre-capture RNA library yield (n=5).

TOP2 analysis demonstrated significant benefits in clarifying the diagnosis of pediatric solid tumors. Of the 204 patients with successful genomic results, 147 (72%) had potentially actionable findings — including diagnostic, prognostic, and therapeutic relevance in 111 (54%), 61 (30%), and 64 (31%) cases, respectively. Oncogenic fusions were detected in 45 (23%) patients, while copy number alterations were identified in at least one genomic region in 170 (83%) patients. Seventeen (8%) patients harbored germline pathogenic or likely pathogenic variants in cancer-predisposing genes.

Additionally, we developed a deep learning-based approach using transcriptome data to determine the most probable cancer type or subtype. The detailed results will be presented.

Cancer Omics

POSTER 102

Visium HD enabled transcriptomic profiling of visceral metastases of prostate cancer reveals changes in tumor microenvironment

Naishitha Anaparthy[1], Grace Waldron[2], Diane Costanzo-Garvey[2], Rachel Z Blumhagen[4], Trey Farmer[4], Chrystal Nathan[3], Borum Ryu[2], Debashish Chitnis[1], Monica Nagendran[1], Chaitanya Aluru[1], Heather C Smith[3,5], Holly A F Stessman[6], Kabhilan Mohan[2,3], Grinu Mathew[2,3]

1. 10x Genomics, Pleasanton, CA, USA
2. University of Nebraska Medical Center, Eppley Institute for Research in Cancer and Allied Diseases, Omaha, NE, USA
3. University of Nebraska Medical Center, Fred and Pamela Buffett Cancer Center, Omaha, NE, USA
4. National Jewish Health, Center for Genes, Environment, and Health, Denver, CO, USA
5. University of Nebraska Medical Center, Department of Genetics, Cell Biology, and Anatomy, Omaha, NE, USA
6. Creighton University, Department of Pharmacology and Neuroscience, Omaha, NE, USA

Prostate cancer (PC) is the second leading cause of cancer-related death among men in the United States and remains a major global health challenge. A subset of advanced, treatment-resistant PC undergoes neuroendocrine differentiation (NEPC), which is frequently associated with visceral metastases and poor clinical outcomes. These metastases arise from disseminated tumor cells that can remain dormant for prolonged periods before reactivation. The mechanisms governing their survival and awakening in visceral niches remain largely unknown.

To address this, we investigated metastatic growth in NEPC by focusing on AXL, a bonafide dormancy marker. We modeled NEPC visceral metastases in an immunocompetent mouse system, using PTEN/TP53/AXL-null genetics (strongly associated with NEPC development), providing a platform to investigate the molecular and spatial cues that drive metastatic adaptation in aggressive PC. Our results show that AXL loss triggers cellular proliferation and metastatic outgrowth, suggesting AXL's critical role in regulating the switch from dormancy to proliferative lesions. Leveraging 10x Genomics' Visium HD Spatial Gene Expression, which utilizes a whole-transcriptome panel (probe-based RNA-templated chemistry), we resolved the tumor microenvironment (TME) at metastatic sites at single-cell scale in a FFPE block of modeled NEPC visceral metastasis. Visium HD XL slides allowed assessment of multiple metastatic sites and cell segmentation allowed for defining cell types with spatiality in TME. We observed changes in innate (increased neutrophils and macrophage infiltration) and adaptive (increased T regulatory cells) immune systems at metastatic locations. Visium HD workflow, compatible with IF staining, allowed us to pinpoint cancer cells when layered with transcriptomic data. Overall, we observed an immunosuppressive TME with altered stromal signaling that promoted metastatic outgrowth. Integration of immunohistochemistry (vimentin, alpha-SMA, trichrome staining) and multiphoton second harmonic collagen imaging on serial sections revealed distinct extracellular matrix remodeling patterns, characterized by increased collagen deposition associated with AXL loss. We are further probing lung metastasis samples derived from NEPC patients with Visium HD XL to confirm AXL loss and TME changes observed. Utilizing Visium HD XL to define the molecular programs underlying dormancy and metastatic reactivation is critical, as these insights may reveal broadly applicable strategies to prevent disease progression across cancer types.

Cancer Omics

POSTER 103

Graph-Based Variant Discovery in African Breast Cancer Genomes Using the HPRC-T2T Pangenome

Alexandria Awai[1,2,8], Praise Idowu[1,2], Rachel Martini[1,2*], Abdulfatai Tijjani[3,4*], Kyriaki Founta[3,4,5], Jason A. White[1,2], Onyinye Balogun[6], Nyasha Chambwe[3,4,5+], Melissa B. Davis[1,2,7,8+]

1. Institute of Translational Genomic Medicine, Morehouse School of Medicine, Atlanta, GA
2. Department of Microbiology, Biochemistry & Immunology, Morehouse School of Medicine, Atlanta, GA
3. Northwell, New Hyde Park, NY
4. Feinstein Institutes for Medical Research, Northwell Health, Manhasset, NY
5. Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY
6. Weill Cornell Medicine, New York, NY
7. Georgia Research Alliance, Atlanta, GA
8. New York Genome Center, New York, NY

Predominantly used human reference genomes, including GRCh38 and even the gapless T2T-CHM13 assembly, remain limited in their representation of African genomic diversity. Recent analyses of the African Pan-Genome (APG) revealed ~296 Mb of African ancestry–specific sequence absent from GRCh38, underscoring that even the most complete linear references incompletely capture African variation. The Human Pangenome Reference Consortium (HPRC) has since released a telomere-to-telomere, graph-based reference that better represents global diversity, yet its impact on variant discovery in cancer genomes from African patients remains largely untested.

To address this, we analyzed whole-genome sequencing data from six African breast cancer patients, aligning reads to both the GRCh38 linear reference and the HPRC-T2T graph using the Variation Graph (VG) Toolkit. We compared alignment quality, variant call sets, and callable bases across workflows to assess the added value of graph-based representation. Preliminary analyses show that pangenome alignment increases read mappability in previously unresolved and ancestry-enriched regions and reveals structural and rare variants not captured by GRCh38-based pipelines. Many of these loci overlap immune and signaling pathways consistent with prior APG functional enrichments, suggesting continuity between contig-level discoveries and raw whole-genome evidence.

This study provides one of the first applications of the HPRC-T2T pangenome to African cancer genomes and demonstrates how graph-based methods can uncover population-specific variation relevant to tumor biology. Ongoing work integrates these variant calls with somatic mutation and transcriptomic profiles to identify ancestry-linked molecular features of breast cancer. Together, our findings emphasize the necessity of incorporating ancestry-enriched genomic sequences into future reference frameworks and highlight graph-based genomics as a key step toward equitable precision oncology.

[1-3]

1. Martini, R., et al., Uncovering Functional Sequence Gaps in Human Reference Genomes using African Pan Genome Contig Sequences. bioRxiv, 2025: p. 2025.08.15.670543.
2. Li, H., et al., Exome variant discrepancies due to reference-genome differences. Am J Hum Genet, 2021. 108(7): p. 1239-1250.
3. Siren, J., et al., Personalized pangenome references. Nat Methods, 2024. 21(11): p. 2017-2023.

Cancer Omics

POSTER 104

Whole-transcriptome multi-omic subcellular imaging in space AND time: bringing spatial biology to life!

Joseph M. Beechem[1], Eric Hobbs[1], Joseph Zenga[2], Margaret Hoang[1], Erin Piazza[1], Ashley Heck[1], Rachel Liu[1], Kimberly Young[1], Mirko Corselli[1], Patrick Danaher[1], Dan McGuire[1]

1. Bruker Spatial Biology, Research, Seattle WA, USA
2. Medical College of Wisconsin, Division of Head & Neck Surgical Oncology, Milwaukee WI, USA

GOAL:
Establish an experimental strategy to "connect" live-cell functional assays with high-plex spatial biology. To achieve this capability, we have developed an experimental approach to bridge-the-gap between live-cell functional assays using single-cell optofluidic Beacon DiscoveryTM technology combined with CosMx® whole-transcriptome Imaging.

EXPERIMENTAL APPROACH:
From cancer excisional biopsies, directly measure live dissociated T-cells ability to kill primary tumor cells (in Beacon Discovery) and perform scRNA-seq and TCR-seq on those T-cells that kill, to establish a unique molecular identifier. From same biopsy material (adjacent section, FFPE), perform CosMx whole transcriptome (with T-cell repertoire clonality capability) to spatially resolve the identical molecular identifier, directly linking spatial-resolution with live-cell image-based functionality.

RESULT:
Head-and-neck cancer biopsy material (live cells) were dissociated and T-cell/Tumor-Cell combinations moved into the ~ 22000 nano-pens (per chip) of the Beacon Discovery system. T-cell/Tumor-cell nano-pen combos were then real-time imaged to determine which T-cell had tumor killing ability. Tumor-killing T-cells were then moved out of the nano-pens and subjected to scRNA-seq and scT-cell repertoire sequencing, generating a unique molecular "identifier". These same biopsy materials were then whole transcriptome imaged with TCR-specific VDJ-recombination probes on CosMx to resolve exactly "where-in-space" T-cells with matching identifiers were located.

Starting with ~ 160,000 single-cells in the biopsy, we spatially resolved 6500 T-cells: 1375 T-regs, 1130 T-naïve, 1009 T-memory, 1463 T-proliferative, and 1523 T-exhausted. We estimated 41 total cells to match the Beacon live-cell molecular identifiers for a single tumor-killing T cell clonotype, of which 3 could be spatially identified at over 90% confidence-level (TCR clonotype = TRAV8-6, TRAJ54, TRBJ1-1). These cells fit the proliferative / non-exhausted immunophenotype expected based on matched single cell data. An additional 126 T-cells were also spatially mapped with high confidence across 7 different live-cell molecular identifiers for terminal-exhausted clonotypes. These T cells corresponded to the "exhausted" and "proliferative" immunophenotypes from scRNAseq.

As far as we are aware, this is the very first case where T-cell "space-and-time-and-function" have been uniquely (and unambiguously) linked together. This experimental approach can be expanded to any study where live-cell function and single-cell spatial resolution are important (e.g., CAR-T cells, NK-cells, Vaccine development, stem-cell studies, etc).

Cancer Omics

POSTER 105

Scalable genotyping in fixed transcriptomes resolves clonal heterogeneity

Sydney B. Blattman[1*], Nabih Maslah[1,2*], Austin Varela[1*], Karolis Kumpaitis[1,2], Benan Nalbant[1], Catherine Snopkowski[1,2], Marisa Mariani[1,2], Laura Kida[1], Meril Takizawa[1,2], Kenny Kwok Hei Yu[3], Sunjay Fernandes[4], Nima Mousavi[4], Erik Borgstrom[4], Derek Vallejo[4], Lorita Boghospor[4], Ruijiao Xin[4], Marco Mignardi[4], Snow Wu[4], Nicholas Scarlott[4], Loruhama Delgado-Rivera[4], Poornasree Kumar[4], Sreenath Krishnan[4], Andrew Kohlway[4], Paul Lund[4], Dana Pe'er[1,5], Ronan Chaligné[1,2], Caleb Lareau[1]

1. Memorial Sloan Kettering Cancer Center, Sloan Kettering Institute, Computational and Systems Biology Program, New York, NY, USA
2. Memorial Sloan Kettering Cancer Center, Single-cell Analytics Innovation Lab, New York, NY, USA
3. Memorial Sloan Kettering Cancer Center, Department of Neurosurgery, New York, NY, USA
4. 10X Genomics, Pleasanton, CA, USA
5. Howard Hughes Medical Institute, Chevy Chase, MD, USA

*These authors contributed equally.

Somatic evolution drives cancer progression and therapy resistance through extensive clonal heterogeneity. While bulk and single-cell sequencing have elucidated the genetic trajectories of tumor evolution, connecting these mutations to their downstream transcriptional consequences remains a major challenge. Existing methods, such as Genotyping of Transcriptomes (GoT), have begun bridging this gap by jointly capturing genotypes and transcriptomes, but they remain constrained by poly(A)-capture bias, limited multiplexing, low throughput, and incompatibility with fixed tissues, highlighting the need for scalable, multi-target, and FFPE-compatible technologies.

Here, we present Genotyping in Fixed Transcriptomes (GIFT)-seq, a probe-based method that uses gap-filling chemistry to enable highly accurate detection of hundreds of targeted genetic variants (up to 10 nucleotides in length) concurrently with full-transcriptome single-cell profiling. GIFT-seq employs customizable probe panels that can seamlessly extend commercial 10x Genomics Flex single-cell and single-nuclei RNA-seq platforms, allowing flexible profiling of mutations across diverse biological contexts, from hematopoiesis to tumorigenesis and metastasis.

After optimizing the probe design and demonstrating the high-throughput capability of GIFT-seq to capture multiple targets simultaneously (detecting up to 364 targets in parallel within a single cell and up to 588 targets in a single experiment or population), we applied the method to samples from patients with myeloproliferative neoplasms (MPNs). Using GIFT-seq on $CD34^+$ PBMCs from 35 patients, we profiled 65 mutations in parallel, including JAK2 c.1849G>T, enabling the simultaneous detection of dozens of mutations and the reconstruction of clonal lineage trees (652,558 cells with single-cell transcriptome and at least one target captured). Finally, we validated the method's capabilities on FFPE glioblastoma samples, establishing GIFT-seq as a powerful and versatile platform for linking genotype to phenotype and enabling clonal tracking in both fresh and archival tissues.

In situ single-cell spatial multiomic profiling of 3D genome organization in fresh frozen colorectal carcinoma tissue in spatially resolved tissue micro-environments

Cancer Omics

POSTER 106

Shyamtanu Chattoraj[1], Pamela Flatley[1], Stacy Elliott[1], David Castillo[1], Kenny Chung[1], Doug Werner[1], Jude Dunne[1], Huy Nguyen[1]

1. Bruker Spatial Genomics, Research and Development, San Jose, CA, USA

The 3D organization of genome structure plays a critical role in regulating gene expression and cellular function. In recent years, disruption in higher order chromatin structure and 3D spatial genome organization has been extensively implicated in disease including various types of cancers. Structural aberrations including copy number variations (CNVs), structural variations (SVs), and topological changes rewire the 3D genome allowing aberrant activation of oncogenes, repression of tumor suppressors and dysregulation in gene expression. While methods like WGS and Hi-C offer bulk-level insights into genomic dysregulation, they disrupt native tissue architecture, limiting the ability to analyze spatial relationships and organizational features essential for understanding cellular context and heterogeneity. Since high degree of tumor cell heterogeneity significantly contributes towards therapeutic resistance with negative impact on disease-free survival of cancer patients, understanding the molecular mechanism behind high intratumoral heterogeneity requires state-of-the-art single cell and spatial technologies that can provide deeper insight into early detection and disease progression.

Here, we present a novel multiomic jebFISH™ protocol on the PaintScape™ platform that can be used to characterize intratumoral heterogeneity of 3D chromatin architecture of fresh frozen normal human colon tissue and colorectal cancer (CRC) samples in different tissue microenvironments at single cell, sub-population and population level. Using the ChromoPaint™ PanChromo MPX Panel, which visualizes over 400 loci across all chromosomes, we show disruption in 3D folding of individual chromosomes such as re-organization of chromosome territories, differential p-q arm interaction and inter-chromosomal proximity at single cell and sub-population level. We characterize chromosomal instability, CNVs and SVs of different colon cancer cell states at sub-chromosomal level in situ within the tissue. By combining the jebFISH protocol with multiplexed immunofluorescence, we characterize 3D genome organization of cancer cell sub-populations along with proteomic signatures in different tissue microenvironment e.g. immune rich vs immune deficient regions within the same CRC tissue section.

We envision the PaintScape system will improve our understanding of the factors driving progression of colon cancer disease and provide deeper insight into the underlying 3D genomic heterogeneity of single cancer cells and sub-populations with unique spatial context within the tissue microenvironment.

Cancer Omics

POSTER 107

Streamlined dual-omic cfDNA analysis for pediatric renal cancer with BrightSeq

Carrie Cibulskis[1], Michelle Ciipicchio[2], Cora Ricker[3], Nenzi Ruzgar[3], Can Kockan[1], Mohammed Tanhaemami[3], Chris Kachulis[1], Micah Rickles-Young[1], Mark Fleharty[1], Kelly Klega[3], Sam Pollock[1], Sophie Low[1], Sean Hofherr[1], Niall Lennon[1], Alanna Church[4], Brian Compton[3]

1. Broad Institute of MIT and Harvard, Boston, MA, USA
2. Broad Clinical Labs, Complex Assay Services, Burlington, MA, USA
3. Dana Farber Cancer Institute, Pediatric Hematology/Oncology, Boston, MA, USA
4. Boston Children's Hospital, Laboratory for Molecular Pediatric Pathology, Boston, MA, USA

For decades, the diagnostic, monitoring, and treatment landscape for rare pediatric cancers has seen few transformative advances, due in part to challenges in powering genomic discovery. Pediatric tumors also harbor distinct genomic and epigenetic features that require innovative sequencing approaches to enable precision care and accelerate discovery. To address these challenges, we established BrightSeq, a partnership among Boston Children's Hospital, Dana-Farber Cancer Institute, and Broad Clinical Labs. BrightSeq is developing a suite of novel clinical cell-free DNA (cfDNA) and tissue-based assays to profile important events in pediatric cancers.

One priority focus is pediatric renal tumors, which comprise diverse histologies with highly variable prognoses. Current diagnostic approaches rely heavily on surgical resection as imaging and core needle biopsies have limited accuracy. Consequently, less invasive circulating-tumor DNA assays capable of distinguishing tumor types and guiding therapy are urgently needed. To explore the possibility of improving BrightSeq cfDNA assays with epigenomics features, we applied Illumina's 5-Base conversion chemistry, which enables simultaneous interrogation of genomic and methylation features from the same cfDNA sample, to a pilot patient cohort (cfDNA) of 8 malignant rhabdoid tumors of the kidney (MRTK), 22 Wilms tumor samples, and 3 non-renal cancer cell lines.

Whole-genome 5-Base sequencing (50x depth) produced high-quality variant calls for both somatic and germline events ((HG002 SNP F1: 0.996, Indel F1: 0.979) vs standard Illumina baseline (HG002 SNP F1: 0.999, Indel F1: 0.997)). Copy number analysis with iChorCNA yielded accurate %ctDNA estimates across samples, with tumor fractions ranging from 3.2-55.1%. Preliminary integration of methylation patterns enabled clustering of MRTK versus Wilms tumor cfDNA samples when % tumor fraction (%TFx) exceeded 5%. The dual-omic data generated from 5-Base libraries also demonstrated strong concordance with the same specimens previously characterized with the NEBNext EmSeq methylation conversion approach (r2 >= 0.97). Together, these results and the operational efficiencies achieved with 5-Base chemistry - including simplified laboratory processing and direct compatibility with Dragen pipelines - underscore the promise of this technology for scalable adoption in the BrightSeq Initiative. Future studies will expand these analyses across additional pediatric tumor samples to establish robust benchmarks for clinical and discovery applications.

Sequencing platform-independent comprehensive cancer genome analysis for precision medicine research and diagnostics using Hartwig WiGiTS

Cancer Omics

POSTER 108

Edwin Cuppen[1], Charles Shale[1], Stephen Watts[2], Hong Wing Lee[1], Daphne van Beek[1], Sandra van den Broek[1], Ewart de Bruijn[1], Junran Cao[1], Matthew Cooper[1], Erik van Dijk[1], Korneel Duyvesteyn[1], Thomas Garrety[1], Rayan Hassaine[1], Nina Jacobs[1], Reece Jones[1], David Koetsier[1], Tim Lavers[1], Joep de Ligt[1], Francisco Martinez-Jimenez[1,3], Matias Mendeville[1], Luan Nguyen[1], Joseph Usset[1,3], Oliver Hofmann[2], Peter Priestley[1]

1. Hartwig Medical Foundation, Amsterdam, Netherlands and North Sydney, Australia
2. University of Melbourne, Melbourne, Australia
3. Vall d'Hebron Institute of Oncology (VHIO), Barcelona, Spain

Comprehensive genomic analysis is increasingly integral to both cancer diagnostics and translational research. As clinical sequencing evolves from targeted panels and exome assays toward whole-genome approaches, there is a critical need for robust, scalable, and integrative computational frameworks. Here, we introduce WiGiTS, an open-source, platform-independent suite of bioinformatics tools designed for end-to-end analysis of short-read cancer sequencing data across diverse assay types—including large panels, exomes, and whole genomes—in both tumor-only and tumor-normal paired configurations. WiGiTS delivers high-confidence variant calling and annotation across all variant classes, while also extracting clinically relevant features and biomarkers such as microsatellite instability, homologous recombination deficiency, HLA genotypes, viral integrations, pharmacogenomic features, and tissue-of-origin predictions. When RNA-seq data is available, WiGiTS integrates transcriptomic features to enhance variant interpretation and tumor profiling.

The pipeline supports data from all major short-read sequencing platforms, including Illumina, Ultima, and Roche and genome reference builds (GRCh37, GRCh38). The tools are also implemented in Nextflow (named OncoAnalyser) as part of the nf-core ecosystem, enabling reproducible deployment across local clusters and cloud infrastructures.

Collectively, WiGiTS provides a validated, vendor-neutral solution for scalable, cost-efficient, and clinically relevant interpretation of cancer genomes, supporting both routine diagnostics and precision oncology research.

Cancer Omics

POSTER 110

Extracellular vesicle and particle–based liquid biopsy detects early-stage lung cancer with performance similar to LDCT

Toumy Guettouche[1], Ibukunoluwapo O. Zabrowski[1], Brittany Grimes[1], Timothy Santos-Heiman[1], Aaron Chevalier[1], Mahmoud M. Bassam[1], Gabrielle N. Barcaskey[1], Katherine S. Yang[1], MacKenzie S. King[1], Cassidy Shultz[1], Eric Szu Pu[2], Troy B. Hawkins[1], Dawn R. Mattoon[1], Daniel P. Salem[1]

1. Mercy BioAnalytics, Inc., Waltham, MA, USA
2. Independent Consultant

Early detection of lung cancer remains a critical unmet need, as most patients are diagnosed at advanced stages, when treatment is less effective. Circulating extracellular vesicles and particles (EVPs) offer a minimally invasive source of tumor-derived biomarkers reflective of underlying tumor biology. Leveraging this biology, we developed an EVP-based liquid biopsy assay for lung cancer detection.

To accelerate biomarker discovery and refine selection, we implemented an NGS-based EVP surface protein mapping platform. The NGS platform was validated using the lung assay. This demonstrated its ability to profile the EVP surface proteome and highlights its potential as a tool for accelerating biomarker discovery and optimizing selection of biomarker combinations across multiple cancer types.

In a case–control evaluation comprising 219 controls and 114 lung cancer cases, with cases strongly biased toward early-stage disease (88% Stage I/II, median tumor diameter 2.7 cm), plasma samples were collected from individuals with current or prior tobacco use, those without a history of tobacco use, patients with COPD, and individuals with benign nodules. EVPs from 500 µl of plasma were analyzed using a proximity ligation immuno-PCR assay with an antibody panel targeting five colocalized protein biomarkers.

The assay achieved 58% sensitivity (66/114, 95% CI: 48.7–66.6) and 90% specificity across all stages. In Stage I adenocarcinoma, sensitivity was 57% at 90% specificity (39/68, 95% CI: 45.5–68.4), with successful detection of tumors <1 cm. Assay signal correlated with tumor size (Spearman $r = 0.541$, $p < 0.001$) but not with smoking history, underscoring the biological relevance of EVP biomarker colocalization.

These findings demonstrate the potential of EVP-based liquid biopsy to identify clinically significant tumors at earlier, more treatable stages and support further validation in independent cohorts. To our knowledge, this is the first blood-based liquid biopsy assay with performance similar to low-dose computed tomography (LDCT). With its low sample requirement and qPCR-based readout, this platform offers a scalable, minimally invasive approach adaptable to diverse clinical environments and a promising technology for improving early cancer detection and outcomes.

Exploring status of drug-resistant cancer cells using the Teton CytoProfiling system

Cancer Omics

POSTER 111

Yasuhiko Haga[1], Yutaka Suzuki[1]

1. The University of Tokyo, Graduate School of Frontier Science, Kashiwa, Chiba, Japan

It is widely recognized that cancer recurrence frequently arises from drug-resistant persistent cancer cells. These cells evade treatment through diverse strategies, including genomic mutations and transcriptomic and epigenomic alterations. Especially in Non-Small Cell Lung Cancer (NSCLC) driven by EGFR mutation, about 50% of recurrence mechanisms remain unclear. Therefore, understanding transcriptomic and epigenomic changes in drug resistance is vital for developing strategies against cancer recurrence. In addition to molecular alterations, morphological plasticity has recently been implicated in drug tolerance and resistance, suggesting that morphological changes may reflect or contribute to underlying molecular reprogramming.

To elucidate drug resistance mechanisms in lung cancer, we subjected PC9 cells to a 3-day gefitinib treatment followed by a 3-day recovery period, repeating this cycle 20 times. The resulting drug-dosed PC9 cells were resistant to both gefitinib and osimertinib without secondary genomic mutations and displayed distinct morphological changes compared to parental PC9 cells. In this study, we investigated the relationship between molecular and morphological changes using the Teton CytoProfiling system, which quantifies 350 RNA targets, 50 protein targets, and morphological features. We collected Teton data from parental, gefitinib-dosed, and resistant PC9 cells across four time points, including the parental state and under drug phases in cycles 1, 11, and 20. The mean transcripts and proteins per cell were 854 and 122, respectively.

To analyze the dynamic changes in cell morphological distributions in response to drug treatment, we constructed a support vector machine model using only morphological features as input and identified key features underlying these alterations. We then explored genes contributing to these morphological changes throughout drug treatment cycles. As a result, we identified genes whose expression correlated with key morphological features distinguishing different time points.

Cancer Omics

POSTER 112

Distinct cancer-associated fibroblast populations induced by epithelial and mesenchymal PDAC subtypes in a xenograft model

Jung Hyun Jo[1], Kyumin Lim[2], Soo been Park[2], Jung-hwan Park[2], Yoo Keung Tae[2], Hee Won Choi[2], Hee Seung Lee[2], Hyungsoon Im[1], Sung-Hwan Choi[2]

1. Massachusetts General Hospital, Boston, MA, USA
2. Yonsei University College of Medicine, Republic of Korea

Background:
Pancreatic ductal adenocarcinoma (PDAC) is classified into epithelial and mesenchymal subtypes associated with varied patient outcomes. Differences in cancer-associated fibroblasts (CAFs) in the tumor microenvironment may contribute to this variation. scRNA-seq studies have revealed heterogeneous CAF populations, although the relationship between tumor subtype and CAF phenotype remains unclear. To investigate this relationship, we utilized epithelial and mesenchymal PDAC cells in xenograft models to examine how these subtypes influence CAF phenotypes.

Methods:
Xenograft models were generated by subcutaneously injecting BALB/c nude mice with PDAC cell lines of either the epithelial subtype (BxPC-3) or the mesenchymal subtype (MIA PaCa-2). Tumors were then harvested and dissociated for scRNA-seq to investigate the TME. Human tumor and murine stromal cells were computationally separated using the Xenocell tool. Subsequent analyses focused on fibroblasts within the TME, which were subclustered to assess CAF heterogeneity between PDAC subtypes. To characterize transcriptional programs and identify CAF features, we performed high-dimensional weighted gene co-expression network analysis (hdWGCNA) and analyzed differentially expressed genes. Finally, spatial transcriptomics was conducted on tumor sections from xenografts to validate CAF clusters observed in scRNA-seq.

Results:
Distinct CAF clusters were identified in the xenograft tumors, including LRRC15 CAFs, RGS5 CAFs, and inflammatory CAFs (iCAFs). LRRC15 CAFs were predominant in tumors derived from the epithelial subtype, while RGS5 CAFs were enriched in tumors from the mesenchymal subtype. LRRC15 CAFs exhibited a myofibroblastic CAF (myCAF) phenotype characterized by extracellular matrix production, with high expression of Lrrc15, Postn, and Fbln1. RGS5 CAFs expressed canonical pericyte markers such as Rgs5 and Cox4i2, along with CAF-associated genes, including Acta2. These CAF clusters also showed distinct expression patterns of collagens and secreted factors. In particular, Col11a1, Col12a1, and Wnt5a were upregulated in LRRC15 CAFs, while Col4a1, Col10a1, and Tgfb1 were upregulated in RGS5 CAFs. Spatial transcriptomics confirmed gene modules associated with Lrrc15 and Rgs5, inferred from hdWGCNA, were differentially expressed between Acta2-positive stromal regions of epithelial and mesenchymal PDACs.

Conclusions:
These results suggest that epithelial and mesenchymal subtypes directly shape distinct CAF populations. Notably, the enrichment of RGS5 CAFs in mesenchymal tumors may underlie their metastatic potential and poor clinical outcomes.

Cancer Omics

POSTER 113

Comparative transcriptomic analysis of UHRF1 expression knockout in different cancer types

Jonathan Kalami[1,2*], Miriam E.L. Gakpey[3], Benson R. Kidenya[4*], Sala Kotochi[5], Benthai Benjamin[6], Nana Yaa Karikari Agyeman[3], Olaitan I. Awe[7]

1. Kwame Nkrumah University of Science and Technology, Department of Biochemistry and Biotechnology, Kumasi, Ghana
2. Kwame Nkrumah University of Science and Technology, National Institute of Mathematical Sciences, Kumasi, Ghana
3. Noguchi Memorial Institute for Medical Research, Legon, Ghana
4. Catholic University of Health and Allied Sciences, Department of Biochemistry and Molecular Biology, Mwanza, Tanzania
5. Kumasi Center for Collaborative Research in Tropical Medicine, Kumasi, Ghana
6. University of Jos, Department of Pharmaceutical Microbiology and Biotechnology, Plateau State, Nigeria
7. African Society for Bioinformatics and Computational Biology, Cape Town, South Africa

Background:
Ubiquitin-like with PHD and RING Finger domains 1 (UHRF1) is an important epigenetic regulator that plays a pivotal role in modulating DNA methylation patterns and chromatin structure. Its role in cancer progression is significant, yet the extent of its impact on gene expression across different types of cancer remains poorly understood. This study focuses on a comparative transcriptomic analysis of UHRF1 expression knockout in four cancers—two solid tumors and two hematological cancers—following UHRF1 expression knockout.

Methods:
To understand the role UHRF1 plays in different cancers, we obtained transcriptomic data from NCBI Gene Expression Omnibus (GEO) database and performed comparative differential gene expression and gene set enrichment analysis of four distinct cancer cell lines: retinoblastoma (Y79), breast cancer (MCF-7), acute myeloid leukaemia (Kasumi-1) and acute monocytic leukaemia (THP-1).

Results:
Among the differentially expressed genes of each cancer cell line, 80 genes were commonly regulated following UHRF1 expression knockout. Gene set enrichment analysis of the 80 overlap genes revealed 179 GO terms significantly enriched, suggesting roles in processes like response to hypoxia, regulation of Wnt signaling pathway, and positive regulation of apoptotic processes with no KEGG pathways found to be enriched. Protein-protein interaction analysis of the 80 overlapping genes revealed a network with 79 nodes and 42 edges, identifying key hub genes *GLUL, SOD2, GPI, TXNDR1*, and *HSPD1*, with *GLUL* and *SOD2* exercising strong functional interactions.

Conclusion:
Our comparative transcriptomic analysis reveals eighty (80) shared differentially expressed genes across solid and hematological cancers, with *GLUL* and *SOD2* being key hub genes with strong functional interactions. Furthermore, our findings suggest that UHRF1 expression knockout leads to significant enrichment in processes such as the response to hypoxia, regulation of the Wnt signaling pathway, and apoptosis, thereby improving our understanding of UHRF1's critical role in pathways essential to cancer progression. This indicates that UHRF1 knockout influences vital cellular functions, potentially identifying novel targets for cancer therapy.

Cancer Omics

POSTER 116

Systematic discovery of non-coding driver mutations in evolutionarily constrained regions: a pan-cancer analysis

Kerstin Lindblad-Toh[1,2,3], Suvi Mäkeläinen[1,2,4], Raphaela Pensch[1,2], Sergey V Kozyrev[1,2], Anna Darlene van der Heiden[1,2], Eric Pederson[1,2], Firoj Mahmoud[1,2], Ananya Roy[1,5], Åsa Karlsson[1,2], Sharadha Sakthikumar[1,2], Mats Pettersson[1,2], Eric S Lander[3], Maja-Louise Arendt[1,2,6], Karin Forsberg-Nilsson[2,4,7]

1. Uppsala University, Department of Medical Biochemistry and Microbiology, Uppsala, Sweden
2. Uppsala University, SciLifeLab, Uppsala, Sweden
3. Broad Institute of MIT and Harvard, Cambridge, MA, USA
4. Swedish University of Agricultural Sciences, Department of Animal Biosciences, Uppsala, Sweden
5. Uppsala University, Department of Immunology, Genetics and Pathology, Uppsala, Sweden
6. University of Copenhagen, Department of Veterinary Clinical Sciences, Copenhagen, Denmark
7. University of Nottingham Biodiscovery Institute, Division of Cancer and Stem Cells, Nottingham, United Kingdom

Cancer is a collective term for approximately 200 diseases, caused by mutations facilitating the replicative and metastatic potential of the cancer cells. Genomic studies have identified hundreds of genes with protein-coding mutations driving the cancer phenotype. The non-coding genome has been less studied but is likely to also contain cancer driver mutations. We analyzed data from 2,539 cancer genomes representing 33 cancer types from the International Cancer Genome Consortium (ICGC). We intersected the non-coding regions of the genome with Zoonomia phyloP scores, an evolutionary constraint metric from 240 mammalian species, to identify non-coding constraint mutations (NCCMs) with regulatory potential.

We developed a statistical framework that models the distribution of NCCMs in regulatory regions up to 100 kbp up- and downstream of each gene, controlling for the background mutation rate and other confounders. This identified 14 genes with significant enrichment of NCCMs across cancer types. These include excellent cancer candidate genes such as *FOXA1*, *TRIM27*, and *RBM4*. Analysis of individual cancer types identified an average of 5 significant genes per subtype and a total of 77 candidates, both overlapping with the pan-cancer analysis and individual subtype-specific genes. The results show that genes with a high NCCM burden often have few protein-coding mutations, suggesting that these proteins are likely required to be intact for cells to be functional.

In addition, we performed a genome-wide 2 kbp sliding-window analysis that identified 45 significant windows in the vicinity of 27 unique genes. Among the NCCMs in significant windows, we further identified mutational clusters where the same NCCM or NCCM within 15 bp regions were shared by multiple patients, sometimes within, sometimes across tumor types. These include both previously seen and novel hotspots, including those near *RMRP*, *ALPK2*, and *FSHR*.

Overall, we identified both known and novel non-coding candidates, and the identified genes were significantly enriched for COSMIC Cancer Gene Census genes. The NCCMs were associated with significant gene expression differences, and we performed massively parallel reporter assays to examine the functional effects of the NCCMs. We suggest the importance of novel driver mutations and driver genes for cancer development.

Reframing RNA-seq as a clinically actionable tool in pediatric precision oncology

Cancer Omics

POSTER 117

Chelsea Mayoh[1,2], Paulette Barahona[1], Angela Lin[1], Lujing Cui[1], Pamela Ajuyah[1], Ann Altekoester[1], Loretta MS Lau[1,2,3], Dong-Anh Khuong-Quang[4,5], Vanessa Tyrrell[1,2], Michelle Haber[1,2], Marie Wong-Erasmus[1], Paul G Ekert[1,2,6], Mark J Cowley[1,2]

1. University of New South Wales (UNSW) Sydney, Lowy Cancer Centre, Children's Cancer Institute, Kensington, New South Wales, Australia
2. University of New South Wales (UNSW) Sydney, UNSW Medicine & Health, School of Clinical Medicine, Kensington, New South Wales, Australia
3. Sydney Children's Hospital, Kids Cancer Centre, Randwick, New South Wales, Australia
4. Royal Children's Hospital, Children's Cancer Centre, Parkville, Victoria, Australia
5. Royal Children's Hospital, Murdoch Children's Research Institute, Parkville, Victoria, Australia
6. Peter MacCallum Cancer Centre, Cancer Immunology Program, Parkville, Victoria, Australia

Despite its widespread availability, RNA-sequencing (RNA-seq) remains underutilised in clinical genomics, often dismissed as non-actionable due to its association with gene expression profiling. This perception has led to an overreliance on DNA sequencing (either targeted, whole exome or whole genome), which, while foundational, fails to capture the functional consequences of many variants. In Australia's national precision medicine program, the ZERO Childhood Cancer Program, we challenge this paradigm by demonstrating the indispensable clinical utility of RNA-seq in paediatric precision oncology.

Across 477 high-risk paediatric cancer samples, we integrated RNA-seq with whole genome sequencing (WGS) to resolve the expressed impact of structural variants (SVs) and single nucleotide variants (SNVs) and small insertions and deletions (Indels). RNA-seq identified 93% of SVs, including complex rearrangements such as multi-hop events, chromothripsis, and intragenic disruptions, often undetectable or uninterpretable by WGS alone. Critically, RNA-seq informed frame calling and protein domain retention analysis, providing functional context essential for assessing oncogenicity and therapeutic relevance. Notably, RNA-seq was essential to the clinical interpretation of 30% of SVs, directly influencing diagnostic and treatment decisions.

Beyond SVs, RNA-seq confirmed 90% of driver SNVs/Indels and reclassified pathogenicity of 22% based on allele-specific expression and splicing outcomes. Importantly, RNA-seq identified clinically relevant SNVs in low purity tumours and SNVs/Indels with low variant allele frequency, in cases where WGS sequencing had failed to detect any drivers. These transcriptional consequences directly impacted variant pathogenicity and treatment decisions, underscoring RNA-seq's role in refining clinical interpretation.

Our findings dismantle the notion that RNA-seq is limited to expression analysis and highlight its power to functionally annotate DNA variants. In paediatric cancers, where mutation burden is low and structural complexity is high, RNA-seq bridges the gap between detection and actionability. We advocate for a shift in clinical genomics where RNA-seq should no longer be viewed as supplementary, but as a core component of multi-omic diagnostics. Its ability to reveal the expressed consequences of DNA alterations makes it a vital tool for precision medicine, particularly in contexts where WGS alone falls short.

Dissecting intratumoral heterogeneity in GIST using spatial transcriptomics

Cancer Omics

POSTER 118

Leonardo A. Meza-Zepeda[1,2], Andrea Terrasi[1], Anine A. Ottestad[1], Susanne Lorenz[2], Ana B. Lid[2], Ingvild V.K. Lobmaier[3], Kjetil Boye[4]

1. Oslo University Hospital, Department of Tumor Biology, Oslo, Norway
2. Oslo University Hospital, Genomics Core Facility, Oslo, Norway
3. Oslo University Hospital, Department of Pathology, Oslo, Norway
4. Oslo University Hospital, Department of Oncology, Oslo, Norway

Gastrointestinal stromal tumours (GISTs) are characterised by driver mutations in the KIT and PDGFRA genes, leading to the abnormal activation of these receptor tyrosine kinases. Approximately 40% of high-risk patients experience relapse after adjuvant imatinib therapy, emphasising the importance of better understanding tumour biology and heterogeneity, and their influence on prognosis and resistance.

We have applied Visium 10x Genomics spatial transcriptomics technology on FFPE samples from nine treatment-naive GISTs. Our approach integrates deconvolution analysis to delineate distinct spatial transcriptomic niches and to identify specific cell types within the tumour and its microenvironment. Preliminary analyses focused on characterising transcriptomic profiles within the tumour area and across patients, supplemented by inferred DNA copy number changes from gene expression data.

The inclusion of normal tissue was essential for enhancing our analyses and defining cell types within the tumour microenvironments. Our findings revealed significant intratumour heterogeneity, with distinct spatial domains displaying unique transcriptional landscapes. Despite the uniformity of KIT mutations among patients, we observed substantial variability in gene expression profiles, suggesting potential differences in the activation of downstream pathways. Additionally, areas of localised chromosome copy number changes were linked to signatures associated with more aggressive disease behaviour.

This study highlights the complex transcriptional and genomic heterogeneity in GISTs, which may have significant implications for patient stratification and the development of personalised therapeutic strategies. A deeper understanding of the molecular mechanisms driving heterogeneity in GISTs will be crucial for enhancing targeted therapies and reducing the risk of relapse in high-risk patients.

Influence of ABC gene variants and patient ancestry on adverse reactions induced by carboplatin and paclitaxel in patients with lung cancer

Cancer Omics

POSTER 119

Patricia Moriel[1], Yasmim Gabriele Matos[2], Cecilia Souto Seguin[2], Giovana Fernanda Santos Fidelis[2], Julia Tiemi Siguemoto[2], Maurício Wesley Perroud Jr.[2], Carolina Luchesi Canizares[1], Julia Camargo Lepore[1], Eder de Carvalho Pincinato[2]

1. Universidade Estadual de Campinas, Faculty of Pharmaceutical Sciences, Campinas, São Paulo, Brazil
2. Universidade Estadual de Campinas, School of Medical Sciences, Campinas, São Paulo, Brazil

Lung cancer, one of the leading malignancies worldwide, is commonly treated with the combination of carboplatin and paclitaxel (CBPX). Although effective, this regimen is often limited by adverse drug reactions (ADRs), which may be influenced by genetic variants (SNVs) and patient ancestry. This study aimed to investigate the impact of SNVs in ATP-binding cassette (ABC) transporter genes and ancestry on chemotherapy-induced ADRs in individuals with non-small cell lung cancer (NSCLC). Patients with NSCLC were recruited at the Hospital de Clínicas, State University of Campinas (UNICAMP), and evaluated for ADRs, SNVs, and ancestry after the first CBPX chemotherapy session. ADRs were classified according to severity using the Common Terminology Criteria for Adverse Events (version 4.03). SNVs in the ABCB1, ABCC1, and ABCC2 genes with a ClinPGx evidence level ≥3 was selected and genotyped via microarray. Only SNVs in Hardy-Weinberg equilibrium were analyzed for the dominant (DM) and recessive (RM) models. The study was approved by the UNICAMP Research Ethics Committee. 109 patients were included, with a mean age of 63.1 years; 52.3% were male, 82.6% self-reported as white, and 40.4% had European ancestry. The most frequent ADRs were anemia (34.8%), nausea (26.6%), and reduced creatinine clearance (CClr; 25.9%), with grade 1 prevalence. Significant associations were observed between specific ADRs and SNVs: vomiting and ABCB1's rs9282564 (DM; p=0,007 OR=6,04; CI95%=1,63-22,42), rs1128503 (RM; p=0,029; OR=0,178; CI95%=0,04–0,84), rs4148738 (RM; p=0,038; OR=0,193; CI95%=0,041-0,910), rs4728709 (RM; p=0,015) and ABCC1's rs45511401 (DM; p=0,027; OR=6,205; CI95%=1,22-31,43); ABCB1's rs10267099 with thrombocytopenia (DM; p=0,003; OR=0,126; CI95%=0,032-0,4890); ABCC2's rs3740066 with reduced CClr (RM; p=0,034; OR=3,476; CI95%=1,096-11,03), and rs2273697 with neutropenia (DM; p=0,005). Ancestry influenced variant prevalence: admixed individuals were less likely to carry rs3740066 (ABCC2; p=0.024; OR=0.206; 95%CI=0.052–0.808), whereas African ancestry was associated with increased likelihood of rs9282564 (ABCB1; p=0.041; OR=8.540; 95%CI=1.092–66.851). Overall, these findings suggest that ABC transporter gene variants modulate susceptibility to CBPX-induced ADRs and that ancestry shapes SNV distribution. Integrating pharmacogenetic and ancestry-informed approaches may enhance chemotherapy personalization, improve efficacy and minimize ADRs.

Cancer Omics

POSTER 120

A scalable platform for single-cell co-profiling of the transcriptome and genotype

Juozas Nainys[1], Rapolas Zilionis[1], Andrius Sinkunas[1], Paulius Matulis[1], Jokubas Tamoliunas[1], Patrick Roelli[1]

1. Atrandi Biosciences, East Amherst, NY, USA

Single-cell RNA sequencing (scRNAseq) has become a routine tool for profiling cellular composition and transcriptional responses across health and disease. Beyond transcriptomic analysis, scRNAseq has also been used to detect expressed genetic variants and infer perturbation identity in CRISPR–Cas-based high-throughput screens through guide RNA sequencing. However, these transcriptome-based measurements face several limitations. They are restricted to transcribed regions, depend on sufficient expression levels to overcome the inherent sparsity of scRNAseq data, and, in the context of genetic perturbations, the detection of a guide RNA transcript does not confirm successful editing at the target locus.

To address these limitations, we developed a high-throughput single-cell RNA & DNA co-assay based on our Semi-Permeable Capsule (SPC) technology, enabling highly parallel, multistep processing of single cells. The method couples whole-transcriptome profiling with targeted genotyping by multiplex PCR amplicon sequencing in the same cells, directly linking genotype to transcriptional state, confirming CRISPR edits at genomic targets, and functionally characterizing engineered or naturally occurring mutations.

We profiled >100,000 primary human cells across multiple peripheral blood mononuclear cell (PBMC) donors using co-sequencing of RNA and amplicons targeting 10 SNP-containing loci. Over 85% of captured cells yielded genotypes, enabling donor deconvolution from variant calls in the amplicons and robust mapping of genetic background to cell-state heterogeneity. Designed for scalability, the assay supports user-defined PCR panels targeting transcribed and non-transcribed loci and is readily extensible in both cell numbers and amplicon breadth. We will present results demonstrating the utility of the platform for dissecting genotype–phenotype relationships at single-cell resolution.

Cancer Omics

POSTER 121

Tahoe-100M: a 100 million single-cell atlas enabling AI-driven cancer drug discovery

Joey Pangallo[1], Vuong K Tran[1], Efthymia Papalexi[1], Simone Marrujo[1], Bryan Hariadi[1], Crina Curca[1], Olivia Kaplan[1], Sarah Schroeder[1], Ajay Sapre[1], Guillermo Gallareta Olivares[1], Maria Nigos[1], Oliver Sanderson[1], Hoai Nguyen[1], Alec Salvino[1], John Thompson[1], Ryan Koehler[1], Aisling Sinclair[1], Gokhan Demirkan[1], Charles Roco[1], Alexander Rosenberg[1]

1. Parse Biosciences, Seattle, WA, USA

Machine learning is transforming the process for drug discovery, but to train powerful predictive models, we need datasets that capture how thousands of compounds impact diverse human cell types. Traditional approaches face a major challenge: batch effects and variability across experiments can mask true biological signals. Recent advances in single cell RNA sequencing (scRNA-seq) using combinatorial barcoding now make it possible to measure transcriptomes from tens of millions of cells in a single, unified workflow—dramatically reducing technical noise while preserving biological detail.

Building on this capability, we created Tahoe-100M, the largest single cell dataset ever assembled for drug discovery: over 100 million cells, spanning 50 human cell line models and 379 drug compounds. Mixed cell lines from Tahoe Therapeutics were co-cultured, grown into 3D spheroids, and treated for 24 hours with compounds across three concentrations. Following treatment, cells were fixed and pooled in batches of ~10 million for barcoding on the Parse Biosciences GigaLab platform. Automated library prep across 96-well plates generated sequencing-ready material for all samples, which were sequenced on the UG 100 (Ultima Genomics). Demuxlet analysis then linked each cell back to its original line, producing high-quality transcriptomes at unprecedented scale.

The resulting Tahoe-100M dataset covers ~56,000 unique cell line–drug–dose combinations and reveals thousands of genes with clear, dose-dependent expression changes. Stratifying these responses by mutational profile uncovers genotype-specific transcriptional programs, while single cell phase analysis exposes compound specific cell cycle effects—such as G1 or G2/M arrest by CDK inhibitors and G2/M accumulation after microtubule inhibition—that bulk assays cannot resolve.

By processing fixed cells in massive pooled batches, we minimized batch effects, enabling direct comparisons across all perturbations. Together, Tahoe-100M provides a scalable blueprint for atlas-scale scRNA-seq and an open resource to accelerate AI-guided cancer drug discovery and genotype-specific response profiling across human cell models.

Cancer Omics

POSTER 122

Evaluating the integration of OvaPrint cfDNA methylation profiling with proteomics for adnexal mass assessment

Bodour Salhia[1,2], Ben Yi Tew[1], Gerald Gooden[1], Lynda Roman[2,3], X. Mona Guo[2,3], Monica Neuman[4], Emma Barber[5], Elisa Romeo[6], Todd Holscher[6], Renee Wunderley[6], Danielle Goldberg[7], Krishna Morampudi[7], Dalia Daujotyte[7], Monique Spillman[8], Thomas Simon[1]

1. University of Southern California, Keck School of Medicine, Department of Cancer Biology, Los Angeles, CA, USA
2. University of Southern California, Norris Comprehensive Cancer, Los Angeles, CA, USA
3. University of Southern California, Division of Gynecologic Oncology, Los Angeles, CA, USA
4. Los Angeles General Medical Center, Los Angeles, CA, USA
5. Northwestern University, Feinberg School of Medicine, Division of Gynecologic Oncology, Evanston, IL, USA
6. CpG Diagnostics, Inc., Pasadena, CA, USA
7. Illumina, Inc., San Diego, CA, USA
8. University of Arkansas Medical Center, Department of Gynecologic Oncology, Little Rock, AR, USA

Background:
Patients with adnexal masses face both overtreatment, when benign disease is misclassified, and undertreatment, when cancers are not managed by gynecologic oncologists. CA125, the most widely used biomarker, lacks specificity. High-grade serous ovarian cancer (HGSOC), the most aggressive subtype of epithelial ovarian cancer (EOC), requires timely specialist management. OvaPrint, a cfDNA methylation liquid biopsy developed to improve adnexal mass evaluation, is currently optimized for HGSOC, with expansion underway. In this prospective multi-site study, we sought to clinically verify OvaPrint, compare it with CA125 and complement it with next generation proteomics.

Methods:
Eligibility included adnexal mass on imaging without evidence of metastatic disease, planned surgery, no prior treatment, and no cancer history. Plasma was collected from six US sites preoperatively and analyzed with OvaPrint. CA125 values were abstracted from medical records. Final surgical pathology was the clinical ground truth. Of 121 cases (7 HGSOC, 16 other EOCs, 5 other cancers, 93 benign), 110 were evaluable (104 had CA125). Endpoints were sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Separately, 80 samples are also being assessed by the Illumina's Protein Preparation (IPP) platform detecting 9500 proteins.

Results:
Median age was 56. OvaPrint had 83.3% sensitivity for HGSOC, 96.6% benign specificity and 97.12% non-HGSOC specificity. Its HGSOC-specific PPV was 62.5% and NPV 99%. CA125 showed 100% sensitivity for HGSOC but had poor specificity (62.2%) and 56.3% sensitivity for other cancers. Its HGSOC-specific PPV was 13%. CA125 specificity differed by age and race while OvaPrint specificity did not. IPP is being evaluated to complement OvaPrint in the discrimination of malignant and benign adnexal masses.

Conclusions:
Accurate tools for EOC diagnosis are urgently needed. This first prospective, multi-site study shows OvaPrint's exceptional specificity. The high HGSOC-specific NPV and PPV of OvaPrint means a positive result is a near-certain indicator of HGSOC, supporting timely and appropriate care. OvaPrint is being expanded to other EOCs and is undergoing CLIA validation. Additional integration of other molecular endpoints could further enhance discrimination of adnexal masses. These findings establish it as the first liquid biopsy in ovarian cancer with potential to transform diagnostic practice and outcomes.

Cancer Omics

POSTER 123

Dissecting high-grade serous ovarian cancer metastasis through single-cell and spatial transcriptional entropy profiling

Robert Sebra[1-4,7], Erika Nemeth[1,2], Kristin Beaumont[1,2], Nadine Schrode[1,2], Olivia Scharfman[1], Sabrina Petri[1,2], Sanjana Shroff[1,2], Maria Padron Rhenals[1,2], Peter Dottino[5], John Martignetti[6]

1. Icahn School of Medicine at Mount Sinai, Department of Genetics and Genomic Sciences, New York, NY, USA
2. Icahn School of Medicine at Mount Sinai, Center for Advanced Genomics Technology, New York, NY, USA
3. Icahn School of Medicine at Mount Sinai, Black Family Stem Cell Institute, New York, NY, USA
4. Icahn School of Medicine at Mount Sinai, Tisch Cancer Institute, New York, NY, USA
5. Yale University School of Medicine, Department of Obstetrics, Gynecology, and Reproductive Sciences, New Haven, CT, USA
6. MDDx Inc., Tarrytown, NY, USA
7. Panacent Bio, Inc. New York, NY, USA

Ovarian cancer is the fifth leading cause of cancer-related deaths among females. High-grade serous ovarian cancer (HGSOC) is the most common and aggressive subtype of ovarian cancer, with >75% of cases diagnosed at an advanced stage with metastasis. Despite improvements in cytoreductive surgery and platinum-based chemotherapy, recurrence is nearly universal, and most patients ultimately develop resistance to standard treatment. Consequently, HGSOC remains highly lethal, with a five-year survival rate <30%. Prior studies have suggested that tumor-initiating cells (TICs) drive tumor growth and may contribute to therapeutic resistance in multiple cancer types, including HGSOC. However, there is a lack of reliable markers of TICs across tumor sites due to tumor microenvironmental (TME) differences, and the ways in which TICs participate in the TME remains poorly defined.

In this study, we aim to evaluate and improve existing TIC markers as well as characterize transcriptional entropy in the tissue landscape among single cells profiled across HGSOC metastatic sites. Transcriptional entropy, a quantitative measure of transcriptional diversity within individual cells, provides a framework to capture cellular plasticity that may underlie metastatic potential of TICs. We performed multi-site single cell RNA-sequencing (scRNA-seq) on >100,000 cells across from five HGSOC patients, as well as orthogonal Visium-HD spatial transcriptomics on paired ovary and omental tumors from matched patients. scRNA-seq data support differences in gene expression profiles across patients and anatomic sites, consistent with the inter-patient and inter-site heterogeneity reported in HGSOC literature. Initial findings also reveal differences in levels of transcriptional entropy among epithelial cells between primary and metastatic tumor sites.

Leveraging our previously published Unicell AI/ML single-cell foundational model and entropy scoring algorithm, pySCE, we anticipate identification of gene expression programs that correlate with cellular transcriptional entropy, which will be validated via analysis of high-powered integrated datasets and via perturbation of HGSOC patient-derived cells in culture. The Unicell model integrates millions of single-cell profiles to contextualize HGSOC data, while pySCE computes entropy-adjusted transcriptional signatures predictive of cellular state transitions. Collectively, this work will enhance understanding of the transcriptomic and spatial dynamics of HGSOC metastasis and identify novel transcriptional entropy-informed candidates for treatment of metastatic HGSOC.

Cancer Omics

POSTER 124

From variability to verifiability: benchmarking gene expression in ATCC's human and mouse cell lines

Ajeet P. Singh[1], Rula R. Khairi[1], Amy L. Reese[1], Jade Kirkland[1], Noah Wax[1], Steve King[1], James Duncan[1], Robert Marlow[1], Jonathan L. Jacobs[1]

1. American Type Culture Collection (ATCC), Manassas, VA, USA

Reproducibility remains a significant challenge in biomedical research, especially when working with in vitro models such as immortalized cell lines. Variations in laboratory protocols, culture conditions, passage number, and experimental workflows can lead to substantial changes in gene expression, often resulting in inconsistent or contradictory findings. This variability undermines the reliability of data and impedes progress in both fundamental research and translational efforts.

To help address this issue, we established a comprehensive reference transcriptome for human and mouse cell lines originally sourced and maintained by ATCC. Through high-quality RNA sequencing and standardized bioinformatics workflows implemented within QIAGEN OmicSoft's ATCC Cell Line Land and ATCC Genome Portal, we produced reproducible, well-annotated gene expression profiles across a wide array of cell lines representing diverse tissues, disease contexts, and research applications.

This dataset provides a robust benchmark for evaluating gene expressions commonly used in vitro models. It supports validation of experimental results, verification of cell line identity, detection of anomalies due to contamination or misidentification, and identification of cell line–specific biomarkers to inform model selection and experimental design. By offering a consistent reference framework, the transcriptome enhances cross-study comparability, enables integration of datasets from multiple sources, and reinforces data quality. This resource promotes reproducibility and accelerates the translation of cell-based research into meaningful biomedical advances.

Our goal in delivering this reliable and accessible transcriptomic resource is to support the scientific community in moving from variability toward verifiability, advancing reproducible research, strengthening collaboration across sectors, and ultimately contributing to improved patient outcomes.

Cancer Omics

POSTER 125

Expanding the landscape of biomarker discovery in a multi-cancer cohort

Donald Skifter[1], Danielle Gutierrez[1], Jessica L. Moore[1], Jaison Arivalagan[1], Tiffany Louie[1], Kyra Richardson[1], Yukta Bhurke[1], Alejandra Stihel[1], Cristian Zimbru[1], Sara Hildebrand[1], Dragana Noe[1], Lina Hultin Rosenberg[2], Jenny Samskog[2], Ryan Lamers[2], Khatereh Motamedchaboki[3], Samad Jahandideh[3]

1. Discovery Life Sciences, Huntsville, AL, USA
2. Olink, part of Thermo Fisher Scientific, Waltham, MA, USA
3. Thermo Fisher Scientific, San Francisco, CA, USA

Advancements in proteomic technologies empower the identification of potential biomarkers from serum and plasma in a scalable manner, providing more comprehensive insight into health and disease states. Approximately 400 plasma samples derived from donors with Stage IV prostate, lung, or colorectal cancer and undiagnosed donors were profiled using high throughput mass spectrometry (MS) and Olink® Explore HT. Proteomic and genomic analysis of matched plasma and FFPE tissue from an additional 80 donors with Stage I or IV of the above cancers and normal adjacent tissues will be performed (Discovery Life Sciences). For mass spectrometry analysis, plasma was enriched using the Seer Proteograph XT workflow with a Thermo Scientific Vanquish Neo HPLC system coupled to an Orbitrap Astral MS. A targeted method was developed using PRM Conductor to transfer discovery findings to targeted analysis on the Thermo Stellar MS. FFPE tissue samples will be processed for whole exome sequencing and MS analysis. The latter will utilize the PreOmics BeatBox and iST kits. MS-based analysis measured an average of 6,600 proteins per sample. In total, greater than 11,000 proteins were measured across all samples by Olink Explore HT and MS. The two technology platforms had ~30% overlap, providing high complementarity, and use of both platforms increased coverage of the druggable proteome. In addition, 424 proteins were measured in common with an analytically driven Stellar MS targeted panel. Significantly changed proteins found in common between Olink Explore HT and MS analysis, and known to be overexpressed in cancer, had similar fold changes across the two platforms. State-of-the-art instrumentation and technology platforms expanded biomarker discovery and linked results with immediate targeted quantification. Proteomic and genomic analysis of matched donor tissue and plasma samples will further drive understanding of the relationship between tissue and soluble biomarkers.

PaintScapeTM enables tissue based unique multiomic protocol for direct in-situ visualization of single cell spatial 3D genome organization of fresh frozen Glioblastoma cells in tumor microenvironments

Cancer Omics

POSTER 126

Matija Snuderl[1], Yuxiu Wang[1], Neha Dhasmana[1], Pamela Flatley[2], Stacy Elliott[2], Doug Werner[2], Jude Dunne[2], Huy Nguyen[2] and Shyamtanu Chattoraj[2]

1. Department of Pathology, NYU Grossman School of Medicine, New York, New York, USA
2. Bruker Spatial Genomics, Research and Development, San Jose, California, USA

Glioblastoma (GBM), the most aggressive form of brain tumor, is characterized by high intratumoral molecular heterogeneity caused by clonal and sub-clonal variability in genomic, epigenomic, transcriptional, and proteomic networks of cancer cells in heterogenous different tumor microenvironments. The 3D organization of genome structure plays a critical role in regulating gene expression and cellular function. Structural aberrations including copy number variations (CNVs), structural variations (SVs), and topological changes rewire the 3D genome allowing aberrant activation of oncogenes, repression of tumor suppressors and dysregulation in gene expression. Although current methodologies such as WGS or Hi-C provide bulk-level understanding of dysregulation, they do not provide insight into spatial distribution and association between molecular subclones and microenvironment. Since tumor cell heterogeneity contributes towards therapeutic resistance and hampers development of successful treatment strategies for GBM patients, understanding the molecular mechanism in a heterogenous spatial context demands state-of-the-art in-situ single cell and spatial technologies that can provide deeper insight into early detection and disease progression.

Here, we present a novel multiomic jebFISH™ protocol on the PaintScape™ platform that can be used to analyze intratumoral heterogeneity of 3D chromatin architecture of fresh frozen GBM cells in different tissue microenvironments, perinecrotic, perivascular, infiltrative, at single cell, sub-population and population level. Using the ChromoPaint™ PanChromo MPX Panel, which visualizes loci across all chromosomes, we show disruption in 3D folding of individual chromosomes such as re-organization of chromosome territory, differential p-q arm interaction and inter-chromosomal proximity at single cell and sub-population level. We characterize chromosomal instability, CNVs and SVs of different GBM cancer cell states at sub-chromosomal level in-situ within the tissue. By combining jebFISH with multiplexed immunofluorescence proteomic detection of hypoxia and microvasculature, we show how different tissue microenvironment e.g. hypoxic vs normoxic regions within the same GBM tissue section have different 3D genome organization in the cancer cell sub-populations.

We envision the PaintScape system will improve our understanding of the progress of glioblastoma disease and provide deeper insight into the underlying 3D genomic heterogeneity of single cancer cells and sub-populations with unique spatial context within the tissue leading to design of better cancer treatment options in the future.

Cancer Omics

POSTER 127

Comparing whole-genome bisulfite sequencing to a novel chemistry for DNA methylation analysis

Ben Yi Tew[1], Thomas Simon[1], Gerald Gooden[1], Seeta Rajpara[1], Danielle Goldberg[2], Ev Chekalin[2], James Baye[2], Lisa Watson[2], Suzanne Rohrback[2], Bodour Salhia[1,3]

1. University of Southern California, Keck School of Medicine, Department of Cancer Biology, Los Angeles, CA, USA
2. Illumina, Inc., San Diego, CA, USA
3. University of Southern California, Norris Comprehensive Cancer, Los Angeles, CA, USA

Background:
Whole-genome bisulfite sequencing (WGBS) is the gold standard for base-resolution DNA methylation analysis; however, its reliance on converting unmethylated cytosines to thymines reduces nucleotide diversity, leading to lower mapping efficiency and limited ability to detect genomic variation, which makes FFPE or cfDNA-derived data analysis especially challenging. The Illumina 5-base sequencing solution employs a novel single-step enzymatic conversion that converts only methylated cytosines, preserving sequence complexity while enabling simultaneous detection of methylation and genomic variants. In this study, we compared the performance of WGBS to the Illumina 5-base solution across sequencing metrics, methylation detection, and variant calling in fresh frozen, FFPE and cfDNA samples.

Methods:
Whole-genome DNA methylation and variant detection were evaluated using the Illumina 5-base solution and compared to established assays: WGBS or Illumina EPIC arrays for methylation and whole-exome sequencing (WES) or whole-genome sequencing (WGS) for SNV and CNV detection. The analysis included 41 fresh frozen samples (29 tumors and 12 matched normals, representing 22 unique cases), 24 FFPE samples with 5-base, WGBS, EPIC arrays WES, and WGS data, and 49 cfDNA samples with 5-base methylation data and targeted bisulfite sequencing data. We compared the alignment rate, duplication, CpG coverage, and conversion efficiency between platforms. 5-base comparison to arrays & WGBS analysis included differential methylation analysis and individual CpGs methylation correlation analysis. Genomic findings (small variants & CNV) were compared to WGS & WES calls.

Results:
Using fresh frozen samples, Illumina 5-base achieved higher alignment (91.6% vs. 73.4%) and lower duplication rates (12.4% vs. 29.0%) compared to WGBS, producing more complete methylomes with greater CpG depth and reduced noise. CpG-level methylation profiles demonstrated strong correlation between platforms ($\Delta\beta > 0.25$ in only 1.2% of CpGs; > 0.5 in 0.1%). Illumina 5-base identified a larger number of DMRs, reflecting enhanced CpG coverage, while shared DMRs revealed concordant biological pathways. Additionally, 5-base data yielded robust CNV and SNV detection comparable to WGBS and WES.

Conclusions:
The Illumina 5-base solution delivers highly concordant methylation profiles with greater coverage, improved efficiency, and simultaneous detection of genetic variants, representing a viable and dual insight alternative to existing methods for comprehensive methylome analysis.

Smarter **Solutions.**
Bigger **Breakthroughs.**

BRAIN BOOST

Espresso's on us!
Swing by and power up!

Tuesday | 10:00am – 12:00pm
Wednesday | 1:00pm – 3:00pm

PRESENTATION

From bottlenecks to breakthroughs: Unlocking high-quality WGS at scale

Wednesday | 3:20pm

POSTERS

TAPS+ paired with methyl-fragment fractionation enables high-sensitivity tissue of origin analysis

Rapid, PCR-free WGS with streamlined normalization for time-sensitive and stringent applications

Challenging samples? Challenge accepted.

Maximize insights from FFPE and liquid biopsy samples with high-efficiency DNA and RNA solutions

One library. Layers of insight.

Unlock integrated genomic and methylation insights from a single library with TAPS+

Let's break bottlenecks.

Generate data at-scale with novel workflows, vetted automation scripts, and pack sizes designed to support high throughput operations

VISIT US IN **MANATEE**

WATCHMAKER GENOMICS

Scan to learn more or visit
watchmakergenomics.com/agbt2026

ABSTRACT CATEGORY 02

Informatics & Computational Biology

Informatics & Computational Biology

POSTER 201

Deep long-read sequencing uncovers subclonal structural variants in LVAD-colonizing Pseudomonas aeruginosa

Abigail Bosco[1], Tiffany Luong[2], Maryam Ahmadi Jeshvaghane[1], Mahshid Fallahpour[1], Samuel Modlin[1], Dwayne Roach[2], Faramarz Valafar[1]*

1. San Diego State University, Laboratory for Pathogenesis of Clinical Drug Resistance and Persistence, School of Public Health, San Diego, CA, USA
2. San Diego State University, Department of Biology, San Diego, CA, USA

Background:
Pseudomonas aeruginosa is a Gram-negative opportunistic pathogen frequently associated with multidrug-resistant infections in clinical settings. The rise of antibiotic-resistant strains and their genomic complexity necessitate advanced tools for comprehensive genetic characterization, particularly in the context of resistance to antibiotic and phage therapies.

Methods:
Here, we use ultra-deep high-fidelity (HiFi) sequencing to analyze strain NRD619, a multidrug-resistant P. aeruginosa isolate that had colonized the driveline of a left ventricular assist device of an 82-year-old male who was identified as a phage therapy candidate. Through HiFi long-read sequencing and bioinformatic analysis, we identified low-frequency subpopulations harboring heterogeneous structural variants (SVs) and investigated their potential phenotypic consequences through functional genomics.

Results:
Small variants inducing frameshifts linked to efflux pump overexpression (mexZ) and missense mutations linked and reduced carbapenem uptake (oprD), and other antibiotic resistance (ampC and mexS) clarified the genetic basis of drug resistance in NRD619. Ultra-deep long-read sequencing captured five low-frequency subpopulations (allele frequency as low as 0.024%) harboring heterogeneous SV with clinically important phenotypic implications. These comprised a reciprocal heterogeneous 32kb deletion and duplication of a cluster of phage proteins and three multi-kilobase insertions. The three insertions affected genes mediating several bacterial functions including fimbriae and an ABC transporter permease, implicating biofilm formation and colonization of medical devices as differential phenotypes in the minority subpopulation.

Conclusions:
This study demonstrated how ultra-deep sequencing can uncover cryptic genetic heterogeneity in bacterial pathogens, offering insights into P. aeruginosa subpopulation composition in clinical isolate unobservable by traditional approaches. Future applications could refine our understanding of multidrug resistance, phage susceptibility profiles, and adaptations influencing host colonization with clinical ramifications. These findings spotlight a strategy to study bacterial adaptation dynamics in response to clinical therapies and potentially inform personalized treatment approaches in certain clinical settings.

Keywords:
Pseudomonas aeruginosa, multidrug resistance, HiFi sequencing, structural variants, heterogeneous variants, efflux pumps, antibiotic resistance

Informatics & Computational Biology

POSTER 202

Tissue and cellular spatiotemporal dynamics in colon aging

Aidan C. Daly[1,2], Francesco Cambuli[1,3], Tarmo Äijö[2], Britta Lötstedt[1,4,5], Nemanja Despot Marjanovic[4,6], Sara Fernandez[1], Olena Kuksenko[4,7], Matthew Smith-Erb[1], Daniel Domovic[1], Nicholas Van Wittenberghe[4], Eugene Drokhlyansky[4], Gabriel K Griffin[4,8], Hemali Phatnani[1,7], Richard Bonneau[2,9,10], Aviv Regev[4,6,10], Sanja Vickovic[1,4,11,12]

1. New York Genome Center, New York, NY, USA
2. Flatiron Institute, Center for Computational Biology, New York, NY, USA
3. Human Technopole, Milan, Italy
4. Broad Institute of MIT and Harvard, Klarman Cell Observatory, Cambridge, MA, USA
5. KTH Royal Institute of Technology, Department of Gene Technology, Science for Life Laboratory, Stockholm, Sweden
6. Massachusetts Institute of Technology, Department of Biology, Cambridge, MA, USA
7. Columbia University Irving Medical Center, Department of Neurology, New York, NY, USA
8. Brigham and Women's Hospital, Department of Pathology, Boston, MA, USA
9. New York University, Center for Data Science, New York, NY, USA
10. Genentech, South San Francisco, CA, USA
11. Columbia University, Department of Biomedical Engineering and Herbert Irving Institute for Cancer Dynamics, New York, NY, USA
12. Science for Life Laboratory, Department of Immunology, Genetics and Pathology, Beijer Laboratory for Gene and Neuro Research, Uppsala University, Uppsala, Sweden

Tissue structure and molecular circuitry in the colon can be profoundly impacted by systemic age-related effects, but many of the underlying molecular cues remain unclear. To address this gap, we build a cellular and spatial atlas of the colon across three anatomical regions and 11 age groups, encompassing ~1,500 mouse gut tissues profiled by spatial transcriptomics and ~400,000 single nucleus RNA-seq profiles. We develop a computational framework, cSplotch, which learns a hierarchical Bayesian model of spatially-resolved cellular expression associated with age, tissue region, and sex, by leveraging histological features to share information across tissue samples and data modalities. Using this model, we identify cellular and molecular gradients along the adult colonic tract and across the main crypt axis, and multicellular programs associated with aging in the large intestine. Our multi-modal framework for the investigation of cell and tissue organization can aid in the understanding of cellular roles in tissue-level pathology.

Conversational analysis of spatial data: empowering investigators to bypass the biostats bottleneck with foundation models and LLMs

Informatics & Computational Biology

POSTER 203

Patrick Danaher[1], Evelyn Metzger[1], Nicole Ortogero[1], Stacey Walker[1], Sayani Bhattacharjee[1], Prajan Divakar[1], Mirko Corselli[1], Joseph M. Beechem[1]

1. Bruker Spatial Biology, Seattle, WA, USA

The scientific impact of spatial transcriptomics is slowed by the limited availability of computational experts. To empower biologists to explore their own data, we are deploying on the AtoMx Spatial Informatics Platform a new kind of analysis framework in which scientists without computational expertise can perform deep and customized analyses.

Two advancements underlie this framework. First, we have trained foundation models on >1 billion cells in the AtoMx database. Prior work like scGPT and GeneFormer showed the power of foundation models in single cell data. Our training database is >10-fold larger than predicate single cell databases and includes spatial information. We use our foundation models to automate annotation of cell type and spatial context, serving results that could take weeks to generate.

Second, we introduce a novel use of large language models (LLMs) in data analysis. Our pipeline creates a carefully-curated package of summary tables, thousands of times smaller than a full dataset, but far richer than a human-generated prompt. Once a LLM ingests this package, it stands ready to answer diverse questions about the dataset, e.g. “what pathways are spatially variable in tumor cells?”, or “what are the spatial contexts T-cells inhabit in this tissue?” LLMs hallucinate, but in parallel with their chat, biologists can use the AtoMx viewer to validate and expand on the LLM’s claims. We propose this chat-then-validate format as a generalizable concept for analysis software. This is distinct from “agentic analyses”, which give LLM’s extensive control over analyses and thereby risk erroneous results.

The experience for a biologist, then, is this: their first glance at their data includes complete cell type and niche annotation, along with a data package for LLMs. They then engage in a wide-ranging exploration, querying an LLM and visualizing pertinent results in the AtoMx viewer. As they develop hypotheses to test, AtoMx provides a GUI to pose exactly the right question, e.g. “how do T-cells change as they get closer to tumor cells”, or “how are microglia different in damaged regions?” The result is an individualized journey through exploratory and confirmatory analysis, all without writing a line of code.

Informatics & Computational Biology

POSTER 204

Synthetic benchmarks enable systematic evaluation of SV callers across short- and long-read technologies

Yueyao Gao[1,2], Mark Fleharty[1,2], Katie Larkin[1,2], Sean Hofherr[1,2], Niall Lennon[1,2], Victoria Popic[1,2]

1. Broad Clinical Laboratories, LLC, Burlington, MA
2. Broad Institute of MIT and Harvard, Cambridge, MA

Structural variants (SVs) are an important class of genomic variation with significant functional and clinical impact, yet their discovery remains an open problem. The performance of SV callers can vary as a function of variant type, size, genomic context, coverage, and sequencing platform. However, real SV truthsets are scarce and typically contain a restricted set of SV types with skewed size and context distributions, hindering our ability to uncover systematic performance bias. To perform a comprehensive evaluation of SV methods, we designed a suite of synthetic benchmarks covering a diverse set of experimental conditions.

We simulated 60 synthetic genomes to evaluate the detection of deletions, insertions, duplications, and inversions across 5 size regimes and 3 genome contexts stratified by mappability. We evaluated 16 SV callers, including alignment-based and assembly-based tools, across multiple sequencing platforms at varying read depth. We found that SV recall varied significantly across different context and coverage regimes. However, some callers maintained high recall across all SV types and sizes, while others performed well only in some regimes. For deletions and insertions, assembly-based tools outperformed alignment-based methods in challenging contexts; nonetheless, some events in highly repetitive regions were missed by all methods. In contrast, for duplications and inversions, the recall of assembly-based tools was low and only certain alignment-based callers found these variants consistently. As expected, we also found that SV detection is impacted by read length.

To stress-test the precision of SV tools in calling simple SVs, we additionally generated a set of synthetic genomes containing only complex variants. We found that most callers report simple events for different combinations of complex SV breakpoints. For instance, in benchmarks containing non-reciprocal translocations, many callers produced numerous deletion, duplication, and insertion calls. Notably, we found that both short-read and long-read callers agreed on many calls. These findings are particularly problematic for SV population studies, where consensus calling is commonly used to improve accuracy.

Overall, our results demonstrate that synthetic benchmarks are a valuable resource for performance evaluation, allowing us to reveal systematic biases and methodological limitations and estimate an upper bound of performance on real datasets.

MGtree: a fast and flexible alignment-based metagenomics pipeline

Informatics & Computational Biology

POSTER 205

John M. Gaspar[1], Samantha L. Sholes[2]

1. Merck & Co., Inc., Data Science & Scientific Informatics, Cambridge, MA, USA
2. Merck & Co., Inc., Data Science & Scientific Informatics, West Point, PA, USA

Metagenomics has improved our understanding of complex microbial samples and ecosystem dynamics through the analysis of genetic material directly from the environment. With the high-throughput nature of next-generation sequencing (NGS), researchers can explore microbial communities without the need for culturing, enabling a wide variety of applications. One such application is the genotyping of viruses in biological samples, which is critical for monitoring viral prevalence and evolution. Compared to other metagenomics applications, the bioinformatics analysis of viral typing is more challenging due to the similarity of genotype reference sequences and the rapid evolution and recombination of viral genomes that frequently occur. For example, the exact k-mer matching approach of the popular metagenomics program Kraken2, though extremely fast, performs poorly when k-mers do not have exact matches or ambiguously match multiple sequences in the reference database.

Here we introduce MGtree, a fast, reliable, Nextflow-wrapped pipeline that utilizes a combination of full-length NGS read alignments and phylogenetic analysis to classify samples of interest. By focusing on full-length alignments, the program ensures maximal use of the sequence information and assignment confidence even in samples with closely related sequences, co-infections, or other complex microbial interactions.

We evaluated MGtree on previously published datasets of norovirus and human papillomavirus (HPV) samples. With the norovirus samples, which are challenging to analyze because of the high mutation rate and the large regions of homology among genotypes, MGtree performed extremely well. MGtree correctly classified an average of 75.3% of norovirus reads down to the genotype level, with negligible incorrect assignments. By contrast, Kraken2 classified an average of only 3.3% of norovirus reads correctly at the genotype level. With HPV, which frequently occurs as co-infections, MGtree similarly outperformed Kraken2 in analyzing both HPV-infected cancer cell lines and cervical cancer samples. MGtree correctly identified all co-infections and had 90% fewer incorrect assignments than Kraken2. These results highlight MGtree's superiority in resolving viral genotypes, and its potential for other metagenomics applications.

Integrating longitudinal imaging and transcriptomics in single cells using CellCage™ technology reveals cellular heterogeneity beyond gene expression

Informatics & Computational Biology

POSTER 206

Pier Federico Gherardini[1], Mehdi Mohseni[1], Shreya S. Deshmukh[1], Shan Sabri[1], Richard Yau[1], Yining Chen[2], Makenzie Sacca[1], Nirvan Rouzbeh[1], Olaia Villa[1], Gaetano Viscido[2], James Walker[3], Joseph Ecker[3], Andrea Califano[2]

1. Cellanome, Foster City, CA, USA
2. CZ Biohub New York, New York, NY, USA
3. Salk Institute for Biological Sciences, Department of Plant Biology, La Jolla, CA, USA

The relationship between gene expression and cellular function is complex. Transcriptomic measurements provide static snapshots, while cellular phenotypes unfold over time. To address this limitation, we developed CellCage™ technology, leveraging light-guided polymerization of hydrogels, to enable longitudinal imaging of cells coupled with endpoint transcriptomics. Within this system, imaging provides unambiguous readouts of complex cellular functions (e.g. cell interactions, phagocytosis, differentiation, senescence), which can be associated with transcriptomic data to investigate underlying transcriptional programs.

Imaging phenotypes are information-rich but challenging to characterize. We leveraged DINOv2, a self-supervised vision transformer trained on massive amounts of natural images, to extract high-dimensional morphological features for interpreting cell images, creating rich descriptors of cellular morphology with dimensionality comparable to gene expression readouts. This combination of imaging and transcriptomics represents a novel data type uniquely enabled by CellCage™ technology.

We validated the accuracy of the imaging features by showing that they can cluster cell images according to their identity, and demonstrating robustness under common augmentations (rotation, blur, etc.). We then investigated morphology-gene expression relationships across multiple biological settings.

Morphological features identified different Hs675.T fibroblast subsets correlating with differentiation state and expression of specific adhesion molecules and cytokines. In plant protoplasts, morphological analysis distinguished three populations: healthy round cells enriched for photosynthesis genes, stressed elongated cells with elevated oxidative phosphorylation, and small guard cells involved in abscisic acid signaling.

Finally, we analyzed THP-1 monocytes differentiated to M0, M1, or M2 macrophages, identifying a specific morphological type preferentially associated with the M2 state. Differential expression analysis identified multiple genes specifically associated with this morphological type and involved in tissue repair, remodeling, and anti-inflammatory activity, revealing phenotypic heterogeneity within M2 populations not apparent by gene expression alone.

In conclusion, the combined imaging and sequencing data uniquely enabled by CellCage™ technology provides new insights into cell state transitions, stress response, and differentiation processes across biological systems, establishing a powerful paradigm for understanding cellular heterogeneity beyond transcriptomics alone.

Dynamic traits concentrate heritability in regulatory regions: scalable functional partitioning in All of Us

Informatics & Computational Biology

POSTER 207

Azadeh Golduzian[1], Davorka Gulisija[1]

1. University of New Mexico, Department of Biology, Albuquerque, NM, USA

Estimating total and functionally partitioned SNP-heritability (h^2) at biobank scale is computationally challenging, particularly for traits with substantial environmental variance. We present a two-path framework on All of Us array genotypes to (i) obtain a robust baseline h^2 and (ii) attribute variance to regulatory (REG) versus non-regulatory (OTHER) annotations without constructing a monolithic genome-wide relationship matrix (GRM).

Path A (baseline REML): Phenotypes are residualized for age, sex, and ancestry principal components; we evaluate age-restricted subsets (<60 and <50 years) to reduce environmental noise. We compute GRMs per chromosome and aggregate them in multi-GRM REML, which recovers total h^2 and enables annotation-aware components while fitting within memory limits. This approach also supports standard sensitivity checks (number of PCs, medication adjustment where applicable).

Path B (scalable partition): We run chunked Bayesian ridge on LD-localized "core+flank" SNP blocks to estimate per-block variance explained and sum across REG/OTHER partitions. This out-of-core design parallelizes efficiently and preserves polygenic signal.

Interim findings support literature-consistent heritability for height, and **we anticipate total $h^2 \approx 0.45$** after age/sex/PC residualization and per-chromosome REML aggregation; finalized point estimates and confidence intervals will be presented. For heart rate and blood-pressure traits, interim estimates behave as expected under covariate control and age restriction; complete REML and partitioned results will be reported at the meeting.

This work contributes: (1) a practical recipe to compute and aggregate per-chromosome GRMs for total and partitioned h^2; (2) a complementary, cloud-friendly chunked approach for functional partitioning; and (3) an age-aware residualization strategy that improves power for environmentally labile traits. Together, these methods enable scalable, interpretable h^2 estimation and functional attribution in large, heterogeneous cohorts such as All of Us.

Expanding and improving the GENCODE human reference

Informatics & Computational Biology

POSTER 208

Toby Hunt[1], Jose Gonzalez[1], Ryan Merritt[1], Zoe Hollis[1], Jane Loveland[1], Jonathan Mudge[1], Adam Frankish[1], on behalf of the GENCODE consortium

1. European Bioinformatics Institute, European Molecular Biology Laboratory, Cambridge, United Kingdom

The GENCODE consortium provides comprehensive reference annotations of all human and mouse protein-coding genes, pseudogenes, long non-coding RNAs, and small RNAs. Accurate gene annotation is essential for both genome biology and clinical genomics, as incomplete or erroneous annotation can propagate significant downstream analytical errors.

Although GENCODE is a well-established community resource, the rapid emergence of new sequencing technologies and diverse biological datasets presents ongoing opportunities to expand and enhance the GENCODE gene set for the benefit of its global user base.

Recent developments include:

1. **Integration of long-read transcriptomic data:** We have developed a suite of bioinformatic pipelines to automatically incorporate large volumes of PacBio and Oxford Nanopore (ONT) long-read RNA sequences into the GENCODE gene set. These efforts have expanded our catalog of novel splice variants and improved the consistency of annotated 5′ and 3′ UTRs.
2. **Improved annotation of small RNAs:** By standardizing our workflows and updating the annotation of small RNA families, we have enhanced the accuracy and usability of GENCODE as a reference resource for this important class of genes.
3. **Identification of non-canonical ORFs:** Using ribosome profiling and immunopeptidomics data, we have generated a high-confidence catalog of “non-canonical” open reading frames (ORFs) that may represent functional translation events within lncRNAs and UTRs of protein-coding genes.
4. **Extension to new genome assemblies:** We have provided GENCODE annotation for the T2T-CHM13 human genome by lifting over and refining models from GRCh38, while also adding novel gene models where appropriate. In parallel, we are developing strategies to project our annotation onto the emerging Human Pangenome reference, thus enabling the exploration of its genetic diversity and previously unresolved genomic regions.

Together, these initiatives extend the depth, accuracy, and scope of the GENCODE catalog, reinforcing its role as a gold-standard resource for both fundamental genome research and clinical genomics applications.

The GENCODE genesets are the default Human and Mouse annotation used in the Ensembl and UCSC genome browsers and can also be downloaded from www.gencodegenes.org.

IsoTag: a universal assembly-free system for deterministic identification of transcript isoforms

Informatics & Computational Biology

POSTER 209

Michihira Tagami[1], Akira Hasegawa[1], Tomoe Nobusada[1], Masayoshi Itoh[1], Takeya Kasukawa[1]

1. RIKEN Center for Integrative Medical Sciences, Laboratory for Large-Scale Biomedical Data Technology, Yokohama, Kanagawa, Japan

The advent of long-read RNA sequencing has revealed vast numbers of previously unannotated transcript isoforms. However, the field lacks a standardized system to identify and track these unannotated isoforms across studies, assemblies, and databases. Here we present IsoTag, a universal and assembly-free identification framework that deterministically encodes transcript structures and sequence variants. IsoTag extends the GA4GH Variation Representation Specification (VRS) to transcripts, employing SHA-512-based cryptographic hashing to generate reproducible, globally unique identifiers. Applied to large-scale long-read RNA-seq data, IsoTag achieved memory-efficient processing, cross-assembly consistency, and seamless integration into existing BAM/SAM workflows via standardized that (i) preserves the isoform Structure (XI) and sequence Variant (XV) identifiers, and (ii) introduces reversible tags Splicetag (XS) and Boundary tag (XB) that encode exon junctions and transcript ends as compact hexadecimal coordinates decodable without a genome FASTA. A Transcript group (XT) identifier supports sample-independent biological clustering with fuzzy 5′/3′ boundaries. By providing the first universal transcript identifier system, IsoTag addresses a foundational barrier to reproducible transcriptomics, enabling scalable integration across single-cell studies, population datasets, and reference transcript annotations.

open-source (MIT license): https://github.com/LSBDT/isotag

A scalable pangenome graph browser for large genomes

Informatics & Computational Biology

POSTER 210

Hiroaki Kamada[1], **Masahiro Kasahara**[1,2]

1. The University of Tokyo, Graduate School of Frontier Sciences, Department of Computational Biology and Medical Sciences, Tokyo, Japan
2. The University of Tokyo, Graduate School of Frontier Sciences, Life Science Data Research Center, Tokyo, Japan

As more human genomes are sequenced to telomere-to-telomere or at comparable quality, visualizing and analyzing human pangenome graphs has become increasingly important for understanding genomic variations of various sizes, ranging from single-nucleotide variants to large structural variants. Viewing a human pangenome graph on a laptop has been a major challenge for existing graph genome browsers because (1) the data size of the graph is large and typically exceeds available memory, and (2) the number of nodes and edges is so high that the rendered graph becomes too complex to interpret.

To address these challenges, we developed a new scalable pangenome graph browser for large genomes such as a human pangenome graph. We implemented a lazy-loading mechanism that loads only the displayed portion of the pangenome into memory, thereby significantly reducing the memory footprint and the loading time. We also introduced semantic zooming, which displays the graph at varying levels of detail: as users zoom out, the browser renders a simplified graph to convey the overall structure, while zooming in reveals finer details such as single-nucleotide variants. The frontend of our pangenome browser is implemented as a web application in TypeScript, while the backend, including modules for precomputing reduced graphs and determining graph layouts, is implemented in C++. With these features, our pangenome graph browser provides smooth user experience comparable to that of Google Maps.

Informatics & Computational Biology

POSTER 211

Leveraging large spatially resolved transcriptomic datasets of 5 million cells (and counting): computing lessons learned

Emily Killingbeck Schneidereit[1], Cindy Reichel[1], Wenyu Zeng[2], C. Dirk Keene[1], Kristyn Galbraith[1]

1. University of Washington, Department of Laboratory Medicine and Pathology, Seattle, WA, USA
2. University of Washington, Department of Biomedical Informatics and Medical Education, Seattle, WA, USA

In recognition of its immense potential to probe biology at previously unprecedented scale, spatially resolved transcriptomics was named the Method of the Year by Nature Methods in 2020. The scale of the data generated with this emerging technology is not without its analysis challenges, however. We describe our approach to analyzing spatial data generated with the CosMx Spatial Molecular Imager (NanoString/ Bruker Spatial Biology (BSB)) in our group, where studies routinely exceed 1 million cells and one large project currently comprises over 5 million single cells with data collection still ongoing.

Our workflow leverages Amazon Web Services (AWS) for compute and data storage, the Python-based GPU-accelerated single-cell data analysis libraries ScaleSC and SCVI-Tools for standard single-cell workflows such as cell QC, normalization, dimension reduction, batch correction, and downstream data analysis, and the Python-based image visualization tool Napari in conjunction with the napari-cosmx plugin developed by BSB for interactive data exploration.

Initial preprocessing of CosMx data, including cell segmentation and assignment of spatially resolved transcripts to cells, is limited to the AtoMx Spatial Informatics Platform (BSB). AtoMx provides a means of exporting the preprocessed data to Amazon Simple Storage Service (S3), and we utilize Amazon Fargate with a Docker image of napari-cosmx to process the CosMx images for visualization in napari-cosmx. Using Amazon CloudFormation, we programmatically generate Amazon Elastic Compute Cloud (EC2) instances for running napari-cosmx. In addition to automating the napari-cosmx image preparation steps, Fargate obviates the need to copy the extensive CosMx files stored in S3 to an Elastic Block Store (EBS) volume on each EC2 instance, thereby saving cost. In summary, our CosMx analysis pipeline is linked by automated steps that limit the potential for user error while leveraging high-performance computing tools to efficiently process the large amounts of data generated by spatial transcriptomic experiments.

Informatics & Computational Biology

POSTER 212

Population-scale merging of structural variants with tandem repeat-aware refinement

Juniper A. Lake[1], Thomas Mokveld[1], Billy J. Rowell[1], Egor Dolzhenko[1], Michael A. Eberle[1]

1. Pacific Biosciences, Menlo Park, CA, USA

Accurate merging of structural variant (SV) calls across large cohorts is hampered by high data volumes, breakpoint uncertainty, and the complexity introduced by tandem repeat (TR) regions. Many SV merging tools err on the side of over-merging or under-merging, distorting allelic representation and downstream population inferences. To address this, we introduce SVX, a scalable, memory-efficient SV aggregation framework. SVX constructs a proximity-based graph over SV calls (linking calls whose breakpoints lie within user-specified distance or similarity thresholds), then iteratively refines clusters to preserve allelic diversity. Crucially, SVX incorporates curated TR catalogs and TR-aware boundary refinement logic, allowing improved resolution in repetitive regions where breakpoint ambiguity is highest. SVX's design minimizes memory footprint and runtime overhead, making parameter sweeps feasible even at cohort scale. In benchmarks on long-read datasets and a 28-sample Platinum pedigree cohort (with iterative cluster refinement), SVX showed superior balance between over- and under-merging, maintained heterozygous allele calls, and improved breakpoint localization in TR contexts relative to baseline greedy merging strategies. The algorithm is modular and extensible. As SVX continues to mature, we expect it to serve as a core module in population SV pipelines for long-read data, providing robust, refined multi-sample SV callsets with reduced merging artifacts.

Informatics & Computational Biology

POSTER 213

Population-scale variant interpretation and clinical impacts in the *All of Us* Research Program using long-read sequencing

Qiuhui Li[1], Kiran V. Garimella[2], Fritz J. Sedlazeck[3] , Michael C. Schatz[4], Michael E. Talkowski[5], Evan E. Eichler[6], All of Us Long Read Working Group

1. Johns Hopkins University, Department of Computer Science, Baltimore, MD, USA
2. Broad Institute of MIT and Harvard, Data Sciences Platform, Boston, MA, USA
3. Baylor College of Medicine, Human Genome Sequencing Center, and Rice University, Department of Computer Science, Houston, TX, USA
4. Johns Hopkins University, Department of Computer Science and Department of Biology, Baltimore, MD, USA
5. Massachusetts General Hospital, Center for Genomic Medicine, and Broad Institute of MIT and Harvard, Harvard Medical School, Medical and Population Genetics, Boston, MA, USA
6. University of Washington School of Medicine, Department of Genome Sciences, and University of Washington, Howard Hughes Medical Institute, Seattle, WA, USA

The *All of Us* Research Program (AoU) is building a comprehensive national biobank to advance genomic medicine in the U.S. While short-read sequencing (SRS) data are available for >500,000 participants, short reads are limited in resolving complex variants, particularly in highly repetitive regions.

To address this, we utilized Phase I long-read sequencing (LRS) of 1,027 self-identified African American/black participants to construct a comprehensive variant resource, including SNVs, indels, SVs, repeat expansions, and more. We identified over 600,000 SVs, > 60% absent from matched SRS data. We identified 473 likely pathogenic SVs overlapping 84 high-priority genes. LRS-based tandem repeat analysis revealed mutations in *FMR1* and other clinical loci, including motif changes missed by SRS. We also detected repeat recombination events affecting *BRCA2*, *TP53*, and *GSTM1*.

We next imputed the SVs in 3,202 samples from the 1000 Genomes Project, identifying 8,249 SVs enriched in African populations. Integration with matched RNA-seq data revealed 889 novel SV-eQTLs, including for *PIK3R1*, a known tumor suppressor, and *SH2B3*, implicated in autoimmunity and cardiovascular disease. Moreover, we found thousands of SVs in strong linkage disequilibrium with significant variants from NHGRI-EBI GWAS catalog and leveraged electronic health records (EHRs) from our long-read samples to further investigate trait associations in Africans. This revealed associations in the same direction but with stronger effects for diseases such as asthma and breast cancer, suggesting African-enriched risk variants. We also identified pleiotropic effects of SVs, including associations with cardiovascular and metabolic traits, such as hypertension, obesity, and metabolic syndrome.

Finally, to further assess the clinical impact of these variants, we imputed SVs in 10,000 short-read African American AoU participants with linked EHRs and performed a phenome-wide association study (PheWAS). We identified 291 SV–disease associations across 226 conditions, with 50.9% involving SVs absent from the matched SRS callset. Fine-mapping revealed SVs were the lead variants over SNVs in 191 associations (65.6%). Additionally, coding-region associations uncovered previously unrecognized protein-altering effects. These findings demonstrate SVs as a major source of medically relevant variation, contributing to missing heritability in complex traits and disease, and underscoring the critical role of long-read sequencing in advancing precision medicine.

Challenging the 30x status quo with a genome-quality vector: a metric-driven, technology-agnostic specification for whole-genome quality, accuracy, and cost

Informatics & Computational Biology

POSTER 214

John E. Major, Jr.[1], Eric Olivares[1], Johnathan C. Whittaker[1], Vincent A. Fusaro[1], Sean P. George[1]

1. Life Science Manufacturing Corporation (LSMC), Genomics, South San Francisco, CA, USA

30x coverage has become shorthand for sufficient genome quality, but it is a blunt, platform-locked proxy. It neither specifies which regions are confidently assayed nor how complex variants are resolved, and obscures cost and turnaround time (TAT). We propose a replacement: a configurable Genome Quality Vector (GQV) that encodes end-use requirements as measurable dimensions, ROI completeness and uniformity; stratified SNV/indel and SV precision and recall (GIAB benchmarks); difficult region recoverability (repeats, segmental duplications, GC extremes); phasing contiguity; allele balance; and cost and time as first-class metrics. Rather than prescribing a single quality proxy, a GQV specifies what must be correct, at what accuracy, for which regions, and at what cost.

We implement GQV in a reproducible, cloud-native benchmarking stack that pins software, hardware and references; logs per-step runtime metrics and cloud spend in real time. Using GIAB samples, we analyze short read datasets (Illumina, Ultima), long read datasets (ONT, PacBio), and hybrid mixtures across diverse aligner-caller combinations. We define parity with a conventional 30x short read baseline as a composite distance in GQV space and trace the Pareto frontier for quality–cost–TAT trade-offs.

Multiple technology paths satisfy a given GQV. Short read pipelines remain strong for small-variant calls in accessible regions at favorable cost and time. Long read workflows improve SV recall, haplotype phasing, and recoverability in difficult regions. Hybrid strategies provide balanced gains, achieving or exceeding 30x short read parity across stratified subsets while reducing blind spots; critically, several hybrid coverage mixes meet GQV targets at lower cost and equal or greater accuracy than is possible by simply increasing short read coverage. Region-specific weights expose clear trade-offs for applications prioritizing SV resolution, phasing, clinical ROIs, or rapid turnaround.

By decoupling quality from platform and replacing coverage targets with metric-driven specifications, GQV enables technology-agnostic procurement, transparent guarantees, and predictive reproducibility of cost, performance, and analytical accuracy: moving the field from 30x to delivering fit-for-purpose genomes.

(Pipelines, results, sample, and reference data made available in our public GitHub repositories.)

A flexible protein-RNA multiomic framework for hierarchical cell typing using marker-derived metagene score

Informatics & Computational Biology

POSTER 215

Dan McGuire[1], Claire Williams[1], Lidan Wu[1], Evelyn Metzger[1], Yi Cui1, Patrick Danaher[1], Bhattacharjee, Sayani[1], Prajan Divakar[1], Mirko Corselli[1], Joseph Beechem[1]

1. Bruker Spatial Biology®, Seattle, WA, USA

Single-cell spatial multiomic assays, which jointly measure RNA transcripts and protein markers, offer unprecedented opportunities for understanding the complex interplay between transcriptional state and cellular phenotype in complex tissues. However, cell typing in these datasets remains challenging: protein signals can bleed across segmented cell boundaries, creating mixed or incompatible marker profiles (e.g., simultaneous detection of CD3 and CD20), while transcriptomic measurements are often sparse and stochastic, limiting the reliability of individual marker genes.

To address these challenges, we developed HieraType, a flexible hierarchical cell typing framework that integrates prior biological knowledge with data-driven modeling. To our knowledge, HieraType is the first dedicated method for hierarchical cell typing in spatial multiomics and incorporates several key advances:

1. It aggregates correlated expression signals across biologically related genes and proteins to construct cell type metagene scores; stabilizing noisy features and preserving interpretability.
2. It formalizes marker-based cell typing through probabilistic modeling, transforming an ad hoc process into a reproducible, quantitative framework.
3. It adaptively weights RNA and protein contributions based on local signal quality, identifying which analytes best distinguish cell type.
4. It requires no external training data or reference atlas, enabling high transferability across diverse tissue types and technologies.

Applied to a spatial breast-cancer multiomics dataset (CosMx® Whole Transcriptome assay and 64-plex CosMx Protein assay, ~147k cells), we benchmarked the agreement between canonical protein and RNA marker expression and the corresponding cell type labels. We found multiomic derived cell type predictions often show stronger alignment with canonical protein markers and RNA markers compared with either modality alone. For example, T cell predictions were more concordant with CD3 protein expression using multiomics (AUC=0.93) compared to cell typing using only protein data (AUC=0.925) or only RNA (AUC=0.87). Meanwhile, CD3D RNA expression also improved using both modalities (AUC=0.69) compared to RNA alone (AUC=0.675) or protein alone (AUC = 0.66). Overall, multiomic cell type labels outperformed single-omic approaches in terms of supporting marker expression across both analytes.

By leveraging both the reliability of validated protein epitopes and the high dimensionality of the transcriptome, HieraType provides a robust, interpretable, and biologically grounded framework for cell type assignment.

Informatics & Computational Biology

POSTER 216

Comprehensive resolution of challenging genomic variation with Oxford Nanopore de novo genome assemblies

Sean McKenzie[1], Philipp Rescheneder[2], Rebecca Stubbs[1], Priyesh Rughani[1], Bryan Leland[1], Phill James[2], Heather Jeffery[2], George Dodd[2], Sissel Juul[1]

1. Oxford Nanopore Technologies, Inc., New York, NY
2. Oxford Nanopore Technologies, plc., Oxford, United Kingdom

Telomere-to-telomere (T2T) near T2T assemblies can be generated robustly and efficiently for arbitrary samples using Oxford Nanopore long and ultra-long reads, providing the most complete picture possible of an organism's genetic sequence. However, the full utility of these assemblies remains to be demonstrated, especially in comparison to raw-read mapping approaches. Here we apply ultra-long read based T2T assembly to six cell lines to demonstrate the power of T2T assembly to resolve potential causes of disease in otherwise un-interrogatable repetitive and structurally divergent regions of the genome.

Using the GIAB benchmark sample HG002, we show that T2T assembly enables accurate genome-wide calling of small nuclear variants (SNVs) and structural variants (SVs), including in challenging medically relevant genes (CMRGs). Compared with read-alignment analysis, we see comparable SNV accuracies and slightly improved SV accuracies.

Balanced chromosomal rearrangements are often difficult to detect with mapping approaches as there are limited or no coverage signatures, and breakpoints are often in repetitive regions. We generated T2T assemblies for two cell lines with chromosomal rearrangements that were previously unresolved following long-read mapping and optical mapping approaches: a Robertsonian translocation and a (balanced) reciprocal translocation. We were able to identify both rearrangements with breakpoint resolution using our assemblies. We also characterized a ring chromosome rearrangement with breakpoint resolution.

Spinal muscular atrophy is a disorder caused by loss-of-function of the SMN1 gene. This gene sits in a variable segmental duplication alongside a nearly identical pseudogene (SMN2). Diagnosis and carrier status are typically assessed with mapping approaches, resolving the copy number of both SMN1 and SMN2. However, the duplication confounds molecular phasing in the locus, resulting in missed "silent carriers" where a duplication of SMN1 on one haplotype masks a deletion on the other. We show that T2T assembly can fully resolve this locus in both haplotypes, allowing the definitive assessment of silent carrier status without extended pedigrees. Application of T2T assembly to two cell lines with silent carrier-associated SNVs revealed functional SMN1 copies in all haplotypes, demonstrating the limited reliability of these markers.

Quantifying spatial resolution and assessing impact on tertiary analysis for ex situ spatial transcriptomics technologies

Informatics & Computational Biology

POSTER 217

Shaobo Zhang[1], Sujai Chari[1], Chris Edlund[1], Darren Segale[1], Kareem Ahmad[1], Mark Wang[1], Cande Rogert[1], Emily Parker[1], Cande Rogert[1], James Han[1]

1. Illumina, San Diego, CA, USA

Spatial Transcriptomics enables a wide range of biological discoveries by capturing gene expression profiles across diverse tissue microenvironments. Common approaches to analyze the transcriptomics data include clustering, cell typing, cell to cell interaction, and pathway enrichment analysis. Achieving high performance with these methods requires technologies that combine high sensitivity along with fine spatial resolution. Illumina Spatial Technology (IST) delivers the combination of whole transcriptome coverage, high sensitivity, and cellular resolution over large sections of tissue samples enabling discoveries. IST achieves cellular resolution through the combination of high density surface features, accurate cell segmentation, and low diffusion in the assay.

Accurate cell segmentation requires identification of cells in tissue images as well as precise alignment of tissue images with transcript data. After using a diverse and curated custom dataset combined with expert annotation of fresh frozen tissue images for training, our AI model achieves high nuclei segmentation accuracy with F1 scores > 0.8. The accuracy of our fiducial based image alignment is validated using feature based alignment of H&E image and spatial transcript data across various sub-regions where all sub-regions show alignment error of < 1 µm.

To address the lack of systematic methods available to assess spatial resolution across tissue types, we introduce a set of versatile tools for evaluating resolution. In addition to the aforementioned feature based alignment, we also provide methods to measure nuclear/intronic transcript diffusion and cell marker gene diffusion measurement, which are applicable across tissue types. For example, using IST applied to mouse brain (coronal) tissue sections, we show high transcript localization by finding that up to 80% of the nuclear transcripts are located within the cell bins. Furthermore, our innovative framework incorporates cell typing confidence and local environment score to enable systematic assessment of annotation accuracy. Lastly, we highlight a diffusion-model based fractional transcript counting strategy that reveals more biologically meaningful cell types.

Using this toolbox, we also demonstrated that IST has superior resolution compared to other ex-situ technologies. Moreover, whole transcriptome combined with sufficient spatial resolution can potentially enable it to deliver biological findings beyond the capabilities of targeted panel in situ technologies.

For Research Use Only. Not for use in diagnostic procedures.

Informatics & Computational Biology

POSTER 218

From fragments to full circles: a scalable strategy for complete phage genome assembly

Joseph R. Petrone[1], Noah Wax[1], Nikhita P. Puthuveetil[1], Jade Kirkland[1], Corina Tabron[1], Stephen King[1], Robert Marlow[1], James Duncan[1], Scott Nguyen[1], Ana Fernandes[1], Briana Benton[1], Jonathan L. Jacobs[1]

1. American Type Culture Collection (ATCC), Sequencing & Bioinformatics Center, Manassas, VA, USA

Phage genome assembly remains a significant challenge in microbial genomics due to fragmented assemblies, contamination from host genomic material, and limitations of current phage genomes available in public databases. Many bacterial reference genomes contain incomplete or misannotated prophage regions, leading to ambiguous or misleading taxonomic classifications when assembling from metagenomic or isolate-derived sequencing data. These issues are compounded when using short-read technologies alone, where tools like SPAdes or Unicycler can often produce fragmented contigs that are difficult to resolve and interpret.

In this study, we present a targeted approach to improve phage genome assembly by leveraging high-quality, authenticated microbial reference genomes from the ATCC Genome Portal. By aligning Illumina reads to these curated bacterial references, we perform precise in silico de-hosting against the bacterial strains the phages were originally isolated from to remove host-derived genetic material prior to assembly. This step significantly reduces background noise and improves the specificity of downstream assembly pipelines.

Despite these improvements, short-read assemblies alone often produce fragmented or incomplete phage genomes. To address this, we incorporate Oxford Nanopore Technologies (ONT) long-read sequencing of phage-enriched material, followed by assembly using Autocycler, a pipeline optimized for circular genome detection and resolution. The resulting assemblies are highly contiguous and frequently yield complete, circularized phage genomes. Subsequent polishing with Illumina data is then leveraged to correct base-level errors and enhance SNP resolution.

This hybrid approach enables the recovery of high-quality phage genomes, even from samples with substantial host background data. Our results demonstrate that combining curated reference-guided de-hosting with long-read sequencing and hybrid polishing produces significantly improved assemblies compared to traditional short-read-only shotgun approaches.

This work underscores the importance of curated reference genomes and tailored assembly strategies in phage genomics. It also provides a scalable framework for researchers seeking to recover complete phage genomes from isolate or metagenomic samples. The resulting phage genomes are actively being published on the ATCC Genome Portal, offering the scientific community access to high-quality phage genomes for enhanced downstream analysis.

OMICS-query: a hybrid AI-powered system for automated interpretation of analysis pipeline results

Informatics & Computational Biology

POSTER 219

Alberto Riva[1], Camilla Callierotti[1]

1. National Facility for Data Handling and Analysis, Fondazione Human Technopole, Milan, Italy

Modern bioinformatics pipelines generate complex, multi-layered results, creating interpretation bottlenecks as omics datasets scale. Researchers must navigate scattered output files and write custom code to access their own data, follow-up questions require analyst intervention, and static reports cannot address the iterative exploration essential to biological discovery.

OMICS-query is an AI-powered system that autonomously orchestrates multiple data analysis tools to answer complex biological questions. OMICS-query is a hybrid system in which structured data retrieval precedes LLM interpretation. This database-first architecture is essential: full-featured analysis pipelines like the ones found in NF-Core produce datasets that routinely exceed millions of rows, far beyond any LLM's context window, and require complex joins and statistical operations that only a relational database can reliably execute.

After an automated ETL pipeline transforms analysis results into a normalized relational database, the system employs LangChain's framework with ReAct logic powered by Google Gemini, using an intent classifier to route queries to specialized prompts that guide the LLM through different reasoning modes. This task-based architectural separation ensures precision for statistical operations while enabling rich interpretations and interactive visual outputs.

Using the NF-core RNAseq pipeline as an example, we show how agentic AI frameworks can democratize access to complex omics datasets, enabling researchers to independently explore their results, while reducing analyst support burden. OMICS-query successfully responds to diverse query types across the analytical spectrum, with an accuracy that in preliminary testing reached 90%: users can discover regulatory patterns ("show me the comparison with the most differentially expressed genes"), explore functional consequences ("what pathways are enriched in downregulated genes in differentiated vs non differentiated cells?"), assess data quality ("detect outlier samples from the correlation matrix"), and seek guidance on their experimental design or analysis ("how should I handle these outliers?"). The system automatically generates publication-ready visualizations and reports while providing scientifically accurate explanations of methods and results in a biological context. The conversational interface maintains context across follow-up questions, enabling iterative data exploration without analyst intervention.

The OMICS-query architecture is readily extensible to other nf-core pipelines and omics data types, offering a scalable solution for modern genomics and bioinformatics facilities.

Informatics & Computational Biology

POSTER 220

eQTL analysis of SLE B cells reveals functional genomic regulatory variation associated with autoimmunity

Christopher Scharer[1], Mathilde Salamon[1], Sakeenah Hicks[1], Toby Amoss[1], Sudhir Kumar[1], Michael Christopher[1], Lou-Ella George-Alexander[1], Mark E. Williams[1], Scott Jenks[1], Arezou Khosroshahi[2], Jeremy Boss[1], Ignacio Sanz[2]

1. Emory University, School of Medicine, Department of Microbiology and Immunology, Atlanta, GA, USA
2. Emory University, Department of Medicine, Lowance Center for Human Immunology, Rheumatology Division, School of Medicine, Atlanta, GA, USA

Systemic lupus erythematosus (SLE) is a complex and heterogenous autoimmune disease characterized by the expansion of extrafollicular pathogenic B cells derived from newly activated naïve cells. Ultimately the production and deposition of autoantibodies results in damage to multiple organ systems. Therefore, further defining the molecular drivers of extrafollicular B cell differentiation and antibody production is important to ultimately prevent and treat SLE. Genetic variants uncovered by GWAS studies only explain a portion of SLE heritability, with most polymorphisms located in non-coding regulatory elements. Expression Quantitative Trait Loci (eQTL) approaches functionally associate non-coding variant alleles with gene expression and cell type abundance changes to facilitate functional annotation of disease associated genetic variation and uncover alterations in gene regulatory networks that contribute to autoimmunity. However, no eQTL study to date has focused on SLE individuals of African ancestry (AFR), despite the higher disease prevalence and burden in this group. Here, we integrated whole genome sequencing with scRNA-seq and scATAC-seq data on B cells obtained from 58 individuals with SLE, including 46 AFR, 7 European ancestry, and 5 healthy controls. We identified thousands of eQTLs where genetic variation correlated with changes in transcription in B cells. Analysis of B cell subsets including extrafollicular, naive, and switched memory cells, revealed eQTLs that were unique to distinct cell types and may promote the altered differentiation profile of B cells observed in SLE. These analyses provide risk alleles and cellular pathways specific to SLE that may contribute to pathogenic B cell responses in autoimmunity.

Simultaneous genotyping and methylation profiling of cell-free DNA using a single assay

Informatics & Computational Biology

POSTER 221

Zohar Etzioni[1], Pascal Hertleif[2], Benjamin Schuster-Boeckler[1]

1. University of Oxford, Ludwig Institute for Cancer Research, Oxford, United Kingdom
2. Softleif AB, Gothenburg, Sweden

Background:
Cell-free DNA (cfDNA) in plasma carries a wealth of molecular information that can be leveraged for cancer detection, monitoring, and minimal residual disease assessment. This information encompasses not only genetic variation, such as somatic mutations and germline polymorphisms, but also epigenetic features, including DNA methylation patterns reflective of tissue of origin and tumor-specific alterations. However, current sequencing methods struggle to detect both methylation and genetic variation in short-fragment, low-input samples such as cfDNA. To address this, we previously developed **TET-assisted pyridine borane sequencing (TAPS)**, a bisulfite-free chemistry that selectively converts methylated cytosine to thymine, while preserving the integrity of the underlying DNA sequence. The assay has now been turned into a commercially available product ("TAPS+"). TAPS+ yields high-complexity, low-error libraries with base-calling and mapping quality comparable to standard whole-genome sequencing (WGS), in principle enabling both methylation and mutation analysis from a single dataset.

Methods:
We have now developed **rastair**, a robust, performance-optimized toolchain that enables **simultaneous genotyping and methylation profiling** from TAPS sequencing data. Rastair employs a set of context-aware machine learning models to accurately distinguish true variants from TAPS-induced base conversions and sequencing artifacts. At the same time, it extracts per-position and per-read methylation information across the genome, allowing direct integration with downstream cell-type deconvolution tools and epigenetic classifiers.

Results:
Applying rastair to cfDNA samples sequenced to a median coverage of 15x, the variant caller achieved **over 95% precision and 95% recall**, on par with current best-in-class genotyping algorithms applied to matched WGS samples. Accurate detection of genetic variants also improved the precision of methylation calls, particularly in regions harbouring CpG-proximal SNPs. Importantly, rastair extends methylation profiling beyond canonical CpG reference sites, identifying methylation at **over 500,000 CpG sites unique to an individual**. Together, these results demonstrate that TAPS combined with rastair provides a unified platform for comprehensive cfDNA analysis, enabling simultaneous characterization of both genetic and epigenetic tumor features from a single sequencing assay.

Informatics & Computational Biology

POSTER 222

Pangenome-aware DeepVariant: enhanced accuracy with the Year 2 HPRC reference and haplotype sampling

Kishwar Shafin[1], Mobin Asri[2], Pi-Chuan Chang[1], Juan Carlos Mier[1], Jouni Sirén[2], Andrew Carroll[1], Benedict Paten[2]

1. Google Inc., Mountain View, CA, USA
2. UC Santa Cruz Genomics Institute, University of California, Santa Cruz, CA, USA

The continual expansion of large-scale genomics datasets provides powerful prior knowledge for sequence analysis. The human pangenome reference from the Human Pangenome Reference Consortium (HPRC) is a prime example, with the **Year 2 release substantially increasing the number and diversity of assembled haplotypes** over the initial version. Utilizing a pangenome reference improves accuracy in read mapping and genotyping, and to fully harness the enriched information in this latest release, we have advanced our pangenome-aware DeepVariant.

DeepVariant is a highly accurate variant caller that creates images of read pileups at candidate variant sites and employs a Convolutional Neural Network (CNN) to classify genotypes. In this work, we enhance our model by integrating the Year 2 HPRC pangenome. A key innovation is the **incorporation of haplotype sampling**, which allows the model to efficiently leverage the vast diversity within the expanded pangenome. This augmentation provides the CNN with a rich set of local haplotype examples, enabling it to more effectively distinguish true genetic variants from alignment artifacts.

Our updated approach demonstrates a significant leap in performance. Building on our previous success, which reduced variant calling errors by up to 25% with the initial pangenome, the integration of the Year 2 reference and haplotype sampling achieves **even greater accuracy**, further reducing errors by more than 10% with respect to the initial pangenome. The improvements are most pronounced when using models trained exclusively on alignments from the graph aligner, vg, further narrowing the gap to perfect variant calling.

This work underscores that as pangenome references grow in scale and quality, so too must the methods that utilize them. Our enhanced pangenome-aware DeepVariant demonstrates a powerful and scalable approach to maximizing accuracy, with direct implications for both genomics research and precision medicine.

Informatics & Computational Biology

POSTER 223

GENE-OM: a web application for visualization and integration of multiomic data in rare Mendelian disease research

Joshua D. Smith[1], Matt Richardson[1], Tom Kolar[1], Chee Hong Wong[1], Karynne Patterson[1], Christina Zakarian[1], Sean McGee[1], Tristan Shaffer[1], Chris Frazar[1], University of Washington Center for Rare Disease Research, Adriana Sedeno Cortes[2], Andrew Stergachis[2], Jessica X. Chong[3], Michael J. Bamshad[3], Ginevra Zanni[4], Chia-Lin Wei[1]

1. University of Washington, Department of Genome Sciences, Northwest Genomics Center, Seattle, WA, USA
2. University of Washington, Division of Medical Genetics, Seattle, WA, USA
3. University of Washington, Division of Genetic Medicine, Department of Pediatrics, Seattle, WA, USA
4. Bambino Gesù Children's Hospital, Unit of Rare Diseases and Medical Genetics, Rome, Italy

Multiomic analysis provides a powerful framework for uncovering causal variants and mechanisms in rare Mendelian diseases. However, the integration and interpretation of heterogeneous datasets remain a major challenge. To address this, we developed **GENE-OM** (Genomic & Epigenomic Network Explorer for OMics), a full-stack web application designed to streamline the loading, organization, and visualization of multiomic data. Built with Nuxt3/Vue 3/TypeScript and supported by PrimeVue and Pinia, GENE-OM offers an intuitive browser-based interface with interactive tables and search functionality. The backend, powered by Spring Boot 3 and PostgreSQL, supports RESTful APIs for secure, role-based access and integration with both local and cloud-based data sources. As a proof of concept, GENE-OM was applied to a rare X-linked congenital ataxia (SCAX5) affecting six males across three generations in a large American family of Norwegian descent. Using our GENE-OM platform, we efficiently reproduced prior results, uncovering a complex structural variant characterized by chimeric insertions from chromosomes 7 and 9 into the Xq25–q27.1 region. These insertions generate novel, tissue-specific chromatin loops that disrupt local 3D genome architecture. This disruption leads to an eightfold upregulation of the neuronal gene *SLITRK4* and a 50% downregulation of the neurodevelopmental transcription factor *SOX3*. The resulting dual gene dysregulation potentially impairs neuronal development through converging mechanisms. Integration of RNA-seq data with ChIA–ATAC-derived differential peaks and loop structures via GENE-OM enabled rapid, mechanistic insights into the regulatory consequences of this variant. Future development will expand GENE-OM's capabilities to integrate orthogonal omics datasets, facilitating the discovery of pathogenic mechanisms driven by noncoding variation in unresolved rare disease cases.

MambaSV: advancing germline and somatic structural variant discovery from long-read genomes through deep sequence modeling

Informatics & Computational Biology

POSTER 224

Zhihan Zhou[1], Tong Zhu[1], Pankaj Vats[1]

1. NVIDIA Corporation, Santa Clara, CA, USA

Structural variants (SVs) represent a major source of human genomic variation, yet accurate detection remains challenging, especially for somatic events in cancer. We present MambaSV, a GPU-native deep learning based SV caller designed to detect germline and somatic SVs from long-reads. MambaSV preserves single-base resolution while modeling megabase context, enabling simultaneous local signal detection and long-range interpretation of alignments.

MambaSV operates on both single BAM (germline) and tumor–normal pairs (somatic) settings. It formulates SV discovery as a per-base classification problem. For each base we compute a 30-feature vector summarizing read statistics (e.g., mapping quality, base quality, CIGAR, supplementary alignment evidence), then feed it into a bi-directional Mamba2 backbone for efficient long-context modeling. We separately compute features across three channels (haplotype 1, haplotype 2, and unphased). A Siamese architecture ingests tumor and normal feature streams with shared weights, and haplotype matched pairs are fused to predict haplotype resolved germline and somatic SVs in a single forward pass. Postprocessing then aggregates contiguous positive bases into breakpoints spanning five SV classes (INS, DEL, INV, DUP, and BND).

On real germline benchmark HG002, we noted instances in which the public GIAB 2024/11 truth SV callset and read alignment (HG002_PacBio-HiFi-Revio_20231031_48x_CHM13v2.0. bam) file was incongruent. We therefore manually curated over 10,000 discordant structural variant calls including false positives and false negatives identified across multiple methods (Sniffles2, CuteSV, DeBreak, Severus, Sawfish, SVision-Pro, PBSV, and SVIM) to construct a high-confidence reference benchmark. On this curated truthset, MambaSV achieved 95.21% F1, while eight baselines ranged from 67–90%F1.

To evaluate breadth across five SV classes in both germline and somatic contexts, we built a comprehensive SV simulator for PacBio HiFi and ONT: matched tumor–normal genomes with SNP/INDEL backgrounds (database-driven + random), reciprocal and fusion translocations, variable allelic fractions, and haplotagging via 3rd party tools. Trained exclusively on HiFi BAMs, MambaSV achieved F1 scores of 96.37% (HiFi) and 94.93% (ONT) on simulated germline datasets with >46,000 SVs, surpassing the best-performing baseline DeBreak (85.66%/85.36%). For somatic calls, MambaSV attained F1 scores of 92.27% (HiFi) and 89.32% (ONT), outperforming Severus (69.18%/70.31%) and additional somatic baselines SVision Pro, SAVANA, and nanomonsv).

Informatics & Computational Biology

POSTER 225

BART series identify functional transcriptional regulators from single-cell and spatial omics data

Hongpan Zhang[1,2,*], Jingyi Wang[1,*], Zhenija Wang[1], Chongzhi Zang[1,2,3]

1. University of Virginia, Department of Genome Sciences, Charlottesville, VA, USA
2. University of Virginia, Department of Biochemistry and Molecular Genetics, Charlottesville, VA, USA
3. University of Virginia, UVA Comprehensive Cancer Center, Charlottesville, VA, USA

*Authors contributed equally to this work.

Transcriptional regulators (TRs) control gene expression programs that define cell identity and functions. Advances in single-cell and spatial omics technologies now enable high-resolution profiling of cellular states and microenvironments, yet identifying functional TRs from such data remains challenging. Built upon our BART framework for functional TR prediction, we develop BARTsc and BART-spatial, computational tools for systematic prediction of TRs from single-cell and spatial omics data, respectively. BARTsc infers TR activity by associating differential genomic features across cell groups with thousands of public ChIP-seq profiles, enabling accurate identification of regulators shaping cell-type-specific signatures. BART-spatial further extends this framework to spatial omics, integrating spatial variability and pseudo-temporal dynamics to capture TR activity even when the gene expression levels are low. Validated on mouse cortex, human PBMCs, pancreatic ductal adenocarcinoma (PDAC), and 10x Visium HD datasets, BARTsc and BART-spatial consistently outperforms existing methods, revealing lineage-defining and disease-driving regulators overlooked by expression-based analyses. Together, BARTsc and BART-spatial provide a unified, interpretable computational strategy for decoding transcriptional regulation from high-dimensional single-cell and spatial omics data, advancing our understanding of cellular identity, tissue organization, and disease mechanisms.

Long-read single-cell transcriptomics with LR_SCAP

Informatics & Computational Biology

POSTER 226

Yongmei Zhao[1,2], Jack Chen[1,2], Sulbha Choudhari[1,2], Biraj Shrestha[1,2], Juan Manuel Caravaca-Guasch[2], Jyoti Shetty[2], Bao Tran[2]

1. Frederick National Laboratory for Cancer Research, Advanced Biomedical and Computational Sciences, Frederick, MD
2. Frederick National Laboratory for Cancer Research, Frederick Sequencing and Genomics Core, Frederick, MD

Single-cell transcriptome sequencing provides high-resolution insights into cellular heterogeneity and gene regulation. However, conventional short-read approaches are limited in capturing full-length transcripts, resolving isoform diversity, and identifying unannotated variants. Long-read sequencing technologies, such as PacBio and Oxford Nanopore, overcome these constraints by enabling full-length transcript characterization and detection of alternative splicing, gene fusions, isoform switching, and single-nucleotide polymorphisms (SNPs) within distinct cellular subpopulations. We developed the Long-Read Single-Cell Analysis Pipeline (LR_SCAP), an integrative framework for comprehensive analysis of single-cell long-read transcriptomic data. LR_SCAP includes modules for quality control, preprocessing, cell calling, barcode correction, clustering, isoform classification, transcript quantification, cell-type annotation, and variant detection. The pipeline allows users to customize workflows by selecting supported tools and automatically generates interactive, HTML-based reports summarizing outputs at each stage. By unifying the long-read single-cell analysis process across sequencing platforms, LR_SCAP provides an efficient, flexible, and accessible solution for high-resolution transcriptomic profiling and characterization of cellular diversity.

Deep learning reveals distinct epigenetic mechanisms of common and rare genetic variants in brain disorders

Informatics & Computational Biology

POSTER 227

Jiyun Zhou[1], Daniel R. Weinberger[1,2], Shizhong Han[1,2]

1. Lieber Institute for Brain Development, Baltimore, MD, USA
2. Johns Hopkins University, Department of Psychiatry and Behavioral Sciences, Baltimore, MD, USA

DNA methylation (DNAm) is essential for brain development and function and may mediate genetic risk for brain disorders. Previous studies have identified genetic variants associated with DNAm levels in the human brain, but largely in bulk tissue and at CG sites, overlooking non-CG (CH) methylation that is enriched in neurons postnatally and crucial for synaptic function. We extend INTERACT, a deep learning model, to predict DNAm regulatory variants across 184 human brain cell types in both CG and CH contexts. INTERACT accurately predicts cell type-specific DNAm profiles and identifies distinct transcription factor (TF) programs, with CH-associated TFs under greater evolutionary constraint than CG-associated TFs. Common variants affecting CG methylation—far more than CH—are enriched in cell type-matched active regulatory regions and show strong heritability enrichment for brain-related traits and disorders. In contrast, noncoding de novo mutations with large effects on CH methylation—but not CG—show greater overlap with high-confidence risk genes for autism spectrum disorder in probands versus unaffected siblings. Our study demonstrates that deep learning can identify functionally relevant, cell type-specific DNAm regulatory variants and reveals distinct epigenetic mechanisms by which common and rare genetic variation contribute to brain disorders.

DRIVE TO SUCCEED

HIGH-PERFORMANCE SEQUENCING WITH DNBSEQ-T7+

Experience Virtual F1 Racing at the Palm Beach Suite

Check out our AGBT workshop/ suite talks

Complete GENOMICS™

ABSTRACT CATEGORY 03

Medical & Biology Omics (Non-Cancer)

Medical & Biology Omics (Non-Cancer)

POSTER 301

Multi-omics analysis reveals genetic, proteomic, and psychiatric correlates of interstitial cystitis/bladder pain syndrome

Elicia Estrella[1], Rosalyn Adam[2], Kristin Cabral[1], David Glahn[3], Ashish Jain[4], Eleanora Melkonian[5], Joshua Motelow[6], Kamil Barbour[7], Ali Gharavi[8], Lindsay Schneider[9], Erika Headrick[9], Michael Diaz[9], Kaylee Mathews[9], John Thompson[9], Catherine Brownstein[1]

1. Boston Children's Hospital, Division of Genetics and Genomics, Boston, MA, USA
2. Boston Children's Hospital, Department of Urology, Boston, MA, USA
3. Boston Children's Hospital, Department of Psychiatry, Boston, MA, USA
4. Boston Children's Hospital, Department of Bioinformatics, Boston, MA, USA
5. University of Toronto, Department of Chemical Engineering and Applied Chemistry, Toronto, Ontario, Canada
6. Columbia University, Division of Critical Care and Hospital Medicine, Vagelos College of Physicians & Surgeons, Department of Pediatrics, New York, NY, USA
7. National Institute of Arthritis and Musculoskeletal and Skin Disease, Bethesda, MD, USA
8. Columbia University, Department of Medicine, Division of Nephrology, Vagelos College of Physicians & Surgeons, New York, NY, USA
9. Nabsys 2.0 LLC, Providence, RI, USA

Interstitial cystitis/bladder pain syndrome (IC/BPS) is characterized by urinary frequency, urgency, and chronic bladder pain. The condition is associated with poor quality of life, significant comorbidities, and limited treatment efficacy due to unclear etiology. This study aimed to better define the biological and clinical landscape of IC/BPS through integrated genomic, proteomic, and psychiatric analyses.

Comorbidities were extracted from medical records, and psychiatric evaluation (N=43) was performed using the Structured Clinical Interview for DSM-5 (SCID-5). Urine proteomics (176 cases, 55 controls) were analyzed using the Olink Explore 384 Inflammation panel. Exome sequencing and electronic genome mapping (EGM) using the Nabsys OhmX Platform (N=248) were combined with data from the Maryland Genetics of Interstitial Cystitis (MaGIC) cohort (N=348 total) and compared to 11,981 controls. Analyses included gene-based collapsing, gene set, exome-wide association (ExWAS), trio, and diagnostic approaches.

Psychiatric testing revealed that 70% of participants met DSM-5 diagnostic criteria for at least one psychiatric disorder, most commonly depressive and anxiety disorders (55%). Proteomic profiling identified 13 differentially expressed proteins and enrichment of cytokine–cytokine receptor interaction, chemokine signaling, and IL-17 pathways. Collapsing analyses highlighted new candidate genes (e.g., *DYX1C1*, *HTR7*) and a novel association with the "small molecule transport" gene pathway. Associations with previously implicated Mendelian disease genes (*ATP2C1, DCAF8, SIX5, ENAM, ATP2A2*) were further strengthened (OR=7.4, CI=1.6–26.4, p=0.005), and comorbid anxiety increased the likelihood of carrying a variant in these genes (OR=29.2, CI=3.0–146.6, p=0.003).

EGM identified a breakpoint fusion supportive of a translocation t(3;9) involving *NOTCH1* in a genome assembly from one individual, which was confirmed using SV-Verify and long-read sequencing. Notably, *ATP2C1* dysfunction has been shown to impair Notch1 signaling, suggesting potential mechanistic overlap.

Together, these findings provide insight into the genetic, proteomic, and psychiatric factors contributing to IC/BPS, supporting a complex, multi-system basis for disease risk and emphasizing the importance of integrated biological and clinical characterization.

Pathogenic virus amplification and sequencing protocol

Medical & Biology Omics (Non-Cancer)

POSTER 302

Gregory A. Buck[1,2,3,4], Myrna G. Serrano[1,2], Katherine M. Spaine[1,2], Christopher J. Oldfield[1,2], Laahirie Edupuganti[1,2], Megan T. Williams[1,2], Andrey Matveyev[1,2], Jingxian Zhao[1,2], Brian Hur[1,2], Kimberly K. Jefferson[1,2], Tom Arodz[2,4], Michael McVoy[2,5], Alexandra Bryson[2,6], Jeffrey R. Donowitz[2,7], Vladimir Lee[3], Kalyan Mallempati[3]

1. Virginia Commonwealth University, Microbiology and Immunology, Richmond, VA, USA
2. Virginia Commonwealth University, Center for Microbiome Engineering and Data Analysis, Richmond, VA, USA
3. Virginia Commonwealth University, Genomics Core, Richmond, VA, USA
4. Virginia Commonwealth University, Computer Science, Richmond, VA, USA
5. Virginia Commonwealth University, Pediatrics, Richmond, VA, USA
6. Virginia Commonwealth University, Pathology, Richmond, VA, USA
7. University of Virginia, Pediatrics, Charlottesville, VA, USA

In a viral pandemic, as was observed for the SARS-CoV-2 event, the etiological agent will likely be derived from a single clone, but variants will occur in response to immunological pressure. As such, a rapid and inexpensive viral genome sequencing protocol will be required in order to follow the development of the virus. In the recent pandemic, we applied a modified viral sequencing protocol developed originally for high throughput profiling of 16S rRNA sequences for taxonomic profiling in microbiome studies. Using this protocol, we were able to sequence thousands of SARS-CoV-2 genomes for a fraction of the standard cost and with a minimum of manual labor. We have now adapted and streamlined this protocol for rapid viral genome sequencing of any viral pathogen. The modified protocol, the Pathogenic Virus Amplification and Sequencing Protocol (PVASP), is quicky applicable to any viral pathogen once the sequence of the original clone is known. It is technically very simple and easily scaled to thousands of samples. Importantly, cost per sample is very low and inversely proportional to the number of samples.

The data included in this study was generated in the Genomics Core at Virginia Commonwealth University. This work was supported through the Pathogen Genomics Center of Excellence program from the Virginia Department of General Services and the Centers for Disease Control.

Integrating spatial proteomics and transcriptomics to map microglial activation in Alzheimer's disease

Medical & Biology Omics (Non-Cancer)

POSTER 303

Ayslan Castro Brant[1], Ekaterina Aladyeva[1], Anjali Garg[2], Iara de Souza[1], Subhodip Adhicary[3], Gordon Meares[3], Oscar Harari[1]

1. The Ohio State University Wexner Medical Center, Neurobiology of Aging & Resilience Center, Division of Neurogenetics, Columbus, OH
2. National Cancer Institute, National Institutes of Health (NCI-NIH), Center for Cancer Research, Bethesda, MD
3. Department of Neurology, The Neuroscience Research Institute, College of Medicine, The Ohio State University Wexner Medical Center, Columbus, OH

Alzheimer's disease (AD) induces complex molecular and cellular alterations that are spatially heterogeneous across distinct brain regions. Emerging spatial omics technologies offer powerful tools to elucidate the spatial context of these changes and provide a deeper understanding of microglial involvement in AD progression. We have recently integrated single-cell RNA sequencing, immunohistochemistry (IHC), and spatial transcriptomics to investigate microglial transcriptional states in relation to amyloid plaques and neurofibrillary tangles (NFTs). Our findings reveal that microglial states in AD brains are heterogeneously distributed, with activated microglia significantly enriched in the external cortical layers and proximal to amyloid plaques, while homeostatic microglia are predominantly located in the internal cortical layers and distal from plaques. We are currently employing advanced technologies to characterize transcriptional changes associated with AD lesions at single-cell resolution. Furthermore, we propose an integrative approach combining high-resolution spatial proteomics using the EVOS S1000 Spatial Imaging System with spatial transcriptomics via 10x Genomics Visium HD, applied to the same tissue section from post-mortem human brain donors with neuropathologically confirmed and clinically diagnosed AD. This dual-modality strategy enables simultaneous mapping of protein and gene activity within the same anatomical regions, preserving cellular architecture and microenvironmental context. Spatial proteomic analysis of immunofluorescence-stained tissue using specific cellular markers (CD68 for activated microglia and P2RY12 for homeostatic microglia) will validate regions of microglial activation and correlate them with AD pathological markers (pTAU and Aβ). In parallel, spatial transcriptomic profiling of the same histopathological regions will generate distinct gene expression signatures associated with microglial activation, neuronal loss, and vascular dysfunction throughout AD progression. By integrating multi-omics data across various brain regions, this approach will uncover convergent molecular pathways and cell-type-specific responses unique to each region and stage of AD. Ultimately, this spatial multi-omics framework will contribute to the development of a comprehensive spatial atlas, offering novel insights into the pathophysiological landscape of AD progression.

Isoform usage in inflammatory bowel disease using long-read single-cell sequencing

Medical & Biology Omics (Non-Cancer)

POSTER 304

Cristina Cotobal Martin[1,2], Omar ElGarwany[1,2], Laura Richardson[1,2], Carl A. Anderson[1,2]

1. Wellcome Sanger Institute, Hinxton, United Kingdom
2. Open Targets, Hinxton, United Kingdom

Using short-read single-cell RNA sequencing, we have previously mapped gene expression dysregulation across large cohorts of healthy and inflammatory bowel disease (IBD) samples (Krzak et al., 2024; Alegbe et al., 2025). Recent advances on Long-read single-cell RNA sequencing (lr-scRNA-seq) now enables direct measurement of isoforms at single-cell resolution, offering an unprecedented view of how alternative splicing contributes to disease mechanisms, gene regulation, and therapy response.

We generated long-read single cell RNAseq data (lr-scRNA-seq) on Pacbio Revio across 61 gut tissue samples, yielding around 90 million reads per sample (median Q30; mean read length 500-700 basepairs). Quality metrics showed high concordance between short and long-read datasets (correlation coefficients: UMIs per cell = 0.88; genes per cell = 0.92). We identified 264,130 transcripts in eight different cell types, 39.5% of which are novel. Novel transcripts show complex patterns of alternative splicing including alternative acceptor/donor splice sites, exon skipping and exonisation of intronic regions. Differential isoform usage analysis revealed 1,818 genes whose isoform composition differs significantly between healthy and IBD individuals. Notably, only 0–26.5% of these genes show differential gene expression within a given cell type, indicating that disease-associated isoform usage dysregulation is largely distinct from gene expression dysregulation.

To systematically chart isoform dysregulation in IBD, we are generating the largest lr-scRNA-seq resource to date across two complementary subcohorts. **IBDverse** aims to uncover genes, cell types, and pathways driving Crohn's disease susceptibility and severity through lr-scRNA-seq of terminal ileum biopsies (n = 115) and PBMCs (n = 121) from patients, plus terminal ileum biopsies from healthy individuals (n = 228). **IBDResponse:** Focuses on identifying and validating predictive models for response to advanced IBD therapies. We will generate long-read scRNA-seq from PBMCs (773) collected before and 14 weeks after the first dose of advanced therapy. These datasets will provide a cell-type resolved map of isoform diversity in IBD, and a framework for linking splicing variation to disease mechanisms and therapeutic response.

References:

Alegbe et al., 2025; Jiang & Chen, 2021; Krzak et al., 2024.

Spatial omics frameworks to guide integration of spatial technology in clinical research

Medical & Biology Omics (Non-Cancer)

POSTER 305

Parambir Dulai[1], Cenfu Wei[1], Alan Robinson[1], Isabella Peshek[1], Michael Patrick[2], Kimberly Young[2], Sanghamithra Korukonda[2], Daniel McGuire[2], Daniel Sanchez[2], Matthew Ingalls[2], Ashley Heck[2], Oliver Braubach[2], Prajan Divakar[2], Joseph Beechem[2]

1. Northwestern University, Chicago, IL, USA
2. Bruker Spatial Biology, R&D, Seattle, WA, USA

Spatial transcriptomics and proteomics help uncover novel biology, guide biomarker discovery, and support drug development. However, data driven frameworks on integration of these technologies in clinical research are lacking. We evaluated multiple spatial single-cell transcriptomic (CosMx, Xenium) and proteomic (CosMx, Cellscape, Phenocycler) platforms for intra- and inter-sample data variance, as well as detection efficiency and confidence, to understand how data quality metrics influenced accuracy of downstream cohort level analyses and biological discovery. Using this spatial single-cell atlas of over 3 million cells from inflammatory bowel disease, with orthogonal in-vitro validation of observations made for cellular interactions, we identified key data confidence metrics which influenced accuracy in downstream analyses for clinical research. Multi-platform spatial single-cell multi-omics was subsequently used to define optimal RNA or RNA + protein based annotation approaches to ensure accurate cellular localization of identified biological pathways. Finally, using pre-post treatment data from multiple cohorts of inflammatory bowel disease patients undergoing treatment with several biologic medications, we applied this data driven framework on a spatial single-cell transcriptomics atlas of over 10 million cells to validate its application for discovery of biological mechanisms of treatment resistance. Through these multi-modal spatial analyses, with orthogonal lab based in-vitro validation of cellular interactions, we have developed and implemented a framework for integration of spatial single-cell transcriptomics and proteomics into clinical research programs. We anticipate this framework will shape future integration of these technologies in clinical research and clinical trials.

Genomic and dietary insights from low-coverage whole-genome sequencing of noninvasive samples of Western Capercaillie in Tyresta National Park, Sweden

Medical & Biology Omics (Non-Cancer)

POSTER 306

Ulf Gyllensten[1], Julia Hedlund Lindberg Julia[1], Jonas Ekstrand[2], Tom van der Valk[3,4]

1. Uppsala University, Department of Immunology, Genetics and Pathology, Uppsala, Sweden
2. Tyresta National Park, Stiftelsen Tyrestaskogen, Tyresta by, Vendelsö, Sweden
3. Swedish Museum of Natural History, Department of Bioinformatics and Genetics, Stockholm, Frescativägen 40, Sweden
4. Centre for Palaeogenetics, Stockholm, Sweden

The Western Capercaillie (*Tetrao urogallus*), Europe's largest woodland bird and a boreal forest specialist, has undergone pronounced declines in southern Sweden due to habitat loss, fragmentation, and isolation. Tyresta National Park and its surrounding forests host one of the last remaining southern populations, yet its genetic status and dietary ecology remain poorly understood. Here, we applied low-coverage whole-genome sequencing to 273 non-invasively collected feather and faecal samples to assess population health, structure, and diet. Feather-derived data provided high-quality genomic information, enabling sex identification, estimates of genome-wide heterozygosity, runs of homozygosity (ROH), and relatedness inference. Heterozygosity levels were comparable to those of other European populations, and overall inbreeding was low, although several individuals exhibited long ROH indicative of recent close-kin mating. Population-structure analyses identified a genetically distinct yet largely panmictic unit, with evidence of female-biased dispersal. Kinship analysis revealed a highly interconnected network, with males having, on average, more relatives than females. Genotype-based recapture simulations estimated a current population size of 164–208 individuals. Faecal metagenomics indicated a diverse diet dominated by key plant taxa such as blueberry (*Vaccinium myrtillus*), lingonberry (*V. vitis-idaea*), and alder (*Alnus spp.*), along with a broad array of invertebrates crucial for chick development. Despite its current genetic robustness, the Tyresta population remains small and isolated, making it vulnerable to future inbreeding and demographic stochasticity. Conservation actions should focus on maintaining habitat quality, promoting gene flow, e.g. by protecting migration corridors to neighbouring populations and ensuring a rich forest invertebrate community to support population resilience. The results underscore the utility of low-coverage genome analyses of non-invasive samples for the wildlife management.

Single-cell 3D spatial transcriptomics reveals neuron–glia communication networks underlying Alzheimer's disease

Medical & Biology Omics (Non-Cancer)

POSTER 307

Qingbo Wang[1,2], Lulu Jiang [1,2*]

1. Virginia School of Medicine, Department of Neuroscience, Charlottesville, VA
2. University of Virginia School of Medicine, Center for Brain Immunology and Glia (BIG), Charlottesville, VA

*Correspondence should be addressed to: jiang.lulu@virginia.edu

Alzheimer's disease (AD) is a complex neurodegenerative disorder driven not only by amyloid and tau pathology but also by dynamic interactions among diverse brain cell types. Understanding how neurons and glial cells coordinate or fail to coordinate within the native three-dimensional (3D) tissue context is essential to uncover mechanisms of neurodegeneration. We developed and applied a single cell 3D spatial transcriptomic framework to systematically map cellular heterogeneity, molecular states, and spatial organization across disease relevant brain regions in AD. Using high resolution 3D spatial transcriptomic profiling of APP/Tau double transgenic mouse brains, we reconstructed the spatial topology of neuronal and glial networks at single cell resolution. Integration of cell type specific transcriptional states with spatial coordinates revealed region specific remodeling of neuron glia communication, including microglial activation gradients, astrocytic network expansion, and oligodendrocyte lineage disruption aligned with amyloid and tau pathology. Through computational inference of ligand receptor signaling within 3D space, we identified pathogenic communication hubs linking glial reactivity to neuronal stress responses and synaptic vulnerability. Cross modal integration with imaging derived pathology maps further connected these signaling networks to amyloid deposition and neuroinflammatory hotspots.

Ongoing work extends this approach to postmortem human AD tissue, enabling cross species comparisons of conserved cellular states and spatial signaling modules. Collectively, this study demonstrates the power of 3D single cell spatial transcriptomics to decode the molecular logic of cell cell interactions in complex brain circuits and provides a scalable framework for spatial systems biology in neurodegeneration.

Keywords:
single-cell, 3D spatial transcriptomics, Alzheimer's disease, neuron–glia interactions, spatial omics, amyloid plaques

Presenter Name and contact information:
Lulu Jiang, MD/PhD
Assistant Professor
University of Virginia, School of Medicine
Center for Brain Immunology and Glia (BIG)
Department of Neuroscience
Email: jiang.lulu@virginia.edu
Phone: 434-924-8229
https://www.lulujianggroup.com/

Integrative multi-omic profiling in heart transplant recipients

Medical & Biology Omics (Non-Cancer)

POSTER 308

Qian Gao[1,2], Alexandre Pereira[3], Michael P. Snyder[4], Brian Piening[5], Brendan J. Keating[1,2,6]

1. New York University (NYU), NYU Langone Health, Langone Transplant Institute, New York, NY, USA
2. New York University (NYU), Grossman School of Medicine, Department of Surgery, New York, NY, USA
3. Harvard University Medical School, Cambridge, MA, USA
4. Stanford University, Palo Alto, CA, USA
5. Institute for Systems Genetics, NYU Langone Health, New York, NY, USA
6. Earle A. Chiles Research Institute, Providence Health, Portland, OR, USA

Heart transplantation provides significant improvement in survival and quality of life for patients with end-stage heart disease. Significant risk of survival of the transplanted heart is attributed to post-transplant rejection phenotypes which can take years of the lifetime of the allograft. Clinical standard-of-care for the evaluation of heart transplant acute cellular- or antibody mediated rejection includes routine endomyocardial biopsy (EMBx) followed by visual assessment by histopathology for immune infiltration and cardiomyocyte damage. Biomarkers such as plasma cell-free DNA adds prognostic and diagnostic capabilities above and beyond conventional histology. We set out to assess spatial transcriptomics and deep proteomics/metabolomics profiling in 384 EMBx and plasma timepoints from 60 heart transplant recipients during the first 12-months post-transplant. We utilized Illumina Spatial Transcriptomics for EMBx, SomaLogic Proteomics (Illumina) of 9,500 protein targets, and Orbitrap metabolomics (Thermo) on all 384 timepoints.

From these data, we observed overall clear discrimination in profiles between clinical ACR and AbMR and non-rejection EMBs (Grade 0R or 1R) as defined by the International Society for Heart & Lung Transplantation (ISHLT), with most significant genes/gene-products comprising immune and inflammatory pathways ($p < 10\text{-}5$) by KEGG/Reactome. Evaluating the non-rejection EMBs in more granularity reveals a wide range of transcriptomic/omic profiles, with a subset exhibiting the gene expression hallmarks of ACR. We show that ACR histologically graded 0R/1R EMBs, which immediately precede an ACR event, exhibit a distinct ACR-like gene expression profile which has a classification accuracy of >95% in these datasets. Fuzzy c-means clustering shows concordance results with rejection related phenotypes. These molecular signatures may significantly benefit diagnoses and prognoses of acute rejection and may have value with conventional histopathology in these highly vulnerable patient populations.

3D reconstruction of the subcellular whole transcriptome in human pancreatic islets

Medical & Biology Omics (Non-Cancer)

POSTER 309

Grant R. Kolar[1], Ashley Heck[2], Aster Wardhani[2], Lidan Wu[2], Mithra Korukonda[2], Margaret Hoang[2], Prajan Divakar[2], Gina L.C. Yosten[1]

1. Missouri University of Science and Technology, Department of Biological Sciences, Rolla, MO, USA
2. Bruker Spatial Biology, Seattle, WA, USA

The pancreatic islet is a complex micro-organ that requires the presence of, and interactions between, multiple cell types within the islet microenvironment for proper functioning. Using spatial molecular imaging (SMI), our group has begun to unravel the genetic bases of these essential cell-cell interactions within the human islet. However, our previous work utilized single sections of human pancreas, with limited tissue depth, to delineate structure-transcriptomic relationships within these micro-organs. Utilizing serial FFPE tissue sections on the CosMx SMI, we have now reconstructed the first known 3D subcellular representations of the human whole transcriptome in non-diabetic and type 1 diabetic islets at sub-micrometer resolution with more than 60 microns of tissue depth. These reconstructions demonstrate significant relationships between the composition and architecture of islet cell neighborhoods and transcriptomic variation within cell types (e.g., between beta cell subpopulations). Notably, interrogation of cells at the border of the endocrine-exocrine interface of the islet reveals differences in transcriptomic, and potentially, functional identities between cells of the same type. These findings represent a significant leap forward in both spatial biology and in understanding of the 3D organization of the pancreatic islet in healthy individuals and in those with type 1 diabetes.

Genome x environment analysis of Sudden Unexpected Infant Death unveils etiologic heterogeneity and strong prenatal cannabis and genetic disease risks

Medical & Biology Omics (Non-Cancer)

POSTER 310

Rishona Lavy[1,2], Stephen F. Kingsmore[1,2], Gretchen Bandoli[3], Daniel Helbling[1,2], Rebecca Baer[3,4], Eric Blincow[1,2], Bryant Cao[1,2], Erwin Frise[5], Alaina Heinen[1,2], Laura Jelliffe-Pawlowski[6], Erica Sanford Kobayashi[1], Lucita Van Der Kraan[1,2], Hugh Kwon[1,2], Barry Moore[7], Danny Oh[1,2], Scott Oltman[4,8], Eric Ontiveros[1,2], Liana Protopsaltis[1,2], Mark Yandell[7], Christina D. Chambers[3]

1. Rady Children's Institute for Genomic Medicine, San Diego, CA, USA
2. Rady Children's Hospital, San Diego, CA, USA
3. University of California San Diego, Department of Pediatrics, San Diego, CA, USA
4. University of California San Francisco, California Preterm Birth Initiative, San Francisco, CA, USA
5. GeneDx, Inc., Oakland, CA, USA
6. New York University, New York, NY, USA
7. University of Utah, Department of Human Genetics, Salt Lake City, UT, USA
8. University of California San Francisco, Department of Epidemiology & Biostatistics, San Francisco, CA, USA

Sudden Unexpected Infant Death (SUID), the third leading cause of infant death, has increasing incidence and multifactorial etiology. Identification of preventative interventions has historically been hindered by etiologic studies limited to genetic or environmental effects in isolation. Here we report a multifactorial genome x environment analysis of SUID risk. Births in San Diego County California from 2005–2018 were linked to hospital discharge summaries and death files, yielding 212 SUID cases and 620,392 infants alive at age 1 year. Whole genome sequencing (WGS) identified probable or possible genetic etiologies in 16% and 48% of SUID cases, respectively. Genetic risks were heterogeneous with 144 loci contributing 173 risks in 57% of SUID cases. Genetic risk was very strong (Prevalence Risk Ratio, PRR >99) or strong (PRR 3.7 – 99) in 12% and 34% of SUID cases, respectively. There was considerable overlap between loci identified by WGS in this study and a previous similar study (e.g. TTN-dilated cardiomyopathy 1G, TSC1-tuberous sclerosis, RYR2-catecholaminergic polymorphic ventricular tachycardia, DSP-dilated cardiomyopathy, NOTCH1-Adams-Oliver syndrome 5, CACNA1H-familial hyperaldosteronism IV, and FLNA-Melnick-Needles syndrome). Additionally, several unifying SUID mechanisms were evident (genetic risk for seizures/SUDEP; cardiac disorders associated with arrhythmias, sudden death or acute heart failure; immunodeficiencies; and breathing or respiratory disorders). Eighty-three percent of the identified genetic risks have preventative or therapeutic interventions available. Six of sixteen significant environmental risks, including very preterm delivery and respiratory distress syndrome, lost significance when SUID cases without strong or very strong genetic risk were compared with infants alive at age 1 year. In contrast, SUID risk associated with prenatal cannabis increased from adjusted hazard ratio (aHR) 3.7 to 6.0, other substance abuse from aHR 2.6 to 3.5, and black race from aHR 1.9 to 2.5. Thus, genome x environment analysis of a large cohort unveiled etiologic heterogeneity, hidden SUID risks, and highlighted cannabis and genetic diseases as strong risk factors. Synthesis of WGS with traditional epidemiologic knowledge has the potential to create a unified understanding to inform new precision preventative interventions. Educational campaigns for SUID should emphasize prenatal cannabis avoidance and newborn screening by WGS has potential for substantial SUID reduction.

Medical & Biology Omics (Non-Cancer)

POSTER 311

Peripubertal remodeling of the human testis mapped in situ with Illumina Spatial Technology (IST)

Michal Lipinski[1], Natalia Brzozowska[2], Amy Guillaumet-Adkins[2], Nicolas Fernandez[1], Amrita Sule[2], Tamara Aleksanyan[1], Jaspreet Ishar[1], Laura Domenech[2], Rebecca L. Linn[3], Tim H. H. Coorens[4], Samouil L. Farhi[1], Kristin G. Ardlie[2]

1. Broad Institute of MIT and Harvard, Spatial Technology Platform, Cambridge, MA, USA
2. Broad Institute of MIT and Harvard, Cambridge, MA, USA
3. Children's Hospital of Philadelphia, Philadelphia, PA, USA
4. European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Cambridge, United Kingdom

The developmental Genotype Tissue Expression (dGTEx) project, an expansion of the adult GTEx project, aims to profile tissue-specific gene expression patterns throughout developmental stages in humans and non-human primates. Among the sampled tissues, the testis follows a distinct developmental trajectory, with substantial structural and functional maturation during infancy, childhood and puberty. This process has to be tightly regulated, as any disruptions in spermatogenesis can lead to life-long infertility. However, our understanding of how testicular tissue is modulated throughout development remains poor.

Here, we applied Illumina Spatial Technology (IST), with a 50 × 15 mm capture area, to map gene expression within intact testicular architecture. In a single assay we profiled testicular tissue sections from six human donors, spanning from infancy to late adolescence, capturing the temporal and spatial molecular mechanisms underlying the onset of spermatogenesis.

We identified spatially variable genes and corroborated marker gene expression using bulk RNA-seq data obtained from the same tissue biopsies. This cross-validation confirmed that the spatial IST technology and bulk RNA-seq profiles were consistent, and subsequent comparison with the 10x Visium HD 3′ platform showed broadly concordant results, albeit with noticeable differences in absolute and relative expression levels. Segmentation of individual seminiferous tubules into distinct neighborhoods allowed us to examine temporal trajectories both across and within time points, enabling computational reconstruction of tubule development at multiple time resolutions. This facilitated the mapping of state-specific signaling pathways and the associated microenvironmental remodeling over developmental time.

This integrated developmental spatial framework provides insight into human testicular maturation and the regulatory programs that shape male reproductive development throughout puberty.

A comparative hydroxymethylome map of cattle and sheep cerebellum

Medical & Biology Omics (Non-Cancer)

POSTER 312

Stephanie McKay[1], Emory Pacht[2], Brenda Murdoch[3]

1. University of Missouri, Division of Animal Sciences, Columbia, MO, USA
2. University of Vermont, Department of Animal and Veterinary Sciences, Burlington, VT, USA
3. University of Idaho, Department of Animal, Veterinary, and Food Sciences, Moscow, ID, USA

Elucidating epigenetic impacts on phenotypic variation involves characterization of multiple epigenetic modifications including DNA 5-hydroxymethylcytosine (5hmC), a stable, transcription permissive mark highly enriched in the brain, where it plays critical roles in neuronal development, synaptic plasticity and gene regulation. Despite its importance, 5hmC remains largely unexplored in agricultural species. Towards that end, a comparative hydroxymethylation map of cattle and sheep cerebellum was generated using whole genome bisulfite sequencing and oxidative reduced representation bisulfite sequencing. The global distribution of 5hmC was similar but quantitatively distinct with cattle (317,136) harboring approximately 2.5x more 5h-mC sites compared to sheep (125,951). Chromosome normalized analysis revealed that orthologous chromosomes BTA25 and OAR24 harbored the highest number of 5-hmC sites with 341.7 and 93.7 sites per megabase, respectively. High resolution comparative mapping of highly hydroxymethylated genes revealed both conserved and species-specific patterns of 5-hmC. Neural and regulatory genes including *GSE1*, *RBFOX1*, *SEPT9*, and *CUX1* exhibited markedly higher 5-hmC abundance in cattle (up to 13×) but greater site-specific hydroxymethylation in sheep. In particular, the highly conserved *CUX1* gene contained 243 5-hmC sites in cattle and 46 in sheep, yet a greater proportion of these sites were highly hydroxymethylated in sheep (37%) than in cattle (18%), reflecting stronger site-specific epigenetic activation despite fewer modification sites. These findings indicate that variation in hydroxymethylation density and intensity across orthologous genes arise from species-specific epigenetic regulation rather than sequence conservation alone. The genome-wide distribution and dynamic function of 5-hmC necessitates further studies into the role of 5-hmC with respect to species-specific regulation and phenotypic diversity. This study introduces new insights into comparative epigenomics and advances understanding of the epigenetic architecture underlying functional diversity across mammalian genomes.

Medical & Biology Omics (Non-Cancer)

POSTER 313

A spatial transcriptomic atlas of autism-associated genes identifies convergence in the developing human thalamus

Fani Memi[1], Alexander Aivazidis[1,2], Koen Rademaker[1], Kenny Roberts[1], Andrew Trinh[1], Tomasz Nowakowski[3], Mahmoud Koko[1], Stanislaw Makarchuk[1], Steven Lisgo[4], Hilary Martin[1], Omer Bayraktar[1]

1. Wellcome Sanger Institute, Cambridge, Cambridgeshire, United Kingdom
2. European Molecular Biology Laboratory (EMBL), Heidelberg, Germany, and Wellcome Sanger Institute, Cambridge, Cambridgeshire, United Kingdom
3. University of California San Francisco (UCSF), San Francisco, CA, USA
4. Human Developmental Biology Resource (HDBR), Institute of Biosciences, International Centre for Life, Newcastle upon Tyne, Tyne, England

Mapping how genes converge on specific brain cell types and neural circuits is critical to understand the complex genetics underlying autism spectrum disorder. Rare loss-of-function mutations strongly predispose susceptibility to autism and additional developmental delays and intellectual disabilities in over 10% of cases. Single cell transcriptomic atlases of developing human brains have implicated enrichment of these genes in cortical neuron subtypes, yet these atlases lack crucial resolution to characterise the cellular architecture of spatially demarcated and functionally distinct brain regions. Here, we present the first spatio-temporal autism gene expression atlas of 250 susceptibility genes, in which over 10 million single cells were profiled using 10x Genomics Xenium from six mid-gestation individuals. We identified five highly spatially regionalised gene expression programs with non-negative matrix factorisation and discovered convergence in the thalamus and germinal zones. We additionally implicate developmental thalamic and cortical neuronal cell types, and progenitor cell populations. These findings implicate the thalamus as a major hub of autism susceptibility in the developing human brain. Overall, we demonstrate future potential to spatially survey expression of trait-associated genes in crucial tissues and developmental stages to better understand genetic pathologies.

Expanded spatial proteomics: integrative same-slide multiomics maps Alzheimer's disease neuropathology at highest plex and resolution

Medical & Biology Omics (Non-Cancer)

POSTER 314

T.C. Orr[1], M.H. Ingalls[2], A. Northcutt[2], O. Braubach[2], M.E. Orr[1,3]

1. Washington University, Department of Neurology, St Louis, MO, USA
2. Bruker Spatial Biology, Research and Development, St. Louis, MO, USA
3. St. Louis VA Medical Center, St. Louis, MO, USA

Neurofibrillary tangles (NFTs) and amyloid plaques define Alzheimer's disease (AD), yet molecular responses to these pathologies are context-dependent and often poorly captured at the transcript level alone. To address this, we developed a same-slide, multiplatform spatial multiomics workflow integrating ultra-high-plex profiling with single-cell and subcellular imaging. Using the GeoMx® Digital Spatial Profiler, we performed whole-transcriptome and 1100-plex protein profiling in formalin-fixed human hippocampal tissue, generating the most comprehensive in situ multiomic dataset reported in neurodegeneration. Regions of interest were centered on neurons with or without tau pathology. The same slides were analyzed with CellScape™ Precise Spatial Proteomics, enabling high-resolution multiplex imaging to map and validate protein biomarkers discovered on GeoMx at single-cell and subcellular resolution. This cross-platform workflow is flexible and iterative: features and morphologies identified in CellScape guided refinement of region selection in GeoMx, allowing targeted profiling of subclassified neurons with the full 1100-plex panel. This strategy revealed tau-bearing neuron subclasses defined by stress response protein expression and cell fate signatures, uncovering distinct trajectories toward apoptosis, necroptosis, or cellular senescence linked to specific tau phosphorylation events. By bridging the highest plex with the highest resolution and enabling iterative cross-platform refinement, this workflow establishes a comprehensive spatial proteomic atlas of AD pathology. These findings illuminate selective neuronal vulnerability to tau pathology and the molecular pathways by which neurons either undergo cell death or evade it through senescence. This integrated platform provides a powerful, flexible framework for mechanistic investigation of neurodegeneration and opens opportunities for targeted exploration of disease-associated cell states at unprecedented depth and resolution.

Multi-omics in 10,000 diverse individuals expands mapping genetic mechanisms in health and disease

Medical & Biology Omics (Non-Cancer)

POSTER 315

Evin Padhi[1], Marianne DeGorter[1], Elizabeth Kiernan[2], Jonathan Nguyen[1], Yulia Mostovoy[2], Alex Lenkin[2], Lee Lichtenstein[2], Namrata Gupta[2], Tara Dutka[3], Konrad Karczewski[2], Niall Lennon[2], Kiran Garimella[2], Michael Talkowski[2], Stephen Montgomery[1]

1. Stanford University, Stanford, CA, USA
2. Broad Institute of MIT and Harvard, Boston, MA, USA
3. National Institutes of Health (NIH), Bethesda, MD, USA

Multi-omics in population biobanks provides new opportunities to map the impact of genetic variation on molecular phenotypes and inform the mechanisms of health traits and diseases. We generate and analyze multi-omics from 10,000 ancestrally diverse participants of the *All of Us* Research Program. These data consist of whole blood transcriptomes using the Watchmaker RNA-seq protocol, detecting both coding and non-coding transcripts, paired with plasma proteome profiling via an expanded Olink Explore HT platform to quantify 5,419 proteins, as well short-read genome sequencing, long-read genome sequencing, and methylation profiling.

The large sample size and ancestrally diverse composition of the *All Of Us* multi-omics cohort enable new opportunities to study common and rare variants impact on gene expression, splicing and protein levels. We perform deep cis-QTL mapping (MAF>0.1%) and identify 70,949 significant (FDR<.05) QTL loci across -omes, with 80% of these signals having reproducible effects. Fine-mapping analysis with SuSiE identified over 700,000 variants across all QTLs and more than 70,000 variants with a PIP > 0.9.

Among fine-mapped QTL variants, we observe comparable numbers of common (>5% MAF), low-frequency (1-5% MAF), and rare (1-0.1% MAF) variants across PIP deciles, demonstrating that rare variants are important drivers of QTL signals that are missed in many studies. We find that these variants are significantly enriched for ENCODE regulatory annotations and coding variant effects similarly across allele frequencies. When examining pathogenicity scores from CADD, we find that causal fine-mapped (PIP >.9) rare variants are significantly enriched for deleterious scores (OR = 6.1,p = 0.00067), suggesting these rare variants may play a role in disease.

We are conducting analyses to integrate transcriptomic QTLs, structural variant and methylation data from *All of Us* long-read sequencing, and performing analyses integrating GWAS data from *All of Us* and external datasets to evaluate the degree that rare fine-mapped variants provide expanded discovery power for trait-relevant genes. Collectively, we have characterized molecular QTLs from *All of Us*, data imminently available on Researcher Workbench, and expanded our identification of rare variants driving molecular phenotypes to human traits and diseases.

Frequencies of clinically important CYP2D6 alleles across Brazil

Medical & Biology Omics (Non-Cancer)

POSTER 317

Eder C. Pincinato[1], Julia T. Siguemoto[1], Nadine Godoy Torso[1], Giovana Fernanda Santos Fidelis[1], Yasmim Gabriele Matos[1], Mariana Vieira Morau[1], Ighor Seiji Okumura Tioda[1], Pedro Gamberini Martins[2], José A.S. Defende[2], João Eduardo Castelano da Silva[2], Sophie Francoise Mauricette Derchain[3], Clarissa Lourenço de Castro[4], Luiz Carlos da Costa Junior[4], Paulo Caleb Júnior Lima Santos[5], Mauricio Wesley Perroud Junior[1], Magnun Nueldo Nunes dos Santos[1], Gabriela da Silva Arcanjo[6], Marcos André Cavalcanti Bezerra[6], Patrícia Moriel[2]

1. State University of Campinas, School of Medical Sciences, Campinas, São Paulo, Brazil
2. State University of Campinas, School of Pharmaceutical Sciences, Campinas, São Paulo, Brazil
3. State University of Campinas, School of Medical Sciences, Division of Gynecological and Breast Cancer, Department of Gynecology and Obstetrics, Campinas, São Paulo, Brazil
4. National Cancer Institute (INCA), Rio de Janeiro, RJ, Brazil
5. Federal University of São Paulo (UNIFESP), São Paulo, SP, Brazil
6. Federal University of Pernambuco, Recife, Pernambuco, Brazil

Brazil, the most genetically admixed country worldwide, remains underrepresented in genomic research. The *CYP2D6* gene is highly polymorphic, with more than 180 defined star-allele haplotypes, and responsible for the metabolism of 21% of drugs. In this study, we aim to analyze the frequency distribution of clinically relevant *CYP2D6* alleles in the Brazilian population, stratified by location and ancestry, in order to provide more precise population allele-frequency estimates. This study included patients from 3 public hospitals in Brazil. Genomic DNA was extracted using a commercial kit and genotyped with the Infinium Global Diversity Array with Enhanced PGx on the iScan platform (Illumina). Variant calling was performed using the DRAGEN™ pipeline, and star alleles were assigned using the StarAllele pipeline (Illumina). Individuals were classified into ethnic groups (European [EUR], American [AMR], African [AFR], East Asian [EAS] and Admixed [MIS]) based on genetic ancestry estimates obtained with the EthSEQ library (Gencode.Exome model, hg38 genome). A total of 1,352 participants were included, originating from Campinas (n=464), Recife (n=282), Rio de Janeiro (n=447), and São Paulo (n=159). Regional allele frequencies ranged from 0.35–0.94% for *3, 10.63–14.01% for *4, 2.30–3.88% for *5, and 0.18–0.67% for *6. The cohort showed broad ancestral diversity, comprising 28.77% admixed (MIS), 28.55% European (EUR) and 25.92% American (AMR) participants. Comparative analyses revealed that individuals from Campinas had a 29.8% higher likelihood of carrying *4 (p=0.034; OR=1.298; 95%CI=1.024–1.645), whereas participants from São Paulo showed a 92.8% higher likelihood of carrying *5 (p=0.019; OR=1.928; 95%CI=1.134–3.277). Regarding ancestry, EUR+EURMIS participants had a 46.3% increased likelihood of carrying *4 (p=0.0019; OR=1.463; 95%CI=1.153–1.855), while MIS individuals had a 30.9% lower likelihood of carrying *4 (p=0.007; OR=0.691; 95%CI=0.526–0.909). AMR participants showed a 58.5% reduced likelihood of carrying *5 (p=0.02; OR=0.515; 95%CI=0.289–0.917). These findings reveal significant regional and ethnic differences in *CYP2D6* allele distribution within the Brazilian population, underscoring the influence of ancestry on genetic variability and reinforcing the importance of ancestry-informed approaches for the advancement of personalized medicine in the national context.

Development of a single comprehensive genomic test based on long-read sequencing technology for the diagnosis of methylation disorders

Medical & Biology Omics (Non-Cancer)

POSTER 318

Siddhartha Sen[1], Hillary Handler[2], Matthew Bower[2], Alec Victorsen[1], Aidan Ellison[1], Bharat Thyagarajan[1]

1. University of Minnesota, Minneapolis, MN, USA
2. Fairview Health, Molecular Diagnostics Laboratory, Minneapolis, MN, USA

Introduction:
We have recently developed and validated a clinically deployable pipeline using Oxford Nanopore Technologies (ONT) long read sequencing that functions as a diagnostic test for patients with suspected genetic disease (PMID: 40385982). We expand on this work by analyzing DNA methylation from ONT data to identify patients with suspected imprinting epigenetic diseases who typically undergo multiple discrete rounds of genetic testing, causing delays in diagnosis and a significant financial burden.

Methods:
We evaluated methylation profiles from patients with known imprinting disorders. Each sample was sequenced on the PromethION-24 targeting a whole genome coverage of 30X or a limited clinical panel at 60X coverage with adaptive sequencing. Data was processed at the Minnesota Supercomputing Institute using a standardized pipeline. Modified bases (CpG methylation) were called with Dorado and aligned to GRCh38 using the ONT wf-alignment workflow. SNVs were identified with Clair3 as the default genotyper. Insertions/deletions/inversions were identified using NanoVar and DeBreak. CNVs on the sex chromosomes were identified using CNVpytor; autosomal CNVs were called by QDNAseq, which is included in the epi2me workflow.

Results:
Our initial studies included patients with either Prader-Willi or Angelman syndrome. When methylation frequency was assessed in the 15q proximal region, a complete loss of methylation at the CpG islands was detected in Angelman syndrome, consistent with loss of the methylated maternal allele. In contrast, 100% methylation was detected in the CpG Islands in Prader-Willi syndrome, indicating a loss of the unmethylated paternal allele. An unaffected sample, as expected, displayed 50% methylation frequency at the same locus, consistent with one methylated maternal allele and one unmethylated paternal allele.

Conclusions and future directions:
Our ONT-based pipeline was able to successfully evaluate known imprinted regions and thus help diagnose diseases of methylation. Our findings reinforce the advantages that long-read sequencing provides as a single comprehensive test for complex clinical testing. As costs decrease and accuracy improves, we anticipate that its use will expand even further. Current work is focused on studying a larger clinical cohort including samples from other methylation disorders like Fragile X syndrome, Beckwith-Wiedemann syndrome, X- inactivation and further optimizing the computational algorithm to match clinical standards.

The long sort of short: comparing Illumina Constellation Mapped Reads with ONT WGS and PacBio long-fragment capture for haplotype phasing in complex blood group loci

Medical & Biology Omics (Non-Cancer)

POSTER 319

Tim A. Steiert[1], Michael Wittig[1], Sören Franzenburg[1], Janina Fuß[1], Björn Brändl[2], Christian Rohrandt[2], Anika Riksted[2], Franz-Josef Müller[2], Christoph Gassner[2], Andre Franke[1]

1. Christian-Albrechts-University of Kiel & University Medical Center Schleswig-Holstein, Institute of Clinical Molecular Biology, Kiel, Germany
2. Private University in the Principality of Liechtenstein (UFL), Institute of Translational Medicine, Faculty of Medical Sciences, Triesen, Liechtenstein

Long-read sequencing technologies have become pivotal for several applications utilizing their benefits, such as structural variation (SV) detection, hybrid allele identification, haplotype phasing, and compound heterozygous mutation analysis. This study evaluates the novel Illumina Constellation Mapped Reads (CMR) technology, which integrates in-flow-cell library preparation with bioinformatic read linking based on flow-cell coordinates to represent long-read sequencing characteristics. We compared CMR performance against established long-read platforms, Oxford Nanopore Technologies (ONT) whole-genome sequencing (WGS), and PacBio targeted long-fragment capture (LFC) workflows for challenging paralogous genomic blood group loci in identical real-world samples. Key metrics assessed include the accuracy of SV and hybrid allele detection, single-nucleotide polymorphism (SNP) genotyping precision, and the length of haplotype phasing blocks. Our results highlight the strengths and limitations of CMR relative to ONT WGS and PacBio LFC, demonstrating its potential as a surrogate for long-read sequencing technologies.

Integrative 3D genomic and spatial multiomic profiling reveals transcriptomic and proteomic consequences of chromosomal reorganization in cellular senescence

Medical & Biology Omics (Non-Cancer)

POSTER 320

Xuehan Sun[1], Johnson Huynh[2], Serdar Tulu[2], Marc Glazer[2], Brian Smart[2], David Castillo[2], Kenny Chung[2], Huy Nguyen[2] Shyamtanu Chattoraj[2], Jude Dunne[2], Doug Werner[2], Taylor Russo[3], Markus Riessland[3], and Miranda Orr[1,4,5,6]

1. Washington University School of Medicine in St Louis, Department of Neurology, St. Louis, MO, USA
2. Bruker Spatial Genomics, Research and Development, San Jose, CA, USA
3. Stony Brook University, Department of Neurobiology and Behavior, Center for Nervous System Disorders, Stony Brook, NY, USA
4. Bruker Spatial Biology, Center of Excellence, Saint Louis, MO, USA
5. St. Louis VA Medical Center, St. Louis, MO, USA
6. Tracy Family SILQ Center, St. Louis, MO, USA

Cellular senescence is a complicated stress response characterized by cell cycle arrest, resistance to apoptosis, and secretion of various proinflammatory molecules collectively known as the senescence-associated secretory phenotype (SASP). In the aging brain, many cell types, including astrocytes, undergo senescence. This process can be further exacerbated in Alzheimer's disease, where toxic protein aggregates have been shown to promote astrocytic senescence and SASP-mediated neuroinflammation.

Emerging evidence suggests that senescence exists along a continuum, accompanied by extensive nuclear and chromatin remodeling. While genome-wide bulk biochemical analyses and lower-throughput imaging studies have revealed broad changes in 3D genome structure, proteomic, and transcriptomic profiles during senescence, these approaches lack single-cell resolution and cannot capture cell-to-cell variability. In this study, we used complementary spatial platforms to characterize in situ changes in genome architecture, chromosomal organization, and proteogenomic profiles of normal and BrdU-induced senescent human astrocytes at single-cell resolution. Senescence was validated by SA-β-gal staining and nuclear size quantification. Notably, BrdU-treated cultures contained a subset of markedly enlarged cells, raising the question of whether this morphological expansion represents a distinct stage of senescence.

The single-cell resolution of both the PaintScape™ and GeoMx® DSP platforms enabled subpopulation-level analysis based on nuclear morphology. Using the PaintScape platform with the ChromoPaint™ PanChromo MPX panel, we visualized over 400 targets across the whole genome and generated chromosome-specific spatial traces and target counts. Parallel GeoMx DSP profiling quantified the whole transcriptome and over 1,000 proteins relevant to neuropathology, immune signaling, and cell-state transitions. Integrating these datasets allowed us to define chromosome-, RNA-, and protein-level signatures associated with the morphological changes unique to a subset of senescent cells. Beyond proteogenomic profiling, we characterized changes in inter-chromosomal proximity, differential chromosome p–q arm interactions, and the nuclear positioning of individual chromosomes and sub-chromosomal loci in situ within single senescent subpopulations.

In summary, we applied an integrated multiomic approach that combines chromosomal architectural mapping with proteomic profiling to uncover the molecular underpinnings of brain cell senescence. These findings provide new insights into how structural and molecular heterogeneity shape the functional diversity of senescent states in the aging and diseased brain.

3D spatial atlas of human lymphatic vasculature

Medical & Biology Omics (Non-Cancer)

POSTER 321

Ioannis S Vlachos[1,2,3,4], Sheethal Umesh Nagalakshmi[1,2,3,4], Angeannie Lefevre[5], Athanasios Ploumakis[1,2,3,4], Angela Chen[5], Antonella Arruda de Amaral[1,2,3,4], Shuoshuo Wang[1,2,3,4], Madeleine Givant[5], Dimitra Karagkouni[2,3,4], Patrick Foley[3,4], Yin Yin[2,3,4], Nikolaos Kalavros[1,2,3,4], Max Itkin[6], Leo Tsai[2,7], Arivarasan Karunamurthy[8], Gloria Pryhuber[9], Liming Pei[10], Katy Börner[11], Luciano Martelotto[12], Fiona Ginty[13], Young-Kwon Hong[14]*, Dhruv Singhal[2,5]*

1. Dana Farber/Harvard Cancer Center/Beth Israel Deaconess Medical Center, Spatial Technologies Unit, Harvard Medical School Initiative for RNA Medicine, Boston, MA, USA
2. Harvard Medical School, Boston, MA, USA
3. Beth Israel Deaconess Medical Center, Department of Pathology, Boston, MA, USA
4. Broad Institute of MIT & Harvard, Cambridge, MA, USA
5. Beth Israel Deaconess Medical Center, Division of Plastic and Reconstructive Surgery, Department of Surgery, Boston, MA, USA
6. University of Pennsylvania, Perelman School of Medicine, Center for Lymphatic Imaging and Interventions, Philadelphia, PA, USA
7. Beth Israel Deaconess Medical Center, Division of Magnetic Resonance Imaging, Department of Radiology, Boston, MA, USA
8. University of Pittsburgh Medical Center, Department of Dermatology, Pittsburgh, PA, USA
9. University of Rochester, University of Rochester School of Medicine and Dentistry, Department of Pediatrics, Rochester, NY, USA
10. University of Pennsylvania, Perelman School of Medicine, Pathology and Laboratory Medicine, Philadelphia, PA, USA
11. Indiana University, Engineering and Information Science, Bloomington, IN, USA
12. Adelaide University, SAiGENCI, ACE, Adelaide, Australia
13. GE Healthcare Technologies, Inc, GE HealthCare Technology and Innovation Center, Chicago, IL, USA
14. Beth Israel Deaconess Medical Center, Department of Surgery, Boston, MA, USA

Lymphatic vessels are essential for lipid transport, immune surveillance, and interstitial fluid clearance yet remain understudied due to technical and biological limitations. They encompass microlymphatics embedded in tissues and lymph nodes and large collecting lymphatics that return peripheral and central lymph. Large collecting lymphatic vessels remain largely uncharacterized at the molecular level in humans.

To address this gap, the Lymphatics Tissue Mapping Center of the Human Biomolecular Atlas Project (HuBMAP) established and optimized single cell and spatial profiling for in depth characterization. We developed intraoperative protocols that preserve RNA/protein integrity, retain precise 3D orientation, and enable downstream single cell and spatial analyses.

In parallel, to provide context and complete a human lymphatic atlas, we assembled the largest human endothelial cell atlas to date (>600,000 single cells; ~20 fold larger than prior efforts) across >15 organs. This dataset identified >50 transcriptionally and functionally distinct endothelial populations and showed similar diversity scaling for blood and lymphatic lineages.

(continue to page 99)

(continued from page 98)

Peripheral, central, and solid organ tissues were assessed in 3D using technologies specifically developed or optimized for this effort. Serial sections were profiled with high plex antibody panels (Akoya PhenoCycler Fusion) and imaging based spatial transcriptomics (10x Genomics Xenium) and registered in 3D. For large dataset generation, we employed the Singular Genomics G4X Spatial Sequencer, which we validated against Xenium. For large format, whole transcriptome mapping at single cell resolution, we used the Illumina Spatial Transcriptomics assay, validated against 10x Genomics Visium HD (targeted and 3′). Finally, lightsheet based VIPAR and Stellaromics Pyxa captured lymphatic vessel micromorphology and transcriptional properties, respectively, natively in 3D.

Integrating these datasets, we find lymphatic endothelial cells (LECs) differ not only between organs or compartments but also within the same niche or even vessel, indicating highly localized specialization and division of labor. Spatial data reveal how LEC specialization is coordinated through cell-cell interactions in a highly localized manner. We extended analyses to vessels and tissues from patients treated for relevant conditions. These findings lay the groundwork to preserve or restore function in lymphatic conditions, which affect more U.S. patients than HIV, Parkinson’s disease, multiple sclerosis, muscular dystrophy, and ALS combined.

Direct RNA sequencing reveals complex patterns of epi-transcriptomic modification following stroke

Medical & Biology Omics (Non-Cancer)

POSTER 322

Zulma Lisseth Reyes-Benitez[1], Elham Amini[1], Andrea Pearson[1], Robert Meller[1]

1. Morehouse School of Medicine, Neuroscience Institute, Atlanta, GA, USA

In the United States nearly 1 million people suffer a stroke each year. The gold standard treatment for ischemic stroke is currently tPA. For patients to receive this treatment, a distinction between hemorrhagic and ischemic stroke must be made within the therapeutic time window for tPA (4.5 hours following stroke onset). Hemorrhage is identified using CT imaging, however diagnostic delays result in only 5% of eligible stroke patients being treated with tPA. Therefore, a biomarker to distinguish hemorrhage from ischemic stroke would have a great effect on population health. In this project, we aim to identify epitranscriptomic biomarkers to distinguish hemorrhagic stroke from other stroke sub types (specifically 5-methylcytosine (5mC), 6-methyladenosine (m6A), Inosine (I), and Pseudo-Uridine (PsU). Whole blood samples were collected from patients at admission (Grady Memorial Hospital, ATL, GA). Following extraction of RNA, 1 ug of RNA was prepared for direct Sequencing using the Oxford Nanopore workflow (RSK004). Libraries were sequenced on flowcells (P24), for 72h. Read data was aligned to GRCh38 using Dorado and modified bases were called. Modkit software was used to extract modified base calls to a bed file, and this data was split by base modification. Data was filtered to only include bases with at least 10x coverage and called present in all samples. For each modification, we determined differential modification rates and corrected for age, sex and race. Our results showed an increase of 5mC and m6A and decrease in I and PsU RNA modifications following hemorrhagic stroke. Differential modification of RNA was observed in targets across all chromosomes, some mitochondrial RNAs showed higher levels of differential modification relative to chromosome size. Some RNAs species showed multiple modifications per target RNA. Pathway analysis using cluster profiler suggested modifications occur in RNAs associated with ATP Synthesis, oxidative phosphorylation, and immune system function. The large and statistically robust changes observed in our results, suggest RNA modifications may have utility as biomarkers for hemorrhagic stroke diagnosis. Future analysis will include defining stroke biomarkers for ischemic and mimic strokes, as well as how our different RNA modifications of interest correlate to both gene expression and splicing.

Multi-omics profiling reveals metal-specific activation of fibrogenesis in human iPSC-derived hepatic stellate cells

Medical & Biology Omics (Non-Cancer)

POSTER 323

Li Tao[1,2], Zhenjie Liu[1], Roger Pique-Regi, Zheyun Peng[1], Yang Jiang[1], Wanqing Liu[1,3,4]

1. Wayne State University, Eugene Applebaum College of Pharmacy and Health Sciences, Department of Pharmaceutical Sciences, Detroit, MI, USA
2. Wayne State University, School of Medicine, Department of Physiology, Detroit, MI, USA
3. Wayne State University, School of Medicine, Department of Pharmacology, Detroit, MI, USA
4. Wayne State University, School of Medicine, Department of Oncology, and Karmanos Cancer Institute, Detroit, MI, USA

Environmental exposure to heavy metals such as cadmium (Cd^{2+}) and lead (Pb^{2+}) has been associated with increased risk of liver fibrosis, while zinc (Zn^{2+}) may have protective effects. However, how low-dose exposures to these metals reprogram hepatic stellate cells (HSCs) - the primary fibrogenic cells in the liver- remains poorly understood. We aimed to determine whether physiological level of Cd^{2+}, Pb^{2+}, and Zn^{2+} based on natural exposure in general population can directly induce fibrogenic activation in human HSCs through transcriptomic and epigenetic remodeling.

HSCs were derived from human induced pluripotent stem cells (iPSC-HSCs) were exposed to each metal at their 1× and 10× blood concentration defined based on the National Health and Nutrition Examination Survey (NHANSE). Activation of iPSC-HSCs was assessed by immunofluorescent staining of key fibrogenic markers, as well as migration and viability assays. Transcriptomic and epigenomic profile changes were assessed using RNA-seq and reduced representation bisulfite sequencing (RRBS), respectively, followed by integrated bioinformatic analyses and Western blotting to identify and verify changes of key pathways and signaling.

Both Cd^{2+} and Pb^{2+} but not Zn^{2+} treatment significantly induced expression of fibrogenic markers α-SMA and collagen, TIMP1, as well as cell migration, without affecting proliferation. Transcriptomic analysis revealed Cd^{2+} and Pb^{2+} induced distinct changes of pathways: while Cd^{2+} treatment is associated with downregulation of mitochondrial and oxidative phosphorylation pathways, as well as upregulation of protein trafficking and cytoskeletal remodeling pathways, Pb^{2+} induces gene expression involved in cell migration and suppression of mitotic genes. On the other hand, Zn^{2+} upregulated stress-adaptive genes and showed dose-dependent cytotoxicity. Epigenomic analysis demonstrated both Cd^{2+} and Pb^{2+} induced dose-dependent promoter hypomethylation, while Zn^{2+} induced higher and lower level of promoter hypomethylation at the 1× and 10× concentration, respectively. Integrated analysis suggested that key genes (*KDM3A* and *KDM3B*) involved in histone H3K9 demethylation, and ephrin receptor tyrosine kinase gene (*EPHA2*) are top genes associated with Cd^{2+} and Pb^{2+} treatment, respectively, which are further verified by Western blotting.

Our study suggests that Cd^{2+} and Pb^{2+} distinctly drive fibrogenic reprogramming in iPSC-HSCs, conferring susceptibility to liver fibrosis. The iPSC-HSC can serve as a patient-derived cell model to study the toxicological impact of environmental stressors.

Medical & Biology Omics (Non-Cancer)

POSTER 324

Adopting a metagenomic and quasimetagenomic approach to comparatively investigate the microbial diversity of bovine feces cultured for *Escherichia coli, Salmonella, Campylobacter, and Listeria spp.*

Kingsley Bentum[1], Rejoice Nyarku[1], Temesgen Samuel[1], Woubit Abebe[1]

1. Tuskegee University, College of Veterinary Medicine, Center for Food Animal Health, Food Safety, and Defense, Department of Pathobiology, Tuskegee, AL, USA

Advancements have been made in the application of metagenomics, with modifications such as quasimetagenomics, greatly aiding in the rapid detection of pathogens. In this study, the fecal microbiomes of cattle known to be shedding Salmonella, Escherichia coli (E. coli), Listeria, and Campylobacter were investigated using metagenomics and quasimetagenomics to examine the dynamics of microbial populations in infected and uninfected cattle. For this study, DNAs extracted directly from fecal samples and culture enrichment samples of 28 apparently healthy cattle kept on grass pastures and a similar feeding regime were sequenced and analyzed. Cattle farms recorded no use of antibiotics for at least 3 months before sampling. The fecal shedding statuses of the animal subjects for Salmonella, E. coli, Listeria, and Campylobacter were determined in a concurrent study. For bacterial enrichment, standard procedures for selectively enriching the respective pathogens were followed. The compositional abundance of microbial populations in the sequenced data was analyzed using a custom k-mer database and Kraken2. Results showed the bovine fecal microbiome comprised large proportions of diverse microbes dominated by Bacteroidetes and Firmicutes. A large proportion of Proteobacteria in a sample positive for simultaneously shedding Salmonella, E. coli, and Campylobacter revealed potential dysbiosis resulting from the multiple-coinfection. For culture-enrichment samples, only E. coli was detected in significantly abundant numbers, even to the serovar level. On the other hand, E. coli and other anaerobes were significant co-enrichers of Salmonella and Campylobacter cultures, respectively, masking the detection of these target pathogens. This study shows that co-enrichers can significantly impact the potential application of quasimetagenomics in studying target microbes from bovine feces. Furthermore, the gut microbiome of animals shedding multiple pathogens can offset the bovine gut microbiota. The implication of these findings for microbial culture and public health deserves further study.

Keywords: Metagenomics, Salmonella, Campylobacter, Escherichia coli, Listeria, cattle

Long-read sequencing of hundreds of diverse brains provides insight into the impact of structural variation on gene expression and methylation

Medical & Biology Omics (Non-Cancer)

POSTER 325

Kimberley Billingsley[1]

1. National Institutes of Health (NIH), Bethesda, MD, USA

Structural variants (SVs) play a crucial role in driving gene expression in the human brain and are implicated in numerous neurological conditions. However, most existing genetic studies rely on short-read sequencing methods, which have limited ability to capture the full spectrum of structural variants across the genome. Long-read sequencing offers a significant advantage in detecting disease-associated and functionally relevant SVs, but until recently, this approach has not been scalable for large genomic studies. Here, we introduce a comprehensive solution using a new scalable wet-lab protocol and Napu, a computational pipeline we developed for whole-genome Oxford Nanopore Technologies (ONT) sequencing. Napu is specifically designed to address the computational challenges associated with the analysis of complex structural variants and methylation patterns in long-read data. By integrating Napu, we applied this approach to neurologically normal control samples from the North American Brain Expression Consortium (NABEC) (European ancestry) and the Human Brain Collection Core (HBCC) (African admixed ancestry) cohorts. Through this work, we generated and present a publicly available long-read resource from the frontal cortex of 361 human brain samples, achieving a median N50 of 26.89 Kbp and 40x genome coverage. This dataset enables the discovery of over a quarter million SVs and allows us to produce locally phased assemblies that cover 95% of all protein-coding genes in the GRCh38 reference genome. Utilizing matched expression datasets, we performed quantitative trait locus (QTL) analyses, identifying structural variants that significantly impact gene expression in post-mortem frontal cortex brain tissue. Further, by leveraging Napu, we determined haplotype-specific methylation rates of millions of CpG sites, identifying key cis-acting structural variants influencing methylation. In summary, this study highlights how long-read sequencing can effectively identify disease-relevant regulatory loci that were previously inaccessible using short-read technologies. This resource provides a critical foundation for understanding the biological effects of genetic variation in the human brain. Expanding on this as part of the NIH's Center for Alzheimer's and Related Dementias (CARD) long-read sequencing initiative, we are currently scaling this framework to sequence thousands of ancestrally diverse case and control brain samples.

Comparative genomic and virulence profiling of Shiga toxin–producing *Escherichia coli* using pangenome analysis tools

Medical & Biology Omics (Non-Cancer)

POSTER 326

A. Elrefaey[1], K.E. Bentum[1], E. Kuufire[1], T. James[1], R. Nyarku[1], V. Osei[1], S. Temesgen[1], W. Abebe[1]

1. Tuskegee University, College of Veterinary Medicine, Department of Pathobiology, Tuskegee, AL, USA

Shiga toxin-producing *Escherichia coli* (STEC) are major foodborne pathogens with high genomic diversity and variable virulence potential. This study aimed to compare pangenome reconstruction and virulence interpretation across two independent workflows, RIBAP and Panaroo, to better understand the genomic variability of STEC. A total of 160 genomes representing eight clinically relevant serogroups (O157, O145, O121, O111, O103, O26, O45, and O104) were retrieved from the NCBI database and analyzed. Both pipelines delineated the core and accessory components of the STEC pangenome but exhibited notable differences in gene clustering and paralog grouping. Panaroo, which incorporates graph-based error correction and stringent quality filters, identified 11,515 gene clusters, including a larger core genome (3,394 genes), compared to RIBAP, which detected 22,238 clusters with 2,967 core genes. Conversely, RIBAP estimated greater soft-core and shell components (225 and 3,328 genes, respectively) than Panaroo (114 and 2,895).

Virulome screening using Abricate against the Virulence Factor Database (VFDB) revealed that most isolates harbored *stx1, stx2, eae, artJ, cadA, cesD, cheA, clpV, csgA, eaeH,* and *entA*. However, serogroup-specific heterogeneity was observed among adhesion- and secretion-associated loci, reflecting differences in host adaptation and pathogenic potential. Integration of pangenome and virulence data demonstrated that clustering algorithms significantly influence gene annotation, leading to differences in downstream interpretation of virulence factors.

Overall, these findings emphasize that cross-tool validation is essential for accurate genomic characterization of STEC. Combining complementary pangenome workflows with virulome analysis enhances understanding of STEC population structure, genetic diversity, and potential virulence mechanisms, contributing to improved surveillance and molecular diagnostic strategies for foodborne pathogen control.

Keywords: *Shiga toxin–producing Escherichia coli* (STEC), pangenome, RIBAP, Panaroo, virulence

Genomic characterization of *Escherichia coli* from ruminants: an examination of diversity, virulence, and antimicrobial resistance patterns

Medical & Biology Omics (Non-Cancer)

POSTER 327

Faith Kadzviti[1], Yesutor Soku[1], Abdelrahman Mohamed[1]

1. Tuskegee University, Pathobiology Department, Tuskegee, AL, USA

Escherichia coli (E. coli) is a complex and ubiquitous commensal organism in both food-producing animals and humans, comprising diverse pathogenic strains. In ruminants, pathogenic E. coli can asymptomatically colonize the intestinal tract, leading to fecal shedding and contamination of the food chain, making them important reservoirs of infection. Notably, studies have attributed approximately 75% of human E. coli O157:H7 outbreaks to food products of bovine origin. In this study, we characterized E. coli isolates from 116 fecal samples collected from cattle (n=48) and goats (n=68) in Tuskegee, Macon County, AL. Presumptive identification of 80 isolates was performed using coliform chromoselect agar and standard biochemical tests, followed by genomic confirmation via whole-genome sequencing (WGS) on the Illumina MiSeq platform. Of these, 55 isolates (26 cattle, 29 goats) were confirmed as E. coli. Genome assemblies analyzed on the GalaxyTrakr platform using the Shovill, E. coli serotyper, ABRicate, and staramr pipelines revealed diverse serotypes, including Shiga-toxin producing E. coli [O93:H28 (20%), O126:H19 (2%), O76:H19 (7%)], extraintestinal pathogenic E. coli O180:H14 (13%), enteroinvasive E. coli O8:H19 (5%), and enterohemorrhagic E. coli O80:H19 (4%). Virulence genes identified included toxins (stx1B, stxA, hlyE), adhesins (fimH, yagZ/ecpA, afaC-VIII), biofilm-associated genes (csgG, csgF, csgE), serum resistance determinants (ompT, iss, traT), and iron acquisition systems (chuX, iroN, iutA). Resistome analysis showed a predominance of the multidrug resistance gene mdf(A)_1 (82%), alongside resistance determinants for β-lactams (blaACT-6), aminoglycosides [ant(3'')-Ia_1], quinolones (oqxA, oqxB), tetracyclines [tet(B)_2, tet(C)_2, tet(C)_3], and sulfonamides (dfrA1, sul2_2). Six plasmid types (IncFIA, IncFIB, IncI1-I, IncFIC, Col156, IncB/O/K/Z) associated with both virulence and antimicrobial resistance were also detected. The occurrence of highly pathogenic E. coli serotypes with multidrug resistance potential highlights a significant public health risk posed by fecal contamination of food and the environment. At the same time, plasmid carriage underscores the potential for horizontal dissemination of virulence and AMR determinants. These findings underscore the importance of ongoing genomic surveillance and rigorous farm-to-fork control measures to mitigate transmission risks.

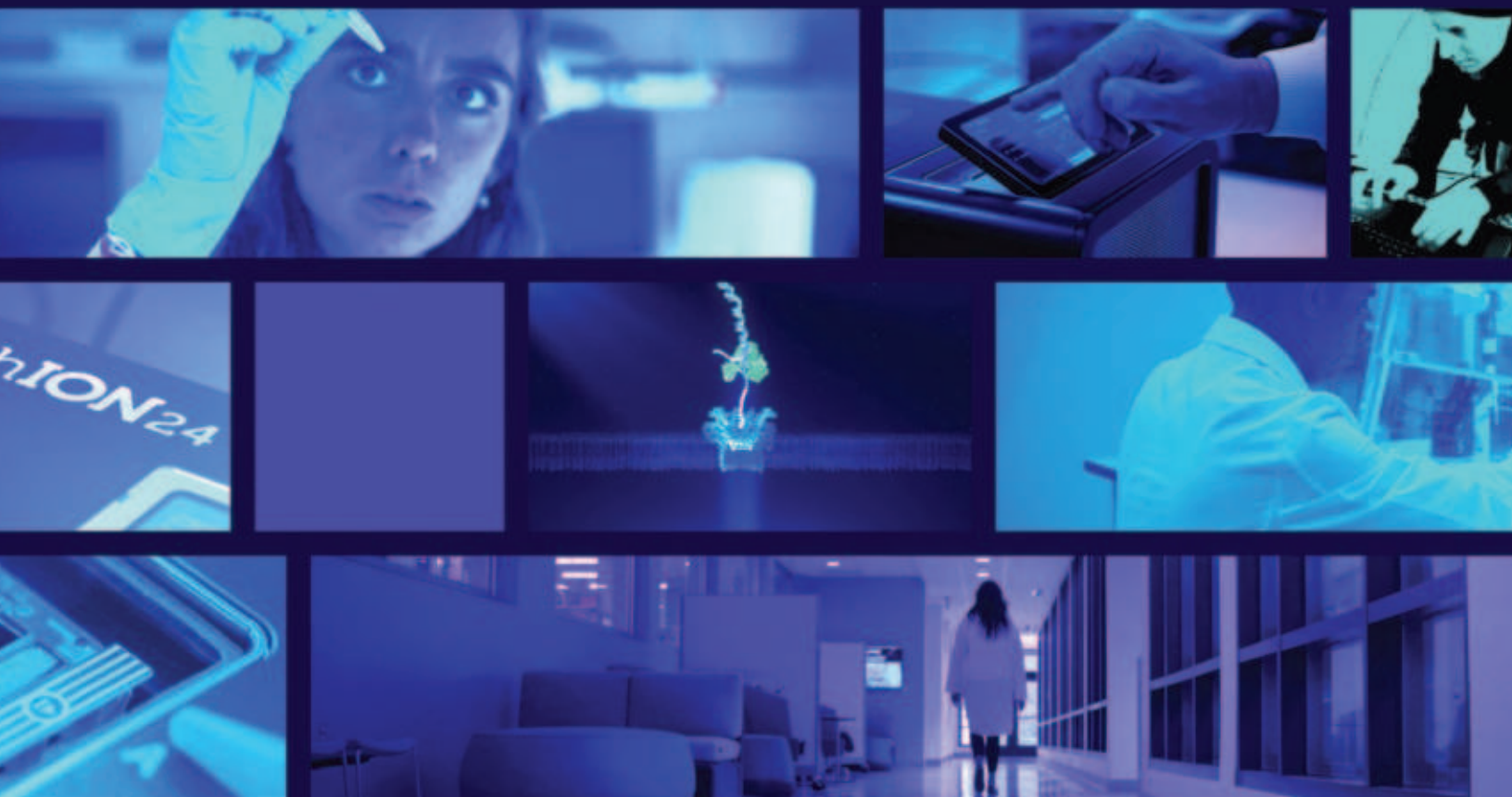
Oxford Nanopore Technologies

Multidimensional insights. In one go.

Who says you can't see it all? With a comprehensive view of the genome, epigenome and transcriptome, Oxford Nanopore technology helps you answer the bigger and bolder research questions you've always wanted to ask.

Visit us in the Lake Suite in the Sponsor Promenade to find out more.

nanoporetech.com



ABSTRACT CATEGORY 04

Technology & Application Development

Technology & Application Development

POSTER 401

Beyond ‘black box’ analysis: a platform-agnostic guide for non–formalin-fixed single-cell RNA sequencing

Karen Abramson[1], Ellora Haukenfrers[1], Vaibhav Jain[1], Stephanie Arvai[1], Khooshbu Patel[1], Simon G. Gregory[2], Devi Swain-Lenz[3,4]

1. Duke University, Molecular Genomics Core, Durham, NC, USA
2. Duke University, Department of Neurosurgery, Durham, NC, USA
3. Duke University, Department of Molecular Genetics and Microbiology, Durham, NC, USA
4. Duke University, School of Medicine, Sequencing and Genomic Technologies, Durham, NC, USA

The rapidly advancing field of single cell RNA sequencing (scRNAseq) offers numerous options for transcriptome profiling. Single cell resolution enables insights into heterogeneous samples. However, questions remain as to which chemistry is ideal for individual experimental goals. Previous benchmarking studies included a mixture of fresh and fixed samples or probe-based and non-probe-based capture methods, which limit the conclusions drawn between analogous technologies. Here, we present a novel comparison of four widely used non-probe-based, non-formalin *fixed* scRNAseq assays. In contrast to past comparisons that used varied analysis pipelines, we have applied company-agnostic cell calling algorithms for unbiased comparison of biological and technical replicates from healthy human PBMCs. Our approach evaluated 10x Genomics, Parse Biosciences, Scale Biosciences, and Illumina scRNAseq assays to examine data QC based on sensitivity, accuracy, precision, power, and efficiency using agnostic and platform-specific cell calling. While metrics vary between assays, there are clear advantages and limitations to each technology, including time and financial costs. In summary, our study highlights the need for carefully considered project design of non-formalin fixed scRNAseq assays, which is based on many factors and dependent on an investigator’s specific research aims and available resources.

Linking high-throughput transcriptomics and imaging for functional drug response profiling

Technology & Application Development

POSTER 402

Alexandre Coudray[1], Daniel Alpern[1]

1. Alithea Genomics, Epalinges, Vaud, Switzerland

Cost-efficient morphological and gene expression profiling are widely used for high-throughput phenotypic analysis and perturbation screens. Previously, comparative data relied on the legacy L1000 probe-based transcriptomics. New technologies, such as DRUG-seq and Total DRUG-seq, offer unbiased whole transcriptome gene detection, providing greater complexity insights and expanding biological discovery in large-scale screens.

In this study, we applied Total DRUG-seq, a new method capable of detecting full length transcripts across hundreds of samples, and Cell Painting to collect gene expression and cell morphology, in U2OS and HepG2 cell lines. Both cell types were exposed to 82 compounds from the JUMP-MoA reference panel to evaluate the capacity of Cell Painting and Total DRUG-seq to uncover mechanistic biological insights. To assess their combined potential, we implemented data integration strategies and compared the added value of multi-omics integration against single-modality analyses. We used unsupervised learning to assess clustering quality and the robustness of Total DRUG-seq and Cell Painting across replicates and conditions. We then applied supervised learning to predict compound mechanisms of action (MOAs) and classify compounds, comparing the predictive power of each assay alone and in combination.

Our results show that integrating transcriptomic and morphological data provides additional insight into drug activity compared to single omics. This highlights the synergy between the assays and underscore the value of multi-omics profiling in experimental design, particularly for identifying cellular responses to chemical or genetic perturbations.

NEBNext® enzymatic solutions for DNA methylation profiling of highly damaged DNA

Technology & Application Development

POSTER 403

Robin L. Armstrong[1], Vaishnavi Panchapakesa[1], Margaret R. Heider[1], Brittany S. Sexton[1], Laura N. Blum[1], V.K. Chaithanya Ponnaluri[1], Bradley W. Langhorst[1], Daniel J. Evanich[1], Louise Williams[1]

1. New England Biolabs Inc., Ipswich, MA, USA

Advances in Next Generation Sequencing (NGS) technologies have enabled large-scale quantification of DNA methylation. DNA isolated from archival material like formalin fixed paraffin embedded (FFPE) tissues is an invaluable resource for genome-wide methylation studies and has advanced the field of cancer genetics. However, FFPE DNA poses many notable challenges for preparing NGS libraries, including low input amounts and highly variable damage from fixation, storage, and extraction methods leading to insufficient coverage and sequencing artifacts.

To investigate the potential to improve DNA methylation profiling from FFPE samples, Illumina compatible libraries were constructed using FFPE DNA with DNA integrity numbers (DIN) ranging from 1.5 to 6.8. DNA was fragmented using mechanical shearing (Covaris) or enzymatic fragmentation (NEBNext UltraShear®). An enzyme-based methylation detection technology, NEBNextÒ Enzymatic Methyl-seq v2 (EM-seqÔ v2), was used to detect DNA methylation. This system overcomes the limitations of chemical-based bisulfite conversion methods and minimizes damage to DNA, enabling longer insert sizes, lower duplication rates, and minimal GC bias resulting in more accurate quantification of DNA methylation. Additionally, DNA methylation was profiled in clinically relevant FFPE DNA from multiple tissue types using EM-seq v2 coupled with UltraShear fragmentation. Results for these challenging DNA types showed that the UltraShear fragmented EM-seq libraries had lower duplication rates, lower chimeras, as well as higher percentages of mapped reads and increased library complexity compared to Covaris sheared libraries. These libraries exhibited higher unique coverage depth and uniformity, leading to more CpGs confidently detected.

EM-seq coupled with UltraShear fragmentation produces reliable methylation profiles and improves library preparation success rates from challenging FFPE samples. The workflow is robust, flexible, and automation friendly, advancing the use of methylation profiling as a cancer biomarker.

Technology developments toward genome-wide single-cell CRISPR screens in hiPS-derived brain cells using 10x Genomics

Technology & Application Development

POSTER 404

Sarah Cooper[1,2], Juliette Steer[1,2], Yixi Chen[1,2], Sam Ogden[1,2], Caia Dominicus[1,2], Diego Peretti[1,2], Aleksander Gontarczyk[1], Omer Bayraktar[1,2], Leopold Parts[1,2], Florian Merkle[2,3], Andrew Bassett[1,2]

1. Wellcome Sanger Institute, Hinxton, Cambridge, United Kingdom
2. Open Targets, Hinxton, Cambridge, United Kingdom
3. Institute of Metabolic Science, University of Cambridge, Cambridge, United Kingdom

In recent years rapid technological developments have enabled pooled single cell CRISPR (scCRISPR) screens that couple perturbation of gene function in individual cells to complex single cell RNA sequencing readouts. However, extending these studies to genome-scale in disease-relevant cell types such as hiPSC-derived brain cells remains constrained by technical bottlenecks in transgene silencing, perturbation efficiency and guide detection.

Here we discuss our technological developments that have enabled pooled scCRISPR screens in hiPSC-derived neurons and astrocytes using direct guide capture with 5' 10x Genomics GEM-X. We discuss how we have overcome the challenges of using these cell types, including developing a dual-guide system to enable better editing and guide detection, and a codon-optimised Cas9 in an 'all-in-one' lenti vector delivery to overcome silencing of Cas9 cell lines during differentiation. We demonstrate the effectiveness of our system by screening a 'NeuroKO' library of ~1000 guide RNA pairs targeting ~250 known Alzheimer's disease targets based on GWAS, expression data, rare variants and network analysis. We have screened more than 1M cells in iNeurons and iAstroctyes at a coverage of over 900 cells per target in each cell type and have revealed differential expression changes in the knockouts that also occur in patients that have allowed a subclassification of the disease targets.

We further show our optimisation of scCRISPR screening in hiPSC derived microglia through improvements in Cas9 and guide vector design and their delivery at a later stage of the differentiation process to avoid bottlenecks allowing coverage of large-scale libraries. Here we present a side-by-side comparison of direct guide detection with 10x Genomics, using either the CRISPR kit in 5' GEM-X or the addition of custom probes towards the guide sequences for detection in FLEX kits. In this experiment we have tested the addition of LNA bases within the custom FLEX probes to gain better guide detection. Overall, the scale of FLEX, the ease of processing and FACS sorting fixed cells and the success of guide detection, are allowing us to scale up to cost-effective genome-wide single cell perturbation experiments in hiPSC-derived brain cells.

From arrays to genomes: a cost-effective TnX transposase-based method for implementing genotyping by low-pass sequencing on the Ultima UG 100™

Technology & Application Development

POSTER 405

David Bays[1], Maura Costello[1], Joseph Mellor[1]

1. seqWell Inc., R&D, Beverly, MA, USA

Genotyping by low-pass sequencing (LPS) offers several potential benefits over the use of conventional genotyping microarrays, including higher resolution, freedom from defined probe sets, and the ability to discover novel variants. However, LPS has not yet fully replaced microarrays due to limitations in sequencing cost and library throughput requirements. The Ultima Genomics UG 100™ is a high-throughput sequencing platform with significant sequencing cost advantages that allow low-pass sequencing to better compete with microarray-based genotyping.

A major challenge in replacing arrays with sequencing is the cost and scalability of NGS library preparation. Traditional workflows rely on mechanical fragmentation, which requires expensive instruments, or enzymatic fragmentation with fragmentase, which can be difficult to control. These methods lack built-in normalization, require multiple enzymatic steps, involve individual PCRs, and necessitate quantification of each library, increasing costs and labor.

Here, we present a fast, efficient, and cost-effective transposase-based library prep workflow using native Ultima adapters optimized for LP-WGS on the UG 100™. The method uses an engineered transposase (TnX™), which improves on standard hyperactive Tn5 by offering lower insertion bias, higher activity, better coverage uniformity, and increased library complexity. A key feature is auto-normalization across a wide range of input DNA, reducing the need for accurate quantification and manual normalization. This enables more libraries per run with less effort. Across a 10-fold DNA input range, this workflow maintains consistent fragment size distributions and achieves a read count coefficient of variation (CV) below 20%, compared to over 80% using a non-normalizing tagmentation method.

By incorporating a native Ultima barcoded adapter during tagmentation, samples can be pooled before amplification. A second index is then added using barcoded primers during PCR, allowing combinatorial multiplexing of up to 1,536 samples. This reduces per-sample costs, improves automation efficiency, eliminates the need to convert incompatible libraries into Ultima libraries, and reduces the library quantification burden.

This flexible TnX-based workflow enhances the utility of the UG 100™ platform for large-scale applications such as LPS, where efficient multiplexing of hundreds to thousands of samples is essential to compete with the costs and throughput of traditional array-based genotyping.

How SPRI chemistry shapes NGS library quality and performance

Technology & Application Development

POSTER 406

Max Boeck[1], Ariele Hanek[1], Josh Haimes[1], Giulia Corbet[1], Eleanor Matrullo[1], Heather Maasdorp[1]

1. Watchmaker Genomics, Boulder, CO, USA

Solid-phase reversible immobilization (SPRI) bead–based cleanup is a core component of next-generation sequencing (NGS) library preparation, enabling DNA and RNA purification, fragment size selection, and removal of unwanted byproducts such as adapter-dimers. However, variability in bead formulation and surface chemistries across manufacturers influences DNA binding and elution behavior. These effects are compounded by interactions between the bead chemistry and the unique buffer composition of different NGS DNA and RNA library preparation chemistries. Consequently, selecting and optimizing bead cleanup ratios for each workflow is critical for maximizing recovery, ensuring accurate fragment sizes, and delivering reliable sequencing performance.

In this study, we systematically evaluated eight commercially available SPRI bead formulations across ratios from 0.5X to 1.0X within three NGS library preparation workflows, including PCR-free whole-genome sequencing (WGS), DNA library prep and RNA library prep as well as post PCR—which are all impacted by insert size and yield tradeoffs. Cleanup performance was assessed using a 100 bp ladder reference, and post-cleanup recovery and fragment distribution were quantified on an Agilent TapeStation D1000 system. We observed substantial differences in fragment retention between bead types, particularly at lower ratios where binding efficiency diverged. Some formulations consistently preserved smaller fragments, while others preferentially retained longer DNA, highlighting strong interactions between bead chemistry and library buffer composition. Notably, bead ratios impact library inserts with reduced read overlap and improved mapping efficiency – key drivers of sequencing efficiency and coverage uniformity in PCR-free WGS.

SPRI bead selection and ratio optimization play a critical role in achieving consistent and high-yield NGS libraries. The observed differences highlight the importance of empirically evaluating bead chemistries within the context of specific library prep workflows. Tailoring cleanup conditions to the chemistry of each kit can substantially improve recovery, fragment size fidelity, and overall library quality across sequencing applications.

High-resolution detection of posttranslational modifications using single-molecule protein sequencing

Technology & Application Development

POSTER 407

Kendrick Nguyen[1], Andrey Reshetnyak[2], Muriel Priault[3], Ajay Vashisht[1], John Vieceli[1], Meredith L. Carpenter[1]

1. Quantum-Si, Branford, CT, USA
2. St. Jude Children's Hospital, Memphis, TN, USA
3. IGBC, Universite de Bordeaux, Bordeaux, France

Post-translational modifications (PTMs) play a critical role in regulating protein function and cellular processes. However, due to the diversity and complexity of the proteome, methods to detect and quantify PTMs have lagged behind advances in DNA sequencing. While conventional mass spectrometry and antibody-based methods provide valuable information, they are limited in their capacity to resolve and quantify PTMs at the amino acid level, particularly for PTMs with minimal differences in mass. The Quantum-Si Platinum® Pro sequencer is a benchtop instrument designed for high-resolution, single-molecule protein sequencing. Platinum Pro uses engineered N-terminal amino acid recognizers to accurately determine the sequence composition of proteins, enabling the detection of PTMs and making advanced proteomic analysis accessible in a standard laboratory setting. In addition to leveraging the recognizer binding patterns, the recent development of PTM-specific binders further enhances the instrument's ability to detect and quantify PTMs by selectively targeting modified residues. This combined approach improves sensitivity and specificity, allowing researchers to quantify PTMs and obtain even more detailed PTM profiles on a single run. In this study, we report the application of single-molecule protein sequencing on Platinum Pro to detect and quantify multiple PTMs across diverse protein types. We demonstrate the detection of phosphorylation order at several sites within the anaplastic lymphoma kinase (ALK) protein, citrullination events on vimentin, asparagine deamidation as a marker of molecular ageing, and the identification and differentiation of aspartic acid stereoisomers in calmodulin sequences. This is the first study demonstrating the ability to resolve and quantify multiple PTMs on the Platinum Pro, enabling differentiation of chemical changes and stereochemical variants that, while subtle, can nevertheless have significant biological impacts. We expect this technology to democratize advanced protein characterization, making the high-resolution detection of PTMs available to a broader scientific community.

Accurate size estimation of short tandem repeat expansions with Illumina Constellation

Technology & Application Development

POSTER 408

Severine Catreux[1], Andrea Finocchio[1], Pascal Grobecker[1], Ningxin Ouyang[1], Daniel Andrews[1], Vitor Onuchic[1], Fenner Macrae[1], Konrad Scheffler[1], Jacqui Weir[1], Manisha Mullen[1], Mitch Bekritsky[1], James Han[1], Fiona Kaper[1]

1. Illumina Inc., San Diego, CA

Short tandem repeat (STR) expansions are associated with a wide range of neurological and neurodevelopmental disorders. While the presence of an STR expansion of size beyond the small ranges observed in healthy population is often a strong indication of pathogenicity, the size of such expansions is known to modulate the presentation of many associated conditions. Short-read WGS is an effective method to distinguish non-expanded from expanded STRs, but its ability to accurately estimate the size of large STR expansions is limited. This limitation can hamper the more nuanced categorization of the effect of STR size on disease presentation.

Illumina Constellation mapped-read technology is a novel approach that provides enhanced genomic insights using standard SBS sequencing by tagmenting long DNA template molecules directly on the flow cell surface. This approach establishes a relationship between the physical distance of nearby reads on the flow cell surface and their genomic distance in an individual's genome. Here we present innovations that leverage the novel proximity information derived from Constellation and allow for more accurate STR expansion size estimation across the full range of observed expansion sizes.

In a cohort of 41 datasets (21 unique cell lines) carrying STR expansions across a set of 10 different disease-associated loci with orthogonally characterized sizes (ranging from 15-4863bp), we demonstrate that Constellation technology achieved 100% sensitivity in distinction of normal vs expanded alleles. Further, Constellation achieved 90% accuracy in the classification of STR sizes into the more nuanced classification bands related to disease phenotype modulation, such as age of onset, disease severity, and carrier status. That included classification of FMR1 expansions into intermediate (6/6 with 40-55 repeats), premutation (4/6 with 55-200 repeats) and full mutation (3/3 with >200 repeats) ranges, as well as classification of DMPK expansions into mild pathogenic (3/4 with 49-100 repeats), classic pathogenic (6/6 with 100-1000 repeats) and congenital disease (3/4 with >1000 repeats).

These findings highlight the power of Constellation technology to overcome current limitations of short read WGS in STR expansion detection, allowing users to glean additional relevant information related to STRs without relying on reflex tests for accurate STR size estimation.

Resolving challenging genes at scale with PacBio PureTarget™: from carrier screening to hereditary disease panels

Technology & Application Development

POSTER 409

Anupam K. Chakravarty[1], Jeff Zhou[1], Davy Lee[1], Shreyasee Chakraborty[1], Guilherme De Sena Brandine[1], Tom Mokveld[1], Dan Nasko[1], William J. Rowell[1], Heather Ferrao[1], Sarah Kingan[1], Michael A. Eberle[1]

1. Pacific Biosciences (PacBio), Menlo Park, CA, USA

Native long-read sequencing preserves DNA methylation and delivers base-level accuracy across complex genomic regions such as repeats, paralogous loci, and structural variants. Although native long reads deliver deep biological insight, adoption for high throughput screening applications has been limited by the cost and throughput of whole-genome sequencing. Targeted long-read sequencing overcomes these constraints by enabling comprehensive analysis of clinically relevant genes in a scalable, cost-efficient manner for applications including carrier screening, repeat expansion disorders, and hereditary disease panels.

Here, we present PacBio® PureTarget™, powered by HiFi sequencing and CRISPR-Cas9 enrichment, which consolidates clinically challenging loci into a single PCR-free assay while retaining native methylation information. The PureTarget carrier panel profiles 12 historically difficult genes, including *FMR1, FXN, SMN1, CYP21A2, GBA, F8* and *HBA1/HBA2*, to resolve repeat expansions, copy number variation, and large structural rearrangements in a single assay. The assay supports manual and automated workflows with 48-plex and 96-plex library preparation, and scales to 100,000 samples per year on the Revio® system and 10,000 on the Vega™ system, enabling efficient high-throughput implementation across diverse laboratory settings.

Furthermore, to demonstrate this platform's flexibility, we utilized PureTarget to profile *CFTR*, a 189 kb locus encompassing thousands of known variants. Because CTFR exceeds a single HiFi read length, we implemented a custom full-gene tiling design that achieved contiguous coverage, allele phasing across variant-rich regions, and methylation analysis while remaining compatible with the carrier panel workflow. CF-positive samples were successfully profiled, confirming detection of both coding and noncoding variants in a multiplexed format. We further applied the tiling strategy to five hereditary cancer genes: *MLH1, MSH2, MSH6, PMS2,* and *EPCAM*, each longer than a single HiFi read, illustrating adaptability to broader hereditary disease panels.

By combining long-read accuracy, native methylation detection, and scalable targeted enrichment, PureTarget enables comprehensive and cost-effective sequencing of clinically relevant loci—streamlining workflows and expanding access to long-read insights across human health research.

A scalable, instrument-free approach to true single-cell spatial transcriptomics

Technology & Application Development

POSTER 410

Christina Chang[1], Bryant Ngo[1], Wanxin Wang[1], Diksha Kool[1], Christina Fan[1], Andrew Farmer[1]

1. Takara Bio USA, Inc., San Jose, CA, USA

Spatial transcriptomics has transformed our ability to map gene expression within intact tissues, yet most existing approaches require specialized instrumentation, limiting accessibility and scalability. We present an instrument-free spatial transcriptomics workflow that combines Trekker™ technology (a barcoded oligonucleotide-tagging approach based on the Slide-tags method) with the Illumina® Single Cell 3′ RNA Prep, T100 assay to deliver high-density, high-sensitivity, and true single-cell resolution spatial data.

In this workflow, photocleavable spatial barcodes on a 10 mm x 10 mm bead array diffuse into a tissue section to spatially tag individual nuclei. Following nuclei isolation, the released spatially barcoded nuclei are captured using the Illumina Single Cell 3′ RNA Prep, T100 kit, enabling recovery of up to 100,000 cells per sample without the need for specialized instrumentation. Libraries are then prepared through the standard single-cell RNA-seq method. Spatial information is thus encoded without the need for microfluidics or imaging hardware.

Applied to both mouse tissue and human tumor samples, this workflow produces high-density, spatially resolved transcriptomes with true single-cell resolution across diverse tissue architectures. We demonstrate robust detection of cell type-specific gene expression, preservation of tissue heterogeneity, and sensitivity that exceeds state-of-the-art instrument-based spatial methods. Importantly, the system integrates seamlessly with existing computational pipelines—without the need for cell segmentation or deconvolution—to facilitate straightforward co-analysis with single-cell datasets.

By removing instrumentation barriers while enabling high-throughput recovery of spatially resolved nuclei per sample, this workflow democratizes access to spatial transcriptomics and expands its utility for both basic research and translational studies. This instrument-free integration of Trekker technology with the Illumina Single Cell 3′ RNA Prep, T100 kit establishes a new paradigm for spatial genomics—delivering true single-cell spatial transcriptomes from routine tissue samples using only standard molecular biology workflows. This method enables widespread application of precise single-cell mapping, driving new developments for genomics and biomarker discovery research.

Live spatial omics for high-resolution, dynamic profiling of tumor tissues

Technology & Application Development

POSTER 411

Miao-Ping Chien[1,2], Lars van Roemburg[1,2], Toba Ajani[1,2], Tsai-Ying Chen[1,2], Maria Lopez-Cascales[1,2], Jelle Storteboom[1,2], Li You[1,2]

1. Erasmus University Medical Center, Department of Molecular Genetics, Rotterdam, The Netherlands
2. Oncode Institute, Utrecht, The Netherlands

Spatial omics technologies have transformed biomedical research in recent years, much like single-cell sequencing did a decade ago. However, most existing spatial methods are restricted to fixed biological samples, limiting our ability to study dynamic cell-cell interactions and temporal changes in living tissues.

To overcome this barrier, our lab has developed a new generation of live spatial omics technologies, multi-FUNseq and in situ FUNseq (*unpublished*), that enable high-sensitivity, whole-transcriptome profiling directly from living tumor tissue slices. These methods extend our original imaging-guided functional single-cell sequencing approach (FUNseq; *Nat Biomed Eng, 2022*), which links cellular phenotypes and behaviors to their molecular states.

Multi-FUNseq and in situ FUNseq allow single-cell multi-omic profiling from specific regions of interest in live tissues. Moreover, in situ FUNseq enables fixation and library preparation directly within the tissue after live-cell analysis, eliminating the need for tissue dissociation and preserving spatial context.

Together, these technologies open new avenues for investigating real-time cellular dynamics and microenvironmental interactions that have remained inaccessible to current spatial omics methods.

Technology & Application Development

POSTER 412

Genomic and epigenetic responses of yeast to Deep Space Radiation during Artemis I

Hamid Kian Gaikani[#1,2], Ashley S. Kleinman[#3,4,5], Tobias Niederwieser[7], Luis F. Paulin[8], Samantha Breaux[1], Eleanor L. Campbell[1], Marjan Barazandeh[1], Pamela Flores[7], Guri Giaever[1], Paul Collier[3], Theodore M. Nelson[3,9], Aaron Berliner[3], JangKeun Kim[3,4], Kirill Grigorev[10,11], Fritz J. Sedlazeck[8,12,13], Denise Freeland[14], Kristine Wilson[14], Kamber Scott[14], Howard G. Levine[14], Deena Najjar[3,15], Krista A. Ryon[3], Ye Zhang[14], Dinah Dimapilis[14], Stefanie Countryman[7], Louis Stodieck[7], Nicanor Austriaco[20,21], Christopher R. Chin*[4,6,7], Corey Nislow*[1], Luis Zea*[7,16], Christopher E. Mason*[3,4,17,18,19]

1. University of British Columbia, Faculty of Pharmaceutical Sciences, Vancouver, British Columbia, Canada
2. University of British Columbia, Department of Chemistry, Vancouver, British Columbia, Canada
3. Weill Cornell Medicine, Department of Physiology and Biophysics, New York, NY, USA
4. Weill Cornell Medicine, The HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine, New York, NY, USA
5. Massachusetts Institute of Technology, Department of Earth, Atmospheric, and Planetary Sciences, Cambridge, MA, USA
6. Weill Cornell Medicine, Department of Medicine, New York, NY, USA
7. Hospital for Special Surgery, Division of Rheumatology, New York, NY, USA
8. University of Colorado Boulder, BioServe Space Technologies, Boulder, CO, USA
9. Baylor College of Medicine, Human Genome Sequencing Center, Houston, TX, USA
10. Columbia University Irving Medical Center, Vagelos College of Physicians & Surgeons, Department of Microbiology & Immunology, New York, NY, USA
11. NASA Ames Research Center, Space Biosciences Research Branch, Moffett Field, CA, USA
12. Blue Marble Space, Seattle, WA, USA
13. Baylor College of Medicine, Department of Molecular and Human Genetics, Houston, TX, USA
14. Rice University, Department of Computer Science, Houston, TX, USA
15. NASA Utilization and Life Sciences Office, Exploration Research and Technology Programs, Kennedy Space Center, FL, USA
16. Albert Einstein College of Medicine, Bronx, NY, UA
17. Jaguar Space, LLC, Erie, CO, USA
18. Weill Cornell Medicine, Tri-Institutional Computational Biology and Medicine Program, NY, USA
19. Weill Cornell Medicine, The Feil Family Brain and Mind Research Institute, New York, NY, USA
20. Weill Cornell Medicine, WorldQuant Initiative for Quantitative Prediction, New York, NY, USA
21. University of Santo Tomas (UST), Research Center for the Natural and Applied Sciences, UST Laboratories for Vaccine Science, Molecular Biology and Biotechnology, Manila, Philippines
22. University of Santo Tomas (UST), College of Science, Department of Biological Sciences, Manila, Philippines

#These authors contributed equally.

*Corresponding authors (e-mail: chc2077@med.cornell.edu corey.nislow@ubc.ca luis.zea@colorado.edu

The Deep Space Radiation Genomics (DSRG) Experiment aboard NASA's Artemis I mission was designed to assess the physiological and genomic responses of eukaryotic cells to the unique environment of Lunar spaceflight. Utilizing the model organism Saccharomyces cerevisiae, we conducted bulk and single-cell phenotyping, Scanning Electron Microscopy (SEM), and single-molecule sequencing techniques on both flight and ground control samples. While spaceflight-associated structural variants (SVs) were detected, surprisingly there were no significant differences in somatic mutation rates, either in frequency or in genomic location.

(continue to page 120)

(continued from page 119)

However, analysis of the three basepair context of the Single Nucleotide Variants (SNVs) shows the first deep space radiation DNA damage signature which is similar, though not identical, to mutational signatures generated from both the NASA Twins Study and Inspiration 4 mission. Additionally, increased DNA methylation was detected, a hallmark of DNA damage repair activity. This methylation was localized near genes related to programmed cell death, while there was a local depletion near genes encoding for both cell cycle and DNA damage repair, suggesting a mechanism of purifying selection. Additional analysis of the transcriptional changes revealed a response that appears specific to UVB radiation. While previous research has characterized spaceflight-induced changes, our study is the first to obtain a combined signature comprising DNA damage profiles, epigenetic modifications and transcriptional changes beyond LEO. The data generated from the DSRG experiment provides a crucial foundation for assessing the molecular risks to DNA and RNA, and for developing strategies for long-duration spaceflight, This first comprehensive analysis of eukaryotic genomic and physiological responses to the unique conditions of deep space provides a resource that will be critical to the success of the Artemis program.

High-speed cell partitioning through reactive machine learning–guided inkjet printing

Technology & Application Development

POSTER 413

Robin Coope[1], Eric Cheng[1], Glenn Chang[2], Haley MacDonald[2], Miguel Ramirez[2], Pamela A. Hoodless[2,3], Adi Steif[1,2,3], Karen C. Cheung[4,5,6]

1. Canada's Michael Smith Genome Sciences Centre, Vancouver, British Columbia, Canada
2. British Columbia (BC) Cancer Research Institute, Department of Basic and Translational Research, Vancouver, British Columbia, Canada
3. University of British Columbia, Department of Medical Genetics, Vancouver, British Columbia, Canada
4. University of British Columbia, Department of Electrical and Computer Engineering, Vancouver, British Columbia, Canada
5. University of British Columbia, School of Biomedical Engineering, Vancouver, British Columbia, Canada
6. University of British Columbia, Centre for Blood Research, Vancouver, British Columbia, Canada

Many single cell assays require partitioning of individual cells so that target molecules may be barcoded, and it is desirable to have high confidence that partitioned cells are truly singlets, to better detect rare cell states. Partitioning individual cells in open nanowells enables both high confidence in single cell occupancy and allows reaction miniaturization to enhance kinetics while providing flexibility in the development of different assays. A challenge for this approach however is to print cells sufficiently quickly to enable experiments of adequate statistical power in a reasonable time. To address this, we developed a single cell dispensing instrument leveraging inkjet technology with continuous real-time optical feedback and a reactive machine learning algorithm for high-throughput single cell isolation. The Isolatrix enables rapid partitioning of cells into open substrates such as nanowell arrays or SBS format plates, permitting high-throughput application of bespoke genomic assays such as direct transposition single cell whole genome sequencing (scWGS). We trained the classifier on manually labelled data with a range of cell lines and applied the instrument to generate scWGS profiles from cultured human cells and primary mouse tissue. Comparison to existing algorithms, as used on commercially available single cell spotters which are based on predicting single cell spotting events, demonstrated that our reactive approach, featuring machine learning classification of events post-dispensing, gives up to a 9.69 times increase in isolation speed. Validation via fluorescent imaging of cell lines confirmed a classification accuracy of 98.7%, at a rate of 0.52 seconds per single cell under tuned spotting parameters. Genomic analysis showed low background contamination and high coverage uniformity across the genome, enabling detection of chromosomal copy number alterations. We are working to deploy the Isolatrix for production cell spotting for single cell genomics and proteomics at our genome centre, and exploring ways to make the technology more widely available.

Pushing the limits: evaluating PacBio AmpliFi performance across diverse invertebrate genomes

Technology & Application Development

POSTER 414

Craig Corton[1], Nicholas Redshaw[1], Fiona Teltscher[1], Caroline Howard[1], Kerstin Howe[1], Jamie Appleyard[2], Ian Johnston[1]

1. Wellcome Sanger Institute, Hinxton, Cambridge
2. Pacific Biosciencies (PacBio), Menlo Park, CA, USA

Long-read sequencing technologies, such as PacBio, have transformed genome assembly by enabling accurate reconstruction of complex genomes. However, generating high-quality libraries from limited DNA remains a major challenge, especially for small or rare specimens. The PacBio AmpliFi protocol was developed to enable low-input long-read sequencing, but its performance across diverse species and genome sizes has not been comprehensively evaluated.

At the Wellcome Sanger Institute, we investigated the robustness of the AmpliFi protocol using a titration experiment across several invertebrate species from the Tree of Life project, representing a range of genome complexities. DNA from each species was prepared and amplified under varying conditions to explore how input quantity and amplification intensity influence library quality. Libraries were assessed for yield, fragment size, and suitability for sequencing.

Across all species, clear patterns emerged linking amplification conditions to the balance between library yield and fragment integrity. These results highlight general principles that can guide the selection of optimal parameters for long-read sequencing from low-input DNA.

This work establishes a practical framework for refining the AmpliFi workflow and improving the efficiency of low-input library preparation. Ongoing optimisation includes testing alternative polymerases, reaction volumes, and size-selection strategies to further enhance performance. By systematically evaluating key protocol variables, we aim to expand the accessibility of long-read sequencing for challenging samples, enabling more comprehensive genomic studies across diverse taxa.

Constellation technology delivers a comprehensive human genome from a variety of DNA inputs leveraging on-flow cell proximity information

Technology & Application Development

POSTER 415

Ali Crawford[1], Louise Fraser[1], Jacqui Weir[1], Mitch Bekritsky[1], Ines Vitoriano[1], Shunhua Han[1], Pascal Grobecker[1], Qing Zhang[1], Drew Ellershaw[1], Vitor Onuchic[1], James Han[1], Peter McInerney[1], Fiona Kaper[1], Joseph Devaney[2], Emily Parker[1]

1. Illumina, Inc., San Diego, CA, USA
2. GeneDx, Inc., Oakland, CA, USA

Constellation is a highly simplified on-flow cell library preparation method on the NovaSeqTM X which delivers an expanded genome by enhancing short read sequencing data with on-flow cell proximity information.

This study evaluates the impact of sample type, DNA integrity, and extraction protocols on the coverage and accuracy of whole genome sequencing. We demonstrate that it is possible to generate highly complete genomes from a variety of different primary samples that have been extracted using different methods. Our findings highlight the importance of using the appropriate collection device combined with the appropriate extraction kit method. Specifically, we show data from blood, buccal swabs and pre-natal testing (CVS, amniocentesis and product of conception) and extracted with a range of on-market extraction kits. We show how DNA fragment size as measured using a TapeStation correlates to coverage in difficult to map regions and phasing performance in the Constellation workflow. We demonstrate that Constellation is still able to significantly improve the completeness of the variant calling above short read sequencing alone for those samples that do not meet quality input requirements for Native Long Read technologies.

SPACE: spatially resolved multiomic analysis for high-throughput CRISPR screening in 3D models

Technology & Application Development

POSTER 416

Yi Cui[1], Mengwei Hu[2], Qianhui Huang[2], Khoi Chu[2], Sierra McKinzie[1], Michael Patrick[1], Sharanya Iyengar[2], Maerjianghan Abuduli[2], Marianne Spatz[2], Nandita Joshi[2], Brendan Miller[2], Shams Vellarikkal[2], Timothy Riordan[1], Danny Bitton[2], Jan Lubojacky[2], Iya Khalil[2], Federica Piccioni[2], Michael Rhodes[1], Alex Tamburino[2], Shanshan He[1], Joseph Beechem[1], Vanessa Peterson[2]

1. Bruker Spatial Biology, Seattle, WA, USA
2. Merck & Co., Inc., Rahway, NJ, USA

High-content single-cell perturbation screens are pivotal for elucidating gene functions and uncovering novel biology, yet conventional methods necessitate cell dissociation, forfeiting critical spatial information essential for dissecting cell-cell interactions and tissue architecture in complex microenvironments. While spatial CRISPR screening mitigates this partially, existing technologies are constrained by hypothesis-driven phenotyping panels limited to sparse RNA or protein coverage, curtailing comprehensive gene function assessment and discovery breadth.

To overcome these barriers, we developed SPACE (SPAtial Cell Exploration), a pioneering platform that fuses whole-transcriptome profiling, CRISPR perturbations, and multiplexed protein detection at single-cell resolution within intact 3D tissue contexts. SPACE delivers unbiased, transcriptome-wide readouts alongside compatibility for up to 76 protein markers, vastly expanding phenotypic landscapes in spatial screens. As the highest-plex multimodal CRISPR assay to date, SPACE achieves this at unprecedented scale and affordability – profiling hundreds of spheroids encompassing tens of thousands of cells per slide – greatly outperforming sequencing-based alternatives in efficiency.

Applying SPACE, we performed CRISPR perturbations across 42 genes in cancer-associated fibroblasts (CAFs) co-cultured with tumor cells in 3D spheroids, yielding multidimensional multiomic datasets from hundreds of spheroids. High-confidence guide RNA detection was coupled with robust endogenous mRNA characterization. Unbiased analyses uncovered new insights on CAF-tumor dynamics: extracellular matrix (ECM) remodeling, spatially resolved ligand-receptor interactions, and perturbation-specific gene variability.

Notably, *ISG20* knockout in CAFs profoundly suppressed multiple matrix metalloproteinases (MMPs) – an unreported link validated orthogonally – implicating *ISG20* in novel ECM regulation and tumor progression. Some perturbations were further revealed to reshape intercellular signaling, revealing knockout-dependent spatial ligand-receptor shifts and coordinated expression signatures that underscore microenvironmental crosstalk in tumor phenotypes.

Culminating in a landmark demonstration, SPACE simultaneously captured whole transcriptomes, CRISPR identities, and 68 protein markers on one slide, enabling holistic perturbation phenotyping. This transformative technology propels spatial CRISPR screening into translational realms, facilitating target and biomarker identification in multicellular models mirroring human tissue intricacy. By merging high-throughput perturbations with spatially resolved multiomics at transcriptome scale, SPACE catalyzes discovery in heterogeneous tissues. SPACE datasets will fuel generative AI models for causal biology inference, accelerating drug discovery and precision medicine.

Technology & Application Development

POSTER 417

Illumina Constellation mapped reads enable high-resolution analysis of genomic diversity and structural complexity

Massimo Delledonne[1], Matteo Orlandi[1], Luca Bertoli[1], Aurelia Iuliana Timis[1,2], Filippo Lucchini[1], Giada Carta[1], Antonina Limongi[1], Drew Ellershaw[2], Jacqui Weir[2], Qing Zhang[2], Melissa Sambrotta[2], Sean Humphray[2], Elena Parrini[3], Enzo Guerrini[3], Davide Mei[3]

1. University of Verona, Department of Biotechnologies, Verona, Italy
2. Illumina Inc., Cambridge, United Kingdom
3. University of Florence, A. Meyer Children's Hospital, Department of Neuroscience, Italy

The pursuit of a single genomic assay capable of delivering complete information across variant classes has remained an aspirational goal in genome biology. To date, researchers and clinicians have relied on a patchwork of technologies, short-read sequencing for SNVs and small InDels, comparative genomic hybridization (CGH) arrays for copy number variation, and optical mapping and long reads for complex rearrangements. This fragmentation creates inefficiencies, increases cost, and limits the full potential of genomic discovery.

Illumina Constellation mapped read technology now provides a powerful one-test-for-all solution, combining the strengths of diverse platforms into a single, streamlined assay. By enabling PCR-free whole-genome sequencing (WGS), the approach delivers high-resolution detection of SNVs and InDels, while also capturing a broad spectrum of copy number variants (CNVs) and other structural variants (SVs). Beyond variant detection, Constellation mapped reads generate long-range proximity information for haplotype phasing, refinement of CNV and SV calls, including very complex rearrangements, and improved mapping in regions of the genome that have historically been intractable with short reads.

We have successfully applied this technology to six clinical samples that were previously resolved using optical mapping, CGH arrays, long reads, and FISH, demonstrating that Constellation mapped reads are not only comprehensive but also clinically actionable. This validation highlights the platform's disruptive potential to consolidate genomic testing pipelines, accelerate diagnosis, and expand access to advanced genomic insights.

Looking forward, the implications extend far beyond clinical genomics. The intrinsic long-range information makes Constellation mapped reads an ideal source of very long-range proximity data, opening new possibilities for telomere-to-telomere (T2T) assemblies and the development of highly resolved pangenomes. These applications will be critical in capturing genomic diversity, resolving structural complexity, and advancing our understanding of genome biology at an unprecedented scale and resolution.

Enabling transcriptome-wide single-cell RNA sequencing from FFPE tissues with Evercode WT combinatorial barcoding

Technology & Application Development

POSTER 418

Ajay Sapre[1], Alec Salvino[1], Guillermo Gallareta-Olivares[1], Vuong Tran[1], Joey Pangallo[1], Vicky Morrison[1], Aisling Sinclair[1], Gokhan Demirkan[1], Charlie Roco[1], Alex Rosenberg[1]

1. Parse Biosciences, Seattle, WA, USA

Formalin-fixed, paraffin-embedded (FFPE) tissues represent a vast and underutilized resource of patient data for retrospective and translational research. While FFPE enables long-term preservation of valuable clinical specimens, RNA degradation and fragmentation have historically hindered its compatibility with single cell RNA sequencing (scRNA seq). Currently, researchers have relied on probe-based FFPE profiling approaches and are limited to predefined target sets, restricting transcriptome coverage and discovery potential.

Here, we introduce a novel split-pool combinatorial barcoding method for FFPE tissue that captures the full transcriptome through dual priming with both poly dT and random hexamer reverse transcription primers. This workflow extends the scalability of Parse's single-cell technology to FFPE tissues, enabling unbiased transcriptome-wide analysis of millions of cells and hundreds of samples within a single experiment. By capturing a broad range of RNA biotypes, this approach delivers deeper insights into cellular function and disease mechanisms.

As a proof of concept, FFPE tissues from multiple organs, including kidney, brain, lymph node, liver, and breast tumor, were processed using our novel Evercode Whole Transcriptome (WT) for FFPE assay. The resulting single cell datasets demonstrated high-resolution identification of cell types across diverse tissue types. Comparative analyses between FFPE and matched fresh frozen samples revealed concordance in gene expression patterns and cell type proportions, validating the approach's complete capture of sample biology.

By enabling transcriptome wide scRNA-seq from thousands to millions of cells from FFPE-preserved tissues across hundreds of samples simultaneously, this method greatly expands access to single cell analyses of archival specimens, offering a powerful platform for translational oncology, biomarker discovery, and precision medicine research.

Enhanced whole-genome sequencing using a single, streamlined workflow builds dynamic, comprehensive coverage for population genomics and oncology research studies

Technology & Application Development

POSTER 419

Nick Downey[1], Katelyn Larkin[1], Ashley Dvorak[1], Tyson Clark[2], Nika Iremadze[2], Ariel Jaimovich[2]

1. Integrated DNA Technologies, NGS Applications Development, Coralville IA, USA
2. Ultima Genomics, Fremont, CA, USA

Whole genome sequencing (WGS) has become a powerful tool for comprehensive genomic analysis, offering an unbiased view across the entire genome. However, for more insightful research, WGS alone often lacks the depth of coverage required in specific genomic regions of interest, such as population genomics, oncological, or disease research-related variants. Conversely, targeted sequencing assays deliver the necessary depth but lack the breadth and scalability of WGS. There is thus a critical need for an integrated approach that unifies the simplicity and comprehensiveness of WGS with the precision and depth of targeted capture, enabling both the utility of genome wide information and reliable coverage of regions required for deeper insight.

Here, we demonstrate a single, tunable Enhanced Whole Genome (eWGS) workflow that eliminates the conventional need to separate libraries. By eliminating the splitting and re-pooling of libraries we provide a method that allows easier sample tracking, and decreases the number of steps and intervention points thus reducing total turnaround time and risk of sample mishandling. The built-in workflow flexibility includes support for a wide range of hybrid capture panel sizes, and tunable coverage levels for both whole genome and targeted regions of interest This is demonstrated for both mechanical and enzymatic fragmentation DNA library prep kits and with cost-efficient high throughput sequencing.

eWGS libraries were sequenced, targeting either 5X mean WGS coverage and 100X mean target coverage for exome to predict population genetics studies, or 30X mean WGS coverage and 1000X mean target coverage of a 536 gene pan-cancer panel for somatic mutation detection. Each enhanced data set shows that tuned, desired sequencing coverage was met for whole genome and targeted regions, and known variants were correctly identified in the majority of samples and replicates. We were able to detect both large rearrangements as well as SNVs with allele frequencies below 2%. Overall, we demonstrate a single-workflow eWGS process that generates high-quality data suitable for population genomics applications and accurate variant detection capabilities for oncology research.

Integrated 3D genome, transcriptome, and proteomic assays: simultaneously observing genome architecture and transcriptional activity at single-cell resolution

Technology & Application Development

POSTER 420

Jude Dunne[1], David King[1], Nagasangeethavani Anandakrishnan[1], Johnson Huynh[1], Serdar Tulu[1], Marc Glazer[1], Brian Smart[1], David Castillo[1], Kenny Chung[1], Huy Nguyen[1], Shyamtanu Chattoraj[1], Doug Werner[1]

1. Bruker Spatial Genomics, Research and Development, San Jose, CA, USA

Variation in chromatin structure including whole or partial chromosome copy gain or loss, genome openness or compartment switching, and TAD boundary disruptions can directly alter transcriptional activity and protein expression. Amplification of oncogenes, deletion of regulatory elements, and misfolding of chromatin domains are known to drive gene expression changes in cancer and other diseases. This link between genome structure and gene expression, while being of fundamental importance, is rarely made on the same sample. Measurements are often made and correlations inferred from data gathered on replicate or parallel samples, often in bulk, losing the resolution coming from viewing both systems simultaneously at the single cell and homolog level.

To address this gap, we present a multiomic assay on the PaintScape™ system using our OncoPaint™ Oncogenic Pathways Panel with associated transcriptome targets. This method enables measurement of in-situ 3D genome architecture and transcript levels in the same cells and on the same chip. Optical barcodes target over 1000 genomic loci across the genome, with dense coverage of 45 gene regions spanning several signaling pathways important in cancer progression Probes for associated transcripts are included. Automated imaging decodes DNA and RNA targets in-situ, preserving nuclear context and enabling integrated analysis of structural and functional features in the same cells.

Expanding this capability, we incorporated spatial proteomics targeting a focused set of cell type, cell state, and structure-associated proteins. These measurements are performed in the same cells, allowing direct comparison of chromatin architecture to transcriptional activity and protein expression.

Assaying HER2+ breast cancer cell lines HCC1954 and SKBR3 we observed variation in chromosome copy number such as gain of Chr8q in HCC1954 and higher ploidy counts in Chr7 and Chr20 in SKBR3. We observed inter- and intra-chromosomal translocations and proximity of multi-loci interactions, some identified as ecDNA (e.g. CCND1 in HCC1954). We visualized 3D folding of oncogenic TADs and disruption of TAD boundaries (e.g. CDK1 in SKBR3). Expression varied across cells and populations, aligning with pathway-level dysregulation in cell cycle, apoptosis, transcriptional regulation, and chromatin remodeling. This integrated approach enables simultaneous measurement of genome structure, transcript expression, and protein context in single cells.

Constellation mapped read technology enables phased variant calling with accurate long-distance phasing in short-read sequencing

Technology & Application Development

POSTER 421

Christopher Edlund*[1], Daniel Andrews[2], Mitch Bekritsky[1], Severine Catreux[1], Ali Crawford[1], Kevin Durant[2], Andrea Finocchio[2], Fiona Kaper[1], James Han[1], Konrad Scheffler[2]

1. Illumina Inc., San Diego, CA, USA
2. Illumina Inc., Cambridge, Cambridgeshire, United Kingdom

*Corresponding author

Determining variant phase and parental haplotype assignment from sequencing data is critical for understanding the complete biological context of genomic variation. This knowledge is crucial for accurately interpreting compound heterozygous variants in autosomal recessive disease, understanding familial inheritance patterns, and determining the combined functional effects of variants in a gene. Standard short-read whole-genome sequencing (srWGS) alone cannot resolve phase in a proband, which necessitates the addition of parental samples or alternative platforms with higher costs and longer turnaround times. To address this, we developed a method that leverages Illumina Constellation mapped read technology to extract long-range haplotype information directly from srWGS data, enabling robust phasing without the need for parental samples or workflow changes.

Traditional approaches to phasing variant calls start by calling unphased variants and adding phasing information in a subsequent step. Here we present a novel approach that first phases reads to inferred ancestral haplotypes and then uses the phased reads to perform phased variant calling. This approach personalizes the reference genome by selecting the most closely related haplotype segment pairs from a haplotype database (hapDB), while considering recombination rates and simultaneously phasing reads to the inferred haplotypes. It leverages long-range Constellation proximity data to provide probabilistic yet informative assignments of reads to specific haplotypes. It also infers long phase blocks, defined as genomic segments within which the mapping from inferred haplotype segments to paternal/maternal origin remain consistent with high probability. Finally, the probabilistically phased reads are used for phased variant calling (e.g. treating 0|1 and 1|0 as distinct genotype hypotheses) where calls are phased with respect to each other if they are phased to the same ancestral haplotype and share the same phase block. Unlike conventional pipelines, this method of early phasing improves both speed and accuracy. We report Hamming error rates below 1% using the HG002 T2T truth set, while fully phasing of over 90% of genes. Collectively, this work represents a paradigm shift in genome analysis workflows, offering a unified, probabilistic framework for mapping, phasing, and variant calling, and for the first time allowing long distance phasing from short-read WGS data.

RNA exome sequencing: maximizing insights into fusion calling paired with gene expression analysis in a single, targeted workflow using low-input RNA

Technology & Application Development

POSTER 422

Amy E. Elias[1], Sandra Theall[1], Nikita Raymond Dsouza[1], Shobana Sekar[1], Mariana Fitarelli-Kiehl[1], Jieqiong Dai[1], Alejandro Quiroz Zarate[1], Lindsey Cambria[1]

1. Roche Diagnostics, Roche Sequencing & Life Science, Wilmington, MA, USA

While widely adopted for comprehensive transcriptome profiling, the methods of poly-A selected or ribodepleted RNA-sequencing may have key limitations, including input mass requirements, inefficient fusion detection, and sequencing capacity consumed by non-coding transcripts. RNA fusions and exon-skipping events, often occurring in cancer cells in low abundance, require targeted enrichment for reliable detection. In this study, we demonstrate a new workflow that can enrich RNA-seq libraries for protein-coding transcripts and identify common fusions by utilizing the KAPA RNA EvoPrep Kit with KAPA HyperExome V2 Probes, which can be successfully used for both DNA and RNA targeted applications. This 1.5-day workflow, with limited hands-on time, was tested with both high-quality RNA and formalin-fixed paraffin embedded (FFPE) RNA using several input amounts (1, 10 , and 50 ng).

Sequencing results demonstrate that this new workflow:

1. Enriches for biologically meaningful genes, with over 94% of genes having protein-coding function, independent of input RNA quality, type, and input mass.
2. Identifies over 89% of genes as not significantly differentially expressed when compared to traditional workflows.
3. Reduces input mass requirements while retaining 90-96% concordance between high-quality and FFPE RNA samples.
4. Enables detection of many common RNA fusions and exon-skipping events such as EGFR variant III and MET 14 exon skipping.

Altogether, these results demonstrate that the RNA-exome sequencing workflow used in this study can be leveraged to enrich exonic sequences from as little as 1 ng RNA and to identify fusions from as little as 10 ng RNA and 30 million read pairs.

For Research Use Only. Not for use in diagnostic applications. The RNA-exome sequencing workflow has been developed with high-quality and FFPE reference materials. Not a validated workflow at this time.

A single-molecule PCR framework bridging low-plex PCR and NGS for biomarker validation at 16-plex and beyond

Technology & Application Development

POSTER 423

Christina Fan[1], Michael Porter[1], Jung Won Keum[1], Huean Lee[1], Vladyslava Ratushna[1], Christina Cheng[1], Andras Saftics[1], Mehdi Molaei[1], Janice Lai[1], Eleen Shum[1]

1. Countable Labs, Inc., Palo Alto, CA, USA

Next-generation sequencing (NGS) has enabled the discovery of thousands of candidate biomarkers across a wide range of diseases and traits. Translating these discoveries into practical, quantitative assays remains a major challenge. Conventional qPCR and dPCR platforms are widely used for low-plex validation, but their limited multiplexing capacity constrains scalability. Moreover, high-plex PCR assays typically require extensive optimization to balance amplification efficiencies and minimize cross-reactivity, limiting their utility for large-scale validation and clinical assay development. A quantitative platform that combines high multiplexing, digital precision, and efficient use of limited samples is needed to close this gap.

We describe a single-molecule quantitative PCR approach, Countable PCR, designed to address these challenges. The method isolates and amplifies up to one million individual DNA molecules in parallel without sample loss, followed by high-resolution 3D light-sheet imaging for digital counting. Operating in the single-molecule regime enables precise quantification and straightforward multiplexing, as each molecule amplifies independently without inter-target competition. Using ten spectrally distinct fluorophores, this system detects ten independent targets simultaneously on a four-color imaging platform. By applying combinatorial fluorophore labeling, multiplexing is extended beyond twenty targets within a single reaction.

We demonstrate applications including a >20-plex respiratory pathogen detection panel and a 16-plex gene expression panel, both developed with minimal optimization. Quantitative results show concordance with single-plex assays and maintain high target specificity. Data from these panels were further analyzed using multivariate and high-dimensional visualization methods such as UMAP, analogous to approaches used in single-cell transcriptomic analysis, enabling intuitive interpretation of multiplexed results.

This single-molecule PCR framework establishes a scalable strategy for high-plex quantitative analysis with NGS-comparable precision. It provides a generalizable approach for accelerating biomarker validation and supports the development of multi-gene assays for cancer subtyping, prognostic scoring, pathogen detection, and other applications requiring rapid, high-plex quantification.

High-throughput single-nucleus spatial mapping of whole-genome copy number profiles

Technology & Application Development

POSTER 424

Andrew Farmer[1], Xuan Li[1], Wanxin Wang[1], Cedric Uytingco[1], Raymond Mendoza[1], Samantha Leong[1], Christina Chang[1], Anaram Shahravan[1], Yue Yun[1]

1. Takara Bio USA, Research and Development, San Jose, CA, USA

In cancer, clonal tumor populations evolve and accumulate heterogenous genetic aberrations, including chromosomal rearrangements, copy number variations (CNVs), and single nucleotide variations (SNVs). Spatially resolving genetic aberrations is essential for understanding the evolution, progression, and migration of cancer. However, current spatial omics platforms cannot yet capture genome-wide aberrations, primarily due to the lack of true single-cell spatial technologies and high-throughput single-cell whole-genome amplification (WGA) methods. Here, we combined TBUSA's TrekkerTM spatial single-cell mapping platform with our Shasta WGA technology to enable spatial mapping of whole-genome CNVs at single-nucleus resolution.

Using Trekker, nuclei are labeled with spatial barcodes released from a microbead array upon UV exposure. Individual tagged nuclei are then dispensed using the Shasta® Single Cell System, which provides automated nanoliter-scale dispensing and library preparation, to generate WGA libraries with unique single-nucleus barcodes for over 1,300 nuclei per chip. The WGA reaction employs quasi-random primers for highly uniform and reproducible genome amplification using only 300 nL of reagents per nucleus. Following sequencing, CNV profiles are reconstructed from single-nucleus barcodes and mapped to their spatial coordinates, enabling direct association of genome-wide variations with pathological features captured in H&E images.

To demonstrate utility of the platform, 1,312 spatially barcoded nuclei from a 25-um fresh-frozen mouse kidney section were processed in one chip. The WGA libraries were sequenced to a median depth of 250,000 reads per nucleus, and single-nucleus CNV profiles were derived at 1-MB bin resolution. Sequencing data showed high mapping rate, low duplicate rate, and comparable genome coverage uniformity as control cell lines. Nuclei were successfully mapped with spatial barcodes, establishing a spatial CNV map of the mouse kidney section. The workflow was further validated on tumor FFPE samples. To enhance throughput, a dual-chip dispensing module was implemented, enabling processing of up to 4,000 nuclei per run. The combined Trekker/Shasta WGA platform provides the first end-to-end solution for spatially mapping whole-genome copy number profiles at single-nucleus resolution, offering a powerful new approach for advancing cancer research.

Spatial colocation-enhanced reconstruction of complex structural variants using Constellation mapped read sequencing

Technology & Application Development

POSTER 425

Jeffrey Fisher[3], Qing Zhang[1], Drew Ellershaw[1], Xinchang Zheng[2], Fabian Klotzl[1], Sera Choi[1], Martina Mijuskovic[1], Bryce Kille[3], Sean Truong[3], Mitchell Bekritsky[3], James Han[3], Fritz Sedlazeck[2], Fiona Kaper[3], Cande Rogert[3]

1. Illumina Cambridge, Cambridge, United Kingdom
2. Baylor College of Medicine, Waco, TX, USA
3. Illumina, Inc., San Diego, CA, USA

Resolving structural variants (SVs) is a significant genome analysis challenge, particularly with regards to interchromosomal and intrachromosomal rearrangements and complex SVs. These events are challenging for any sequencing technology, with only recent studies demonstrating the capacity of long-read sequencing to reconstruct complex SVs. When extending beyond sequencing, molecular assays, including FISH, chromosomal microarrays, and karyotyping, are limited to gross structural rearrangements. The absence of reliable assays and methods to resolve these rearrangements prevent a complete understanding of the mechanisms by which they contribute to disease. This impedes molecular diagnosis, extends diagnostic odysseys, and impairs a more comprehensive picture of the contribution of SVs to genetic disease.

We present a novel method for reconstructing complex SVs using Constellation mapped read technology ("Constellation"). This method excels in resolving complex SV architectures, including multi-breakpoint events and heterogeneous rearrangements, which are typically refractory to short-read-based pipelines. We present complex SV reconstruction with Constellation across a range of samples with challenging rearrangements, including 30 complex multibreakpoint SVs, 6 ring chromosomes, 40 translocations, and 25 inverted insertions/duplications and inversions, comprising both balanced and unbalanced events.

Included in this set are 3 previously characterized MECP2 trios with a tandem duplication, an inverted duplication, and a complex duplication-triplication-duplication structures, where Constellation successfully confirms the disease-associated rearrangements. Constellation's long-range information enables reconstruction of multi-kilobase templates, offering a more comprehensive view of genomic architecture.

Additionally, the absence of spatially proximal reads in colocation plots can serve as a negative indicator, filtering false positives events which rely on read-level evidence alone. We used this property to improve intrachromosomal and intrachromosomal rearrangement calling, reducing the triage burden needed to identify critical SVs, including balanced and unbalanced translocations. In a set of 40 samples, Constellation can reduce the number of candidate translocation events by up to 80%, while maintaining high sensitivity.

Constellation enables long-range resolution that is broadly applicable to population genomics and rare disease research. By expanding the capabilities of short-read sequencing, Constellation mapped read represents a transformative advance in complex structural variant detection and curation.

For research use only. Not intended for diagnostic purposes.

Benchmarking short-read sequencing technologies using donor-specific assemblies

Technology & Application Development

POSTER 426

Christian D. Frazar[1], Sean R. McGee[1], Mitchell R. Vollger[5], Youngjun Kwon[2], Katherine M. Munson[2], Kendra Hoekzema[2], Erica Ryke[1], Joshua D. Smith[1], Nancy L. Parmalee[3], William Harvey[2], Jimmy Bennett[3,4], Andrew B. Stergachis[2,5,7], Evan E. Eichler[2,6,7], Chia-Lin Wei[1]

1. University of Washington, Northwest Genomics Center, Department of Genome Sciences, Seattle, WA, USA
2. University of Washington, Department of Genome Sciences, Seattle, WA, USA
3. Seattle Children's Research Institute, Center for Developmental Biology and Regenerative Medicine, Seattle, WA, USA
4. University of Washington, Department of Pediatrics (Genetics), Seattle, WA, USA
5. University of Washington, Department of Medicine, Division of Medical Genetics, Seattle, WA, USA
6. Howard Hughes Medical Institute, Seattle, WA, USA
7. University of Washington, Brotman Baty Institute for Precision Medicine, Seattle, WA, USA

The rapid advancement of sequencing technologies has delivered unprecedented insights into the biology of human genome. Most strikingly, the proven single-molecule long-read sequencing technologies have delivered the complete, gapless T2T human genome reference, while significant improvement of efficiency in the production of high-output, highly accurate short-read data has accelerated the scale and resolution of genome variation analyses. We have evaluated the performance of sequencing chemistries from six commercial and emerging platforms for their ability to accurately identify human genetic variation. We produced 18 high quality sequencing datasets from Element AVITI, PacBio ONSO, Roche SBX, Illumina NovaSeq6000/NovaSeqX, Ultima UG100/ppmseq and MGI T7 using same batch benchmarking DNA from HG002 and COLO829BL as part of the Somatic Mosaicism across Human Tissues consortium. Sequencing data were mapped to a GRCh38 reference genome as well as high-quality diploid donor specific assemblies (DSA), respectively for systematically evaluating the sequencing error profiles. The use of a DSA effectively removes majority genetic variation and allows us to examine true sequencing errors. The sequencing output, quality, bias, error rates and modes were objectively evaluated. We found that PacBio ONSO has the lowest single base mismatch rate at less than 0.5 per Kb. As well, the duplex sequencing modes of both Ultima ppmSeq and Roche SPX-D platforms significantly improve accuracy. Small indel mismatch rates were also studied. The Element AVITI platform with UltraQ chemistry showed the lowest indel mismatch rates at approximately 0.002 per Kb. If base quality is taken into consideration, mismatch rates improve by almost an order of magnitude when considering only the highest quality base calls. Finally, we show that base score recalibration using GRCh38 worsens the base quality scores for selective sequencing technologies. These results provide valuable insight for studies looking to accurately study low allele fraction variants.

Next-generation karyotyping: WGA-free single-cell CNV profiling

Technology & Application Development

POSTER 427

Emilis Gegevicius[1], Domas Rupkus[1], Emile Pranauskaite[1], Paulius Matulis[1], Vaida Kurmauskaite[1], Jan Bergmann[1], Simonas Jocys[1], Zana Zdaniene[1], Rapolas Zilionis[1]

1. Atrandi Biosciences, New York, NY, USA

Copy number variation (CNV) is an important source of genetic variability in humans, influencing both normal and deleterious phenotypic traits. In cancer, the prevalence of CNVs has prompted comprehensive studies of CNV landscapes across tumor types. In the field of cell and gene therapy, genomic stability of engineered cell lines is a key indicator of clinical safety. Conventional karyotyping methods, such as G-banding or molecular techniques based on hybridization arrays, often fall short due to low throughput or reliance on bulk genomic DNA, which obscures cellular heterogeneity. Recent advances in next-generation sequencing and single-cell sample preparation have significantly improved data resolution and depth. Single-cell DNA sequencing protocols that employ whole-genome amplification (WGA) are now available; however, their reliability and cost are not yet optimal.

Here, we present a genome-wide CNV profiling technique based on the use of semi-permeable capsules (SPCs) for cell isolation. SPCs enable easy multi-step processing of cellular genetic material and support a convenient and scalable combinatorial barcoding. Importantly, we demonstrate the feasibility of single-cell whole-genome library preparation without DNA pre-amplification, thereby alleviating data analysis challenges related to amplification bias.

As a proof of concept, we conducted six barcoding experiments, processing approximately 60,000 cells from the GM12827 and K562 cell lines, and deeply sequenced a subset of 1,800 cells. GM12827 exhibits a stable diploid karyotype, while K562 is hypotriploid. We used the AneuFinder tool to analyze CNV profiles of individual cells and compare the results with the pseudo-bulk dataset. The best concordance in karyotype calls across per-cell coverage depths was observed at a 5 Mb resolution. The resulting karyotypes aligned well with expected cellular states. Notably, high-throughput single-cell analysis revealed rare subpopulations with CNV profiles that deviated from the pseudo-bulk consensus. For biological validation, we applied this protocol to nearly 2,000 cells from acute myeloid leukemia (AML) patient samples and identified rare populations with clinically significant aberrations that would likely be missed using bulk approaches.

Collectively, this study expands the molecular genotyping toolkit, offering a robust and scalable method for exploratory genotyping that supports both disease characterization and quality control of engineered cell lines.

High-purity single-cell RNA sequencing of circulating trophoblasts directly from whole blood

Technology & Application Development

POSTER 428

Martin Cienciala[1], Christopher Kan[1], Daniel Georgiev[1]

1. Sampling Human Inc., Department of Research & Development, Berkeley, CA, USA

Circulating trophoblasts (cTBs) provide a direct, cell-resolved view of fetal and placental biology, yet their extreme rarity and fragility have made them inaccessible by standard single-cell workflows.

We developed a one-hour protocol that captures and barcodes mRNA from 10 mL of whole blood without red-blood-cell (RBC) lysis, RBC depletion, or separate encapsulation. The chemistry unifies trophoblast isolation and reverse transcription within a single reaction, eliminating cell loss and preserving fragile fetal cells.

In a pilot study of five pregnant donors (gestational ages 9–15 weeks), we recovered an average of 0.8 cTB per mL of whole blood (approx. 8 cells per 10 mL sample) with an average trophoblast purity of 52% and an estimated maternal transcript contamination below 0.9 transcripts per cell, well below the detection threshold. The high purity of the genetic material enabled direct sequencing without any additional enrichment or purification steps, further minimizing handling artifacts. Median expression depth reached 8,340 UMIs per cell and 1,960 genes per cell, approximately fourfold and 1.7-fold higher than paired PBMC controls (3,010 UMIs per cell, 1,180 genes per cell), respectively. Differential-expression analysis identified distinct trophoblast populations. Fetal origin was confirmed from heterozygous single-nucleotide variants and Y-linked transcripts in male pregnancies, derived from cDNA sequence data.

These results demonstrate that lineage-resolved single-cell transcriptomes can be obtained directly from unprocessed maternal blood with high purity and sensitivity. This capability enables non-invasive access to intact fetal cells and establishes a scalable foundation for next-generation single-cell prenatal diagnostics.

A novel high-plex spatial multi-omics platform coupled with temporally precise depletion reveals a new role for the ASD-linked gene *SETD5* in human stem cells

Technology & Application Development

POSTER 429

Alon Goren[1], Dube Umber[1], Eugin Destici[1], Courtney Sawyer[1], Angels Almenar[1], Andrew Boddicker[2], Semyon Kruglyak[2], Ben Song[2], Lingzhi Zhang[1]

1. University of California at San Diego (UCSD), The Center for Neural Circuits and Behavior, San Diego, CA, USA
2. Element Biosciences, San Diego, CA, USA

Haploinsufficiency of the chromatin regulator SETD5 causes intellectual disability and an autism spectrum disorder. However, the immediate cellular consequences of acute SETD5 loss in a human developmental context remain poorly understood. To address this, we engineered human induced pluripotent stem cells (hiPSCs) using CRISPR-Cas9 to endogenously fuse a degradation tag (dTAG) to the *SETD5* gene. This system enables rapid, specific, and reversible depletion of SETD5 protein upon addition of the small molecule dTAG-13, providing precise temporal control over protein function. To characterize the effects of acute SETD5 loss in hiPSCs, we employed a novel high-plex spatial multi-omics platform that allows simultaneous quantification of protein and RNA expression with subcellular resolution. Inducible degradation of SETD5 resulted in profound and rapid alterations to the cellular landscape as well as changes to nascent transcription – these are detected by using the ratio of nuclear/cytoplasmic transcripts and the observed changes were validated by csRNA-seq (capped small RNA-seq), an orthogonal, bulk level approach for capturing nascent transcription. Spatially resolved analysis revealed a significant disruption of the cell cycle as a primary consequence of SETD5 depletion. Specifically, treated hiPSCs exhibited a notable accumulation of cells in the G2/M phase, coupled with a corresponding decrease of cells in the G1/S phases. This cell cycle arrest was underpinned by the coordinated downregulation of key proliferative drivers, (e.g., CDK1, CDK2, CDKN1B and RB) and the upregulation of cell cycle inhibitors at the protein levels. Further we detected changes in the transcript levels of multiple ASD-linked genes such as *SEMA5A*, *ROBO2* and *NLGN1* following the depletion of SETD5. These changes were also appreciable at the level of cell morphology. Our findings demonstrate that acute SETD5 loss directly impairs cell cycle progression and nascent transcription of ASD-linked genes in hiPSCs, providing a potential mechanistic basis for the neurodevelopmental defects observed in patients. Altogether, our experimental system, combining inducible degradation with high-resolution spatial profiling, presents a powerful framework for dissecting the function of critical developmental genes.

Single-step library preparation workflow enables high-throughput low-pass whole-genome sequencing without compromising accuracy

Technology & Application Development

POSTER 430

Sabina Gude[1], Samuel Widmayer[2], Mary Barter[2], Qingchang Meng[2], Dan Gatti[2], Maura Costello[1], Yanyan Liu[1], Michelle Rahardja[1], Rebecca Feeley[1], Jenna Couture[1], Stella Huang[1]

1. seqWell, Inc., Beverly, MA
2. The Jackson Laboratory, Bar Harbor, ME

Accurate genotyping is fundamental to population genetics, enabling the characterization of genetic diversity within and between populations. In recent years, genotyping-by-sequencing (GBS) has increasingly replaced microarray-based approaches due to its ability to detect a broader spectrum of genetic variants and the declining cost of sequencing technologies. Low-pass whole-genome sequencing (lpWGS), performed at 0.1–5× coverage and combined with statistical imputation, has emerged as a cost-effective strategy for genotyping large cohorts while still capturing genome-wide variation.

A major bottleneck in scaling lpWGS to hundreds or thousands of samples lies in the complexity of library preparation workflows. Traditional multi-step protocols are labor-intensive, time-consuming, and susceptible to contamination and batch effects, which limit scalability and reproducibility in large-cohort studies.

ExpressPlexTM Plus addresses these challenges with a streamlined, single-step library preparation workflow that integrates tagmentation and amplification within a single thermal cycling reaction, minimizing hands-on time. The method also automatically normalizes libraries across a tenfold range of input DNA concentrations, eliminating the need for pre-quantification and manual normalization.

In this study, we benchmarked genotyping accuracy from ExpressPlex Plus libraries at 0.5× coverage against those generated using a conventional multi-step workflow. Genotype imputation sensitivity and accuracy, as assessed by F1 scores and related metrics, were comparable between methods. However, the ExpressPlex Plus workflow was completed in under 100 minutes, compared to over three hours for the multi-step protocol, demonstrating a significant improvement in efficiency without compromising genotyping performance.

These findings demonstrate that ExpressPlex Plus provides a technically robust and time-efficient approach for low-pass WGS library preparation. The comparable imputation accuracy to conventional workflows, combined with a substantially reduced preparation time, supports its application in large-scale genotyping studies where throughput, reproducibility, and cost-effectiveness are critical.

Technology & Application Development

POSTER 431

Haplotype-resolved variant calling in paralogous 'dark' regions with Illumina Constellation

Shunhua Han[1], Vitor Onuchic[1], Varun Jain[1], Arun Subramaniyan[1], Zarko Tasev[1], Stephen Gaffney[1], Fabian Klotzl[1], Drew Ellershaw[1], Sean Humphray[1], Ali Crawford[1], Mitch Bekritsky[1], Jacqui Weir[1], James Han[1], Joseph M Devaney[2], Aaron Johnson[2], Keith Nykamp[2], Julianna Spangler[2], Leah Matzat[3], Dalia Dhingra[3], Ashley Sirard[3], Joy Hwang[3], Nathan Usem[3], Arun Ramani[3], Nirav Malani[3], Deborah Boles[3], Robert Daber[3], Katie Larkin[4], Yueyao Gao[4], Michelle Cipicchio[4], Mark Fleharty[4], Victoria Popic[4], Sean Hofherr[4], Marina DiStefano[4], Diana Toledo[4], Niall Lennon[4], Inga Nagel[5], Inga Vater[5], Malte Spielmann[5]

1. Illumina, Inc., San Diego, CA, USA
2. GeneDx, LLC, Gaithersburg, MD, USA
3. Labcorp, San Francisco, CA, USA
4. Broad Clinical Labs, Cambridge, MA, USA
5. University Hospital Schleswig-Holstein, Kiel & Lübeck, Germany

Paralogous, medically relevant loci remain difficult for standard short-read sequencing because of high homology. Clinical workflows often require targeted assays (MLPA, LR-PCR) that rarely provide haplotype resolution at large distances, or long-read reflex testing, adding cost and days of turnaround. We present Illumina's Constellation mapped read technology with new informatics that can utilize long-range proximity from the original DNA molecule and generate haplotype-resolved genotypes in "dark" regions *de novo*, using short reads, within three days from DNA extraction.

We evaluated Constellation across 31 clinical samples spanning *SMN1/2, PMS2, CYP21A2, STRC, NCF1, IKBKG, CFC1,* and *HYDIN* and assessed pathogenetic variant detection performance under various genome complexities, including recombination, gene conversion, and copy number changes. These regions have 98.2-99.9% homology and sizes vary from 18-357kb. Constellation matched or exceeded long-read sequencing and standard-of-care (SOC) performance while providing fully phased haplotypes.

In nine spinal muscular atrophy cases, we produced fully phased *SMN1* and *SMN2* copies and accurately quantified *SMN1* copy number in homozygous (n=4) and heterozygous (n=5) deletion backgrounds.

In seven Lynch syndrome cases, we resolved all haplotypes across the high-homology *PMS2/PMS2CL* region and correctly assigned pathogenic variants to *PMS2* versus *PMS2CL*. A tandem duplication in *PMS2CL* was resolved which MLPA could not.

In five congenital adrenal hyperplasia cases, we reconstructed the RCCX locus, distinguished *CYP21A2, CYP21A1P,* and fusion copies, and delivered haplotype-resolved calls exceeding MLPA and LR-PCR resolution. In one previously unresolved affected proband, we detected two *CYP21A1P* and one *CYP21A2* copies on one haplotype, and two *CYP21A2* copies on the other, all carrying the pathogenic variants, providing molecular confirmation.

In a hearing loss case, the phenotype indicated an affected individual, yet the SOC assay reported a carrier based on a copy-number deletion of *STRC*. Constellation detected a complete deletion of one *STRC* allele and a pathogenic small variant in the other, confirming the affected status.

Collectively, Constellation resolved known pathogenetic variants in 31/31 cases across eight paralogous genes. These data support that Constellation provides a unified, genome-wide assay that brings historically "dark" regions into routine clinical research while reducing the need for orthogonal workflows with associated complexity, and time to result.

Pre-validation of HiFi libraries for transitioning PacBio long-read sequencing to clinical space

Technology & Application Development

POSTER 432

Maria Elena Hernandez Gonzalez[1], Thanaphat Auwattanamongkol[1], Huachun Zhong[1], Lakshmi Rao Venkata[1], Eileen Stonerock[1], Katherine Miller[1], Tracy Bedrosian[1], Daniel Koboldt[1], Catherine Cottrell[1], Benjamin Kelly[1], Richard Wilson[1], Elaine Mardis[1], Anthony Miller[1]

1. Nationwide Children's Hospital, The Institute for Genomic Medicine, Columbus, OH, USA

Structural variants are increasingly recognized as a cause of pediatric genetic disease but remain difficult to resolve with short-read sequencing. Long-read genome sequencing, such as Pacific Biosciences (PacBio) methodology, improves detection of these events and enables haplotype phasing and methylation analysis by sequencing long native molecules in a PCR-free manner.

The Institute for Genomic Medicine at Nationwide Children's Hospital, has implemented PacBio sequencing in the research and translational space. To evaluate clinical feasibility, we compared PacBio recommended library preparation workflow against an in-house protocol designed to optimize recovery and balance turnaround time with sequencing quality. PacBio recommends a 500 ng of input DNA, short-read eliminator treatment, shearing to 15–20 kb, 30 min incubations for end repair and ligation, 15 min nuclease treatment, and final AMPure PB bead size selection (>5 kb), with reported average final library recovery of ~15% from input DNA.

In contrast, our in-house method doubled the input to 1,000 ng, incorporated shearing on the Megaruptor3 targeting 27-30 kb when needed. We performed extended incubations: end-repair 60 min, overnight ligation, 45 min nuclease treatment, and BluePippin size selection 8.5-50 kb. We evaluated this protocol across 15 pediatric samples (oncology, rare disease, epilepsy, and neuromuscular disease) derived from blood, bone marrow, tissue, and cells. Starting gDNA quality varied widely, with mode sizes ranging from 13.6 kb to >60 kb, reflecting the true variability of clinical samples, yet strong recoveries were consistently achieved. We observed an average library recovery of ~22.4% (11.7–31.5%), higher than the PacBio protocol despite the more stringent size cutoff. Sequencing on the Revio system with SPRQ chemistry at 101–175 pM loading achieved HiFi yields of 63.6–130.5 Gb (mean 103.5 Gb), mean read lengths of 12.6–19.5 kb (average 16 kb), and mapping coverage of 20–41x, with 11 of 15 samples (73.3%) surpassing 30x genome coverage.

This work demonstrates that targeted protocol modifications can increase library recovery and sequencing performance, even when starting material is of suboptimal quality. By benchmarking against the manufacturer's SOP, we provide a framework for clinical laboratories to consider trade-offs in input, recovery, and turnaround time when adopting long-read sequencing for pediatric genomic medicine.

Mapping organelle-associated mRNA localization in Alzheimer's disease brain via spatial multiomics

Technology & Application Development

POSTER 433

Ashley Heck[1], Kimberly Young[1], Rachel Liu[1], Lidan Wu[1], Aster Wardhani[1], Claire Williams[1], Margaret Hoang[1], Sayani Bhattacharjee[1], Courtney Anderson[1], Prajan Divakar[1], Mirko Corselli[1], Joseph Beechem[1]

1. Bruker Spatial Biology, Seattle WA, USA

High-plex spatial imaging of brain cells remains a major challenge due to the extraordinary complexity and scale of neural morphology. Neurons possess axons and dendrites that extend far beyond the cell body, forming vast, branching networks. This organization requires precise control of energy metabolism, ionic gradients, and protein localization to support essential functions like signal transmission and neural adaptability. These processes depend on localized RNA translation and organelle positioning in specialized regions such as synapses and dendritic spines. Mitochondria, for example, are enriched at synaptic sites, supporting neuronal communication through energy production and calcium buffering. Dysfunction of mitochondria and other organelles is strongly implicated in neurodegenerative diseases such as Alzheimer's disease (AD), highlighting the need for spatially resolved sub-cellular mapping to understand neural function and pathology.

Existing cell segmentation algorithms primarily focus on the soma, leaving much of the biology-rich architecture of brain cells unprofiled in spatial omics workflows. The CosMx® Spatial Molecular Imager (SMI) addresses this limitation by enabling high-plex detection of proteins and RNAs on the same tissue section. This multiomic approach improves segmentation of neural projections and subcellular features, allowing more accurate assignment of transcripts to individual cells and compartments. The assay begins with detection of 76 proteins using oligonucleotide barcode-conjugated antibodies, followed by protease digestion and transcript detection via barcoded RNA probes. The protein portion targets key organelles (e.g., mitochondria, lysosome), regulators of translation (e.g. P-bodies, ribosomes), cell type markers, and Alzheimer's pathology, while RNA provides a comprehensive view of 18,000+ protein-coding genes.

We demonstrate this multiomic approach through analysis of 6 human AD and healthy brain sections across various brain regions. By integrating protein and RNA, results reveal unprecedented segmentation of neurons and glia, increasing counts per cell and improving annotation of neuronal, glial, and vascular subtypes. Organelle-specific markers enable evaluation of subcellular compartments and associated mRNAs. We observe that mitochondrial abundance and organization differ across brain cell types and are further shaped by spatial proximity to Alzheimer's disease pathology, suggesting local microenvironments influence mitochondrial homeostasis. These findings mark a key advance in quantifying subcellular, spatially resolved gene expression in Alzheimer's pathology at unparalleled multiplexing.

Comparative analysis using three proteomics methods

Technology & Application Development

POSTER 434

Akinori Kanai[1], Yukie Kashima[1], Yutaka Suzuki[1]

1. The University of Tokyo, Graduate School of Frontier Sciences, Kashiwa, Chiba,d Japan

High-throughput proteomics technologies enabling comprehensive quantification of plasma and serum proteins are rapidly evolving as powerful tools for disease biomarker discovery and immune response analysis. This study compared the measurement characteristics and data correlation across three representative platforms—Olink's HT system, SomaLogic's SomaScan 11K, and Illumina Protein (SomaScan 7K ver)—using the same plasma sample cohort.

Analysis of plasma samples from the same subjects (n=75) using three methods showed a Spearman correlation coefficient of r=0.42 between SomaScan and Illumina Protein, indicating good agreement between the two aptamer-based approaches. In contrast, the correlation between Olink HT and SomaScan was r=0.36, and between Olink HT and SomaScan was r=0.27, reflecting differences in detection principles and antibody/aptamer selection, yielding moderate correlations. The correlation coefficients comparing these three methods all exhibited a bimodal distribution, splitting into groups with high and low correlation coefficients. Among the 2,551 proteins common to all three methods, 510 proteins showed a correlation coefficient of $r \geq 0.6$ in all comparisons, indicating high reliability regardless of the system used. These results further suggest that the quality of results obtained may vary depending on the method and detection system used, necessitating careful consideration in method selection and data handling.

We collaborated with Bio Bank Japan, a disease-oriented cohort, to perform over 10,000 Olink assays on serum/plasma samples. The average %CV for Olink HT was 10.1, with r=0.86, yielding highly reproducible data. Data from 3,000 individuals has already been released, representing the largest proteome dataset for the Japanese population. Results from large-scale Olink testing suggest that this method enables highly sensitive and reproducible quantification in multi-sample analyses using clinical samples. Its complementary use with other methods allows for more precise biological protein profiling. This presentation reports detailed correlation analysis results, including comparisons of measurement characteristics, and detection limits across platforms.

Evaluation of Illumina 5-base sequencing for integrated methylation and variant analysis in FFPE samples and its potential for tumor-informed MRD applications

Technology & Application Development

POSTER 435

Stephen Gross[1], Joshua Beytebiere[1], James Baye[1], Antoine Dejong[1], Christine Eng[1], Khai Luong[1], Anthony Mai[1], Mike Mehan[1], Suzanne Rohrback[1], Jeff Tsai[1], Lisa Watson[1], Fiona Kaper[1], James Han[1]

1. Illumina, San Diego, CA, USA

Formalin-fixed, paraffin-embedded (FFPE) tissue represents an abundant and clinically valuable source of genomic material, yet nucleic acid degradation and chemical modifications inherent to FFPE processing pose significant challenges to downstream analyses. The limited quantity of extracted DNA further restricts the ability to perform multiple assays for comprehensive biomarker evaluation. The Illumina 5-base workflow integrates genetic and epigenetic information in a single assay, enabling simultaneous whole-genome variant calling and methylation detection from a single sample prep. Here we present an assessment of 5-base compatibility with FFPE samples.

Genomic and epigenomic variant detection accuracy was evaluated using genomic DNA (gDNA) from fresh frozen tissue or FFPE inputs processed through whole-genome sequencing (WGS) and 5-base workflows. All sample types generated high-complexity libraries with mapping rates exceeding 85%, independent of FFPE DNA quality (ΔCq 1–4). Methylation annotations were concordant across technical replicates, with minimal introduction of false-positive or false-negative methylation conversion events in FFPE. Somatic variant detection sensitivity was >80% for both 5-base with FFPE samples, with minimal reduction observed for 5-base compared to WGS (≤ 3%).

When evaluated in a tumor-informed MRD use case, 5-base successfully captured tumor fingerprint mutations from FFPE that were also detected in matched cfDNA, while additionally identifying cancer-informative methylation signatures. This combined variant and methylation information substantially increased the total informative signal in cfDNA compared to 4-base workflows.

These results highlight the technical feasibility of using Illumina 5-base to obtain multiomic, clinically actionable profiles from FFPE samples. By unifying methylation and variant detection in a single workflow, 5-base enables more comprehensive molecular profiling and supports development of high sensitivity applications spanning oncology research and translational biomarker discovery.

High plex spatial single-cell microRNA profiling of human colon cancer tissue using Spatial Molecular Imaging (SMI)

Technology & Application Development

POSTER 436

Rachel Liu[1], Joe Phan[1], Rustem Khafizov[1], Kimberly Young[1], Ashley Heck[1], Margaret Hoang[1], Dwayne Dunaway[1], Sayani Bhattacharjee[1], Courtney Anderson[1], Prajan Divakar[1], Mirko Corselli[1], Joseph Beechem[1]

1. Bruker Spatial Biology, Seattle WA, USA

Spatially resolved microRNA (miRNA) profiling remains a key frontier in the rapidly advancing field of spatial omics. Until now, no imaging-based platform has achieved omics-scale detection of these small conserved noncoding RNAs, which regulate up to 30% of human genes through translational repression or mRNA degradation. miRNAs exhibit distinct spatial, temporal, and cell type-specific patterns. Dysregulation of these patterns is linked to a wide range of diseases. In cancer, miRNAs can act as oncogenes (e.g., miR-21), tumor suppressors (e.g., miR-200 family), or immune modulators (e.g., miR-155), underscoring their relevance for both mechanistic and translational research.

miRNAs have been largely inaccessible to spatial transcriptomics due to short sequence length, which poses a challenge for conventional in situ hybridization probe design. CosMx® SMI addresses this limitation by using a strategy to attach readout barcodes to miRNA targets in tissue. Detection is achieved through multiple rounds of reporter binding and fluorescence imaging. Although over 38,000 miRNA sequences are cataloged, only a subset is considered biologically or clinically relevant per the latest miRBase database release. The CosMx miRNA Expression Panel targets these most impactful 827 human miRNAs, including key molecules such as miR-21, miR-200 isoforms, and miR-155, among others.

Here, we showcase miRNA detection at the highest plex ever achieved at subcellular resolution with an imaging-based platform in sections of formalin-fixed paraffin embedded (FFPE) colon adenocarcinoma. By integrating over 800 miRNAs and 68 proteins or over 1000 RNAs in a multiomics workflow, we enabled precise cell typing and tumor microenvironment profiling. Spatial analyses revealed distinct miRNA patterns, such as miR-21 enrichment in tumor cells, and uncovered miRNA-mRNA interactions, highlighting regulatory networks across tissue compartments. These results underscore the power of high plex spatial miRNA imaging for translational and clinical research.

A multi-dimensional strategy for uncovering associations between variants of the epithelial sodium channel (ENaC) subunits and cardiovascular and renal phenotypes

Technology & Application Development

POSTER 437

Patrick Knutson[1], Jeffrey Brodsky[2], Srikant Sarangi[1], Andy Cosgrove[1], Jack Wimberley[1], Chris Garay[1], Chris Guerriero[2], Katie Nguyen[2], Aidan Porter[2], Thomas Kleyman[2]

1. Paradigm4, Waltham, MA, USA
2. University of Pittsburgh, Pittsburgh, PA, USA

The epithelial sodium channel (ENaC) plays a critical role in regulating sodium homeostasis and blood pressure. Rare and common variants in ENaC subunit genes (*SCNN1A, SCNN1B, SCNN1D, SCNN1G*) have been implicated in cardiovascular and renal conditions (e.g., hypertension, hypotension), yet their mechanism of action remains incompletely characterized.

In this work, results are presented from extensive analyses of ENaC subunit variants across multiple population-scale datasets using a multi-dimensional strategy. The REVEAL biomedical analytics platform was used for single variant and gene-based analyses via REGENIE and REGENIE-GENE. This study utilized the 470K whole exome sequencing dataset from UK Biobank and clinically relevant phecodes for hypotension, hypertension, hyperkalemia, hypokalemia, and glomerular filtration rate (GFR). This analysis was tested and replicated in the All of Us Research Program.

Other large summary statistics datasets (e.g., FinnGen, OpenTargets, eQTL Catalogue) were cross-referenced to assimilate further evidence on variant-trait associations in REVEAL's *BioLens* application. For further orthogonal validation, the ESM1b protein language model was employed to predict variant functional impact and conduct variant knockout assays in *Xenopus oocytes* to assess functional consequences on ENaC activity. To evaluate potential causal relationships, two-sample Mendelian randomization (GSMR) and pairwise conditional colocalization (PWCoCo) was applied to variants and traits with the strongest evidence of association.

This multi-dimensional approach, integrating population genetics, functional genomics, computational protein structure prediction, and causal inference analysis, enables more robust target identification, validation, and exploration of mechanism of action. By leveraging the REVEAL platform's capabilities for multi-modal data integration and provenance tracking, it demonstrates a generalizable framework for moving from genetic association to mechanistic understanding in complex trait biology.

Targeted enrichment with Watchmaker TAPS+ distinguishes immune cell types by methylation signal in concordance with WGS

Technology & Application Development

POSTER 438

Can Kockan[1,2], Carrie Cibulskis[1,2], Michelle Cipicchio[1,2], Fabiana Duarte[3], Mark Fleharty[1,2], Chris Kachulis[1,2], Sophie Low[1,2], Sam Pollock[1,2], Micah Rickles-Young[1,2], Indranil Singh[3], Ruochi Zhang[3], Jason Buenrostro[3,4], Sean Hofherr[1,2], Niall Lennon[1,2]

1. Broad Clinical Labs, Complex Assay Services, Burlington, MA, USA
2. Broad Institute, Genomics Platform, Burlington, MA, USA
3. Broad Institute, Epigenomics Program, Cambridge, MA, USA
4. Harvard University, Department of Stem Cell and Regenerative Biology, Cambridge, MA, USA

Methylation patterns are key biomarkers involved in many disease processes, but current gold standard methods for methylation profiling have many limitations. Recently, approaches such as TET-Assisted Pyridine Borane Sequencing, commercialized by Watchmaker Genomics as TAPS+, have addressed these limitations with a novel chemistry that converts methylated rather than unmethylated cytosines. TAPS+ preserves DNA integrity, enabling integrated dual-omic analysis with accurate detection of genomic variants, copy number changes, and methylation states within the same assay.

To evaluate the TAPS+ technology for methylome-enabled cell-type deconvolution, we generated whole genome (WGS) and promoter-region target-enriched sequencing for a cohort of PBMCs and sorted immune cells. Application of the TAPS+ technology in this dataset demonstrated several compelling advantages over standard methods, as well as strong concordance between the WGS and targeted results. Since TAPS+ leaves a much larger fraction of the genome unmodified, we were able to use a standard enrichment kit, instead of one designed specifically for methylation converted libraries. This feature is highly desirable for developing cost effective methylation sequencing assays that can be deployed to accurately profile methylation patterns of large biobank collections. With the TAPS+ method and a Twist Biosciences custom panel, enrichment resulted in a >99% mapping rate, with over 80% of reads mapping to the targeted regions. Methylation rates were highly concordant between WGS and enrichment samples, with an observed mean r2 over 0.99. We also found good agreement between TAPS+ WGS methylation rates and matching previously generated whole genome bisulfite sequencing data across CpG densities. Hierarchical clustering using Ward's method on both WGS and targeted panel TAPS+ datasets accurately clustered the 11 replicates into their respective 4 cell types. Finally, we observed germline variant calling accuracy (SNP F1: 0.989, Indel F1: 0.959) on TAPS+ WGS of the HG002 cell-line, demonstrating the dual-omic utility of this technology. These results support continued evaluation of the TAPS+ technology for a variety of clinical and research applications in our platform.

Direct genome-scale mapping of endonuclease activity of the human LINE-1 ORF2p endonuclease

Technology & Application Development

POSTER 439

Hanna Kodama[1,2,3,4], Syndi Koltz[5], Kaylee Mathews[5], John Thompson[5], Barrett Bready[5], Gerwald Jogl[3,4], Martin S. Taylor[1,2,3]

1. Brown University, Department of Pathology and Laboratory Medicine, Providence, RI, USA
2. Brown University, Legorreta Cancer Center, Providence, RI, USA
3. Brown University, Brown Center on the Biology of Aging, Providence, RI, USA
4. Brown University, Department of Molecular Biology, Cell Biology, and Biochemistry, Providence, RI, USA
5. Nabsys 2.0 LLC, Providence, RI, USA

Rationale:
The LINE-1 retrotransposon is a virus-like DNA parasite that has written at least a third of the human genome and is our only remaining active protein-coding transposon. LINE-1 attempts to "copy and paste" itself in the genome through an RNA intermediate via retrotransposition, a process mediated by sequential activities of the encoded ORF2 protein (ORF2p) endonuclease (EN) and reverse transcriptase (RT) domains. This insertional mutagenesis can cause sporadic genetic disease, and around 5% of humans have a new LINE-1-mediated genomic insertion not present in either parent. LINE-1 also contributes to the pathologies of common diseases, including cancer, by causing inflammation and DNA damage. DNA damage is EN-dependent, and the number of γH2AX damage foci far exceeds the number of completed retrotransposition events in model systems, suggesting DNA damage may occur outside of retrotransposition. Canonical LINE-1 insertions occur preferentially at 5′-TTTTT↓AA consensus motifs, but it is unknown whether this preference comes from site-specific EN DNA nicking or by sequence requirements in the subsequent RT priming step, in which the cut gDNA flap must pair with the poly(A) RNA template. What sequences EN cuts in the human genome or in isolated DNA in general is not known because only completed events can be recovered from cells, due to technical limitations in purifying ORF2p, and challenges mapping nicks in a complex DNA mixture.

Approach:
To address these knowledge gaps, we purified ORF2p EN to homogeneity from *E. coli*, as well as a catalytically inactive mutant as a control and assayed their cleavage patterns on Lambda, *E. coli*, and human DNAs using electronic genome mapping, a technique that directly measures nick positions on long DNAs.

Results:
Analytical methods were modified to minimize false positives and enable the detection of low-level nicking. The frequency of nicks increased in an EN activity- and concentration-dependent manner. Mapping of low-frequency nicks to specific sites is underway.

Conclusions:
These results reveal a powerful new method to directly measure genome-scale endonuclease nicking. Future applications of this technology may also be used to directly assay other DNA modifications and damage at scale.

A million cells is just the beginning: new GEM-X Flex for massive-scale single cell profiling

Technology & Application Development

POSTER 440

Xue Wu[1], Carolyn Morrison[1], Yinzhou Zhu[1], Tingsheng Yu Drennon[1], Jawad Abousoud[1], Payam Shahi[1], Shaun Jackman[1], Chris Macklin[1], Ann Renschler[1], Sini Reponen[1], Paul Lund[1], Peter Smibert[1], Andrew Kohlway[1]

1. 10x Genomics, Pleasanton, CA, USA

Single cell profiling at massive scale has become critical for advancing biomedical research, offering unprecedented insights into cellular heterogeneity, tissue architecture, and disease mechanisms. Landmark initiatives such as the Human Cell Atlas and the Billion Cells Project are driving an urgent demand for higher-throughput single cell technologies. Simultaneously, there is increasing need for platforms that can accommodate a wide range of clinically relevant sample types including blood, fresh or frozen tissues, and FFPE blocks, while enabling efficient sample collection, transport, storage, and batch processing. GEM-X Flex is currently the only platform that addresses all these demands, and enables profiling of up to 320,000 cells per GEM-X microfluidic chip lane with the capability of 16-sample multiplexing. Building on this foundation, we are excited to introduce GEM-X Flex v2, a major upgrade featuring dramatically expanded multiplexing capabilities that enables simultaneous processing of up to 384 samples and profiling of up to 100 million cells per kit.

To demonstrate the expanded multiplexing capacity of the GEM-X Flex platform, we first processed 96 FFPE sections derived from a variety of human healthy and cancer tissues within a single 96-well plate using a novel xylene-free deparaffinization protocol. The resulting nuclei were processed using the GEM-X Flex v2 workflow with 96 unique multiplexing barcodes, and loaded onto a single GEM lane of a chip for a final recovery of 1 million nuclei.

In a separate experiment, we profiled perturbation responses in four colorectal cancer cell lines treated with a panel of 24 small-molecule compounds intended to enhance Wnt signaling. Perturbations were performed across two time points using a 384-plex configuration. Importantly, adherent cell lines were maintained attached to cell culture plates throughout the probe hybridization process, before being detached, pooled, washed and then loaded into a single GEM lane of a chip.

Our study highlights the power of the GEM-X Flex v2 workflow with two high-throughput applications for both clinical sample profiling of FFPE samples and anti-cancer drug mechanism studies on adherent cells. With improved sensitivity, higher sample throughput, and reduced cost, GEM-X Flex v2 is a powerful, versatile platform for large-scale single cell transcriptomic studies.

A novel solution for enhanced Whole Genome Sequencing (eWGS) that provides a simple, tunable workflow across a range of target panel sizes and desired coverage depths

Technology & Application Development

POSTER 441

Katelyn Larkin[1], Nick Downey[1], Ashley Dvorak[1]

1. Integrated DNA Technologies, NGS Applications Development, Coralville, IA, USA

As sequencing technology advances researchers find themselves torn between juggling relatively low pass WGS and deeper targeted coverage. WGS approaches give the most complete data, while targeted enrichment allows deep coverage to provide insight into rare events, but it is limited by the target space. As a result of this, there is a growth in experiments that look to mix and match the two approaches and get the benefits of both. This is often referred to as Enhanced Whole Genome Sequencing (eWGS). Although several approaches are used, they all require splitting samples or library and often some sort of repooling before sequencing. This may result in fragmented data that needs to be sewn together or risk of mixing up samples.

We introduce a novel, tunable, single-library approach to simplify eWGS experiments. After adapter ligation, libraries are amplified with a conditioning reagent and then taken through a regular capture protocol with a standard hybrid capture panel. By adjusting the eWGS conditioning reagent we show we can change the WGS vs hybrid capture target coverage while still using the same streamlined workflow. This approach can utilize different adapter strategies, including full length adapters as well as stubby adapter with indexing primers. The resulting libraries represent molecules that have both whole genome coverage but also enriched regions of interest. The eWGS conditioning reagent is used to determine the ratio of WGS to capture coverage, and the sequencing of the libraries can be scaled to achieve the desired depth.

We have tested our system across a range of panel sizes and relative coverage depths. We show that we can predict coverage for panel sizes of approx. 840 kb, 2 Mb and 39 Mb over a range of 10X-5000X enriched coverage (normalized to 1X WGS coverage). We have also tested multiplexing up to 12 libraries that were pooled after conditioning and prior to the capture protocol and see no negative impact to coverage when compared to a single-plex capture. The ability for high multiplexing and tunability of coverage across panel sizes suggests a robust system that can be expanded for multiple applications in the future.

Technology & Application Development

POSTER 442

SWAN Genomics: demonstration of a novel plasmonic nanoantenna-based DNA sequencing platform

Lawrence Lee[1,2,3], Morassa Mohseni[1], Min Huang[1], James Brown[1,2,3], Remy Ripardelli[1], Nirukshan Shanmugam[1], Simao Coelho[1], Calum Fusco[1], Chu Wai Liew[1], Brian Ee[1], Daryan Kempe[1], Jonathan Berengut[1,2], Slobodan Jergic[1], Paulina Rudolffi-Soto[1], Winton Wu[1], Nava Whiteford[1], Michael Finney[1], William Hyun[1,4], Brock Siegel[1]

1. SWAN Genomics, Sydney, New South Wales (NSW), Australia
2. University of New South Wales, School of Biomedical Sciences, Sydney, NSW, Australia
3. University of New South Wales, ARC Centre of Excellence in Synthetic Biology, Sydney, NSW, Australia
4. University of California, San Francisco, Department of Laboratory Medicine, San Francisco, CA USA

We present proof-of-concept data demonstrating the function of a new class of DNA sequencing technology built on a plasmonic nanoantenna platform. The plasmonic nanoantenna achieves over 1000-fold enhancement of signal to noise ratio emitted from individual nucleotides during polymerase-driven incorporation events, enabling single-molecule sequencing in high background fluorescence. Leveraging a patented method for the large-scale chemical synthesis of billions of plasmonic nanoantenna structures with an extremely low cost of materials (<$1).

These data establish the foundation for our next stage of development where we aim to demonstrate the potential for accurate (Q30), single molecule, single-pass, real time sequencing for the first time. Without the need for consensus sequencing or complicated data processing, our platform can be agnostic to read length, drastically reduce workflow complexity and process both short and long fragments (>10kb) simultaneously in a single run, while lowering cost per genome and without compromise to throughput, which may exceed commercial systems.

SWAN's technology has the potential to revolutionize the future of genomics. Enabling comprehensive solutions, such as long-read whole genome sequencing, clinical diagnostics, and personalized medicine, at a fraction of the current cost. This affordable approach accelerates the advancement of personalized medicine and genomics, making them more accessible. It enables population-scale studies that can identify strong disease correlations, support early interventions, improve clinical outcomes, and significantly reduce global healthcare costs.

Transforming adventitious virus detection in manufacturing: optimized NGS workflows for enhanced sensitivity and speed

Technology & Application Development

POSTER 443

Ka-Wai Leong[1], Reiko Nakashima[2], Qinyu Sun*[2], James Ashall[3], Siti Ayuni Mohamed Yussof[4], Brandye Michaels[1], Nasrin Salehi[1]

1. Analytical Research and Development, Microbiology Strategy and Testing Department, Biotherapeutics Pharm. Sci., Pfizer Inc., Andover, MA, USA
2. Pfizer Inc., Medicine Design, Machine Learning & Computational Sciences, Cambridge, MA, USA
3. Pfizer, Inc., Internal Medicine, Research & Development, London, United Kingdom
4. Pfizer Inc., Biotherapeutics Pharm. Sci., VRD Early Bioprocess Development, New York, NY, USA

*Former Pfizer colleague

Adventitious virus detection is a regulatory requirement for clinical CHO manufacturing, ensuring product safety and compliance with global standards. Traditional methods for virus testing are slow, limited in scope, and often rely on animal testing. In contrast, next-generation sequencing (NGS) offers faster, more sensitive detection without animal use, streamlining batch release and improving safety assurance. However, assay development for adventitious virus detection via NGS faces challenges, including viral genomic diversity, the overwhelming presence of host-derived sequences and requires a strong data analysis pipeline. Optimizing wet-lab and bioinformatics workflows is essential for reliable sensitivity.

This presentation outlines an efficient workflow for virus detection, evaluating sample treatments, library preparation methods, and bioinformatics tools to ensure a streamlined and sensitive approach.

Integrating long-read structural variant analysis with snRNA-seq to elucidate effects on gene expression in disease

Technology & Application Development

POSTER 444

Joshua Z. Levin[1,2], Kwanho Kim[1,2], Zechuan Lin[1,3,4,5,6], Sean K. Simmons[1,2], Jacob Parker[1,3,4,5,6], Matthew Kearney[1,3], Zhixiang Liao[1,8], Idil Tuncali[1,8], Daniel N. El Kodsi[1,8], Monika Sharma[1,3,4,5], Beatrice Weykopf[1,3,4,5], Xianjun Dong[1,3,4,5,7], Victoria Popic[9], Clemens R. Scherzer[1,3,4,5,6]

1. Aligning Science Across Parkinson's (ASAP) Collaborative Research Network, Chevy Chase, MD, USA
2. Broad Institute of MIT and Harvard, Stanley Center for Psychiatric Research, Cambridge, MA, USA
3. Yale University School of Medicine, Stephen & Denise Adams Center for Parkinson's Disease Research, New Haven, CT, USA
4. Yale University School of Medicine, APDA Center for Parkinson Precision Medicine, New Haven, CT, USA
5. Yale University School of Medicine, Department of Neurology, New Haven, CT, USA
6. Yale University School of Medicine, Department of Genetics, New Haven, CT, USA
7. Yale University School of Medicine, Department of Biomedical Informatics and Data Science, New Haven, CT, USA
8. Harvard Medical School, Precision Neurology Program of Brigham & Women's Hospital, Boston, MA, USA
9. Broad Institute of MIT and Harvard, Broad Clinical Labs, Cambridge, MA, USA

Structural variants (SVs) are a diverse class of genomic alterations and are critical contributors to genetic diversity and disease, including Parkinson's disease (PD). Their characterization remains challenging with conventional short-read sequencing, which struggles to accurately span larger SVs and resolve variants in highly repetitive genomic regions, often missing a large fraction of an individual's SVs. Compounding this, the unique algorithmic designs of various variant callers mean no single tool is perfect, leading to distinct advantages and an incomplete picture if only one caller is used. Our study addresses these challenges by employing high-fidelity (HiFi) long-read whole genome sequencing from the PacBio Revio platform. To maximize the sensitivity of SV calling, a multi-stage ensemble strategy was established utilizing three distinct algorithms—Cue2, PBSV, and Sniffles2. SVs identified by each caller were harmonized, filtered by quality and read support, and comprehensively annotated with functional information, such as population allele frequency and overlapping regulatory elements, by querying public databases. Of 40,117 SVs we identified, 61.6% had similar variants in the Human Pangenome Reference Consortium database.

We applied this method to our Parkinson Cell Atlas in 5D cohort of 100 individuals. Focusing on regions surrounding recently identified, causally associated "eGenes" of PD, we identified 321 high-confidence SVs within a 100 kb window of these eGenes. Functional annotation as described above will be presented. Integrating these findings with single-nucleus RNA sequencing (snRNA-seq) data from the same individuals, cell type-specific expression quantitative trait loci (SV-eQTL) and allele-specific expression (ASE) analyses identified 86 and 31 SVs, respectively, with significant, often highly cell type-specific, regulatory effects on gene expression relevant to PD, notably at the chromosome 17q21.31 locus. These results highlight SVs as a potent and under-recognized source of *cis*-regulatory variation in the human brain, influencing gene expression in a cell-type-specific manner, and can help prioritize SV-gene associations for future therapeutic investigation in PD and other diseases.

Optimization of Watchmaker RNA-seq for clinical adoption

Technology & Application Development

POSTER 445

Jocelyn M. Lucyshyn[1], Anthony R. Miller[1], Saranga Wijeratne[1], Jenell Lewis[1], Emilio Soriano Chavez[1], Benjamin J. Kelly[1], Richard K Wilson[1], Elaine R. Mardis[1]

1. Nationwide Children's Hospital, Institute for Genomic Medicine, Columbus, OH, USA

RNA sequencing (RNA-seq) enables transcriptomic profiling in pediatric disease by detecting gene fusions and expression alterations that provide complementary insights to DNA analyses. The Institute for Genomic Medicine's clinical laboratory currently utilizes the Archer FusionPlex assays for targeted fusion detection, but these assays are limited to a subset of known disease-causing fusions in particular cancers. To expand diagnostic capabilities, we chose to compare bulk and cDNA capture RNA-seq methods to provide a broader view of the human transcriptome.

We compared bulk RNA-seq and cDNA capture workflows prepared with the Watchmaker RNA Library Prep Kit, coupling the latter with the IDT xGen Exome v2 panel for cDNA capture enrichment. Commercial reference RNA containing known fusions from both high-quality and FFPE sources was used to evaluate performance. Multiple parameters, including hybridization time, hybridization "plexity", rRNA depletion, and capture kit version, were systematically varied to evaluate their impact on data quality and fusion detection. Performance was assessed using key RNA-seq metrics such as duplication rate, library complexity, and fusion sensitivity to identify an optimized, clinically applicable workflow.

The cDNA capture RNA-seq protocol with 2-hour hybridization using the IDT Hyb and Wash v3 kit demonstrated several advantages over bulk RNA-seq, including reduced RNA input requirements, improved sequencing metrics and enhanced sensitivity for fusion detection. For example, an input of 250ng FFPE into bulk RNA-seq generated 51 million protein coding reads and 14 of 16 expected fusions, while 100ng of the same FFPE sample for cDNA capture generated 82 million protein coding reads and 16 of 16 expected fusions at a lower total sequencing read depth. These findings support the robustness and clinical utility of the cDNA capture workflow.

Adoption of a cDNA capture RNA-seq protocol enables broader fusion detection in a clinical setting. Clinical adoption of this protocol will complement existing targeted assays and improve clinical decision-making.

TAPS+ paired with methyl-fragment fractionation and hybrid capture enables high-sensitivity detection of cell-type–specific DNA fragments in cfDNA

Technology & Application Development

POSTER 446

Craig Marshall[1], Travis Sanders[1], Sam Ortiz[1], Martin Ranik[1], Kristina Giorda[1], Edu Casas[1], Thomas Harrison[1], Doug Wendel[1], Brian Kudlow[1]

1. Watchmaker Genomics, Boulder, CO, USA

CpG methylation serves as a key regulator of cell identity by shaping gene expression programs that define cell form, fate, and function. Additionally, understanding the dysregulation of this epigenetic modification offers a powerful readout for detecting and prognosticating the course of tumor progression. Combining methylation detection with enrichment technologies such as hybrid capture can increase the sensitivity for detecting cell-type specific DNA fragments in clinically relevant samples like cfDNA. However, existing 5mC detection methods rely on a negative readout - converting unmethylated cytosines to thymine - which requires deeper sequencing and custom hybrid capture panels. This conversion affects roughly 20% of nucleotides, reducing the sequence complexity of the human genome to a three-base code and diminishing the richness of downstream analysis.

Watchmaker Genomics' TAPS+ overcomes these limitations by directly converting 5mC to T. This positive readout chemistry preserves native DNA complexity, allowing the use of standard sequencing adapters, standard hybrid capture panels, and fractionation of methylated DNA library fragments using methyl binding proteins. Additionally, because of the direct conversion, only 1-2% of cytosines in the human genome are converted, resulting in the retention of other valuable genomic information such as SNVs and CNVs.

Here, we demonstrate the unique application of TAPS+ in combination with an engineered methyl-binding protein and hybrid capture to detect DNA fragments associated with a specific cell type in a background of cfDNA. By fractionating the DNA prior to TAPS+ treatment, cell-type specific and stably hypomethylated loci are separated from ubiquitously hypermethylated regions. Importantly, both fractions contained valuable genomic and methylation information. Following TAPS+ treatment and amplification, hybrid capture further enriched cell-type specific and clinically informative fragments. TAPS+ then provides per-fragment methylation haplotypes, enabling methylation haplotype load (MHL)-based cell-of-origin calls that disambiguate signals where coverage enrichment alone is insufficient. Additionally, genomic variant information was preserved in the captured loci. Ultimately, TAPS+ paired with methyl-fragment fractionation and hybrid capture provides a sensitive and multimodal tool for detecting cell-type specific DNA methylation while retaining genomic information in clinically relevant cfDNA.

Cross-platform evaluation of long-molecule genomic technologies in pediatric patients

Technology & Application Development

POSTER 447

Anthony R. Miller[1], Meng Wang[1], Iftekhar A. Showpnil[1], Maria E. Hernandez Gonzalez[1], Lakshmi Prakruthi Rao Venkata[1], Eileen Stonerock[1], Allison Daley[1], Catherine E. Cottrell[1,2,3], Benjamin J. Kelly[1], Katherine E. Miller[1,2], Daniel Koboldt[1], Elaine R. Mardis[1,2,4], Richard K. Wilson[1,2]

1. Nationwide Children's Hospital, Steve and Cindy Rasmussen Institute for Genomic Medicine, Abigail Wexner Research Institute at Nationwide Children's Hospital, Columbus, OH, USA
2. The Ohio State University College of Medicine, Department of Pediatrics, Columbus, OH, USA
3. The Ohio State University College of Medicine, Department of Pathology, Columbus, OH, USA
4. The Ohio State University College of Medicine, Department of Neurosurgery, Columbus, OH, USA

Accurate reconstruction of long-range genomic information remains a challenge in pediatric research where structural variants (SV), repeat expansions, and haplotype phasing often underlie disease but are difficult to resolve with conventional short-read sequencing. To address these limitations, we evaluated long-molecule workflows, including short-read based informatic reconstruction (Illumina Constellation mapped read technology), long-read sequencing (PacBio HiFi), and optical genome mapping (Bionano OGM), across eight samples (six patient-derived and two controls). All samples underwent Constellation, however, HiFi and OGM were applied to select samples, based on material availability, enabling cross-platform comparisons of molecule lengths and SV resolution.

For Constellation technology, DNA extraction method influenced long-template reconstruction. Ultra-high molecular weight (uHMW) samples reconstructed substantially more long templates (75% >10kb; 62% >20kb; up to 33% >60kb) than column-based preps (<32% >20kb). Phasing performance followed a similar trend, with uHMW samples achieving ~99% phased SNVs and >94% gene-level phase block coverage, compared to gene-level coverage as low as 56% in column-based preps.

PacBio HiFi enables direct sequencing of native long molecules, offering an empirical view of molecule length distributions. However, the need for multiple polymerase passes to generate consensus reads can limit recovery of ultra-long molecules. In one uHMW sample, Constellation technology reconstructed more templates >20kb than PacBio (66.9% vs. 36%) and uniquely recovered templates >60kb (19.8%), reflecting the influence of input quality and informatic reconstruction. Beyond sequencing, Bionano OGM provides a complementary genome-wide view of large SVs but lacks single-base resolution, limiting precise breakpoint detection. In one uHMW sample, a balanced 3;12 translocation and an unbalanced 17;22 translocation identified by OGM were confirmed in Constellation colocalization plots, with defined breakpoint coordinates supported by VCF outputs.

Our analysis provides a real-world view of pediatric samples processed with Constellation technology, capturing performance metrics across diverse sample types and extraction methods. These findings reinforce that uHMW DNA delivers further improved long-template reconstruction and phasing over standard DNA extraction. Cross-platform evaluation highlights distinct yet complementary strengths of Constellation, PacBio HiFi, and Bionano OGM, each offering trade-offs in resolution, molecule length, input requirements, and variant detection, underscoring the value of multi-platform strategies for comprehensive genomic characterization in pediatric research.

5hmC detection in mouse tissue using long- and short-read sequencing methods

Technology & Application Development

POSTER 448

Luke Morina[1], Jessica Hosea[1], Paul Hook[1], Winston Timp[1,2]

1. Johns Hopkins University, Department of Biomedical Engineering, Baltimore, MD, USA
2. Johns Hopkins University, Department of Molecular Biology and Genetics, Baltimore, MD, USA

The epigenome plays a key role in gene regulation in mammals and often becomes dysregulated in disease. However, it is still not fully understood how epigenetic patterns are regulated - especially how methylation is lost at specific loci. Most studies of methylation in mammalian samples have focused on 5-methylcytosine (5mC) - 5-hydroxymethylcytosine (5hmC) is less characterized due to its low prevalence in samples and difficulty in measurement. But 5hmC is a key intermediate in demethylation because it is not recognized by DNA methylation maintenance enzymes (e.g., DNMT1).

Despite the development of methods for profiling 5hmC, control samples with appreciable levels remain scarce. We reasoned that biological tissues from well-characterized inbred mouse models offer the closest approximation to a “Genome in a Bottle,” with realistic levels of 5hmC, which is rarely observed in actively dividing cultured cells. We used short- and long-read sequencing methods to generate a high-coverage (~100X) dataset across four mouse tissues: cortex, cerebellum, liver, and heart.

Our analysis revealed strong concordance between short-read enzymatic methylation data and long-read basecalling results. We defined three classes of differentially methylated regions between tissues: (1) differentially modified regions (DMRs), reflecting total cytosine modification and serving as a proxy for conventional bisulfite-based methylation; (2) differentially methylated regions (5mC-DMRs), specific to 5-methylcytosine; and (3) differentially hydroxymethylated regions (DhMRs), specific to 5-hydroxymethylcytosine. Genome-wide comparison of these sets revealed approximately 75% of DhMRs were undetectable when considering total cytosine modification alone. For example, we observed differential gene-body 5hmC in brain at *Adam22*, a synaptic receptor critical for neuronal communication, and at *Apob* in liver, a central mediator of lipoprotein assembly and transport. This suggests a connection between active regulation/transcription and hydroxymethylation. The long reads also enable alignment in repetitive genetic elements and phasing of reads to detect allele-specific methylation and hydroxymethylation across tissues.

Our study underscores the importance of 5hmC in different tissues for gene regulation and epigenetic control, while also providing a benchmark resource for future method development and standardization in the field of hydroxymethylation. Our work significantly advances our understanding of 5hmC in mammalian tissues and will facilitate future epigenetic research.

‘Constellation genomes’: a pilot study utilizing Illumina constellation mapped read technology

Technology & Application Development

POSTER 449

Marcel Nelen[1], Bram de Bruijn[1], Karlijn Kusters[1], Carola Roelands[1], Martin Elferink[1], Hanneke van Deutekom[1], Koen van Gassen[1], Wouter de Laat[1], Bert van der Zwaag[1]

1. University Medical Center Utrecht, Department of Genetics, Pharmacy and Biomedical Genetics, Division Laboratory, Utrecht, the Netherlands

We present a pilot study exploring the utility of constellation mapped read technology—a novel framework for integrating high-resolution, phased genomic assemblies across multiple individuals in a disease context. This approach aims to enhance variant interpretation, structural variant resolution, and haplotype-aware diagnostics. Our first experiment utilized the Genome in a Bottle (GIAB) HG002 reference trio, a complex spinal muscular atrophy (SMA) family and patients with MUC1-associated kidney disease caused by a cytosine insertion in the variable number of tandem repeats (VNTR) region.

We generated phased diploid genomes for each individual. HG002 served as a benchmark for assembly accuracy and variant calling fidelity, enabling calibration of our analytical workflows. The SMA-trio analysis revealed compound heterozygous variants and complex rearrangements in the SMN1/SMN2 locus, demonstrating the power of constellation genomes in resolving paralogous regions and inheritance patterns. For the MUC1 patients, we identified pathogenic frameshift mutations within the VNTR region of exon 2, a notoriously difficult-to-sequence locus, confirming the advantage of complete assemblies over traditional short-read approaches.

Overall, this pilot study shows that *constellation genomes* could play a bigger role in clinical genetic testing. By providing more complete and accurate genetic information, this technology could improve genetic testing and diagnosis for a range of conditions. Our next steps will explore how this method could be used in non-invasive prenatal diagnostics (NIPD) and whole genome sequencing in a clinical setting.

Integrating spatial open chromatin and DNA methylation profiling with spatial ATAC-TAPS+

Technology & Application Development

POSTER 450

Colin Ng[1], Katelyn Noronha[1], Martin Ranik[4], Doug Wendel[4], Kristina Giorda[4], Alissa Nolan[3], Jeff Sabina[1], Brian Kudlow[4], Matija Snuderl[2]

1. AtlasXomics Inc., New Haven, CT, USA
2. New York University (NYU) Langone Health and NYU Grossman School of Medicine, Department of Pathology, New York, NY, USA; and Laura and Isaac Perlmutter Cancer Center, NYU Langone Health, New York, NY, USA
3. Ultima Genomics, Newark, CA, USA
4. Watchmaker Genomics, Boulder, CO, USA

Background:
Bulk DNA methylation profiling is a cornerstone of tumor classification and prognostication, particularly in brain tumors such as medulloblastoma. However, conventional approaches average across heterogeneous regions, obscuring the epigenetic diversity that drives tumor progression, therapy resistance, and clinical outcome. Spatial approaches are needed to directly link epigenetic modifications with histological and microenvironmental context.

Innovation:
We developed **Spatial ATAC-TAPS+**, a dual modality assay to simultaneously maps chromatin accessibility and DNA methylation in intact tissue at near single-nucleus resolution. The platform integrates spatial ATAC-seq (AtlasXomics) (>10,000 fragments per spot, FRiP >0.3) with TAPS+ chemistry (Watchmaker Genomics) to enable integrated profiling of open chromatin and cytosine modification states. Unlike bisulfite-based methods, TAPS preserves DNA integrity and identifies both **5-methylcytosine (5mC)** and **5-hydroxymethylcytosine (5hmC)** signals, enabling direct analysis of active, repressed, and poised regulatory states.

Results:
Combined accessibility and methylation signals are co-registered with tumor morphology, while also identifying unique spatial niches. TAPS-derived 5mC and 5hmC profiles demonstrated high concordance with bulk assays from the same tissue blocks, validating the accuracy of in situ detection. In medulloblastoma, integrated analysis revealed >70,000 differentially methylated and accessible loci across subclones, including poised and bivalent regulatory regions whose activity can only be interpreted by combining accessibility and methylation. In gliomas, spatial ATAC-TAPS+ generated the first methylation–accessibility atlases distinguishing IDH-mutant and wildtype tumors, showing >90% concordance with HM450 classifiers while resolving spatially restricted resistant niches.

Impact:
Spatial ATAC-TAPS+ is a first-in-class technology delivering co-registered accessibility and methylation maps within a single tissue section. By resolving 5mC/5hmC states in open chromatin, it links epigenetic modifications to functional gene regulation, providing unprecedented insight into tumor heterogeneity, clonal evolution, and therapy response. Spatial ATAC-TAPS+ establishes a foundation for next-generation spatial epigenomics and biomarker discovery across diverse tissue types.

In situ 3D spatial characterization of genome architecture and whole-transcriptome imaging in single cells with the PaintScape™ system and CosMx® SMI on ER+ breast cancer cells

Technology & Application Development

POSTER 451

Huy Nguyen[1], Yi Cui[2], Andrea Floris[1], David King[1], Serdar Tulu[1], David Castillo[1], Sierra McKinzie[2], Shanshan He[2], Shyamtanu Chattoraj[1], Jude Dunne[1], Joseph Beechem[2]

1. Bruker Spatial Genomics, Research and Development, San Jose, CA, USA
2. Bruker Spatial Biology, Research and Development, Seattle, WA, USA

Disruption of 3D genome organization can profoundly affect gene expression, contributing to disease and cancer development. Mechanisms such as oncogene amplification, deletion or chromatin domain misfolding are known drivers of gene expression changes in cancer. However, conventional methods like whole genome sequencing (WGS), RNA-seq, and Hi-C typically infer these changes at the bulk level, obscuring cell-to-cell heterogeneity.

To unravel the complex mechanisms underlying cancer, approaches with single-cell and spatial resolution are essential. Here, we leverage the PaintScape™ platform to visualize 3D genome architecture across multiple length scales, in situ, within single cells. We then correlate changes in genome structure to gene expression and phenotype using the CosMx® Spatial Molecular Imager (SMI)'s multiomic capabilities to generate a comprehensive multi-modal dataset.

Using the PaintScape system and CosMx SMI, we characterize differences in genome structure, gene expression and phenotype at subcellular resolution between normal (MCF10a) and cancerous (MCF7) breast cell lines. The PaintScape system enables high-resolution visualization of gene region environments involved in cancer-related pathways, including cell cycle regulation, apoptosis, transcriptional control, and chromatin remodeling. We observed cell type variations in chromosome territories, large scale structural chromosome aberrations such as translocations and ecDNA, locus copy number, and topologically associating domain (TAD) boundaries. CosMx multiomic imaging reveals distinct transcriptome-wide signatures across cell types/states, and subcellular localization of key markers. In addition, we expose both cell lines to Estradiol, an Estrogen receptor (ER) agonist, with subsequent upregulation of transcription of MYC, CCND1, and E2F1 in ER+ MCF7 cells. We contrast this with ER- MCF10a cells and use both platforms to compare changes in genome structure with associated changes in transcription programs upon ER activation. After exposure to Estradiol, PaintScape analysis revealed spatial clustering of Distant Estrogen Response Elements (DEREs) around Chr17q and Chr20q with CosMx SMI quantifying local gene transcription activities, highlighting the complementary strengths of the platforms.

The single-cell capabilities of the PaintScape system and CosMx SMI reveal populations of cells with distinct genome organization, transcriptional profile and proteomic features. This integrated multi-modal approach interrogating DNA, RNA and protein in spatial context, offers a powerful framework for understanding how cellular heterogeneity shapes cancer progression.

Enabling community-driven genetics of the domestic cat at large scale

Technology & Application Development

POSTER 452

Chad Nusbaum[1,2], Kasia Bryc[1,3], Benoit Molinie[3], Jimmy Harold[3], Tricia Rogers[3], Rachel Johnston[3,4], Danica Wolfe[3,4], Maxim Elmaleh[4], Aydin Golabi[3], Charlie Lieu[1], Elinor Karlsson[1,3,5]

1. Darwin's Ark, Edmonds WA, USA
2. Gentelechy LLC, Newton MA, USA
3. Broad Institute of MIT and Harvard, Cambridge, MA, USA
4. Zoo New England, Boston MA, USA
5. University of Massachusetts (UMass) Chan Medical School, Worcester MA, USA

Cats and dogs have been our companions for millennia. While extensive genetic research has focused on dogs, cats remain comparatively understudied. To address this gap, Darwin's Ark, a community science nonprofit catalyzing groundbreaking research by engaging pets and their people in scientific discovery, launched Darwin's Cats in 2025, with the long-term goal of generating and analyzing data from 100,000 cats. Ultimately, by empowering pet owners to actively participate in scientific research, Darwin's Cats aspires to expand our understanding of feline and human health.

Achieving these goals requires very large scale collection of DNA samples coupled with detailed information on physical traits, behavior and health. Online surveys to collect phenotypic data are remarkably effective, with over 18,000 cats enrolled and over 1 million survey responses. However, to collect DNA samples, we needed to develop a non-invasive method appropriate for cats, who are notably less cooperative than dogs when using traditional DNA collection methods such as saliva swabs or blood.

To address this challenge, we have developed a novel workflow, Fur-seq, which delivers DNA sequencing libraries sufficient for whole-genome sequencing from small samples of cat fur with an exceptionally low fail rate. In the first 200 fur samples sent by owners, 199 yielded sufficient DNA. Comparison with blood-derived DNA shows it delivers good representation and is effective for genome-wide variant calling. With nearly 1,000 fur samples collected to date, we have implemented an automated, scalable version of Fur-seq that minimizes sample handling through robotic processing.

Given Fur-seq's robust performance in cats, we are exploring its broader application for genetic analysis of archival samples and in other species. This approach could enable genomic studies using fur from deceased pets, often saved by owners, and could be particularly valuable in biodiversity conservation. Blood or saliva collection is often impractical in wild populations and fecal samples yield little host DNA. We have confirmed Fur-seq's success on samples stored for up to 6 years and are now testing it across diverse mammalian species in collaboration with our partner Zoo New England.

Iterative spatial multi-omic analysis combining short- and long-read sequencing for ultra-deep characterization of tissue using the novel EXPO molecular archive

Technology & Application Development

POSTER 453

Linda D. Orzolek[1], Sara Agee Le[2], Jonas Hansen[2], David Symer[2], Chris Barbazette[2]

1. OMAPiX, Inc., Frederick, MD, USA
2. Exponential Molecular, Mountain View, CA, USA

Very recently, long-read sequencing (LRS) is gaining use to elucidate isoforms and other long-read information from fresh frozen tissue; however, short-read sequencing of spatial libraries is still the norm. Although both types of sequencing can provide unique and potentially beneficial information, using combinatorial sequencing approaches is not often performed due to limited input material from a small number of sections.

To investigate the potential for combining both long- and short- read information from spatial transcriptomics analysis, we tested a series of samples using a novel click-chemistry-based method to construct an EXPO molecular archive from spatial cDNA from lung tissue samples. This novel method preserves long-read molecules in a template that can be used for multiple generations of non-destructive downstream sequencing and molecular assays, including short-read sequencing (SRS), long-read sequencing (LRS), hybrid capture, targeted panel, or single gene amplification.

Using the Element AVITI System for SRS and Oxford Nanopore for LRS, we show that an EXPO molecular archive can be prepared from pre-existing cDNA from commercially available spatial transcriptomics methods. Spatial applications which produce full length cDNA can be archived to preserve long molecules for LRS, while also allowing SRS to occur from the same template using standard fragmentation library methodologies.

After comparing the sequencing data outputs of both technologies, we show that the EXPO molecular archive can be assayed repeatedly and reliably. While Nanopore LRS preserves the full-length cDNA for isoform detection, the Element AVITI platform was an expedient and cost-effective way to sequence cell-type classification, as well as targeted genes and panels.

In conclusion, we show that we were able to sequentially interrogate full spatial transcriptomes, cell-types, isoforms, targeted panels, and specific genes of interest in a series of consecutive analyses originating from a single sample using the EXPO molecular archive technology. Further studies are underway to investigate a broader set of tissue types and disease pathologies, as well as how to incorporate this new technology and ultra-deep tissue characterization into our service offering.

ReGenSeq: an ecosystem for high-throughput high-content imaging using decommissioned sequencers

Technology & Application Development

POSTER 454

Kunal Pandit[1,2], Craig Fouts[1,2], Sarah Rodwin[1,2], Sanja Vicković[1,2,3,4]

1. New York Genome Center, New York, NY, USA
2. Columbia University, Fu Foundation School of Engineering & Applied Science, New York, NY, USA
3. Columbia University, Irving Institute of Cancer Dynamics, New York, NY, USA
4. Uppsala University, Royal Swedish Academy of Engineering Sciences, Uppsala, Sweden

The analysis of spatially resolved proteomic data is critical for linking protein localization to cellular function and disease mechanism, yet current iterative multiplexed imaging platforms are often cost-prohibitive and lack scalable analytical tools for complex high content datasets. We developed ReGenSeq, an open-source high-content imaging ecosystem, by repurposing decommissioned Illumina sequencers to make high-content imaging accessible to any research laboratories. For a proof-of-concept, we automated ImmunoSABER, a multiplexed immunofluorescence assay utilizing DNA conjugated antibody to study the progression of amyotrophic lateral sclerosis (ALS) in a SOD1-G93A transgenic mouse model. An antibody panel targeting 29 early ALS biomarkers and relevant spinal cord cell types was developed to characterize nearly 100 spinal cord sections from transgenic and wild type animals across 3 time points. To process all the images we developed a Snakemake pipeline which segments single cells, and measures interpretable features such as antibody intensity, antibody texture, and cellular morphology. Interpretable features can also optionally be augmented with embeddings of cells from a vision transformer. With a combined set of over 10,000 features, we discovered spatial microenvironments using a single-cell embedded latent Dirichlet allocation (sceLDA) model that combines the feature scalability of deep generative models with the interpretability of topic modeling. The sceLDA model efficiently categorizes cells and assigns them into anatomical regions to determine the underlying cell type mixtures associated with disease states. Using sceLDA we discovered pre-symptomatic ALS cellular interaction niches characterized by molecular and morphological features. High-content imaging assays like ImmunoSABER are easily adaptable to the ReGenSeq ecosystem, which offers a complete solution for assay automation, image processing, and spatial microenvironment discovery.

Multi-site benchmarking studies for PRECYSE, a novel DNA single-molecule resolution imaging and analysis QC platform for long- and short-read NGS libraries

Technology & Application Development

POSTER 455

Brian Patterson[1], Katia Sol-Church[2], Kelly Cribari[3], ChewYee Ngan[4], Yongjun Huang[4], Sophie Low[5], Michelle Cipicchio[5], Evelyn Ng[6], Jack Wheeler[6], Guilin Wang[6], Karolina Aberg[7], Jason Reed[1]

1. Evizia, Inc., Richmond, VA, USA
2. University of Virginia School of Medicine, Charlottesville, VA, USA
3. Harvard University, Cambridge, MA, USA
4. The Jackson Laboratory, Farmington, CT, USA
5. The Broad Institute of Harvard and MIT, Cambridge, MA, USA
6. Yale University School of Medicine KMSR, New Haven, CT, USA
7. Virginia Commonwealth University, Richmond, VA, USA

Quality control (QC) remains a fundamental component of complex DNA sequencing workflows, ensuring the integrity of nucleic acid samples prior to downstream analysis and validating intermediate constructs during library construction. Despite its importance, current QC technologies have not kept pace with the rapid evolution of sequencing platforms and applications, presenting a barrier to broader clinical implementation.

In this study, we report findings from an extensive benchmarking evaluation of a novel biomolecular imaging and analysis system, **PRECYSE**, designed for rapid sizing and visualization of DNA fragments ranging from 100 to over 250,000 bp. Conducted across six prominent genomics core facilities in the United States, the study assessed over 200 DNA and RNA samples representing a broad spectrum of sample types and quality. These included short- and long-read sequencing libraries, gDNA derived from blood, saliva, and FFPE tissue, ultra-high molecular weight DNA for Oxford Nanopore sequencing, size-selected cfDNA, enzyme-bound libraries, and sheared genomic DNA.

For short-read libraries, PRECYSE achieved sizing accuracy on par with traditional automated electrophoresis-based instruments, while offering a significantly faster workflow from sample loading to data output in under 10 minutes, compared to 20+ minutes. The platform also enabled detection of sample impurities undetectable using conventional methods. For long DNA fragments (>10 kb), including genomic and sheared DNA and long-read libraries, PRECYSE outperformed electrophoresis QC, delivering improved speed (<10 minutes per sample vs. 30+ minutes) and enhanced accuracy (absolute error <15% versus >15%). Notably, electrophoresis systems were found to consistently overestimate larger fragment sizing, as confirmed by downstream sequencing read lengths, and significantly underrepresent smaller DNA fragment populations.

The platform facilitated direct observation of polymerase and helicase binding to DNA in fully prepared libraries, distinguishing between ligated constructs with and without enzyme treatment. RNA integrity assessments further validated the system's capabilities, showing degradation-related reductions in 18S and 28S rRNA with performance comparable to electrophoresis. Challenging DNA samples such as FFPE-derived DNA were clearly visualized and sized. Across all sample types, PRECYSE revealed previously undetected contaminants and structural variants in nucleic acids that had passed standard electrophoresis QC, highlighting its potential to improve overall sequencing workflow reliability.

Linked Target Capture (LTC) sequencing technology for the detection of fetal cell-free DNA from a heterozygous background

Technology & Application Development

POSTER 456

Joel Pel[1], Aaron N. Bruns[1]

1. Natera, Inc., Austin, TX, USA

The non-invasive analysis of cell-free DNA (cfDNA) from blood has emerged as a powerful tool with broad applications in oncology, organ health, and women's health. Advances in targeted next-generation sequencing (NGS) have enabled the simultaneous detection of multiple cfDNA biomarkers, supporting high-performance assays across diverse indications. Despite this progress, significant challenges remain, such as detecting fetal cfDNA mutations when they exist within a heterozygous maternal background of the same variant. The core difficulty lies in distinguishing the fetal mutational signal from the inherent variability of a nominal 50% maternal allelic contribution.

Linked Target Capture (LTC) is a novel targeted sequencing technology that addresses these challenges. By covalently linking PCR primers to custom capture probes, LTC enriches full cfDNA fragments through amplification of targeted molecules and priming from library adapter sequences. This workflow can be completed within a single day, requires minimal hands-on time, and achieves high sensitivity and specificity due to its linked probe–primer capture strategy. Importantly, LTC preserves the native fragment start/stop positions and consistently delivers high on-target rates and uniform coverage, independent of panel size.

Here, we demonstrate the potential of using LTC to detect fetal cfDNA mutations associated with cystic fibrosis (CF) in a heterozygous parental background. Using titrations of fetal and parental cell line mixtures harboring known CF mutations (both SNVs and indels) to simulate cfDNA, we show that LTC can distinguish fetal mutations from a heterozygous background at low fetal fractions, levels commonly observed in early pregnancy. Furthermore, we demonstrate the ability to detect fetal mutations against a homozygous wild-type background.

A ligation-free capsule-based combinatorial indexing platform for high-resolution single-cell genotype–phenotype mapping in somatic mosaicism

Technology & Application Development

POSTER 457

Joe Pelt[1], Elliot Eton[1], Abigail Thress[1], Dan Landau[1,2*], Ivan Raimondi[1,2*]

1. Weill Cornell Medicine, New York, NY, USA
2. New York Genome Center, New York, NY, USA

Somatic mosaicism is a fundamental driver of aging and disease. Defining the functional consequences of these somatic mutations is essential for understanding how clonal evolution shapes cellular behaviour and disease risk. However, existing single-cell sequencing approaches that link genotype to phenotype remain limited by low throughput and poor genotyping efficiency.

We developed a ligation-free, capsule-based combinatorial indexing platform that enables simultaneous scRNA-seq and genotyping across >100,000 single cells with unprecedented high genotyping efficiency. Our high-throughput approach uniquely enables detection and transcriptional characterization of rare mutant cells with excellent scRNA data quality.

Benchmarking against leading single-cell sequencing platforms, our method demonstrated enhanced performance across key metrics, including genotyping precision, molecular efficiency, and RNA data quality, setting a new standard for integrated single-cell multi-omics.

By combining throughput, precision, and molecular completeness, out technology provides an enhanced framework for mapping genotype to phenotype in somatic mosaicism, enabling systematic discovery of how rare mutations shape cellular function in aging and disease.

Technology & Application Development

POSTER 458

A sample-to-results workflow for agnostic or targeted pathogen sequencing from wastewater

Christine Sumner, Kaylinnette Pinet[1], Patrick Acer[2], Anagha Kadam[1], Lauren Saunders[2], Brittany S. Sexton[1], Matthew Angel[1], Patrick Andersen[2], Ganit Pricer[2], Ben Lepene[2], V.K. Chaithanya Ponnaluri[1], Keerthana Krishnan[1]

1. New England Biolabs, Ipswich, MA, USA
2. Ceres Nanosciences, Manassas, VA, USA

Wastewater sequencing is an efficient and cost-effective approach for the early detection of established and emergent pathogens. Population-level pathogen surveillance can be achieved through agnostic or targeted sequencing of wastewater samples. Agnostic pathogen detection via metagenomic sequencing supports surveillance for unknown and novel pathogens. Targeted sequencing permits the monitoring of whole genomes or key genomic regions to track variants for a pathogen of interest. However, wastewater is a challenging substrate that requires optimized enrichment, extraction, and sequencing protocols. Here we present effective and streamlined sample-to-results workflows for agnostic or targeted pathogen sequencing from wastewater.

We employed Nanotrap® Microbiome Particles to concentrate intact viruses and bacteria from contrived wastewater samples in a semi-automated, high-throughput method. Nucleic acids were then isolated using the magnetic bead-based Monarch® Mag Viral DNA/RNA Extraction Kit. From these total nucleic acid extracts, we ran agnostic or targeted pathogen sequencing.

When applying the agnostic sequencing approach, we detected expected control pathogens and organic population compositions consistent with wastewater substrates. For the targeted pathogen sequencing workflows, we focused on influenza A (Flu A) and respiratory syncytial virus (RSV). The targeted sequencing of viral pathogens like Flu A and RSV from wastewater is technically challenging due to low viral titers, fragmented RNA, and high background from environmental and host nucleic acids. However, sequencing data following these robust workflows showed high genome coverage and accurate strain identification. Thus, these workflows provide scientists and public health labs with streamlined and cost-effective solutions for monitoring pathogens from population-level wastewater sources.

Evaluation, optimization, and implementation of Ultima for sequencing of a large human disease sample cohort

Technology & Application Development

POSTER 459

Michael Quail[1], Sarah Field[1], Carl Anderson[2], Laura Fachal[2], Ruth Eberhardt[2]

1. Wellcome Sanger Institute, Sequencing Operations, Hinxton, Cambridgeshire, United Kingdom
2. Wellcome Sanger Institute, Human Genetics Programme, Hinxton, Cambridgeshire, United Kingdom

We have carried out an extensive evaluation of the Ultima UG100 sequencing system to benchmark its performance for a range of sequencing applications. We demonstrate equivalent performance to currently used alternative short read platforms with SNP and indel precision and recall for analysis of variants in Human whole genome data and concordant data within single cell studies for cell type classification, quantitation of differential gene expression as well as transcript and gene numbers per cell. Having confidence in the data we have embarked on the sequencing of a 10,000 human IBD cohort. We have modified our existing short read automated library prep pipeline to accommodate the requirements of this platform where we make PCR-free libraries from 500ng input in 96 well plates using our suite of automated liquid handlers. To obtain maximal utilisation of sequencing capacity we make an equivolume pool of each batch of 96 libraries and after sequencing an initial wafer we review the coverage obtained for each sample then make a rebalanced pool that is calculated to generate equal coverage per sample when run over a further two wafers. This has generated ~9Tb over 3 wafers for 96 samples with a CV of <5% and mean per sample coverage of 28x.

SNARE (Simultaneous Nucleic Acid Recovery from Embedded tissue): a protocol for rapid extraction of DNA and RNA from formalin-fixed paraffin-embedded sections

Technology & Application Development

POSTER 460

Sitharam Ramaswami[1], Grittney Tam[1], Juan Blanco Heredia[1], Tessara Baldi[1], Caryn Hale[1], Mark Ewalt[1], Brian Loomis[1], Michael F. Berger[1]

1. Memorial Sloan Kettering Cancer Center, Marie-Josée and Henry R. Kravis Center for Molecular Oncology, New York, NY USA

FFPE (formalin-fixed paraffin-embedded) samples are widely used in clinical settings because they preserve tissue morphology and allow long-term storage. However, formalin fixation can degrade nucleic acids, necessitating large amounts of input material, making efficient use of tissue essential for downstream molecular analyses. As multi-omic approaches become standard practice, extracting both RNA and DNA from the same tissue section, or very few sections, would reduce variability caused by tumor heterogeneity and enhance the reliability of integrative analyses. We developed and tested a rapid, spin-free protocol for Simultaneous Nucleic Acid Recovery from Embedded tissue (SNARE) by adapting an existing commercially available DNA extraction kit from Roche-KAPA Express Extract. This protocol uses limited proteolysis to release cytoplasmic ribonucleic acid (RNA) from the tissue, followed by bead capture and lysis of nuclei to release deoxy ribonucleic acid (DNA). We compared DNA and RNA sequence libraries prepared post SNARE extraction to that after using the gold-standard, the AllPrep DNA/RNA kit from Qiagen. Libraries prepared post SNARE showed equivalent performance across all evaluated metrics. We compared library complexity, coverage across GC content bins, and GC bias of libraries constructed from nucleic acids extracted using SNARE and show strong concordance with the Qiagen kit. From a logistical standpoint, SNARE helps streamline laboratory workflows, reduces the number of reagents, processing steps, and time required, which in turn lowers costs. It could be particularly valuable in retrospective studies and clinical trials where archived samples are limited.

Rapid, PCR-free whole-genome sequencing with streamlined library normalization for time-sensitive and stringent applications

Technology & Application Development

POSTER 461

Martin Ranik[1], Bryce Foster[1], Philip Benson[1], Giulia Corbet[1], Travis Sanders[1], Craig Marshall[1], Shelby Lennon[1], Josh Haimes[1], Amy Emerman[1], Lee French[1]

1. Watchmaker Genomics, Boulder, CO, USA

As the cost of sequencing continues to decline, whole genome sequencing (WGS) is increasingly being adopted for comprehensive genomic analysis across a wide range of clinical and research applications. PCR-free WGS workflows are particularly valued in contexts where accuracy and minimal bias are paramount, such as rapid newborn screening, genetic testing for inherited disease, and population-scale genomics studies. In highly time-sensitive applications, minimizing sample-to-sequencer time is essential. Here, we present a rapid, PCR-free WGS workflow capable of converting extracted DNA into sequencing-ready, pooled libraries in under three hours. This method preserves high data quality, incorporates tunable and consistent enzymatic fragmentation, and features a novel, non-destructive library normalization method to overcome a common challenge in PCR-free workflows: accurate library quantification and pooling.

Performance of this rapid PCR-free workflow was evaluated across a range of genomic and FFPE DNA input amounts. Consistent fragment sizes were achieved using a single workflow, maximizing efficiency for PE150 sequencing. The novel normalization technology yielded well-balanced libraries and enabled analysis of the PCR-free libraries by TapeStation for precise library sizing and instrument loading. Furthermore, this approach enriched functional, sequencing-compatible library molecules, reducing index hopping and increasing clusters identified relative to hand-pooled, qPCR-quantified libraries. Coverage uniformity was maintained even in GC-rich and otherwise challenging promoter regions, supporting high precision and recall for both SNVs and indels. Compared to amplification-based workflows, PCR-free libraries demonstrated improved indel fidelity in repetitive regions, resulting in higher F1 scores. With efficient library construction and streamlined pooling, the workflow further reduced input requirements to under 50 ng, expanding PCR-free WGS to input-limited sample types.

The rapid PCR-free WGS workflow presented here combines speed, accuracy, and robust library handling to deliver high-quality, sequencing-ready pooled libraries in under three hours. By addressing common challenges in PCR-free library quantification and pooling, it maximizes sequencing economy and ensures reliable variant detection, making it ideal for time-sensitive clinical applications and large-scale genomic research requiring unbiased, high-fidelity genome coverage.

Effect of pre-analytical variables on cfDNA fragmentomics studies

Technology & Application Development

POSTER 462

Varsha Rao[1], Andrew Mishin[1], Tobin Groth[1], Chris Troll[1]

1. Claret Bioscience, LLC, Scotts Valley, CA, USA

Abnormal genomic DNA fragmentation due to dysfunctional DNA repair and cell-death result in unique circulating tumor DNA profile in the blood. Non-invasive liquid biopsy assays that evaluate pathogenic mutations harbored in circulating tumor-derived cell-free DNA have emerged as novel landscape for identification disease-specific diagnostic and prognostic mutational markers in cancer. However, they require a priori knowledge of the variants of interest and need cost-prohibitively high depth of sequencing. Alternatively, fragmentomic approaches that profile cfDNA fragment length and/or native end motifs do not require high depth of sequencing and may be performed agnostically. However, liquid biopsy approaches are sensitive to a multitude of preanalytical variables such as blood collection tube type, storage and transport conditions, plasma preparation and DNA extraction method. Here, using specialized NGS library preparation approaches that retain native ends and capture DNA as low as 20 bp, we demonstrate the change in fragmentary profile of cfDNA during plasma processing. We developed a specialized control oligo pool that harbors unique overhangs on 5´ and 3´ end that enables accurate measurement of cfDNA native termini loss. We spiked-in this control oligo into various blood collection tubes with/without specialized reagents to prevent DNA degradation, cell-lysis. We assayed cfDNA integrity by preparing plasma from the spiked-in whole blood at 0,4,24 hours to mimic typical duration required to process whole blood in the clinic. We extracted DNA from 3 ml plasma and prepared NGS libraries with 1ng of cfDNA. We observed the formation of multi-nucleosomal cfDNA fragments in tubes that did not prevent cell-lysis (Citrate tubes and Potassium EDTA tubes), especially by 24 hours. Furthermore, we observed complete loss of the spike-in control in tubes that did not have any nuclease inhibitors (Citrate tubes). We further observed that there was higher, progressive loss of the 3´ overhangs than 5´ in the spike-in controls. In comparison, tubes with both nuclease and cell-lysis inhibitors demonstrate maximum DNA stability with minimized gDNA contamination. These observations highlight both the importance of sample preparation methods and the evaluation of sample integrity prior to performing diagnostic tests with liquid biopsies.

Optimizing optical genome mapping for detecting somatic structural variants in human brain tissue

Technology & Application Development

POSTER 463

Lakshmi Prakruthi Rao Venkata[1], Katherine E. Miller[1,2], Maria E. Hernandez Gonzalez[1], James J. Anderson[1], Eileen Stonerock[1], Lauren Mashburn-Warren[1], Daniel C. Koboldt[1,2], Elaine R. Mardis[1,2,3], Richard K. Wilson[1,2], Tracy A. Bedrosian[1,2], Anthony R. Miller[1]

1. Abigail Wexner Research Institute at Nationwide Children's Hospital, Steve and Cindy Rasmussen Institute for Genomic Medicine, Columbus, OH
2. The Ohio State University College of Medicine, Department of Pediatrics, Columbus, OH
3. The Ohio State University College of Medicine, Department of Neurosurgery, Columbus, OH

Optical Genome Mapping (OGM) has recently gained popularity for its ability to detect large-scale structural variants. While OGM as a technology has been widely used and validated for hematological malignancies, its use in non-tumor brain tissues has not been as widely explored, particularly in epilepsy. There has also been limited use of brain tissue, in general, as a source material for OGM.

Malformations of cortical development and other developmental lesions are a major cause of pediatric drug-resistant epilepsies referred for surgical intervention. Somatic variants are a prominent cause of epilepsy-associated cortical malformations. However, most studies focus on single nucleotide variants (SNVs) or small indels that are more easily detectable using short-read sequencing technologies. These short-read technologies often fail to capture large and complex somatic structural variants.

While OGM is a promising tool for the detection of structural variants, it requires high quality, ultra-high molecular weight DNA which can be challenging to obtain from limited frozen clinical specimens. In our study, we successfully optimized a protocol to use OGM to detect somatic structural variants in surgically resected non-tumor brain specimens. We applied this protocol to brain specimens from four patients with focal epilepsy. Three of the four patients had pathogenic or likely pathogenic germline variants in *NPRL3, DEPDC5,* or *TSC1* – all of which are believed to require a second somatic variant to cause disease. All four patients had previous short-read exome sequencing of brain tissue which resulted in no causal findings.

OGM detected a somatic structural variant in one patient which was a 12 kb deletion within *DEPDC5* at 20% variant allele fraction in affected brain tissue; this variant deletes the first several exons of *DEPDC5* and is expected to result in loss-of-function. PacBio and Sanger sequencing further confirmed the deletion and its breakpoints. Taken together, these results verify the somatic second hit in *DEPDC5* identified by OGM and provide a genetic explanation for the patient's focal seizures. Building on this success, future efforts will focus on applying our optimized OGM workflow to additional unsolved epilepsy cases and other neurodevelopmental disorders to expand its diagnostic utility.

High-throughput, automated library preparation with the PacBio HiFi Plex Prep Kit enables the generation of reference-grade, closed chromosomes from microbial isolates at scale

Technology & Application Development

POSTER 464

Nicholas Redshaw[1], Emma Dawson[1], Iraad Bronner[1], Fiona Teltscher[1], Ben Farr[1], Shelly-Ann Coutts[1], Nicole Sevilleno[2], Lesley Shirley[1], Stacey Price-Weight[1], Mark Stares[2], Bonface Gichuki[2], Ellen Higginson[2], Sally Kay[2], Jukka Corander[2], Trevor Lawley[2], Yan Shao[2], Ian Johnston[1]

1. Wellcome Sanger Institute, Sequencing Operations, Hinxton, Cambridge, United Kingdom
2. Wellcome Sanger Institute, Parasites and Microbes Programme, Hinxton, Cambridge, United Kingdom

The recently released PacBio HiFi Plex Prep-96 Kit is designed for cost-effective, automatable, high-throughput processing of up to 96 samples. It has recommended input amounts of 50-300ng gDNA per sample and 20-250ng per amplicon and is ideally suited for microbial, metagenomics, targeted or low-coverage Whole Genome Sequencing (WGS) applications.

Here we present how our HiFi Plex Prep-96 pipeline has been used to generate reference-grade, high accuracy, closed chromosomes for microbial isolates, including human gut microbial and clinical strains. In a head-to-head benchmark with hybrid Oxford Nanopore Technologies (ONT)/Illumina sequencing, HiFi Plex Prep PacBio data gave superior performance across a range of both raw read and assembly quality metrics.

We also describe how the HiFi Plex Prep-96 method has been operationalised as a production pipeline in the Wellcome Sanger Institute Sequencing Operations Laboratories, with the implementation of an automated workflow from gDNA extraction to library preparation, including 96-well plate based methods for gDNA shearing (FastPrep-96 and Hamilton pipette shearing).

The scalable, cost-effective workflow offered by the HiFi Plex Prep-96 kit unlocks new opportunities to deliver high-accuracy long read PacBio sequencing to a range of sample types that were previously limited by cost and labour bottlenecks.

Genomic demultiplexing enables convenient, scalable batching of genetically diverse samples in high-throughput single-cell population studies

Technology & Application Development

POSTER 465

Mike Mehan[1], Sophia Bevans[1], Andrea Manzo[1], Eyal ben David[2], Robert Meltzer[1], Fiona Kaper[1], Cande Rogert[1], Mark Wang[1], Emily Parker[1], Carolyn Conant[1]

1. Illumina, San Diego, CA, USA
2. Illumina, Foster City, CA, USA

Large-scale population genomics studies are rapidly evolving into multiomic investigations, demanding scalable, efficient workflows for transcriptomic profiling. Traditional sample multiplexing strategies such as cell hashing and combinatorial indexing introduce complexity, cost, and potential sample loss. Illumina's Integrated Single Cell Prep (ISCP) workflow, combined with DRAGEN™ Genotype Demultiplexing, offers a streamlined alternative that leverages natural genetic variation and existing genomic databases to enable high-throughput, pooled single-cell RNA-seq (scRNA-seq) analysis.

ISCP simplifies upstream processing with a single-tube, instrument-free workflow that eliminates the need for extensive labeling, washing, or secondary enrichment steps. Downstream, NovaSeq™ X delivers high-throughput sequencing (up to 25B reads per flow cell), supporting million-cell reactions without the need for multiple flow cells. GT-Demux provides rapid, accurate sample assignment with no additional downstream processing, automatically generating individual sample outputs.

The expected performance of genotype demultiplexing was benchmarked using a pooled cell line study. Six distinct human cell lines were combined into a single 100,000 cell reaction using the ISCP T100 kit. This design allowed ground truth for cell identities and doublets to be established using expression differences between the cell lines. Leveraging variant calls from matched bulk RNA-seq, genotype demultiplexing achieved PPVs of greater 99.9% for all cell lines, with a mean no-call rate of less than 3% per cell line. This demonstrates the precision of genotype-based demultiplexing in controlled, mixed-population settings.

Genotype demultiplexing was performance was also evaluated on a test set of Peripheral Blood Mononuclear Cells (PBMC) samples from 10 donors. Performance was similar to the benchmark dataset that used a mixed cell line population, with a mean no-call rate of less than 5% per sample. The false positive rate was less than 1%, which was estimated by including VCFs for PBMC samples not present in the pool and counting the number of cells assigned to samples absent from the pool. These results demonstrate that the integration of million cell kits, sample batching, and genotype-demultiplexing provide an efficient end-to-end solution for high-throughput single-cell transcriptomics, supporting its application in atlas-scale and population-level studies.

For Research Use Only. Not for use in diagnostic procedures.

Spatial imaging omics facilitate multimodal profiling of both cells and tissues

Technology & Application Development

POSTER 466

Kenny Roberts[1], Luke Wylie[1], Andrew Lawson[1], Qianqian Zhang[1], Julie Matte[1], Aleksander Gontarczyk[1], Yixi Chen[1], Tannia Gracia[1], Gosia Trynka[1], Inigo Martincorena[1], Omer Bayraktar[1], Andrew Bassett[1]

1. Wellcome Sanger Institute, Hinxton, Cambridgeshire, United Kingdom

As the scope of spatially-resolved transcriptomics data continues to grow, spatial approaches become increasingly appealing not only for the analysis of complex tissue structures but also single-cell samples including blood/PBMCs and dissociated tissue suspensions. Concurrently, the field is increasingly focusing on multimodal approaches that allow direct integration of transcriptomics, genomics, proteomics, and other omics to maximise our understanding of complex biological environments. By leveraging several technology platforms as well as sequential combinations thereof and bespoke application to varied sample types, we are building an ecosystem of spatial assays designed to marry together multiple modalities in tissues, cultured cells, and cell suspensions.

Using a novel protocol for application of 10x Genomics Xenium to PAXgene-fixed paraffin-embedded (PFPE) tissue, in concert with laser capture microdissection (LCM), we integrate single-cell spatially-resolved 5000-plex RNA imaging with whole exome or targeted nanorate sequencing (NanoSeq) to interrogate the genotype and transcriptome of cancerous or pre-cancerous clones across diverse human tissues and disease conditions including thyroid, synovium, adrenocortical carcinoma, and osteosarcoma. Furthermore, application of the Akoya PhenoCycler-Fusion platform post-Xenium adds 60-plex proteomics thus facilitating assay of DNA, RNA, and protein in the same cells or clones.

In parallel, by either culturing live cells directly on slides or adhering them post-fixation using STAMP or simple gravity settling, we are able to apply Xenium to hundreds of thousands or millions of cells on a single slide, including iPSC-derived microglia, patient T-cells, and glioblastoma stem cells. This facilitates not only the use of spatial transcriptomics as an efficient read-out following culture or genetic perturbations but also integration with other imaging assays such as cell type-specific Cell Paint.

Meanwhile, the Element Biosciences AVITI24 platform allows simultaneous assay of 350 RNA targets and 50 proteins, with a gasket architecture specifically designed for live cell culture under multiple experimental conditions. The newly developed direct in sample sequencing (DISS) assay allows for de novo in situ read-out of unknown sequences including CRISPR guide RNAs (gRNAs).

Collectively, these advances define a flexible framework for multimodal spatial profiling that delivers a more comprehensive view than any single assay and thus accelerates both discovery and pre-translational research.

Solving amplification of high-GC expanded repeats to enable rare disease detection using long-read sequencing

Technology & Application Development

POSTER 467

Isidora Arriagada[1], Javier Gimpel[1], Jennifer Hsieh[1], Julian Sáez[1], Javier Cáceres[1], Catalina Valdivia[1], Isa Marx[1], Manuel Rozas[1]

1. Kura Biotech, Puerto Varas, Los Lagos, Chile

Long-read sequencing technologies can often capture complex genomic regions without prior amplification. However, in applications where sample input is limited, or when targeting specific loci such as disease-associated repeat expansions, amplification is required to generate sufficient material and recover the full structure of the region. This is especially important when repeats are long, GC-rich, or structurally unstable, as incomplete or error-prone amplification can compromise downstream analysis.

To address this challenge, a high-fidelity amplification method was developed to robustly generate long, GC-rich, and repetitive genomic regions without introducing errors or dropouts. This approach enables the enrichment of difficult targets before sequencing while preserving their full sequence integrity.

This method was tested on genomic targets containing repeat expansions in *C9orf72* (Amyotrophic Lateral Sclerosis), *FXN* (Friedreich's ataxia), *FMR1* (Fragile X syndrome), and the *IT15 gene* (Huntington's disease), as well as large non-repetitive regions in *BRCA1* (8.4–36.4 kb) and *SMN1* (28.2 kb). DNA from affected and unaffected individuals (Coriell Institute) was used to evaluate the performance of the amplification method.

The amplification method successfully recovered long and GC-rich repeats, including a 6.5 kb *FMR1* allele with a (CGG)n repeat and 79.26% GC content, and a 5.7 kb *C9orf72* allele with a (GGGGCC)n repeat. It also enabled amplification of large genomic regions up to 36.4 kb (*BRCA1*). Sequencing with Oxford Nanopore showed high coverage and full-length spanning reads across expanded alleles, allowing accurate repeat sizing.

This approach supports reliable recovery of expanded repeats and large genomic fragments in contexts where amplification is necessary before sequencing. It may facilitate improved characterization of clinically relevant regions that are difficult to access using standard protocols.

FLeCS-seq: a method for complete full-length cDNA sequencing and its application to targeted analysis

Technology & Application Development

POSTER 468

Masahide Seki[1,2], Ayako Suzuki[1], Yutaka Suzuki[1,2]

1. The University of Tokyo, Department of Computational Biology and Medical Sciences, Kashiwa, Chiba, Japan
2. The University of Tokyo, Life Science Data Research Center, Kashiwa, Chiba, Japan

Long-read sequencing has enabled reading the full length of RNA as a single read. Existing long-read RNA sequencing methods, such as cDNA-seq utilizing the SMART method, often co-synthesize cDNAs derived from fragmented RNA, making it difficult to selectively obtain genuine full-length cDNAs (fl-cDNAs). Although several methods rely on the 5' cap structure exist to selectively synthesize fl-cDNA, they require micrograms of RNA or involve highly labor-intensive procedures, limiting general applicability. We previously developed an improved oligo-capping method that specifically substitutes the cap structure with an RNA oligo.

In this study, we developed FLeCS-seq (fl-cDNA-selective sequencing) by combining the improved oligo-capping method and poly(A) tail linker ligation for selective fl-cDNA synthesis, followed by Nanopore sequencing. FLeCS-seq successfully read full-length RNA sequences up to 20 kb and was applicable to total RNA input as low as 5 ng. The 5' and 3' ends of the reads mapped with high efficiency around known transcription start and termination sites, respectively, demonstrating highly efficient capture of full-length RNA sequences. Leveraging this high selectivity, the method was applicable even to degraded RNA. Furthermore, we applied FLeCS-seq to target sequencing using hybridization capture. We performed gene-specific, deep fl-cDNA analysis on eight lung cancer cell lines, utilizing a panel targeting 679 cancer-related genes, demonstrating its capability to detect multiple mutations and gene fusions within a single fl-cDNA read. FLeCS-seq allows for relatively easy implementation of fl-cDNA analysis and is expected to contribute to the determination of full-length RNA sequences in various and limited quantity samples.

Quantifying reverse transcription efficiency through single-molecule analysis with Countable PCR for enhanced NGS library preparation optimization

Technology & Application Development

POSTER 469

Eleen Y. Shum[1], Jung Won Keum[1], Christina Fan[1]

1. Countable Labs, Inc., Palo Alto, CA, USA

Optimization of enzymatic reactions in next-generation sequencing (NGS) library preparation is critical for maximizing yield and sensitivity, particularly in low-input, single-cell, and spatial transcriptomics workflows. Small differences in enzyme performance in molecular conversion efficiency can have a large impact on transcript detection sensitivity and overall assay reproducibility. Traditionally, assay optimization relies on direct NGS readouts to quantify these effects, but sequencing-based evaluation is costly, time-consuming, and limits iteration speed during development.

Here, we present an alternative approach using Countable PCR, a quantitative, single-molecule linkage analysis method that enables rapid and precise measurement of enzymatic efficiency without sequencing. Countable PCR can count and characterize millions of individual DNA molecules in parallel, allowing fine-grained quantification of molecular conversion events, in a single PCR reaction.

We demonstrate the use of Countable PCR to evaluate reverse transcriptase (RT) performance across a panel of commercially available enzymes. This approach provides direct measurement of mRNA-to-cDNA conversion efficiency, positional bias along transcripts, and the rate of template switching—key parameters that govern library complexity and bias. We showed that the mRNA-to-cDNA conversion efficiency differed from within ~20% to ~6-fold across reverse transcriptases. By studying the linkages of multiple regions along the transcript in each cDNA molecule, we revealed differences in the processivity and template switching efficiency across different RTs. By contrast, when the same assessment was performed using qPCR, it was inconclusive due to confounding factors associated with qPCR, such as amplification bias across multiple targets.

Beyond reverse transcription, the same framework using Countable PCR can be extended to assess other enzymatic steps central to NGS library preparation, such as ligation. By eliminating the need for sequencing during optimization, the use of Countable PCR significantly reduces experimental cost and turnaround time, eliminating complicated bioinformatics analysis, thereby accelerating NGS assay development cycles.

mcPCR: PCR for methylation

Technology & Application Development

POSTER 470

Geoff Smith[1], Chiara Airoldi[1], Jason Betley[1], Harriet Chidwick[1], Christopher Hodson[1], Marina Lobon[1], David McBride[1], Arthur Norman[1], Helen Northen[1], Gary Skinner[1], Arif Surani[1], Maria Tothova[1], Olivier Lan Chow Wing[1]

1. Syndex Bio Ltd., Foxton, Cambridgeshire, United Kingdom

DNA methylation is an important biomarker for the diagnosis and prognosis of many diseases and for assessing the progression of treatment. However, measuring the methylation status of individual DNA fragments is complex. Most methods are indirect, involving base conversion chemistries which can be destructive to the sample itself. Because PCR cannot copy the underlying DNA methylation signatures, the low level of DNA present in many clinical samples poses a significant technical challenge, especially for non-invasive or small sample biopsies.

Here, we show for the first time the amplification of DNA in which both the underlying 4-base code *and* the 5-methyl cytosine base modifications are faithfully copied from the original sample. We achieve this with *mcPCR* (which stands for methyl-copying PCR), our proprietary DNA amplification technology that synchronizes base extension and methylation using novel enzymes designed by AI-driven protein engineering.

Our *mcPCR* workflow is compatible with many sample types and amplification formats. The amplification products can be sequenced on both short- and long-read sequencers, allowing both the direct or indirect reading of nucleotide variants and methylation marks from the same sample.

In this study, we demonstrate the application of *mcPCR* to the targeted amplification of specific DNA fragments from the human genome. In the future, we see *mcPCR* as a powerful method for the sensitive and accurate detection of DNA methylation with broad application in research and disease prognosis and diagnosis.

Technology & Application Development

POSTER 471

A platform for integrated live-cell imaging and RNA-seq enables functional single-cell profiling

Magali Soumillon[1], Caroline Sartain[1], Amara Alexander[1], Nicolas Chevrier[1,2]

1. Flexomics, Waltham, MA, USA
2. The University of Chicago, Pritzker School of Molecular Engineering, Chicago, IL, USA

Single-cell sequencing has transformed molecular biology but remains largely limited to suspended cells captured in droplets or hydrogels. These methods cannot readily support adherent cells, colonies, or dynamic cell–cell interactions, key to understanding tissue organization, immune synapses, and drug responses.

We present a new single-cell workflow that integrates live-cell imaging and RNA-seq from the same adherent or interacting cells, bridging functional and molecular biology in one scalable system. The platform supports both high-density microscope slide–format consumables (8 samples per run) and a 96-well configuration for perturbation and screening studies, adaptable to a wide range of adherent and co-culture systems.

Cells are cultured, imaged, and assayed directly within picowell arrays before RNA capture on optically encoded barcoded beads, enabling direct linkage between image-based phenotypes and transcriptomic data at single-cell level. The workflow accommodates live-cell assays, cell-painting–style imaging, and co-culture experiments, including colony growth and T cell–Antigen Presenting Cell interaction models. After imaging, cells are stabilized and processed for barcode deciphering and RNA sequencing, enabling seamless integration of functional imaging and transcriptomic profiling.

Across thousands to hundreds of thousands of wells per sample, we observe high occupancy, consistent recovery of single cells and small multicellular structures, and robust RNA quality (>2,500 genes per cell). The system enables precise mapping between morphology, interaction dynamics, and molecular state. In the 96-well format, it supports systematic perturbation designs for high-content single-cell “drug-seq” and response profiling.

Unlike droplet or hydrogel-based methods, this approach captures spatially organized and adherent biology, supports real-time functional assays, and integrates seamlessly with existing microscopes and sequencing workflows. By connecting imaging-rich functional data with transcriptomics in a unified, microscopy-compatible format, this platform expands what can be analyzed at single-cell resolution, linking molecular profiles to real cellular behaviors and interactions.

Integrated B cell receptor and spatial profiling with long-read sequencing maps clonotypes to lung immune niches

Technology & Application Development

POSTER 472

Devjanee Swain Lenz[1,2], Alan Smith[3], Emily Hocke[3], Michael Aksu[3], Karen Abramson[3], Graham Alexander Jr.[1], Vaibhav Jain[3], Robert M. Tighe[4], Jadee Neff[5,6], Simon Gregory[3,7], Mary H. Foster[4]

1. Duke University, School of Medicine, Sequencing and Genomic Technologies, Durham, NC, USA
2. Duke University, Department of Molecular Genetics and Microbiology, Durham, NC, USA
3. Duke University, Molecular Genomics Core, Durham, NC, USA
4. Duke University, Department of Medicine Durham, NC, USA
5. Duke University, Department of Pathology, Durham, NC, USA
6. Duke University, BioRepository & Precision Pathology Center (BRPC), Durham, NC, USA
7. Duke University, Department of Neurosurgery, Durham, NC, USA

The adaptive immune response provides precise, robust, and antigen-specific protection in animals. However, dissection of these responses is difficult due to responses' high diversity, tissue dependency, and dynamic transcriptional regulation. Current methods to resolve spatial lymphocyte clonotypes must compromise between spatial resolution and full-length sequencing of receptor transcripts. To this end, we developed a targeted long-read spatial transcriptomic profiling approach to recover full-length B cell receptor (BCR) sequences and map clonotypes within immune niches. We applied this method in a mouse model exposed to crystalline silica dust (cSi), which induces lung tertiary lymphoid structures and autoimmunity. Briefly, we generated 3'-barcoded 10x Genomics Visium cDNA libraries from lung sections of mice exposed to cSi dust and enriched for full-length Ig transcripts containing BCR sequences using a custom-designed hybrid capture panel from Twist Biosciences. We constructed and sequenced long-read sequencing libraries using the Template Prep Kit 3.0 and SPRQ chemistry for the PacBio Revio. Data were processed and analyzed with PacBio Iso-Seq within the SMRT Link platform, Pigeon tools, and IsoQuant and compared to data from unenriched Visium short-read cDNA libraries using Space Ranger, Seurat and Bansky. Our results reveal diverse heavy and light chain variable regions in lungs with mild-to-modest cSi-induced histopathologic damage, including lymphoid clusters. Among 142 unique productively rearranged heavy chains, IgA, IgG2b, and IgG2c isotypes are prominent (76%), indicating local immune activation with isotype switch. Ig heavy chain CDR3 length is significantly longer than the mouse population reference, possibly indicating an autoimmune response. Several IGHV and D1-1 gene segments are used recurrently. In one lung, a clone defined by use of the same V4-1/D2/J4 gene segments with an identical long 16 a.a. HCDR3 includes both IgA and IgG2b, suggesting contributions of a single clone to both local mucosal and systemic immunity. Based on spatial profiling, Ighm, Ighg2b, Igha/J-chain, and Ighg2c gene expression associated with different clusters, indicating a dynamic immune response and presence of multiple distinct immune niches within cSi-damaged lung. In sum, these results demonstrate the utility of our method to profile tissue immune microenvironments in concert with local B cell repertoires elicited by toxicant exposure.

Optimizing FFPE DNA library preparation for Twist Exome Capture: comparative evaluation of NEBNext UltraShear™, Twist enzymatic, and mechanical fragmentation methods

Technology & Application Development

POSTER 473

Hongli Tang[1], Tabish Hussain[1], Xiaoping Su[2], Rijith Jayarajan[1], Clay Brasuell[1], Erika Thompson[1]

1. The University of Texas MD Anderson Cancer Center, Advanced Technology Genomics Core, Department of Genetics, Houston, TX, USA
2. The University of Texas MD Anderson Cancer Center, Department of Bioinformatics and Computational Biology, Houston, TX, USA

Formalin-fixed, paraffin-embedded (FFPE) DNA remains a challenging input for next-generation sequencing due to degradation and fragmentation artifacts. The chemical modifications and crosslinking effects in FFPE tissues can compromise downstream analyses, making optimization of library preparation strategies critical for ensuring data quality and consistency, particularly in targeted capture workflows such as Twist Exome Capture. We evaluated three library preparation approaches to identify the most robust method for FFPE DNA. Human FFPE DNA samples (n=4) were processed using: (i) NEBNext UltraShear™ FFPE DNA Library Prep Kit, (ii) Twist Enzymatic Fragmentation (EF 2.0) kit, and (iii) Twist Mechanical Fragmentation (MF) kit. All libraries underwent Twist Exome Capture 2.0 target enrichment and sequencing on NovaSeq X. Key performance metrics included normalized depth of coverage, PCR duplicate rate, and on-target rate. NEBNext libraries exhibited the highest PCR duplication rates, leading to markedly reduced normalized coverage compared to Twist EF and MF. Both EF and MF libraries produced approximately 50X higher normalized coverage than NEBNext. While EF libraries displayed variability in coverage among samples, MF libraries achieved consistent coverage across all samples. All three methods showed comparable on-target rates, indicating no major differences in capture specificity. These findings indicate that library preparation strategy substantially influences capture efficiency, with Twist MF demonstrating the most reproducible performance. For FFPE DNA sequencing with Twist Exome Capture 2.0, the Twist MF kit outperforms alternative methods by minimizing PCR duplication and ensuring consistent coverage. Thus, adopting Twist MF based preparation can improve the reliability of FFPE sequencing studies, especially in translational and clinical genomics applications where consistency and accuracy are critical.

Building agility into high-throughput genomics: modular laboratory automation at the Wellcome Sanger Institute

Technology & Application Development

POSTER 474

Scott Thurston[1], Lesley Shirley[1], Ben Farr[1], Shelly-Ann Coutts[1], Tobi Ajenifuja[1], Abdulrahman Tuameh[1], Emma Fraser[1], Rhona Long[1], Ian Johnston[1]

1. The Wellcome Sanger Institute, Hinxton, Cambridge, United Kingdom

The growth and diversification of genomic technologies regularly present new opportunities for advancing Sanger science. To realise these opportunities, Sanger must rapidly develop and deploy new automated laboratory workflows utilising our existing automation fleet.

Historically, Sanger used a bespoke approach to lab automation, designing highly optimised pipelines tailored to specific workflows. Whilst maximising walk-away time, these rigid workflows often did not maximise instrument utilisation and created significant barriers to adapting processes, limiting responsiveness to evolving scientific priorities.

To address this challenge, Sanger has adopted a modular automation strategy, breaking workflows into simple, reusable, process-focused components (e.g. SPRI bead purification, enzymatic reactions, library normalisation). Common programs are used as shared "nodes" to support multiple genomics pipelines. By redeploying and reshaping these nodes we can swiftly develop and implement new workflows, ensuring pipelines keep pace with the fast-evolving genomics landscape at Sanger.

This approach has already demonstrated impact across a diverse range of projects:

- PacBio Sequel IIe SPK3 workflow modularisation enabled rapid deployment of the Kinnex and HiFi Plex workflows. Both pipelines were operational within days, a timescale unimaginable with our previous approach.
- Sanger's viral bait capture workflow (RVI-seq) was modularised when expanding capacity to another lab, enabling the repurposing of under-utilised equipment and minimising the need for capital expenditure.
- A PCR-free WGS workflow for Ultima UG100 was spun up quickly using existing program elements to generate proof-of-principle data for the sequencing of Inflammatory Bowel Disease (IBD) samples. This workflow now supports the processing of a 10,000 sample IBD cohort.

Although it requires a cultural change in method development and scheduling, modularisation of lab automation has already started to prove its value by unlocking capacity, accelerating proof-of-principle data generation, and providing agility to respond to scientific opportunity by enabling rapid adoption of new assays without wholesale redevelopment.

Optimization of high-quality HMW DNA extraction protocol from formalin-fixed paraffin-embedded tissue for long-read sequencing applications

Technology & Application Development

POSTER 475

Pascal Uchehara[1], Nourddine Djeddar[1], Rachel Martini, Ph.D.[1], Melissa Davis, Ph.D.[1]

1. Morehouse School of Medicine, Institute of Translational Genomic Medicine, Atlanta, GA, USA

Formalin-fixed paraffin-embedded (FFPE) remains a widely used and foundational method for long-term preservation of biospecimens, serving as one of the largest global archives of patient-derived tumor material. Decades of FFPE biobanking have yielded valuable archival biobanks from patients, offering diverse clinical and molecular profiles across various cancer tumor types and populations. Despite their vitality, FFPE samples carry significant challenges for genomic analysis due to formaldehyde-induced crosslinking (methylol adducts), DNA fragmentation, and chemical modifications during fixation that compromise nucleic acid integrity. These limitations have restricted most sequencing applications from FFPE-derived DNA to short-read whole genome sequencing (srWGS). While short-read sequencing applications have advanced cancer genomics considerably, they struggle to accurately resolve complex structural variants, repetitive regions, and epigenetic changes that underlie tumor heterogeneity and therapeutic response in the cancer genome.

Recent advancements in long-read whole genome sequencing (lrWGS) technologies now enable comprehensive detection of structural variants, transcript isoforms, and epigenetic modifications, offering unprecedented insight into cancer genome architecture. However, these platforms have been largely inaccessible for FFPE samples due to the difficulty of recovering intact, high-molecular-weight (HMW) DNA.

We have developed an optimized extraction protocol that consistently yields HMW DNA from FFPE tumor samples, with fragment sizes exceeding 10 kilobases (kB) and quality metrics compatible with lrWGS library preparation. We aim to evaluate our workflow across multiple tumor types and archival storage duration to demonstrate consistent recovery and minimal batch-to-batch variability. Preliminary sequencing data indicate successful long-read library preparation from FFPE-derived DNA, with read length and coverage profiles within expected ranges for high-quality performance. Downstream library metrics and sequencing performance

This advancement represents a critical step toward unlocking the full potential of archival FFPE biobanks for high-resolution cancer genome analysis. By enabling lrWGS from these materials, our approach expands opportunities for retrospective genomic discovery, structural variant reannotation, and integrative epigenomic profiling across clinically annotated cohorts. This method provides a scalable framework for extending long-read sequencing technologies to legacy cancer biorepositories worldwide, transforming how preserved specimens can be leveraged for translational and precision oncology research.

Low-temperature transcriptomics: improved data quality in short-read and long-read RNA-seq from an ultraprocessive reverse transcriptase that copies at 30°C

Technology & Application Development

POSTER 476

Jason Underwood[1], Li-Tao Guo[1], George Maio[1], Ryan Muldoon[1]

1. RNAConnect, Branford, CT, USA

Despite advancements, transcriptomics still faces an inherent challenge; the fragile nature of RNA. RNA is prone to spontaneous hydrolysis at elevated temperatures in the presence of divalent cations and alkaline pH. Transcriptomics often measures RNA by proxy through DNA sequencing, where conversion of RNA to cDNA is carried out by reverse transcriptase (RT), usually at slightly alkaline pH and requiring magnesium ions. Retroviral-derived RTs used in these reactions employ higher temperatures (42-60˚C) in an effort to unfold RNA secondary structures and compensate for their lack of processivity and helicase activity.

We hypothesized that RT reaction conditions may contribute to chemical hydrolysis, resulting in a race between cDNA copying versus RNA degradation. As a test, we incubated a 7.6 kb RNA in three different 1X RT reaction buffers at their recommended temperatures and found that 50% of this long RNA degraded in 30 minutes. With reaction times up to 1.5 hours in RNA-seq workflows, RNA degradation may present a barrier to complete and accurate transcriptomes, especially for transcripts above 10 kb.

To address this, we performed three different standardized transcriptomic workflows using a non-retroviral RT, UltraMarathonRT (uMRT), that is highly processive with intrinsic helicase activity and high activity at 30˚C. RNA-seq libraries generated by template switching and PCR (similar to SMART-seq workflow) profiled on a short-read sequencer show improved gene detection, gene body coverage and improved quantification of long RNAs. These libraries also uncover nearly double the intron retention events. Long read cDNA libraries sequenced by PacBio SMRT sequencing also show improved synthesis of long cDNAs resulting in 20% longer isoforms.

To more directly test how RT conditions affect RNA quality, we used uMRT for ONT direct RNA-seq where a cDNA copy is synthesized, but only the RNA strand is sequenced. We find that uMRT provides nearly double the mapped reads of very long RNAs (≥10 kb) and increases the maximum transcript sequence to 30 kb. Together, these results from multiple sequencing modalities indicate that the high processivity and low reaction temperature of uMRT can improve RNA profiling by making long RNA copying more efficient, likely due to reduced random hydrolysis.

Automated HMW DNA extraction and library preparation on Volta Labs' Callisto for routine ONT long-read whole genome sequencing at UMC Utrecht

Technology & Application Development

POSTER 477

Bert van der Zwaag[1], Rama Krishna Simhadri[2], Rumo Jansen[1], Carola Roelands[1], Nayan Agarwal[2], Sam Bae[2], Mirjan Albring[1], Koen van Gassen[1], Martin Elferink[1], Abdul Mohammed[2], Udayan Umapathi[2], Marcel Nelen[1]

1. University Medical Center Utrecht, Department of Genetics, Division Laboratory, Pharmacy and Biomedical Genetics, Utrecht, the Netherlands
2. Volta Labs, Inc., Boston, MA, USA

The genome diagnostics section at UMC Utrecht has recently transitioned from short read whole exome sequencing (srWES) to short read whole genome sequencing (srWGS) as the method of choice for the detection of genetic variants in patients with heritable diseases. The broader perspective of WGS improves the detection of genetic variants linked to these diseases, while reducing the complexity and costs of handling multiple parallel workflows. We are now exploring long read WGS as the next step toward further workflow integration, aiming to retain or improve sensitivity for clinically relevant variants. As part of this exploration, we will sequence up to 1,000 genomes using the Oxford Nanopore Technologies (ONT) PromethION 24 platform and library preparation kit. A considerable portion of these samples will be used for retrospective analysis of DNA's with known clinically relevant variants, enabling us to assess the sensitivity and specificity of ONT long read sequencing.

For a project of this scope, and to meet ISO15189 accreditation standards for diagnostics, high-quality input DNA and a robust, reproducible automated library preparation are essential. We compared DNA isolated using our routine workflow with high molecular weight (HMW) DNA isolated via the Volta Labs Callisto system. These samples were then used to prepare sequencing libraries for the ONT PromethION 24, also using the Callisto system. We found that combining HMW DNA isolation with automated library prep on the Callisto system enabled the consistent generation of up to 24 high-quality sequencing libraries per batch. These libraries produced stable sequencing results with high N50 values (15–20 kb range) over 72-hour runs on PromethION flow cells.

This workflow, featuring high-quality HMW DNA and automated library preparation, will form the foundation of our evaluation of ONT long read sequencing for clinical diagnostics.

Enhanced spatial transcriptomics enables high-resolution profiling of neuronal transcription factors in the developing nervous system

Technology & Application Development

POSTER 478

Angela Vasaturo[1], Renchao Chen[1], Manisha Ray[1], Justin He[1], Bin Wang[1], Bing Yang[1], Timothy Wiggin[1], Lizi Maziashvili[1], Alexander Genshaft[1], Peter Reinhold[1], Sudhir Tattikota[1], Yvette Cajigas[1], Lorenz Rognoni[1], Jiang He1, George Emanuel[1]

1. Vizgen, Inc., Cambridge, MA, USA

The advent of imaging-based high-plex spatial transcriptomics tools with single-cell resolution has transformed our understanding of the structure, function, and development of the nervous system. These technologies provide the ability to visualize gene expression in its native spatial context, enabling deeper insights into cellular identity, organization, and communication. However, detecting genes expressed at low levels—especially in samples with compromised RNA quality—remains a challenge for downstream biological analysis. Transcription factors, in particular, are difficult to profile due to their typically low abundance, transient expression, and tight regulation across developmental stages.

To address this, the Multiplexed Error-Robust Fluorescence In Situ Hybridization 2.0 (MERFISH 2.0) chemistry and sample preparation workflow were developed to enhance transcript detection efficiency for up to 1,000 genes. This technology enables highly sensitive, spatially resolved RNA profiling directly in tissue, while preserving cellular architecture and gene expression patterns.

In this study, we applied MERFISH 2.0 to comprehensively profile transcription factors during mouse development. Two panels, each containing 815 genes, were designed to cover the full repertoire of known mouse transcription factors. Multiple developmental stages of mouse embryos and postnatal brains were sectioned onto MERSCOPE® Ultra™ slides and analyzed using both panels. The resulting data enabled detailed spatial and single-cell analyses, demonstrating that MERFISH 2.0 provides robust detection of transcription factors throughout embryonic and early brain development.

High-resolution, spatially resolved profiling of transcription factors at the single-cell level opens new avenues for studying neuronal development, function, and disease. This approach will facilitate exploration of transcriptional dynamics that drive lineage commitment, neurogenesis, and brain regionalization, offering a powerful tool for investigating gene regulation in complex tissues.

A multi-tissue atlas of spatial gene-expression adaptations in late pregnancy

Technology & Application Development

POSTER 479

Darren Segale[1], Sujai Chari[1], Kareem Ahmad[1], Niranjan Vissa[1], Mats Ekstrand[1], Yao Xiao[1], Sarah Nainar[1], Mark Wang[1], Cande Rogert[1], Emily Parker[1], Reza Riahi[1]

1. Illumina, San Diego, CA, USA

Pregnancy in mammals is associated with a range of physiological changes in multiple organ systems. In addition to reproductive organs, adaptations supporting pregnancy occur in the nervous, cardiovascular, renal, digestive and hematological systems, among others. Disruptions of these critical changes can have serious impacts on maternal and fetal health and ultimately the viability of the pregnancy.

To provide a comprehensive view of spatially defined gene expression changes in pregnancy, we used Illumina spatial technology with fresh frozen sections from virgin and late pregnancy (E18.5) mice. This technology uses a single large (15x50 mm), barcoded capture surface to append spatial addresses to individual transcripts at a high resolution (1µm feature size). The assay is fully compatible with H&E histological staining and has built in AI-supported cell segmentation. We leveraged the large, uninterrupted, capture areas to mount 4-6 sections from multiple organs known to undergo adaptive changes during pregnancy and that spans a range of sizes. Spatially barcoded cDNA was sequenced on a NovaSeq™ X instrument and analyzed using Illumina Connected Multiomics (ICM). We generated 1-2 million single-cell gene expression profiles per spatial slide and detected ~500-1100 genes per cell depending on tissue type. Finally, we performed cell typing and differential gene expression analyses for all studied tissues, including heart, kidney, liver and small intestine.

In livers of pregnant mice, our data revealed gene expression changes consistent with hepatic growth where hepatocytes exhibited spatially heterogenous adaptations indicative of an altered developmental state. We were able to identify all key cell types of the intestinal epithelium (ileum) and detect pregnancy specific gene expression in a subset of enterocytes, consistent with metabolic changes and altered nutrient uptake. In cardiomyocytes of pregnant mice we found gene expression changes related to contractile function, extracellular matrix remodeling, protection from oxidative stress and modulation of steroid receptor signaling. Finally, we were able to detect changes in cell signaling pathways related to metabolism and cellular growth in the kidney medulla of pregnant mice. In summary, we are able to measure spatially localized gene expression changes directly related to known physiological adaptations that are critical to a viable pregnancy.

For Research Use Only. Not for use in diagnostic procedures.

HiFi sequencing with multi-use SMRT Cells

Technology & Application Development

POSTER 480

Aaron Wenger[1], Matthew Boitano[1], Richa Pathak[1], Andrei Fedorov[1], Daniel Roitman[1], Satwik Kamtekar[1]

1. Pacific Biosciences, Menlo Park, CA, USA

The cost of sequencing consumables, particularly SMRT Cell microchips, has limited the scale of studies based on HiFi long-read sequencing, constraining adoption in large population and disease genomics projects. As a result, many large-scale efforts have been unable to leverage the comprehensive insight into genome structure and methylation status provided by HiFi sequencing. To address these cost and scalability constraints, we developed multi-use SMRT Cells and a compatible sequencing chemistry that improve both efficiency and sustainability. Advances in surface durability, flowcell design, and a new washing chemistry allow each SMRT Cell to be reused for multiple sequencing runs while maintaining optical and biochemical integrity. We evaluated this capability across whole-genome and Kinnex RNA sequencing applications. Base yield (>100 Gb), read length (>15 kb), variant-calling accuracy (SNV >99.8% F1, indel >98% F1, SV >95% F1 at 20X coverage), and methylation detection (both native 5mC and exogenous 6mA from Fiber-seq) remained consistent across sequential uses. Genome assembly contiguity (>80 Mb N50) was also stable, confirming uniform data quality across runs. Experiments with barcoded libraries demonstrated that sample carryover between uses is negligible (<1 in 8,000 reads) following automated wash and reload cycles. Collectively, multi-use SMRT Cells reduce the effective cost of HiFi sequencing by approximately 30–40% while maintaining the accuracy, completeness, and methylation fidelity that characterize the technology. This advance broadens access to long-read sequencing, enabling larger and more comprehensive genomic and multiomic studies.

svCapture and HiFiRe3: single-molecule SV and SNV detection through practical and comprehensive approaches to error correction

Technology & Application Development

POSTER 482

Thomas E. Wilson[1,2], Joseph Stewart[1,2], Jeanmarie Mishler[1], Samreen Ahmed[1,2], Thomas W. Glover[1,2]

1. University of Michigan, Department of Pathology, Ann Arbor, MI, USA
2. University of Michigan, Department of Human Genetics, Ann Arbor, MI, USA

Detecting ultra-rare/single-molecule genetic variants has become critical to genome analysis in somatic mutation assessment, genetic toxicology, and experimentation to define mutagenic mechanisms. Duplex sequencing enables single-nucleotide variant (SNV) error correction, and our probe-based capture method, svCapture, enables single-molecule structural variant (SV) detection through tagmentation-based suppression of ligation artifacts and strand-family correction of chimeric PCR. However, few methods achieve all goals. We report a deep analysis of the sources, frequencies, and sequence characteristics of SV artifacts across common short and long-read sequencing platforms, which include chimeric PCR, ligation and transposon artifacts, mapping errors, and platform-specific crossover reads. We deployed an initial error correction approach based on "forced restriction enzyme ends" to characterize single-molecule SVs in mouse brains using targeted nanopore sequencing and ligation-mediated libraries. While successful, residual artifacts in these experiments prompted derivative library methods that add strict 1N to 2N pre-ligation size selection to achieve High Fidelity sequencing via Restriction Enzyme-mediated Reduced Representation (HiFiRe3). We describe how HiFiRe3 (i) provides SV call accuracy consistent with single-molecule analysis, (ii) can be extended to calling both mosaic SVs and SNVs using PacBio sequencing, and (iii) can support targeted sequencing using Element Trinity capture or Oxford Nanopore adaptive sampling. Together, we report a comprehensive analysis of the artifacts that impede rare SV detection and describe practical error suppression and correction approaches that enable both SV and SNV single-molecule detection and that, with minor modifications, could be adapted to accurate and efficient whole genome sequencing.

Mapping immune repertoires in FFPE tissue with Direct-Seq™: in-situ sequencing of IgH and TCRβ transcripts at subcellular resolution

Technology & Application Development

POSTER 483

Daan Witters[1], Tung T. Le[1], Ryan Shultzaberger[1], Ashley Tsue[1], Zane Hiatt[1], Nathan Ing[1], Kenneth Gouin III[1], Michael Lawson[1], Eli Glezer[1]

1. Singular Genomics Systems, Inc., San Diego, CA, USA

Spatial multi-omics is transforming our understanding of the tumor microenvironment, immune responses, and disease mechanisms. While spatial transcriptomics has made major advances, a critical gap remains: the ability to sequence variable transcript regions in-situ at single-cell and sub-cellular resolution. This is especially important for immune repertoire profiling, where mapping B- and T-cell clonotypes within their native context provides unique biological insight.

Here, we present Direct-Seq™, a novel in situ sequencing technology built on the G4X™ platform that enables in situ sequencing of RNA at subcellular resolution. Direct-Seq uses probes designed to target the variable (V) and joining (J) regions flanking the highly diverse CDR3 domain of IgH and TCRβ transcripts, sequencing both the J region and the CDR3 sequence. The workflow combines in situ reverse transcription, amplification, and sequencing-by-synthesis chemistry from Singular Genomics. In this work, 5-µm formalin-fixed paraffin-embedded (FFPE) tonsil and renal cell carcinoma (RCC) sections were transferred to G4X slides, deparaffinized, and subjected to antigen retrieval. Similarly, 10-µm fresh frozen (FF) tonsil sections were permeabilized and processed using the same Direct-Seq workflow.

In FFPE tonsil sections, up to 9% of B cells were profiled, while in FF tissue approximately 20% of both B and T cells were detected. Across both tissue samples, extensive CDR3 sequence diversity was observed. Clonally expanded B cells with highly similar CDR3 sequences were spatially localized within germinal centers, highlighting the high-resolution clonotyping capabilities of Direct-Seq.

We applied Direct-Seq to an RCC sample to explore immune repertoire organization in the tumor microenvironment. Across nine serial sections of kidney tissue, we identified clonally expanded B-cell populations enriched around tumor regions, and several clonally expanded T-cell groups. One dominant T-cell clone was consistently detected across all serial sections analyzed, revealing persistent clonal expansion and infiltration within the tumor.

Finally, Direct-Seq was combined with multiplexed protein detection on the same tissue sections, confirming spatial correlation between transcript identity and phenotypic protein expression.

In summary, Direct-Seq enables spatially resolved sequencing of IgH and TCRβ transcripts, uncovering immune clonality and distribution with subcellular precision, and offering a powerful approach for advancing spatial immunology, cancer biology, and translational research.

Low-cost, high-resolution spatial transcriptomics with Seq-Scope: validation in mouse kidney

Technology & Application Development

POSTER 484

Yanxia Wu[1], Jessica Kepple[1], Bobby Bojovic[1], Rui Qi[1], Katherine Bull[1], Paolo Piazza[1]

1. University of Oxford, Centre for Human Genetics, Oxford, United Kingdom

Spatial Transcriptomics (ST) method can broadly be categorized into two main types: probe-based imaging platforms such as Xenium, and sequencing-based technologies including Visium, Slide-seq, and others. While sequencing-based ST is target-agnostic and well suited for discovery-driven research, it typically relies on the production of chip with an array of capture probes, mostly poly-dT sequence. These chips determine spatial resolution and are a primary driver of the high cost associated with commercially available platforms.

Seq-Scope (Cho *et al.,* 2021, Kim *et al.,* 2024, Schott *et al.,* 2024) introduces a game-changing approach by repurposing Illumina Novaseq 6000 S4 sequencing flow cell to generate ST chips in-house. This enables significantly reduced cost while achieving ultra-high-resolution datasets.

We validated Seq-Scope using mouse kidney tissue, from chip production to library sequencing. Our results indicate successful chip generation with optimal cluster density. The base composition at each cycle reflects the design of the HDMI, indicating that the HDMI randomer is successfully clustered. Sequencing yielded ~544 million paired-end reads, covering 32 bp spatial barcode and 115 bp transcript, sufficient for downstream validation. Unique barcode per tile is ~2.5 million, further confirming the fully optimized chip production.

Using a 12-μm hexagonal grid, we visualized the gene counts per spital unit. The resulting gene expression map aligns precisely with kidney tissue architecture. Key renal markers appear in their expected anatomical locations: *Umod* and *Slc12a1* co-localize in the tubular region of the nephron; *Gatm* and *Slc34a1* are both highly expressed in proximal tubules; *Nphs2* and *Podxl* are expressed in podocytes within discrete small regions of the cortex.

Downstream analysis, including cell segmentation and cell type clustering, are in progress. Together, these results demonstrate that Seq-Scope enables cost-effective, ultra-high-resolution spatial transcriptomics with accurate tissue architecture mapping. We believe Seq-Scope will be transformative for both research and clinical translation.

Single cell multi-omics sequencing in Alzheimer's disease reveals distinct regulatory and aging signatures

Technology & Application Development

POSTER 485

Bohan Zhu[1], Anurupa Ghosh[2], Zhe Wang, Agata Kurowski[3], Ka lung Cheung[2], Martin Walsh[3], Sai Ma[1]

1. Icahn School of Medicine at Mount Sinai, Department of Genetics and Genomic Sciences, New York, NY, USA
2. Icahn School of Medicine at Mount Sinai, The Mount Sinai Center for RNA Biology and Medicine, New York, NY, USA
3. Icahn School of Medicine at Mount Sinai, Department of Pharmacological Sciences, New York, NY, USA

Alzheimer's disease (AD) is a devastating age-related neurodegenerative disorder that affects approximately 1 in 9 individuals aged 65 and older in the United States and remains one of the greatest biomedical challenges of our time. Despite decades of research, the molecular drivers of hallmark AD pathologies—such as amyloid-beta plaques and tau tangles—are still not fully understood. This is in part due to a critical gap: the lack of cell type-specific insights into epigenetic regulation during disease progression.

Emerging evidence points to DNA methylation as a major player in AD. Several bulk tissue studies have attempted to identify methylation differences in neuronal and non-neuronal cells during AD progression. However, these bulk approaches studies obscure the nuances of how DNAme shapes gene regulation in individual cell types. Understanding these relationships at single-cell resolution is essential to unraveling the complex molecular networks underlying AD.

To address this, we developed ME-seq, a groundbreaking, cost-effective single-cell multi-omic technology that simultaneously captures gene expression, chromatin accessibility, and DNA methylation in individual cells, at a massive scale. This innovative method enables unbiased discovery of regulatory pathways, transcription factors, gene programs that are linked to DNA methylation at unprecedent resolution. Using ME-seq, we generated a comprehensive single-cell dataset from multiple brain regions that are strongly affected by AD. The samples encompassed several developmental and aging stages and included both wild-type and AD model mice. Leveraging the power of the new technology and next generation sequencing, we dissect cellular and molecular diversity of AD with transformative depth and precision, transforming our understanding of epigenetic regulation in Alzheimer's disease and leading to new therapeutic strategy.

svArcher: a reproducible structural variant calling pipeline module for plant genomes

Technology & Application Development

POSTER 486

Trevor Mugoya[1], Arun Sethuraman[1]

1. San Diego State University, Department of Biology, San Diego, CA, USA

Structural variation refers to differences in genomic regions that are larger than 1 Kbps between individuals, and can arise as a result of errant DNA repair mechanisms, whole genome duplications and transposable element activity across the genome. Recent advances, optimizations and cost reductions in next generation sequencing technologies have facilitated the exponential increase in the amount of available short and long read genomic data. Subsequently, the number of plant species sequenced has grown over the past 20 years, paving the way for a better understanding of their genetic diversity. Plant genomes also vary widely in ploidy, size, and amount of repeat content. This presents challenges in detecting structural variation, especially with short read genomic sequence data. Numerous methods and bioinformatics tools have been developed to detect signatures of structural variation from short read sequence data. Here we present *svArcher*, a reproducible structural variant calling pipeline. *svArcher* implements an ensemble suite of structural variant callers (*lumpy-sv, delly* and *wham*) in addition to an integrated benchmarking with *Truvari* as a module within the *snpArcher* pipeline. We apply *svArcher* to short read whole genome plant sequence data using *Arabidopsis thaliana* as a study species for benchmarking against a set of 155,440 structural variants (14116 duplications, 81490 deletions and 3501 inversions) and report F1-Score, precision and recall. We utilize 1163 samples from BioProject(PRJNA273563) to validate pipeline workflow. Preliminary evaluation metrics show we achieve an F1 score of ~0.5 for duplications and deletions and 0.04 for inversions, precision scores of 0.4, 0.57 and 0.02 for duplications, deletions and inversions respectively, and recall scores of 0.71, 0.47 and 0.66. Further development is currently underway to optimize call merging and filtering parameters and *svArcher* can be found at the snpArcher fork https://github.com/ChabbyTMD/snpArcher under the svArcherRebase branch.

10x GENOMICS

AI WORKSHOP

Building biology's digital twin

Shape your AI future with the precision of 10x

AI models that transform biological discovery start with the highest quality data. 10x single cell and spatial assays deliver unmatched sensitivity and the most efficient scaling so you can build a solid foundation for more impactful AI models.

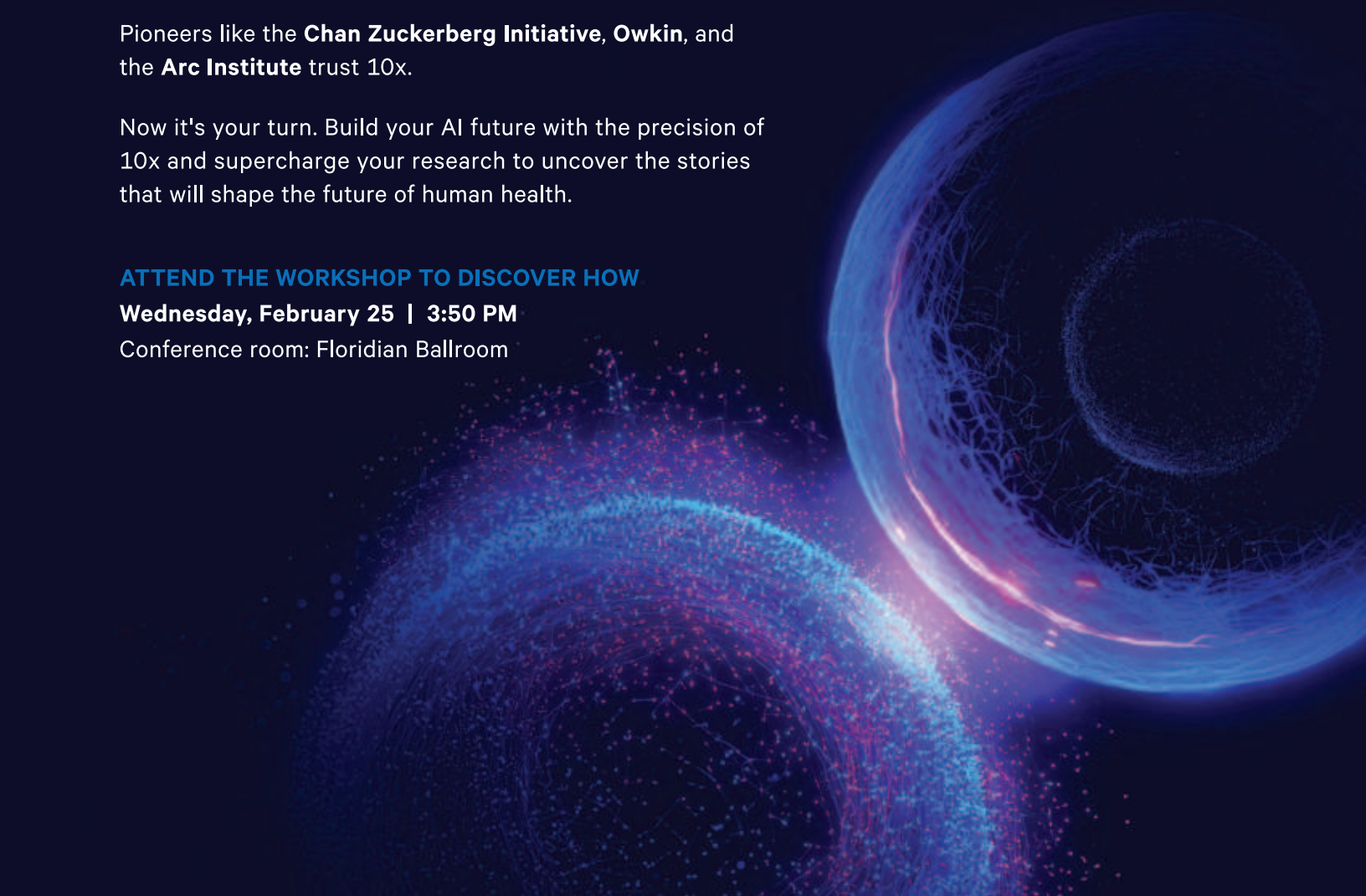
Pioneers like the **Chan Zuckerberg Initiative**, **Owkin**, and the **Arc Institute** trust 10x.

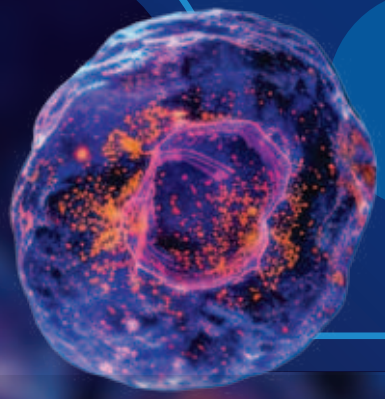
Now it's your turn. Build your AI future with the precision of 10x and supercharge your research to uncover the stories that will shape the future of human health.

ATTEND THE WORKSHOP TO DISCOVER HOW

Wednesday, February 25 | 3:50 PM

Conference room: Floridian Ballroom



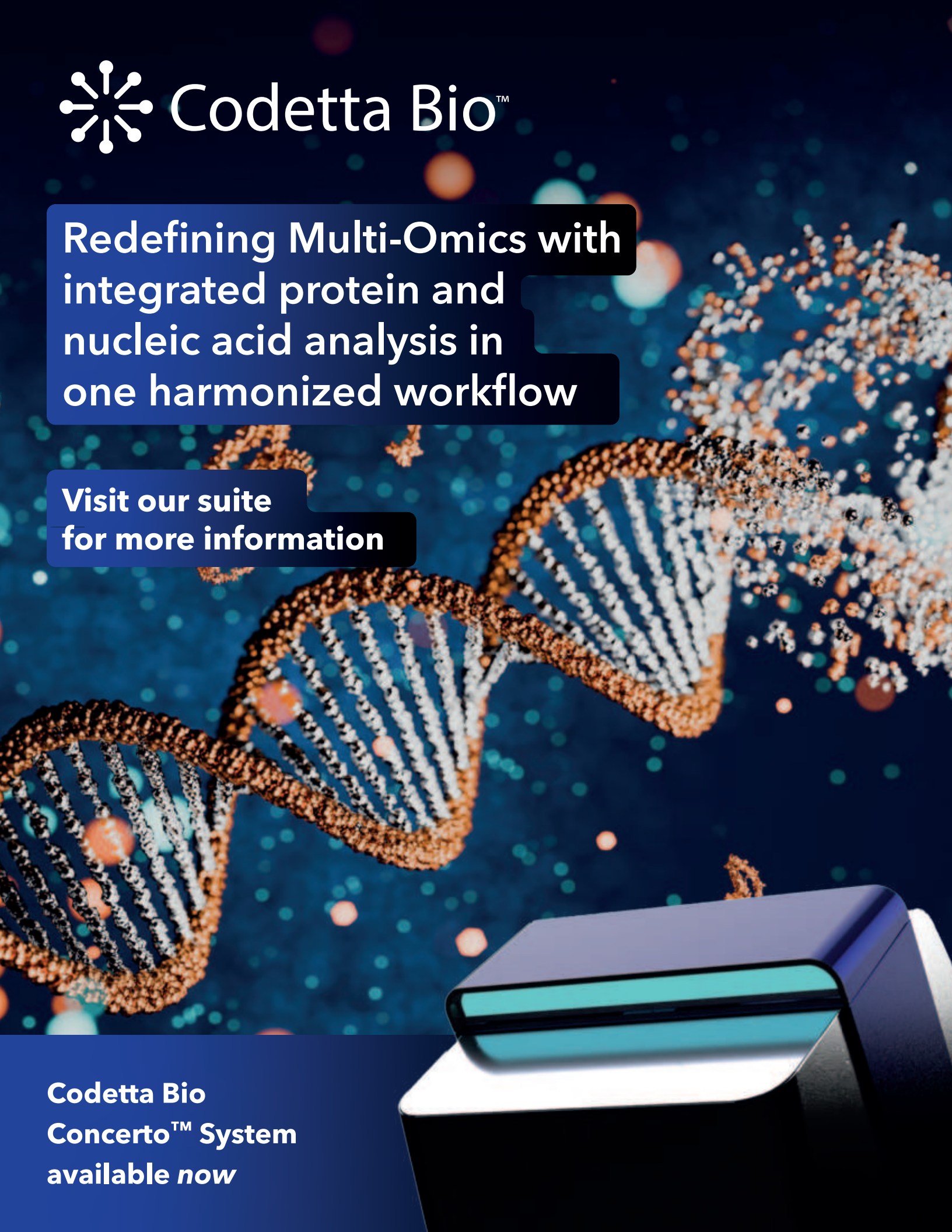


Experience more biology in the Hamilton Suite

High-Confidence Multiomics + Deep Discovery at 1 Cent per Cell

- Gene Expression: WTA & Targeted
- Protein Expression: CITE-seq
- Gene Regulation: ATAC-seq (with Multiomics)
- Immune Repertoire Profiling
- Workflow Automation Solutions
- Sample Multiplexing
- Sample Preservation Solutions
- Data Visualization & Analysis

BD


Codetta Bio™
Redefining Multi-Omics with integrated protein and nucleic acid analysis in one harmonized workflow
Visit our suite for more information
Codetta Bio Concerto™ System available *now*

Index

Index

Index

Index

Index

AGBT 2026™

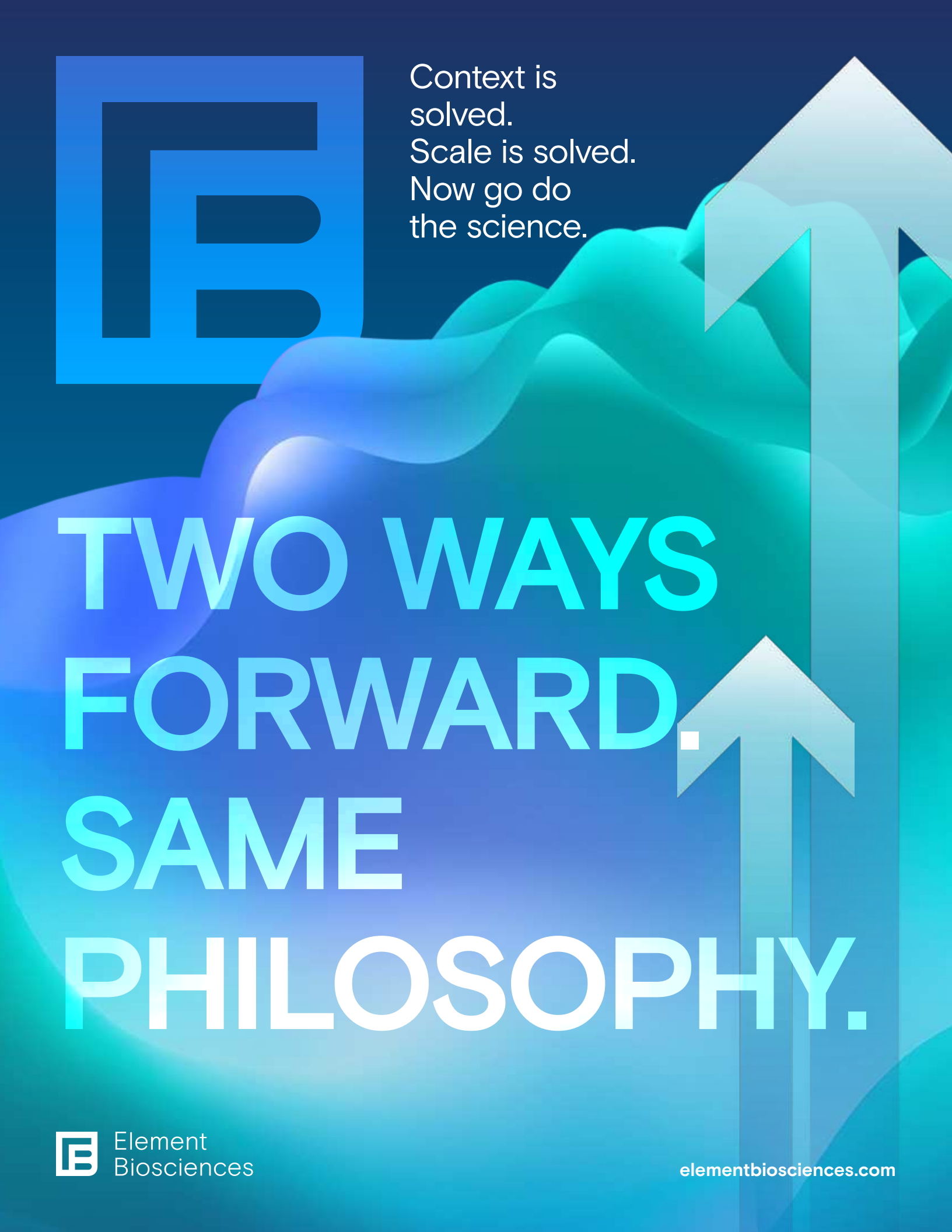
PRECISION HEALTH

AGBT RETURNS TO CORONADO ISLAND IN 2026

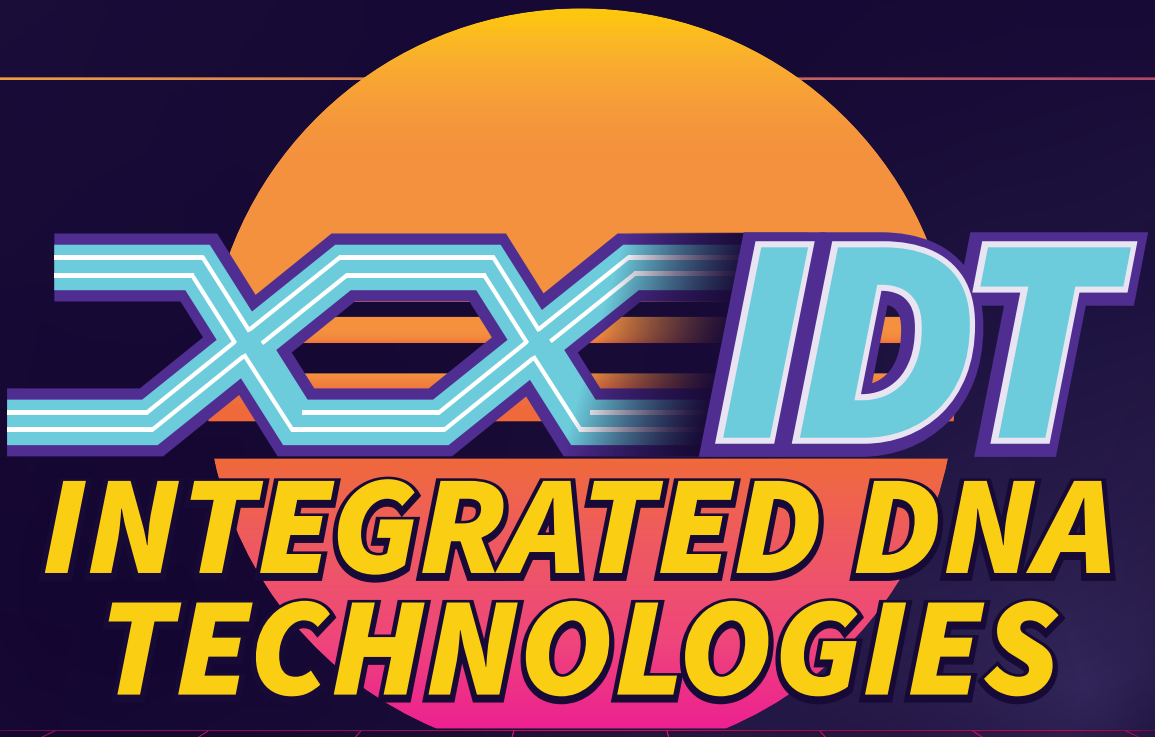
LOEWS CORONADO BAY RESORT, SAN DIEGO, CA

SAVE THE DATE: SEPTEMBER 14 – 16, 2026

LEARN MORE AT WWW.AGBT.ORG


Context is
solved.
Scale is solved.
Now go do
the science.
TWO WAYS
FORWARD.
SAME
PHILOSOPHY.
Element
Biosciences
elementbiosciences.com

COLLABORATION AT SCALE—INNOVATING TOGETHER



Join us for an unforgettable 80s throwback party—
honoring the decade that IDT first began accelerating science.
Party with us Tuesday, February 24, at 9:30 PM in the Nassau Suite.

Don't miss our bronze sponsor talk with IDT President Ajay Gannerkote

ACCELERATING ENHANCED GENOMIC SOLUTIONS

February 25 | 3:00 PM

danaher

IDT™
INTEGRATED DNA TECHNOLOGIES

The path from **insight to impact**

Explore how Illumina continues to **redefine the limits of what's possible.**

Visit Illumina in the Orange Lounge to explore our latest innovations in library prep, sequencing systems, and analysis.

Pre-register to watch the Illumina Gold Sponsor session on-demand after the event.

illumina®