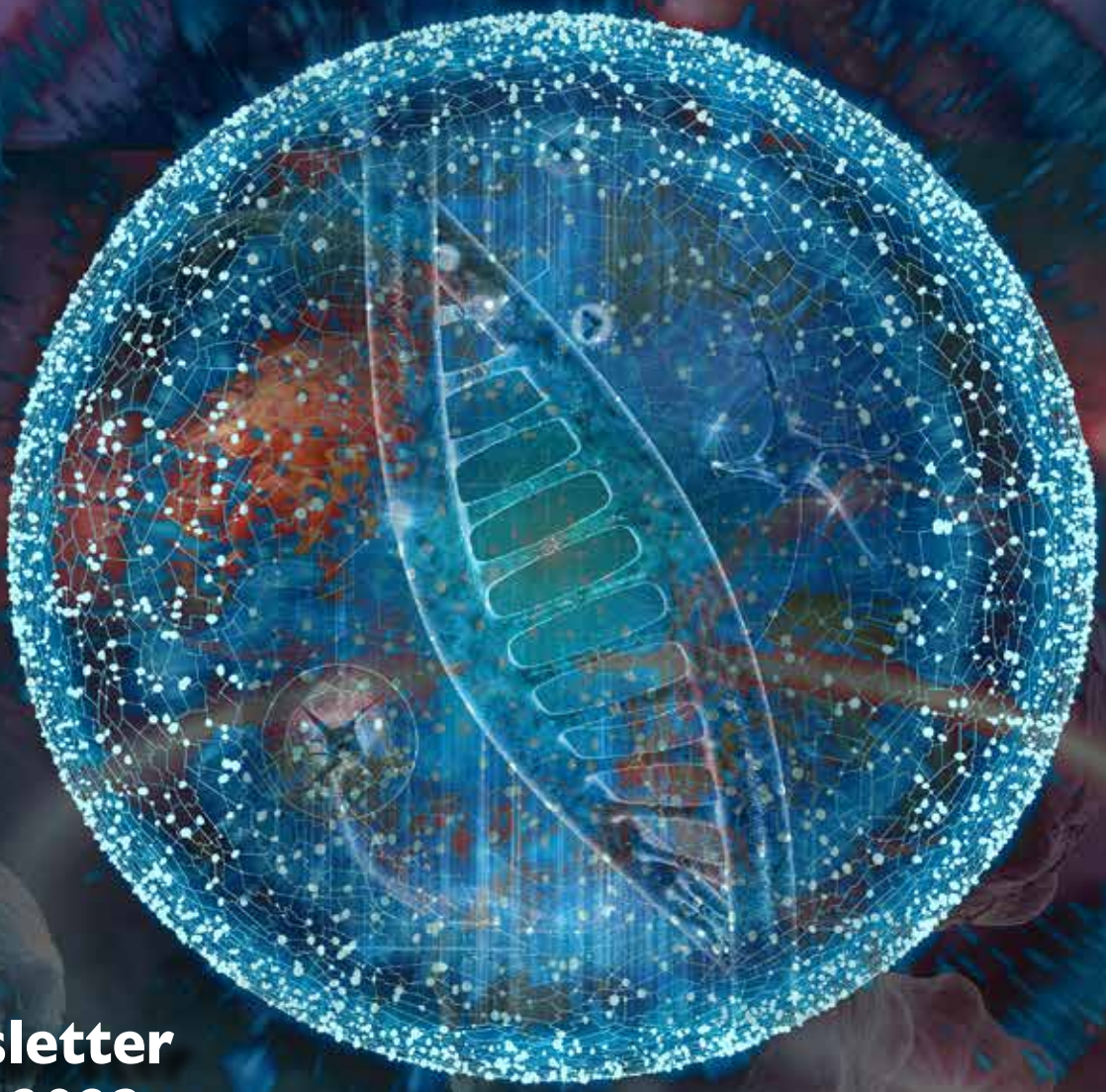
Department of Systems Biology

Columbia University Irving Medical Center



Newsletter 2022–2023

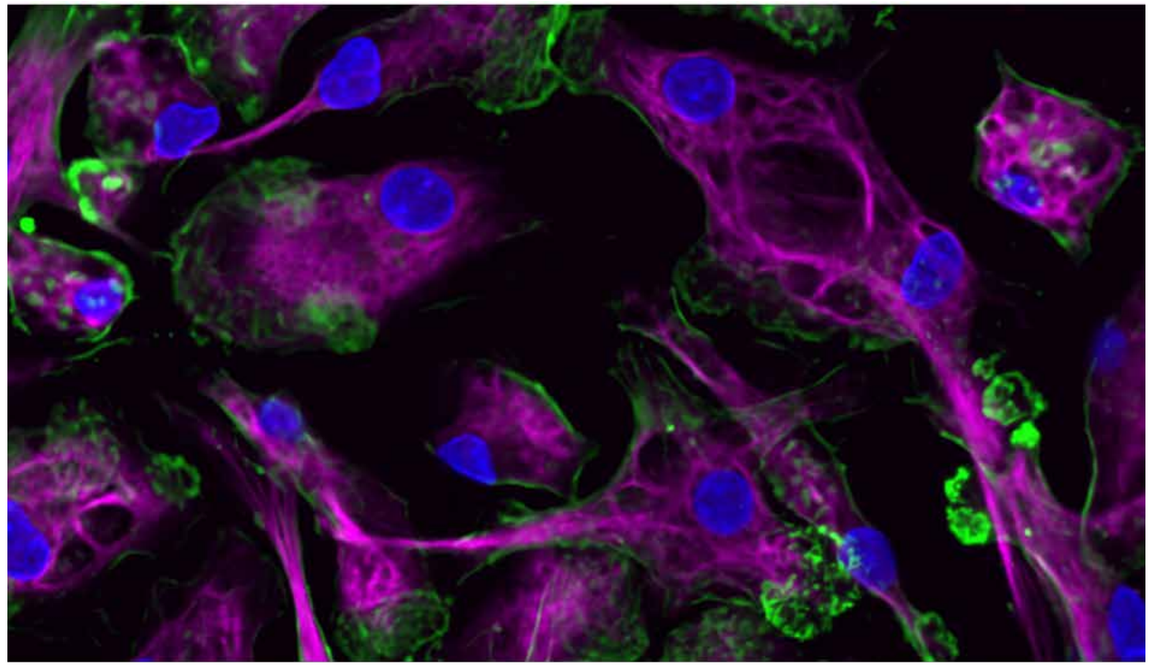
COLUMBIA UNIVERSITY
IRVING MEDICAL CENTER
Department of Systems Biology

Table of Contents

Highlights 2022–2023

Inside this issue, learn about the latest research, noteworthy grants and milestones from the Department of Systems Biology.

Andrea Califano to Lead New Biohub



Immune cells derived in the laboratory from human-induced pluripotent stem cells. Image courtesy Gordana Vunjak-Novakovic/Columbia University.

Andrea Califano, Dr, who just stepped down as the Department of Systems Biology chair, is to lead the new Chan Zuckerberg Biohub NY (CZ Biohub NY). The Biohub, which is part of the CZ Biohub Network, brings together scientists from Columbia, Rockefeller, and Yale.

The scientists will attempt to engineer the cells of the immune system to act like miniature doctors in the bloodstream, detecting and eradicating diseases in their earliest states, years before they may produce detectable symptoms. The Biohub will not be experimenting on patients; its goal, rather, is to create the scientific understanding, technology, and bioreagents that will enable startups and pharmaceutical companies to turn their discoveries into clinical reality.

The researchers will start with cancers that are currently too difficult to detect until they are largely untreatable, like ovarian and pancreatic cancer. They will also try to detect incipient neurodegenerative diseases such as Parkinson's and Alzheimer's. Though a number of therapies are starting to emerge that can slow progression of these diseases, they work only in the very early states, which typically go undetected.

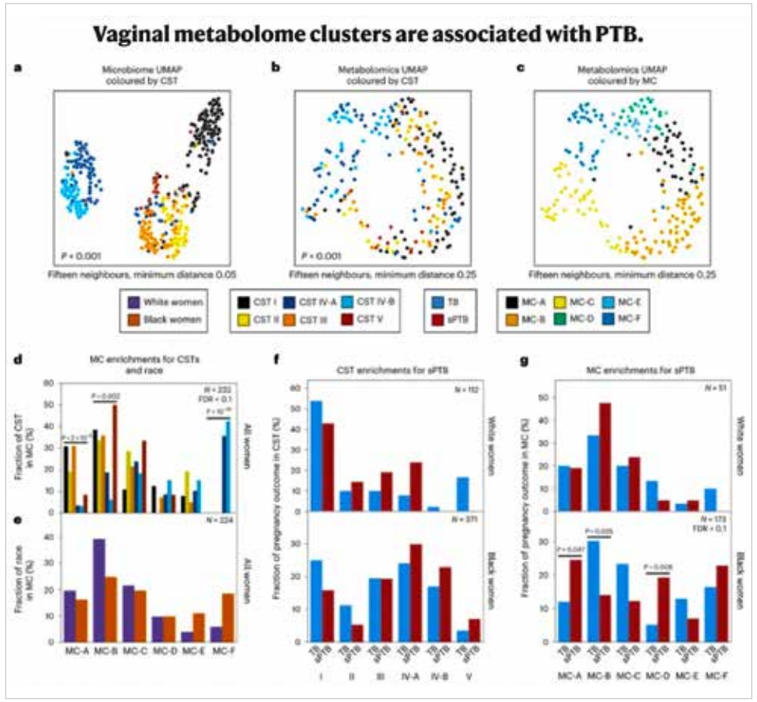
Preterm Birth Linked to Chemicals Found in the Vagina

Cosmetics and hygiene products contain non-biological chemicals that accumulate in the vagina and may contribute to spontaneous preterm birth, according to a study co-led by Tal Korem, PhD, Systems Biology and Obstetrics and Gynecology, and Maayan Levy, PhD, of the University of Pennsylvania. The study was published January 12 in *Nature Microbiology*.

Preterm birth, defined as childbirth before 37 weeks of pregnancy, is the primary cause of neonatal death and can lead to a variety of lifelong health issues. Two-thirds of preterm births occur spontaneously, but despite extensive research, clinicians have no way of predicting or preventing spontaneous preterm birth.

The research team looked at the metabolome of the vaginal microenvironment. The metabolome is the complete set of small molecules found in a particular biological niche; it includes metabolites produced by local cells and microorganisms, as well as molecules from external sources. The researchers measured more than 700 different metabolites in the second-trimester metabolome of 232 pregnant women, including 80 from pregnancies that ended prematurely.

A number of metabolites were significantly higher in women who had delivered early than in those who had delivered at full term. Several of the metabolites were chemicals that are not produced by humans or microbes; these included diethanolamine, ethyl-beta glucoside, tartrate, and ethylenediaminetetraacetic acid—all of which are found in cosmetics and personal hygiene products. The researchers also developed an algorithm based on metabolite levels that can predict preterm birth, potentially paving the way for early diagnostics.



a–c, UMAP ordination of microbiome (a, N = 503) and metabolomics data (b and c, N = 232), coloured by CSTs (a and b) or de novo clustering of metabolites data (c, MCs; Methods). The vaginal microbiome and metabolome are significantly separated by CSTs (PERMANOVA $P < 0.001$ for both), yet the separation is less clear in the metabolome. For similar plots coloured by maternal race, see also Extended Data Fig. 4c,d. d, The fraction of women whose metabolite profiles clustered to each MC, shown for each CST separately. e, Similar to d but shown for Black and White women separately. f, The fraction of White (top) and Black (bottom) women whose microbiomes belonged to each CST, separated by pregnancy outcome. g, Similar to f, for the fraction of women whose metabolomes clustered to each MC. We show a significant association of sPTB with MCs A, B and D among Black women ($P = 0.047$, $P = 0.025$ and $P = 0.006$, respectively, $q < 0.1$). Number above horizontal lines in d–g is two-sided Fisher's exact P, $q < 0.1$..

REFERENCE

Kindschuh WF, Baldini F, Liu MC, Liao J, Meydan Y, Lee HH, Heinken A, Thiele I, Thaiss CA, Levy M, and Korem T. Preterm birth is associated with xenobiotics and predicted by the vaginal metabolome. *Nature Microbiology* 2023 Jan 12;8:246-259. doi.org/10.1038/s41564-022-01293-8.

Illuminating Tumor Cells with Dark Proteins

Columbia researchers have shined new light on how the "dark" part of the genome allows cancer cells to be detected by the immune system, which could lead to better immunotherapies. How tumor cells display fragments of these "dark" proteins was an open question, now answered in a new study published in *Nature* by Xuebing Wu, PhD, and his team.

The immune system recognizes cancer cells by the cells' tumor-specific antigens, fragments of degraded proteins found uniquely on the surface of cancer cells. Previous studies have shown that the vast majority of tumor-specific antigens are produced from the noncoding genome, the "dark" part of the genome that until recently scientists believed did not code for any protein.

"Many previous studies have tried to describe this process, but each of those focused on just a single or a handful of DNA sequences," Wu says. "They have made very important discoveries, but there is little agreement with each other. It's unclear whether the mechanisms they discovered applied broadly to all types of noncoding sequences in the genome."

Seeking a more comprehensive answer, Wu and his team, led by graduate students Jordan Kesner and Ziheng Chen, used massively parallel analysis involving about 10,000 human noncoding genetic sequences and millions of synthesized random sequences. "What we discovered is a unified mechanism applicable to all types of noncoding sequences," says Wu.

That led them to uncover a signal unique to noncoding sequences that triggers the degradation of those nonfunctioning protein, the first step toward making antigens. The researchers then knocked out every human gene using the CRISPR technology, which revealed the molecular machinery that recognizes this signal—the BAG6 complex.

"The BAG6 complex thus represents a potential drug target for tuning the visibility of cancer cells to the immune system," says Wu.

A Fundamental Finding

"The tumor-specific antigen is one of the reasons why our research is important," Wu adds, but the paper also uncovered something more fundamental about the way cells deal with nonfunctioning proteins and how some of these nonfunctioning proteins eventually evolve into functional ones.

The majority of "dark" proteins are likely nonfunctional or even toxic in cells and need to be removed. The new research reveals that the BAG6 complex identifies a large percentage of these proteins and sends them to be destroyed.

The production of nonfunctional proteins increases with age and may contribute to the development of neurogenerative disorders and other diseases, as well as cancer. Drugs that target BAG6 function could potentially treat these conditions. "Once you figure out the molecular mechanism inside the cell, you open the door for developing new therapies," says Wu.

REFERENCE

Kesner JS, Chen Z, Shi P, Aparicio AO, Murphy MR, Guo Y, Trehan A, Lipponen J, Recinos Y, Myeku N, and Wu X. Noncoding translation mitigation. *Nature* 2023 Apr 12;617:395-402. doi.org/10.1038/s41586-023-05946-4.

A Better Way to Decontaminate Samples for Early Cancer Detection

Unique microbial signatures in blood had been identified for certain types of cancer, suggesting that blood-borne microbial DNA could be used to distinguish between those patients with cancer and those without, as well as to discriminate between cancer types. However, when clinicians draw blood from a cancer patient, that sample may be "contaminated" by, for example, microbes from the reagents or from the skin of the person processing the samples. In addition, biological material can transfer from one sample to another during sample processing, a phenomenon called "cross-contamination."

A microbiome-based diagnostic tool that requires only a patient's blood sample could be used to detect cancer earlier and more easily than traditional methods. Though such a diagnostic had been proposed (Poore et al., *Nature* 2020), it performed much better for certain types of cancer, e.g., prostate and lung cancer, than for melanoma. Tal Korem, PhD, and his team thought that the disparity had to do with issues of sample contamination.

Korem and his team developed a method called Source-tracking for Contamination Removal in microBiomes (SCRuB), for high-precision decontamination of microbial data using control samples. Essentially, SCRuB incorporates shared information across multiple samples and controls to precisely identify and remove contamination.

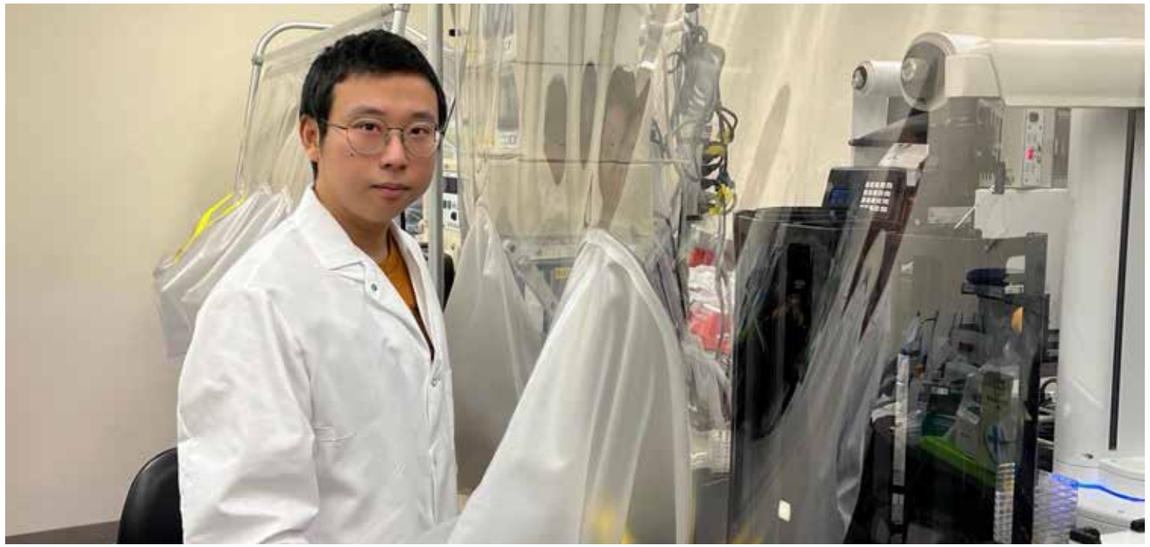
SCRuB was found to be highly predictive for melanoma; for other cancer types, it performed comparably to other methods. The researchers were also able to use SCRuB to develop predictors for immunotherapy treatment response using microbial DNA for the melanoma itself.

The study, co-led by Korem and Liat Shenhav, PhD, of Rockefeller University, was published in *Nature Biotechnology*.

REFERENCE

Austin GI, Park H, Meydn Y, Seeram D, Sezin T, Lou YC, Firek BA, Morowitz MJ, Banfield JF, Christiano AM, Pe'er I, Uhlemann A-C, Shenhav L, and Korem T. Contamination source modeling with SCRuB improves cancer phenotype prediction from microbiome data. *Nature Biotechnology* 2023 Mar 16;41:18201-828. doi.org/10.1038/s41587-023-01696-w.

AI-Guided Robotic Culturomics for Microbiome Studies



Postdoctoral research scientist Yiming Huang, PhD, with the CAMII platform.

The Wang lab has published a paper in *Nature Biotechnology* describing their automated AI-guided robotic microbiome culturomics system, called CAMII (Culturomics by Automated Microbiome Imaging and Isolation).

Culturomics refers to high-throughput strategies for systematically culturing the many bacterial species in a sample. Because isolating and characterizing individual bacterial colonies by hand is far too slow and error-prone for performing large-scale microbiome studies, the Wang lab developed the CAMII system, which can isolate and analyze thousands of microbial colonies in a day.

The CAMII platform includes an anaerobic chamber that encloses the robotic arms that handle the samples and isolate the bacterial colonies. The selection of isolated colonies by the robotic arms is guided by machine learning algorithms that can infer different species based on colony color, size, density, shape, and other related morphological features. These isolated colonies are then processed through an ultra-high-throughput genomic sequencing pipeline and biobanked along with a searchable database with corresponding rich genotypic and phenotypic information.

The technology has allowed the lab to build personalized microbiome strain collections to investigate healthy and diseased conditions. CAMII has already been applied to fecal, skin, oral, and vaginal microbiome samples from humans and animal models. The Wang lab is now using the CAMII system to study the microbiomes of environmental sources such as soil and fabric samples, as well.

The CAMII platform is propelling complex microbiome studies forward and powering the discovery of new insights into interactions between humans and the microorganisms that influence our lives. Researchers worldwide have reached out to train on the platform so they can build their own CAMII systems in their laboratories to advance their microbiome research efforts.

REFERENCE

Huang Y, Sheth RU, Zhao S, Cohen LA, Dabaghi K, Moody T, Sun Y, Ricaurte D, Richardson M, Velez-Cortes F, Blazejewski T, Kaufman A, Ronda C, Wang HH. High-throughput microbial culturomics using automation and machine learning. *Nature Biotechnology* 2023 Feb 20;41:1424–1433. https://doi.org/10.1038/s41587-023-01674-2.

FACULTY SPOTLIGHT

Q&A with Dian Yang

The advent of CRISPR-Cas9, the Nobel Prize-winning gene-editing technology, has opened up countless new possibilities for researchers in biotechnology and medicine. The lab of Dian Yang, PhD, leverages CRISPR-Cas9—along with single-cell genomics, in vivo cancer modeling, and computational analysis—to investigate tumor evolution. A CRISPR-based molecular recorder, which Yang developed as a postdoc with Jonathan Weissman at MIT, has the ability to track the lineage relationships of cancer cells as they evolve.

In September 2023, Yang joined Columbia as an assistant professor in the Departments of Systems Biology and of Molecular Pharmacology and Therapeutics. His lab plans to examine the intrinsic and extrinsic mechanisms that govern cell-state transitions, with the goal of developing a comprehensive and quantitative roadmap of tumor evolution.

Q: *How did you become interested in tumor evolution?*

A: As a cancer biologist, I wanted to understand the process of how a normal cell becomes an aggressive, metastatic tumor. Over the past decade, many studies using single-cell sequencing have identified an incredible amount of heterogeneity in a tumor. So, we know that there are different cell states in the tumor that all come from one progenitor cell, but it remains unclear how these cell states are related to one another.

In our lab, we uncover lineage relationships that can help us understand the evolution process. With phylogenetic analysis—similar to how researchers traced the evolution of SARS-CoV-2 variants during the pandemic—we can track back in history to understand what happened in the past. So many questions can be answered by reconstructing the lineage tree, such as how different clones grow in the tumor, where specific driver mutations occur that drive clones to become metastatic or evade therapies, and many others.

Q: *How does the CRISPR-based molecular recorder work?*

A: We use CRISPR-Cas9, which you can think of as a pen that writes information into a DNA sequence. It generates a double-strand break that can lead to different repair outcomes and basically writes in a mutation—either an insertion or deletion of bases, or "indels." We have different pairs of Cas9 and guide RNA that link to synthetic DNA target sites, along with a barcode associated with each site, to progressively generate heritable indels.

In an experiment, we first turn on CRISPR-Cas9, which starts generating mutations in two daughter cells. These cells further divide, and those daughter cells inherit the same mutations from the parental cells. At the same time, because Cas9 keeps cutting, new mutations specific to the daughter cells are generated, as well. At the end of the experiment, we collect the final cells, sequence them based on the cumulative pattern of mutations, and reconstruct the cell history.

Another key feature of the system is that the DNA recording site can be transcribed into mRNA, so we can decode this information with single-cell RNA sequencing. Usually, single-cell RNA sequencing is used to get the transcriptome, or the gene-expression state of the cell. But now, single-cell RNA sequencing can give us the evolution history of the cell. So, we can beautifully couple a cell's current state with its evolutionary history.

Q: *Can you give an example of how you applied this technology to study cancer?*

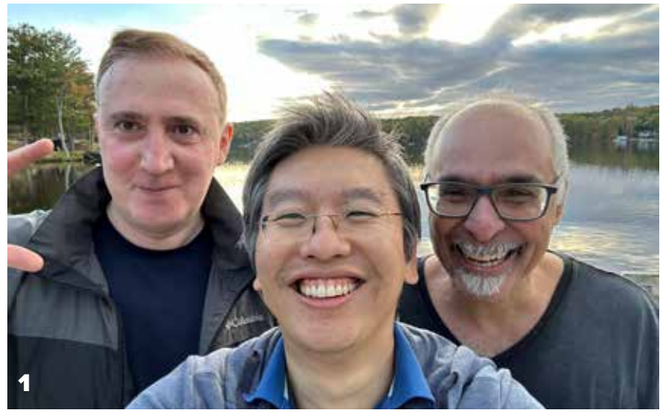
A: We are collaborating with Tyler Jacks' lab at MIT. We introduced the technology into a genetically engineered mouse model of lung cancer. The mouse starts out completely normal. Upon delivery of a special enzyme, we activate the CRISPR-based mutations, and at the same time, we turn on the lineage tracing. This allows us to track a normal epithelial cell in situ as it develops into an aggressive tumor, while at the same time, the lineage tracing follows tumor division and starts tracking all of the cells in the tumor. At the end of the experiment, we can reconstruct the evolution history of the tumor to understand how all the cells in it are related to one another phylogenetically.

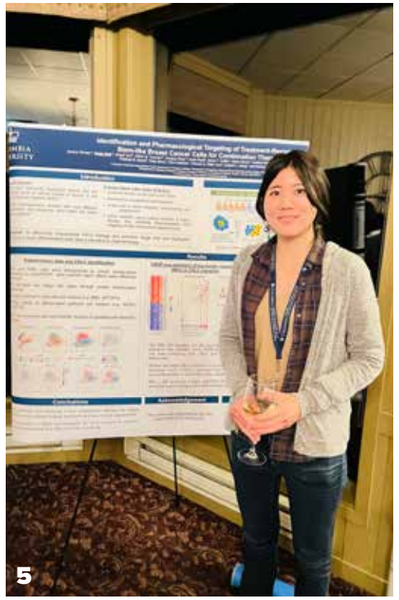
In some tumors, almost all the clones grew in a balanced structure, but in others, one clone starts taking over the tumor. So, there are different clonal dynamics. With the CRISPR-based molecular recorder, we can identify these clones, determine their specific subgroups, and analyze their gene expression. Then we can identify interesting gene signatures from these groups compared with the other, non-expanding groups. We find, interestingly, that the gene signature associated with those aggressive, expanding clones actually very well predicts poor outcomes in human lung cancer patients.

ANNUAL RETREAT

The Department of Systems Biology's 2023 annual retreat was held in October at Woodloch Pines Resort, PA.

The three poster winners were Yocelyn Recinos for "Two-step adaptive evolution of MAPT exon 10 splicing in primates"; George Austin for "Experimental bias correction and domain adaptation with DEBIAS-M improves cross-study generalization of microbiome-based prediction models"; and Jiayu Su for "Smoother: a unified and modular framework for incorporating structural dependency in spatial omics data."





1) From left: Dennis Vitkup, Harrris Wang, and Aris Floratos. **2)** Postdocs (left) Alejandro Buendia (Rabadan lab) and Haiqing Zhao (Honig lab). **3)** An onlooker. **4)** From left: Dian Yang, Chaolin Zhang, Dennis Vitkup, Tal Korem, Mohammed AlQuraishi, and Harris Wang. **5)** Postdoc Heeju Noh (Califano lab).

Around the Department, 2022-2023

Grants, Awards, and More

Harris Wang, PhD, has been appointed interim chair of the Department of Systems Biology.

The Department has established the **Genome Engineering Core (GECO)** at the JP Sulzberger Columbia Genome Center. Tools available to collaborators include CRISPR knockout, interference, base editing, CRISPRoff, and ORF overexpression.

The **Department of Systems Biology Information Technology (DSBIT)** received $2,000,000 from NIH to upgrade its main high-performance computing cluster.

Mohammed AlQuraishi, PhD, was awarded $2,056,250 over five years from the National Institute of General Medical Sciences for "Machine Learning of Biomolecular Interactions and the Human Signaling Networks They Comprise."

Andrea Califano, Dr, received the 26th Alfred G. Knudson Award in Cancer Genetics from the National Cancer Institute (NCI) for his exceptional contributions to the field of cancer research.

The Center for Cancer Systems Therapeutics (CaST), PIs **Andrea Califano,** Dr, and **Barry Honig,** PhD, Systems Biology, Biochemistry and Molecular Biophysics, and Medicine, was awarded $12,416,129 for 2023–2028 from NCI.

Andrea Califano, Dr, received $4,893,902 over five years from the NCI for "Elucidating and Targeting Tumor Dependencies and Drug Resistance Determinants at the Single Cell Level."

Graduate student **Karin Isaev** (Knowles lab) was awarded a Lead Teaching Fellowship from the Columbia Center for Teaching and Learning.

Ben Izar, MD, PhD, Medicine, was awarded the 2023 Pershing Square Sohn Prize for Young Investigators in Cancer Research.

Brian Joseph, joint postdoc in the Zhang and Wichterle labs, was named a 2023 New York Stem Cell Foundation Druckenmiller Early Postdoctoral Fellow.

Jordan Kesner received the 2023 Dean's Award for Excellence in Research in recognition of the high quality and significance of his PhD thesis research in the Coordinated Doctoral Programs in Biomedical Sciences.

Tal Korem, PhD, will receive $804,804 over two years for a subaward from the National Institute of Diabetes and Digestive and Kidney Diseases for "Microbiome Based Biomarkers of Wound Healing."

Tal Korem, PhD, received a Repro Grant for "Biogeographic and Longitudinal Study of Microbial and Phenotypic Interactions in Endometriosis."

Laura Landweber, PhD, Biochemistry and Molecular Biophysics, will receive $3,858,485 over five years from the National Institute of General Medical Sciences for "Understanding Complex Genome Editing and RNA Biology in Oxytricha."

Rodney Rothstein, PhD, received the Alumni Achievement Award, the highest honor bestowed by the University of Illinois at Chicago Alumni Association.

Robert Schwabe, MD, PhD, Medicine, and Andrea Califano, Dr, received the Columbia University 2023 Precision Oncology and Systems Biology Award 2023–2024 for "Master Regulators as Therapeutic Targets in Hepatocellular Carcinoma." Total award amount $50,000/ Califano lab $12,500.

Yufeng Shen, PhD, received a 5-year grant from NIGMS for "Computational Methods to Interpret Genomic Variation and Integrate Functional Genomics Data in Genetic Analysis of Human Diseases."

Yufeng Shen, PhD, received a 3-year grant from SFARI for "Triangulation of Missense Variant Impact through Multimodal Modeling and Functional Assays."

Harris Wang, PhD, Systems Biology and Pathology and Cell Biology, was named a finalist for the Blavatnik National Awards for Young Scientists.

Harris Wang, PhD, and Samuel Sternberg, PhD, Biochemistry and Molecular Biophysics, will receive a $1,499,999 subaward over five years from the Army Research Office for "CHARMME: Center for Harnessing Microbiota from Military Environments."

Harris Wang, PhD, will receive up to $8,479,837 over four years from DARPA for "IMPEDE: Inhibiting Molds with Probiotic Ensembles from Diverse Environments."

Harris Wang, PhD, will receive an NIH renewal of $2,891,590 over five years for "Micron-scale Spatial Metagenomic Mapping of Microbial Biogeography in the Gastrointestinal Tract."

Harris Wang, PhD, will receive $452,375 over two years from the National Human Genome Research Institute (NIH) for "Rapid and Efficient Generation of Sequence Variants by Templated Synthesis."

Harris Wang, PhD, will receive $359,101 over one year from the Defense University Research Instrumentation Program (DURIP) for "Automated Biobanking for DoD-Relevant Biorepository for Synthetic Biology and Microbial Culturomics."

Harris Wang, PhD, received a Young Investigator Award from the Columbia University Asian Faculty Association.

Xueling Wu, MD, PhD, Medicine, received $2,661,540 over three years from the National Institute of Allergy and Infectious Diseases for "Characterization of HIV-1 IgA bNAbs and ADCP Function."

Xuebing Wu, PhD, Systems Biology and Medicine, received one of the inaugural Glenn Foundation Discovery Awards, $525,000 over three years, for "Aging as a Self-Reinforcing Feedback Loop: Investigate the Role of Noncoding Translation."

Xuebing Wu, PhD, received a Longevity Impetus Grant of $275,000 over 18 months for "Awakening a Heart Ribosome in Brain and Other Tissues to Enhance Proteostasis and Delay Aging."

Xuebing Wu, PhD, with Jianwen Que, MD, PhD, Medicine, received $100,000 for one year from the HICCC Inter/ Intra-Programmatic Pilot Program (IPPP) for "Targeting MYC-Driven Cancer with CRISPR/Cas13 Collateral Activity."

Sara Zaccara, PhD, received the Edward P. Evans Center for Myelodysplastic Syndromes 2023 Pilot Award.

Chaolin Zhang, PhD, received $2,432,096 from NIH/ NHGRI, 6/2023–05/2027, for "Mapping Proximal and Distal Splicing-Regulatory Elements."

Chaolin Zhang, PhD, received a $150,000 Scientific Innovations Award, 1/2023–12/2024, from the Brain Research Foundation for "Human-Specific Alternative Splicing, Brain Development, and Ciliopathies."

Chaolin Zhang, PhD, received $1,999,299 from IH/ NICHD, 4/2023–3/2026, for "Developing RNA Therapeutics for Rare Neurodevelopmental Disorders."

Postdocs Receive Faculty Postitions

Xiao Fan (Chung and Shen labs), assistant professor at the University of Florida

Carlotta Ronda (Wang lab), principal investigator at Innovative Genomics Institute and the University of California, Berkeley

April Shu (Chung and Shen labs), assistant professor at the University of Southern California

PhD Graduates

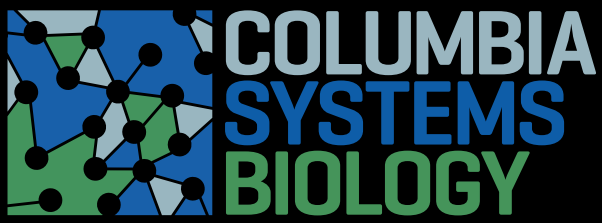
Congratulations to our recent grads!

Shuonan Chen (Paninski lab)
Jenna Kefeli (Tatonetti lab)
William Kindschuh (Korem lab)
Baihan Lin (Kriegeskorte lab)
Zhouzerui Liu (Sims lab)
Daniel Moakley (Zhang lab)
Yocelyn Recinos (Zhang lab)
Deirdre Ricaurte (Wang lab)
Miles Richardson (Wang lab)
Stephen Trudeau (Honig lab)
Dinara Usmanova (Vitkup lab)
Lukas Vlahos (Califano lab)

And congratualtions to:

Aleks Obradovic upon receiving his MD!

COLUMBIA UNIVERSITY
IRVING MEDICAL CENTER



Columbia University Department of Systems Biology
Irving Cancer Research Center
1130 St. Nicholas Avenue
New York, NY 10032

To learn more about our research and programs,
visit **systemsbiology.columbia.edu.**

Design by: Nicoletta Barolini
Edited by: Ann Rae Jonas