When AI meets Biology

JULY 2024 WEBINAR

Integrated analysis of single-cell data across technologies, modalities, and perturbations

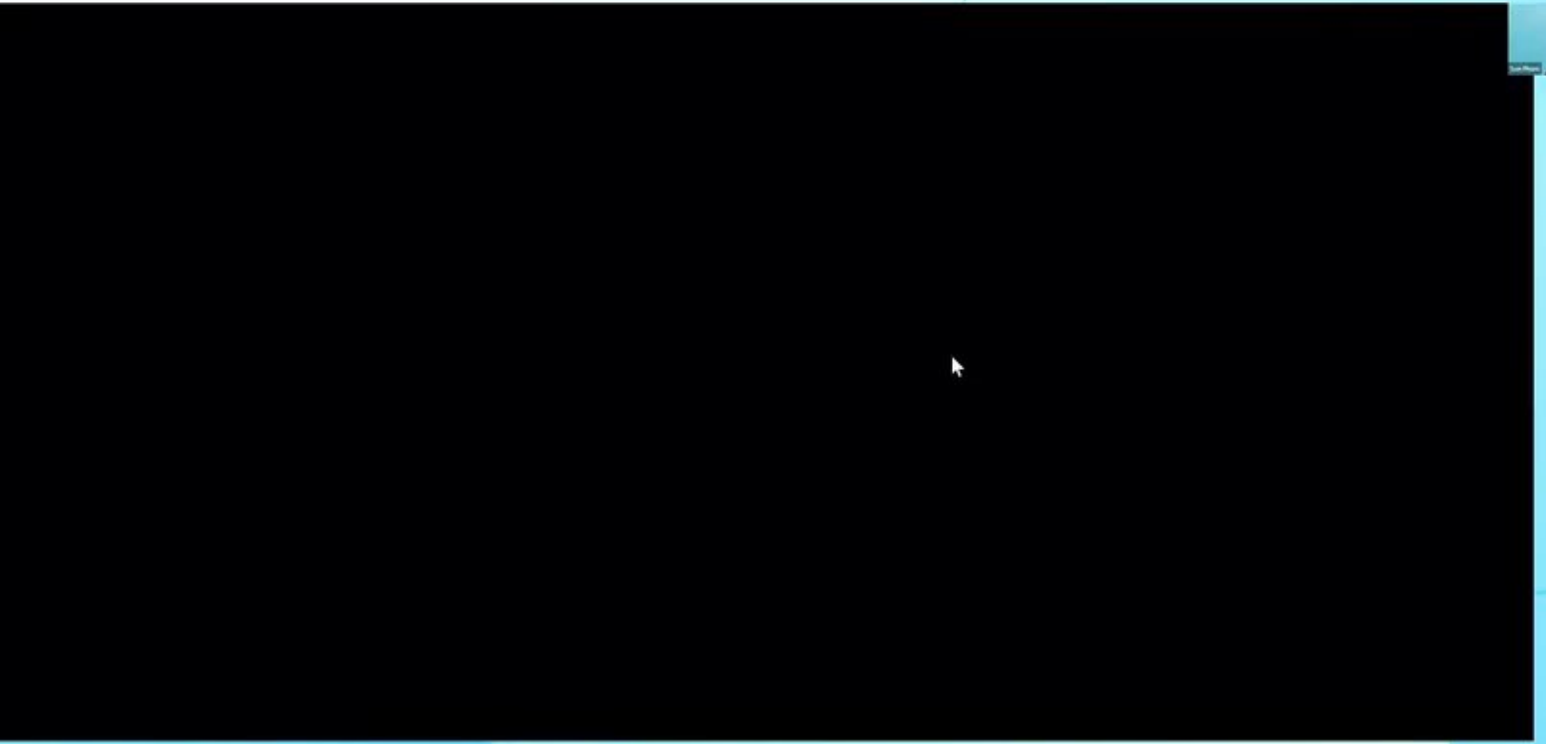
SPEAKER

RAHUL SAT

Core Member, New York Genome Center;

Associate Professor

www.bioturing.com



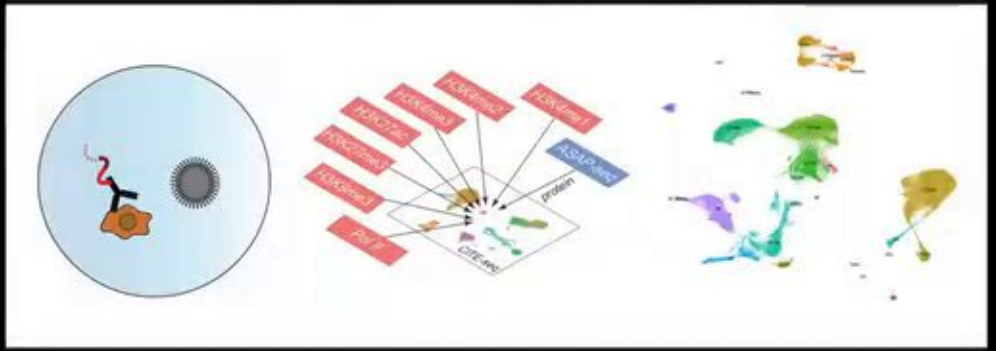
Dr. Rahul Satija

Associate Professor of Biology, Genomics & Systems Biology, NYU

Core Faculty Member, NYGC

Associate Faculty, Institute for Systems Genetics,
NYU Langone Medical Center

Integrative analysis of single-cell data across modalities and technologies



Rahul Satija
Core Faculty Member, New York Genome Center
Associate Professor, Biology, New York University

@satijalab

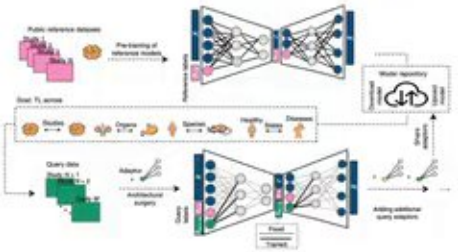
Reference 'maps' for single-cell genomics

nature biotechnology ANALYSIS

OPEN

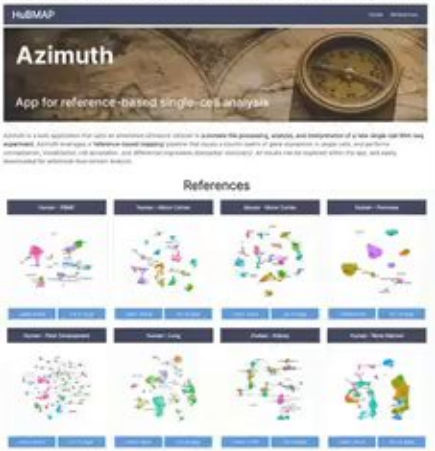
Mapping single-cell data to reference atlases by transfer learning

Mohammad Lotfollahi, Mohsen Naghipourfar, Malte D. Luecken, Matin Khajavi, Maren Büttner, Marco Wagenstetter, Žiga Avsec, Adam Gayoso, Nir Yosef, Marta Interlandi, Sergei Rybakov, Alexander V. Misharin and Fabian J. Theis



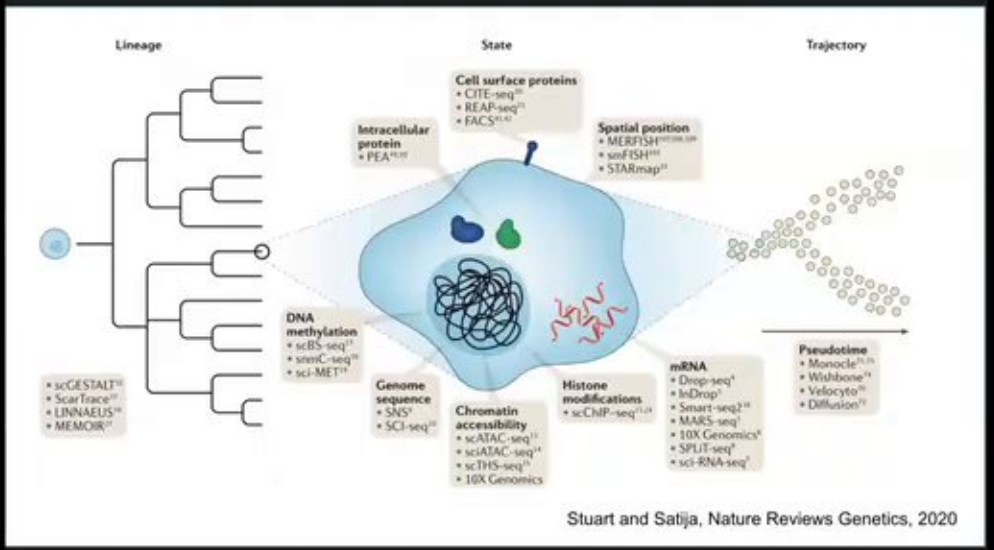
Lotfollahi et al, NBT 2021 (Theis Lab)

azimuth.hubmapconsortium.org



Hao et al, Cell 2021 (Satija Lab)

Multimodal cellular analysis



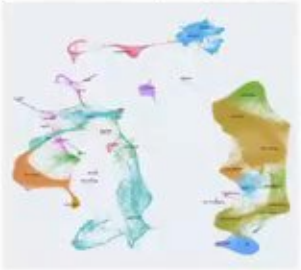
Stuart and Satija, Nature Reviews Genetics, 2020

Goal: fill in the missing gaps (modalities) for cell types in the human body

Bridge Integration via Dictionary Learning

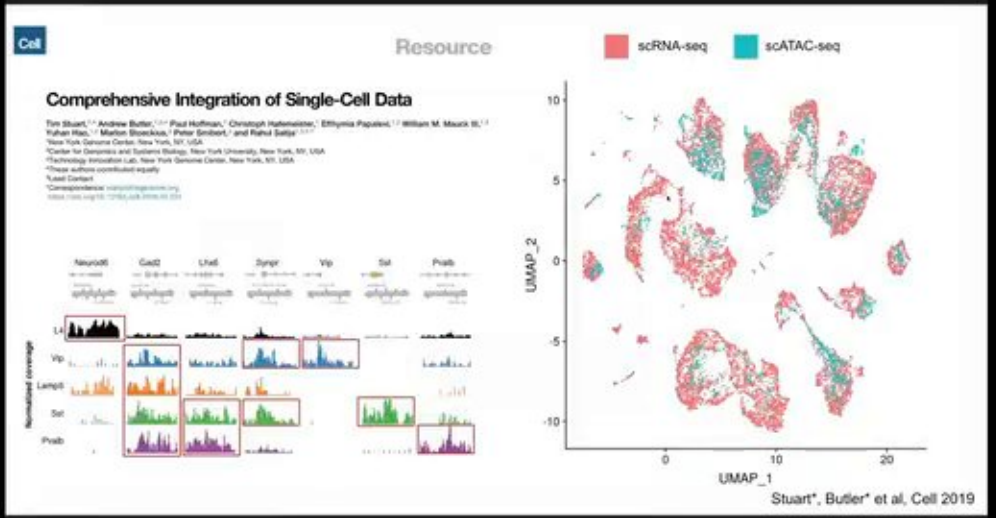
- How can we map datasets when **they measure different things**
- Example: mapping scATAC-seq (peaks) to scRNA-seq (genes)

Azimuth scRNA reference
(bone marrow; 320,000 cells)



ATAC-seq query dataset
(bone marrow; 35,000 cells)

Data Integration across Modalities



Integration of ATAC-seq using ‘gene activity scores’

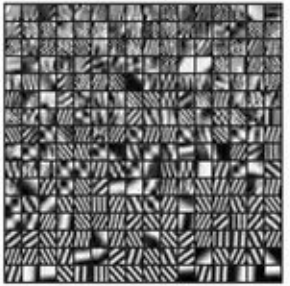
Dictionary Learning

Dictionary learning develops a set (dictionary) of representative elements from the input

Each data point can be represented as a weighted sum of the dictionary elements.

Noisy image

'Dictionary elements'



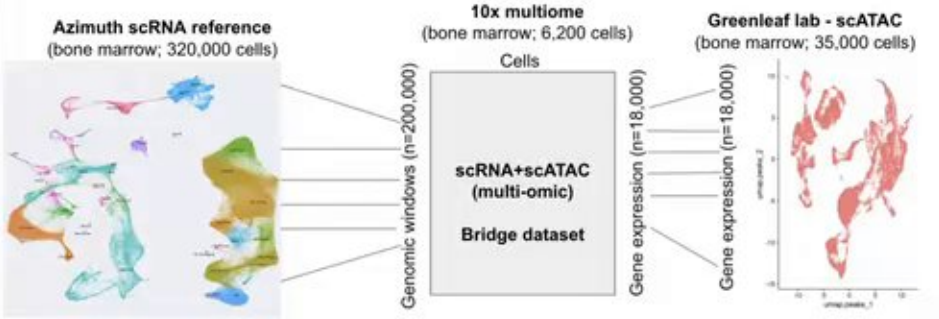
Denoised image

Elad and Iharon, IEEE Transactions on Image Processing, 2006

Dictionary elements can be learned 'within' or 'across' datasets

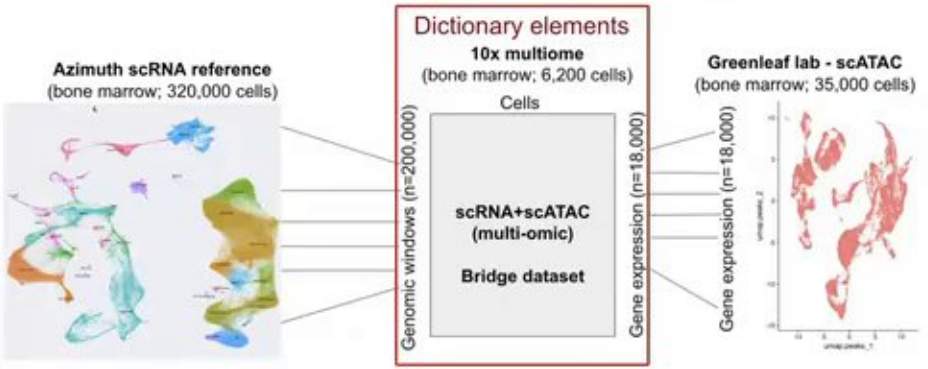
Bridge Integration via Dictionary Learning

Idea: Treat 10x multi-ome dataset (scRNA+scATAC) as a 'bridge'



Bridge Integration via Dictionary Learning

We represent **both** scRNA-seq and scATAC-seq datasets as linear combinations of multi-omic cell profiles (dictionary elements)

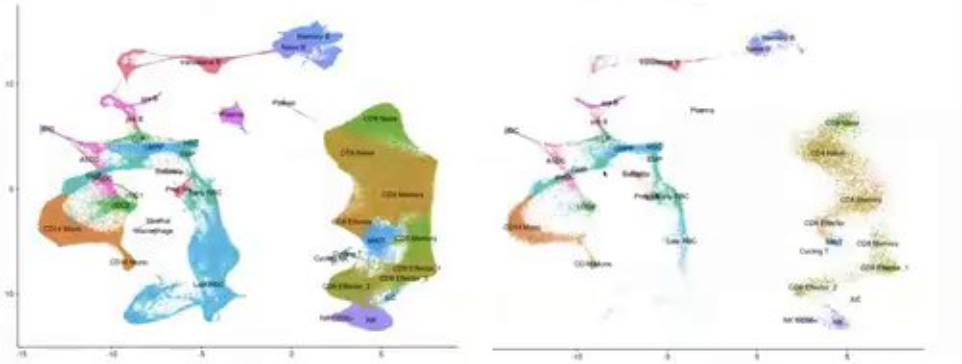


Dictionary learning maps different features into a shared space

Integrating scATAC-seq + scRNA-seq

Azimuth scRNA reference
(bone marrow; 320,000 cells)

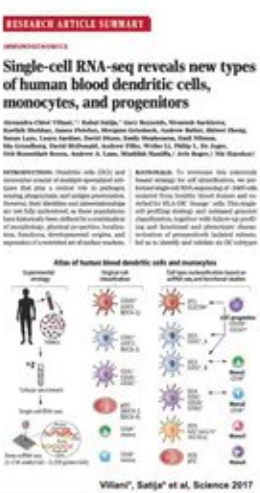
ATAC-seq query dataset
(bone marrow; 35,000 cells)



Hao et al, Nat Biotech 2024

Bridge integration enables accurate cross-modality integration

Integrating scATAC-seq + scRNA-seq

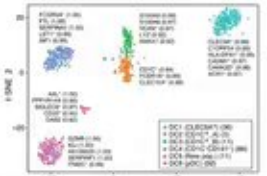

RESEARCH ARTICLE SUMMARY

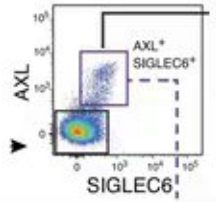
IMMUNOGENOMICS

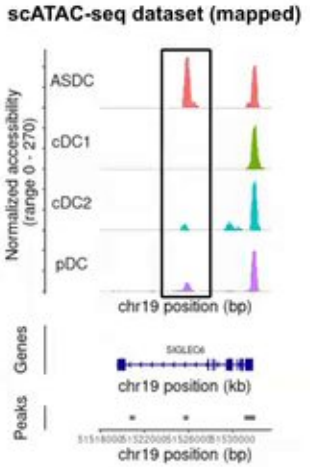
Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors

Atlas of human blood dendritic cells and monocytes

Villani*, Satija* et al, Science 2017

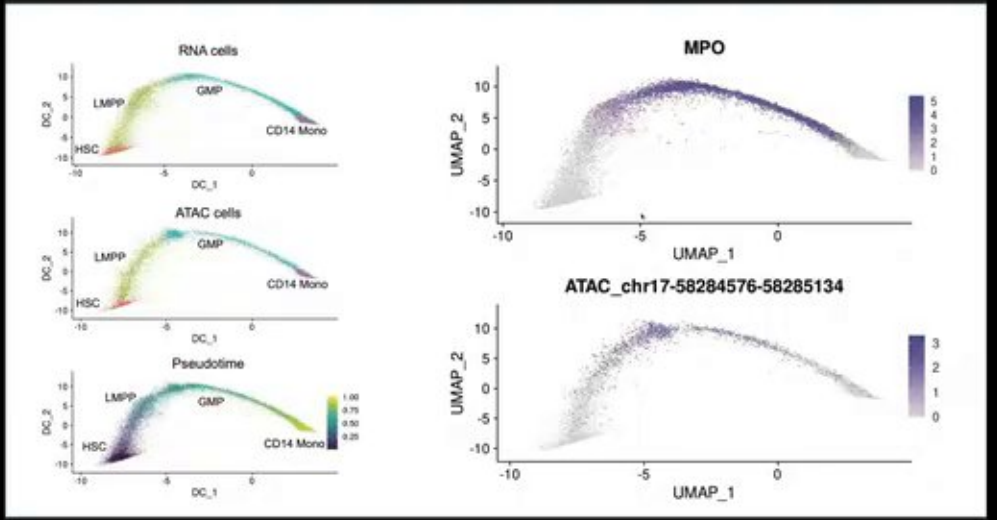






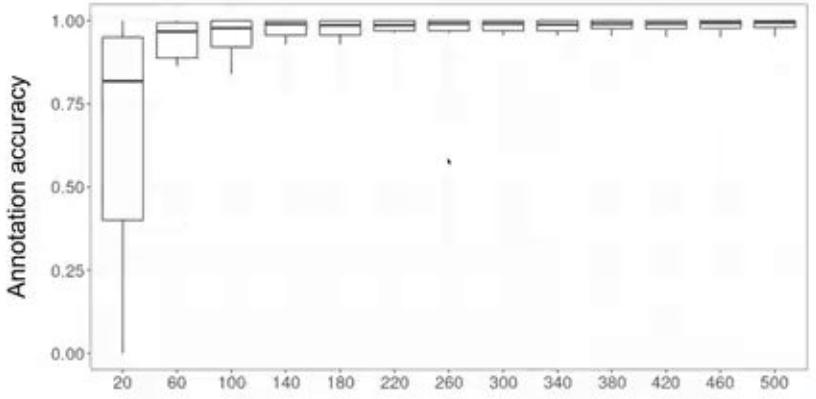
Bridge integration can map extremely rare groups (0.04%)

Bridge Integration via Dictionary Learning



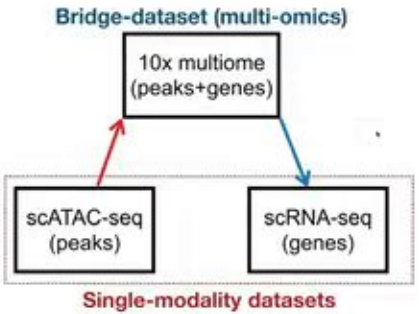
Identifying genes whose expression ‘lags’ behind chromatin state

Integrating scATAC-seq + scRNA-seq



Accurate annotations when the bridge dataset is ‘sufficient’

Integrating scATAC-seq + scRNA-seq

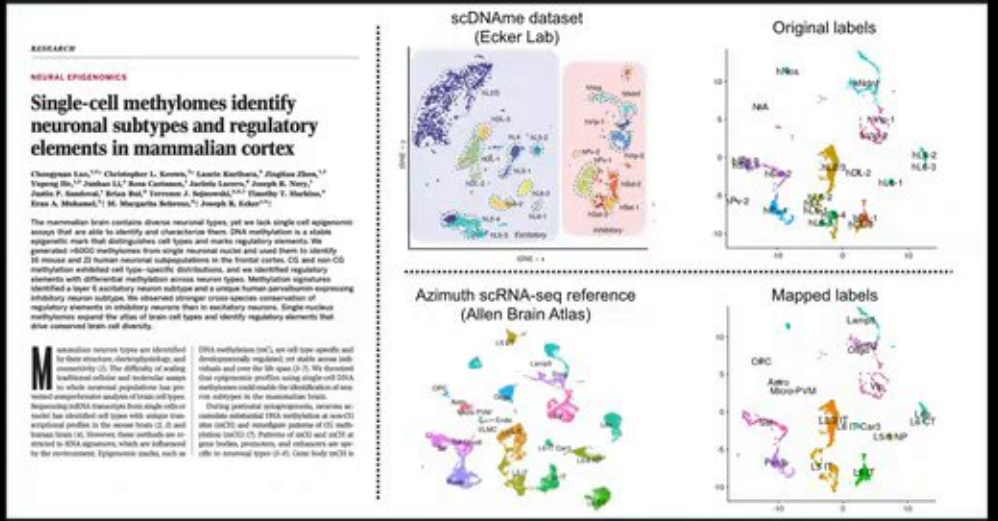


Example bridge technologies

Modality	Bridge Technology
Chromatin accessibility	10x Genomics 'multiome' kit (ATAC+RNA)
Chromatin modifications	scCUT&Tag-pro[30] (chromatin mods+protein)
DNA methylation	snmCT-seq[31] (methylation + RNA)
Genome organization	scMethyl-HiC[32] (HiC+methylation)
Protein levels	CITE-seq[33], DOGMA-seq[344] (RNA+ATAC+protein)
Lipid metabolism	scLipid[35] (MALDI-MSI + RNA)

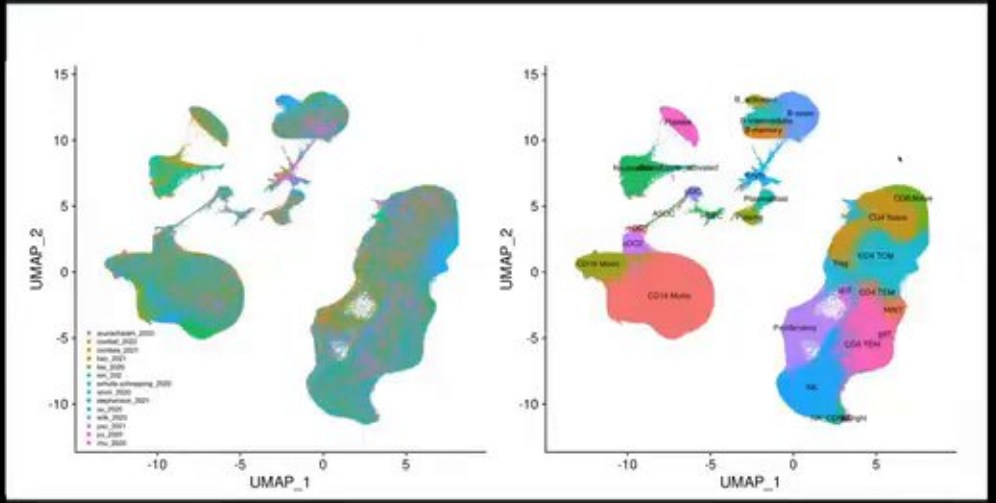
Bridge integration can be broadly applied to multiple modalities

Integrating scDNAme + scRNA-seq



Bridge integration can be broadly applied to multiple modalities

An ‘atlas’ of PBMC from COVID patients



Integrated dataset of 3.46M human PBMC (19 studies; 693 individuals)

Bridge Integration FAQ

Question: What are the requirements of the bridge dataset?

- Needs to be representative (20-50 cells per cell type/state)
- Can be (substantially) lower quality than reference or query

Question: What if cell types are missing in the bridge datasets

- Dictionary representation detects high reconstruction error
- Confidence scores for mapping sharply decrease

Question: Why not always process with multi-omic technologies?

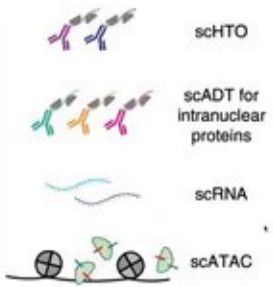
- Compatibility, data quality, and throughput
- Bridge integration enables more flexible experimental design

Multimodal profiling technologies

NEAT-seq

snATAC + snRNA + intra-nuclear ADT

10x multiome kit

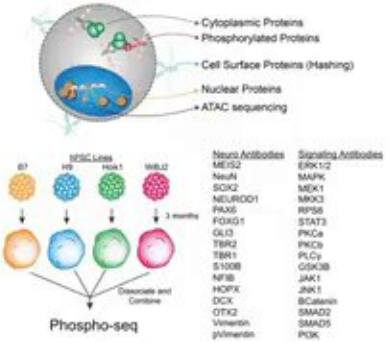


Chen et al, Nature Methods

Phospho-seq

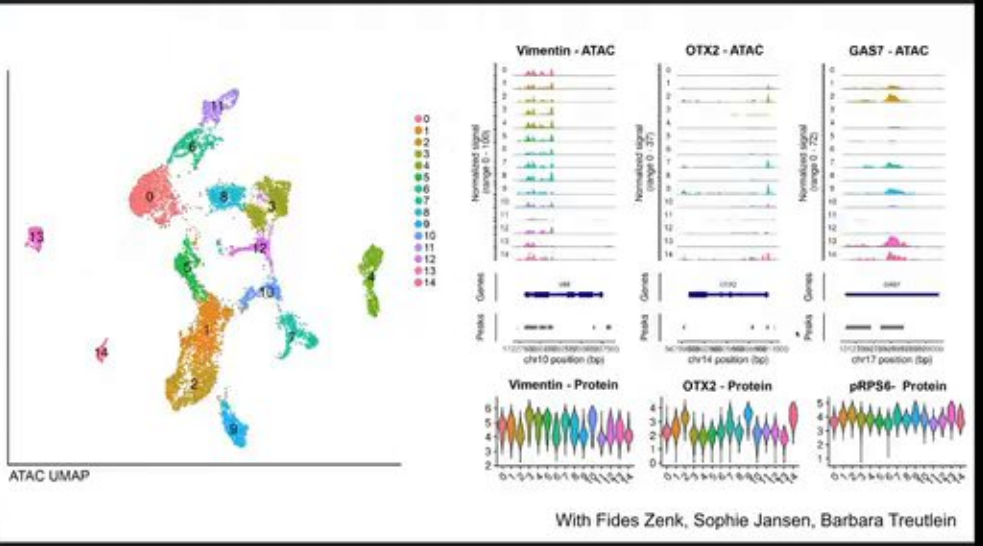
scRNA-seq + intra-cellular/nuclear ADT

10x scATAC-seq kit



Blair et al, In Revision

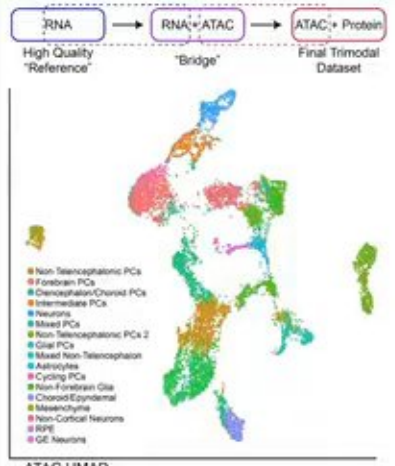
Intracellular protein profiling + scATAC

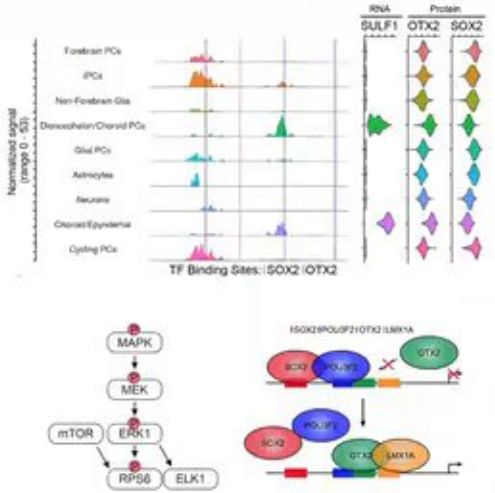


With Fides Zenk, Sophie Jansen, Barbara Treutlein

Phospho-seq datasets from human cerebral organoids

Intracellular protein profiling + scATAC





Bridging provides high-quality, integrated, multimodal profiles

COVID vaccine-induced T cell responses

Bingjie Zhang
Postdoctoral Fellow
Satija / Littman Labs

Rabi Upadhyay
Heme/Onc Fellow
Littman Lab

Dan Littman
Professor
NYU SoM / HHMI

Mark Mulligan
Professor
NYU Vaccine Center

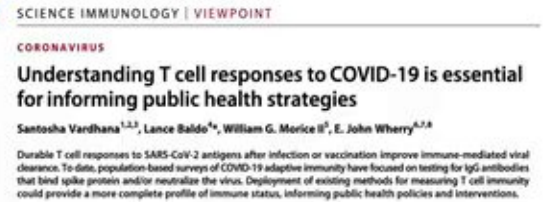
Zhang*, Upadhyay* et al, Nat Immunology 2024

COVID vaccine-induced T cell responses


NIH RESEARCH MATTERS

T cells protect against COVID-19 in absence of antibody response

At a Glance


SCIENCE IMMUNOLOGY | VIEWPOINT

CORONAVIRUS

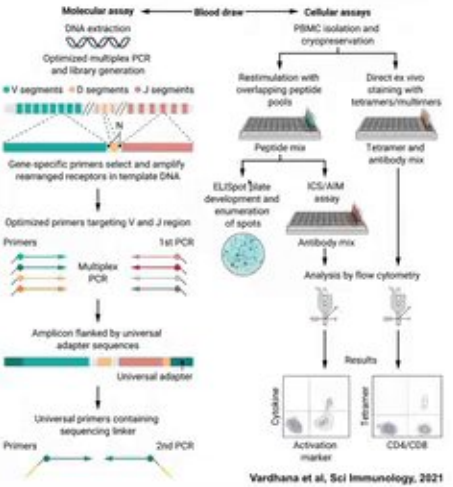
Understanding T cell responses to COVID-19 is essential for informing public health strategies

Santosha Vardhana[1,2,3], Lance Baldo[4]*, William G. Morice II[5], E. John Wherry[6,7,8]

Durable T cell responses to SARS-CoV-2 antigens after infection or vaccination improve immune-mediated viral clearance. To date, population-based surveys of COVID-19 adaptive immunity have focused on testing for IgG antibodies that bind spike protein and/or neutralize the virus. Deployment of existing methods for measuring T cell immunity could provide a more complete profile of immune status, informing public health policies and interventions.

1. **T cell contributions are required for early and durable SARS-Cov-2 protection**
2. **T cell responses are linked to disease severity and progression**
3. **New technologies are needed to identify, characterize, and track T cell responses during infection and vaccination**

Assays for characterizing T cell responses



Cellular assays

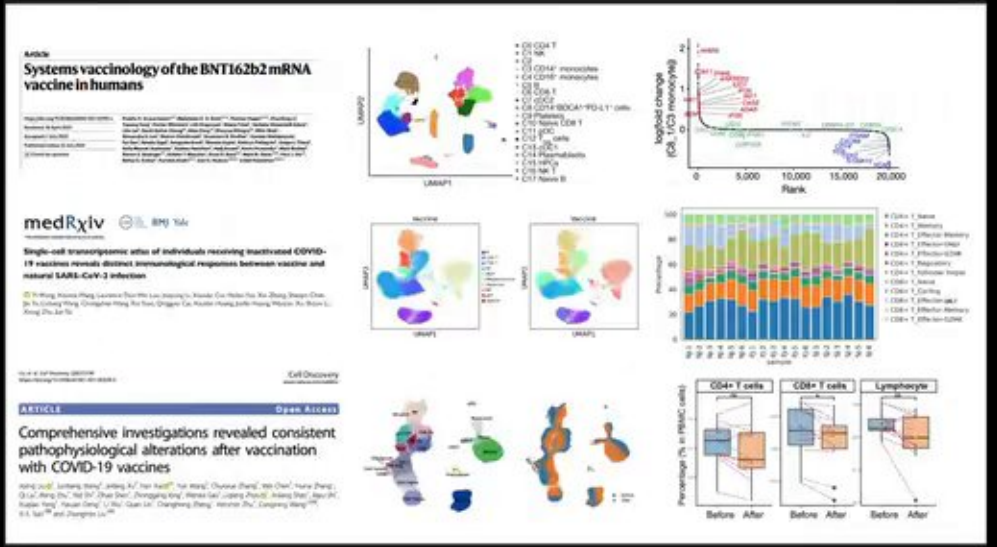
- Cytokine production after antigen staining
  - Requires in-vitro restimulation
  - Low-cost and broadly applicable
- Tetramer-based assays
  - Direct ex-vivo characterization
  - Restricted for individual haplotypes

Molecular assays

- Immune repertoire analysis
  - Use of NGS provides flexibility when analyzing new variants

Existing assays each have individual strengths and limitations

COVID vaccine-induced T cell responses



Previous scRNA-seq studies have not identified antigen-specific T cells

Characterizing molecular responses to BNT162b2

Initial dataset (6 individuals)

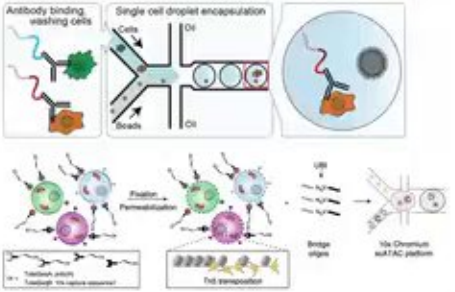
Modalities (initial dataset)

3' RNA (CITE-seq)

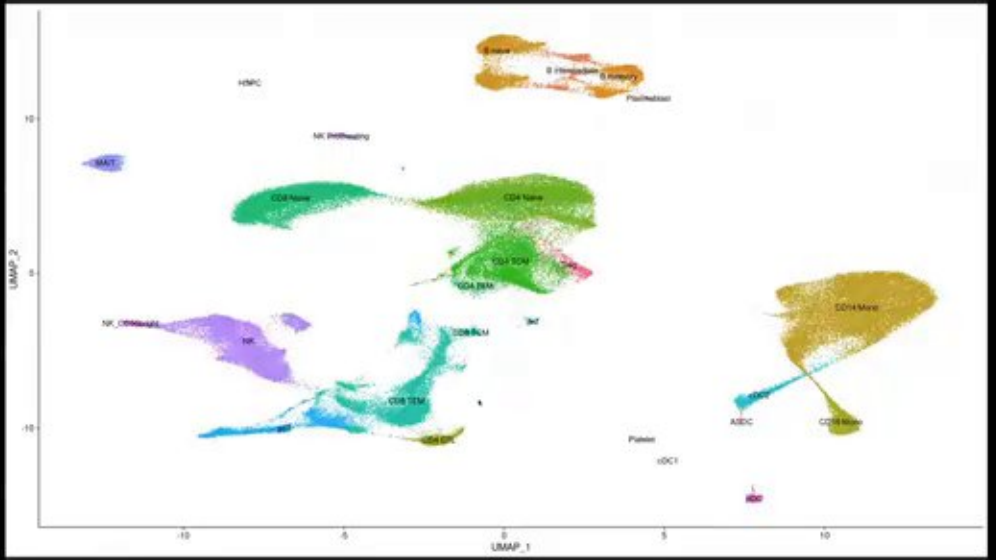
ATAC (ASAP-seq)

175 Surface protein (CITE/ASAP-seq)

Total: 195,790 single cells

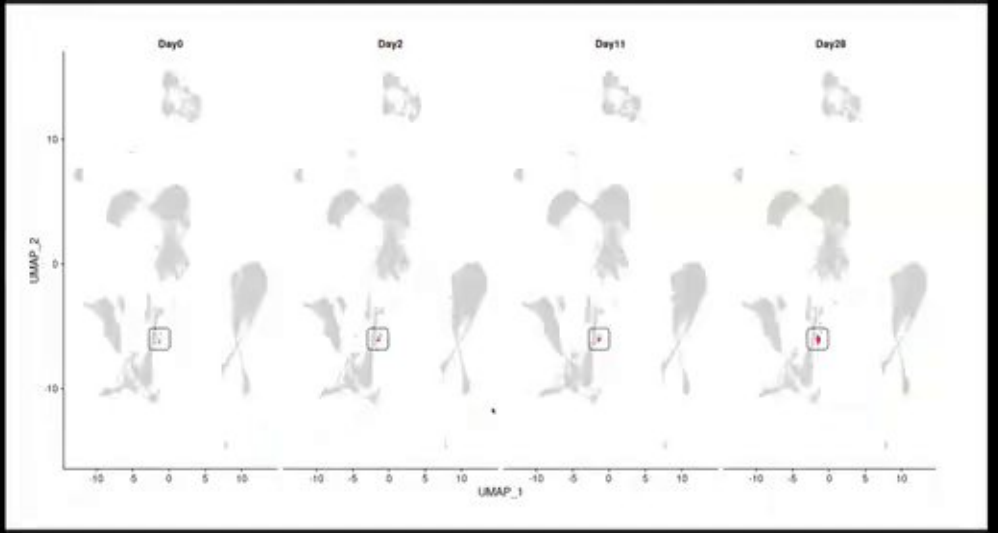


Characterizing molecular responses to BNT162b2



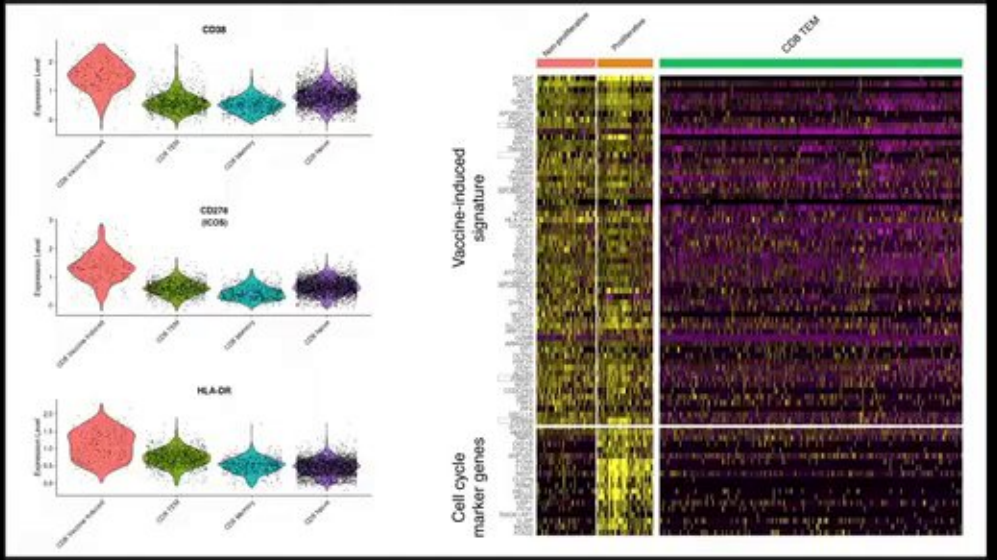
CITE-seq data reveals high-resolution immune subpopulations

Vaccine-induced CD8 T cells



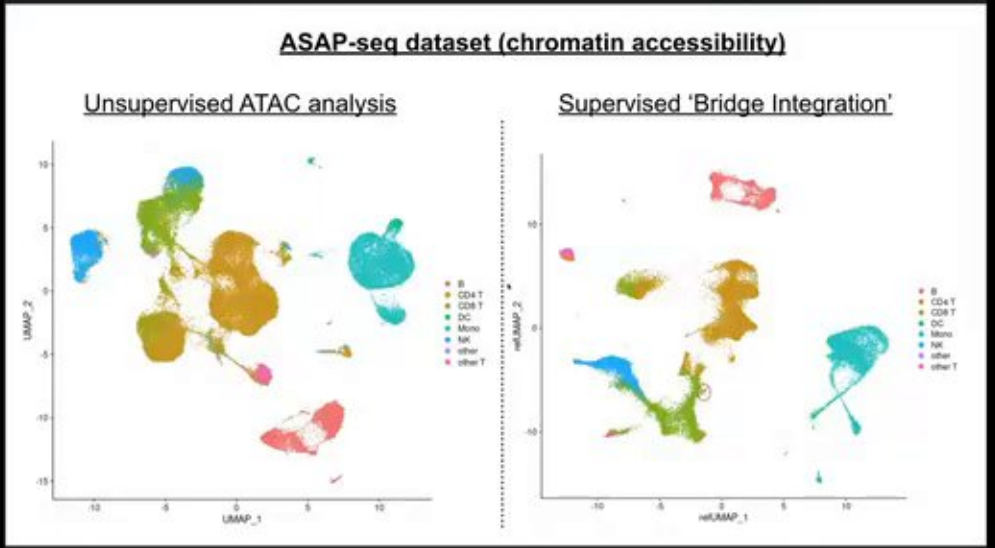
CITE-seq data reveals high-resolution immune subpopulations

Vaccine-induced CD8 T cells



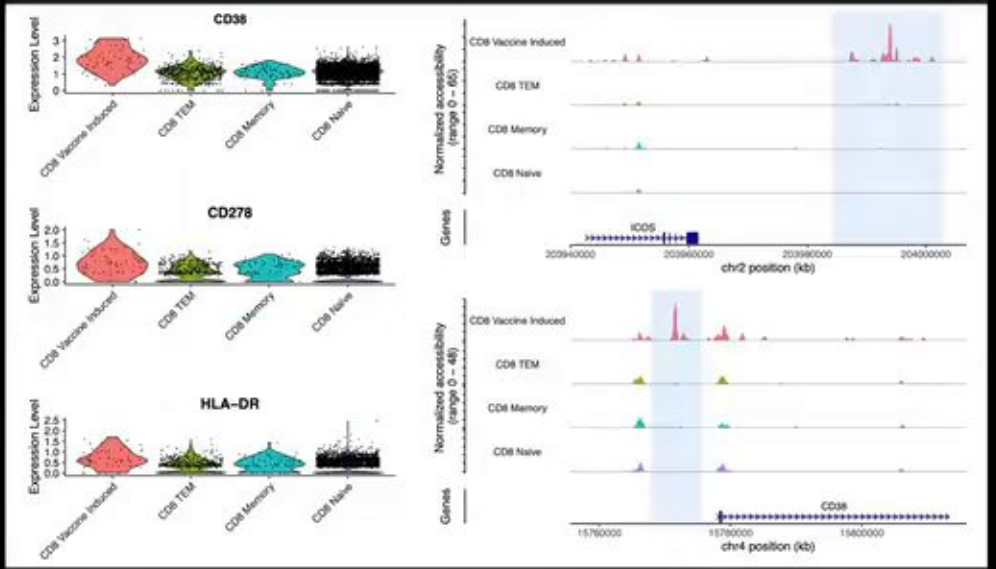
RNA and protein molecular signatures of vaccine-induced CD8 cells

Vaccine-induced CD8 T cells



Bridge integration maps integrated CITE-seq and scATAC-seq data

COVID vaccine-induced T cell responses



ASAP-seq reveals vaccine-induced enhancers in CD8 T cells

COVID vaccine-induced T cell responses

nature immunology
ARTICLES

BATF3 programs CD8⁺ T cell memory

Marco A. Ataide, Karl Komander, Konrad Knöpper, Annika E. Peters, Hao Wu, Sarah Eickhoff, Tea Gogishvili, Justus Weber, Anika Grafen, Axel Kallies, Natalio Garbi, Hermann Einsele, Michael Hudecek, Georg Gasteiger, Michael Hölzel, Martin Vaeth and Wolfgang Kastenmüller

Antiviral CD8⁺ T cell responses are characterized by an initial activation/priming of T lymphocytes followed by a massive proliferation, robust differentiation, population contraction and the development of a stable memory pool. The transcription factor BATF3 has been shown to play a central role in the development of conventional dendritic cells, which in turn are critical for optimal priming of CD8⁺ T cells. Here we show that BATF3 was expressed transiently within the first days after T cell priming and had long-lasting T cell-intrinsic effects. T cells that lacked *Batf3* showed normal expansion and differentiation, yet succumbed to an aggravated contraction and had a diminished memory response. Vice versa, BATF3 overexpression in CD8⁺ T cells promoted their survival and transition to memory. Mechanistically, BATF3 regulated T cell apoptosis and longevity via the proapoptotic factor BIM. By programing CD8⁺ T cell survival and memory, BATF3 is a promising molecule to optimize adoptive T cell therapy in patients.

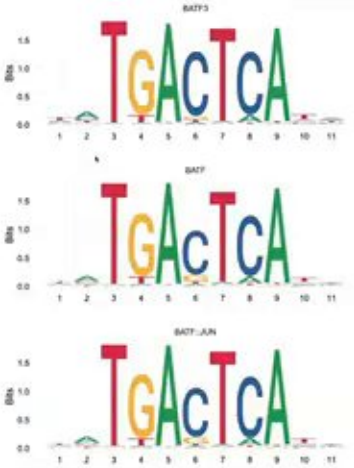
Upon encounter with dendritic cells (DCs) in the context of infections, naive antigen-specific CD8⁺ T cells undergo initial activation (priming) and consecutively, a massive expansion to generate the required effectors to target (also to effectively eliminate infected cells in the body. At the peak of the response, expanded CD8⁺ T cells comprise a spectrum of differentiation states ranging from short-lived effector cells to long-lived memory cells[1]. Notably, the quality of the developing memory T cell response is already determined during the first days of activation, among other factors by cognate CD8⁺ T cells[2], which transmit their signals via a subset of conventional DCs (cDC)[3–5]. Revealing the underlying programs and the transcriptional elements that promote CD8⁺ T cell fitness, functionality and memory, ultimately determines the efficacy of successful immunotherapy in humans.

The family of AP-1 (activator protein-1) transcription factors comprises numerous basic region-leucine zipper (bZIP) proteins that belong to different families (JUN, FOS, ATF, basic leucine zipper transcriptional factor ATF-like (BATF) and MAF) and form heterodimers to execute transcriptional activity. AP-1 transcription factors play a central role in every tissue and cell type and regulate cell death, survival and proliferation[6]. The overall regulation of AP-1 activity and its molecular elements is highly complex and involves transcriptional co-activators, modulation by MAP kinases and transcriptional inhibitors such as the BATF family of proteins[7]. BATF proteins play important roles in the development and function of both myeloid and lymphoid cell populations[8]. They act as negative regulators of the AP-1 complex, but also promote transcriptional activity by interacting with interferon (IFN) regulatory factor (IRF) family members[9]. BATF3 was first identified in human (129 versus C57BL/6), the commensal communities and the infection history[10]. Thus, *Batf3*-deficient mice became a key model for elucidating the in vivo functions of cDC1.

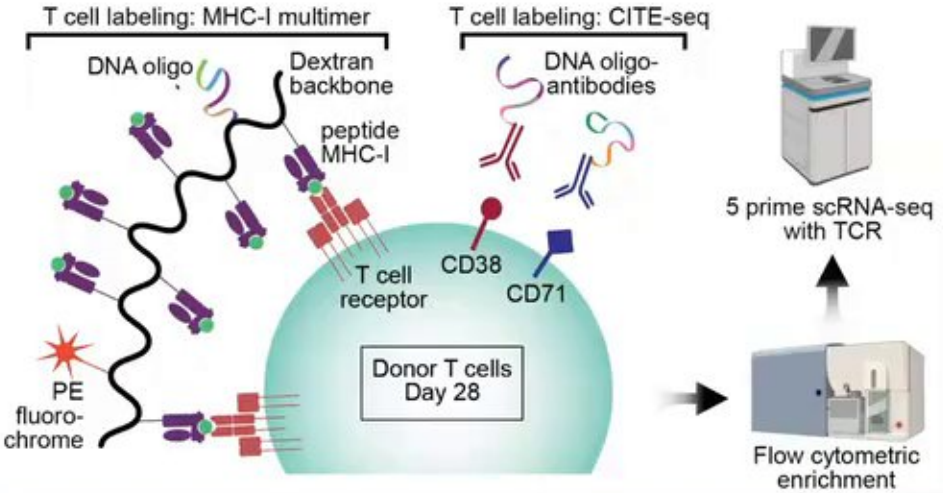
In murine lymphoid cells, *Batf3* rather than *Batf* plays a central role in their function, development and differentiation. In particular, BATF regulates various CD4⁺ helper T cell subsets such as interleukin (IL)-17-expressing helper T cells, follicular helper T cells and T regulatory type 1 cells[11–13]. In cytotoxic CD8⁺ T cells, BATF is required for their expansion and differentiation, but during chronic antigen exposure BATF also drives functional exhaustion[14–17]. Mechanistically, BATF functions within the AP-1 complex, interacts with cJUN and executes transcriptional activity together with IRF4, leading to reduced anabolic metabolism and a lowered mitochondrial membrane potential. Thereby, BATF regulates a large network of key transcription factors for CD8⁺ T cell differentiation, epigenetic modification, T cell metabolism and survival[14,15].

In this study we identified a specific and CD8⁺ T cell-intrinsic role for BATF3. In contrast to *Batf* deficiency, loss of *Batf3* did not impact CD8⁺ T cell differentiation or their capacity to produce inflammatory cytokines. Instead, *Batf3*-deficient CD8⁺ T cells showed a specific defect in T cell memory formation following viral or bacterial infections. The underlying molecular mechanism was due to aggravated apoptosis caused by a dysregulated expression of the proapoptotic factor BIM (BCL2L11) in *Batf3*-deficient T cells. Conversely, overexpression of BATF3 augmented CD8⁺ T cell metabolic fitness, inhibited cell death via suppression of *Bim* and promoted their memory development. By increasing the lifespan of both murine and human CD8⁺ T cells, BATF3 is a promising candidate to enhance adoptive immunotherapy against cancer and infections.

Motifs differentiating vaccine-induced populations

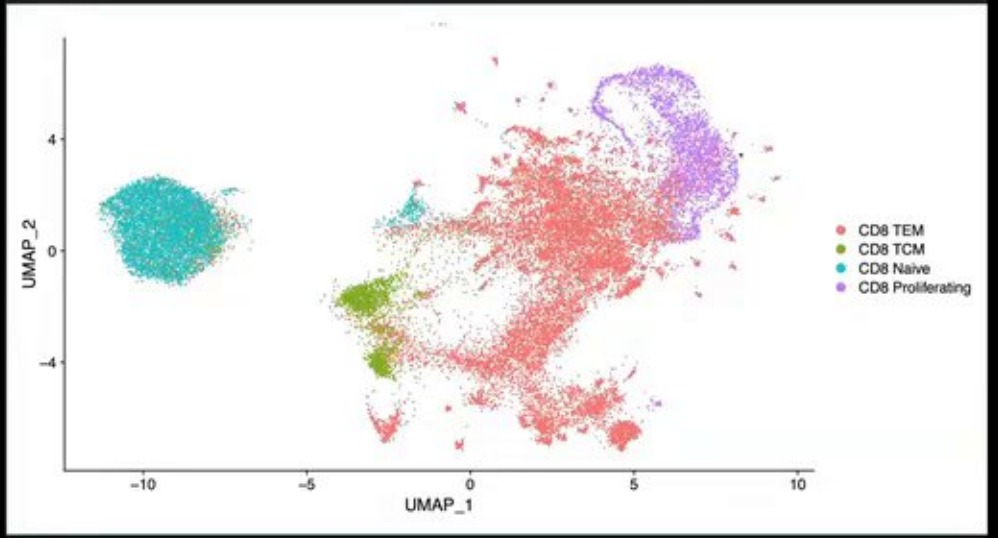


Antigen-specific or passenger response?



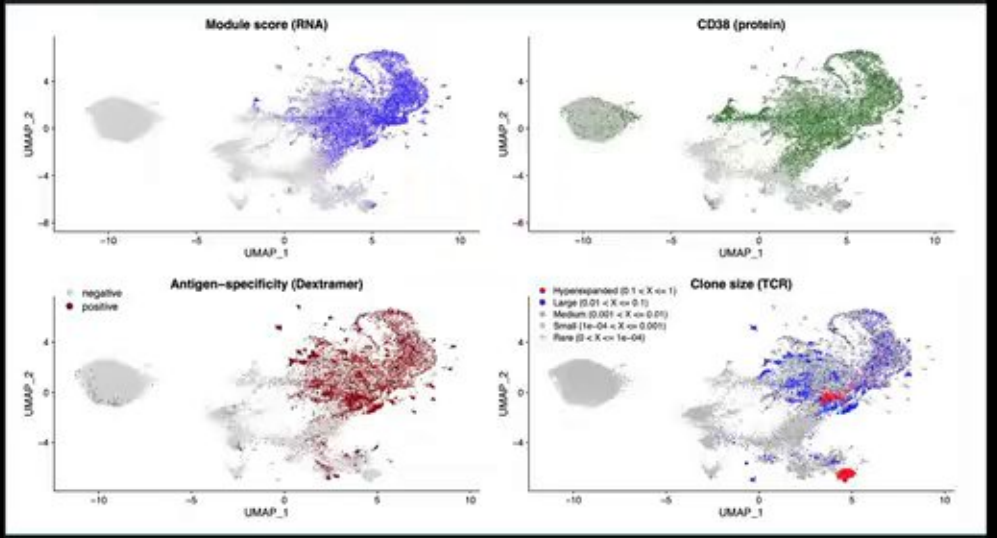
ECCITE-seq can capture antigen specificity and immune repertoires

Antigen-specific or passenger response?



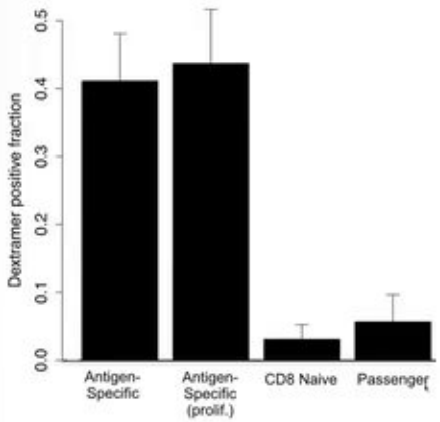
ECCITE-seq can capture antigen specificity and immune repertoires

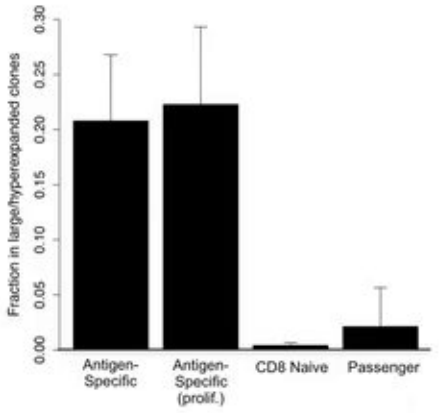
Antigen-specific or passenger response?



ECCITE-seq can capture antigen specificity and immune repertoires

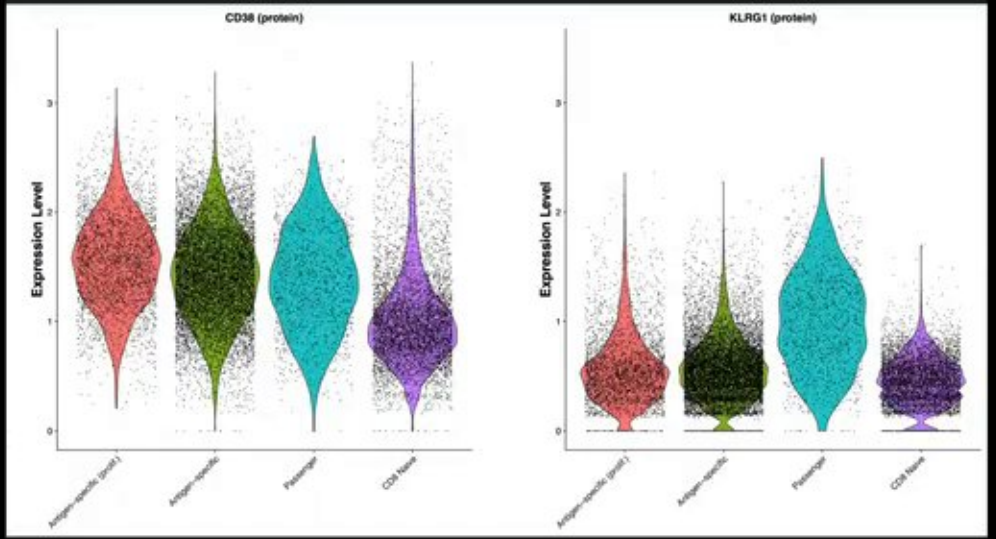
Antigen-specific or passenger response?





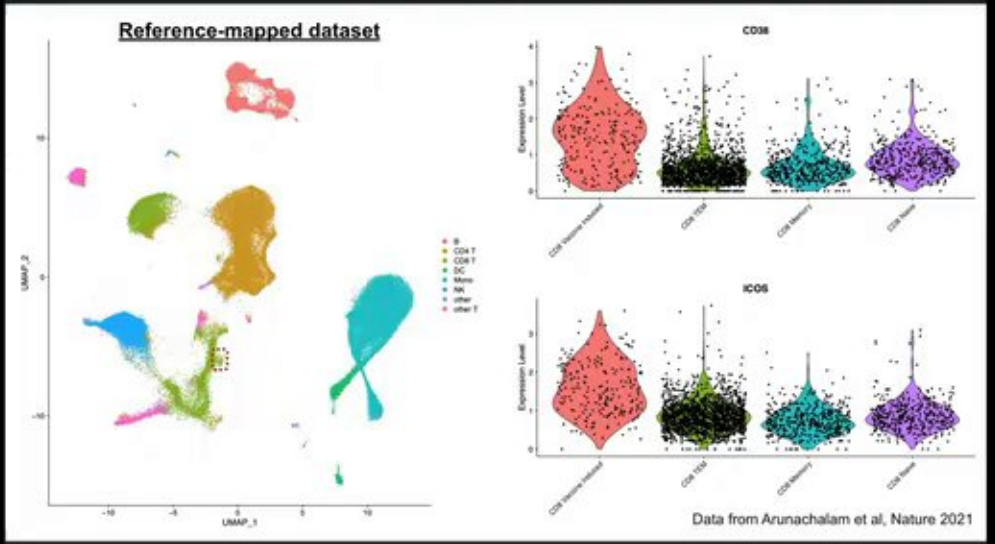
A subpopulation of CD38+ CD8+ T cells are not antigen specific

Antigen-specific or passenger response?



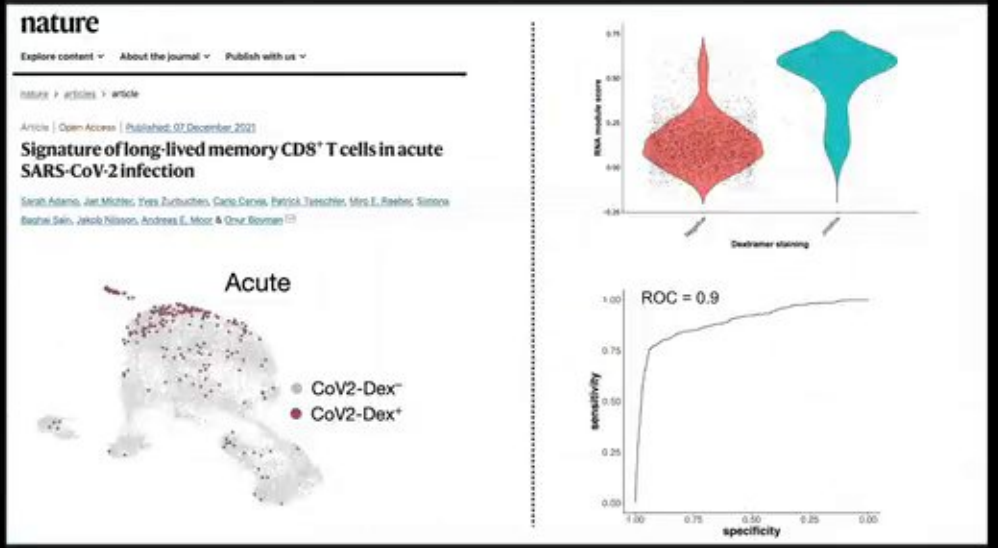
A subpopulation of CD38+ CD8+ T cells are not antigen specific

Identification of vaccine-induced cell populations



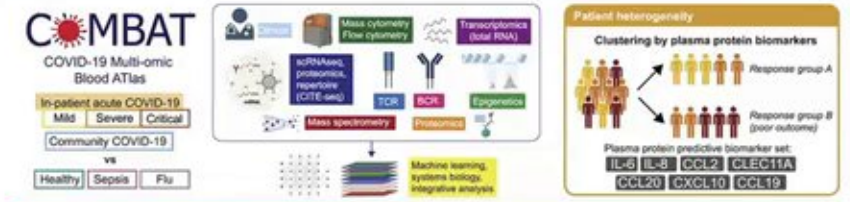
Vaccine-induced cells are reproducibly identified across datasets

Antigen-specific cells in COVID datasets



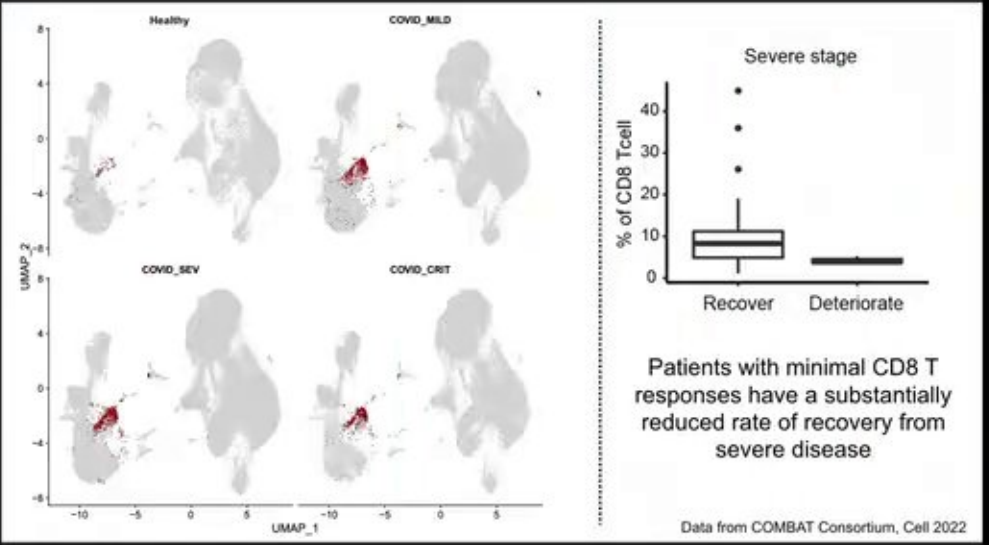
Module score accurately classifies antigen specificity during acute COVID

Antigen-specific cells in COVID datasets



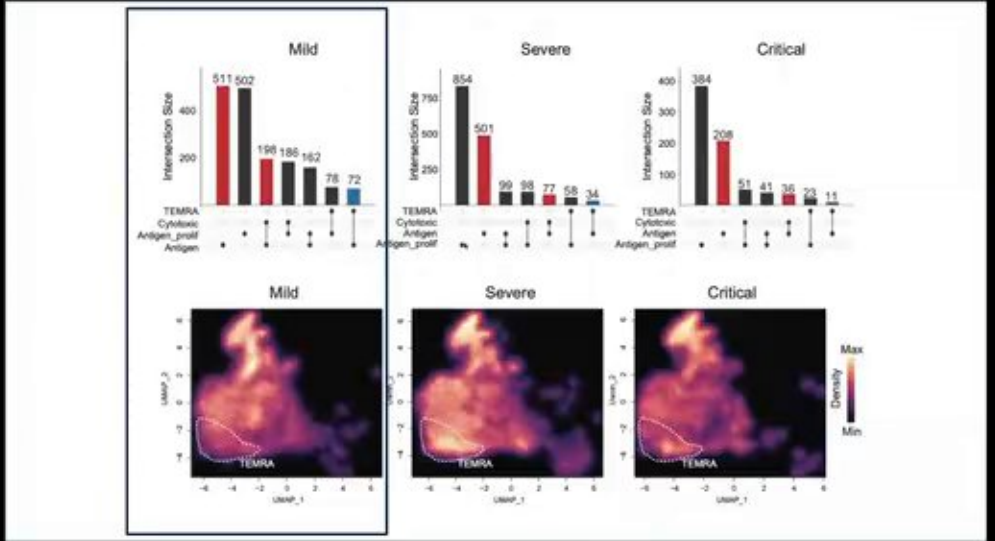
Data from COMBAT Consortium, Cell 2022

Antigen-specific cells in COVID datasets



CD8 T cell responses correlate with patient outcomes

Antigen-specific cells in COVID datasets



In mild disease, antigen-specific cells differentiate into cytotoxic states

Summary

Identification of vaccine-induced CD8 T cells in single-cell datasets

- Multimodal data and analysis is key for identification
- Clear separation of proliferating and non-proliferating populations
- Leads to identification of rich gene signature

Acknowledgements

Bingjie Zhang
Gesmira Molla
Jimmy Yun
Madeline Kowalski
Carol Dalgarno
Alex Bradu
Saket Choudhary
John Blair
Tracy Smith
Isabella Mascio
Nathan Nakatsuka
John Blair
Izzy Grabski
Sarah Curtiss


BioTuring