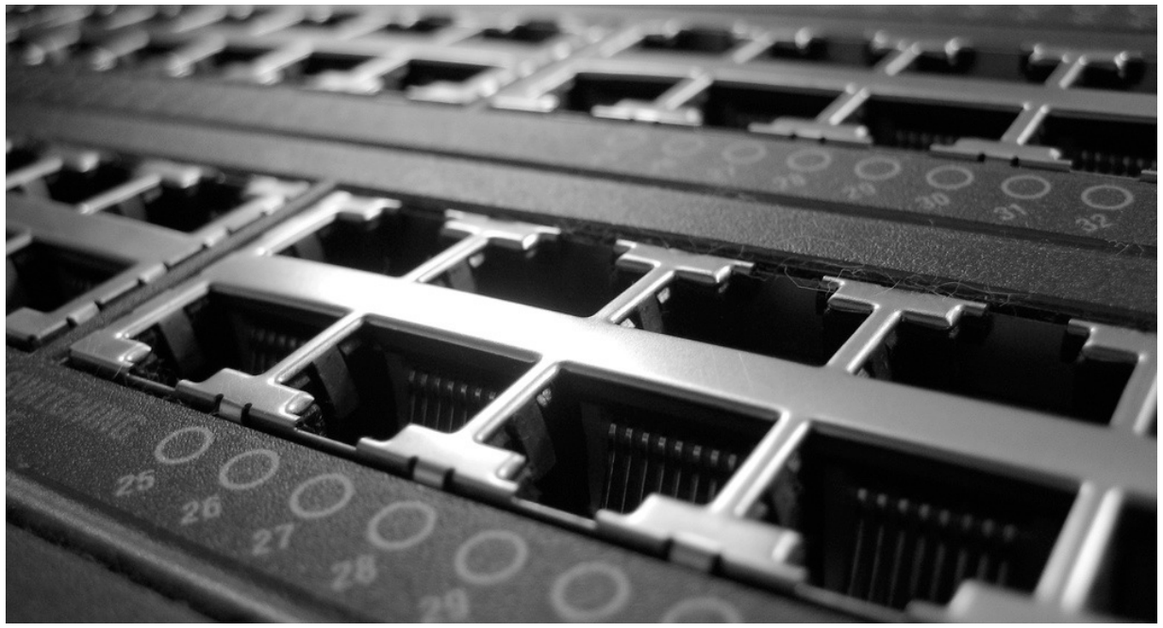


Supporting Single Cell RNA-seq Analysis: A Core's Perspective

Shannan Ho Sui

Harvard Chan Bioinformatics Core

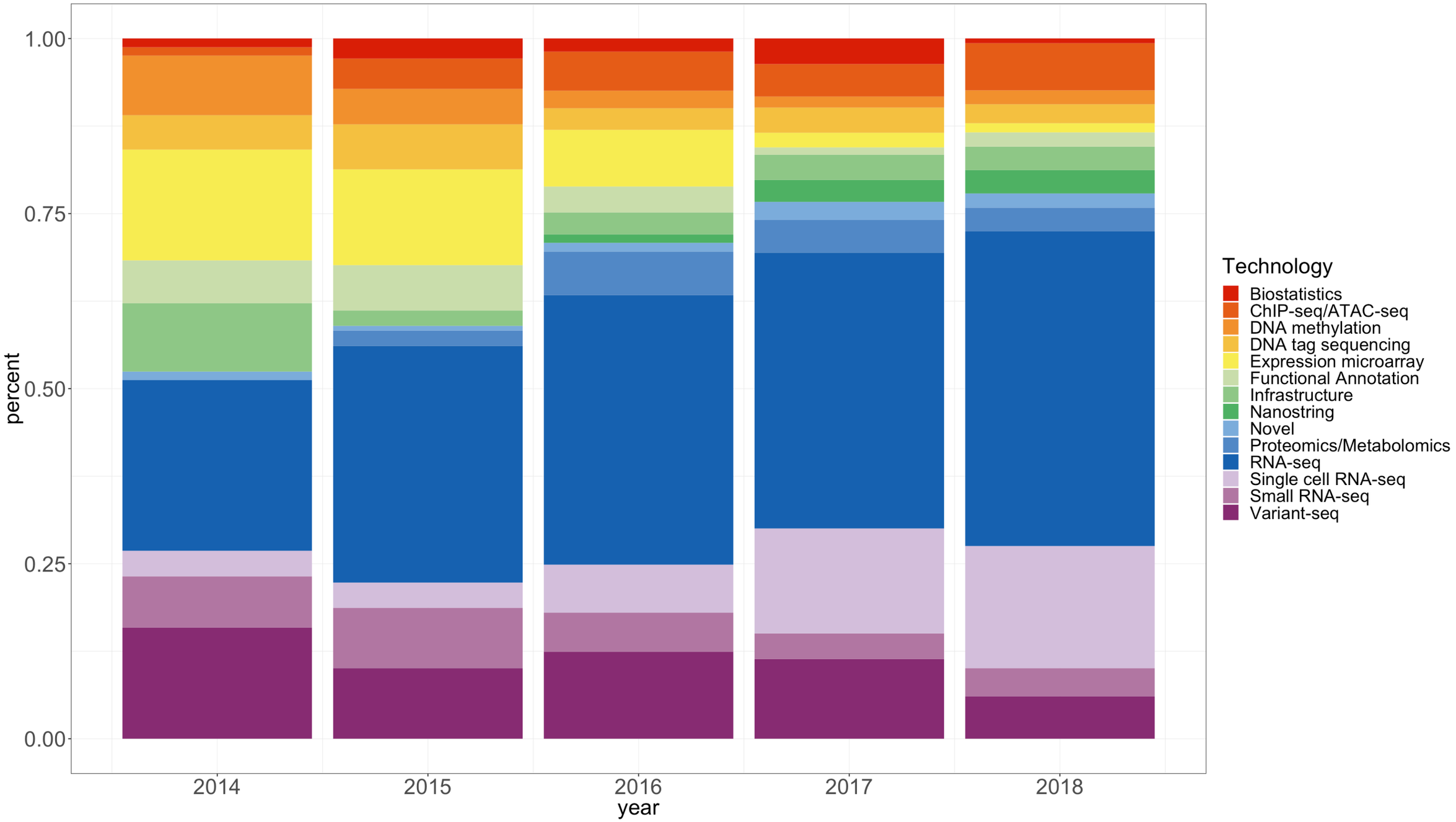
Webpage: http://bioinformatics.sph.harvard.edu

Email: bioinformatics@hsph.harvard.edu

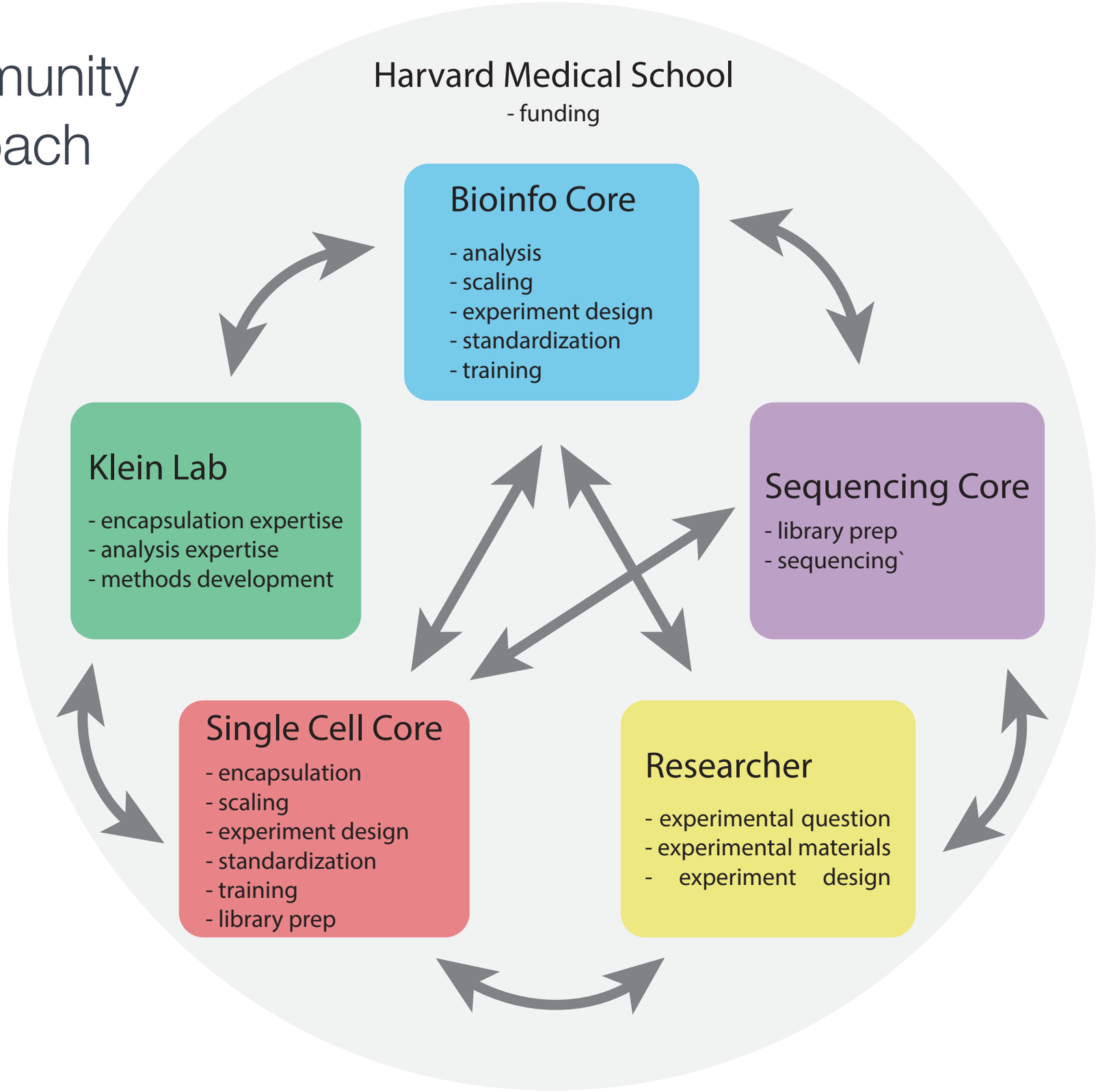
Common applications of scRNA-seq

- Explore which cell types are present in a tissue
- Identify unknown/rare cell types or states
- Elucidate the changes in gene expression during differentiation processes or across time or states
- Identify genes that are differentially expressed in particular cell types between conditions (e.g. treatment or disease)
- Explore changes in expression among a cell type while incorporating spatial, regulatory, and/or protein information

Growing demand for scRNA-seq support



A Community Approach



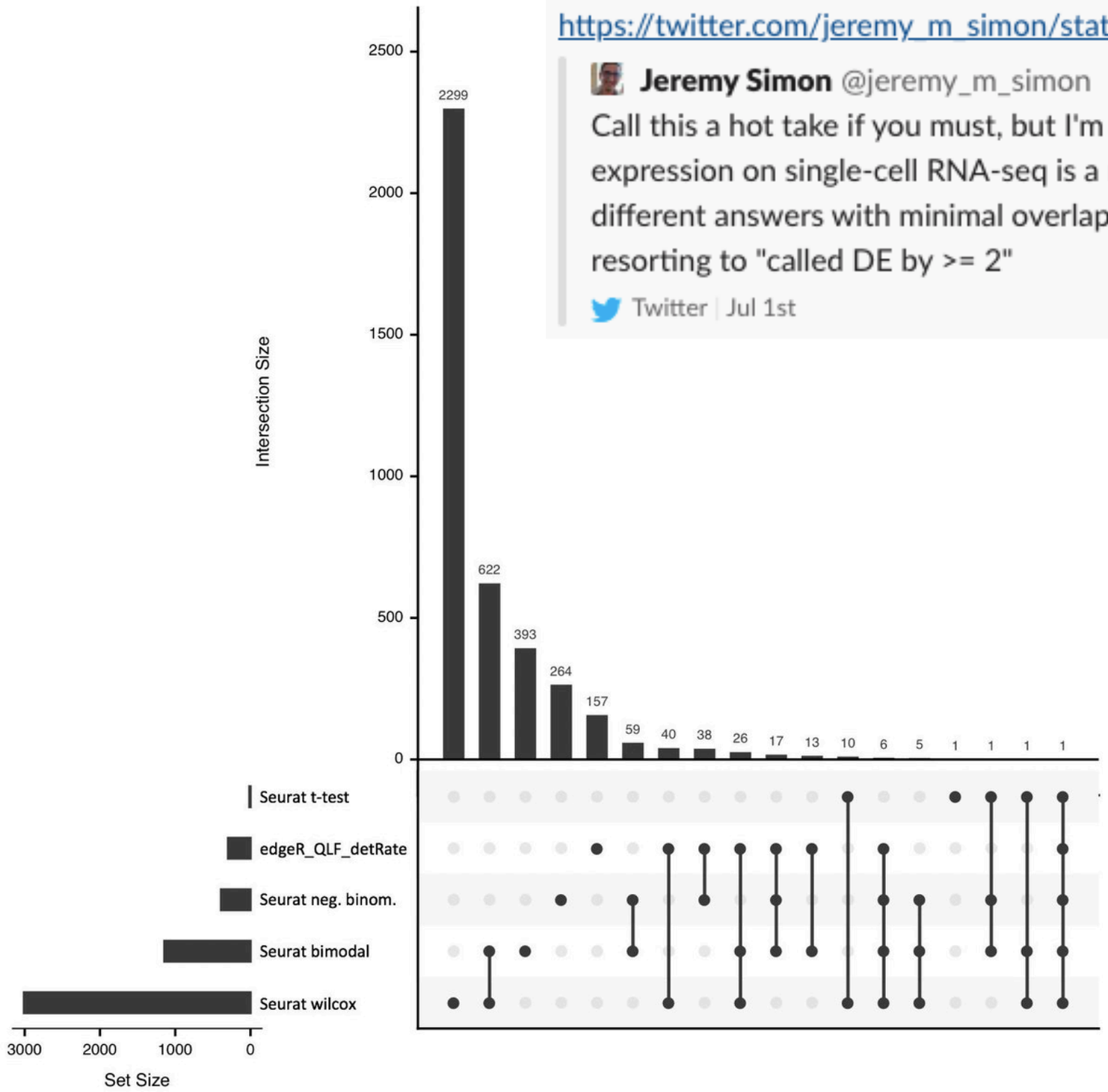
Standardizing using reproducible, scalable, validated best practice workflows

bcbio-nextgen

Python toolkit to automate best practice NGS pipelines

Challenges and Opportunities

- Complex designs - replicates, batches, technologies
- Close collaborations to allow for rapid, iterative analyses
- Rapidly emerging methods and evolving tools
  - Which ones to use?
  - Keeping versions consistent/synchronized (esp. among computing environments)
  - Different results from different methods
  - Lots of open questions

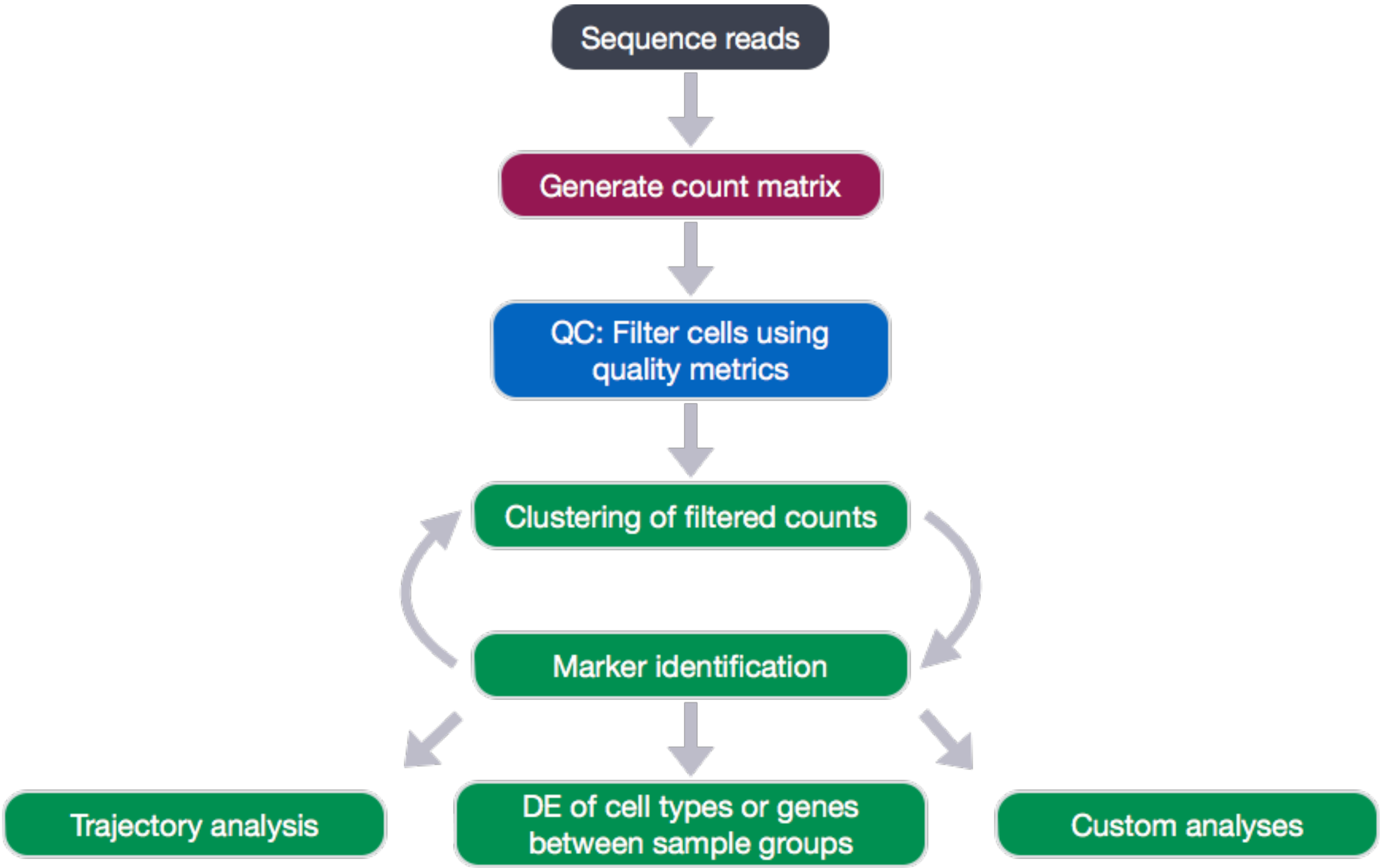


https://twitter.com/jeremy_m_simon/status/1145754750110990336?s=19

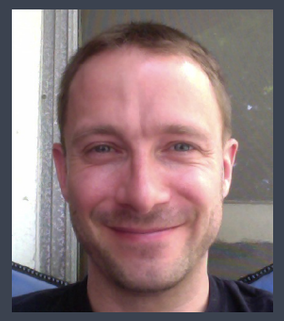
> **Jeremy Simon** @jeremy_m_simon
> Call this a hot take if you must, but I'm going on record and saying it: differential expression on single-cell RNA-seq is a mess. 10 different methods = 10 wildly different answers with minimal overlap. Hard to believe it's 2019 and we're likely resorting to "called DE by >= 2"
> Twitter | Jul 1st

Challenges and Opportunities

- Projects take longer to complete
- Practical approach to training
    - Internal training through retreats, development of materials, group discussions
    - Community training through our Bioinformatics Training Program


Sequence reads
Generate count matrix
QC: Filter cells using quality metrics
Clustering of filtered counts
Marker identification
Trajectory analysis
DE of cell types or genes between sample groups
Custom analyses

John Hutchinson
Associate Director



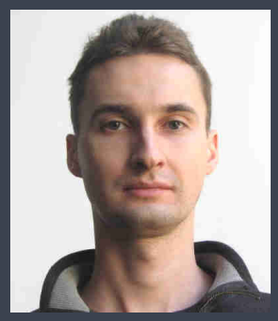
Rory Kirchner

Victor Barrera

Meeta Mistry

Mary Piper

Radhika Khetani
Training Director



Ilya Sytchev

Zhu Zhuo

Sergey Naumenko

James Billingsley

Peter Kraft
Faculty Advisor

Funding

Harvard Chan School of Public Health
Harvard Medical School
Harvard Stem Cell Institute
Harvard Catalyst
National Institute of Environmental Health Sciences (NIEHS)
Harvard University Center for AIDS Research (CFAR)
AstraZeneca
Boehringer Ingelheim