

tradeSeq: TRAjectory-based Differential Expression for SEQuencing data

Talk given at the retreat of
the Berkeley Stem Cell
Center on 04/14



Hector Roux de Bézieux
Group in Biostatistics
Sandrine Dudoit's lab

GitHub: HectorRDB

Website: <http://hectorrdb.github.io>

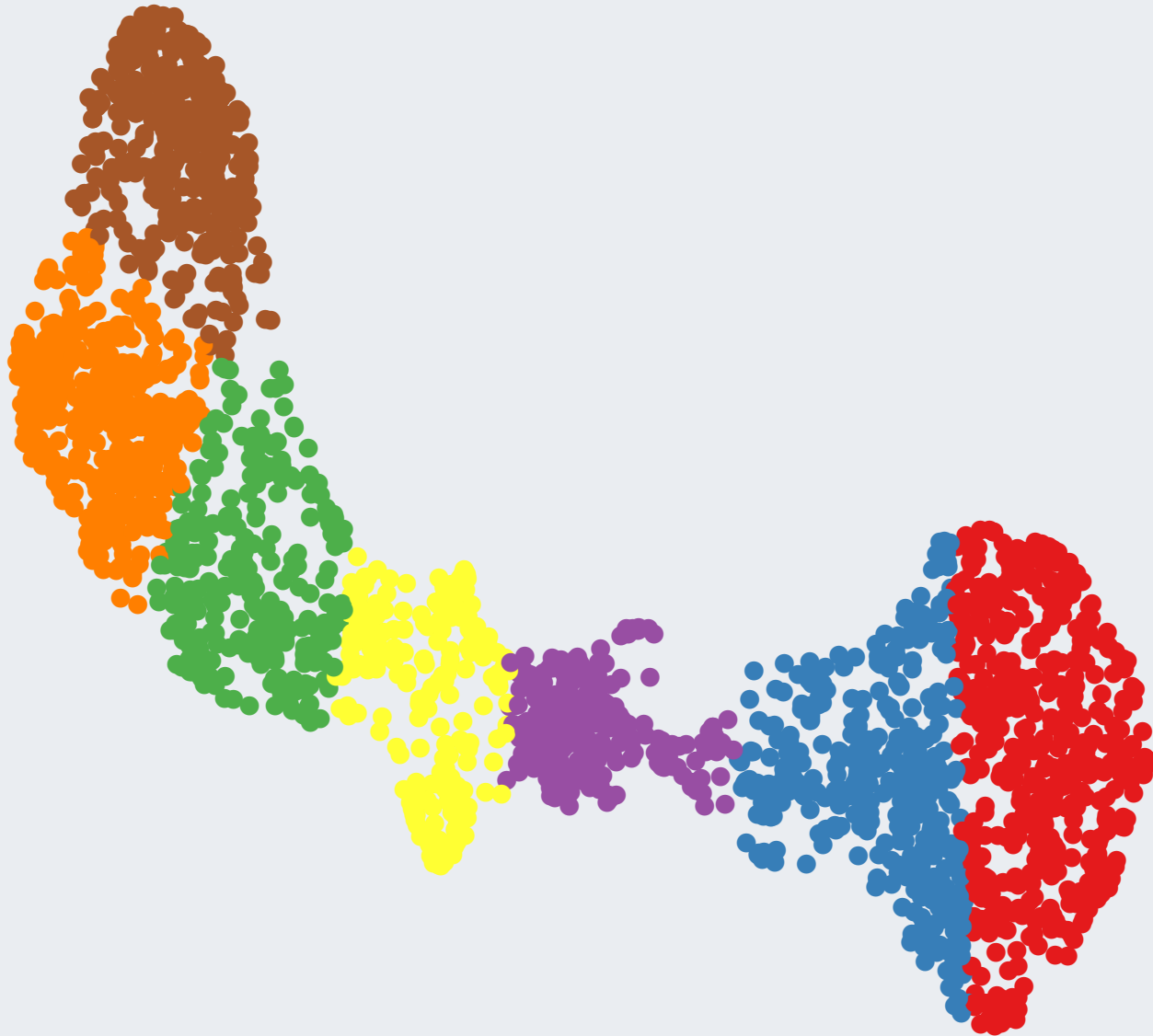
Single-cell RNA-Seq



Unmixing the smoothie

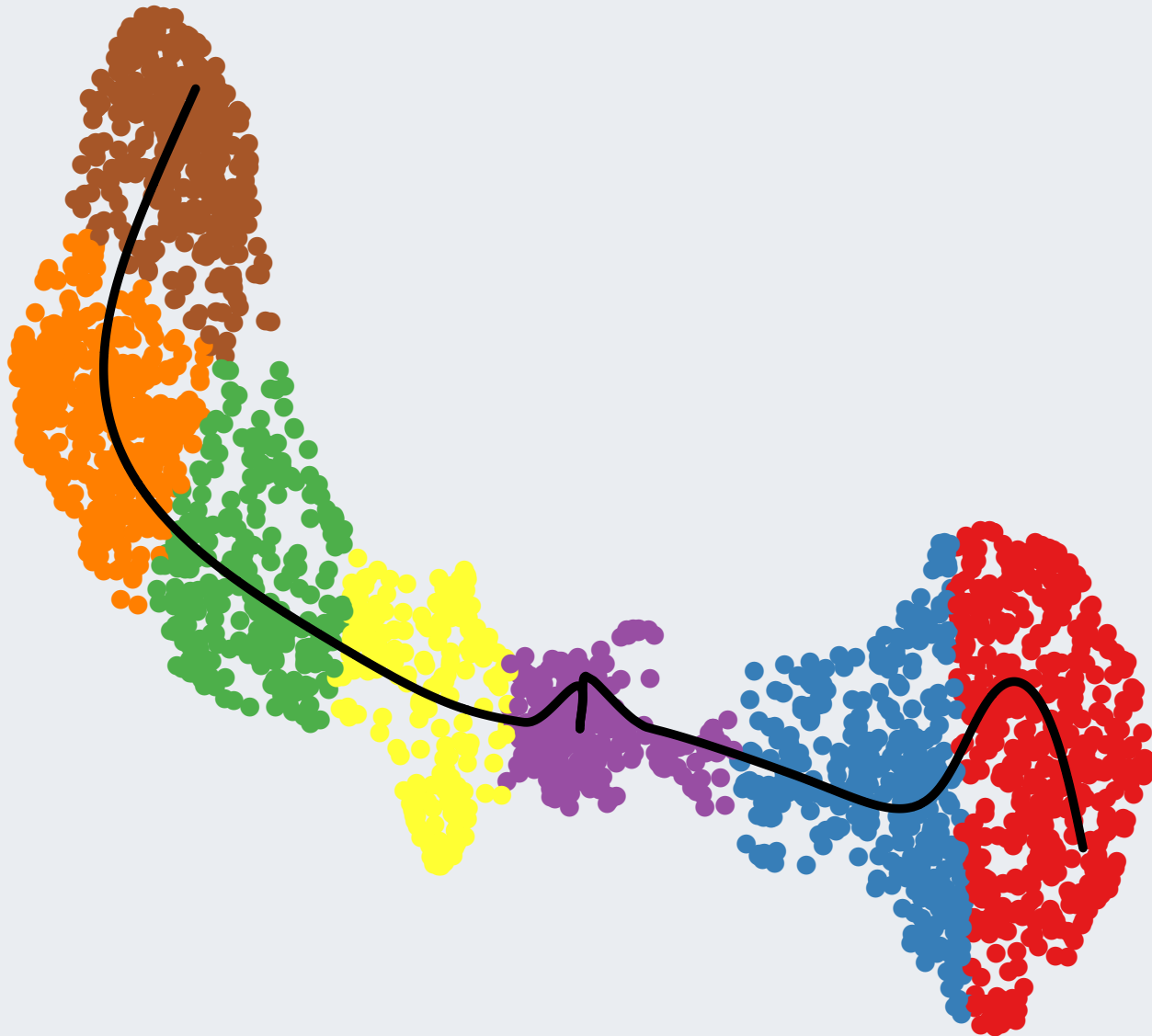
Bulk RNA - Seq	Single-cell RNA - Seq
 A glass jar filled with a thick green smoothie, topped with a yellow and white striped straw.	 
<p>VS</p>	<p>VS</p>

Dimensionality reduction



- Bone-marrow stem cells from the [monocle 3 vignette](#)
- 2660 cells and 3004 genes
- Now 2660 cells in two dimensions
- Many methods exist (PCA, t-SNE, UMAP)

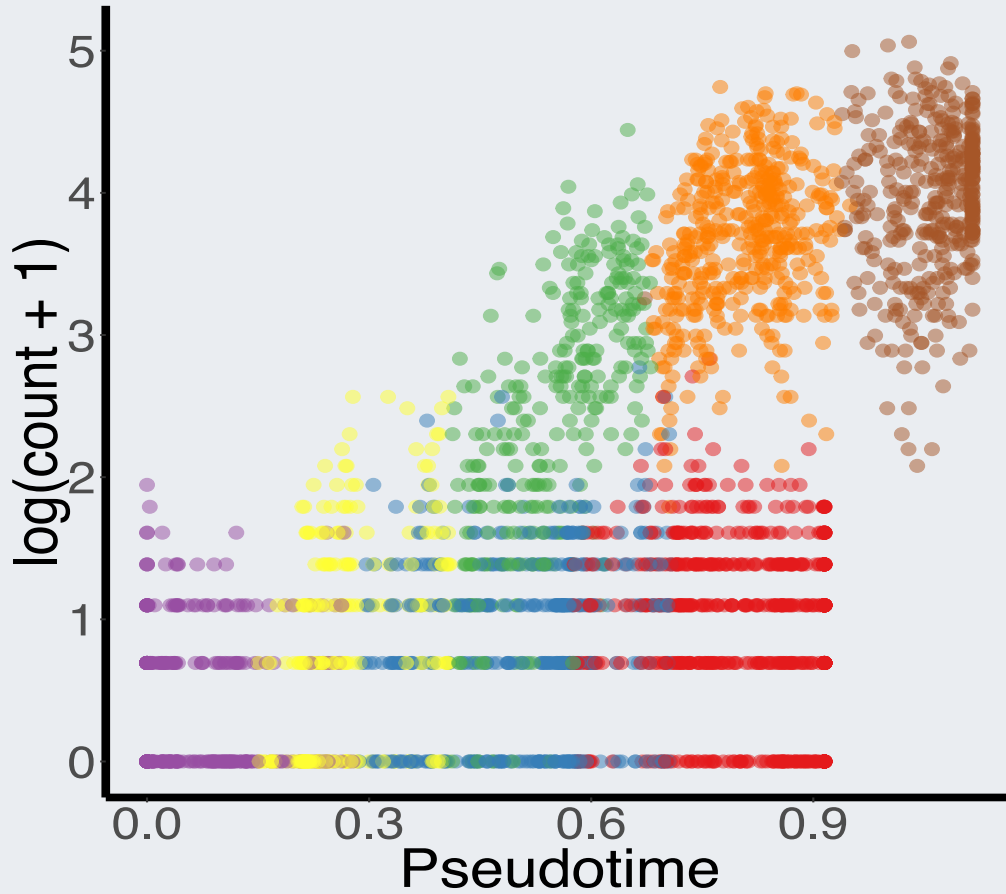
Trajectory inference



➤ Finding developmental paths

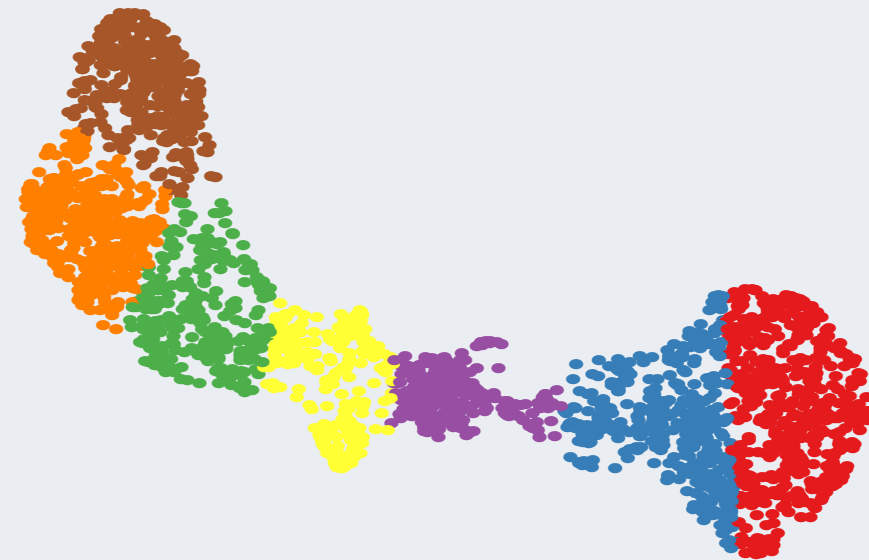
Each cell has a pseudotime, which measure how far along it is in the developmental process

cluster-based DE is artificial



Genes are now expressed in a continuous manner (since 2014)

Differential Expression is still cluster-based, i. e. discrete.



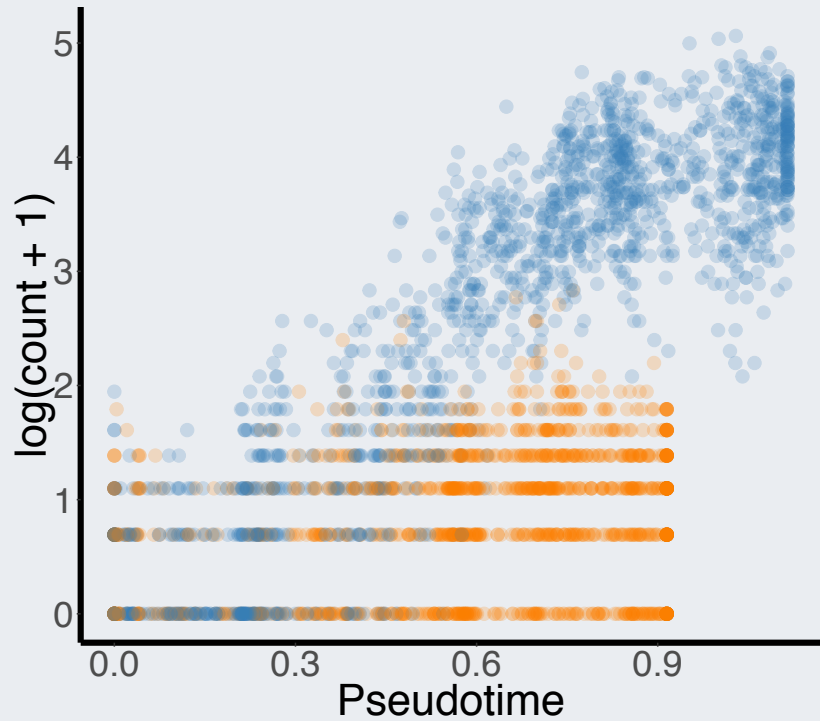


Trajectory-based DE

We developed *tradeSeq*, an algorithm that leverages the continuous nature of scRNA-Seq.

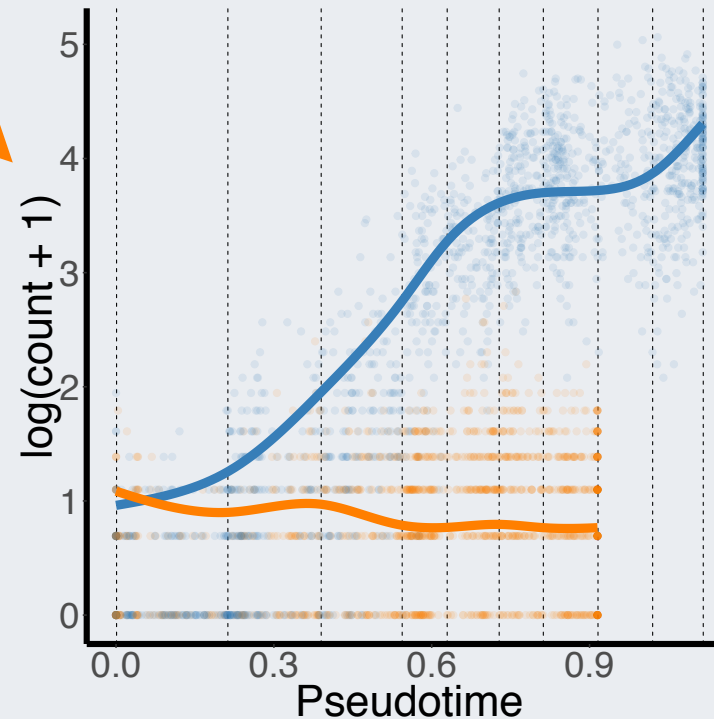
- Available as an **R** package on Github ([statOmics/tradeSeq](https://github.com/statOmics/tradeSeq)). Soon on Bioconductor.
- Modular tool that work with any dimensionality reduction and trajectory inference method.

Statistical model



Can accommodate

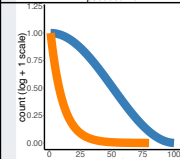
- Design matrix
- Different sequencing depth
- Weights



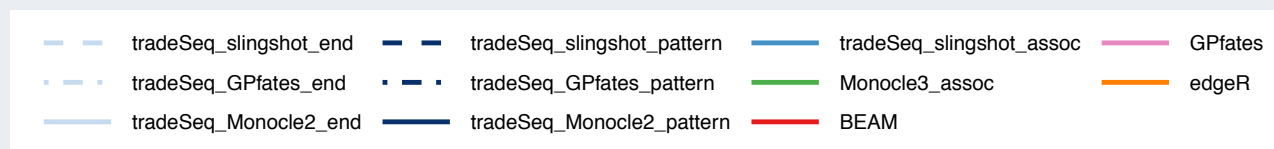
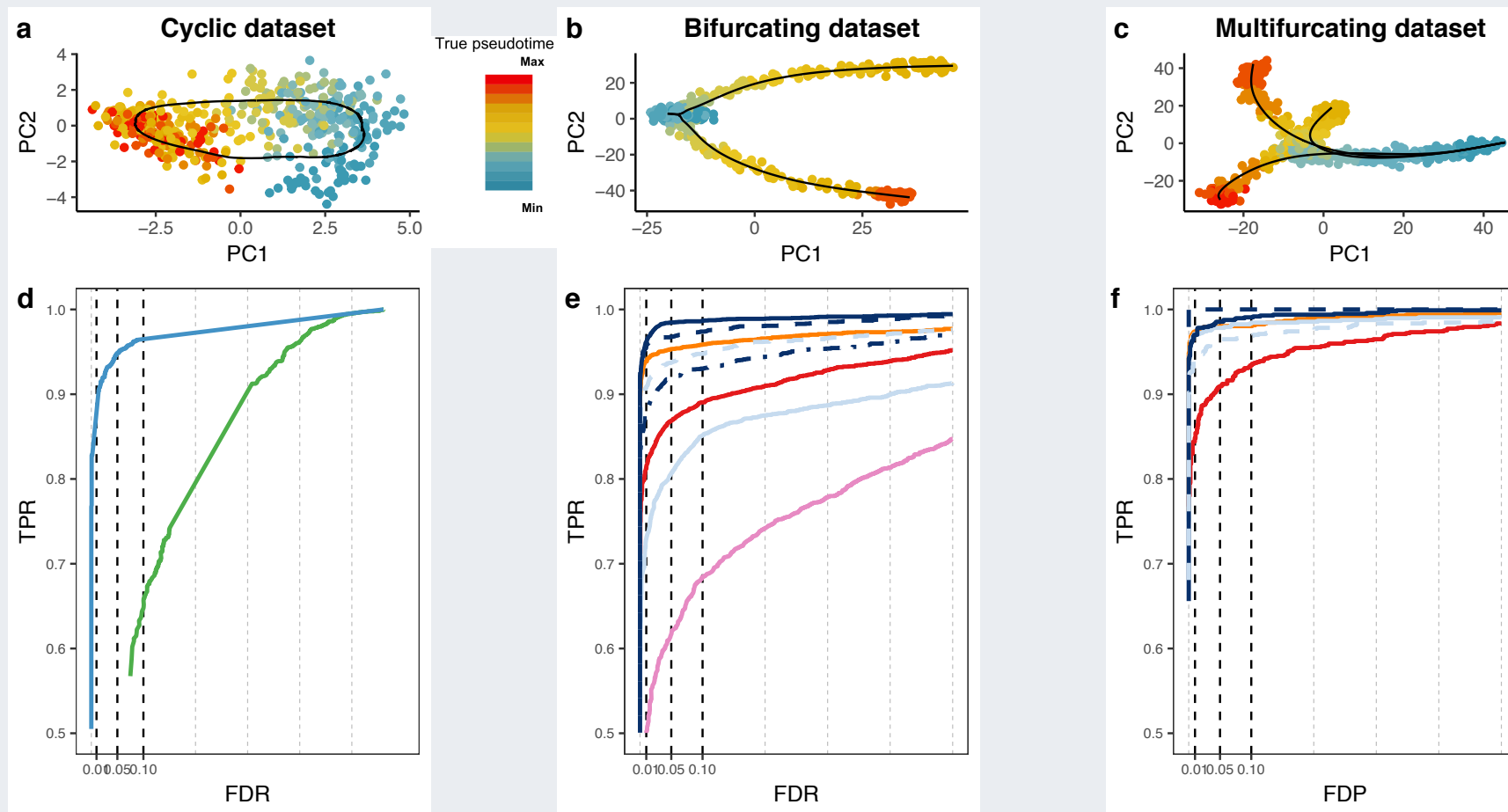
An investigation tool



Differential Expression Tests

Lineages	Within the orange lineage		Between the orange and blue lineages		
	associationTest	startVsEndTest	diffEndTest	patternTest	earlyDETest
	DE	DE	Not DE	Not DE	Not DE
	Not DE	Not DE	DE	DE	DE
	DE	Not DE	Not DE	Not DE	Not DE
	DE	DE	DE	DE	Not DE
	DE	DE	Not DE	DE	DE
	DE	DE	Not DE	DE	Not DE

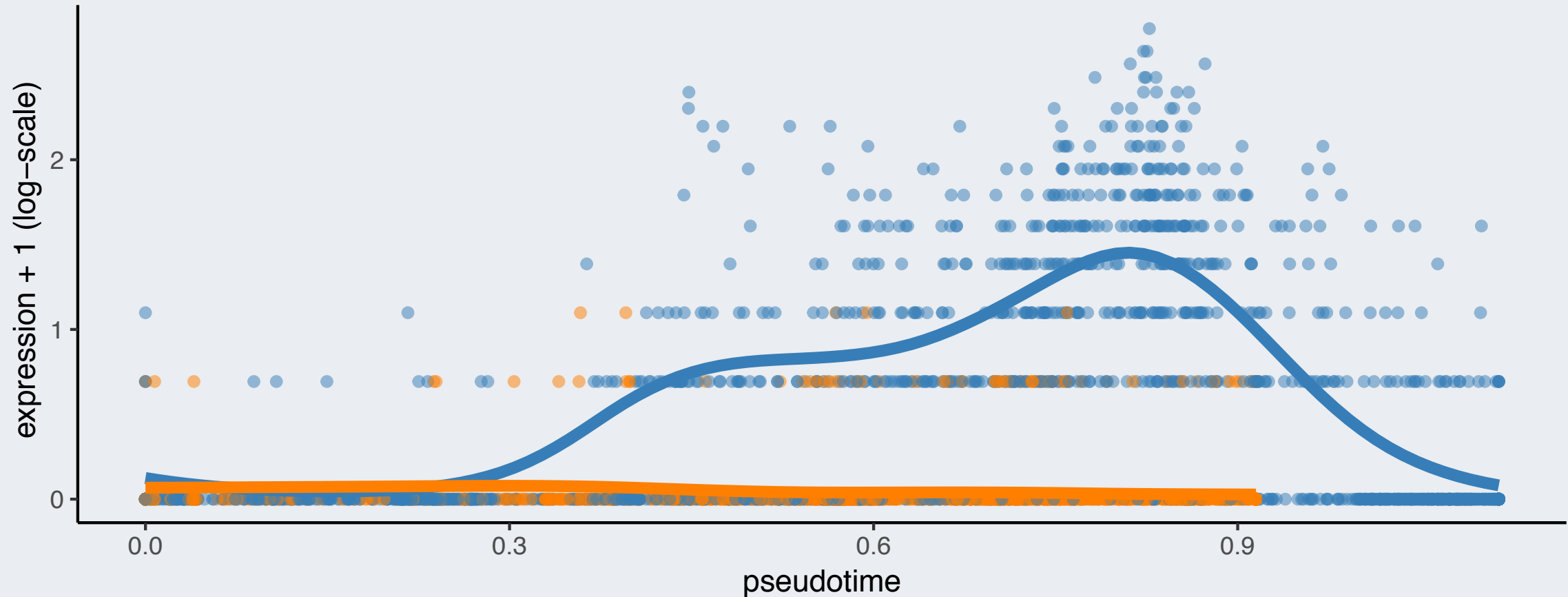
Outperforms existing methods



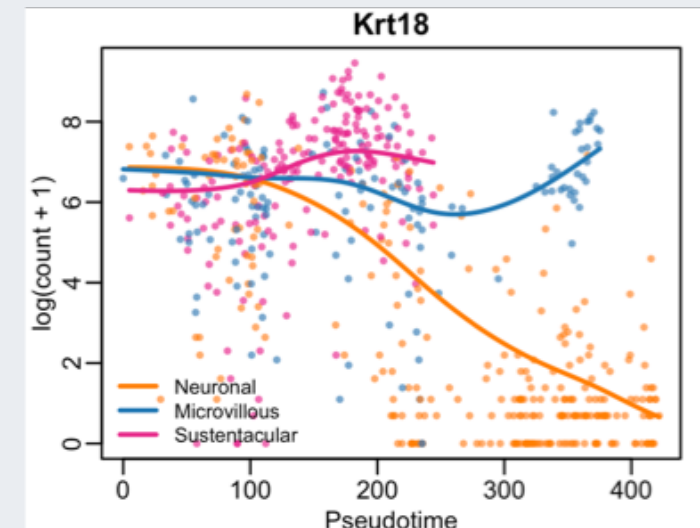
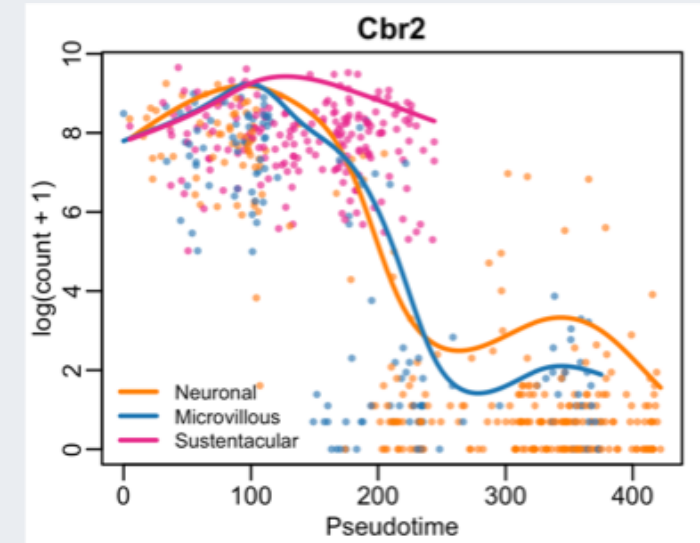
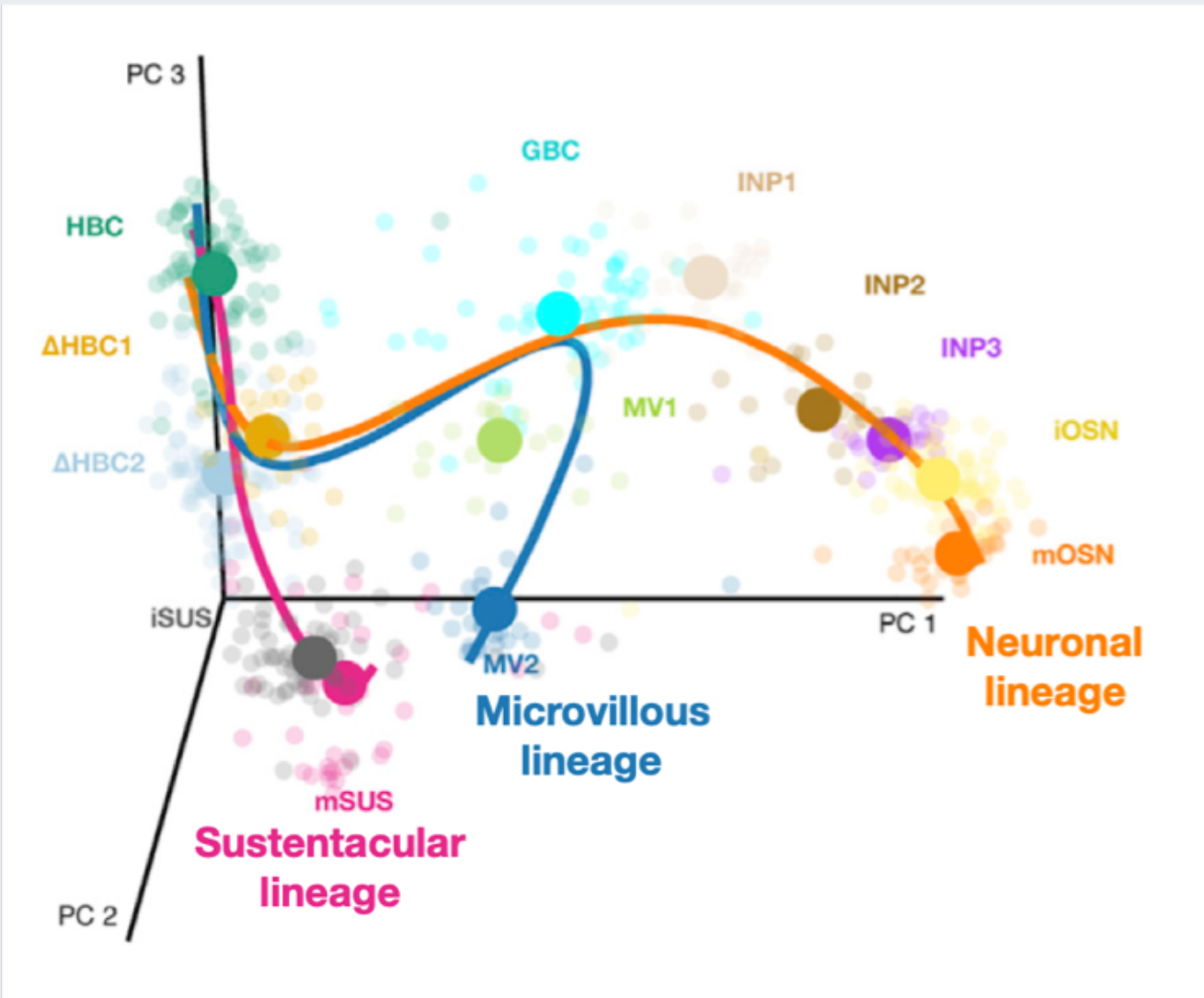
Provides unique insights



Gene Irf8 in the bone marrow dataset



Also works with multiple lineages





Take-aways

trade*Seq* is

- An open-source **R** package to investigate differential expression along and between trajectories
- A modular tool that work with any dimensionality reduction and trajectory inference method.
- An investigation tool that can answer various biologically relevant questions and provide unique and novel insights into complex datasets.



Thank you for listening

Any questions?