

SCENIC & cisTopic

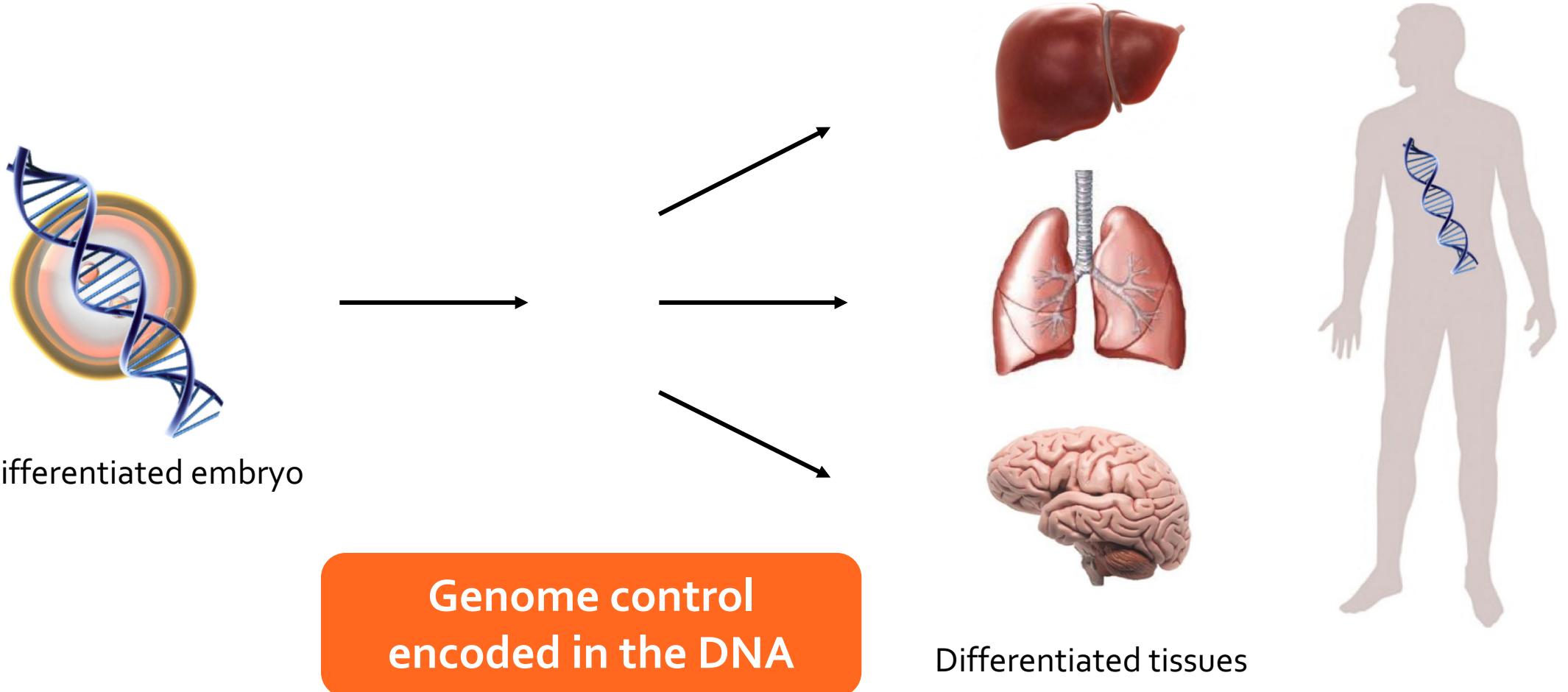
*Gene Regulatory Network inference
from single-cell transcriptomics & epigenomics data*

3 february 2019 – SingleCellTE - Roscoff

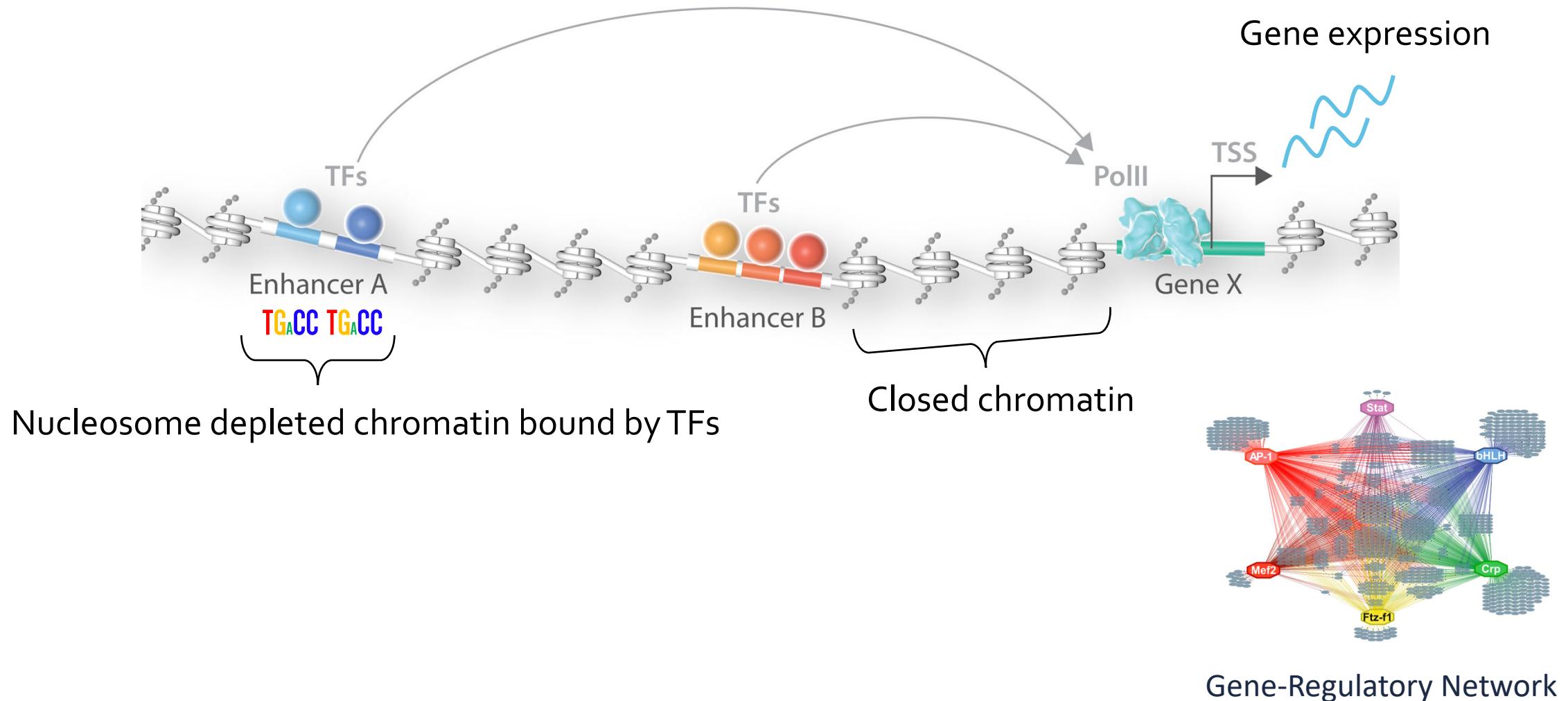


Sara Aibar & Carmen Bravo
Laboratory of Computational Biology (Stein Aerts lab)
VIB-KU Leuven Center for Brain & Disease research

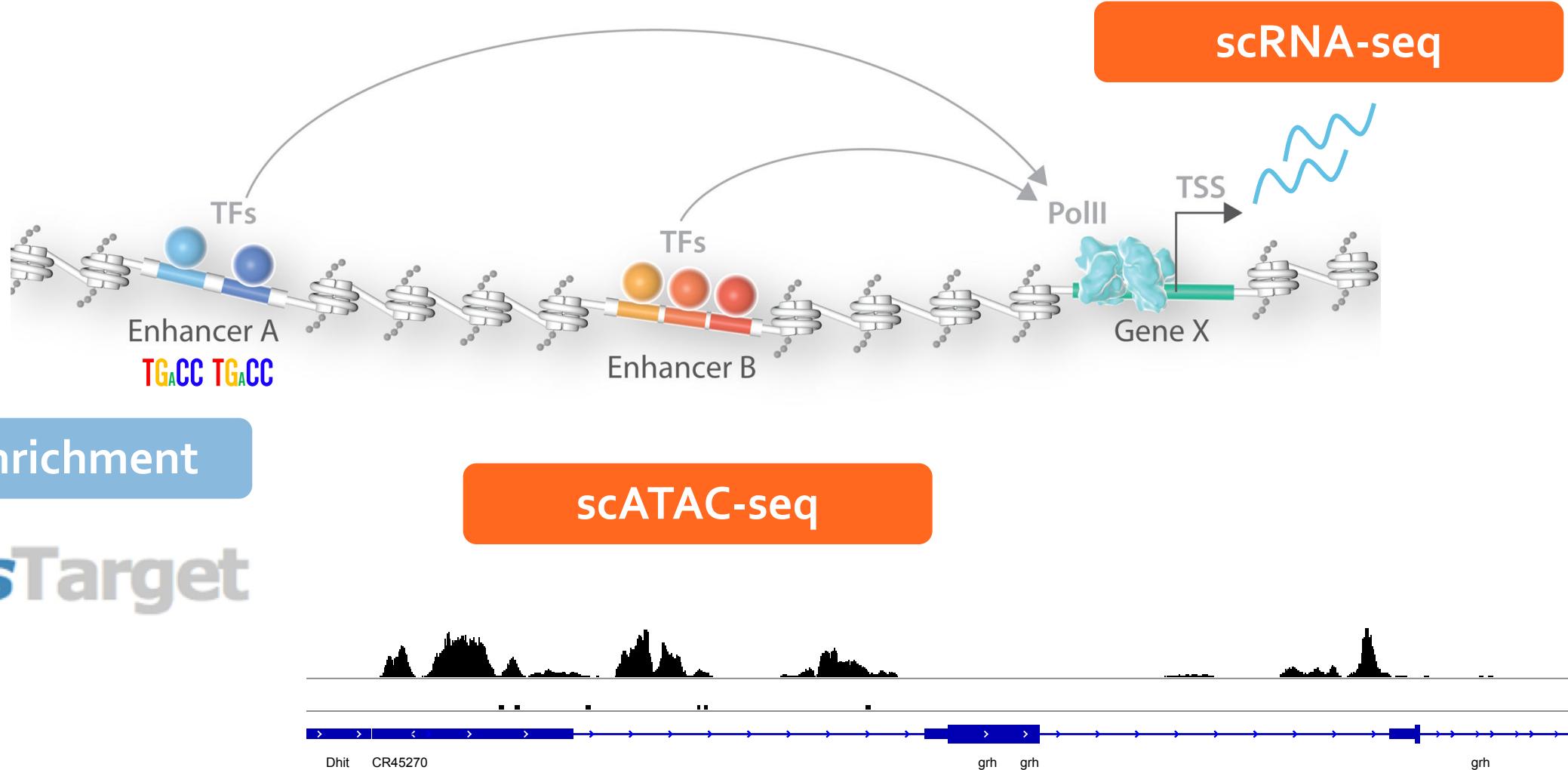
Deciphering the genomic regulatory code



Gene Regulatory Programs

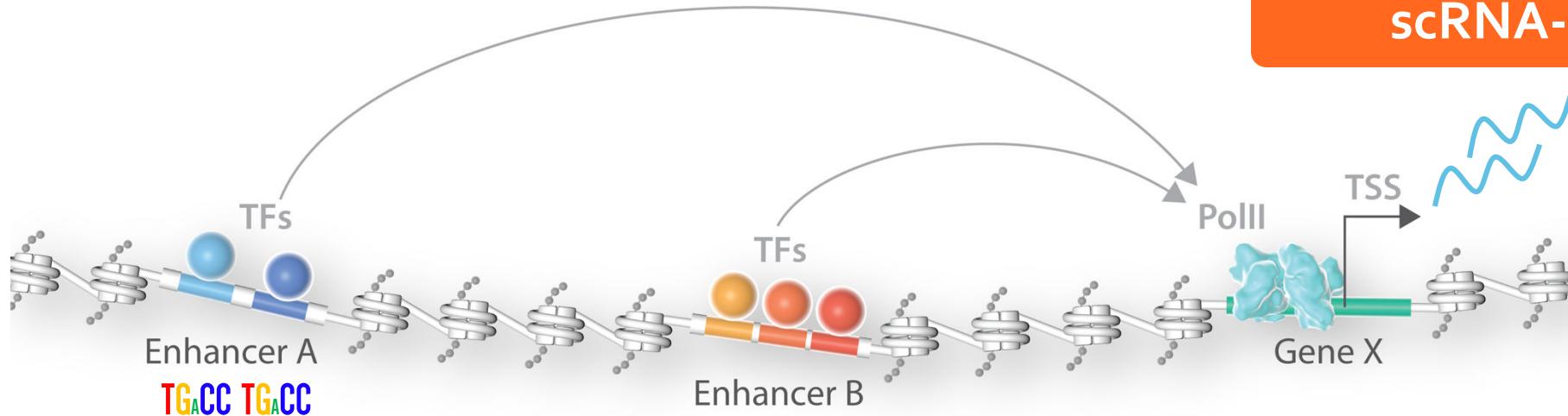


Regulatory genomics at single-cell resolution



Regulatory genomics at single-cell resolution

(our tools)



Motif enrichment



scATAC-seq

cisTopic

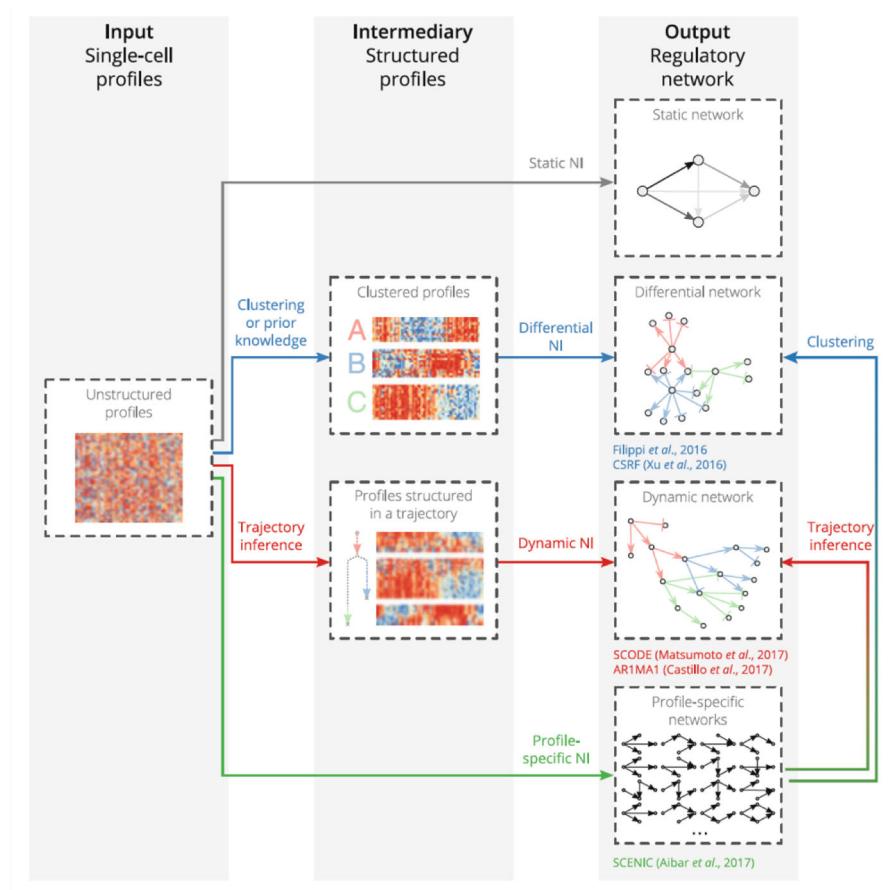
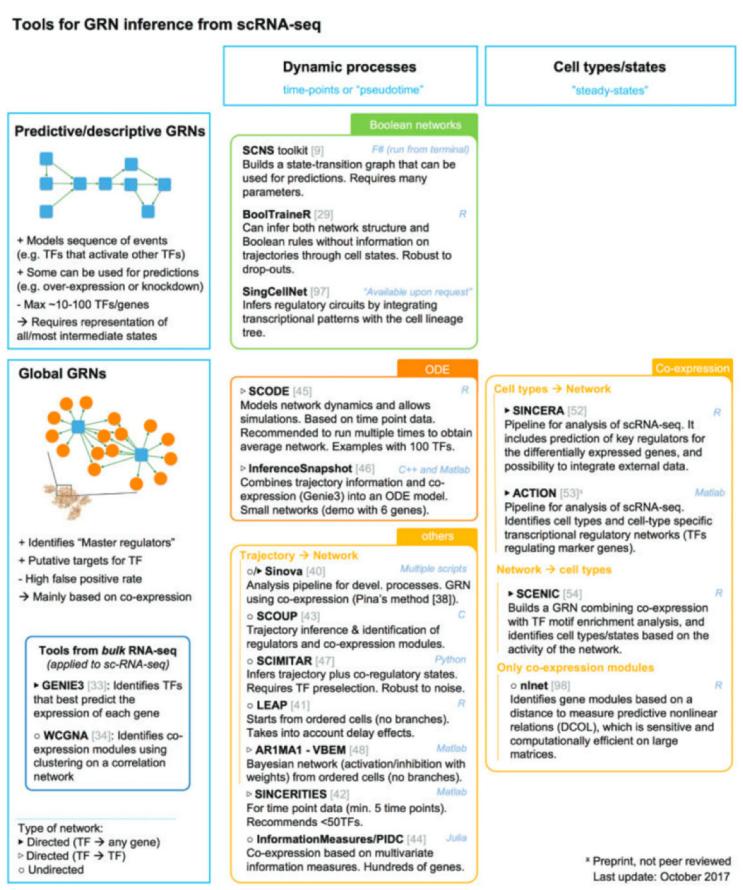
SCENIC

scRNA-seq

Arboreto
AUCell
SCope
pySCENIC
SCopeLoomR

Visualization

Other methods?



Mapping gene regulatory networks from single-cell omics data

Mark W E J Fiers, Liesbeth Minnoye, Sara Aibar, Carmen Bravo González-Blas,

Zeynep Kalender Atak, Stein Aerts Author Notes

Briefings in Functional Genomics, Volume 17, Issue 4, 1 July 2018, Pages 246–254,

Chapter 10

Network Inference from Single-Cell Transcriptomic Data

Helena Todorov, Robrecht Cannoodt, Wouter Saelens, and Yvan Saeyns

What is your objective?

Aim: GRNs



Plan experiments a-priori

Which (sc)-omics?

Which conditions?

(cell types, perturbations,...)

Complementarity!!

I already have
scRNA-seq...



Let's also check gene regulation...

"Master" regulators for cell types

Regulatory states

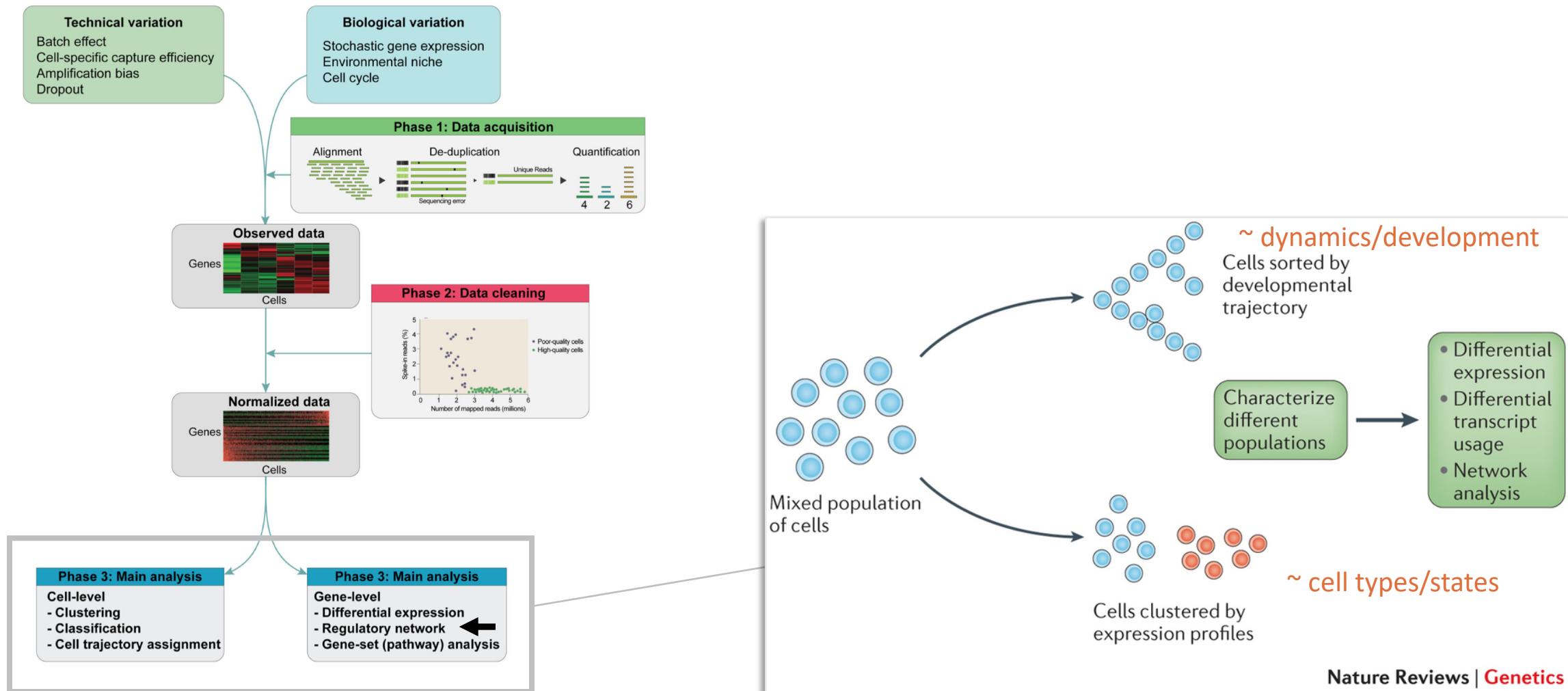
GRNs

...

SCENIC

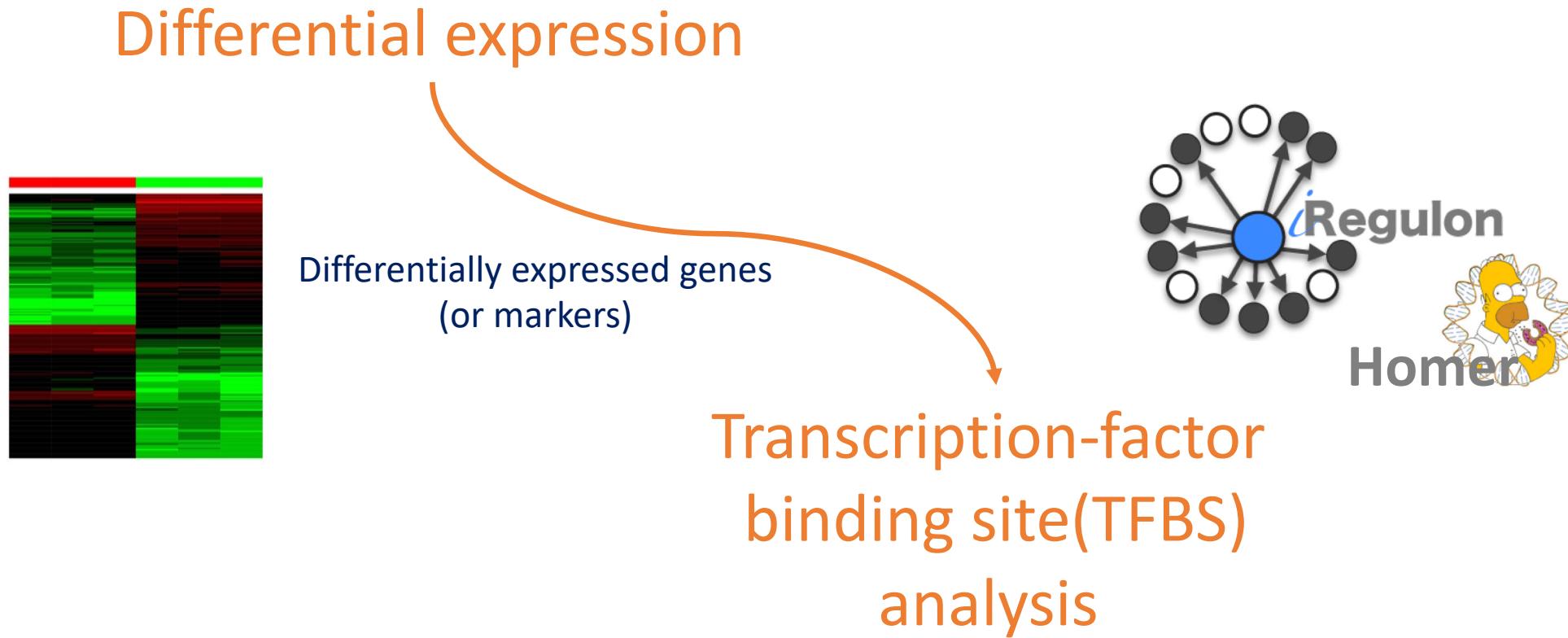
Gene regulatory network inference from **single-cell RNA-seq** data

Standard analysis of single-cell RNA-seq



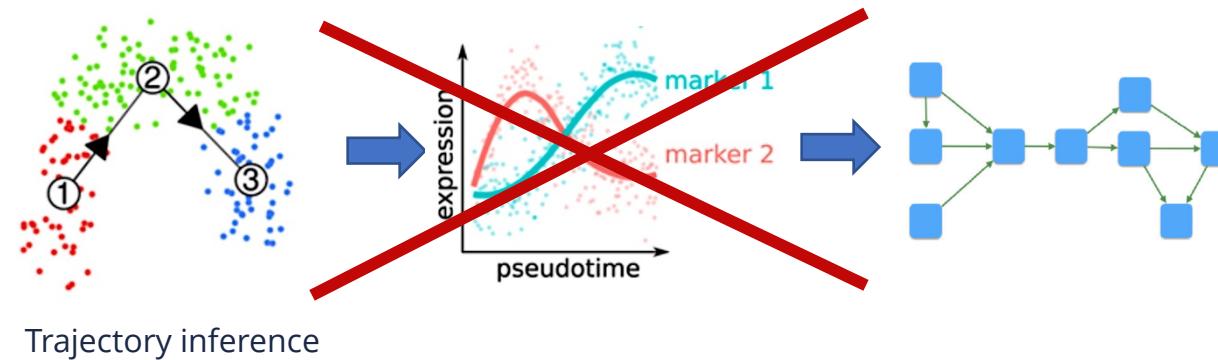
Exploring gene-regulation in sc-transcriptomics

Typical approach in bulk RNA-seq analyses:

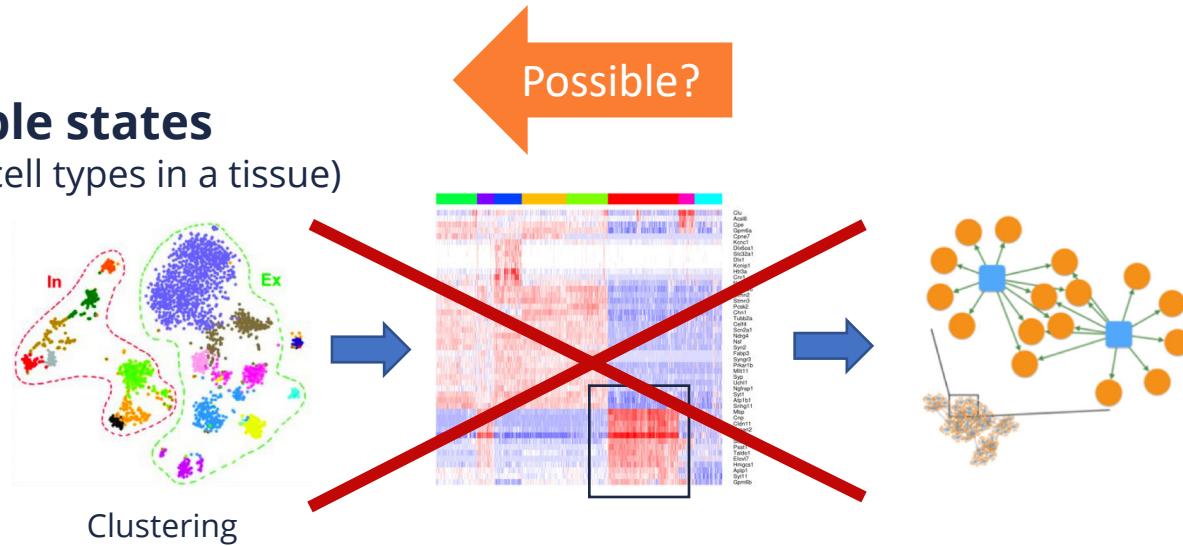


Tools for GRN inference from scRNA-seq

Dynamic processes (e.g. differentiation)



Stable states (e.g. cell types in a tissue)



Possible?

Boolean networks

SCNS
BoolTrainR
SingleCellNet

ODEs

SCODE

Other

Sinova
SCOUP
SCIMITAR
LEAP
SINCERITI
ES

► Model dynamics/predict effects

SINCERA
ACTION

► Regulators for marker genes

More info:

Fiers et al. (Briefings in Functional Genomics, 2018)
Todorov et al. (Methods in Molecular Biology, 2019)

Single Cell rEgulatory Network Inference and Clustering

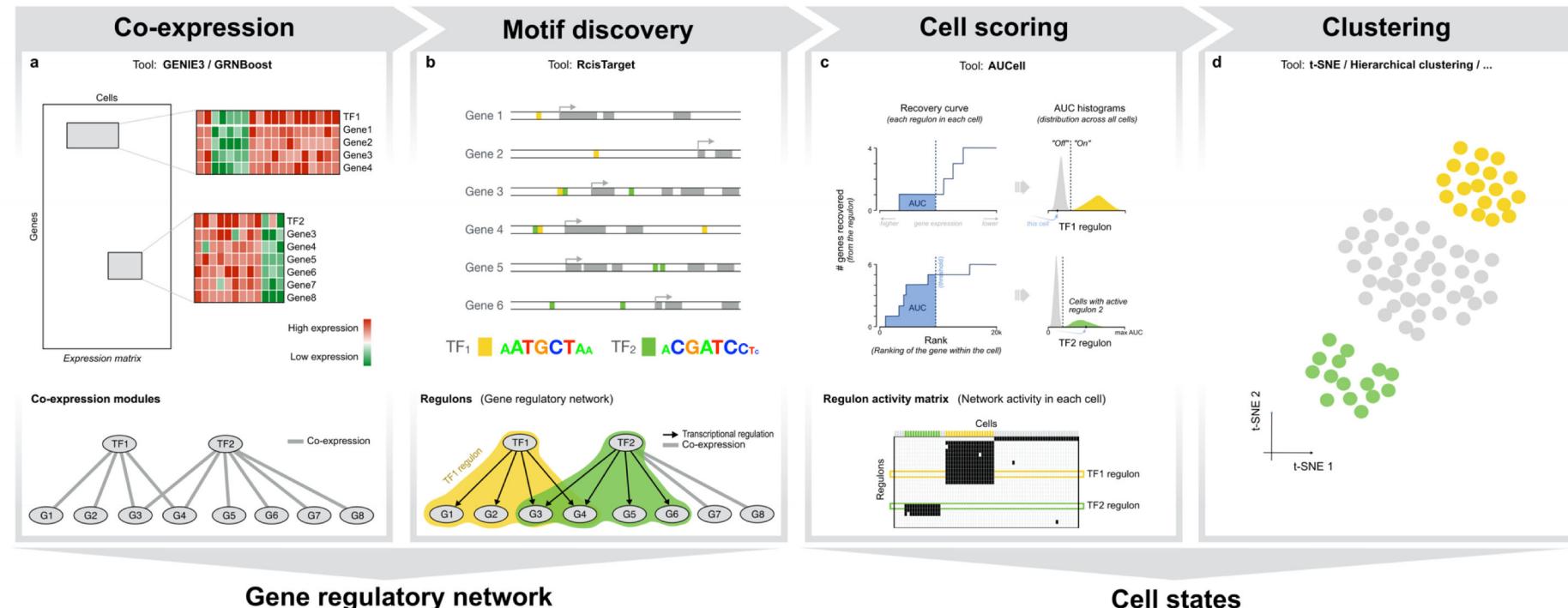


Table of contents

SCENIC

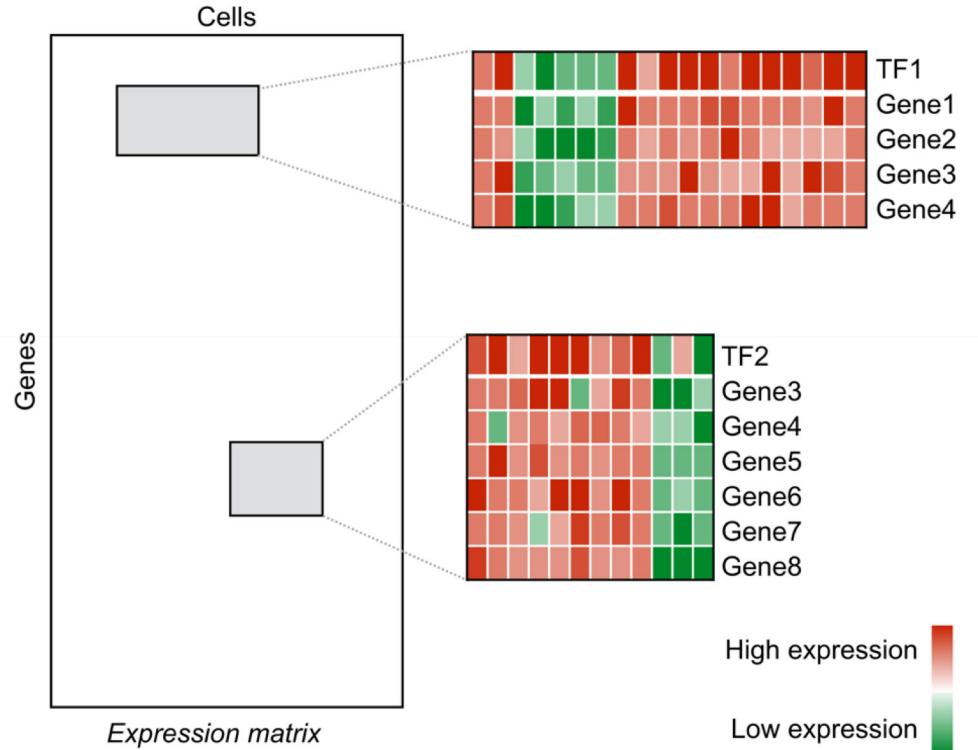
- Algorithm
- Validation
- Usage examples
- How to run it?

NATURE METHODS | VOL.14 NO.11 | NOVEMBER 2017

SCENIC: single-cell regulatory network inference and clustering

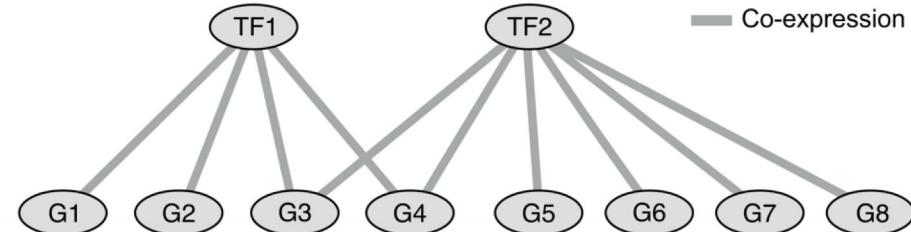
Sara Aibar, Carmen Bravo González-Blas, Thomas Moerman, Ván Anh Huynh-Thu, Hana Imrichova, Gert Hulselmans, Florian Rambow, Jean-Christophe Marine, Pierre Geurts, Jan Aerts, Joost van den Oord, Zeynep Kalender Atak, Jasper Wouters & Stein Aerts 

1. TF-based co-expression network



GENIE3
(Random Forest)
or
GRNboost
(Gradient Boosting)

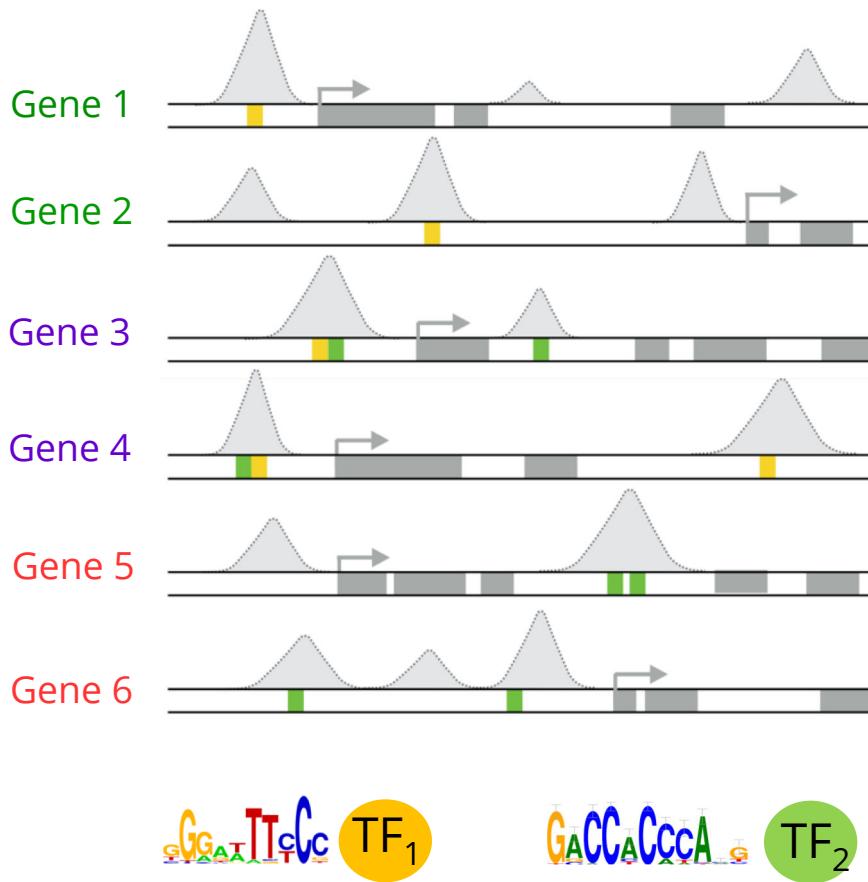
Co-expression modules



GENIE3: Best performer in DREAM4 and DREAM5 challenges.

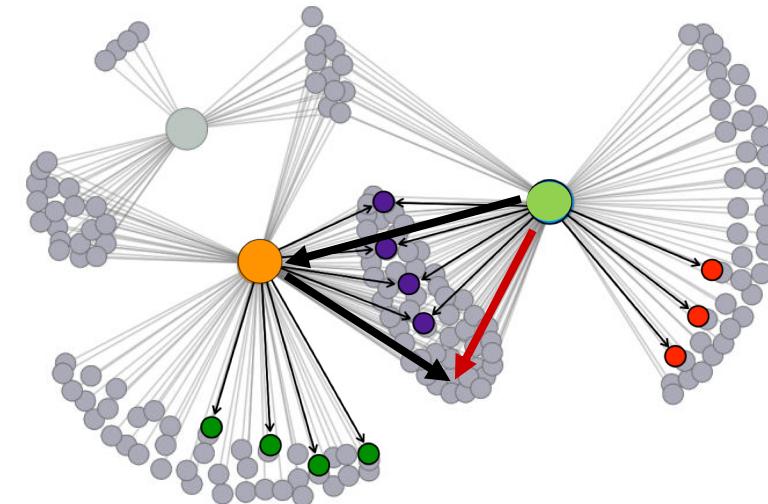
Huynh-Thu et al. PloS ONE 2010. Inferring regulatory networks from expression data using tree-based methods

2. Gene Regulatory Network



RcisTarget
cis-regulatory sequence analysis

Gene regulatory network

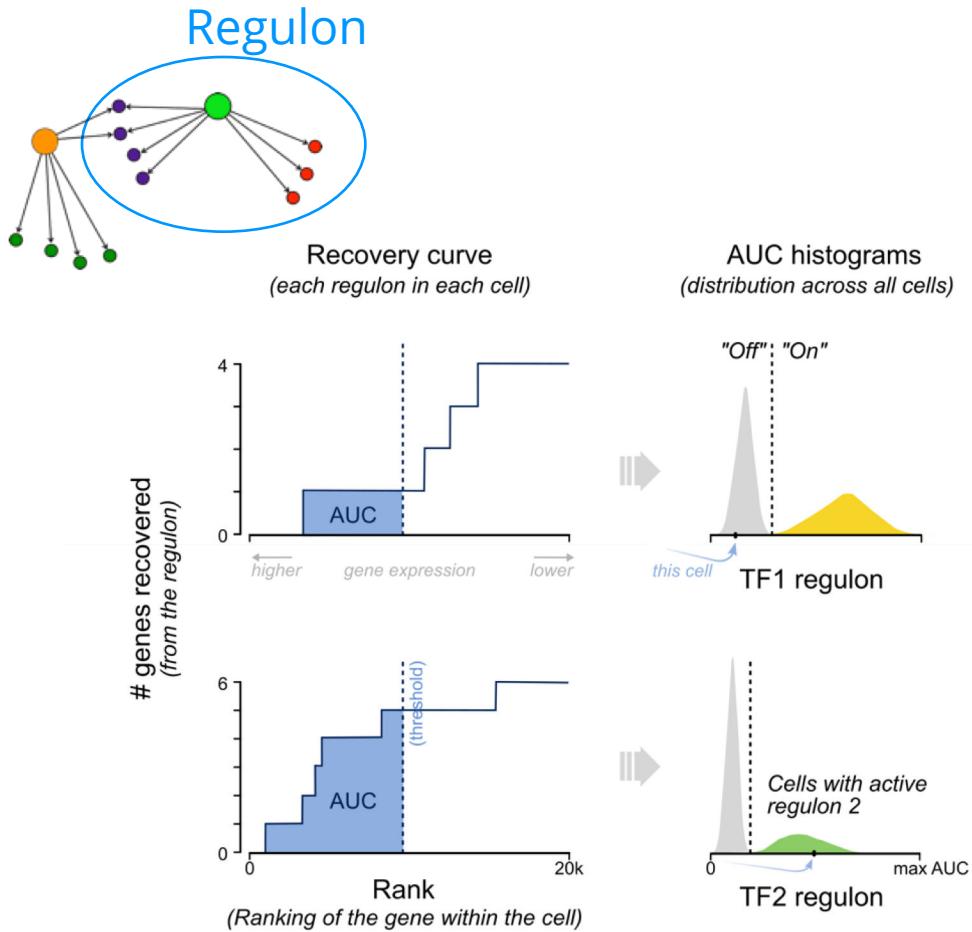


Other implementations: **i-cisTarget** (web version), **iRegulon** (cytoscape)

Method applied and validated in **Potier et al.** Cell Reports 2014:

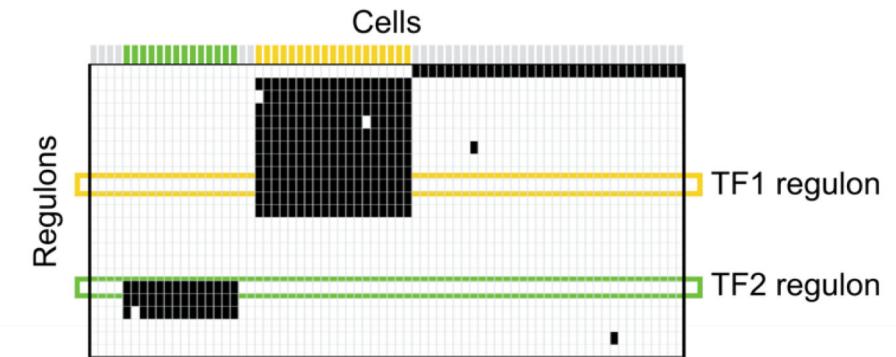
Mapping Gene Regulatory Networks in Drosophila Eye Development by Large-Scale Transcriptome Perturbations and Motif Inference

3. Activity of the network in each cell



AUCell
*Identifying cells
with active gene-sets*

Network activity in each cell



(regulon activity score)

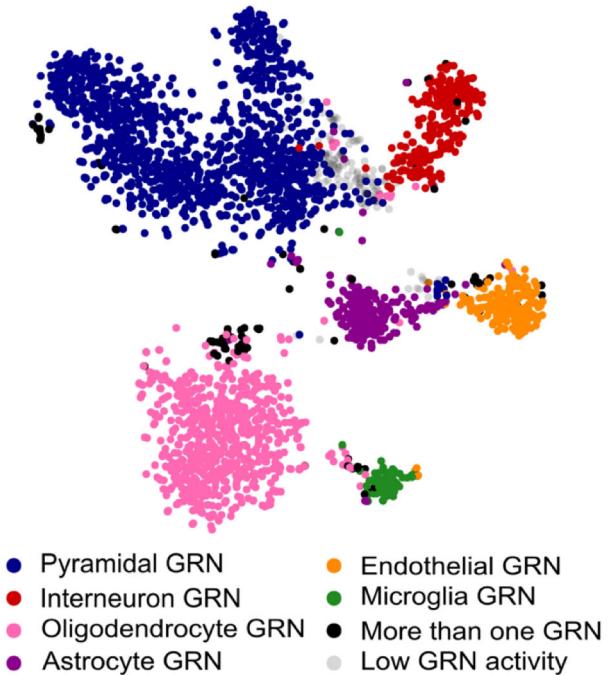
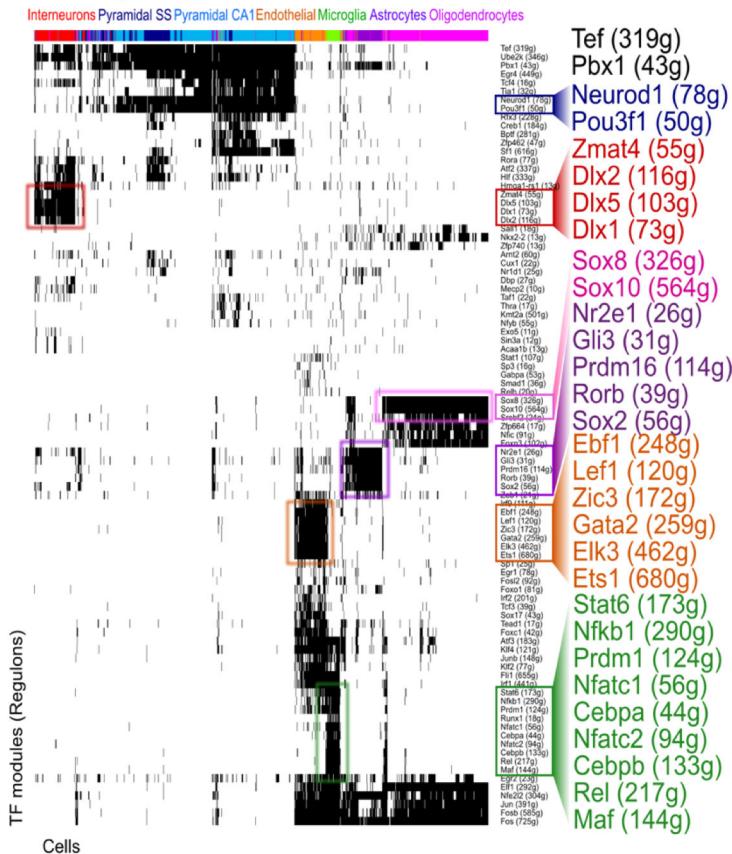
Application on data sets

(2017)

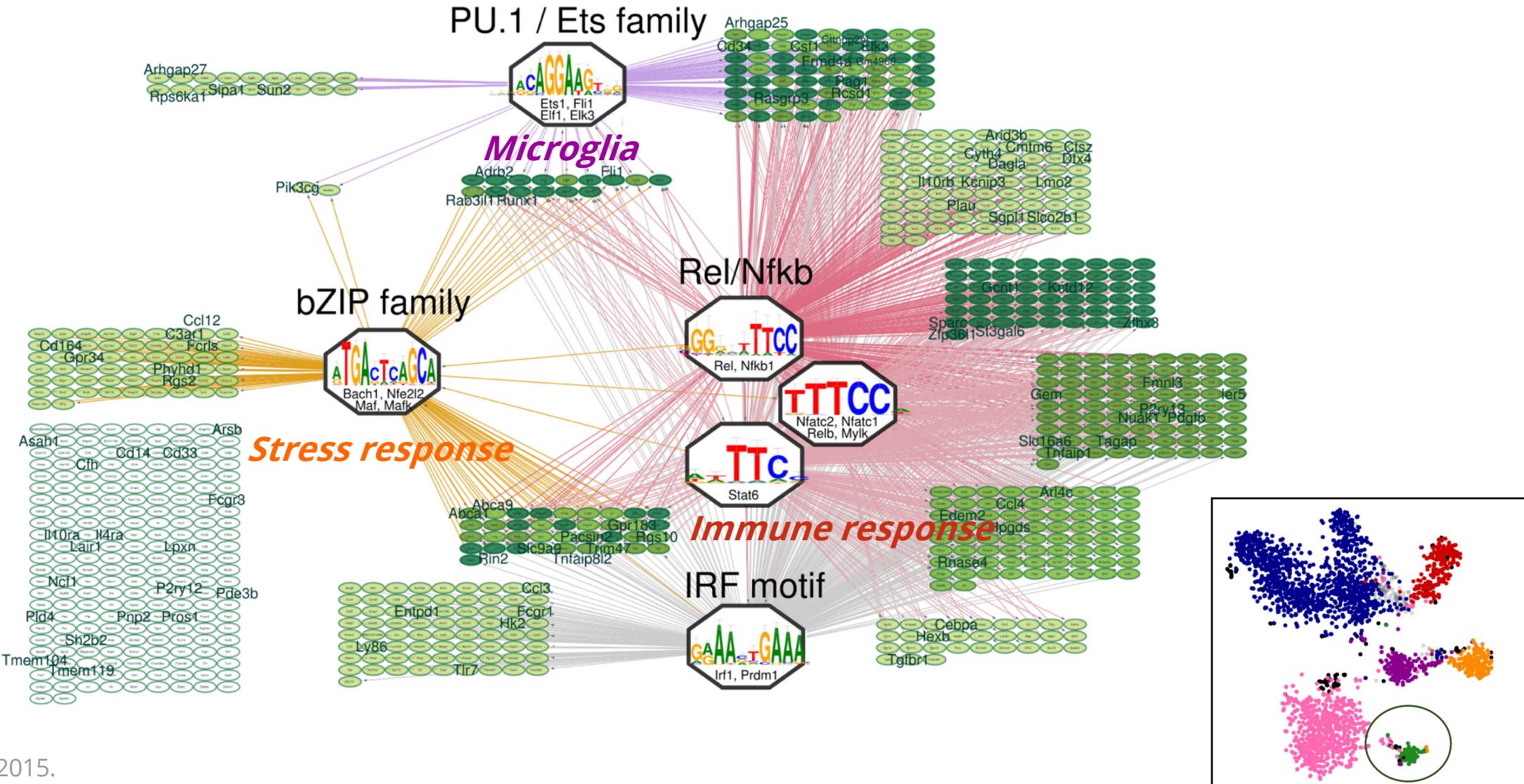
Dataset		# cells	Feature
Brain (Zeisel et al.)	Mouse	3005	✓ Identification of cell types and master regulators
Neurons (Lake et al.)	Human	3083	✓ Cross-species GRN comparison ✓ Data integration across datasets, technologies and species
Brain (Darmanis et al.)	Human	466	
Oligodendrocytes in development (Marques et al.)	Mouse	5069	✓ Analysis of continuous processes
Oligodendrogloma (Tirosh et al.)	Human	4347	✓ Application on cancer / other types of states
Melanoma (Tirosh et al.)	Human	1252	✓ Overcoming tumor-effect
Retina (Macosko et al.)	Mouse	44808	✓ Scalability to larger datasets
Embryonic brain (10X Genomics)	Mouse	1,3M	

Validation: Mouse Brain (3005 cells)

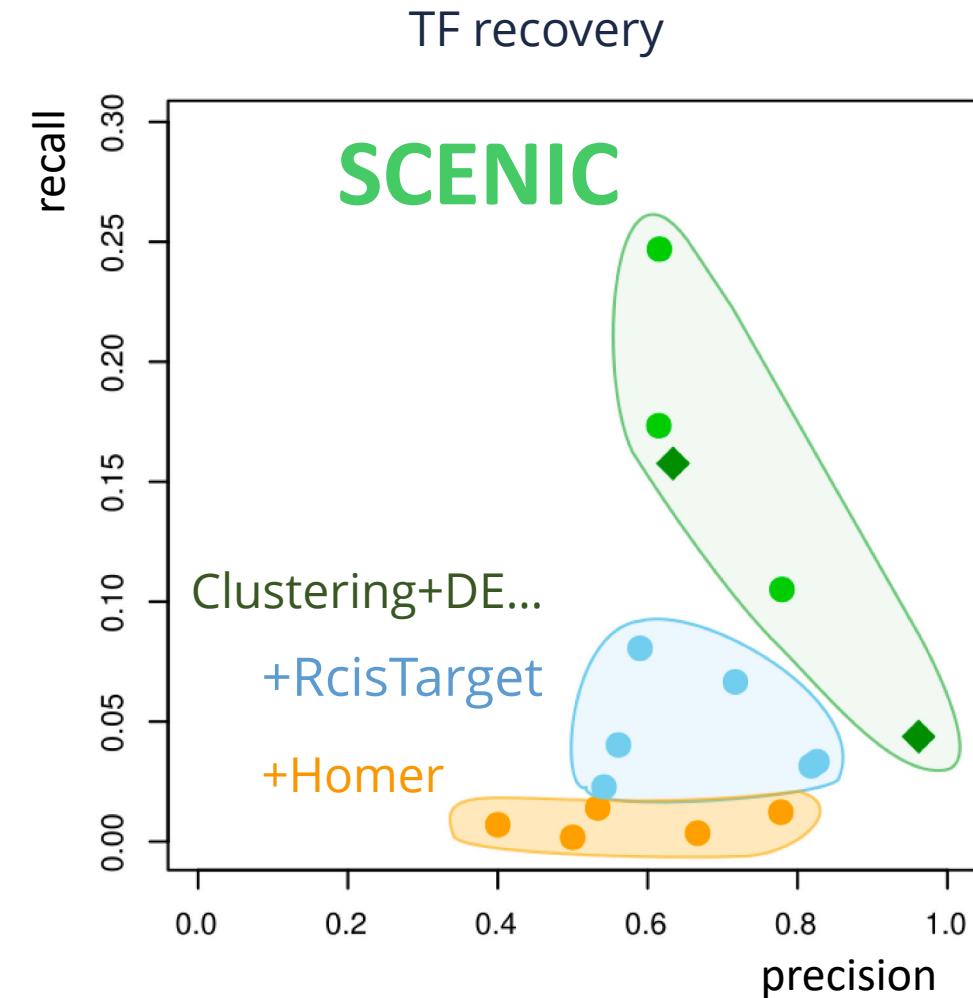
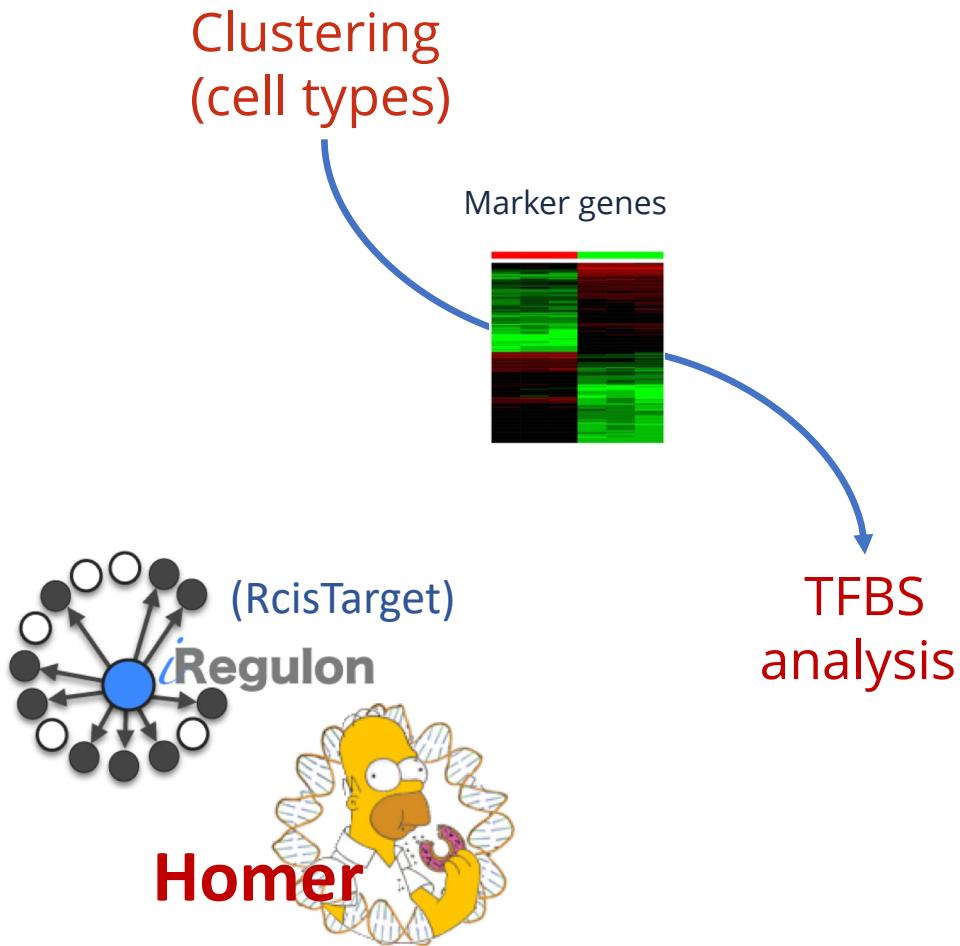
Binary regulon activity matrix



Microglia network

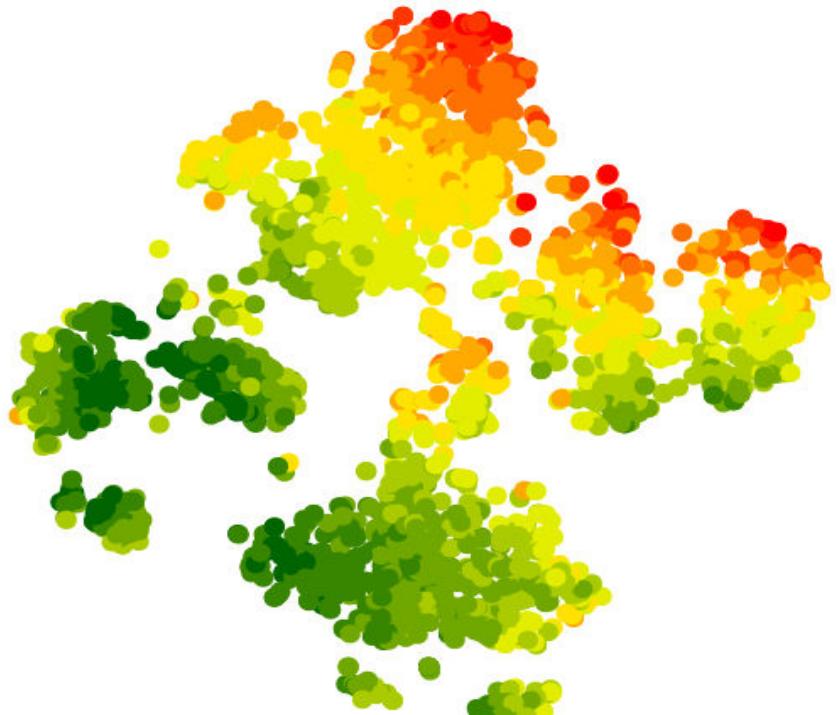


Is it better than standard approaches?



Reduction of bias/Batch effect

t-SNE on
expression matrix

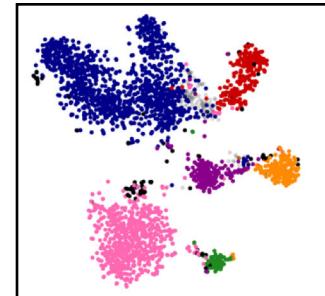


Color: Number of genes detected per cell

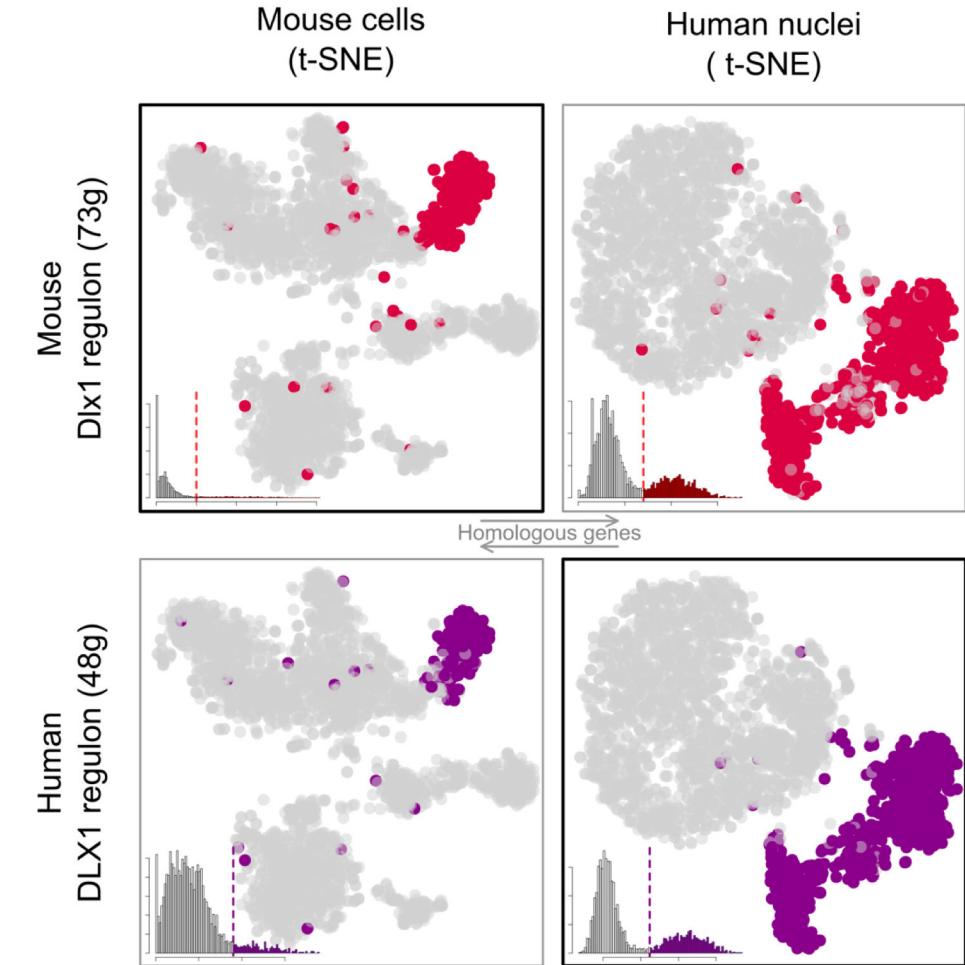
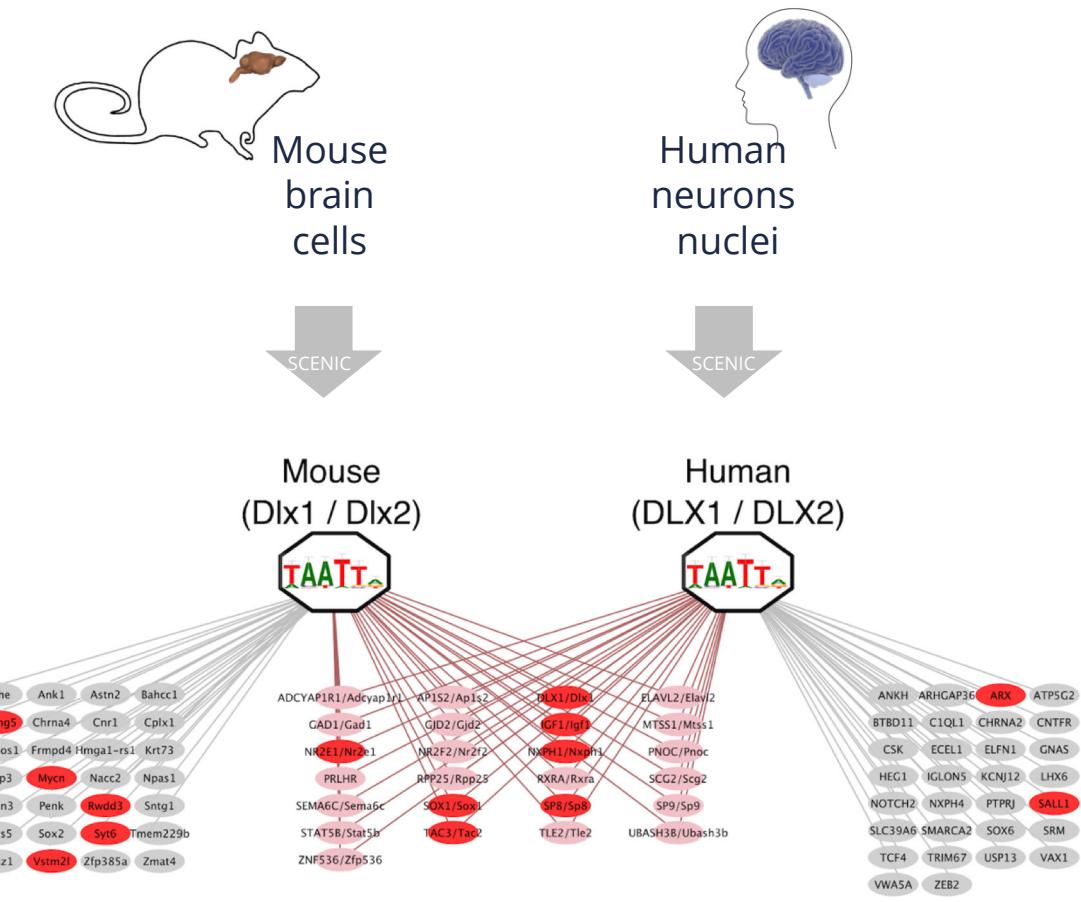


Dataset: Zeisel et al. 2015.

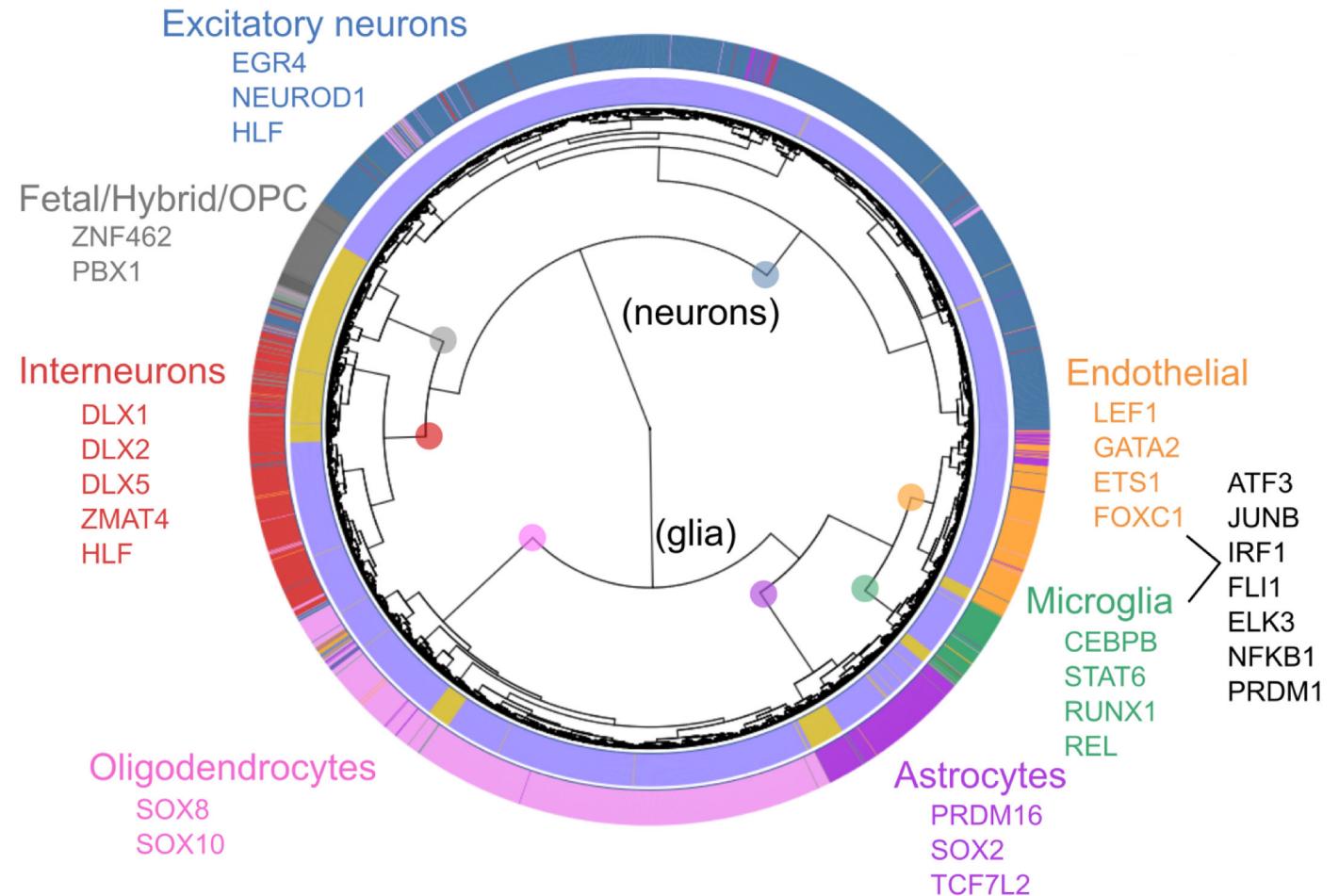
t-SNE on
binary regulon activity



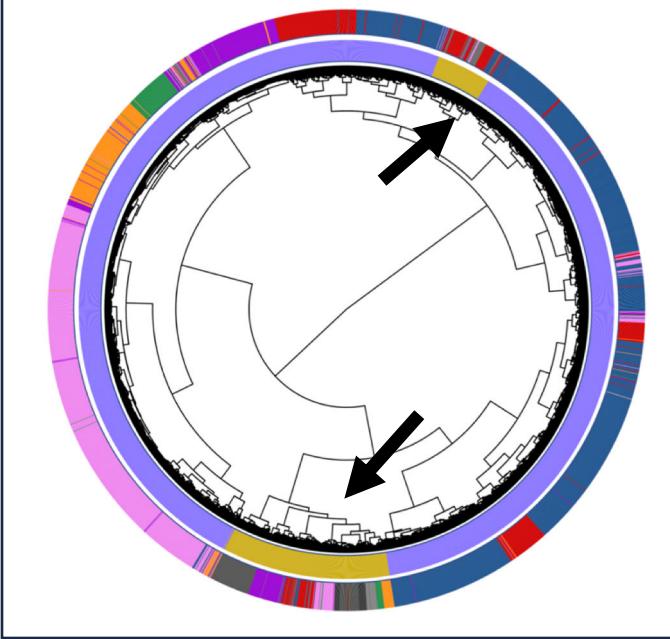
Cross-species GRN comparisons



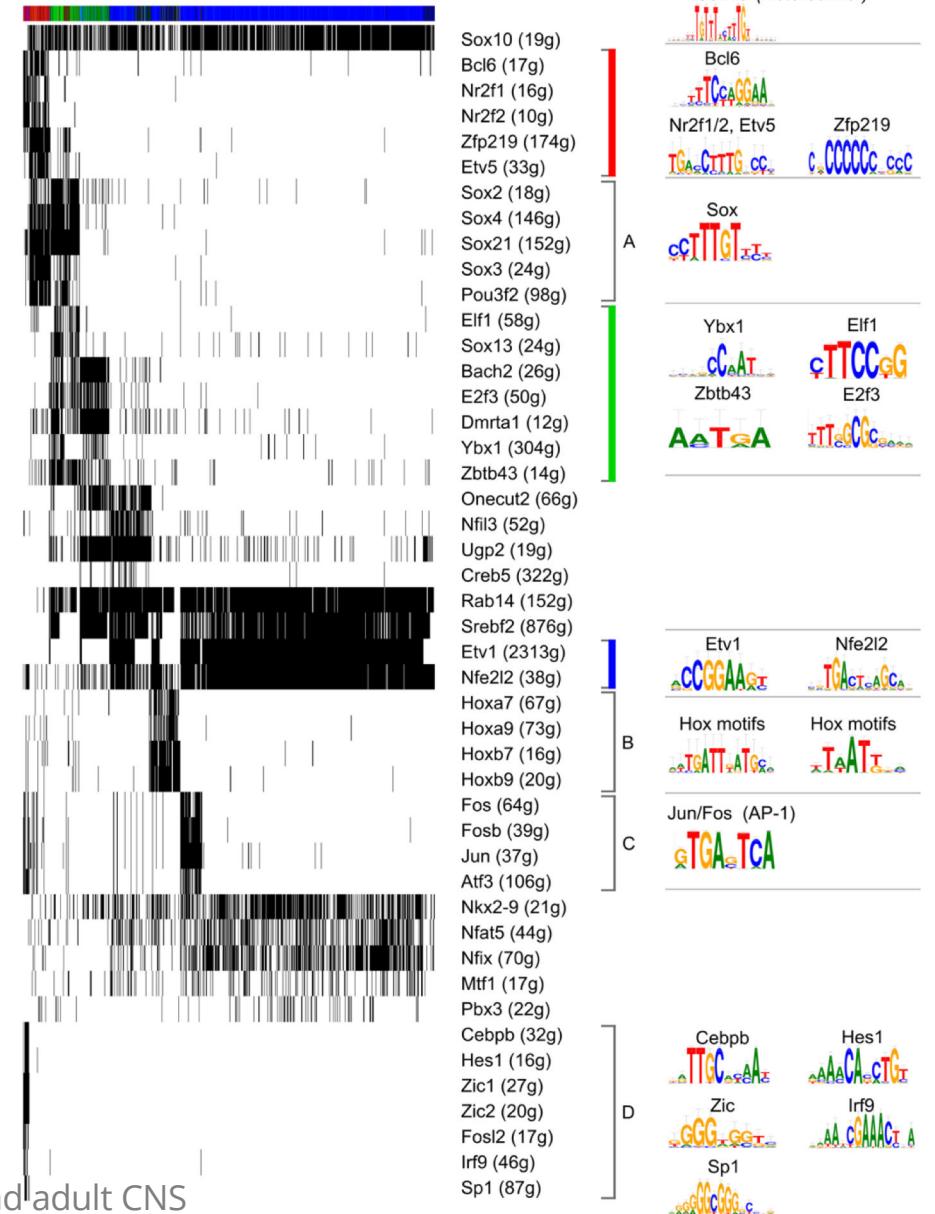
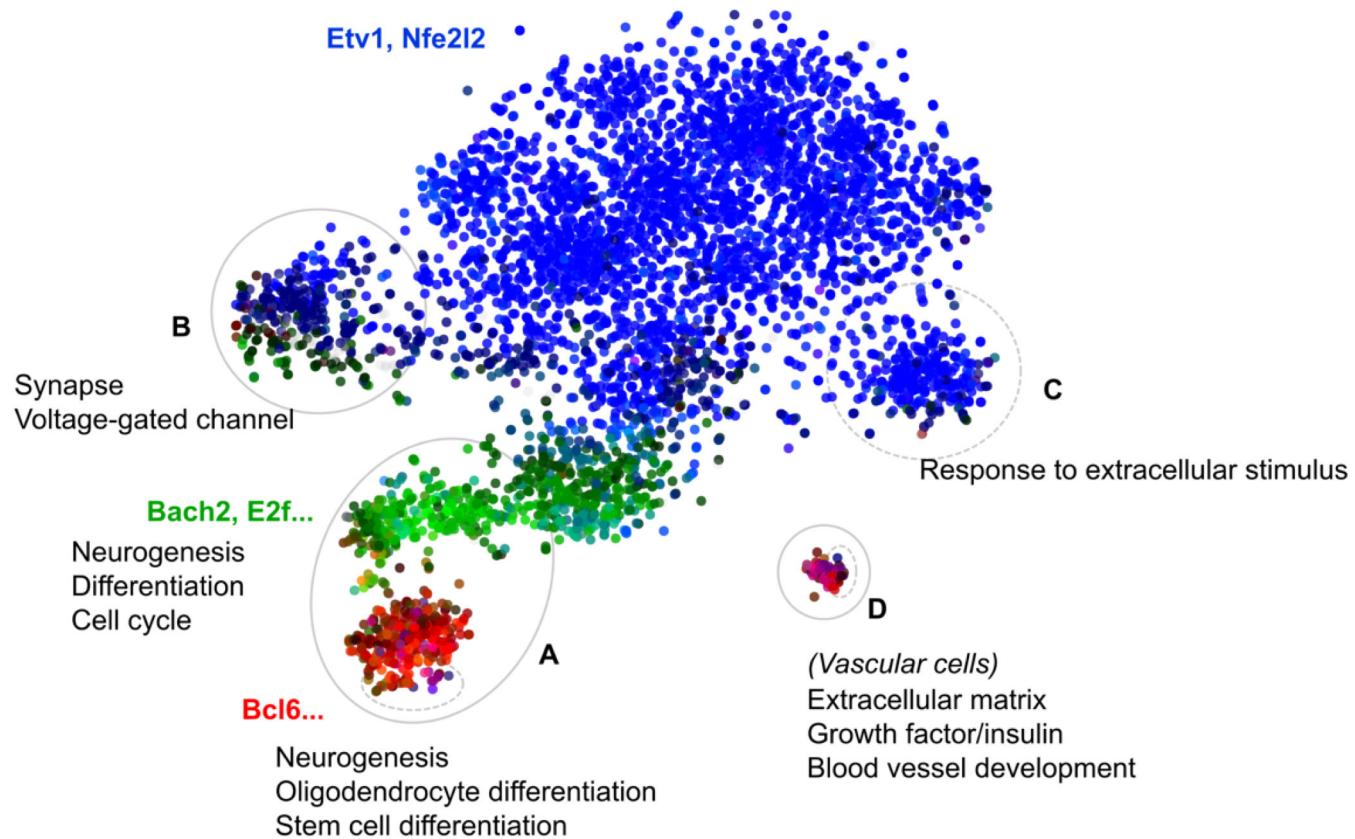
GRN-based clustering across species



Expression-based clustering:

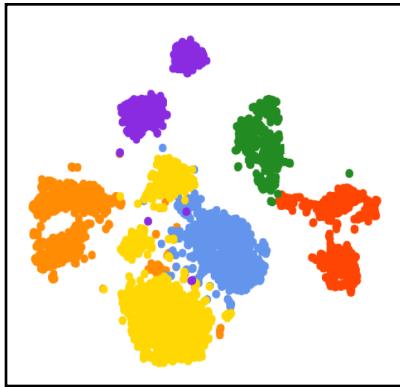


Application: 5069 oligodendrocytes

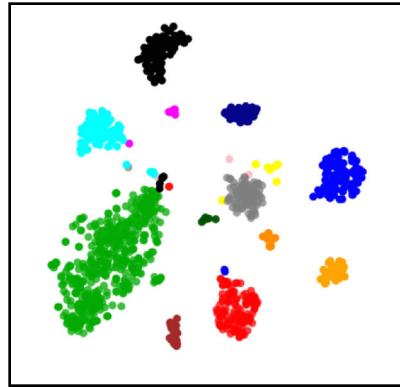


Application on cancer

Oligodendrogloma



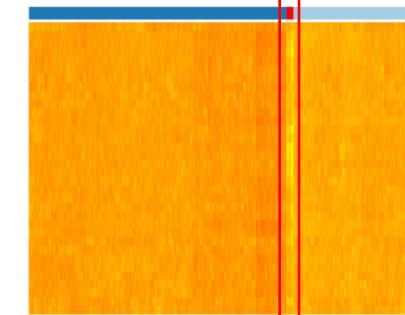
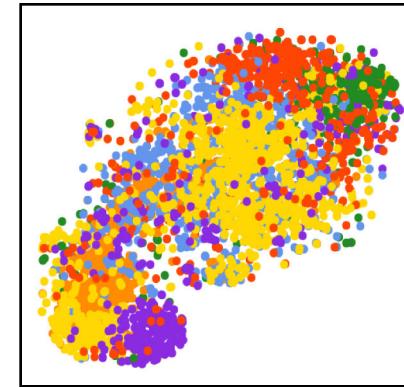
Melanoma



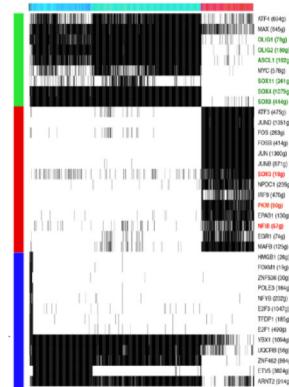
Expression-based tSNE

Color: Patient of origin

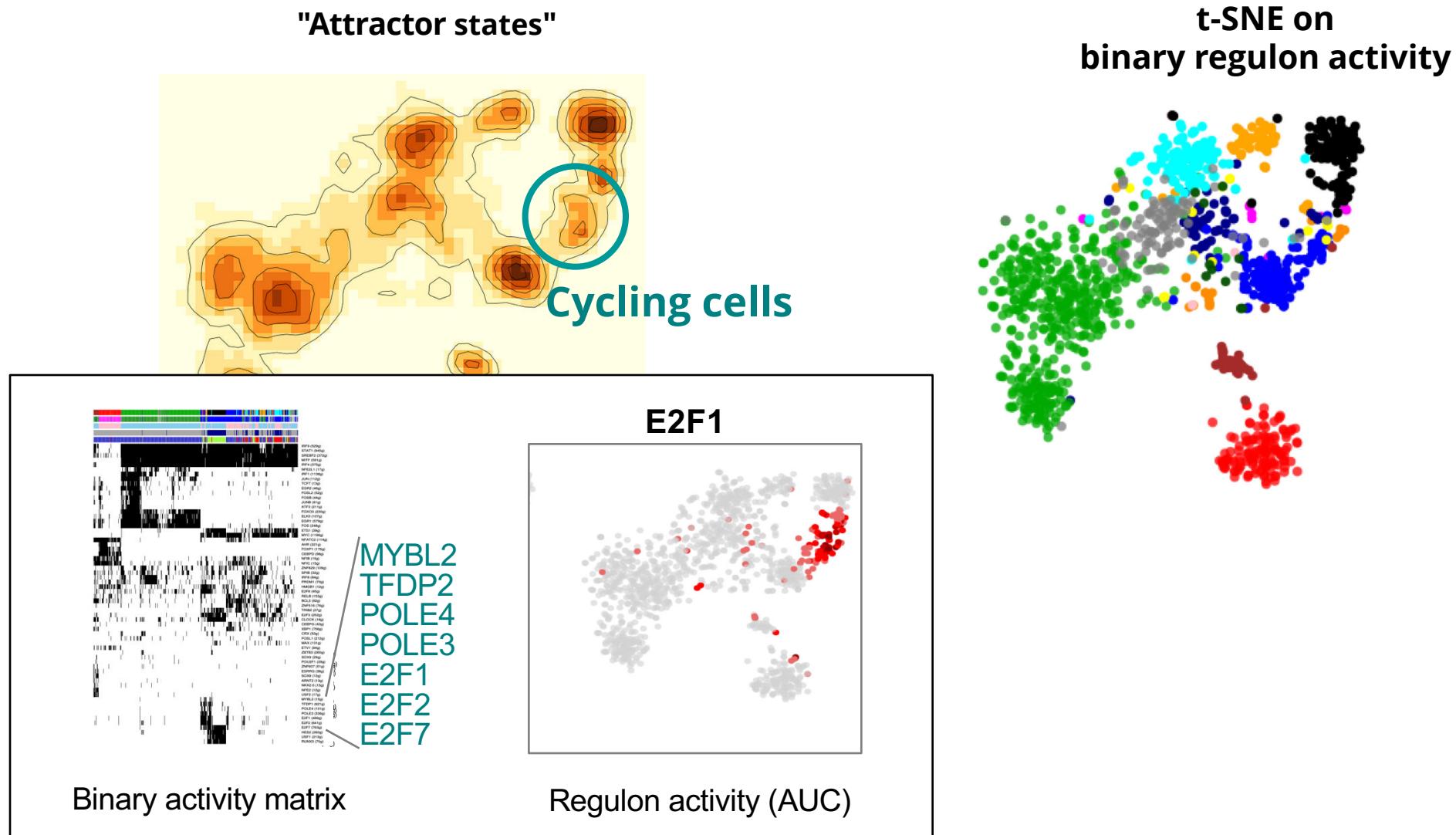
SCENIC



Cell-cycle gene sets

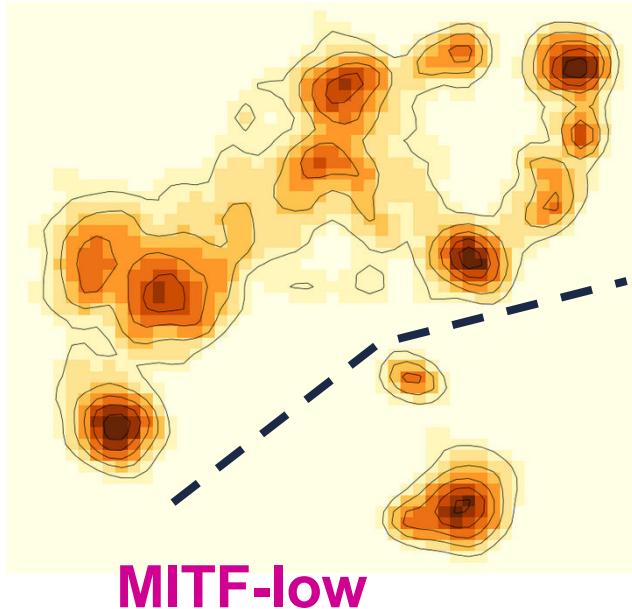


Application: Melanoma (1252 malignant cells)



Application: Melanoma (1252 malignant cells)

"Attractor states"



MITF-low

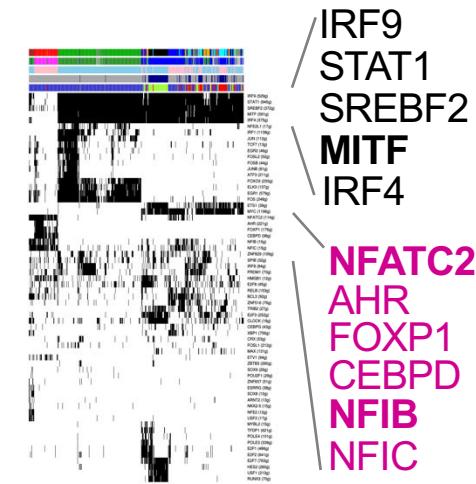
Enriched GO terms:

- Extracellular matrix
- Cell adhesion

Low:

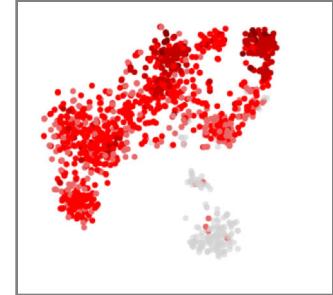
- Melanocyte differentiation
- Immune response

Module activity matrix

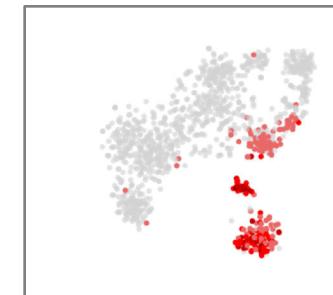


Module activity (AUC)

MITF

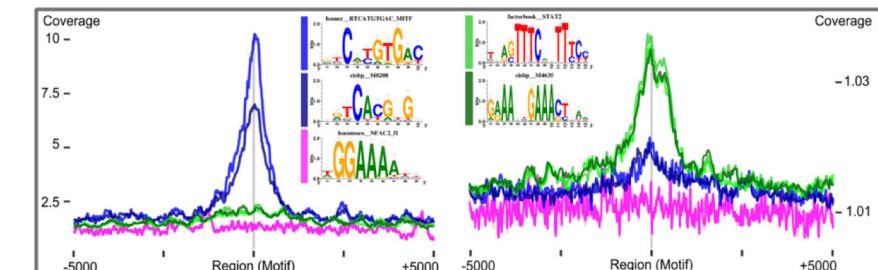


NFATC2



ChIP-seq

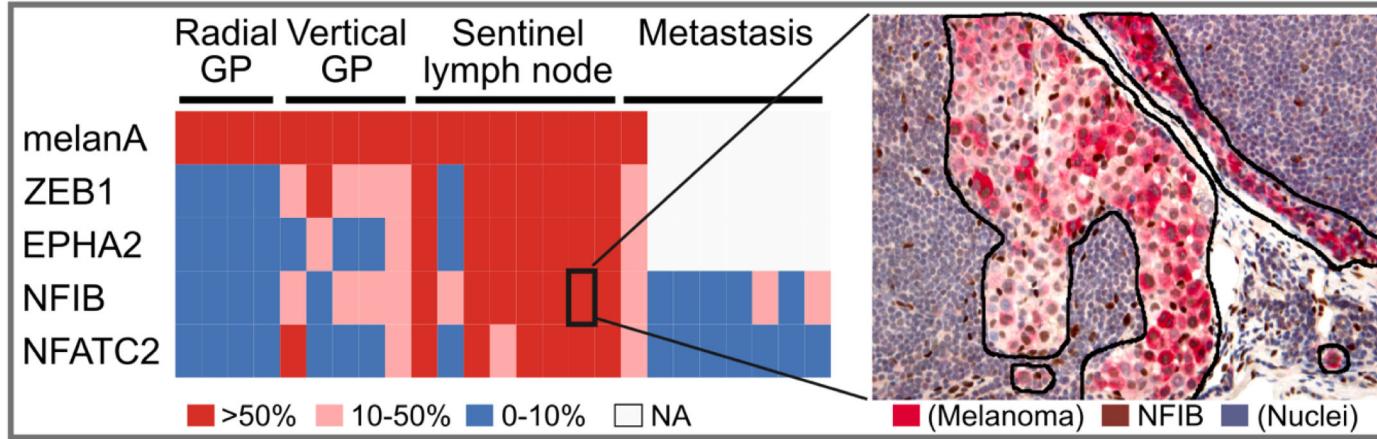
MITF



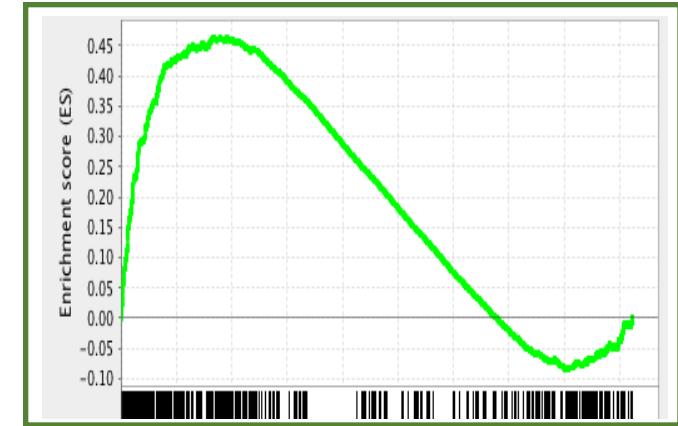
STAT

Application: Melanoma (1252 malignant cells)

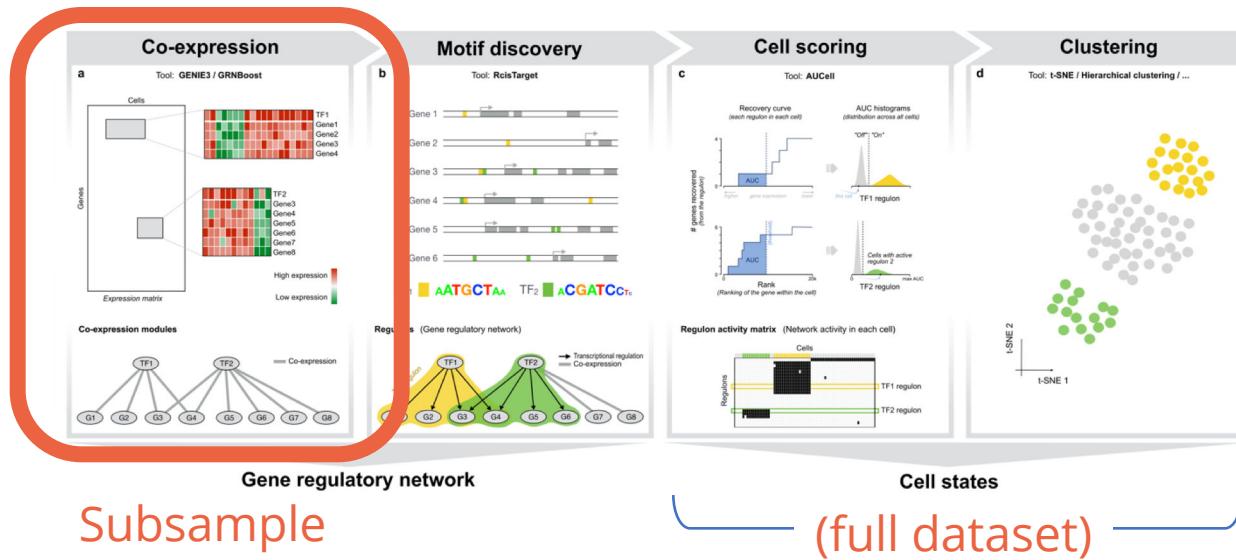
Immunohistochemistry



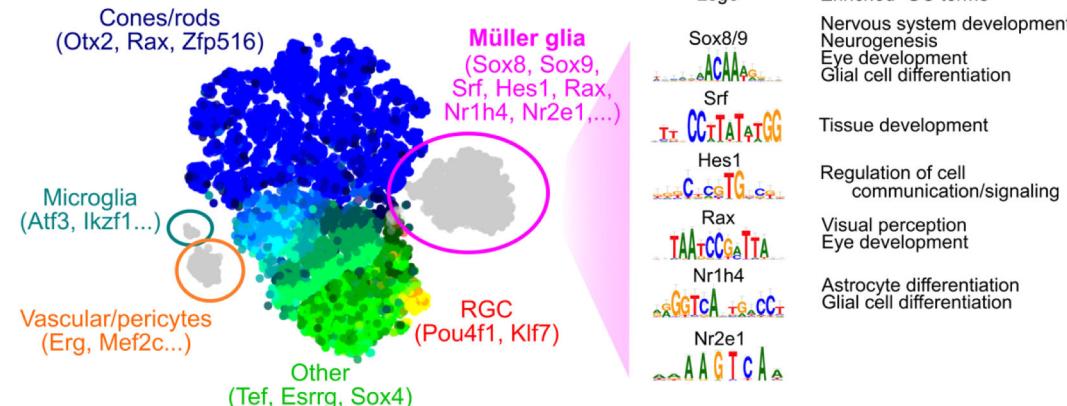
NFATC2 knock-down



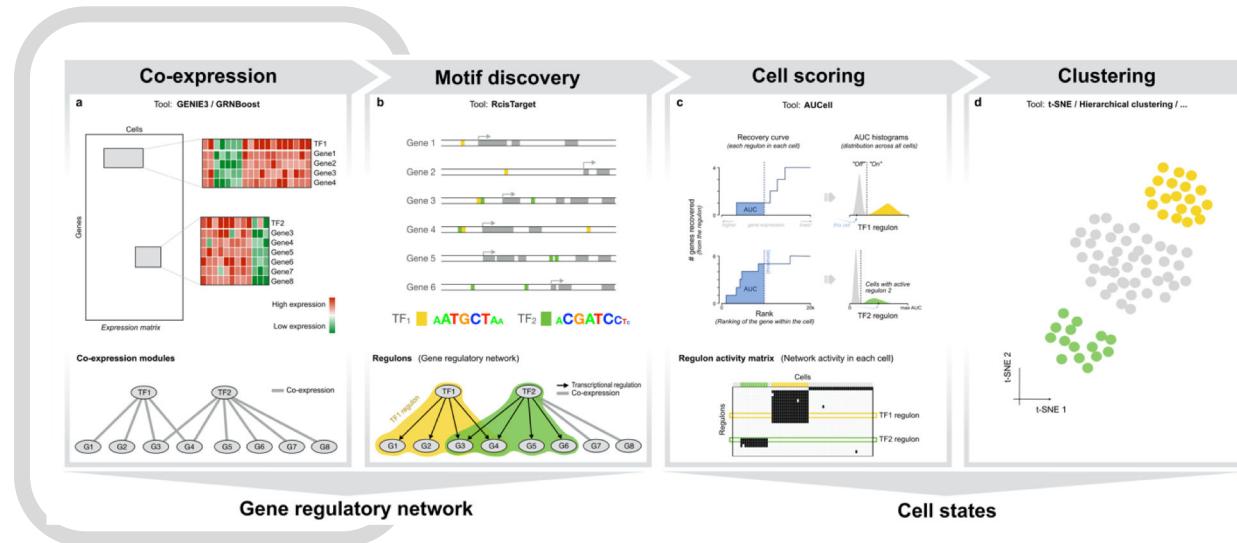
Application on large scale data sets



Drop-seq
49k mouse retina cells (Makosko et al. 2015)

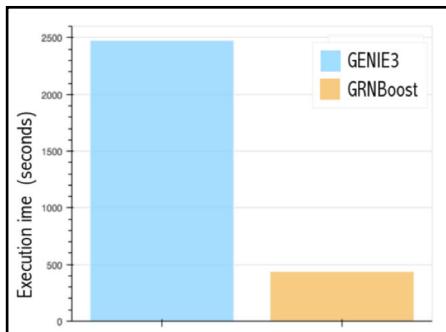


Application on large scale data sets



GRNboost (arboreto)

Regulator prediction based on gradient boosting

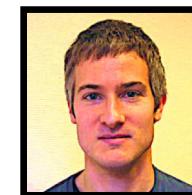


Thomas Moerman

GRNBoost2 and Arboreto: efficient and scalable inference of gene regulatory networks. Moerman et al. Bioinformatics 2018

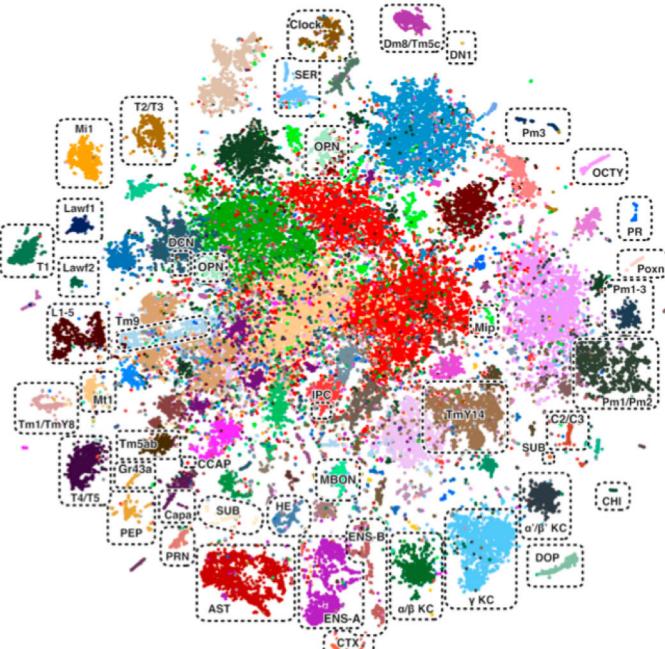
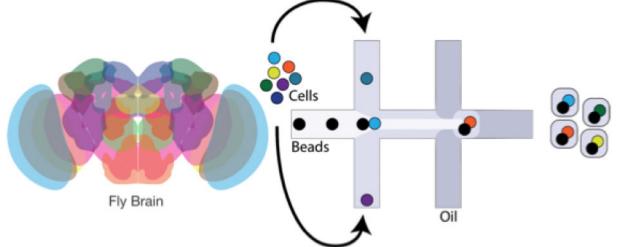
pySCENIC

Lightning-fast implementation in python



Bram Van De Sande

Since then...



<http://scope.aertslab.org>

A Single-Cell Transcriptome Atlas of the Aging Drosophila Brain

Kristofor Davie,^{1,2,9} Jasper Janssens,^{1,2,9} Duygu Koldere,^{1,2,9} Maxime De Waegeneer,^{1,2} Uli Pech,^{1,3} Lukasz Kreft,⁴ Sara Albar,^{1,2} Samira Makhzami,^{1,2} Valerie Christiaens,^{1,2} Carmen Bravo González-Blas,^{1,2} Suresh Poothongal,¹ Gert Hulselmans,^{1,2} Katina I. Spanier,^{1,2} Thomas Moerman,^{5,6} Bram Vanpouwelen,⁷ Sarah Geurs,² Thierry Voet,² Jeroen Lamertyn,⁷ Bernard Thienpont,² Sha Liu,^{1,3} Nikos Konstantinides,⁸ Mark Fiers,^{1,3} Patrik Verstreken,^{1,3} and Stein Aerts^{1,2,10,*}



Cel

Toward Minimal Residual Disease-Directed Therapy in Melanoma

Florian Rambow,^{1,2,15} Aljosja Rogiers,^{1,2,15} Oskar Marin-Bejar,^{1,2} Sara Albar,^{3,4} Julia Femel,⁵ Michael Dewaele,^{1,2} Panagiotis Karras,^{1,2} Daniel Brown,⁶ Young Hwan Chang,⁷ Marie Debbie-Rychter,⁸ Carmen Adriaens,^{1,2} Enrico Radella,^{1,2} Pascal Wolter,¹⁰ Oliver Bechtler,¹⁰ Reinhard Dummer,¹¹ Mitchell Levesque,¹¹ Adriano Piris,¹² Dennis T. Frederick,¹² Geneviève Boland,¹² Keith T. Flaherty,¹³ Joost van den Oord,¹⁴ Thierry Voet,⁶ Stein Aerts,^{3,4} Amanda W. Lund,⁵ and Jean-Christophe Maring^{1,2,16,17*}

Resource | Published: 09 July 201

nature
medicine

Phenotype molding of stromal cells in the lung tumor microenvironment

Diether Lambrechts , Els Wauters, Bram Boeckx, Sara Aibar, David Nittner, Oliver Burton, Ayse Bassez, Herbert Decaluwé, Andreas Pircher, Kathleen Van den Eynde, Birgit Weynand, Erik Verbeeken, Paul De Leyn, Adrian Liston, Johan Vansteenkiste, Peter Carmeliet, Stein Aerts & Bernard Thienpont 

> 65 citations

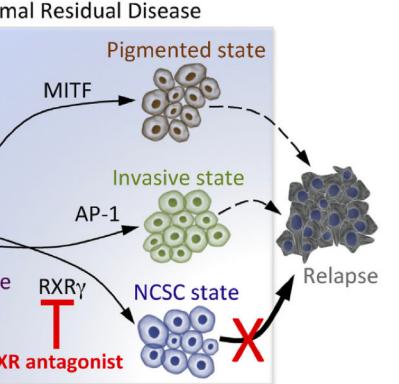
Cell Reports

Revealing the Critical Regulators of Cell Identity in the Mouse Cell Atlas

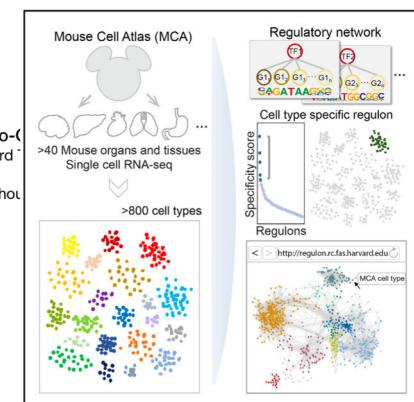
Shengbao Suo,¹ Qian Zhu,^{1,3} Assieh Saadatpour,^{1,3} Lijiang Fei,² Guoji Guo,² and Guo-Chen Cai^{1,3}
¹Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Harvard
T.H. Chan School of Public Health, Boston, MA, USA

Boston, MA 02215, USA
2Center for Stem Cell and Regenerative Medicine, Zhejiang University School of Medicine, Hangzhou,

<http://requlon.rc.fas.harvard.edu>



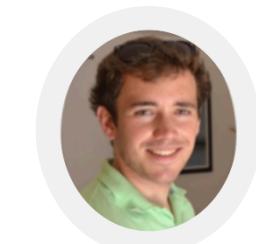
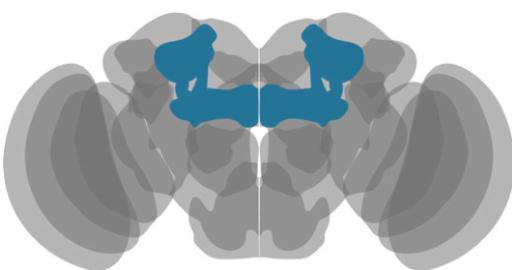
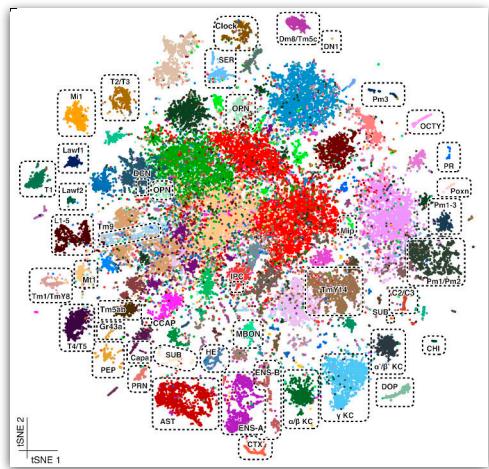
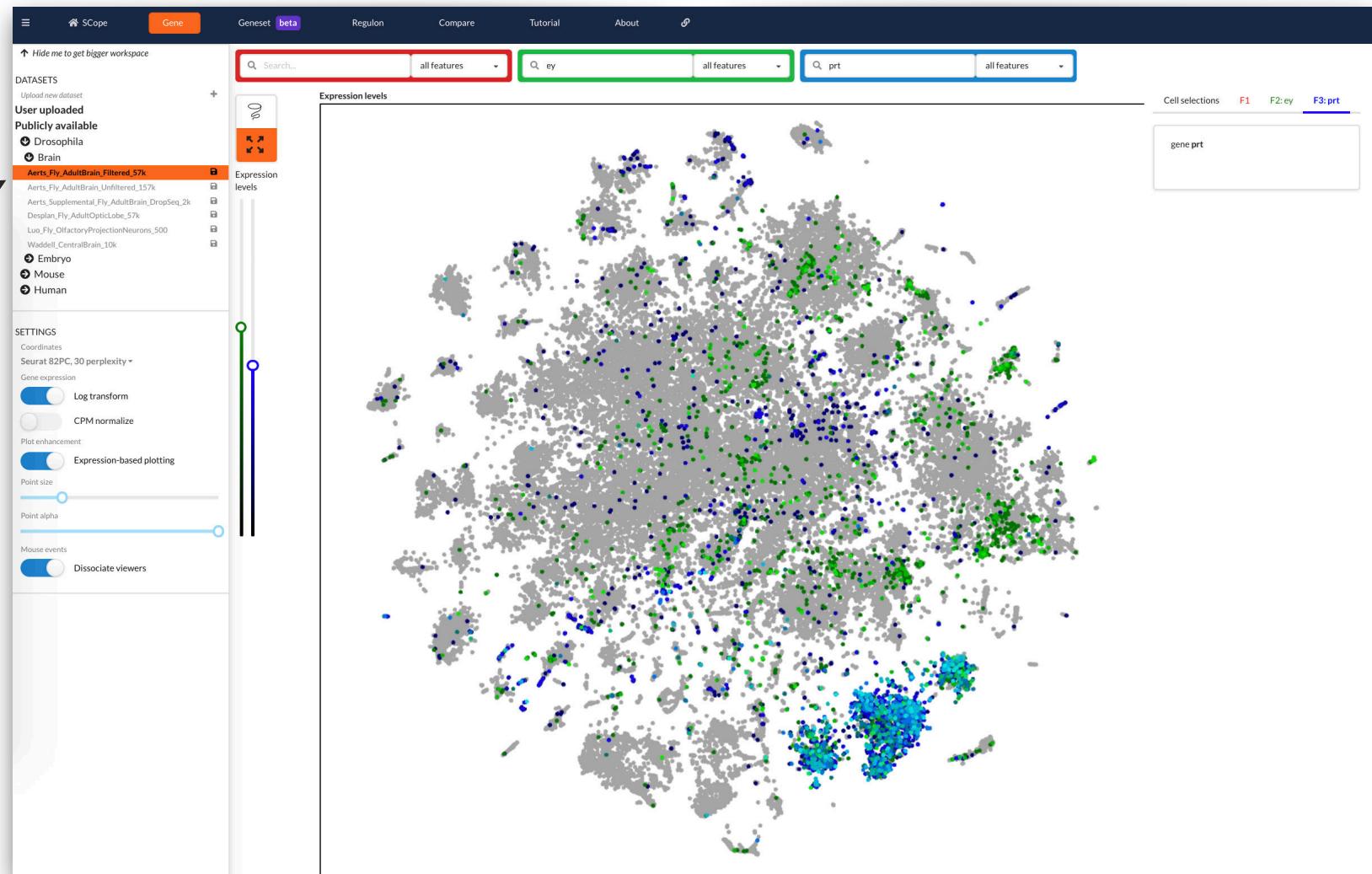
Report





Fly brain

+ Other datasets



Max



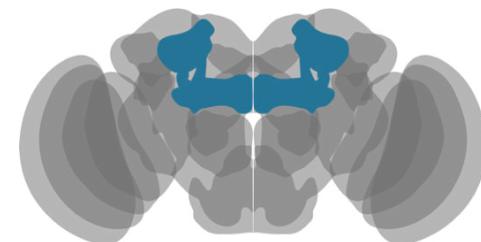
Kris



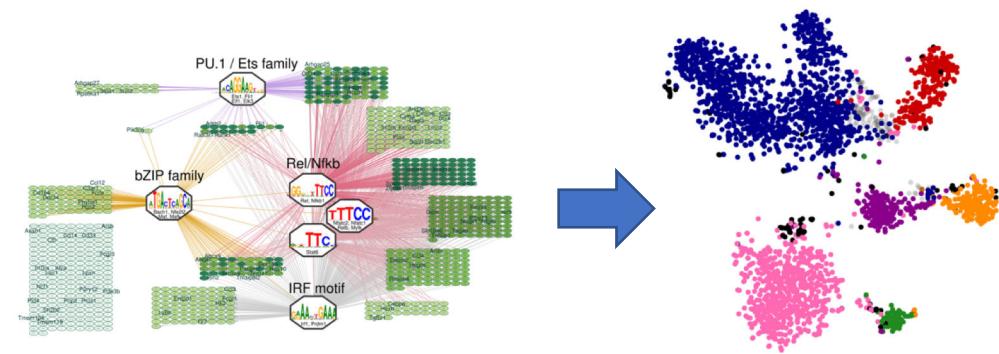
SCENIC regulons



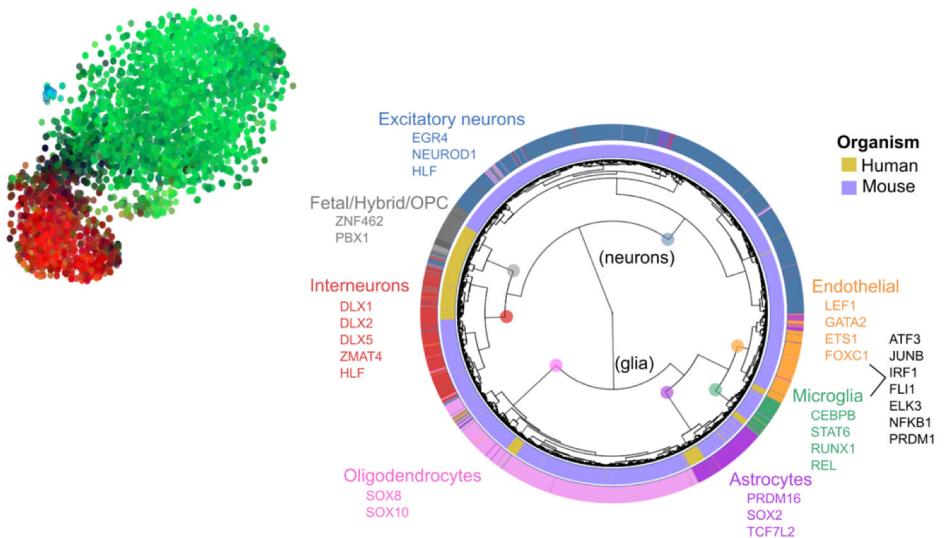
Repo regulon: **Glia** cell types
Ey regulon: **Kenyon cells**



Summary



- ✓ Identification of cell types and key regulators
- ✓ Biologically-driven dimensionality reduction



- ✓ Robust to batch effect and technical bias
- ✓ Data integration across datasets, technologies and species
- ✓ Overcomes tumor-effect

- ✓ Stable cell types/states
- ✓ Cancer / Disease/ Other types of states
- ✓ Analysis of continuous processes
- ✓ Large datasets

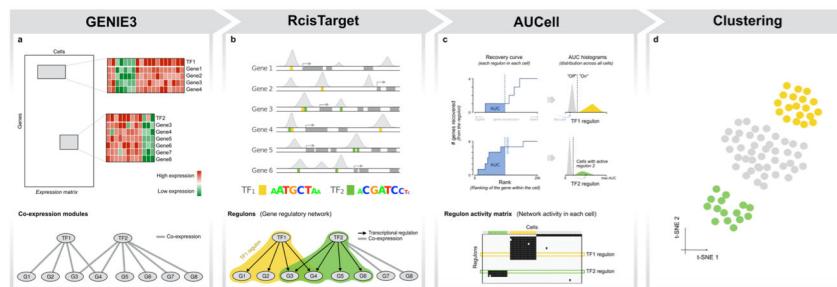
How to run it?

All links in scenic.aertslab.org

R implementation

- Very detailed tutorials
- Flexibility

SCENIC
SCopeLoomR
RcisTarget }
AUCell }

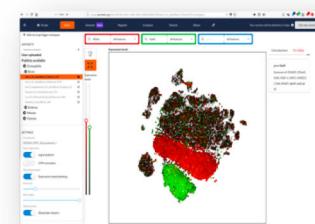


Python implementation

- “Lightning fast”
- pySCENIC
Arboreto>GRNBoost

Visualization

SCope



RcisTarget - Transcription factor binding motif analysis

Input:

Gene/region list (e.g. cluster markers, co-expressed genes)

Motif score database (provided for human, mouse and drosophila)

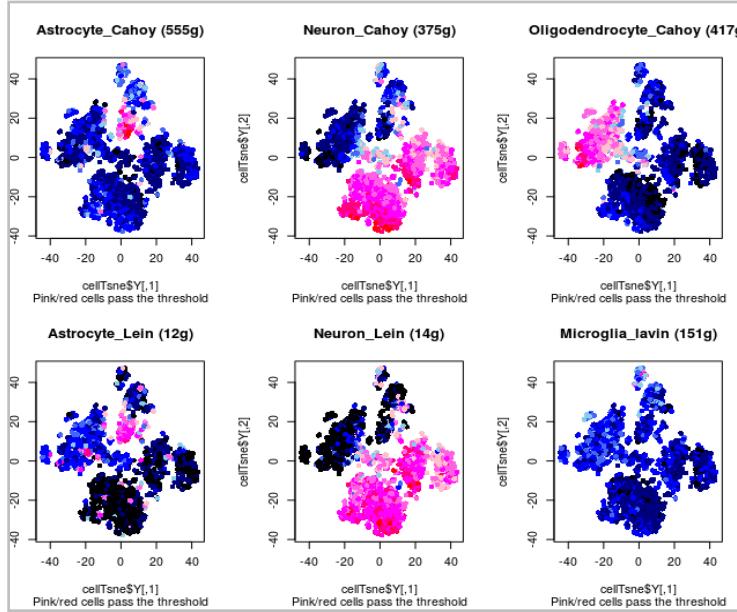
Show: 10 entries					
	motifDb	logo	NES	geneSet	TF_highConf
	All	All	All	All	All
1	500bp		4.44	Ahr_w001	Ahr; Ahrr; Arnt; Arnt2; Hif1a (directAnnotation).
2	500bp		3.13	Arnt2_top10perTarget	Atf3 (inferredBy_Orthology).
3	500bp		3.83	Arnt2_top50perTarget	Bhlhe40 (inferredBy_Orthology).
4	500bp		3.54	Arnt2_top50perTarget	Bhlhe40 (inferredBy_Orthology).

<https://bioconductor.org/packages/RcisTarget>

AUCell - Identifying cells with active gene sets

Input:

Gene list (e.g. known cell type markers, pathway)
Expression matrix



<https://bioconductor.org/packages/AUCell>