

Sc-Type: A novel method for prioritization of genes for scRNA-seq based automatic cell-type labeling

Available at <http://sctype.fimm.fi/>

6th October 2022

Dr. Anil K Giri
Finland Institute for Molecular Medicine



scRNA-seq has accelerated cell types identification

nature
biotechnology

Low-coverage single-cell mRNA sequencing reveals
cellular heterogeneity and activated signaling pathways
in developing cerebral cortex

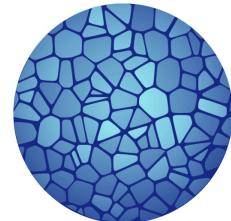
Alex A Pollen^{1,2,4}, Tomasz J Nowakowski^{1,2,4}, Joe Shuga^{3,4}, Xiaohui Wang^{3,4}, Anne A Leyrat³, Jan H Lui^{1,2}, Nianzhen Li³, Lukasz Szpankowski³, Brian Fowler³, Peilin Chen³, Naveen Ramalingam³, Gang Sun³, Myo Thu³, Michael Norris³, Ronald Lebofsky³, Dominique Toppani³, Darnell W Kemp II³, Michael Wong³, Barry Clerkson³, Brittnee N Jones³, Shiquan Wu³, Lawrence Knutsson³, Beatriz Alvarado³, Jing Wang³, Lesley S Weaver³, Andrew P May³, Robert C Jones³, Marc A Unger³, Arnold R Kriegstein^{1,2} & Jay A A West³

ARTICLE

<https://doi.org/10.1038/s41586-019-1373-2>

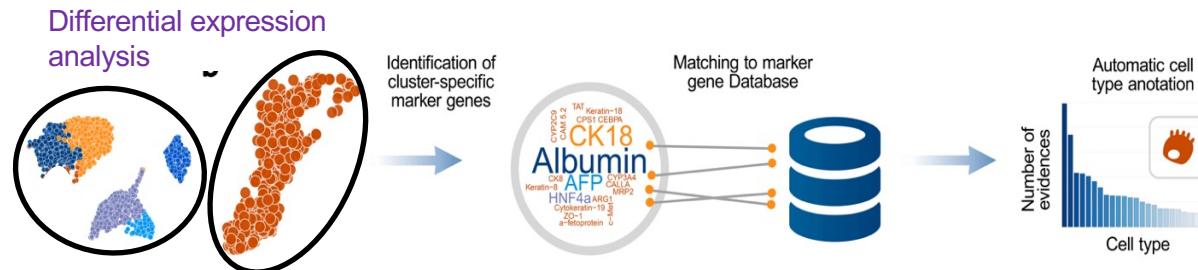
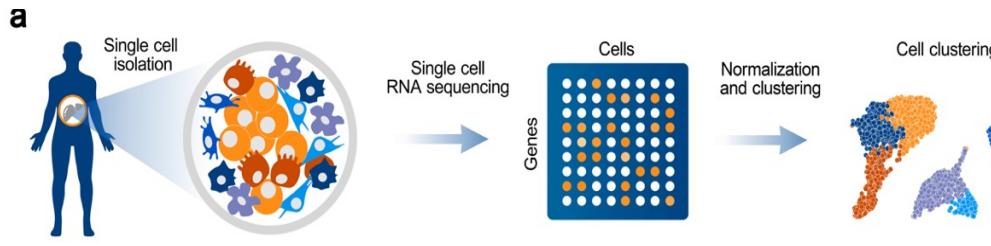
A human liver cell atlas reveals heterogeneity and epithelial progenitors

Nadim Aizaranj^{1,2,3}, Antonio Saviano^{4,5,6,8}, Sagar^{1,8}, Laurent Mailly^{4,5}, Sarah Durand^{4,5}, Josip S. Herman^{1,2,3},
Patrick Pessaix^{4,5,6}, Thomas F. Baumert^{4,5,6*} & Dominic Grün^{1,7*}



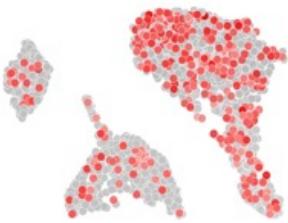
HUMAN
CELL
ATLAS

Cell types identification using scRNA-seq



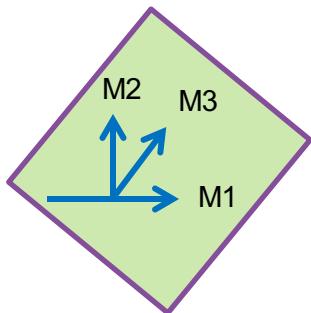
Limitation of the genes prioritized based on DEA

Gene 1 (Low spec.)



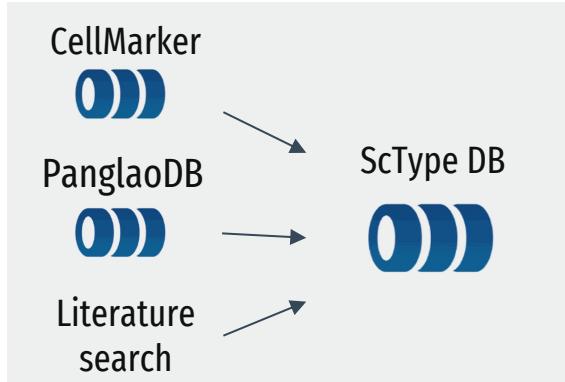
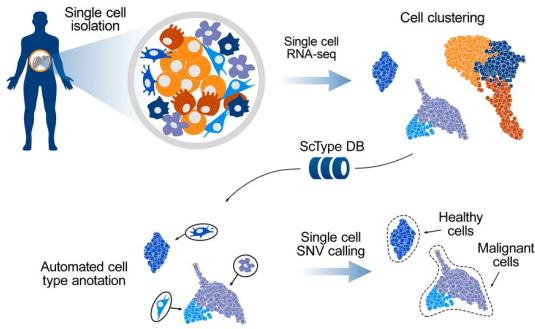
Gene specificity across cell types, s_{type}

Cell type	Gene 1 (Low spec.)	Gene 2 (High spec.)	Known cell type marker:
Blue bird	●	●	● True
Orange circle	●	●	● False
Purple flower	●	●	● False
Yellow star	●	●	● True
Red flame	●	●	● False



- Use of limited number of markers to assign cell type

ScType database



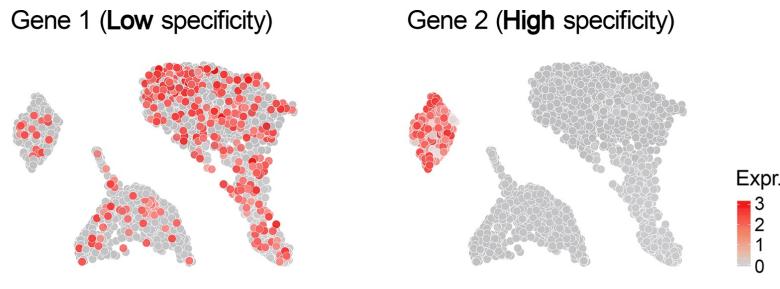
ScType comprises markers for all major cell types from **17** tissues



3980 markers for **194** human cell types
4212 markers for **194** mouse cell types

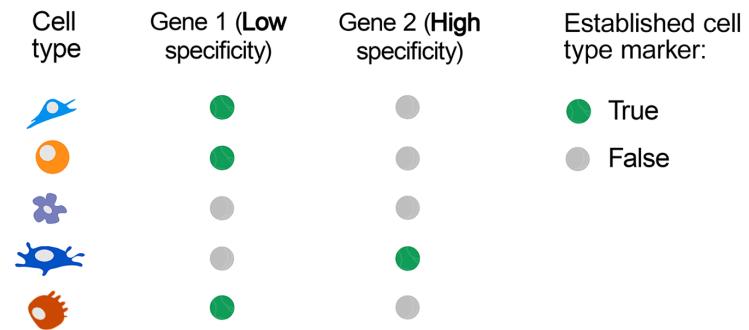
Two aspects of marker specificity

Across clusters



The higher score is given to genes specifically expressed in 1 cluster

Across cell types



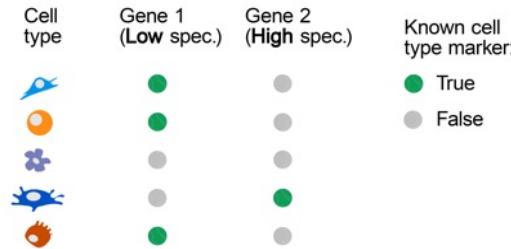
The higher score is given to genes being an established marker of 1 cell type

Sc-Type algorithm to select marker gene and cell annotation

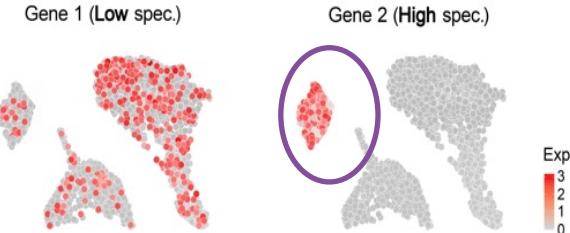
Selection of top 30 positively expressed marker



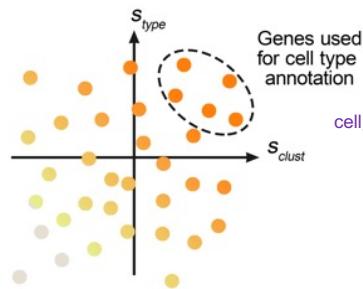
Gene specificity across cell types, s_{type}



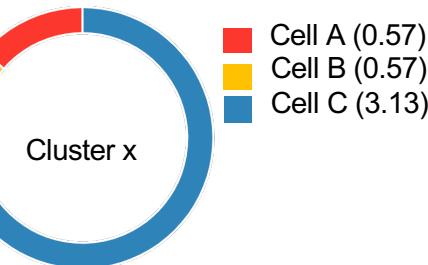
Gene specificity across clusters, s_{clust}



Combined gene specificity

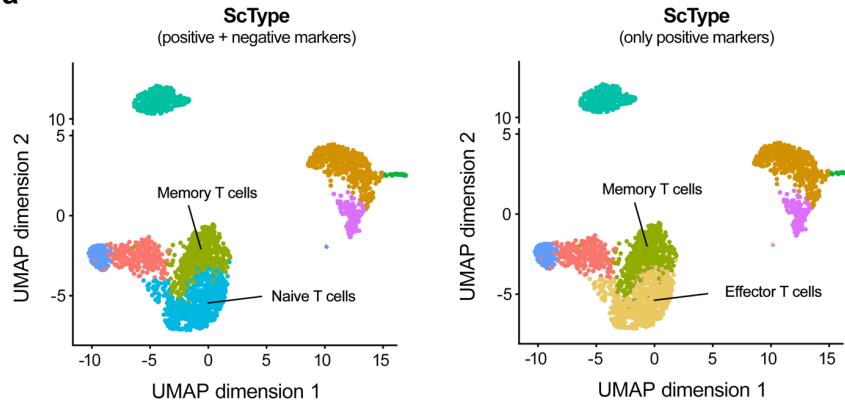


$$cellS_{scstype} \text{ score} = \sum_1^n \sqrt{(S_{clust} * S_{type})}$$



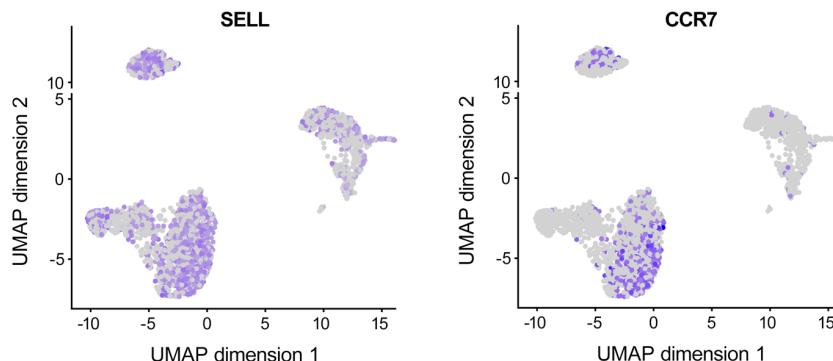
ScType utilizes both positive and negative markers

a



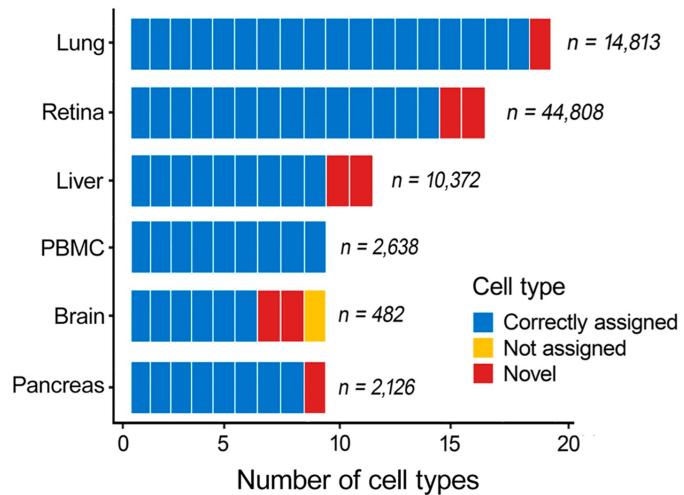
ScType allows using *negative marker genes*, which provide evidence against a cell being of a particular type to distinguish between very closely-related subtypes.

b



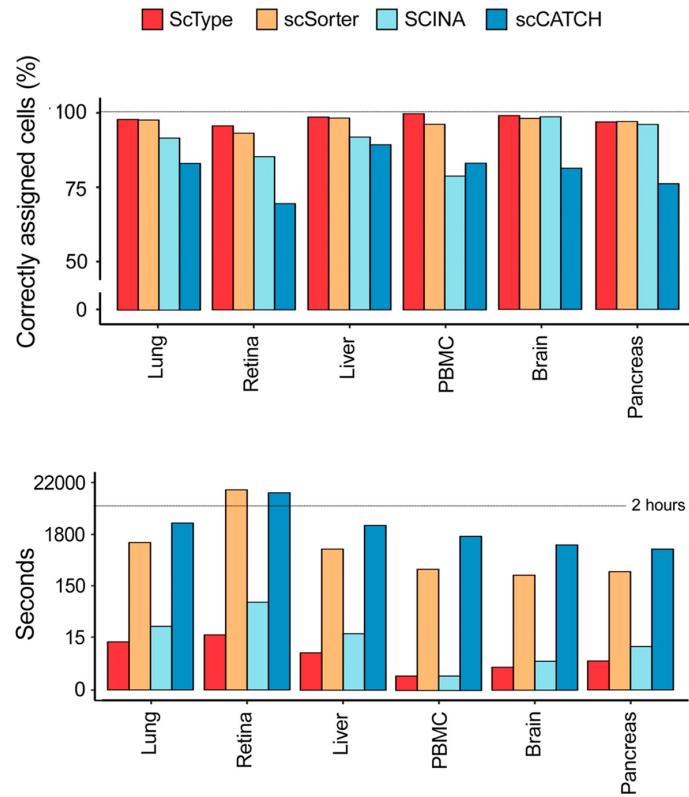
Expression of CCR7 and SELL genes, which are required for lymph node migration, are not expected to be expressed in the effector T cells (i.e., negative markers).

Overall ScType performance

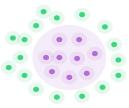


ScType correctly annotated a total of **72** cell types out of **73** cell-types (98.6% accuracy), including **8** correctly reannotated cell-types that were originally incorrectly or non-specifically annotated

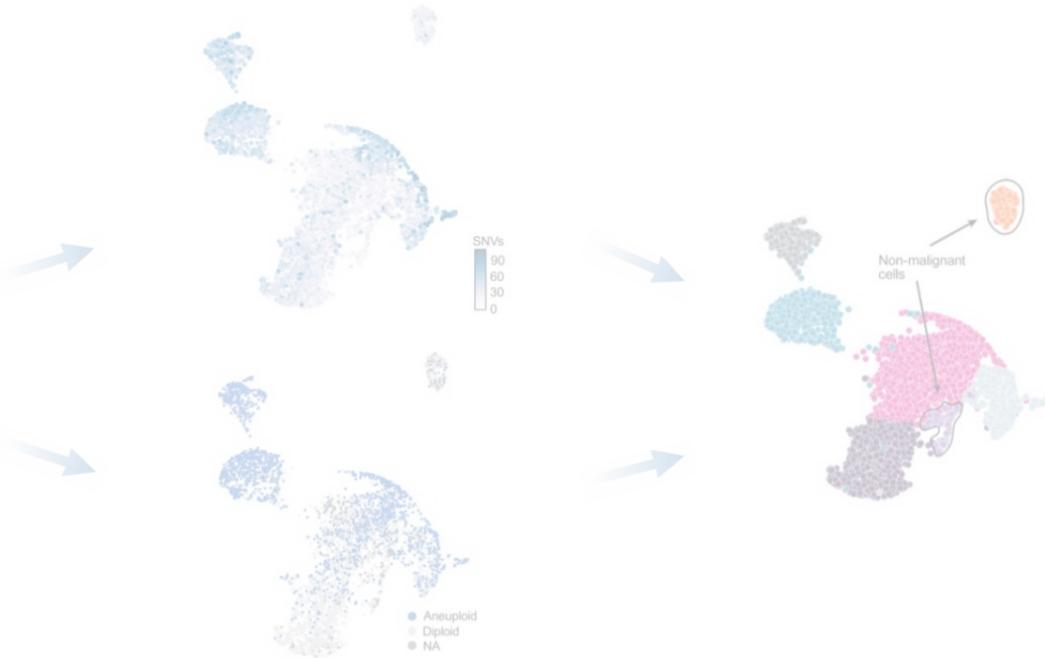
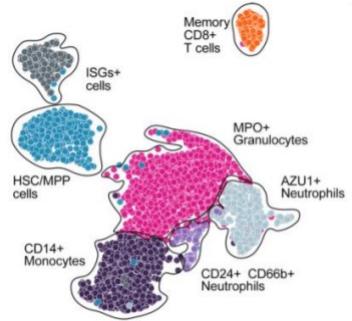
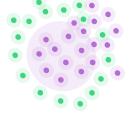
Comparison of ScType to other methods



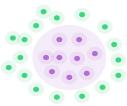
ScType distinguishes between malignant and “healthy” cells



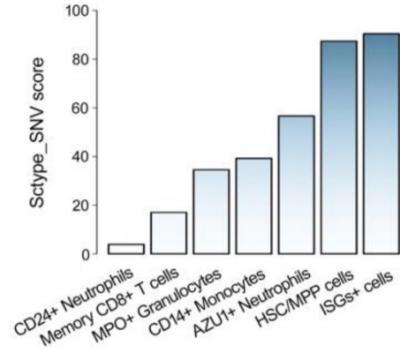
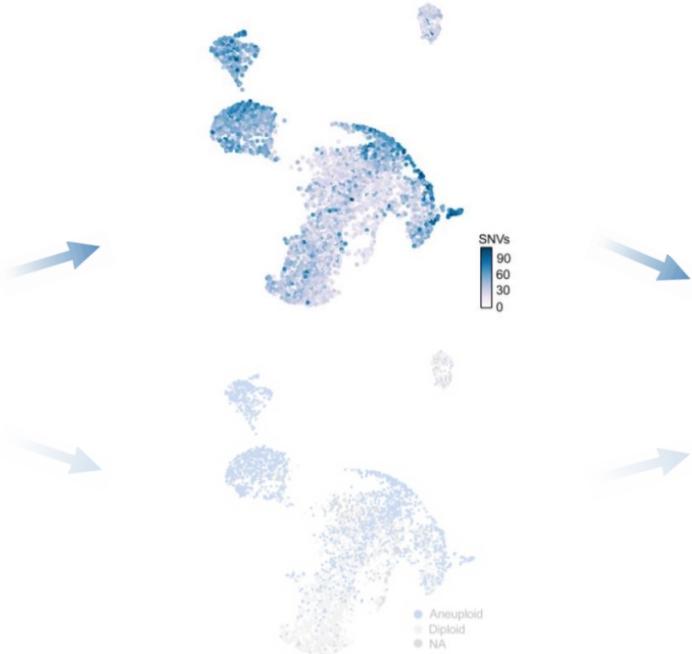
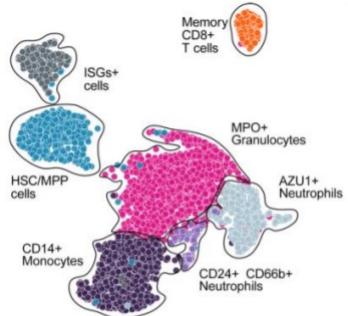
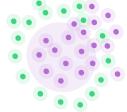
VS.



ScType distinguishes between malignant and “healthy” cells

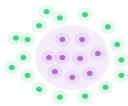


VS.

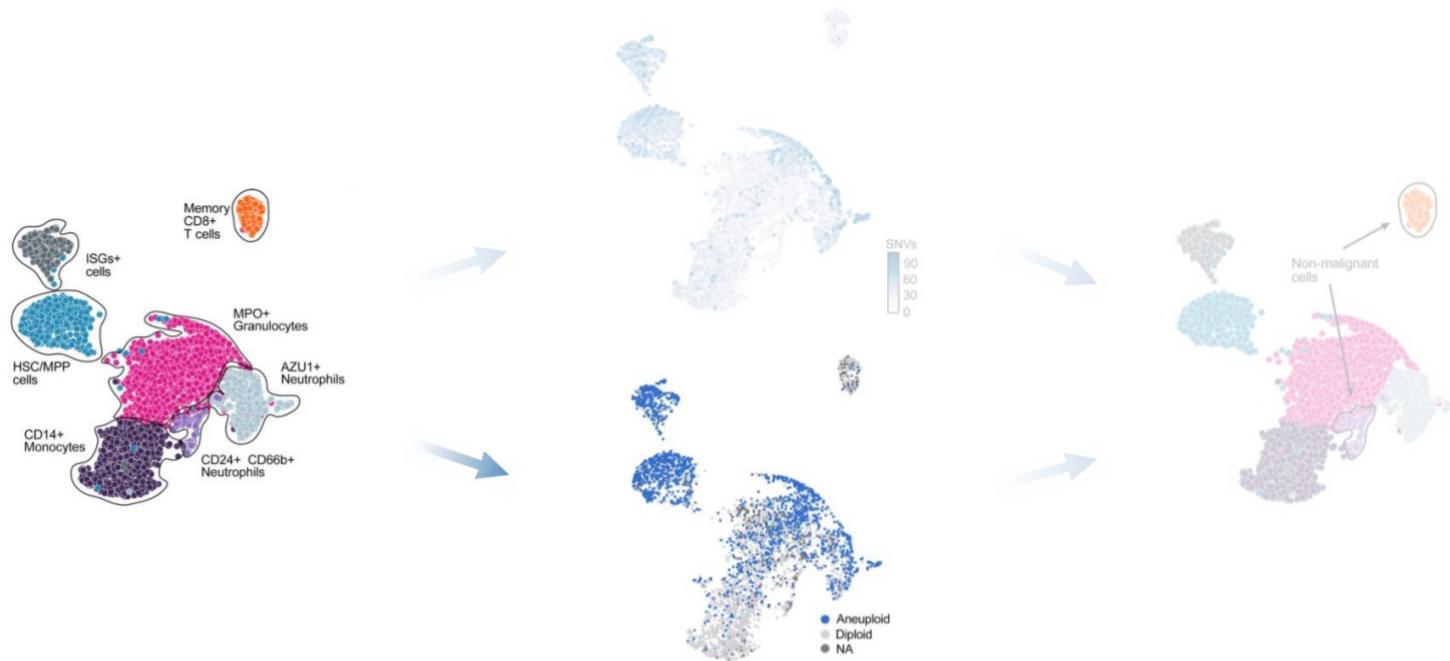


ScType calculates the sum of total number of SNV present in the COSMIC cancer census genes as the combined SNV score (ScType_SNV score) for each cluster in a cancerous tissue profile

ScType distinguishes between malignant and “healthy” cells

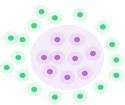


VS.

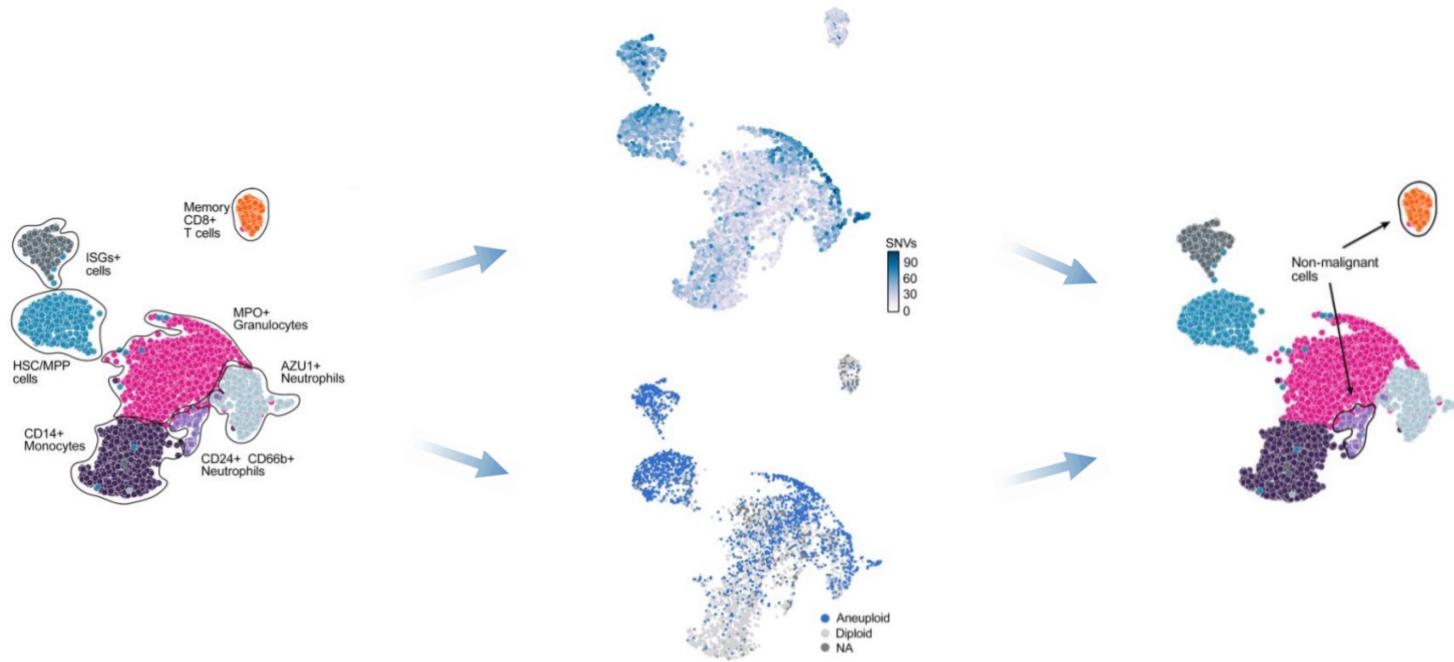
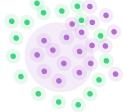


ScType also incorporates a recently implemented Bayesian segmentation approach, called CopyKAT, to distinguish between aneuploid and diploid cells. Aneuploidy (presence of an abnormal number of chromosomes in a cell) common (88%) in most human tumors CNV calling

ScType distinguishes between malignant and “healthy” cells

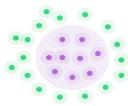


VS.

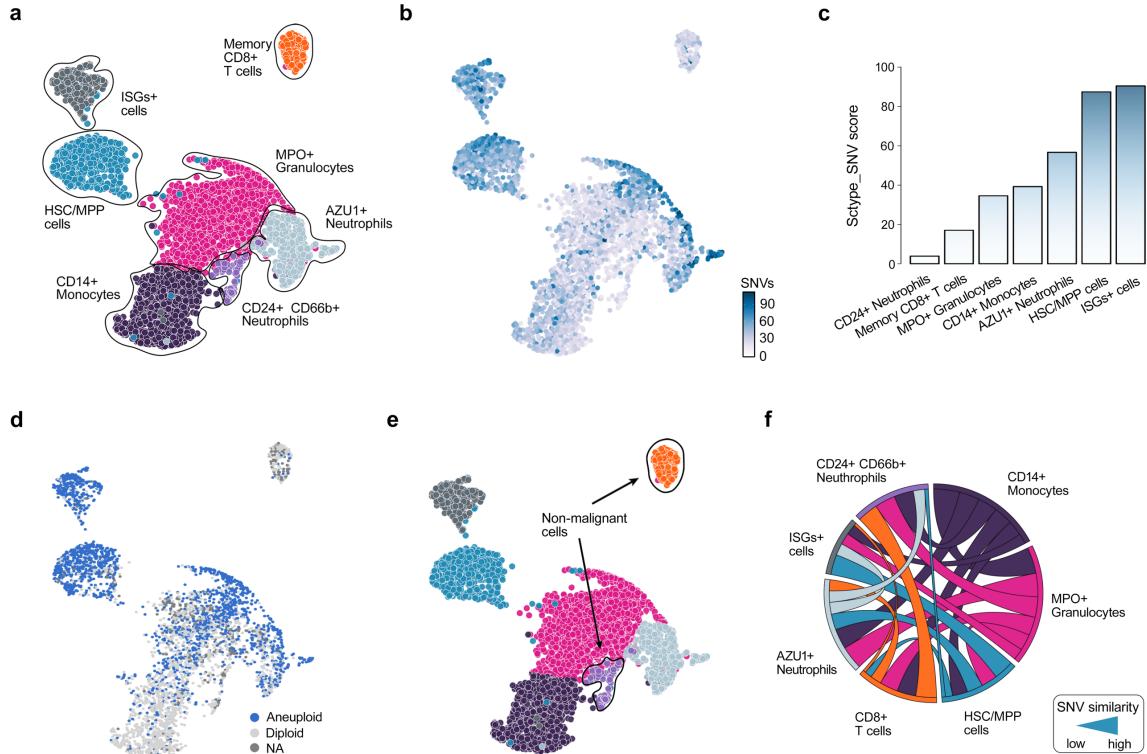
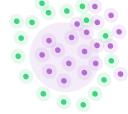


ScType automatically assigns cells as non-malignant if ScType_SNV score < 20 and the majority of cell within the celltype (>50%) are classified as diploid and the others as malignant.

ScType distinguishes between malignant and “healthy” cells



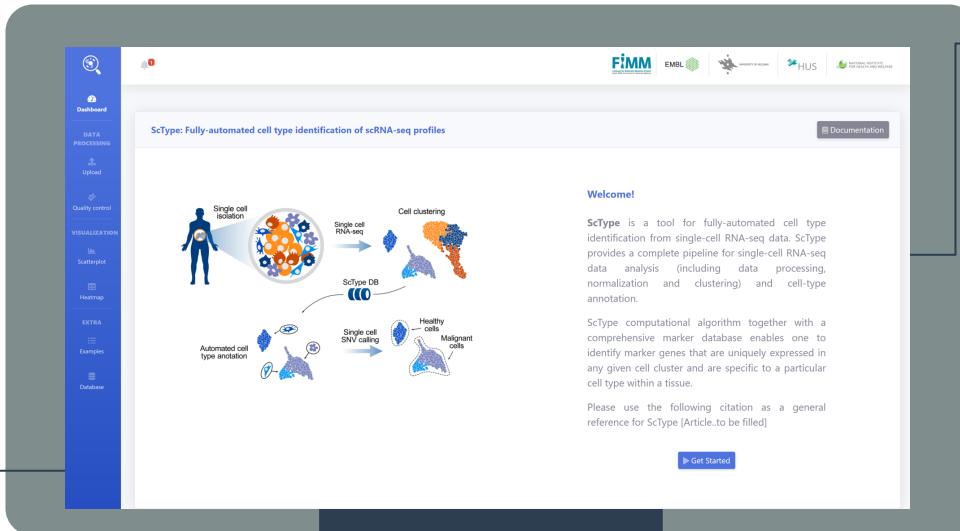
VS.



ScType web-application

sctype.app

Platform
independent



Interactive
analysis

Thank you!