



CyteGuide: Visual Guidance for Hierarchical Single-Cell Analysis

T. Höllt♥♠ N. Pezzotti♥ V. van Unen♠
F. Koning♠ B.P.F. Lelieveldt♠♥ A. Vilanova♥

♥ TU Delft

♠ Leiden University MC

A decorative graphic at the bottom of the slide consists of multiple thin, wavy lines in shades of orange and yellow, creating a sense of movement and depth.



CyteGuide: Visual Guidance for Hierarchical Single-Cell Analysis

T. Höllt♥♠ N. Pezzotti♥ V. van Unen♠
F. Koning♠ B.P.F. Lelieveldt♠♥ A. Vilanova♥

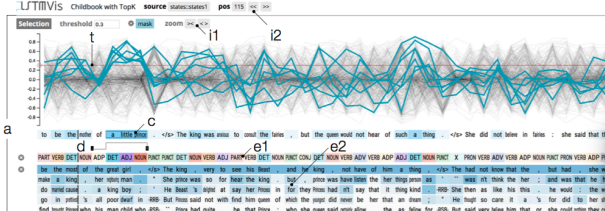
♥ TU Delft

♠ Leiden University MC

A decorative graphic at the bottom of the slide consists of multiple thin, wavy lines in shades of orange and yellow, creating a sense of movement and depth.

LSTMVis: A Tool for Visual Analysis of Hidden State Dynamics in Recurrent Neural Networks

Hendrik Strobel, Sebastian Gehrmann, Hanspeter Pfister, and Alexander M. Rush



ACTiVis: Visual Exploration of Industry-Scale Deep Neural Network Models

Minsuk Kahng, Pierre Y. Andrews, Aditya Kalra, and Duen Horng (Polo) Chau

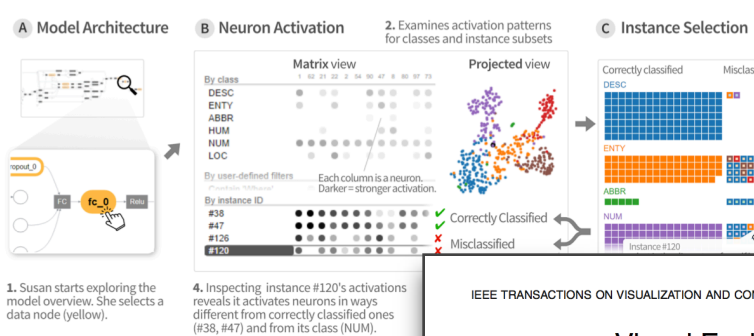


Fig. 1. ACTiVis integrates several coordinated views to support exploration and analysis of industry-scale deep neural network models.

Visual Exploration of Semantic Relationships in Neural Word Embeddings

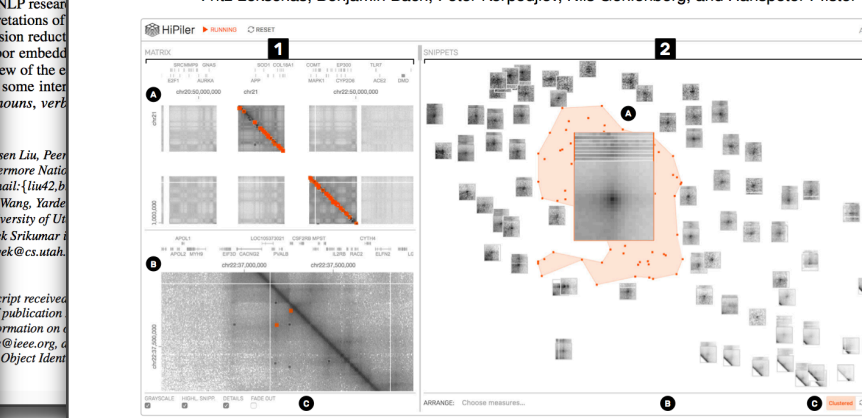
Shusen Liu, Peor-Tim Bremer, Jayaraman J. Thiagarajan, Vivek Srikrumar, Bei Wang, Yarden Liatov and Valerio Pasoucci

Abstract—Constructing distributed representations for words through neural language models and using the resulting vector spaces for analysis has become a crucial component of natural language processing.

Neural Networks (DNNs) have shown outstanding performance in various tasks, including image and speech recognition.

HiPiler: Visual Exploration of Large Genome Interaction Matrices with Interactive Small Multiples

Fritz Lekschas, Benjamin Bach, Peter Kerpedjiev, Nils Gehlenborg, and Hanspeter Pfister

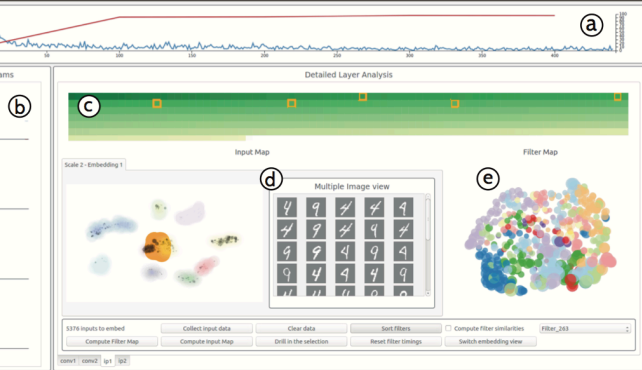


Abstract—Single-cell visualization through mass cytometry has become an increasingly important tool for immunologists to study the immune system in health and disease.

Abstract—Dimension reduction algorithms and clustering algorithms are both frequently used techniques in visual analytics.

DeepEyes: Progressive Visual Analytics for Designing Deep Neural Networks

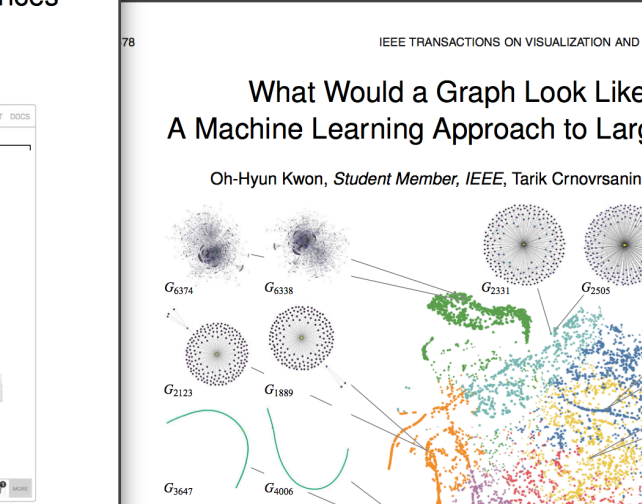
Nicola Pezzotti, Thomas Höllt, Jan van Gemert, Boudewijn P.F. Lelieveldt, Elmar Eisemann, Anna Vilanova



Abstract—DeepEyes is a Progressive Visual Analytics system for the analysis of deep neural networks during training.

LDSScanner: Exploratory Analysis of Low-Dimensional Structures in High-Dimensional Datasets

Jiazhi Xia, Feijin Yi, Wei Chen, Yusi Wang, Welfeng Chen, Yuxin Ma, and Anthony K.H. Tung



Abstract—Exploratory analysis of low-dimensional structures in high-dimensional datasets is a challenging task.

Abstract—Using different methods for laying out a graph can lead to very different visual appearances, with which the viewer perceives different information.

Skeleton-based Scagnostics

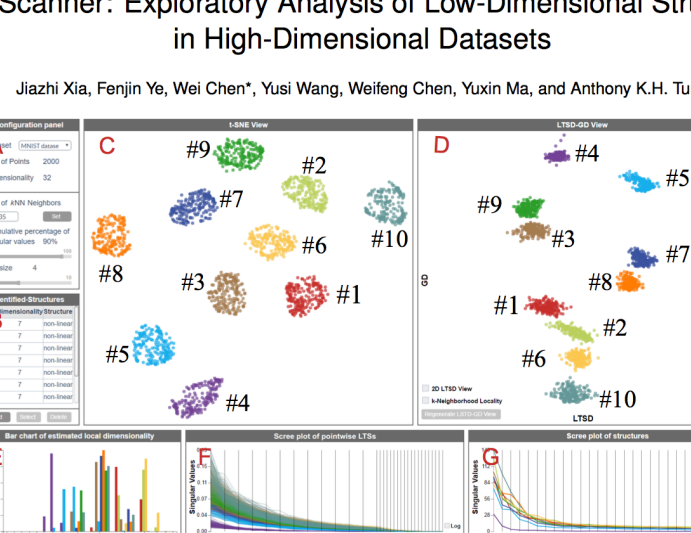
José Matute, Alexandru C. Telea, and Lars Linsen

Abstract—Scatterplot matrices (SPLOMs) are widely used for exploring multidimensional data.

Abstract—The oldest, and still frequently used, method for visualizing multidimensional data is the scatterplot matrix.

VIGOR: Interactive Visual Exploration of Graph Query

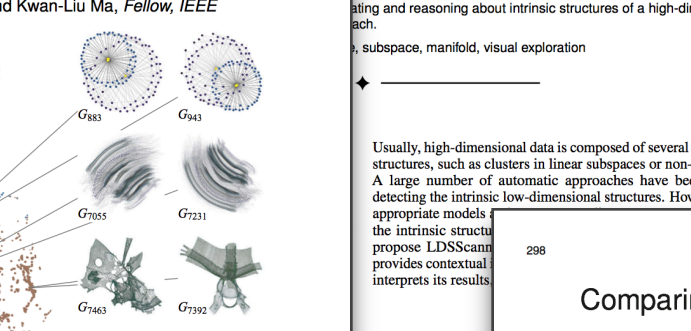
Robert Pienta, Fred Hohman, Alex Endert, Acar Tamersoy, Kevin Rounoy, Chris Gates, Shankant Navathe, Duen Horng Chau



Abstract—The tree boosting, which combines weak learners (typically decision trees) to generate a strong learner.

SkyLens: Visual Analysis of Skyline on Multi-dimensional Data

Xun Zhao, Yanhong Wu, Weiwei Cui, Xinnan Du, Yuan Chen, Yong Wang, Di Lu, Lun Lee, and Huamin Qu

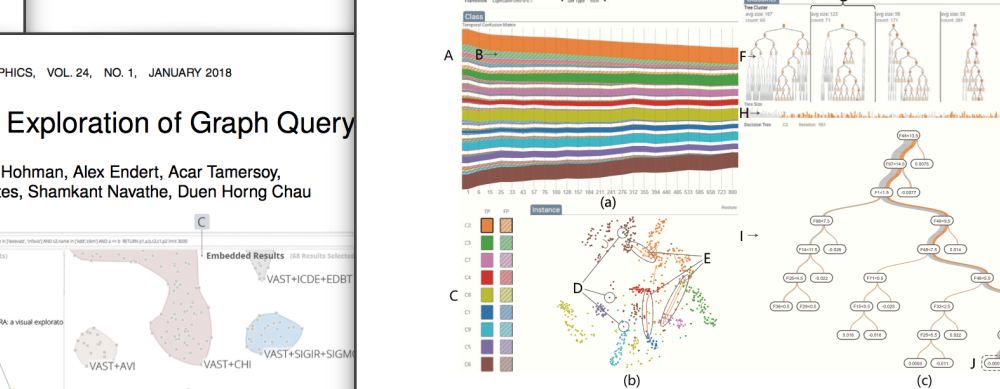


Abstract—Skyline queries have wide-ranging applications in fields that involve multi-criteria decision making.

Abstract—Labeling data instances is an important task in machine learning and visual analytics.

Visual Diagnosis of Tree Boosting Methods

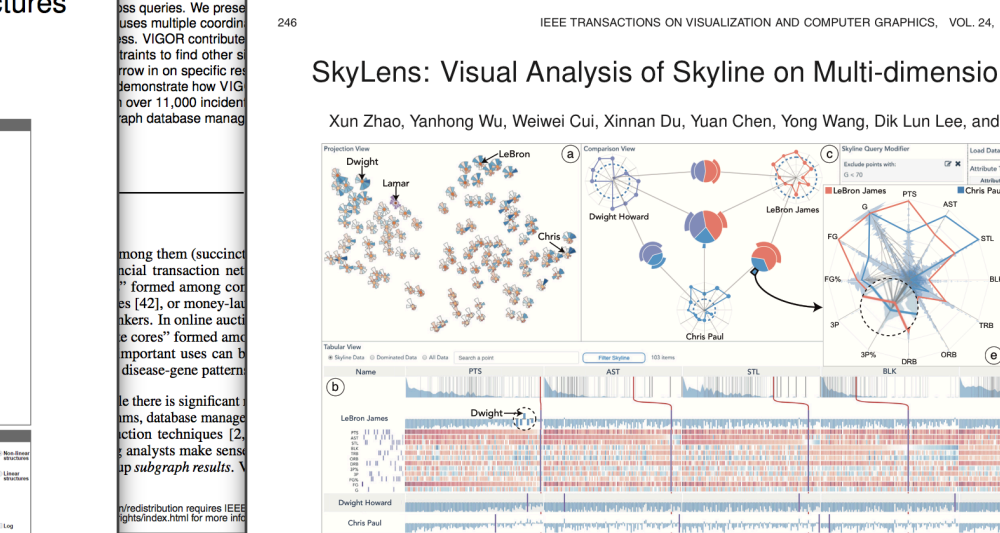
Shixia Liu, Jiantan Xiao, Junlin Liu, Xing Wang, Jing Wu, Jun Zhu



Abstract—The tree boosting, which combines weak learners (typically decision trees) to generate a strong learner.

ConceptVector: Text Visual Analytics via Interactive Lexicon Building using Word Embedding

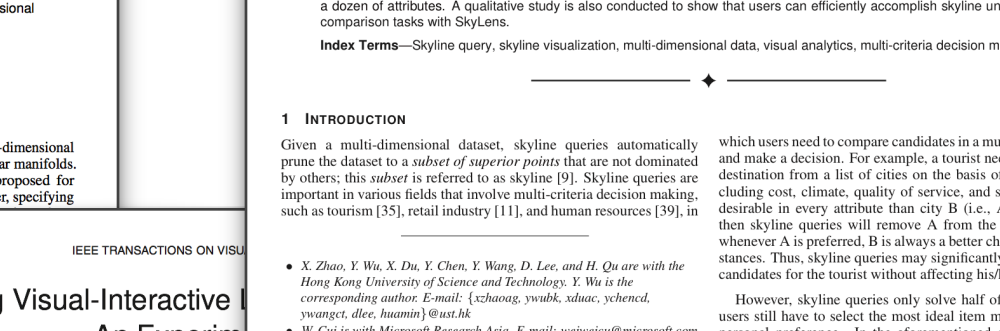
Deokgun Park, Seungyeon Kim, Jurim Lee, Jaegul Choo, Nicholas Diakopoulos, and Niklas Elmqvist, Senior Member, IEEE



Abstract—Central to many text analysis methods is the notion of a concept: a set of semantically related keywords characterizing a specific object, phenomenon, or theme.

What Would a Graph Look Like in This Layout? A Machine Learning Approach to Large Graph Visualization

Oh-Hyun Kwon, Student Member, IEEE, Tarik Cirovsanin, and Kwan-Liu Ma, Fellow, IEEE

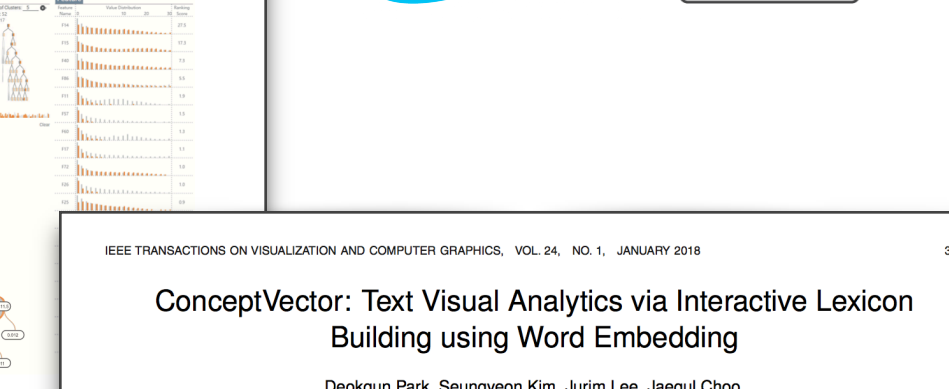


Abstract—Using different methods for laying out a graph can lead to very different visual appearances, with which the viewer perceives different information.

Abstract—Labeling data instances is an important task in machine learning and visual analytics.

Visual Supervision of Unsupervised Clustering

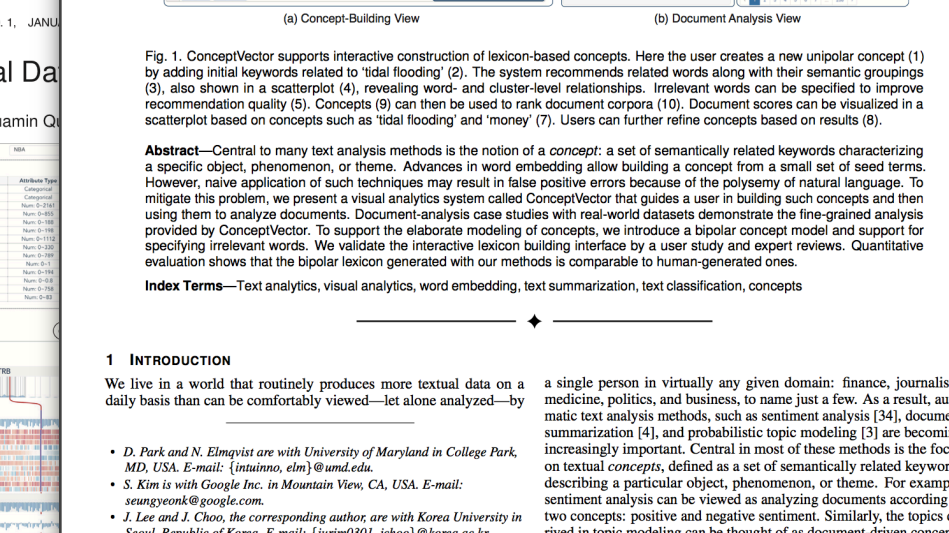
Bun Chul Kwon, Ben Eysenbach, Jana Verma, Ng, Christoffer deFlippi, Walter F. Stewart, and Adam Perer



Abstract—Clustering, the process of grouping together similar items into distinct partitions, is a common type of unsupervised machine learning that can be useful for summarizing and aggregating complex multi-dimensional data.

Comparing Visual-Interactive Labeling with an Experiment

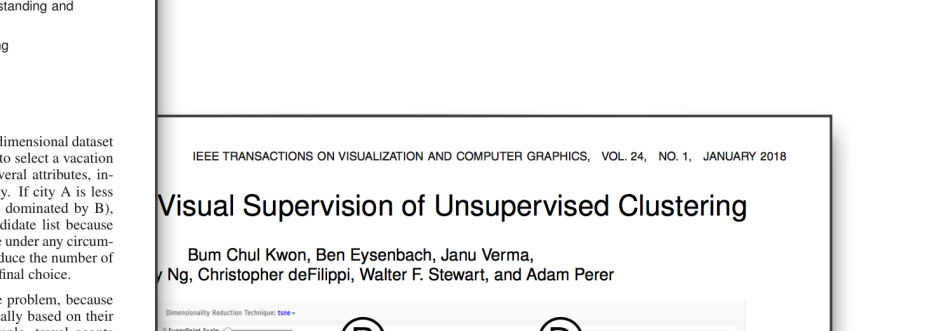
Jürgen Bernard, Marco Hutter, Matthias Zepp



Abstract—Labeling data instances is an important task in machine learning and visual analytics.

Towards a Systematic Combination of Dimension Reduction and Clustering in Visual Analytics

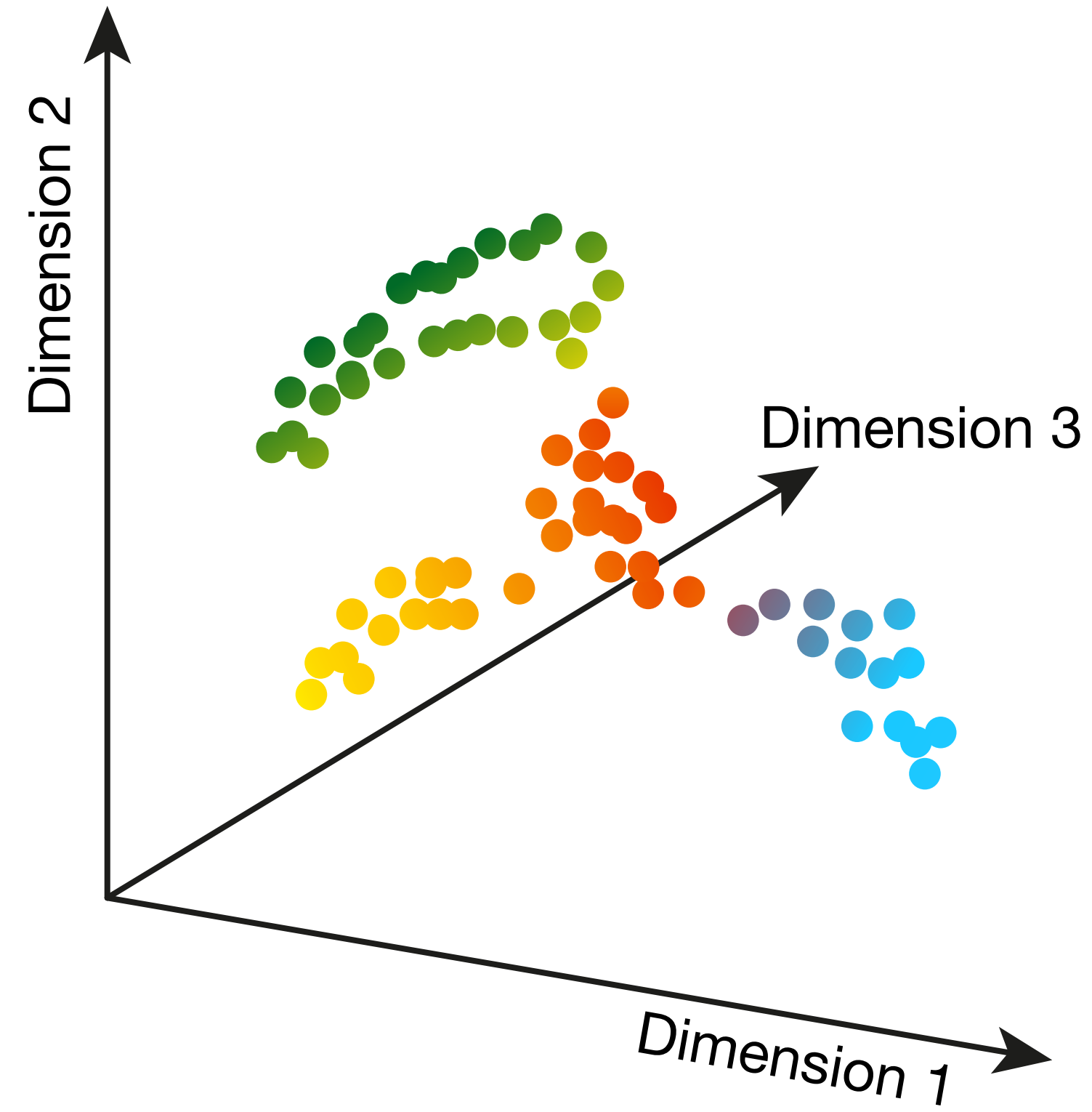
John Wenskovich, Student Member, IEEE, Ian Crandell, Naren Ramakrishnan, Member, IEEE, Leanna Howe, Scotland Laird, Chris North



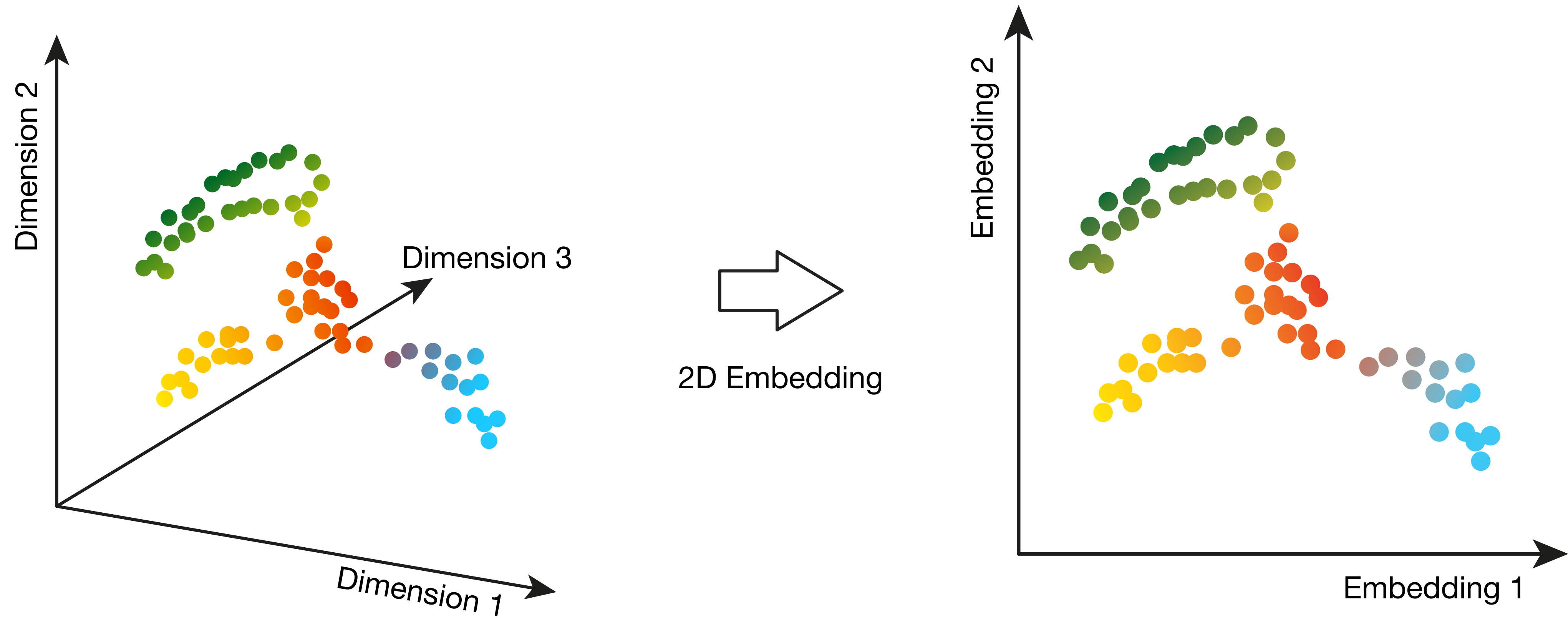
Abstract—Dimension reduction algorithms and clustering algorithms are both frequently used techniques in visual analytics.

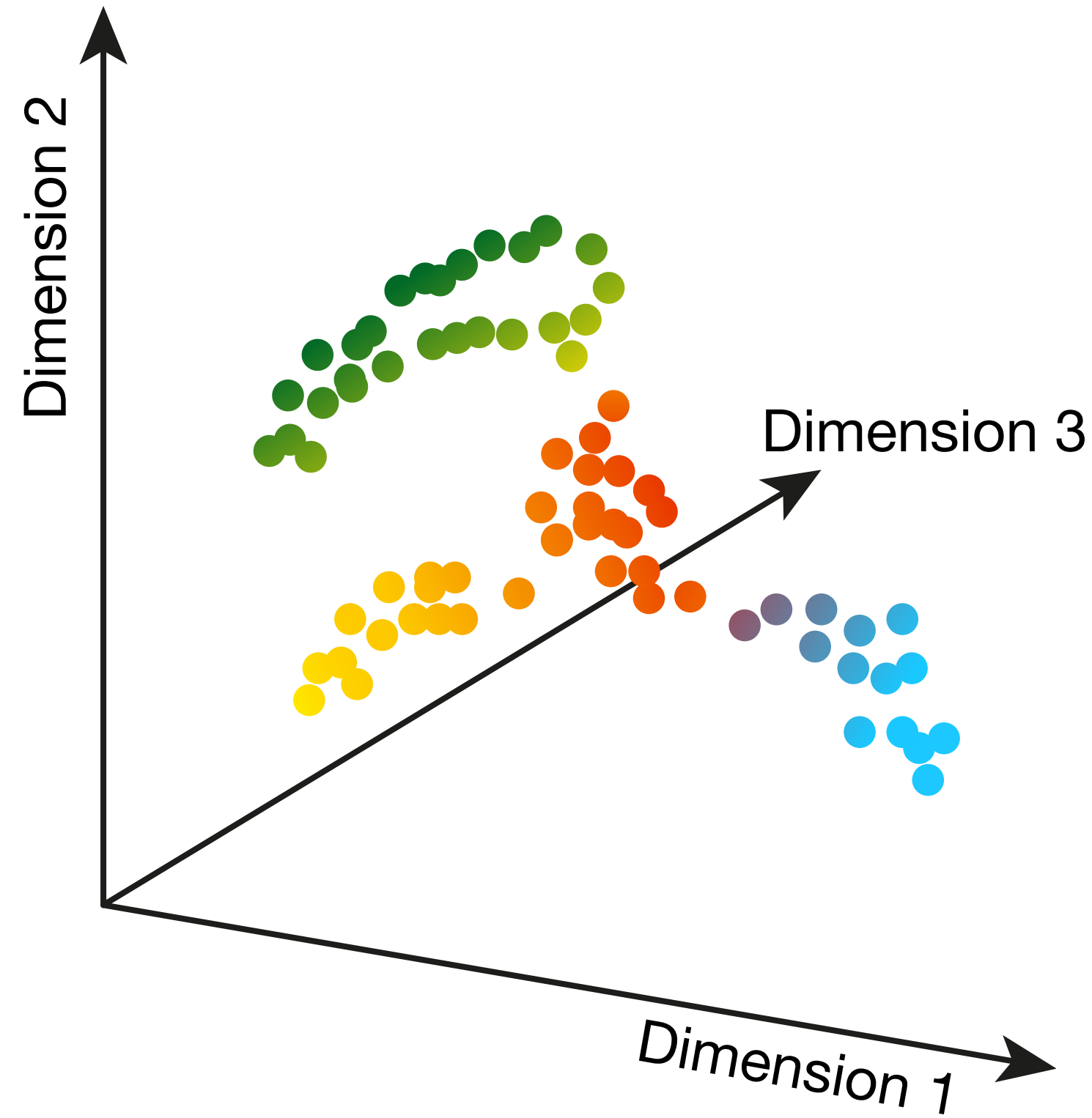
Abstract—Clustering, the process of grouping together similar items into distinct partitions, is a common type of unsupervised machine learning that can be useful for summarizing and aggregating complex multi-dimensional data.

Dimensionality Reduction

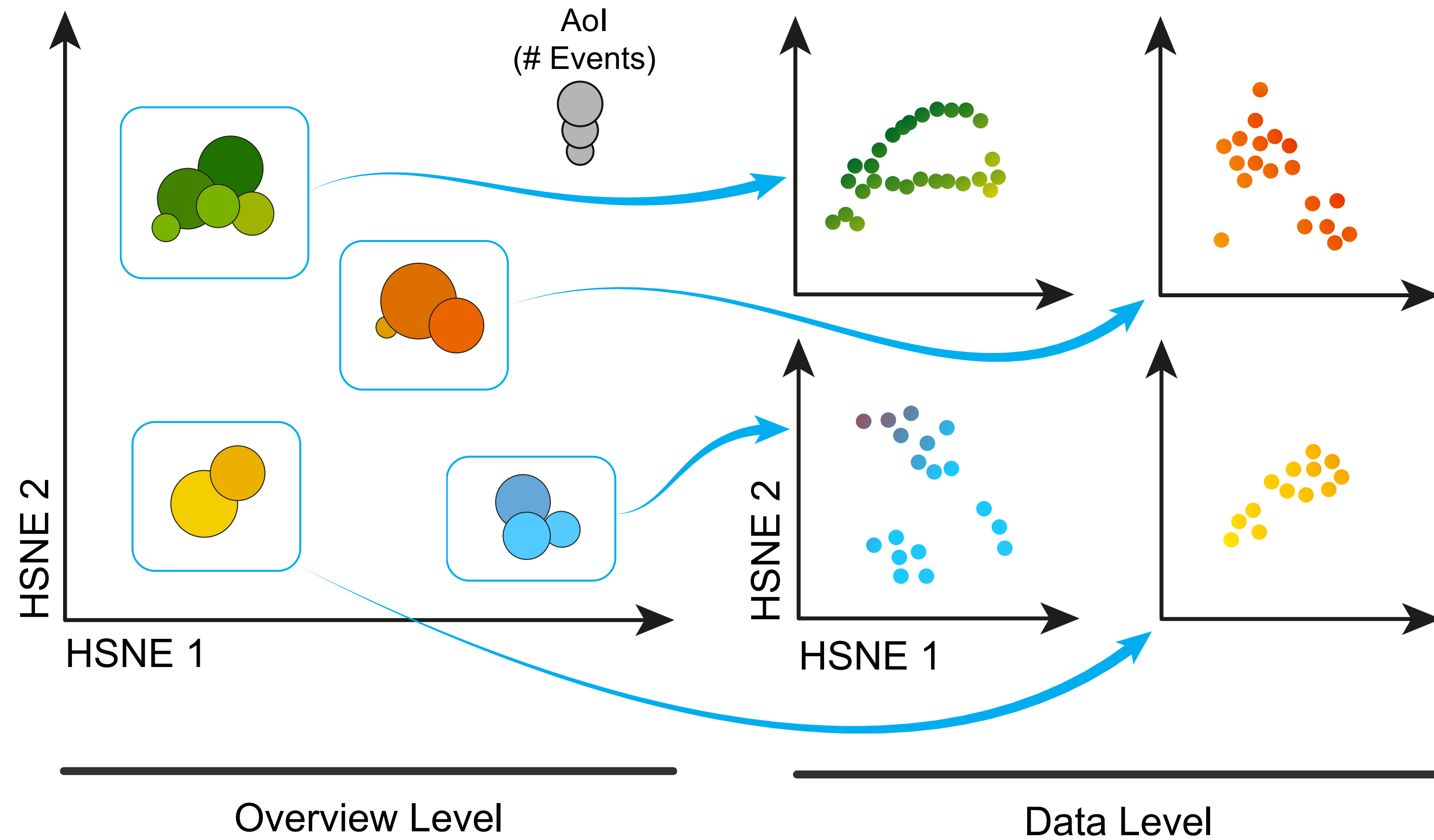


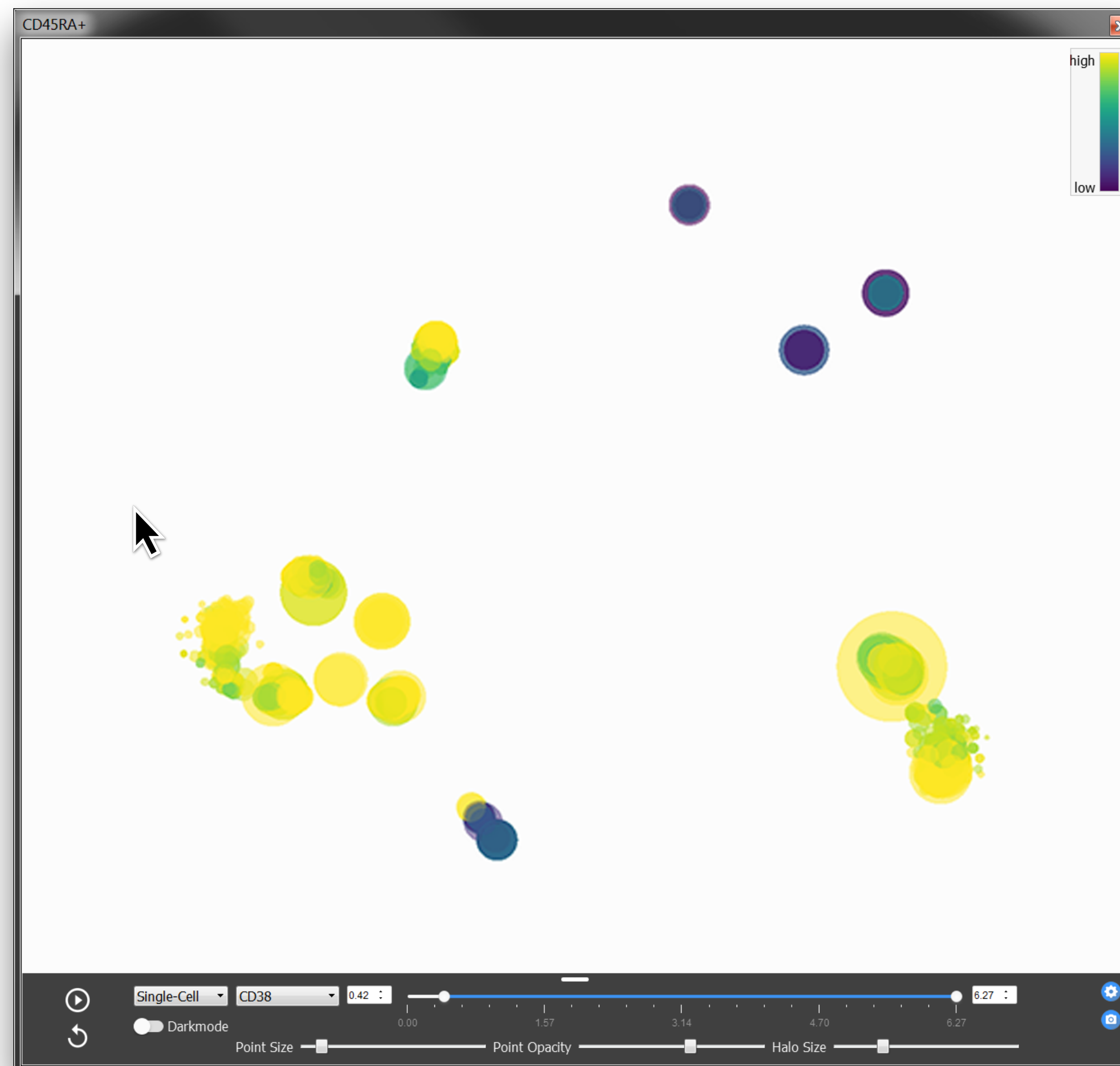
Dimensionality Reduction

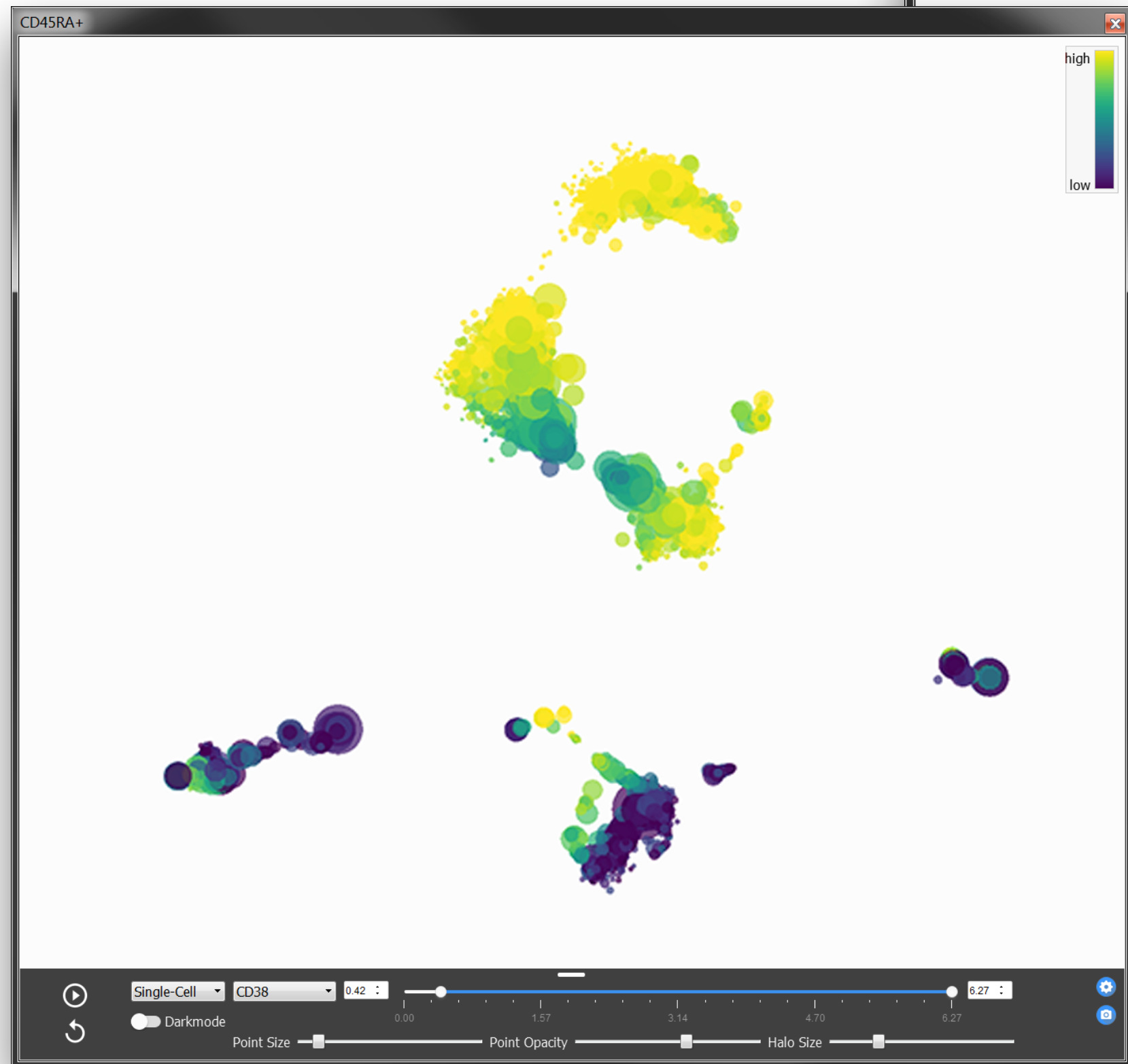
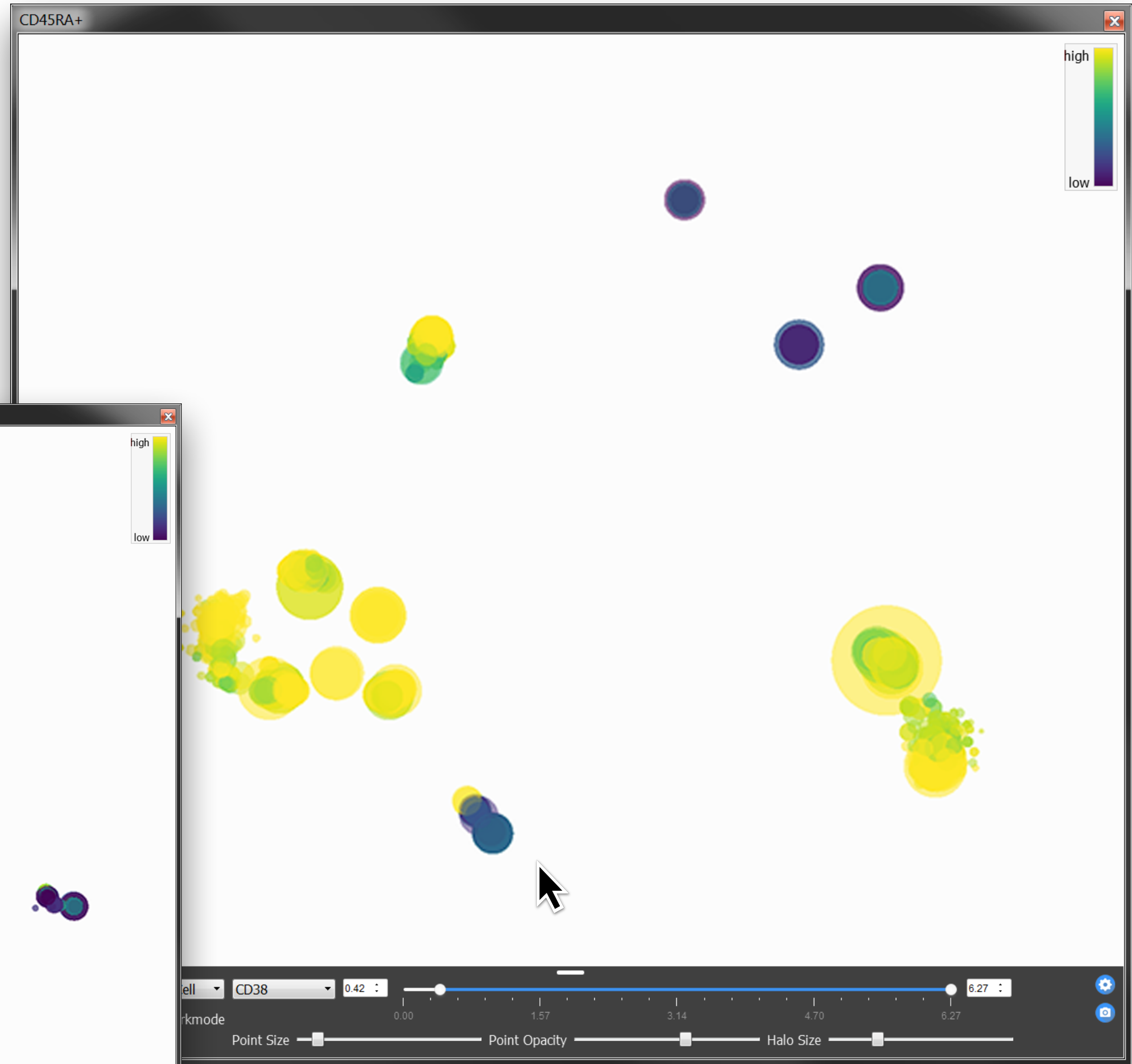


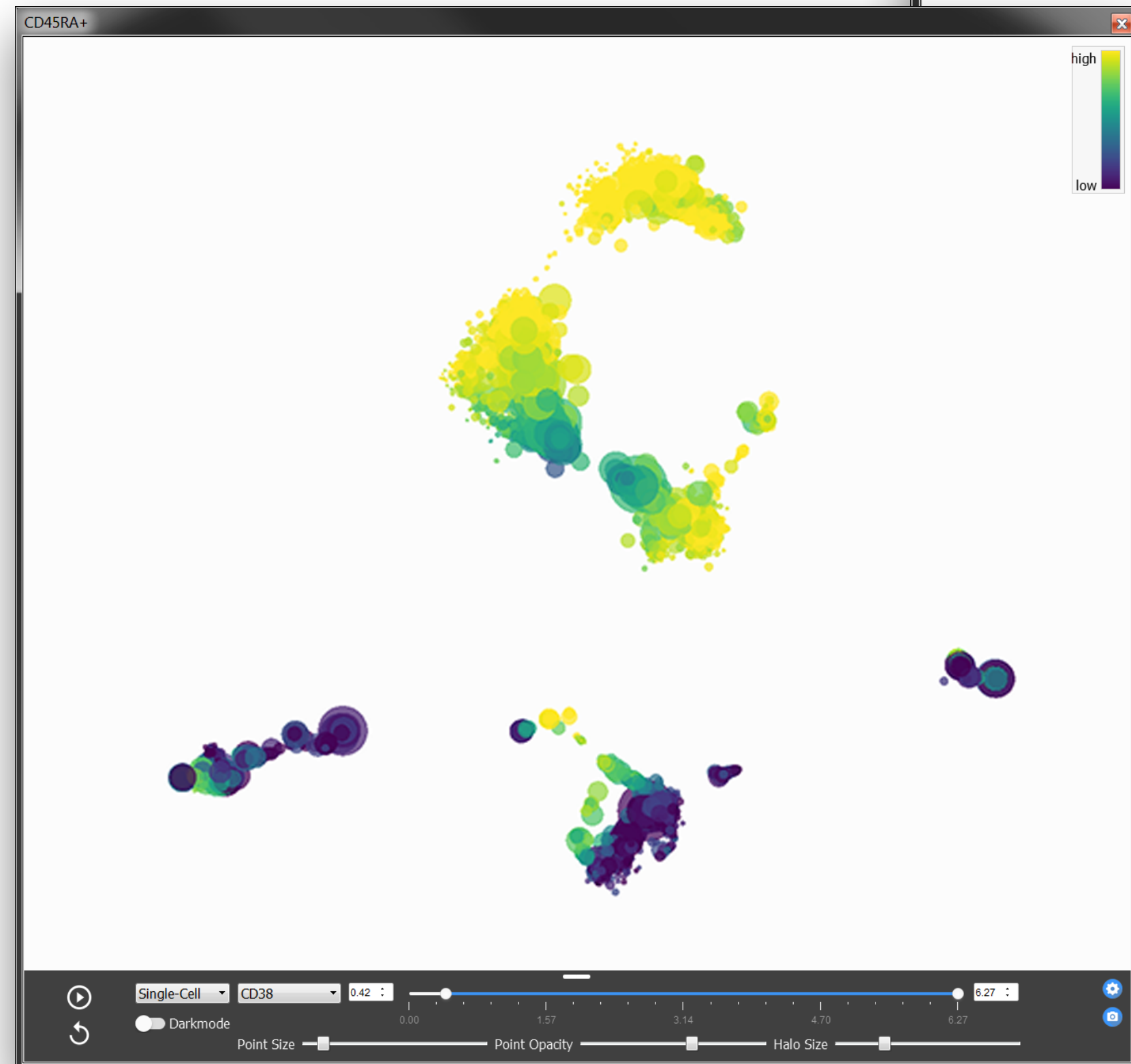
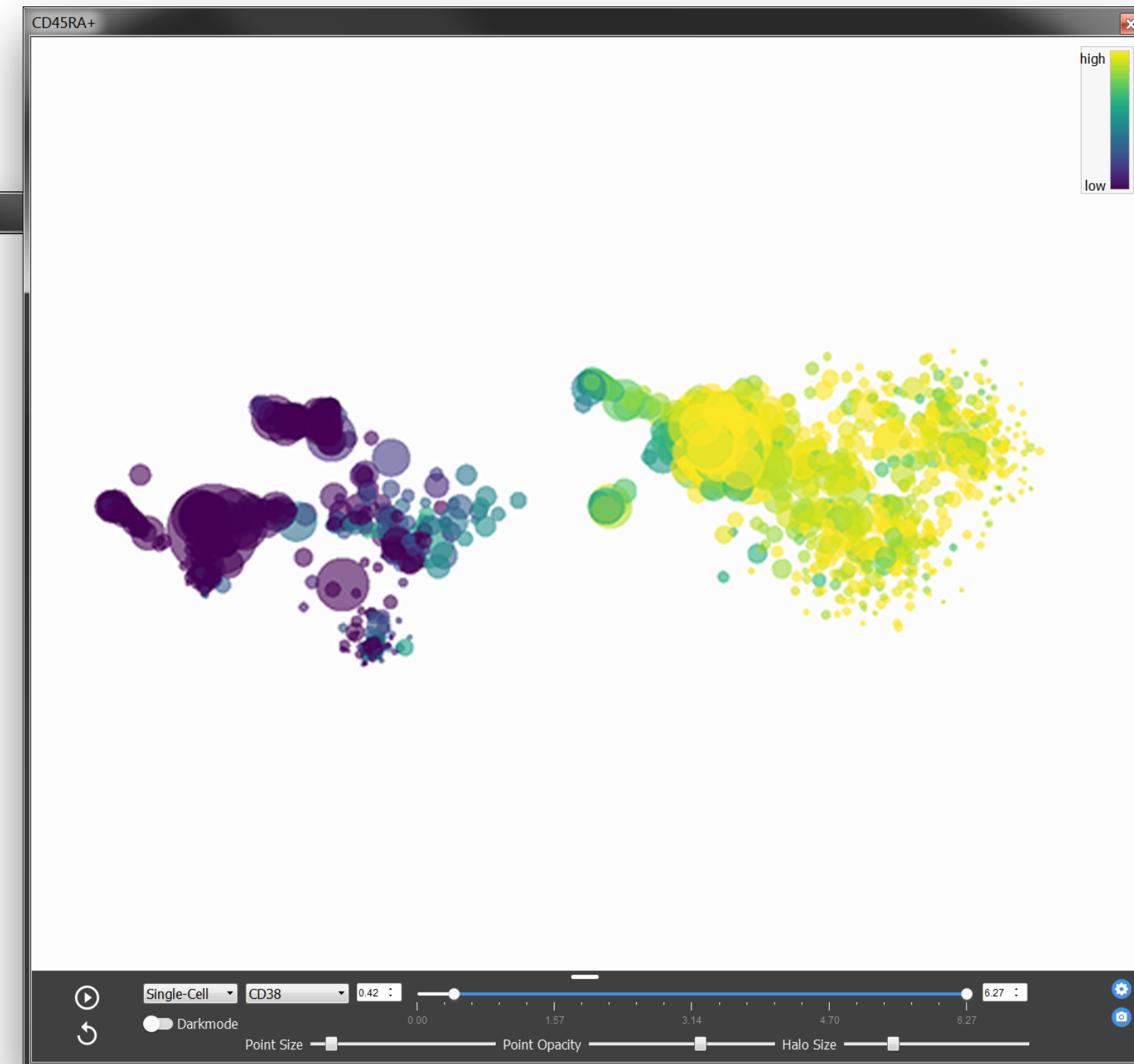
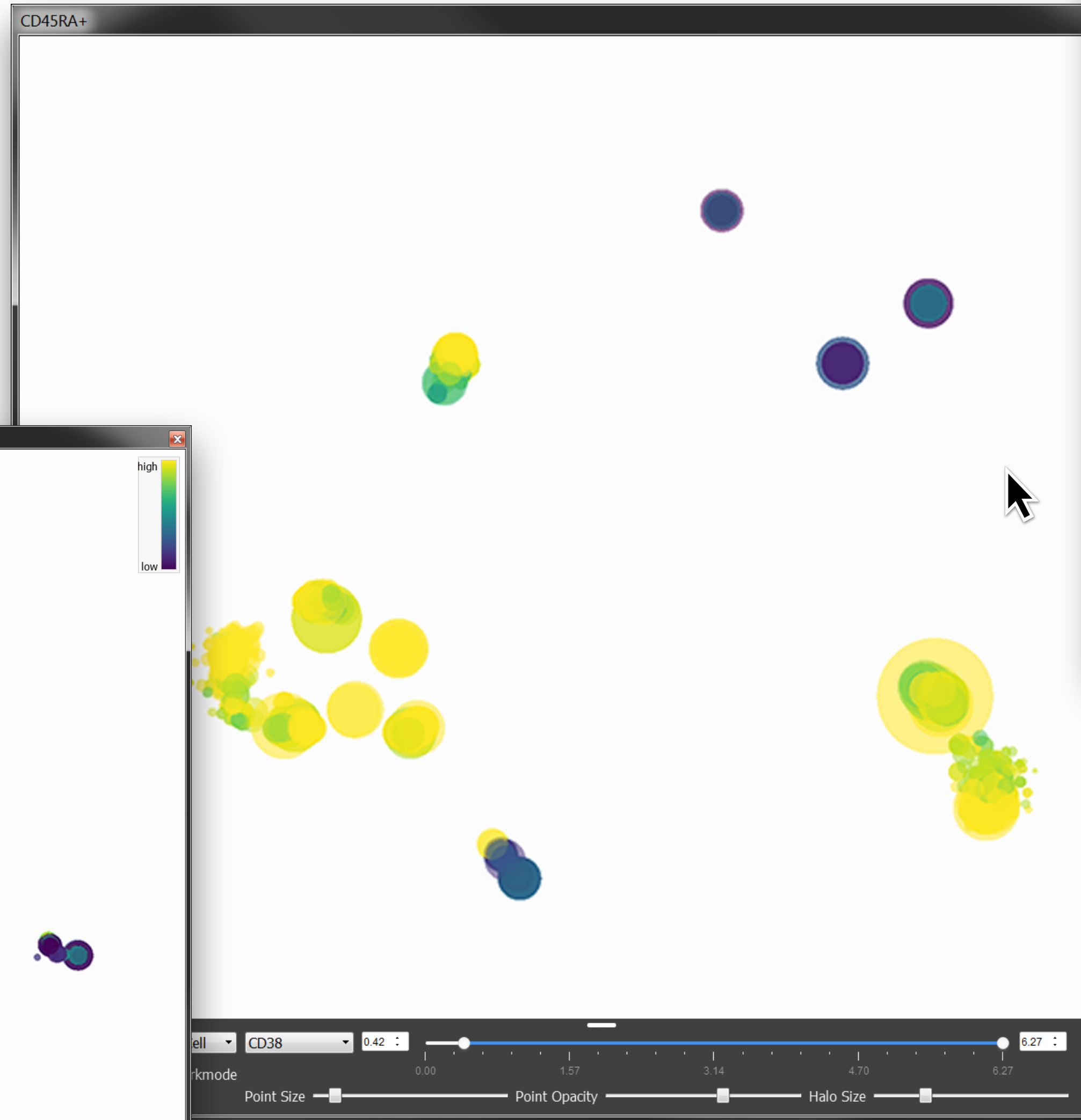


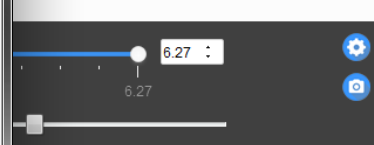
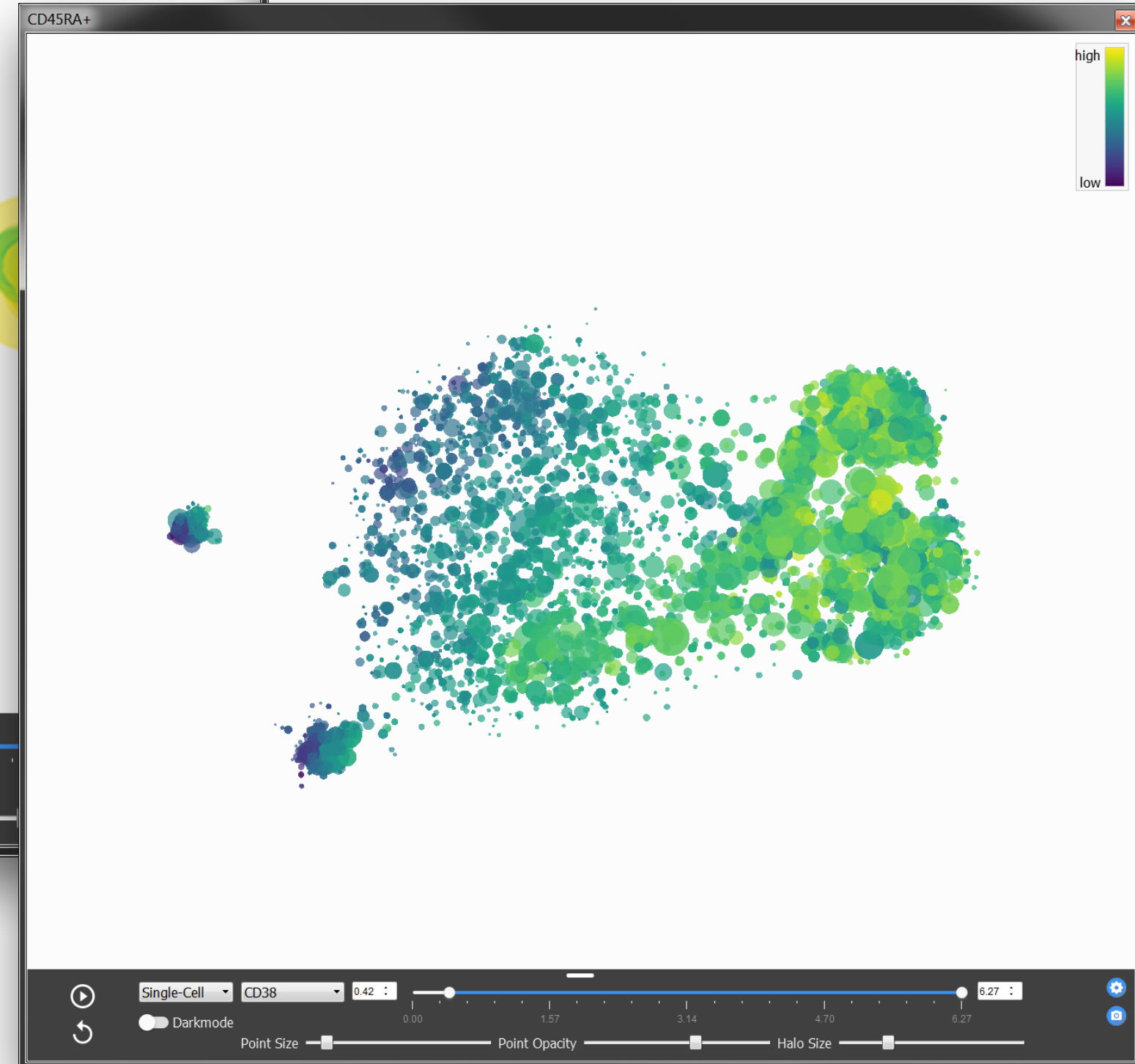
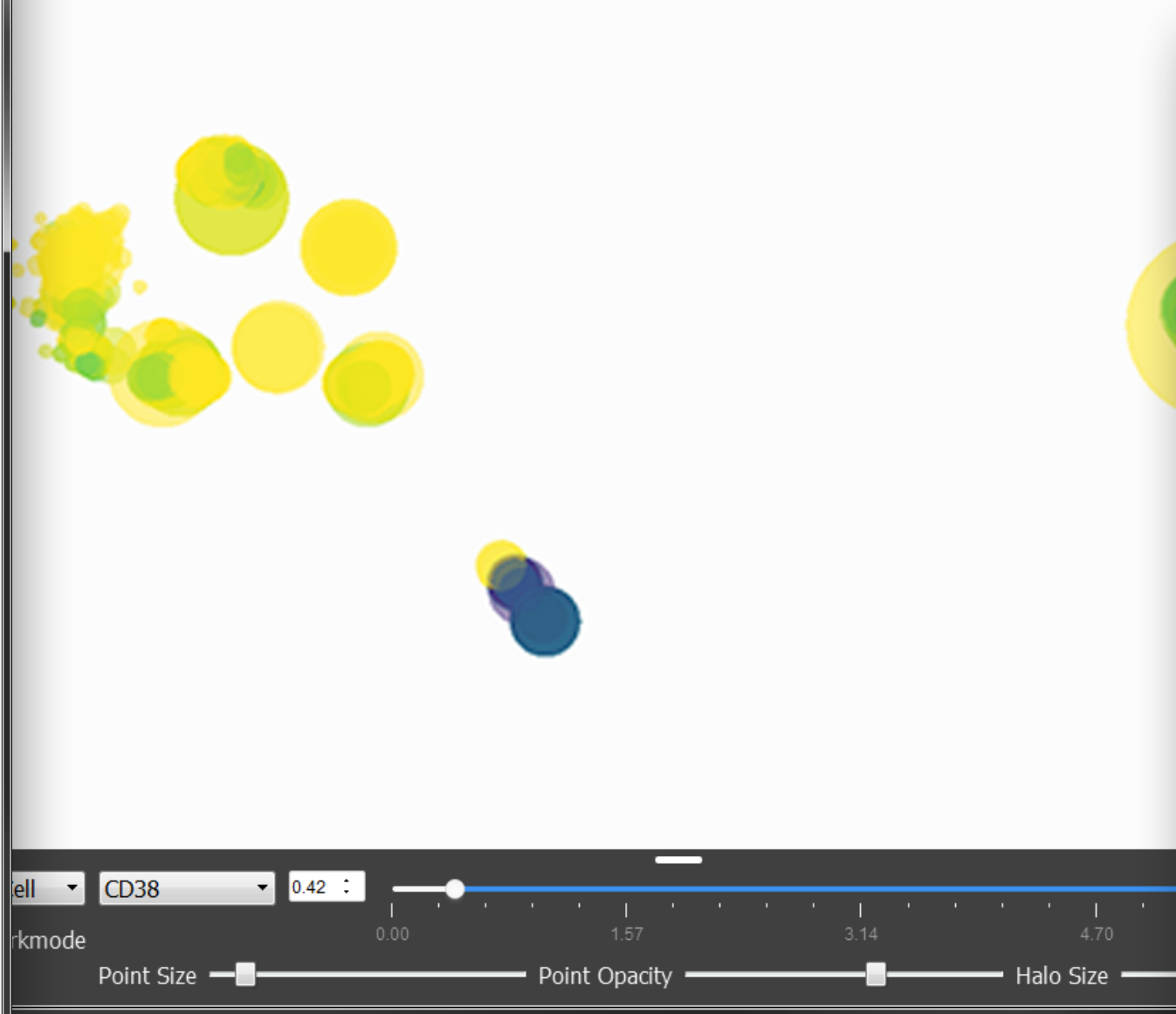
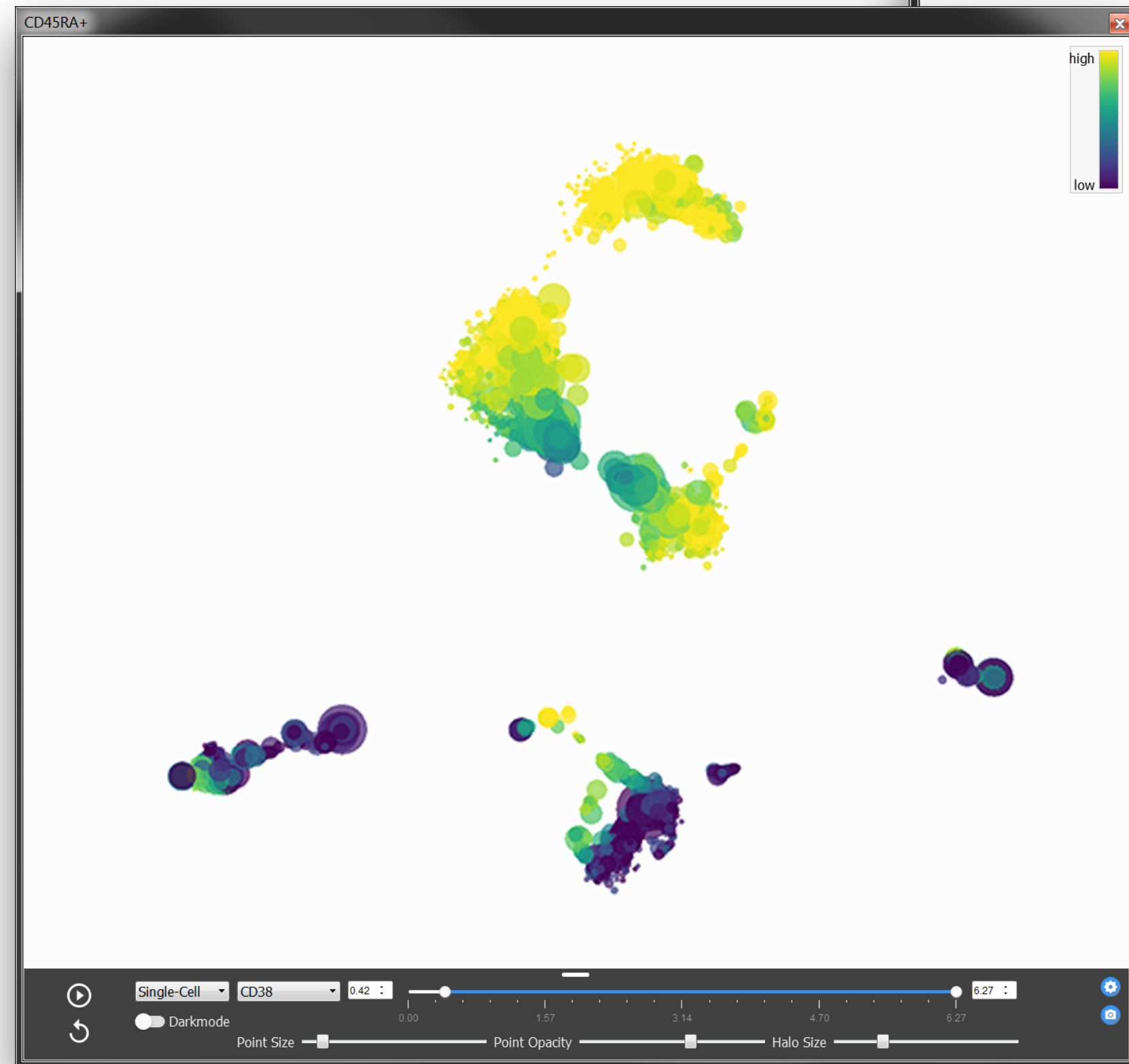
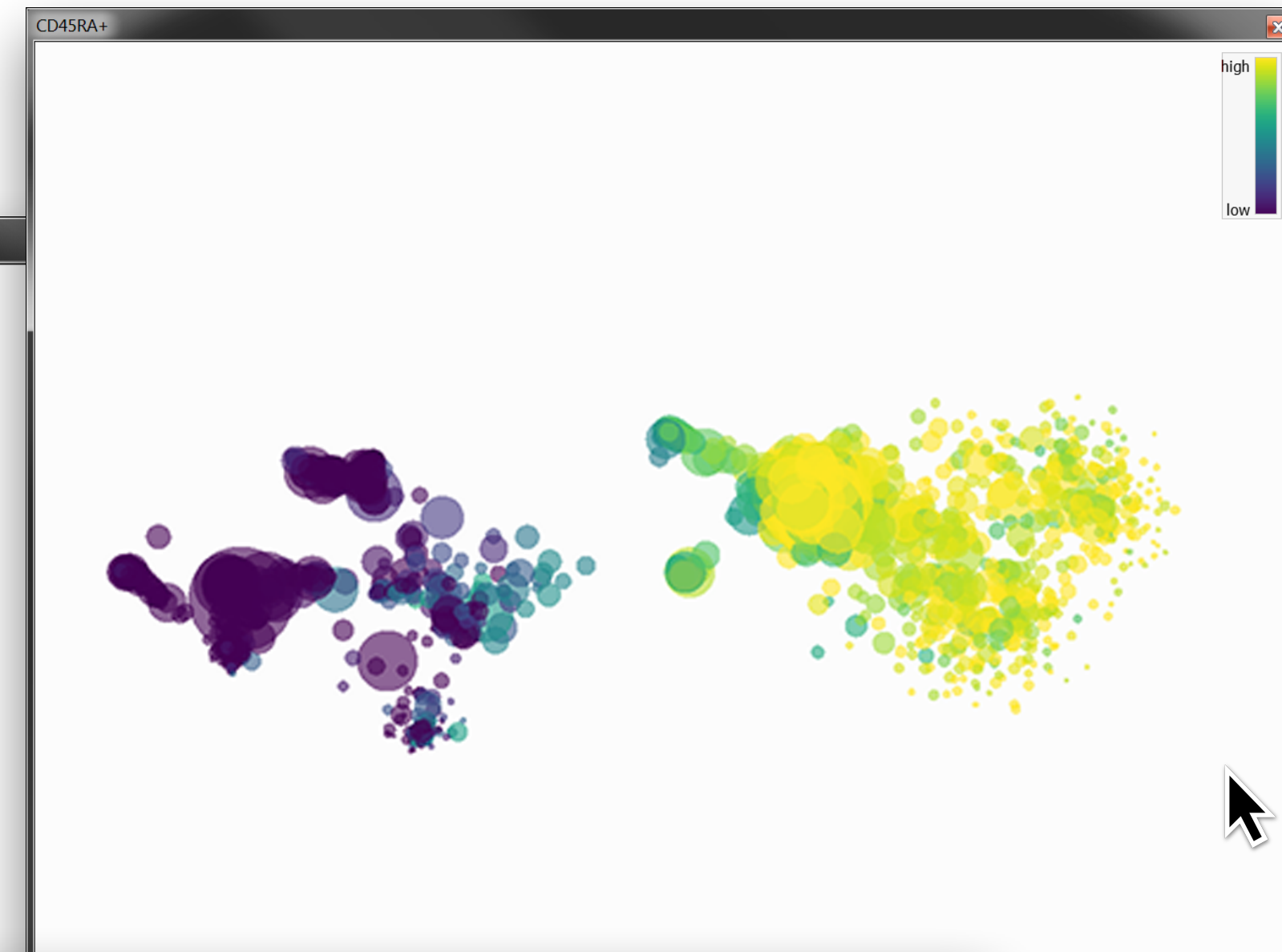
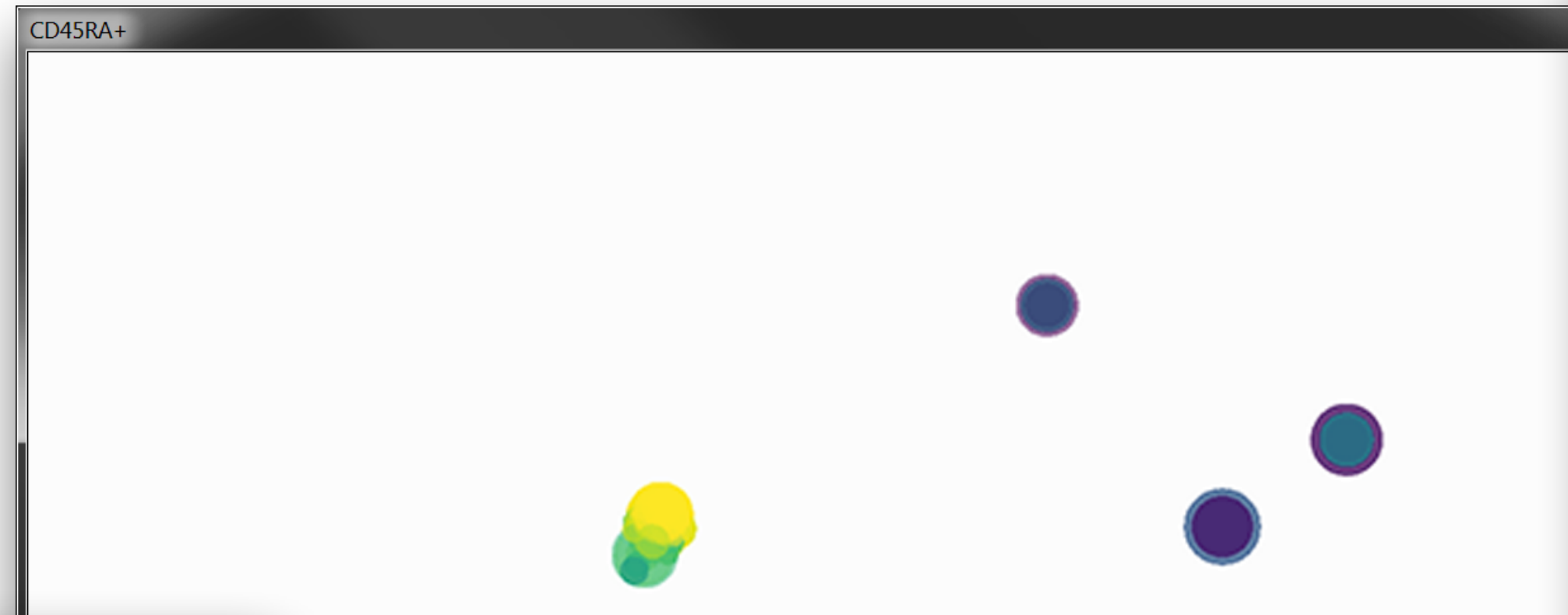
HSNE









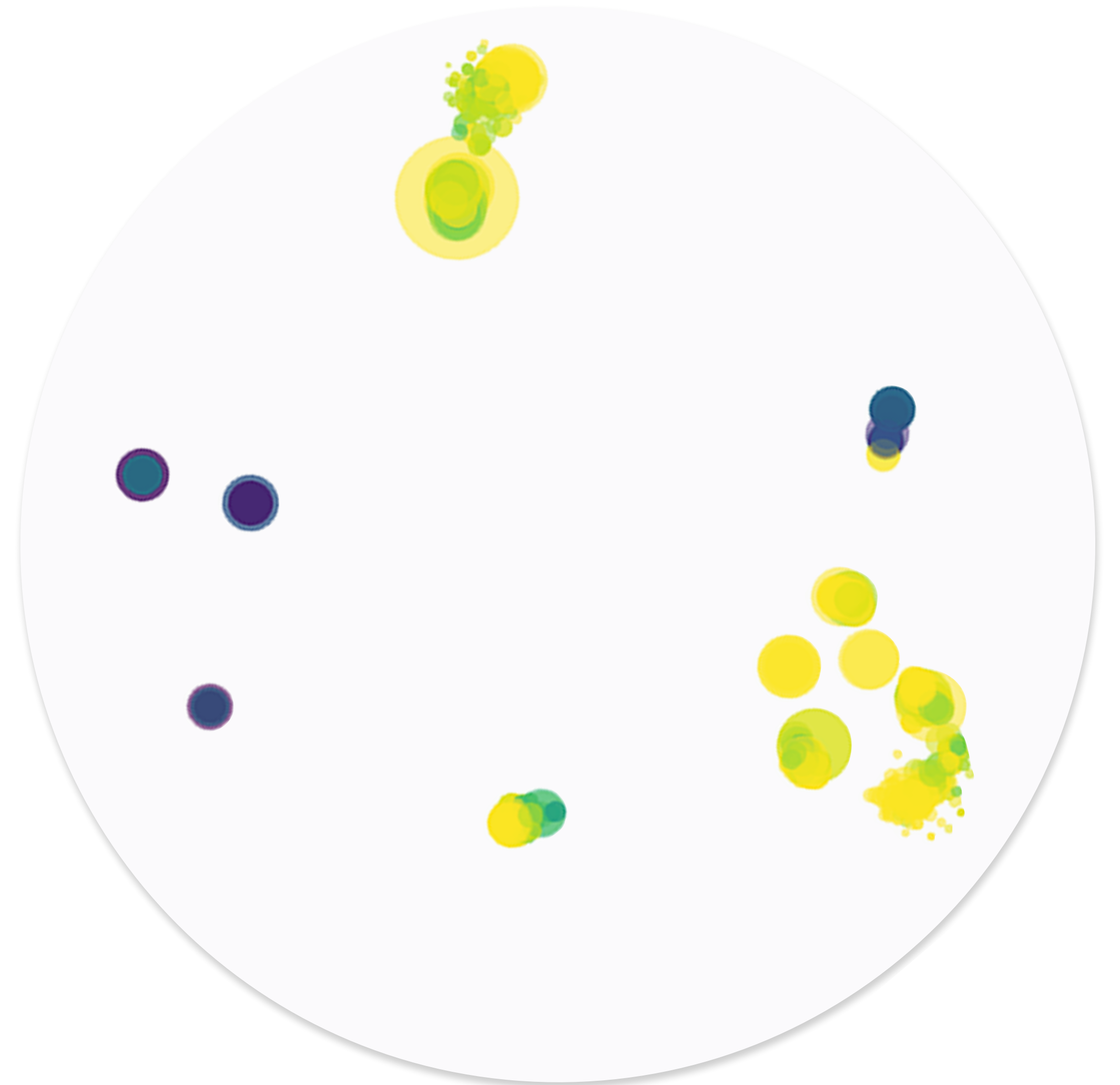


CyteGuide

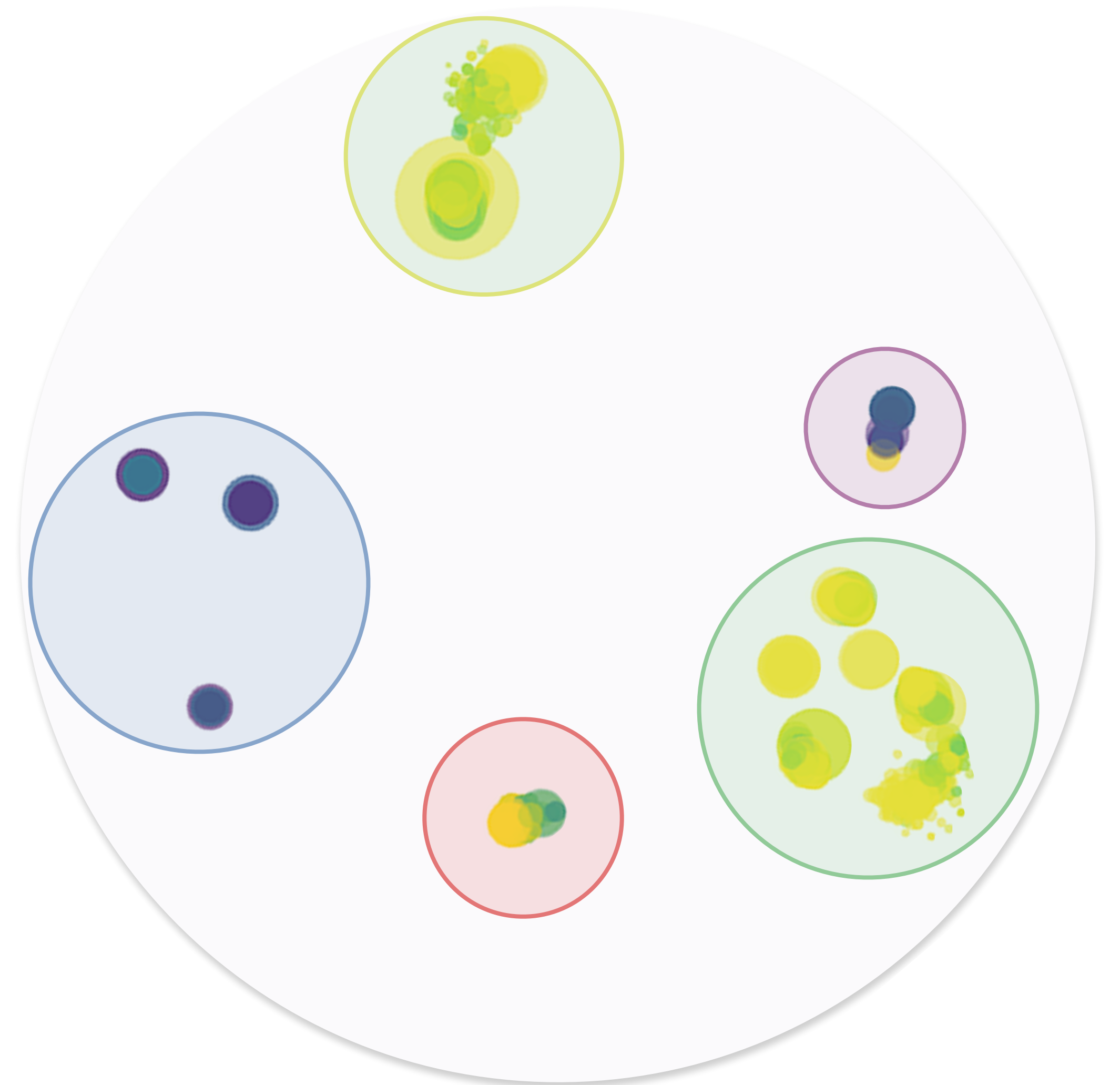
- Recent development in cell analysis
(Commercially available since 2014)
- Protein expression for single cells
 - Each cell: 40 dimensional data point
 - Millions of cells per study
- Identify new cell types



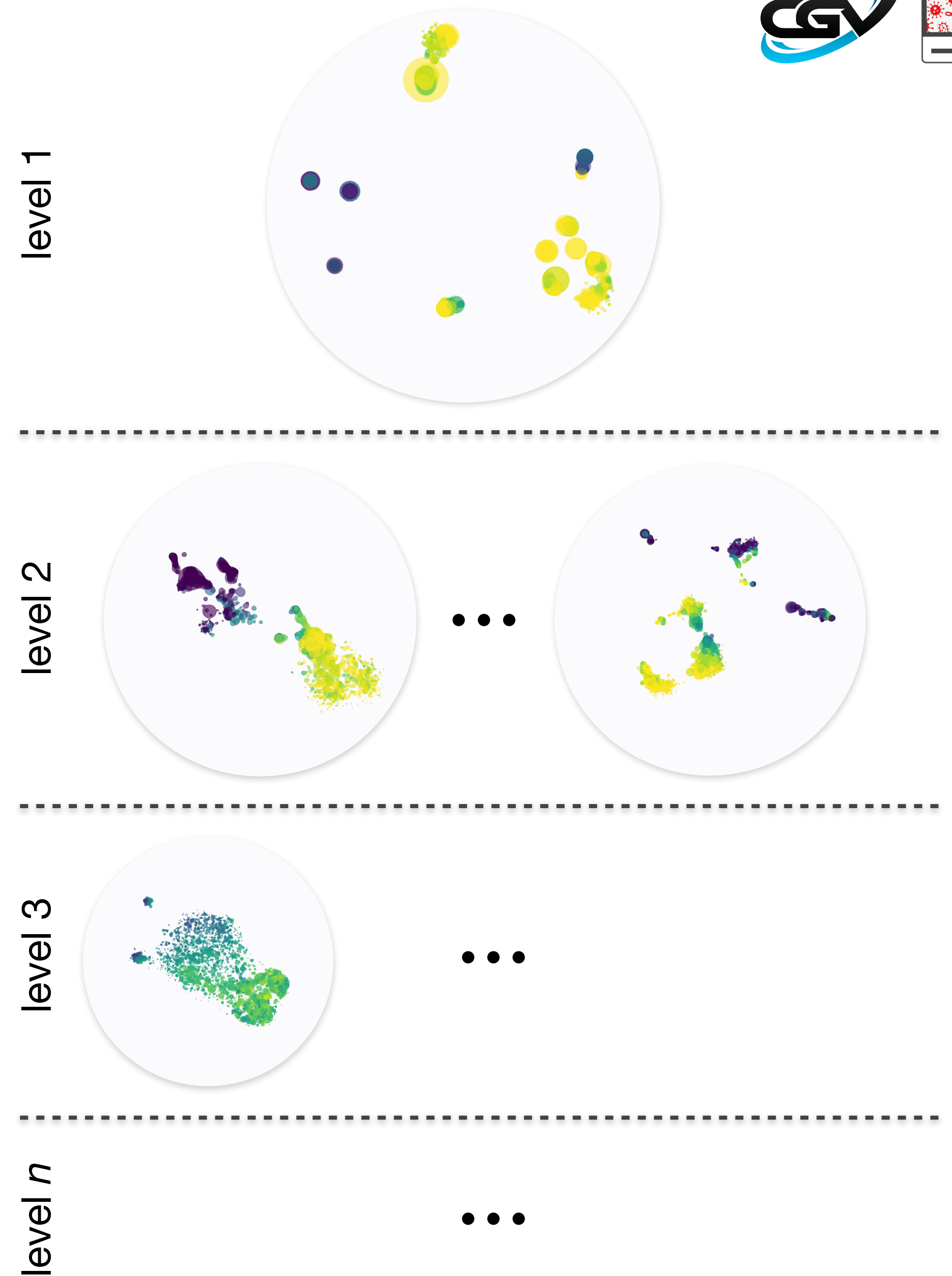
- Representation of the exploration
- Exploration through clusters
 - 3 to 5 clusters in each embedding
- Up to 5 hierarchy levels



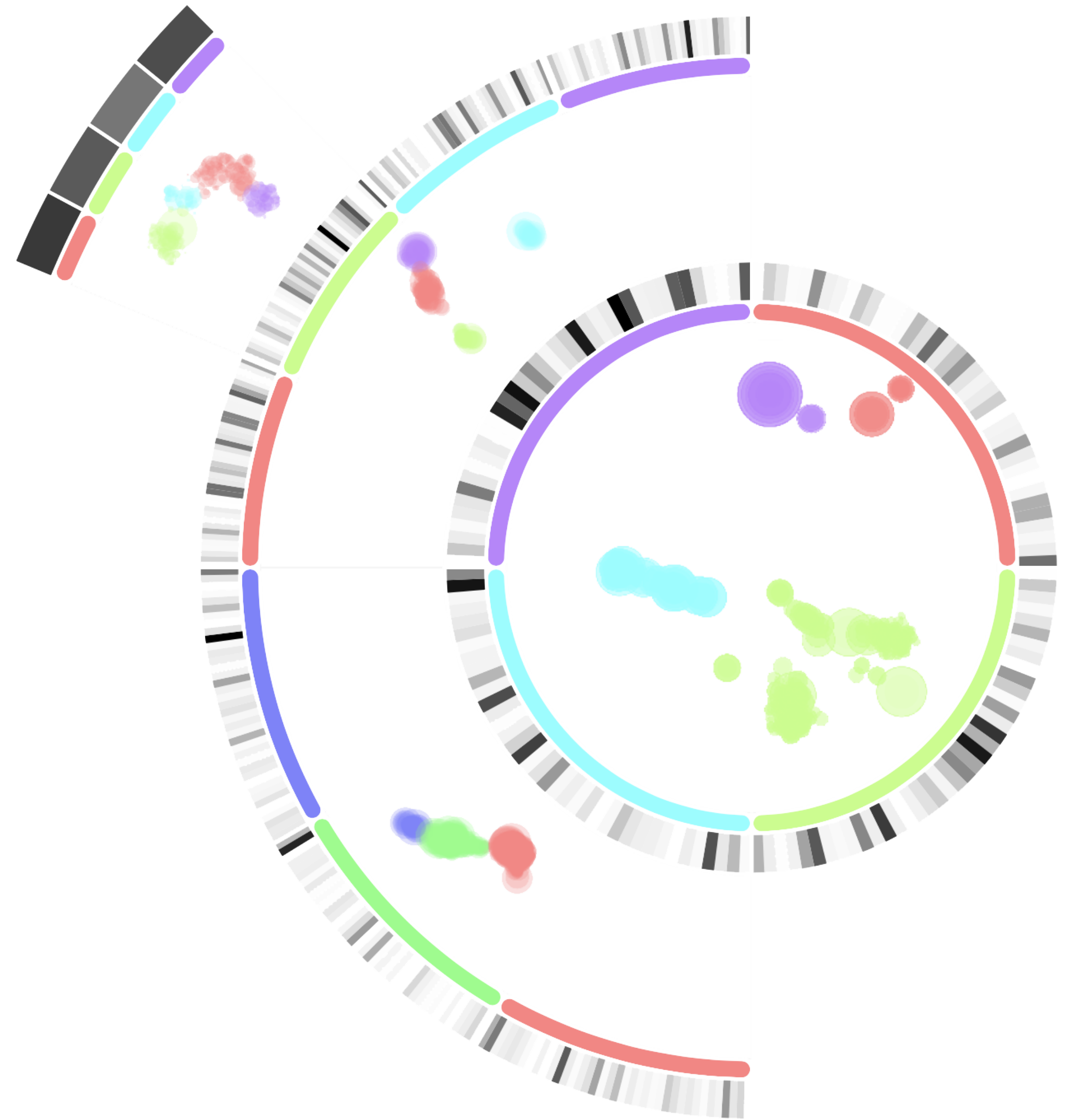
- Representation of the exploration
- Exploration through clusters
 - 3 to 5 clusters in each embedding
- Up to 5 hierarchy levels



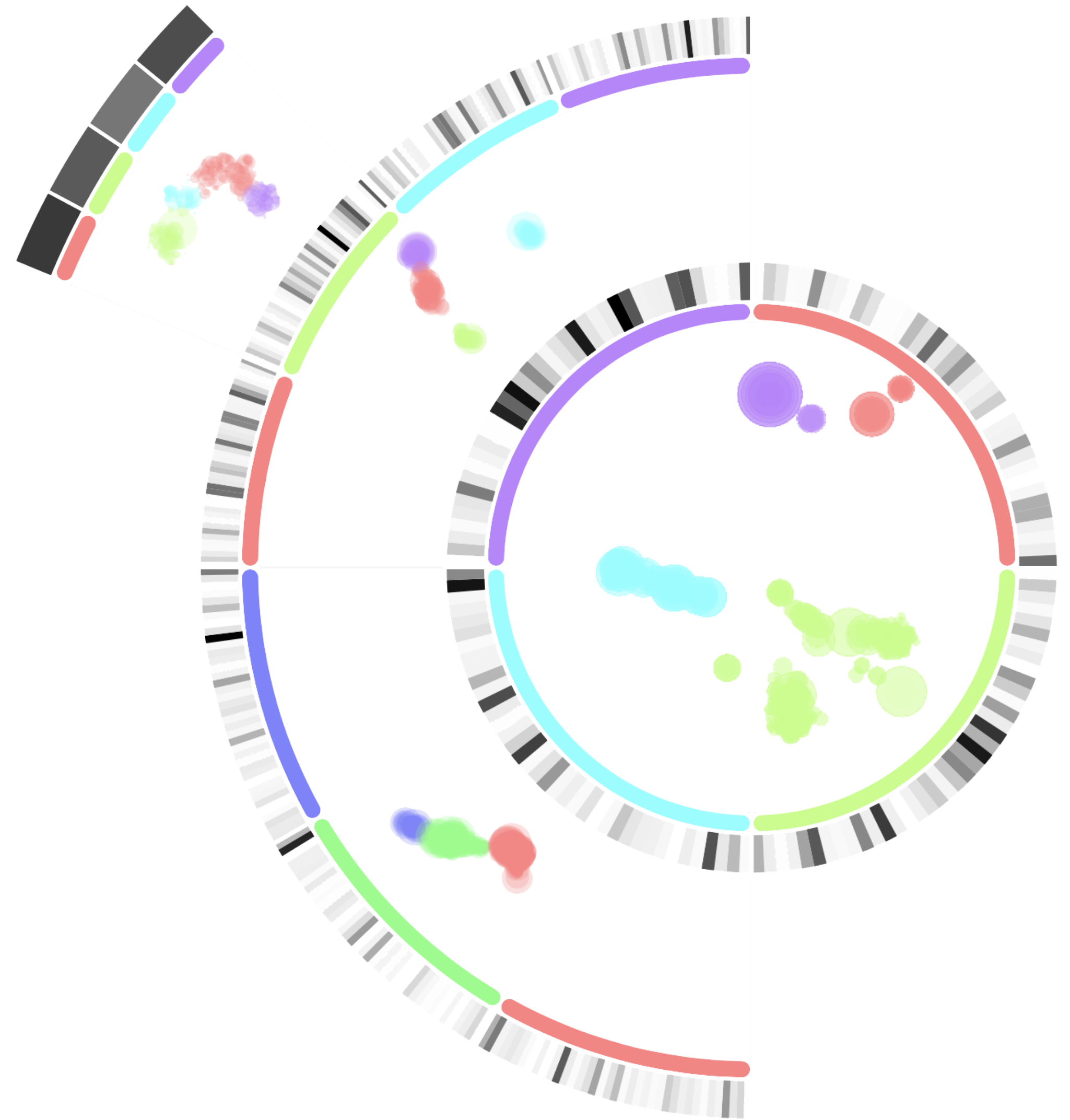
- Representation of the exploration
- Exploration through clusters
 - 3 to 5 clusters in each embedding
- Up to 5 hierarchy levels



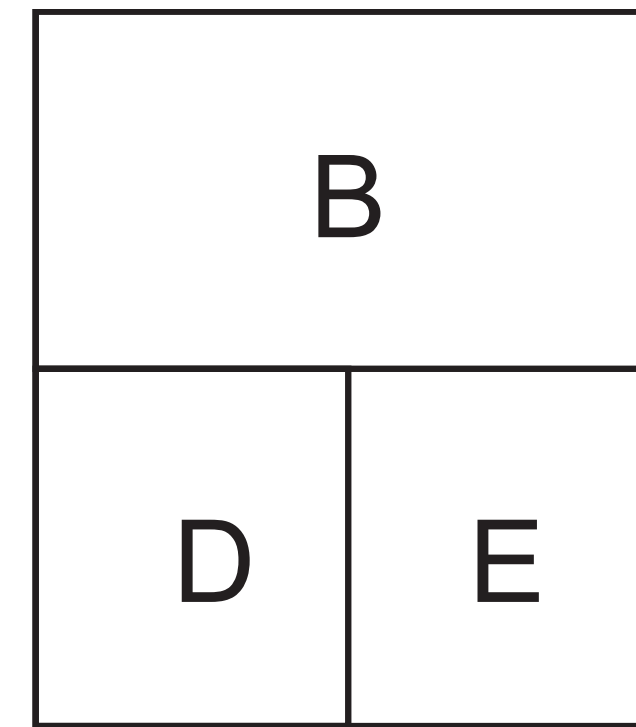
- Single, integrated visualization
- Guides the hierarchy exploration
- Presents the hierarchy exploration



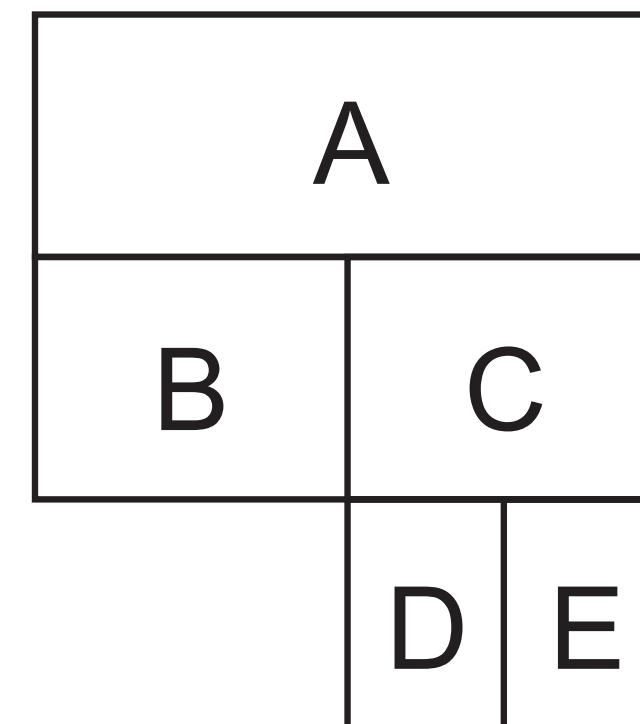
- **Exploration**
- Navigate the hierarchy
- Identify interesting regions
 - Structural variation
 - Functional variation
- Identify analyzed regions



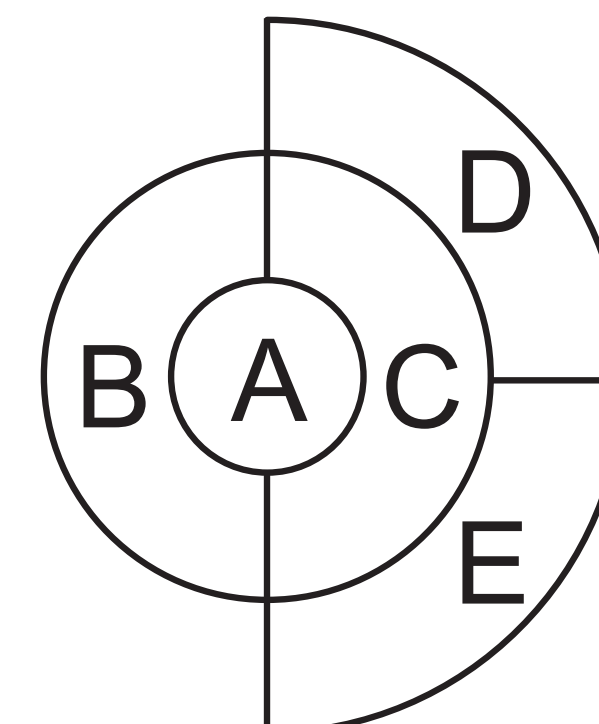
- Hierarchy
- Up to 5 levels
- Up to 5 branches per node
- Embedded vis in nodes



Treemap



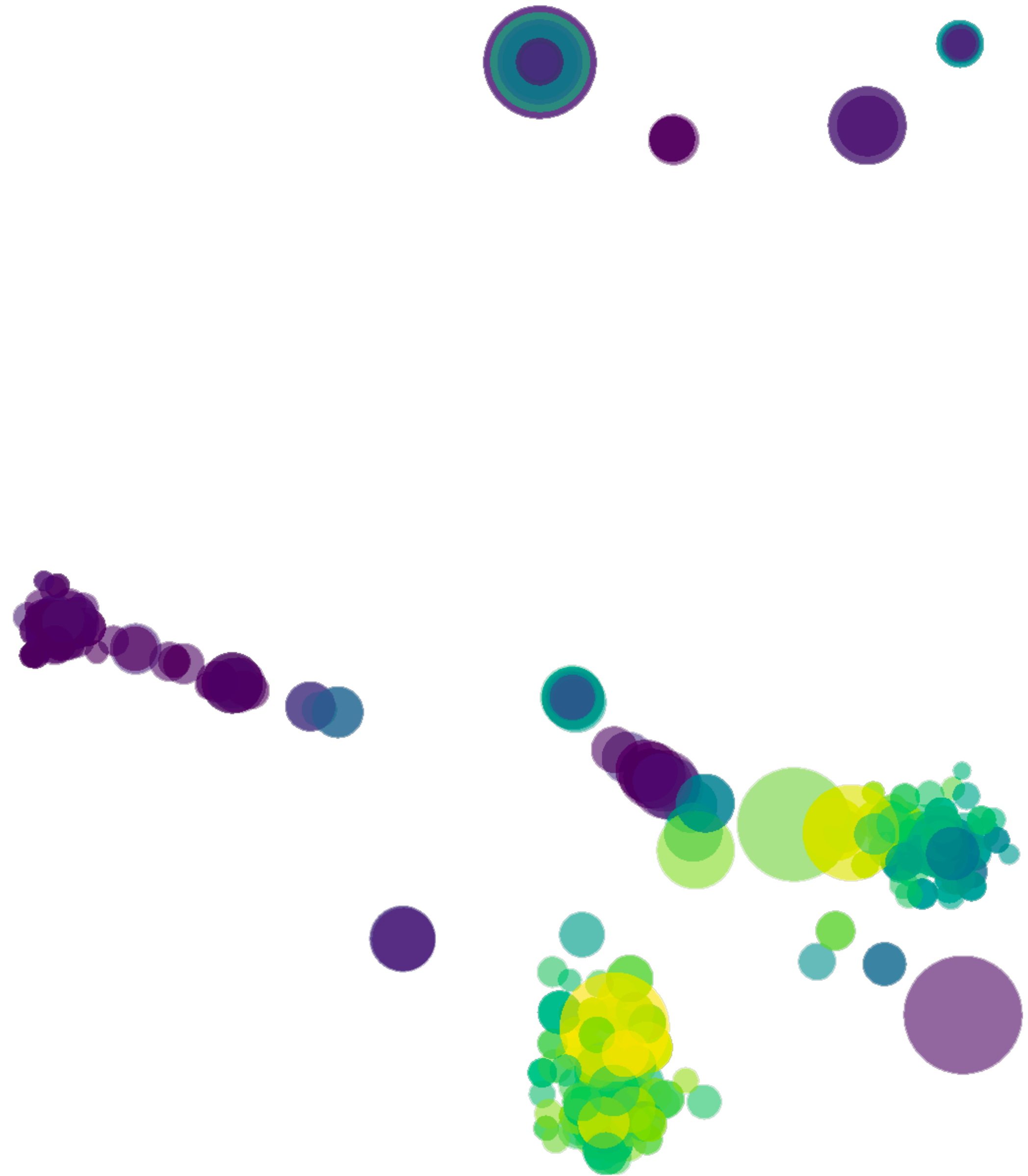
Icicles

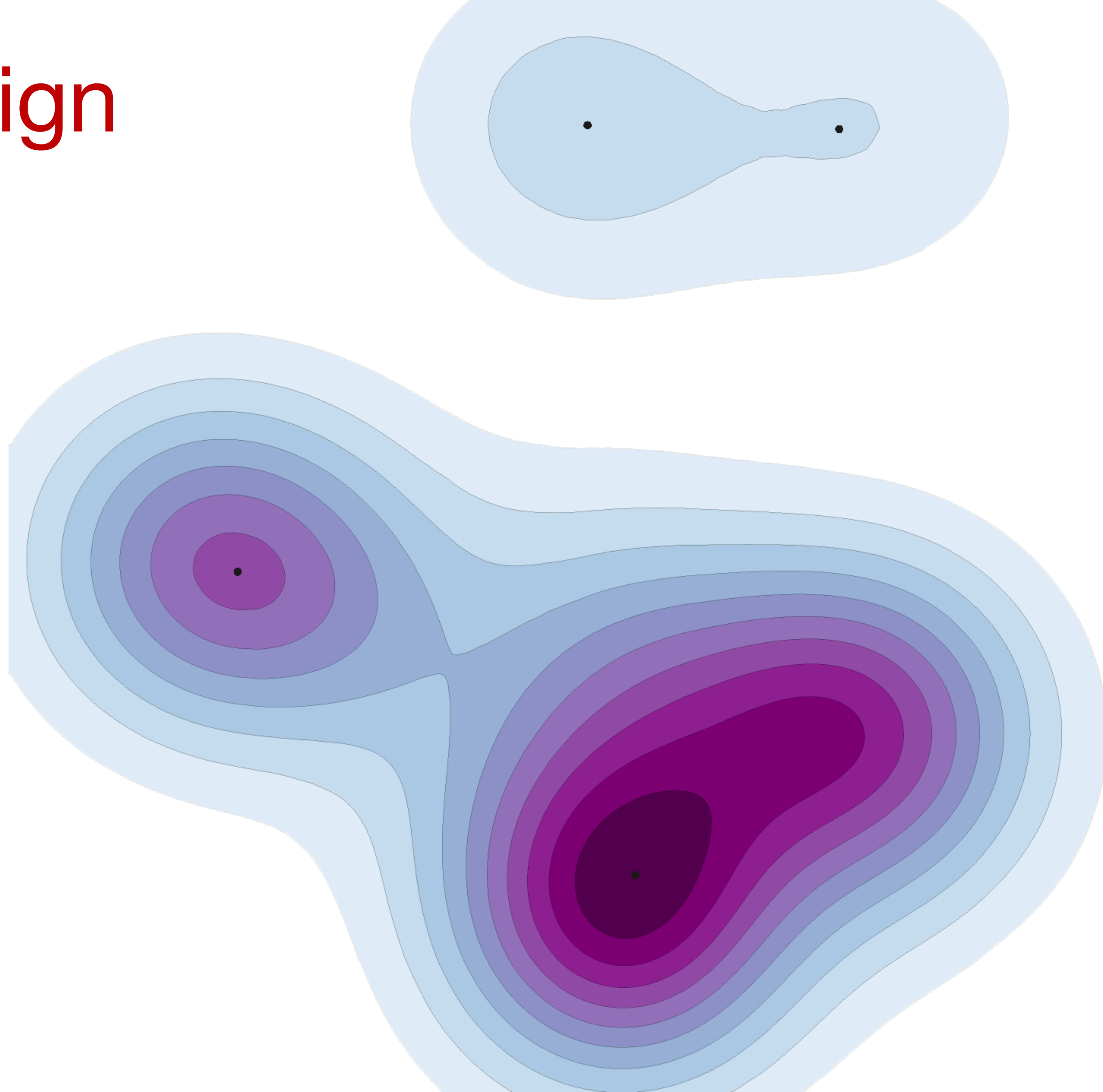


Sunburst

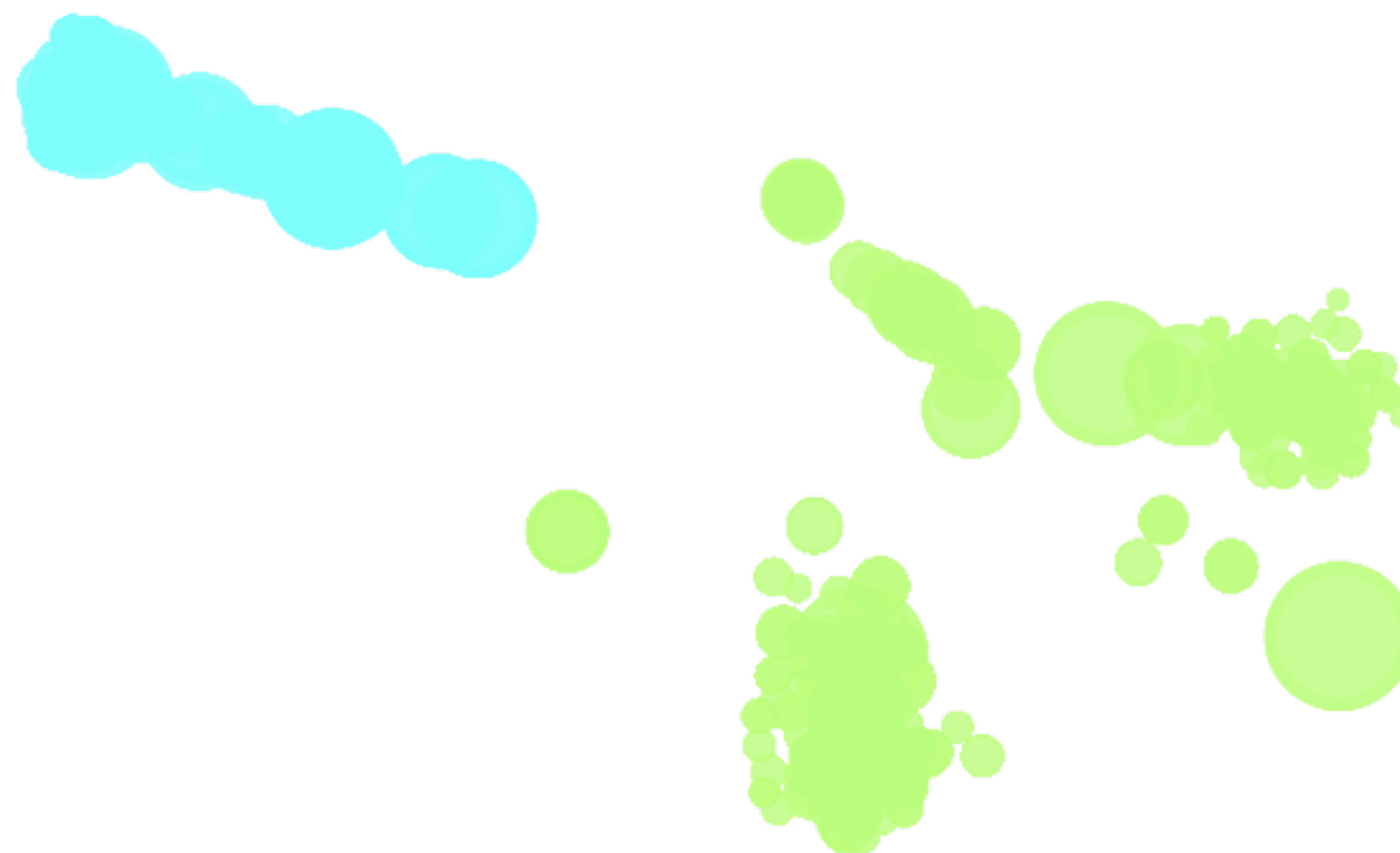


Design



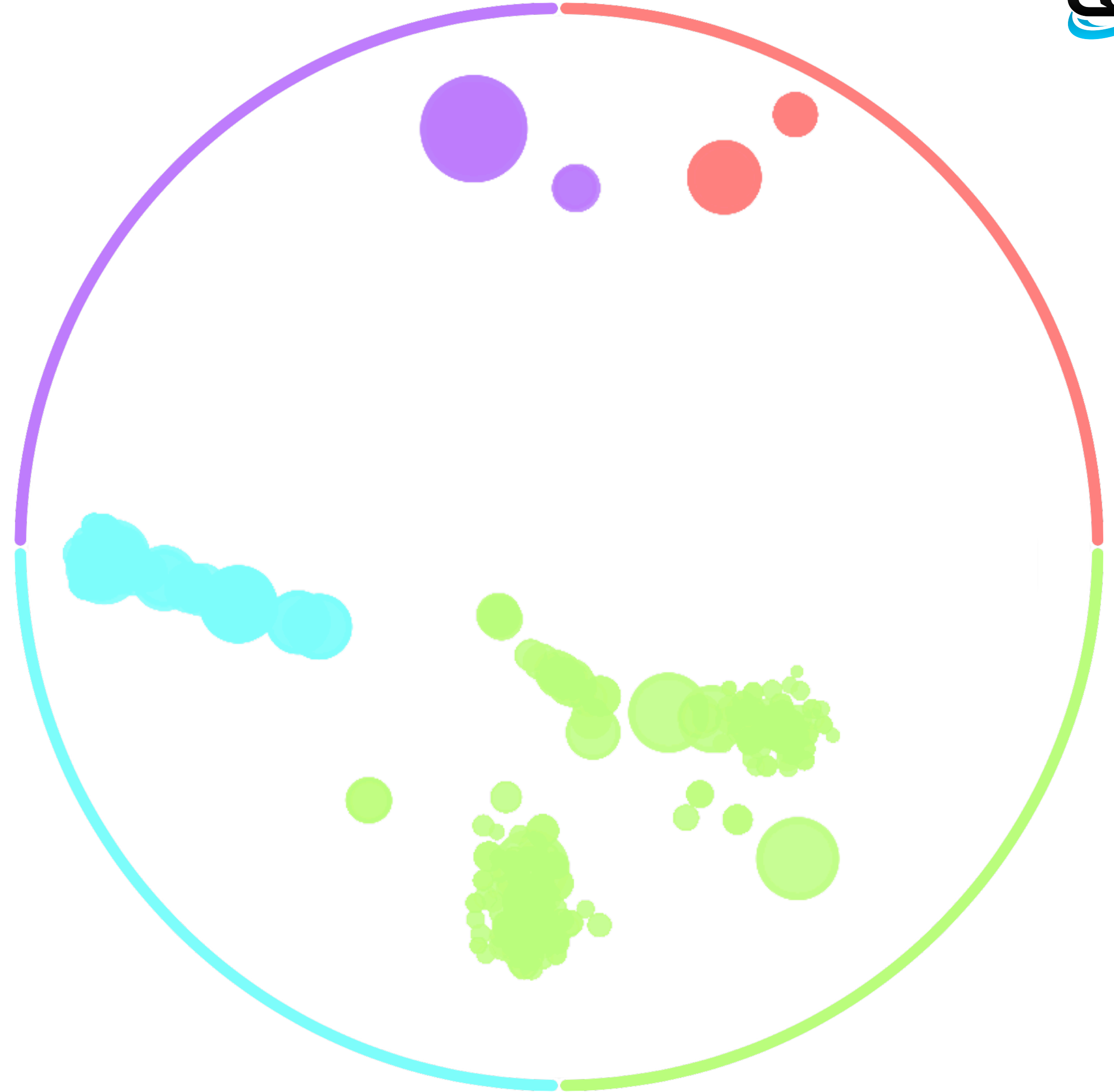


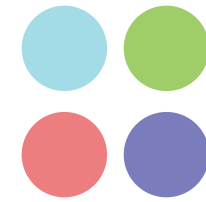


 clusters

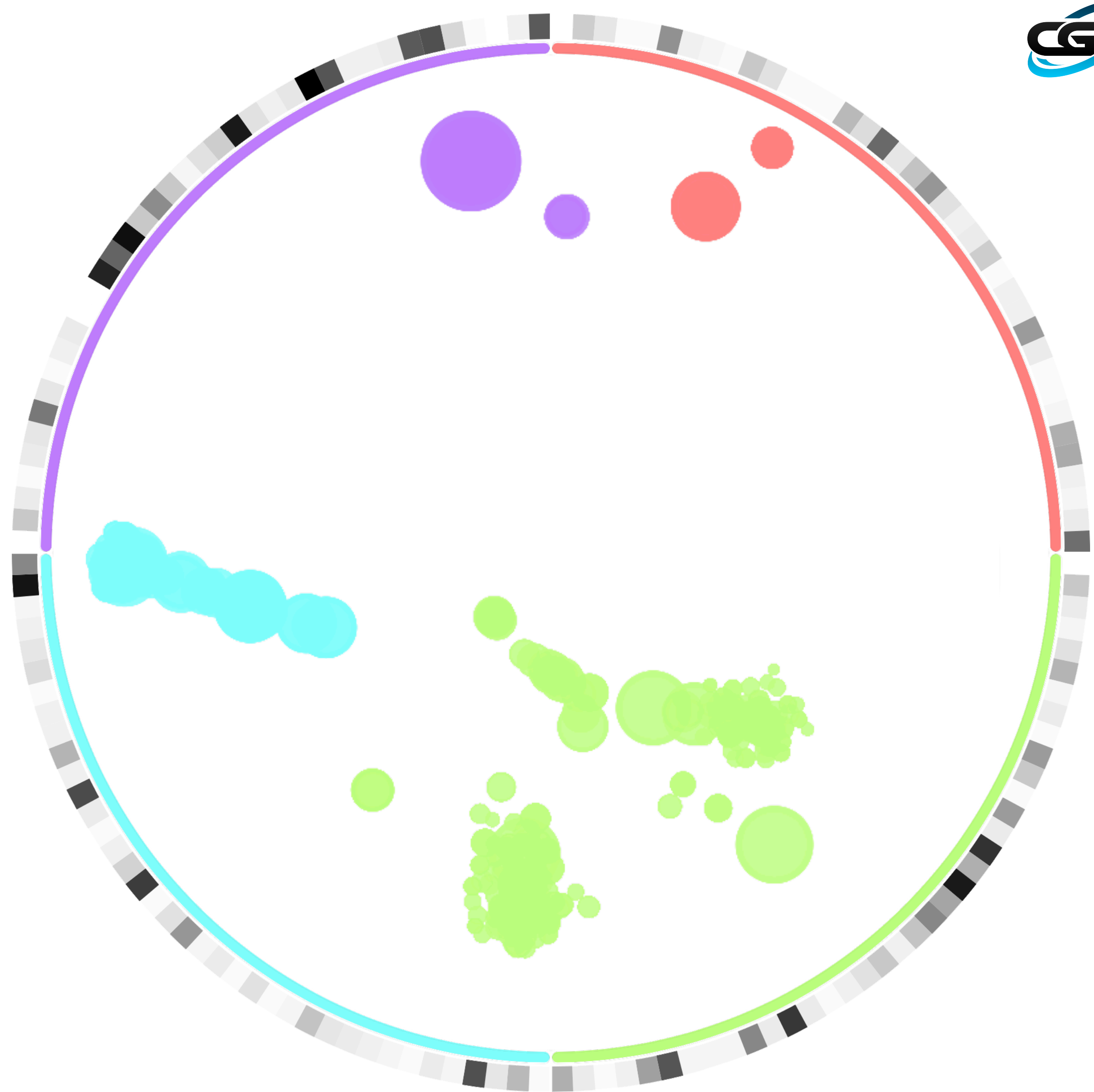


 clusters


 arc segment

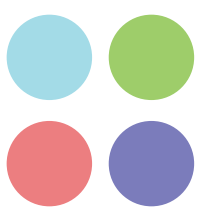




-  clusters
-  arc segment
-  marker heatmap

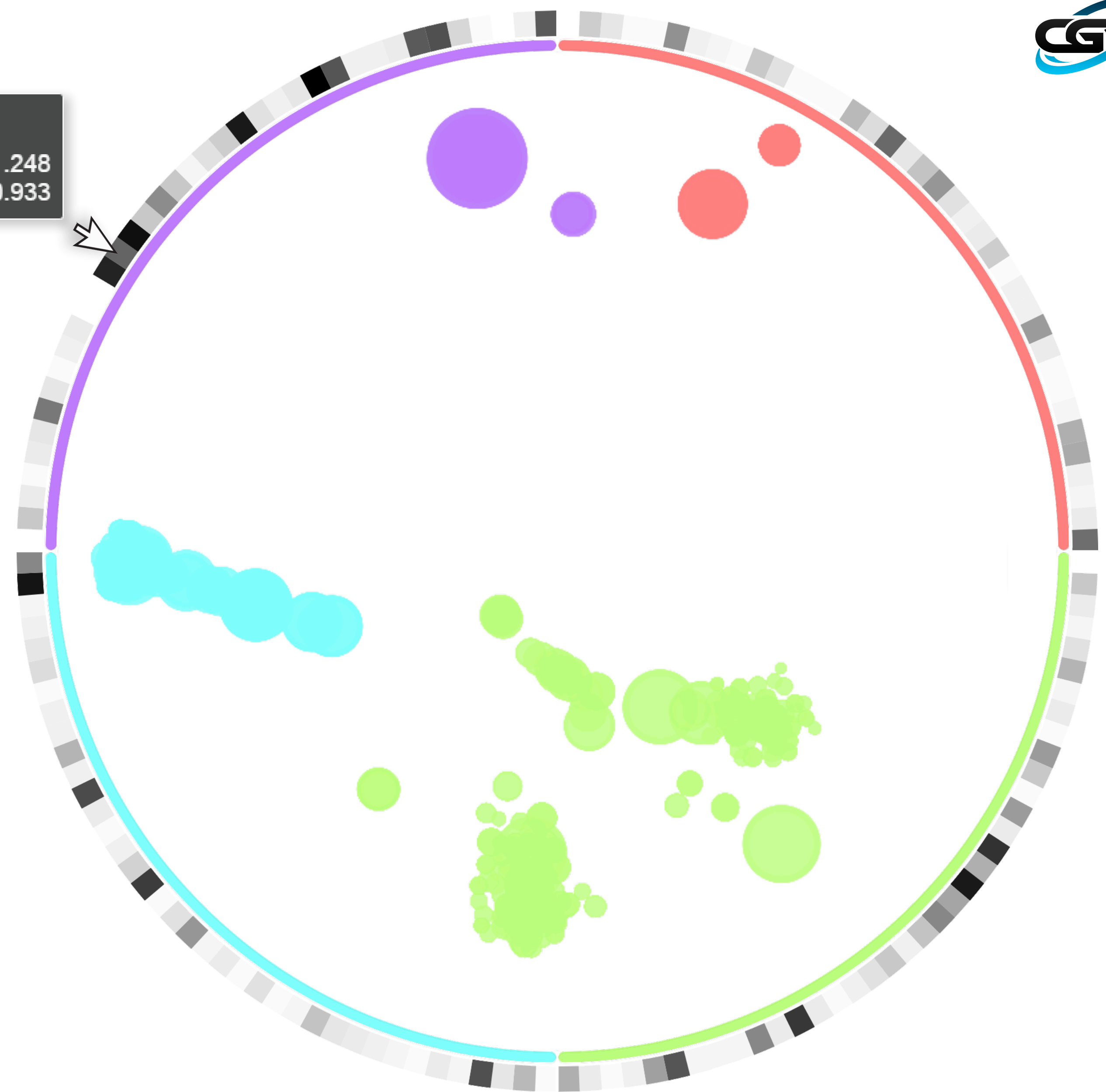


std-dev




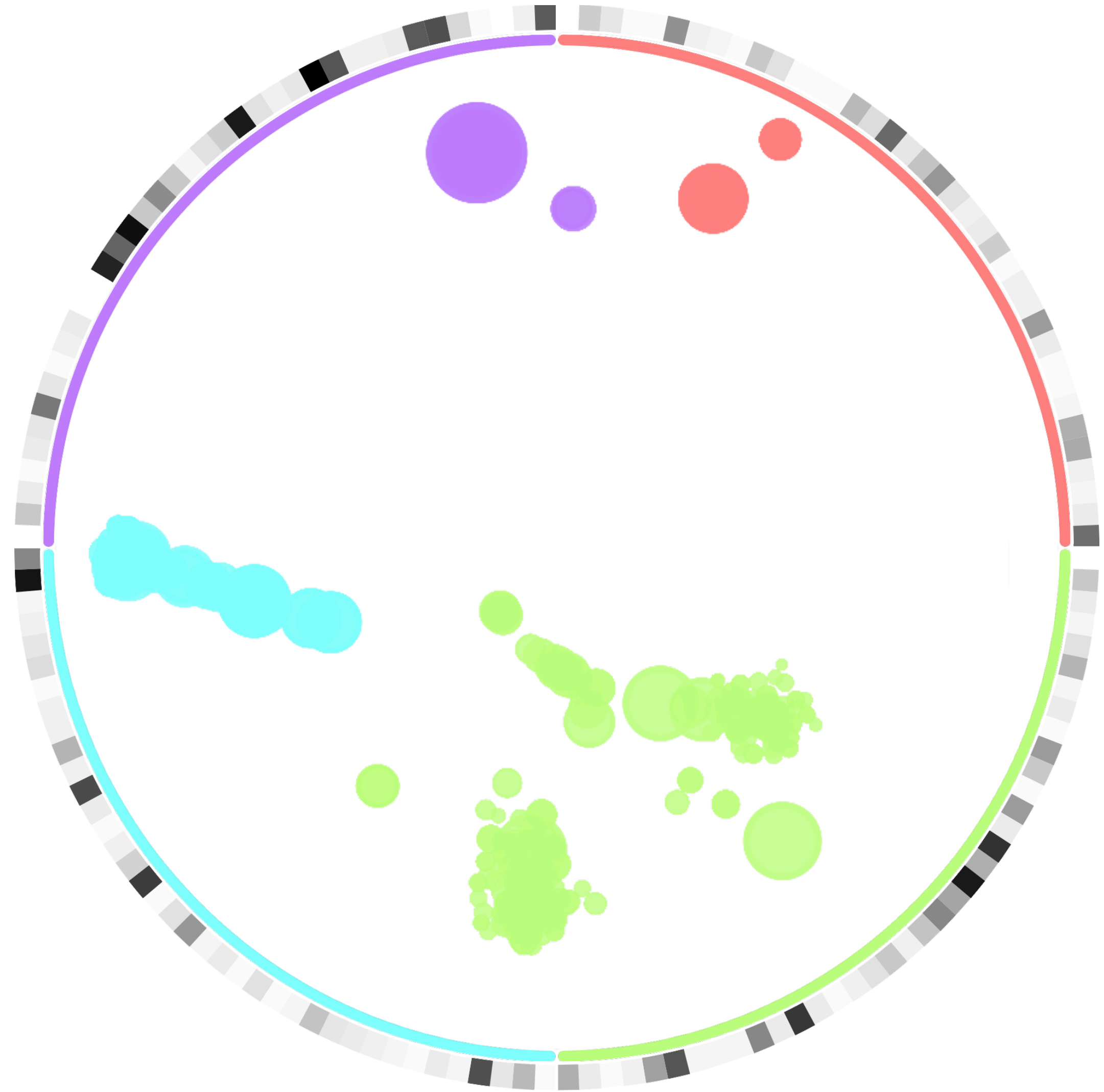
-  clusters
-  arc segment
-  marker heatmap

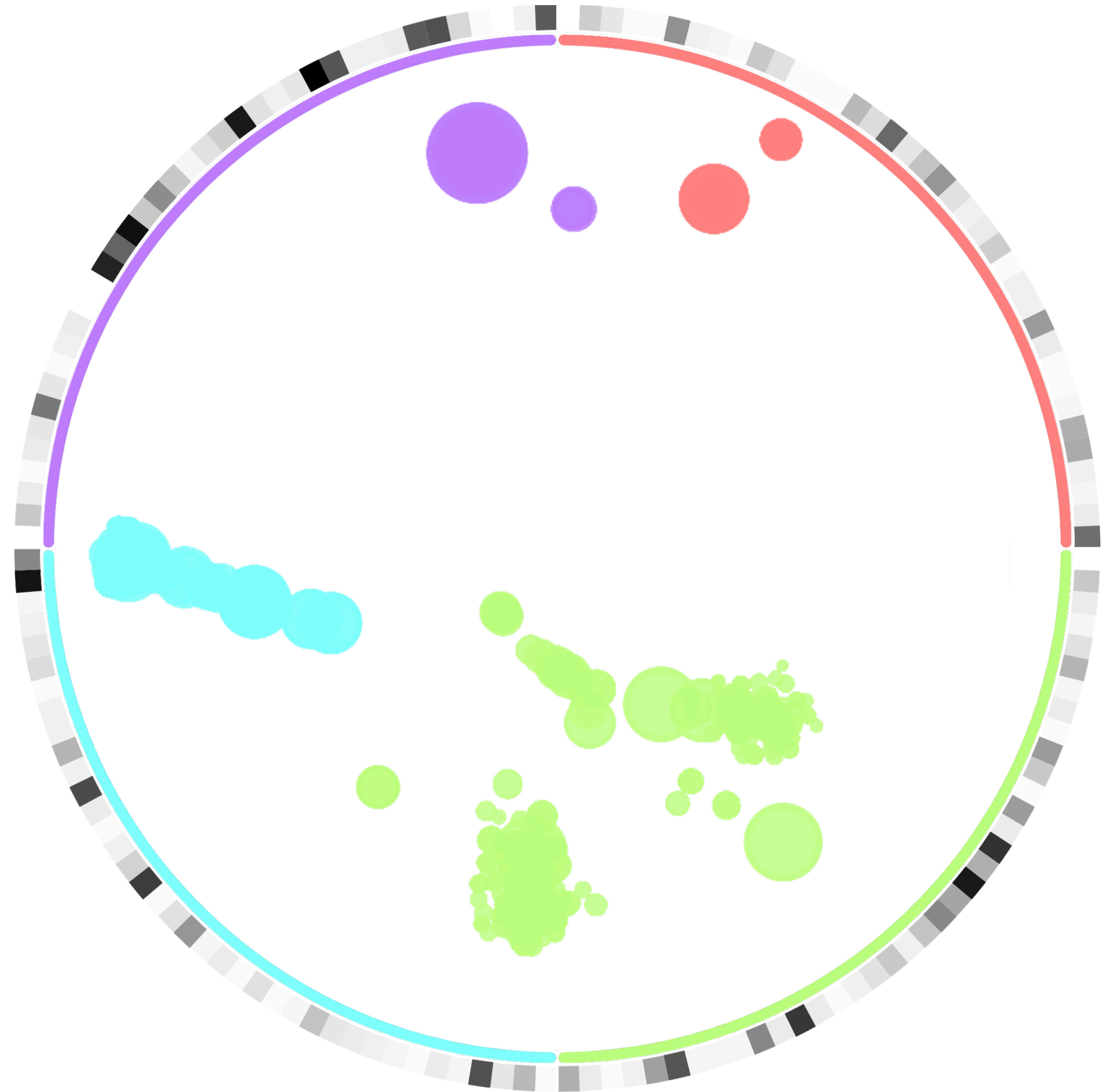
TCRgd
Expression: 1.248
Std-Dev: 0.933

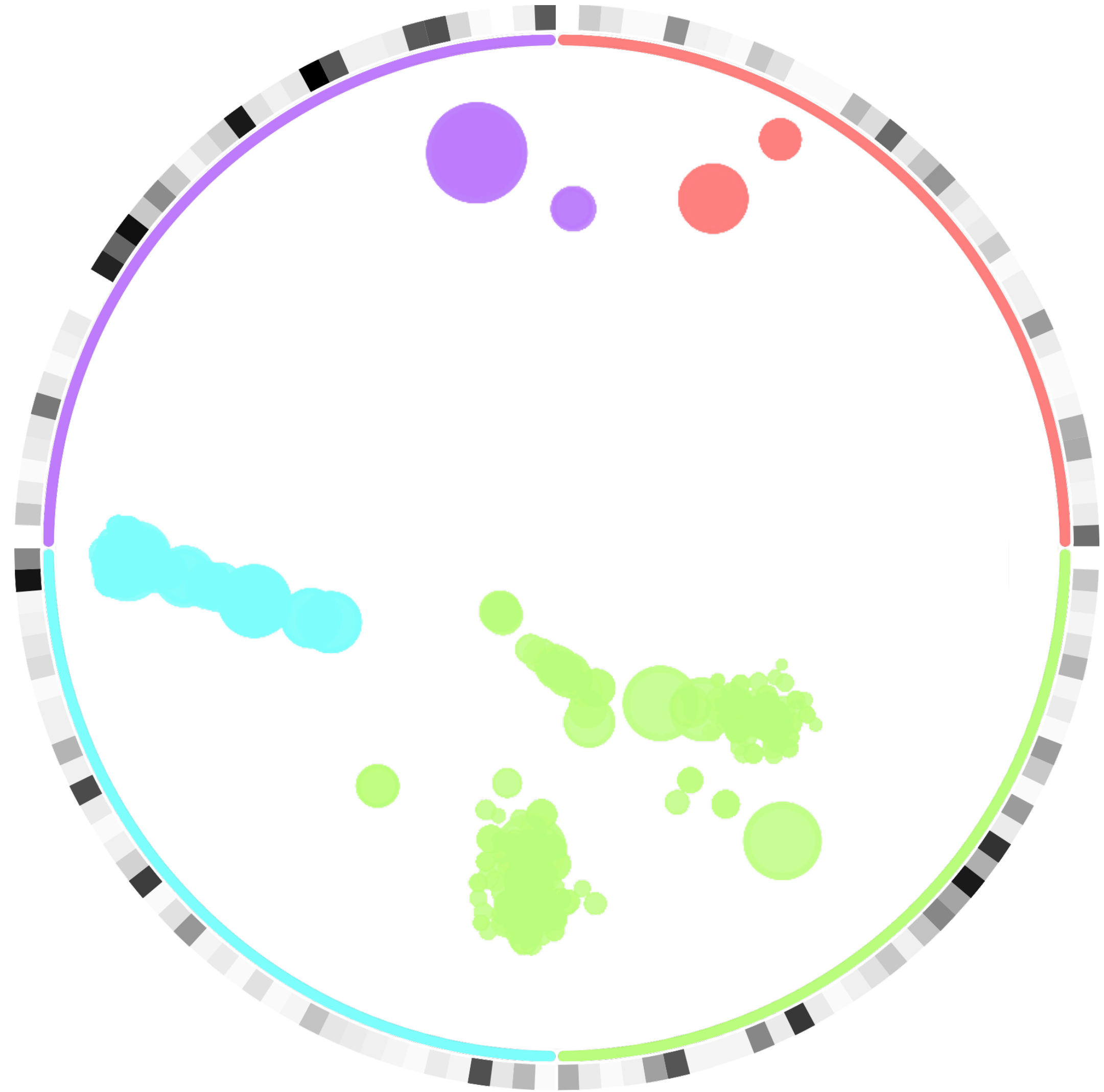
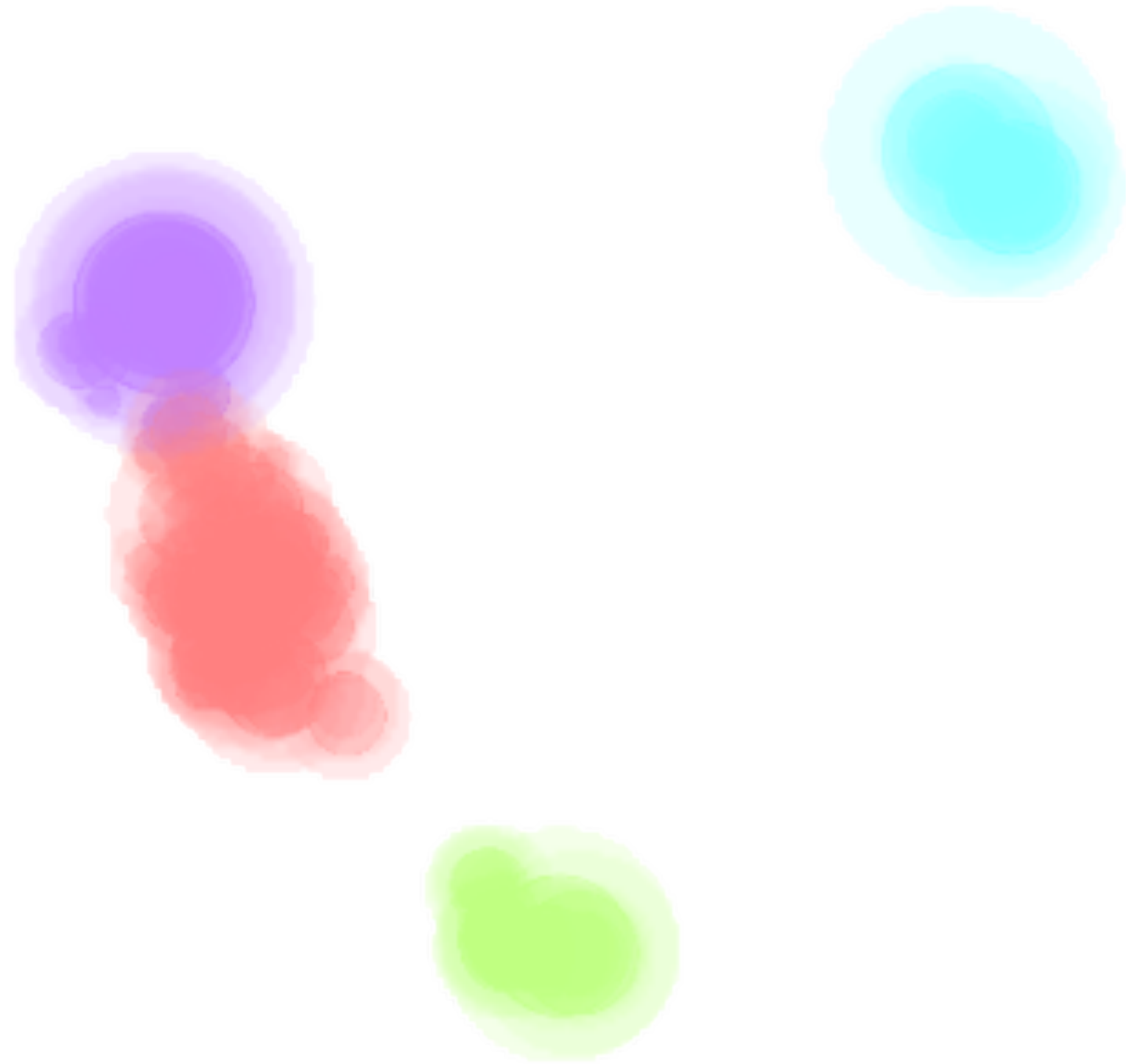


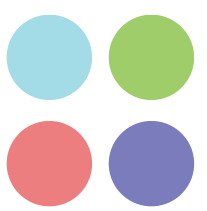


std-dev

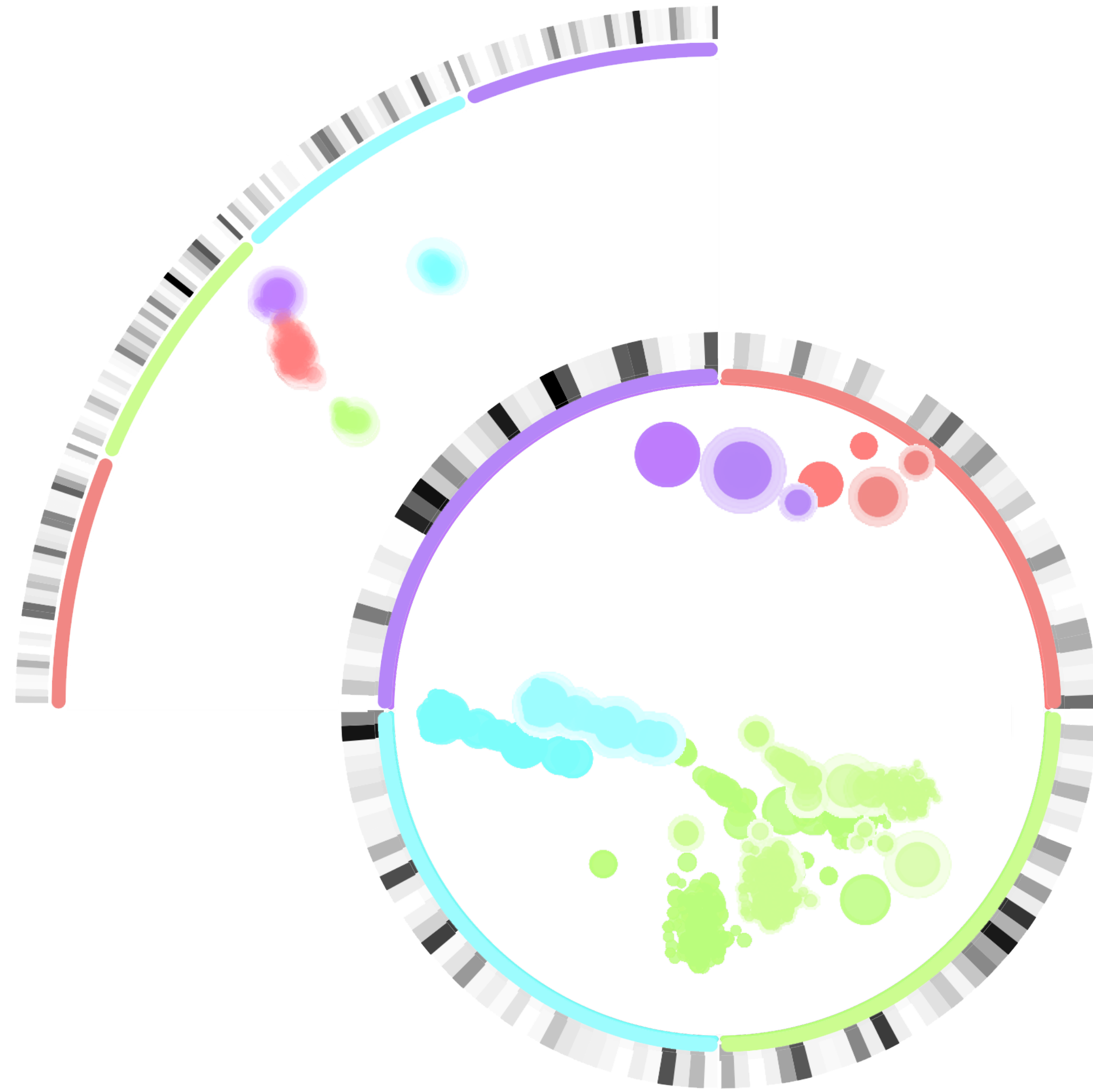


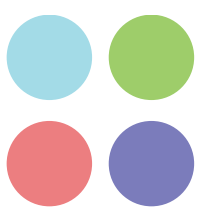




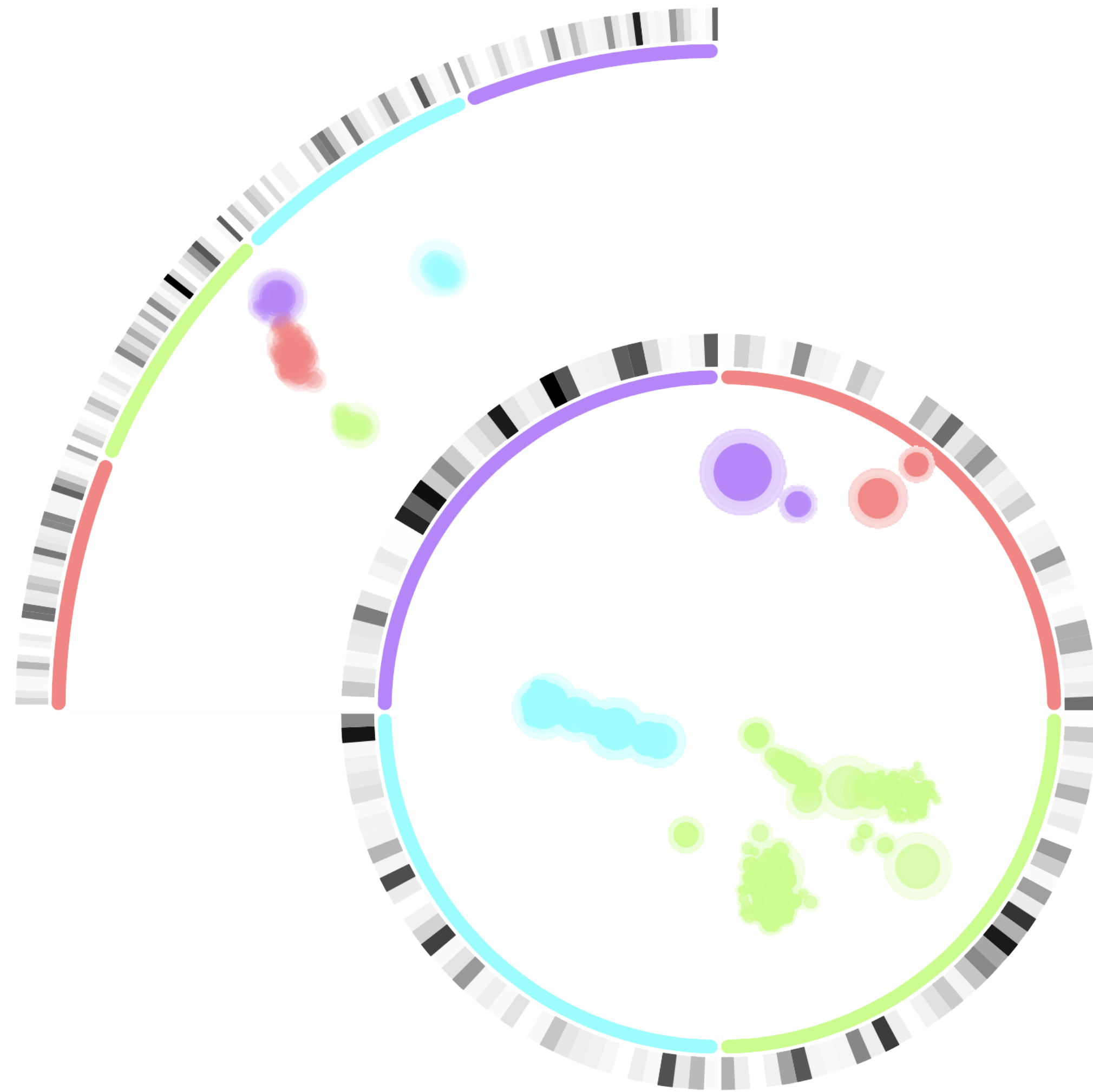


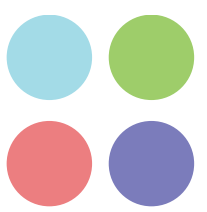




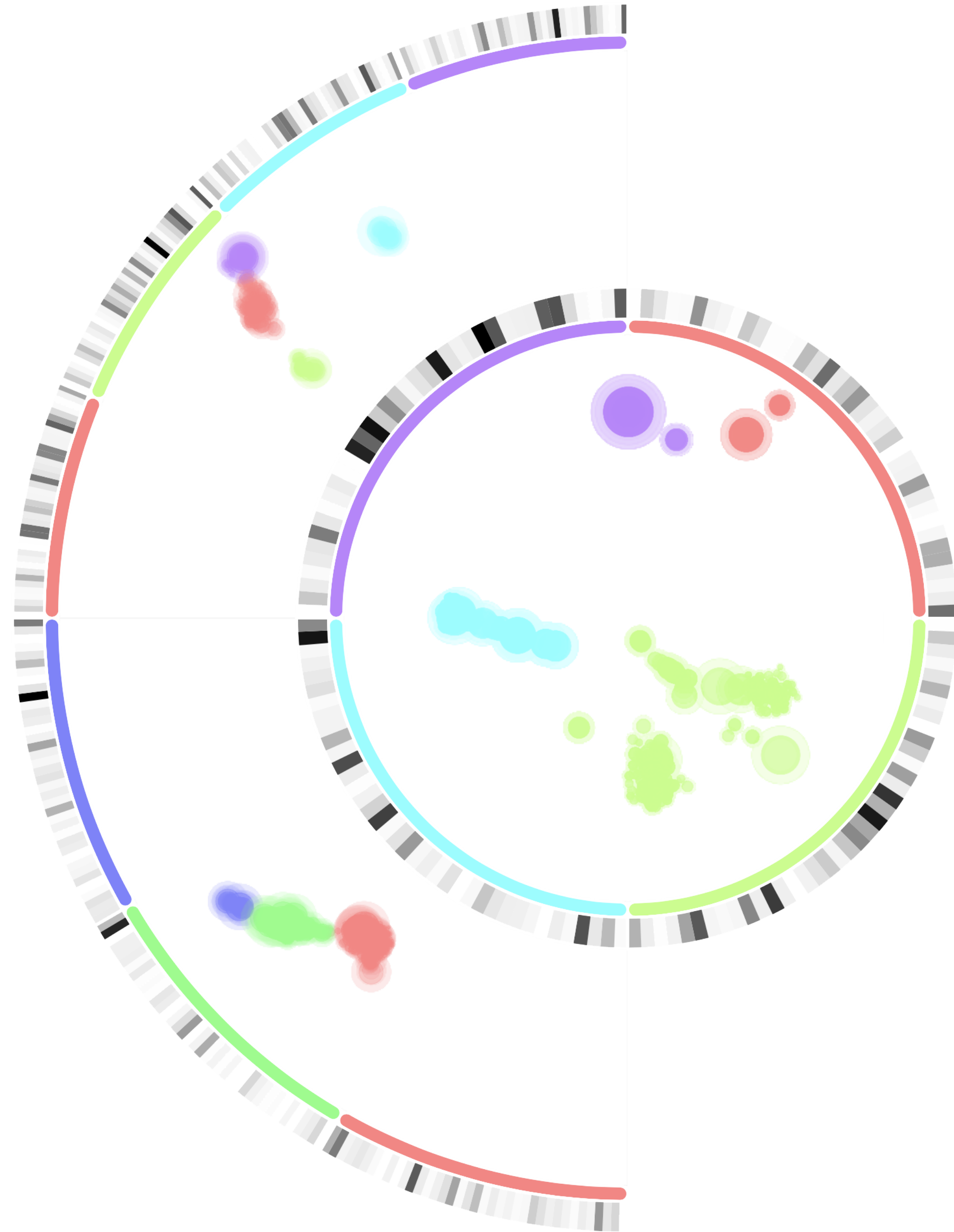
-  clusters
-  arc segment
-  marker heatmap

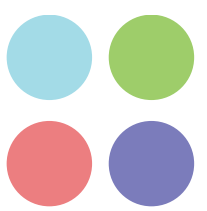




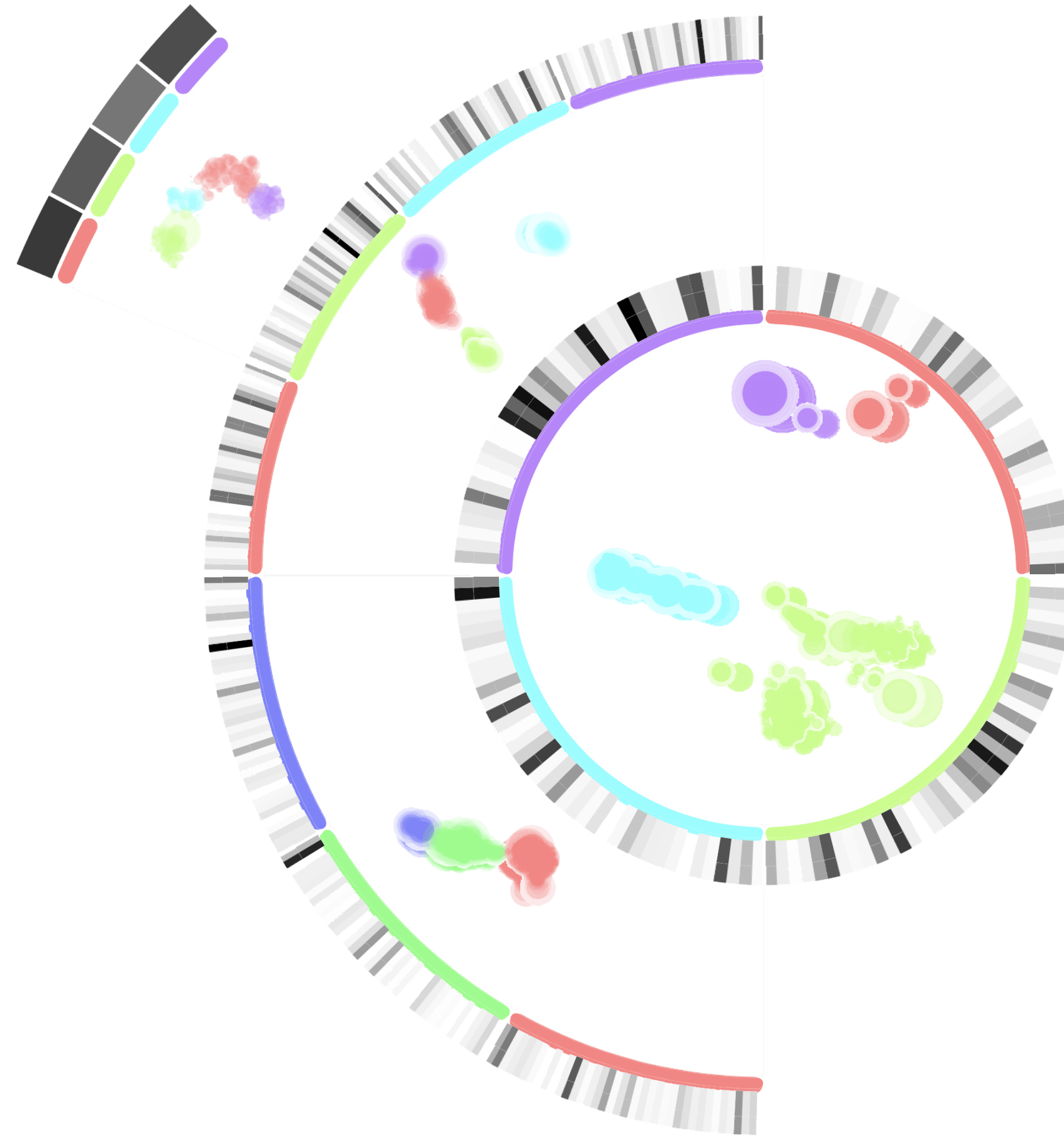
-  clusters
-  arc segment
-  marker heatmap

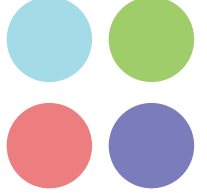




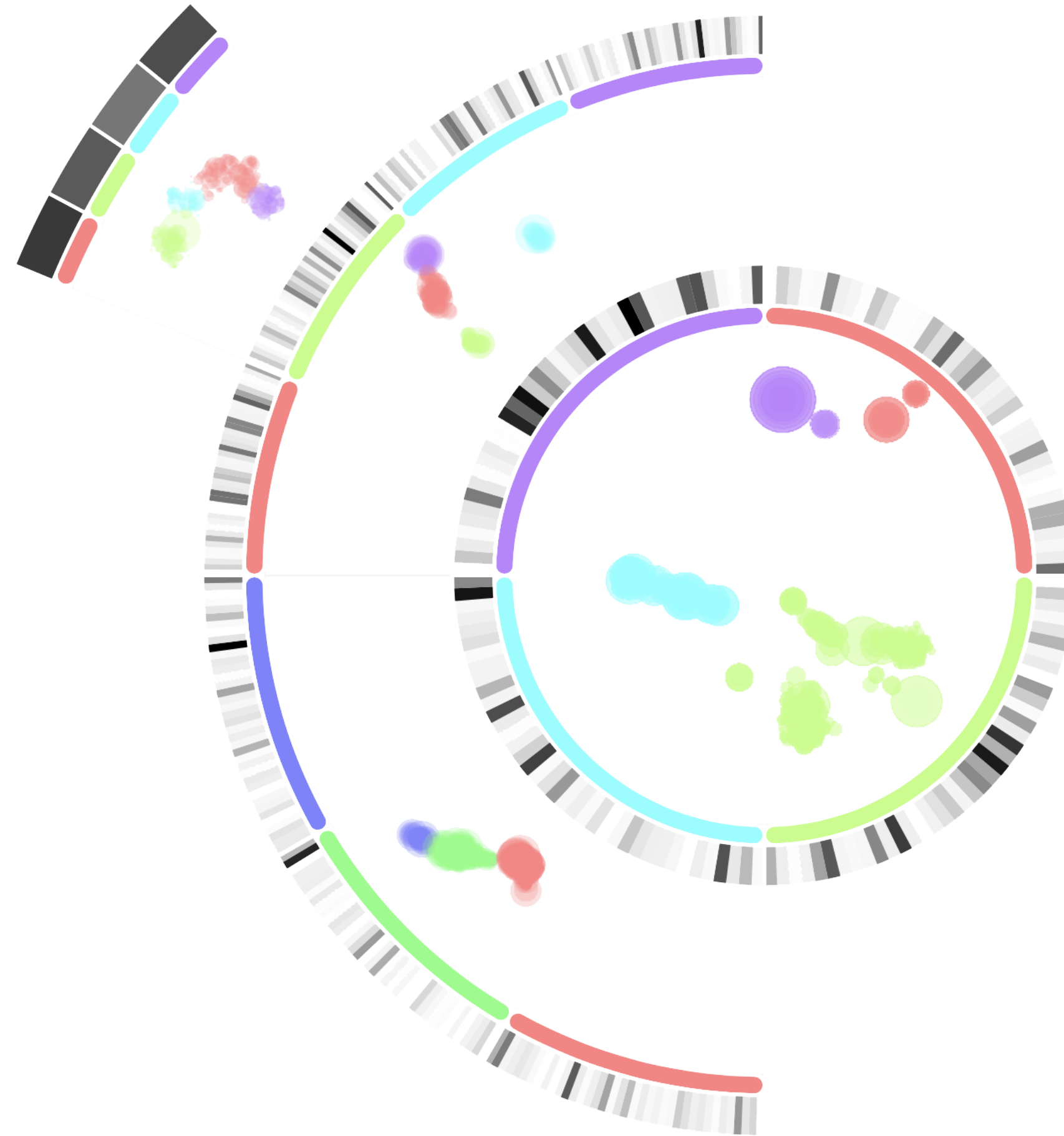
-  clusters
-  arc segment
-  marker heatmap

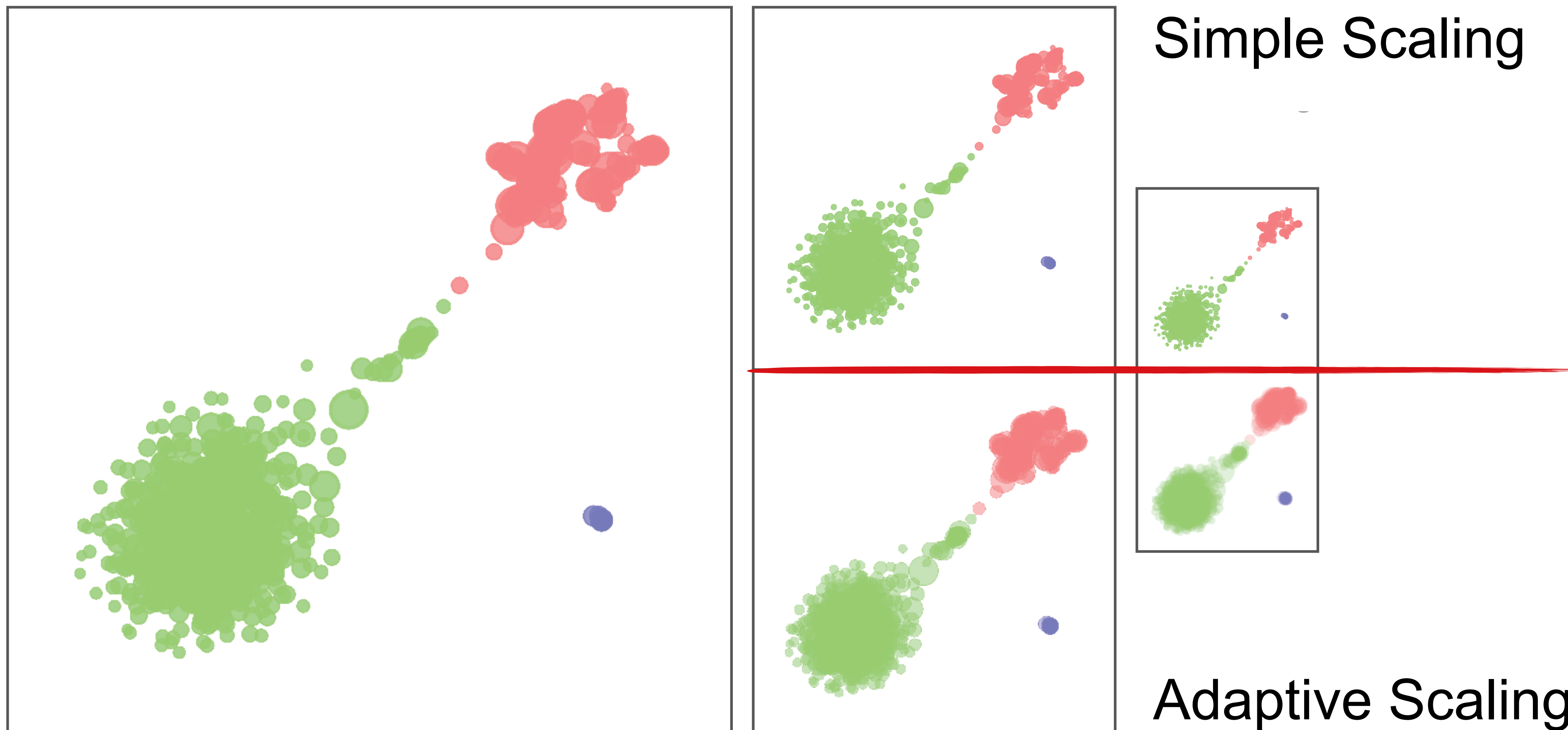


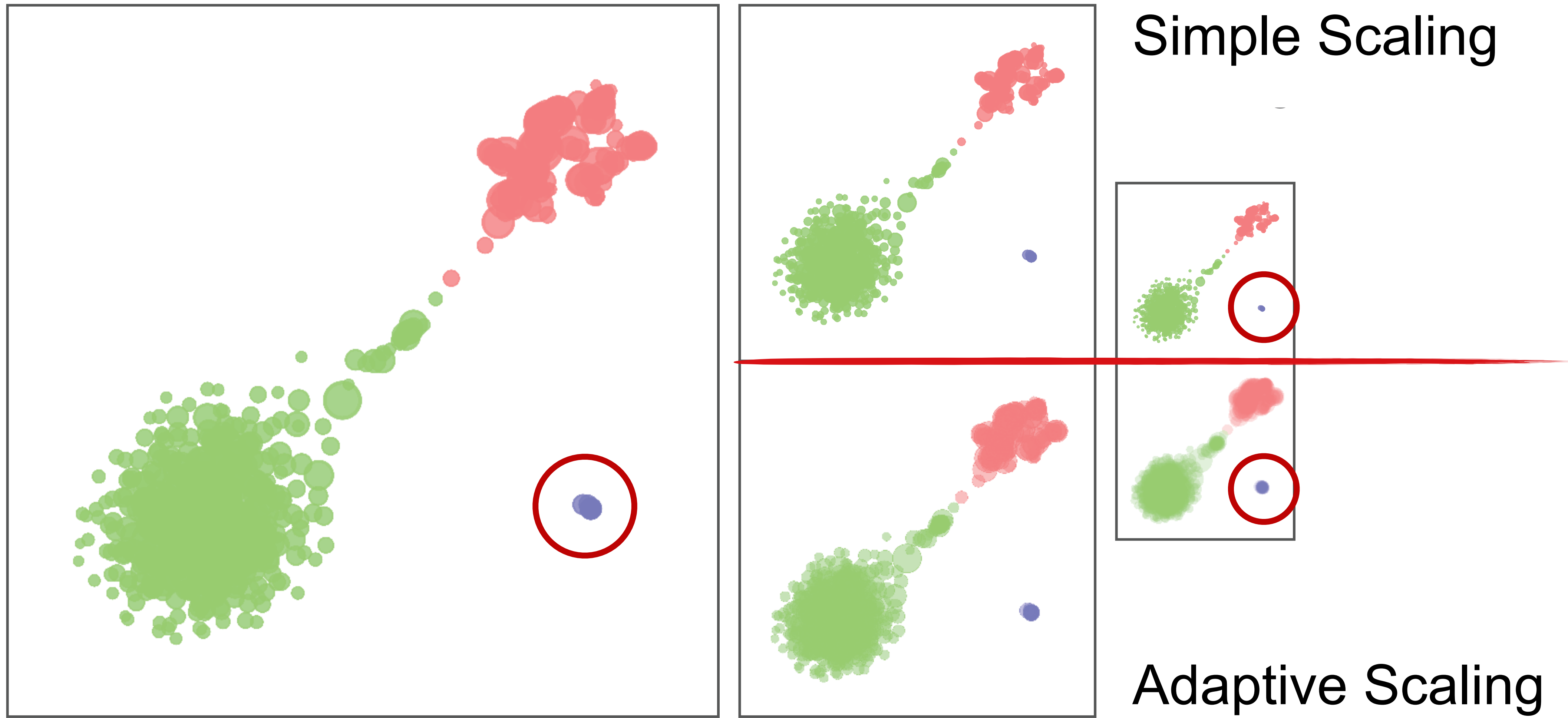
-  clusters
-  arc segment
-  marker heatmap



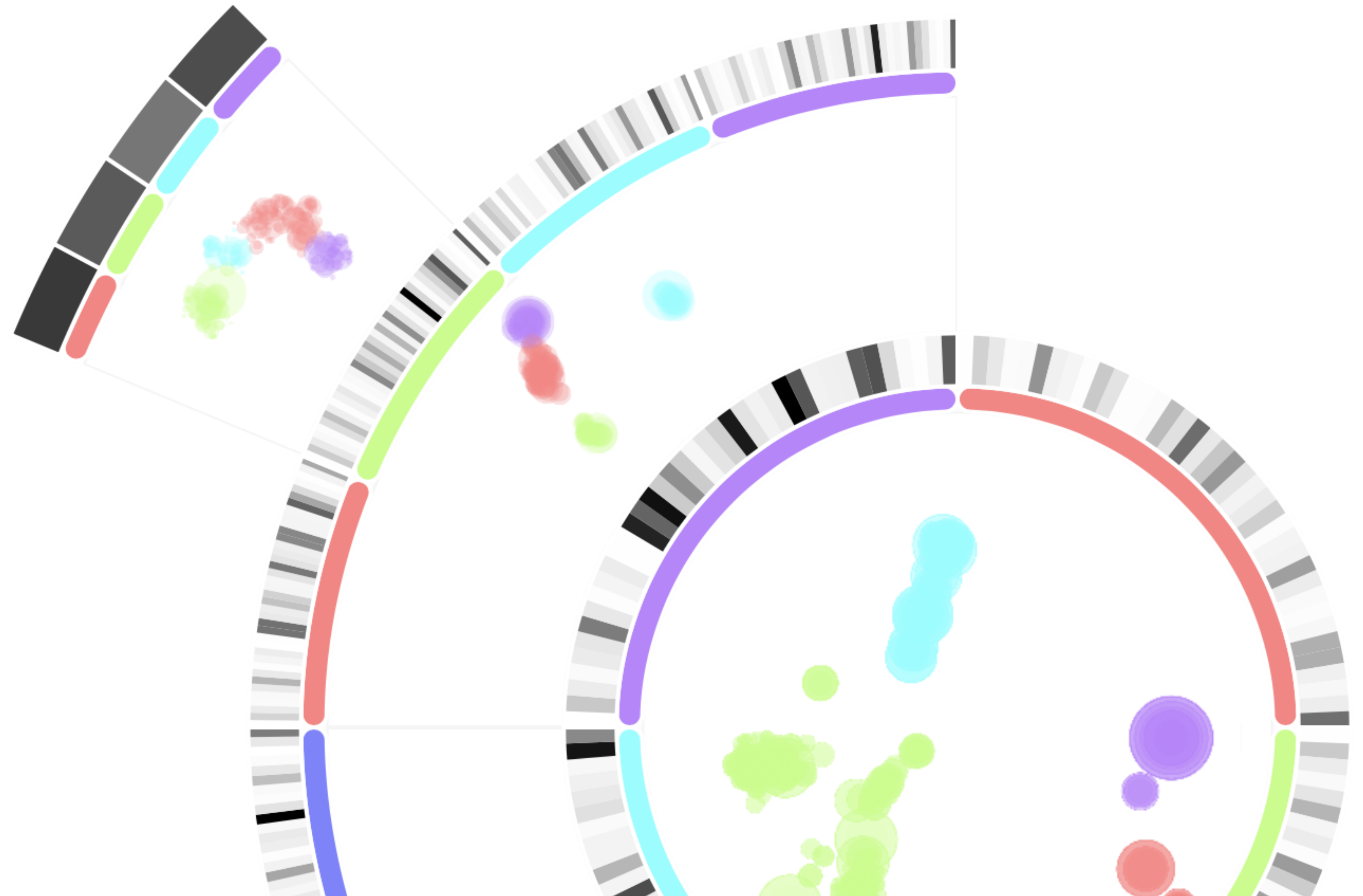
-  clusters
-  arc segment
-  marker heatmap



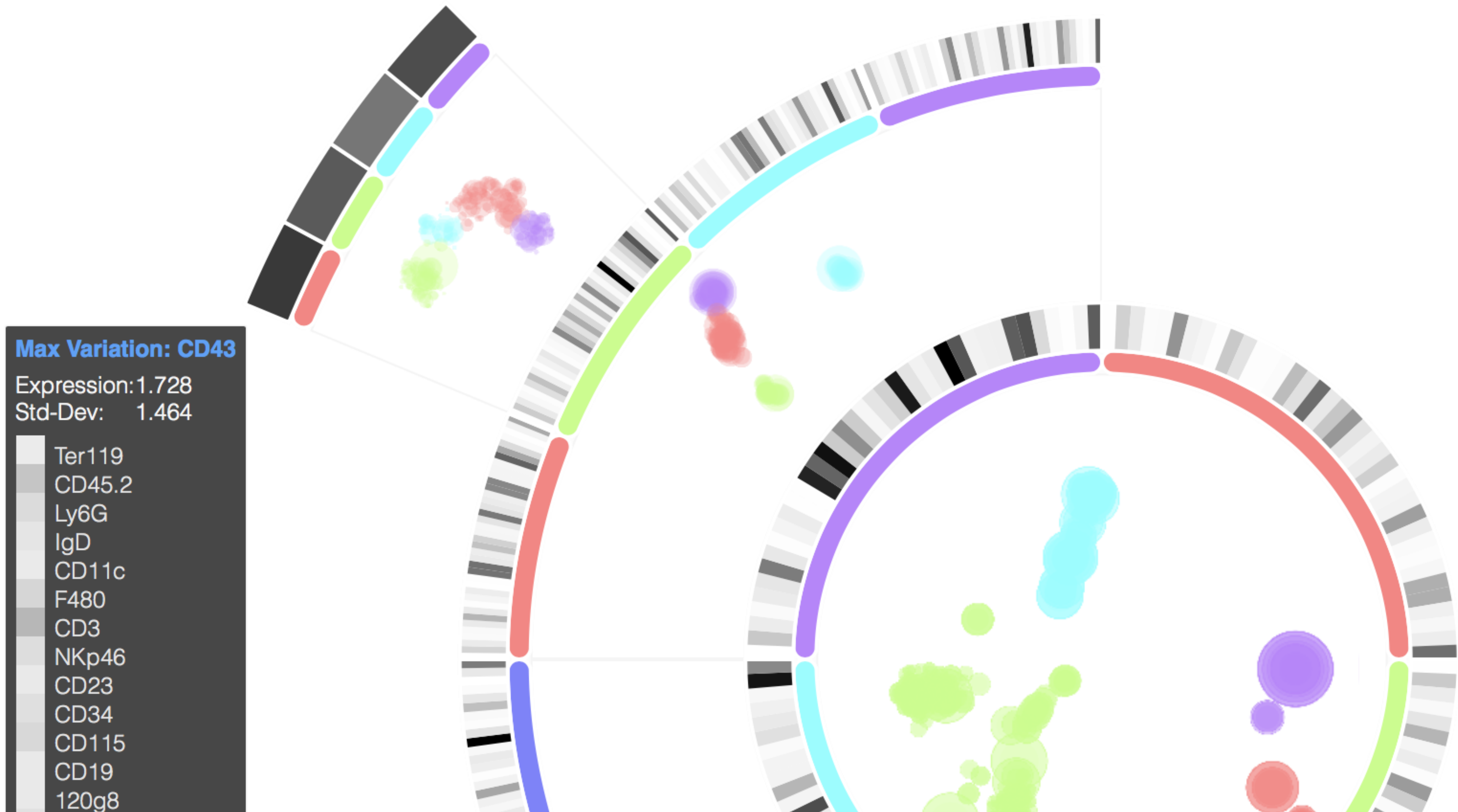




Adaptive Design - Heatmaps



Adaptive Design - Heatmaps



HSNE Analysis

Input Data

Active Markers

- Time
- Cell_length
- BC1
- BC2
- BC3
- BC4
- BC5
- BC6
- Ter119
- CD45.2
- Ly6G
- IgD
- CD11c
- F480
- CD3
- NKp46
- CD23
- CD34
- CD115
- CD19
- 120g8
- CD8
- Ly6C
- CD4
- CD11b
- CD27
- CD16_32
- SiglecF
- Foxp3
- B220
- CD5
- FceR1a
- TCRgd
- CCR7
- Sca1
- CD49b
- cKit
- CD150
- CD25
- TCRb
- CD43
- CD64
- CD138
- CD103
- IgM
- CD44
- MHCII
- DNA1
- DNA2
- Cisplatin
- beadDist

Inverse Markers

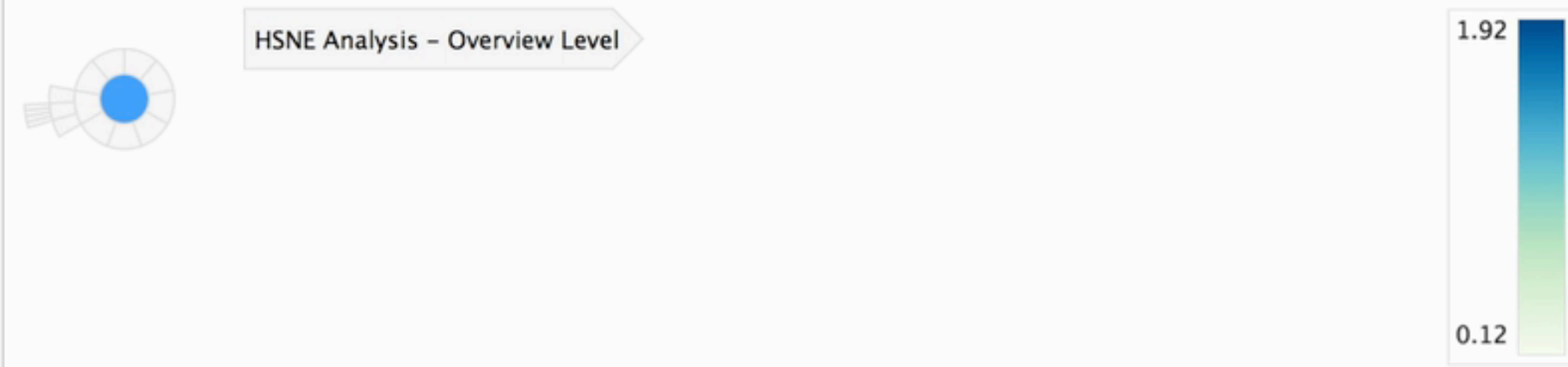
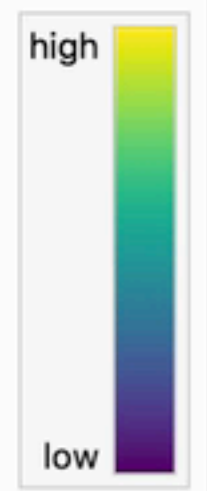
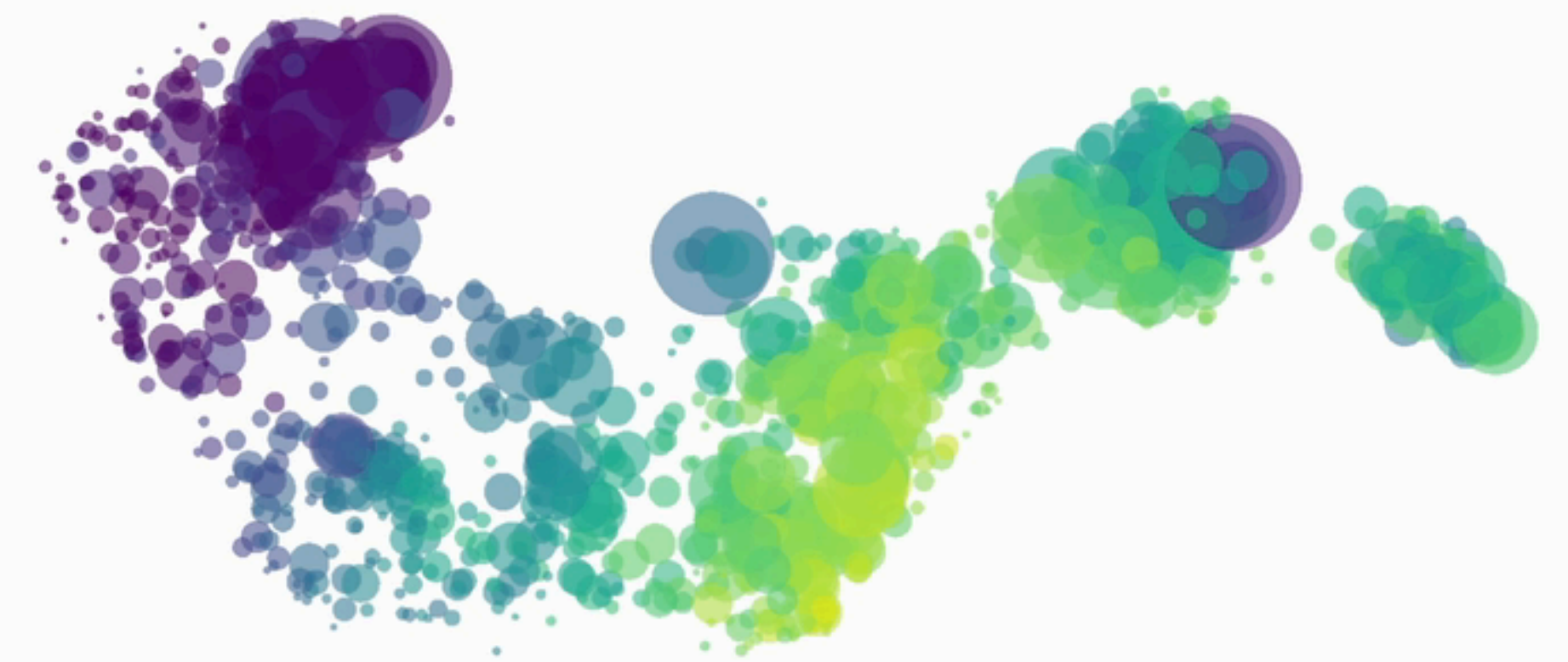
Load Selection Set Default

Advanced Parameters

Parameters

Number of Scales 4

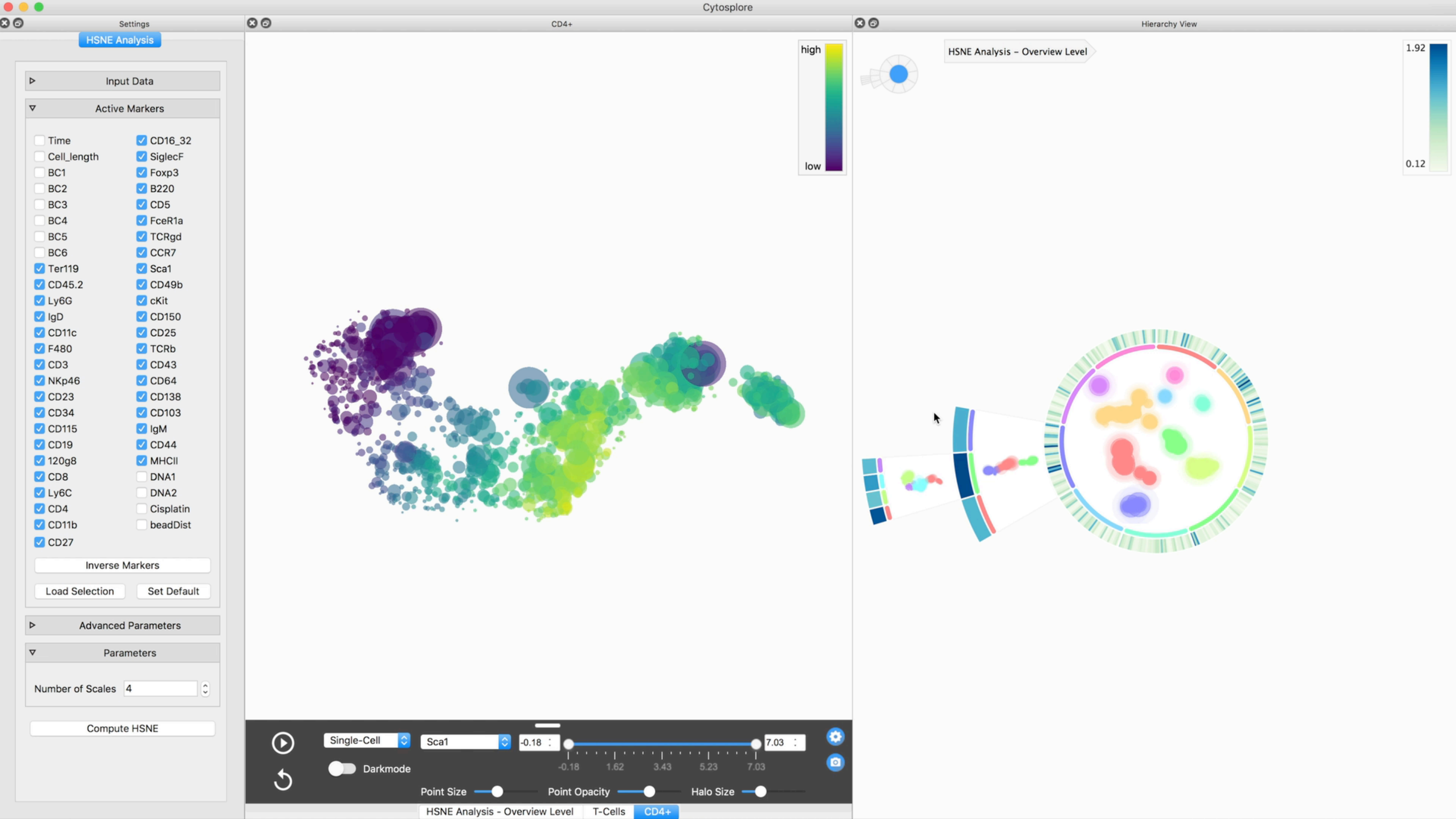
Compute HSNE

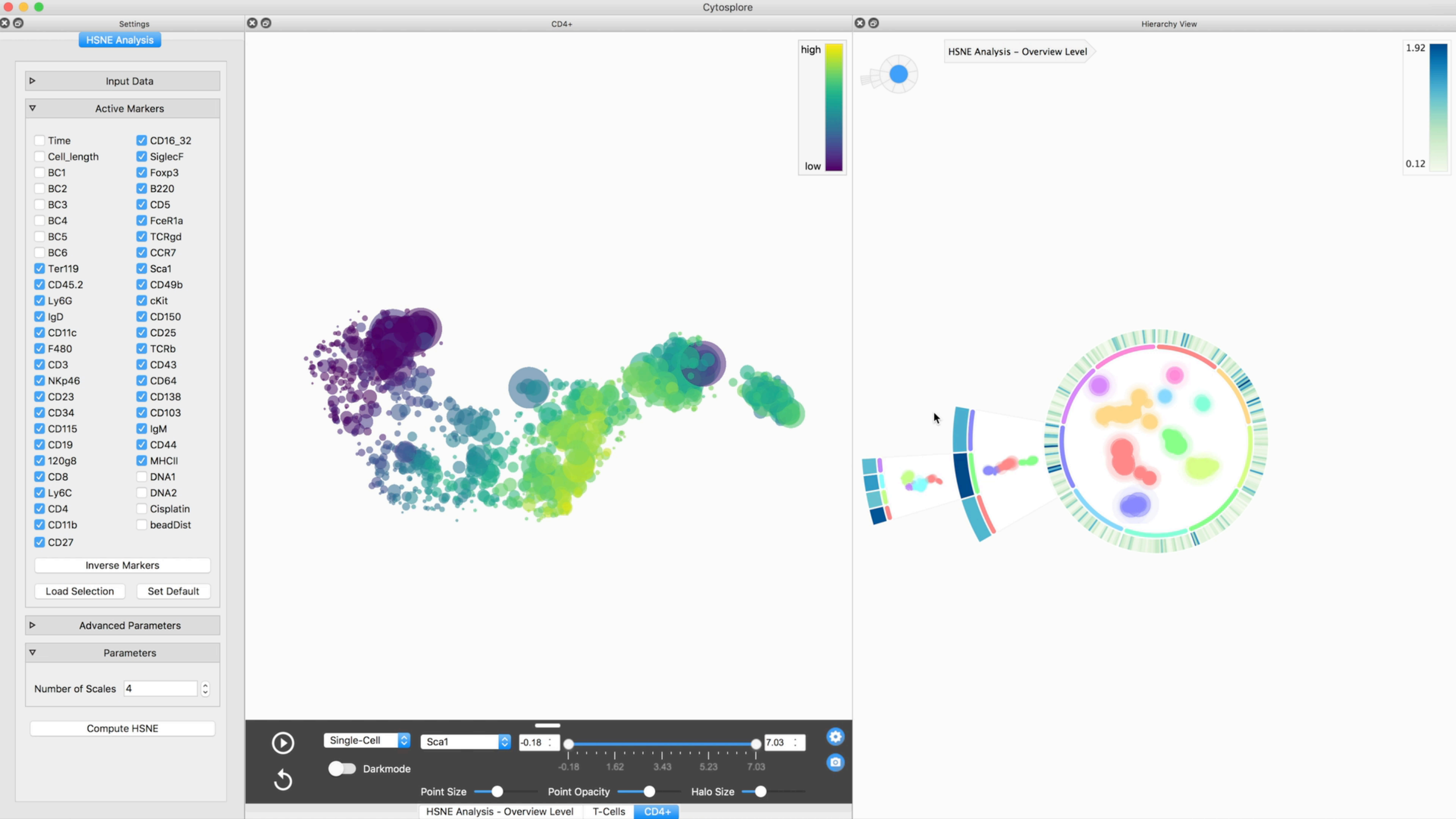


Single-Cell Sca1 -0.18 : 7.03

Darkmode

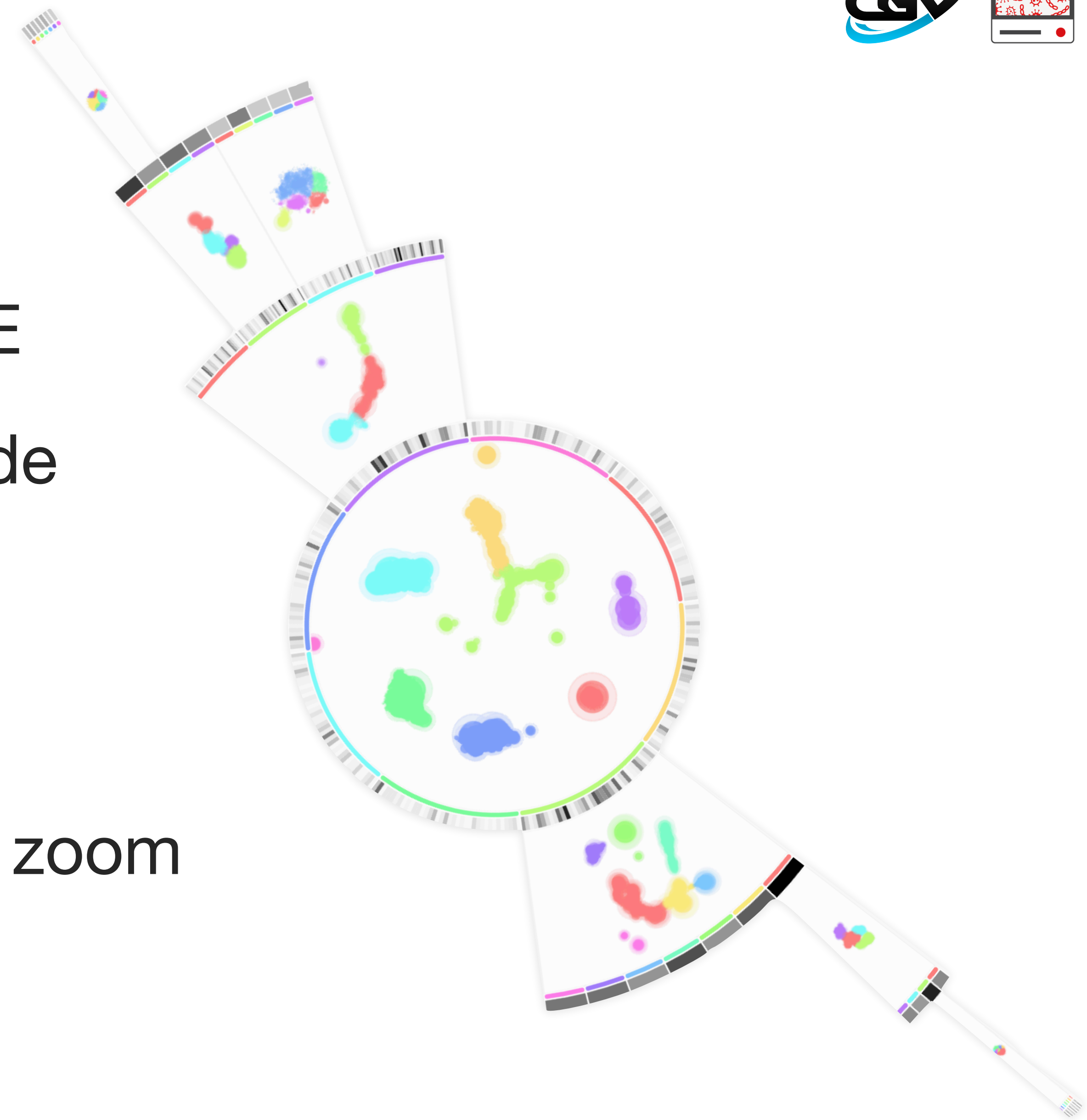
Point Size Point Opacity Halo Size



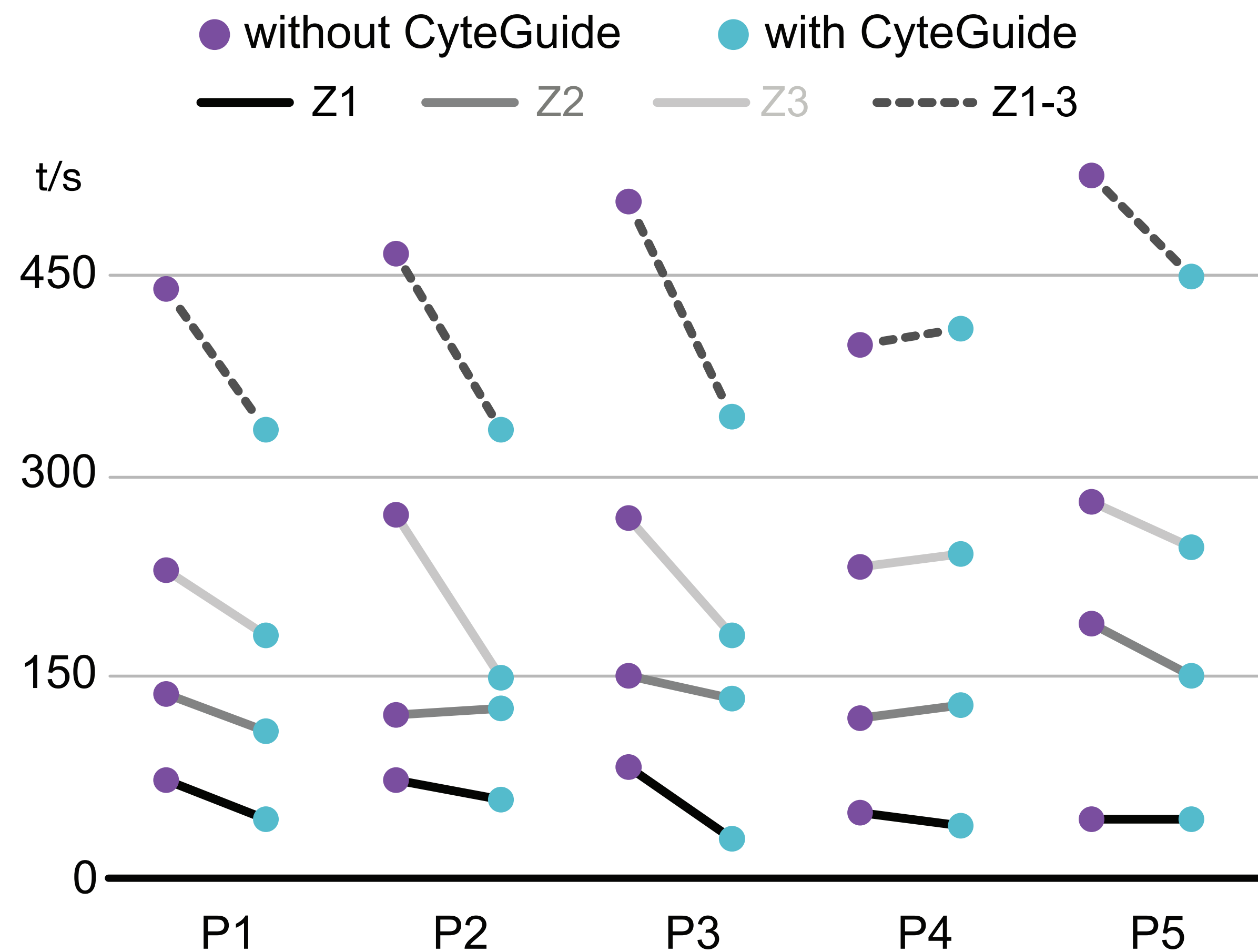


Evaluation

- 5 domain expert participants
- Familiar with Cytosplore and HSNE
- 30 minute introduction to CyteGuide
- Exemplative, small exploration
- ~800k cells, 39D, 4 level hierarchy
- 3x repeated cluster of interest id + zoom



- Generally navigation is faster
- Outlier P4
- not part of the field study
- Participants more confident
- Easy to keep track



- Generally navigation is faster
- Outlier P4
 - not part of the field study
- Participants more confident
- Easy to keep track

“I enjoyed analyzing the data.”

*“I am eager to start analyzing
my data with CyteGuide.”*

*“It’s easier [..] to keep the overview
of the different scales and clusters.”*

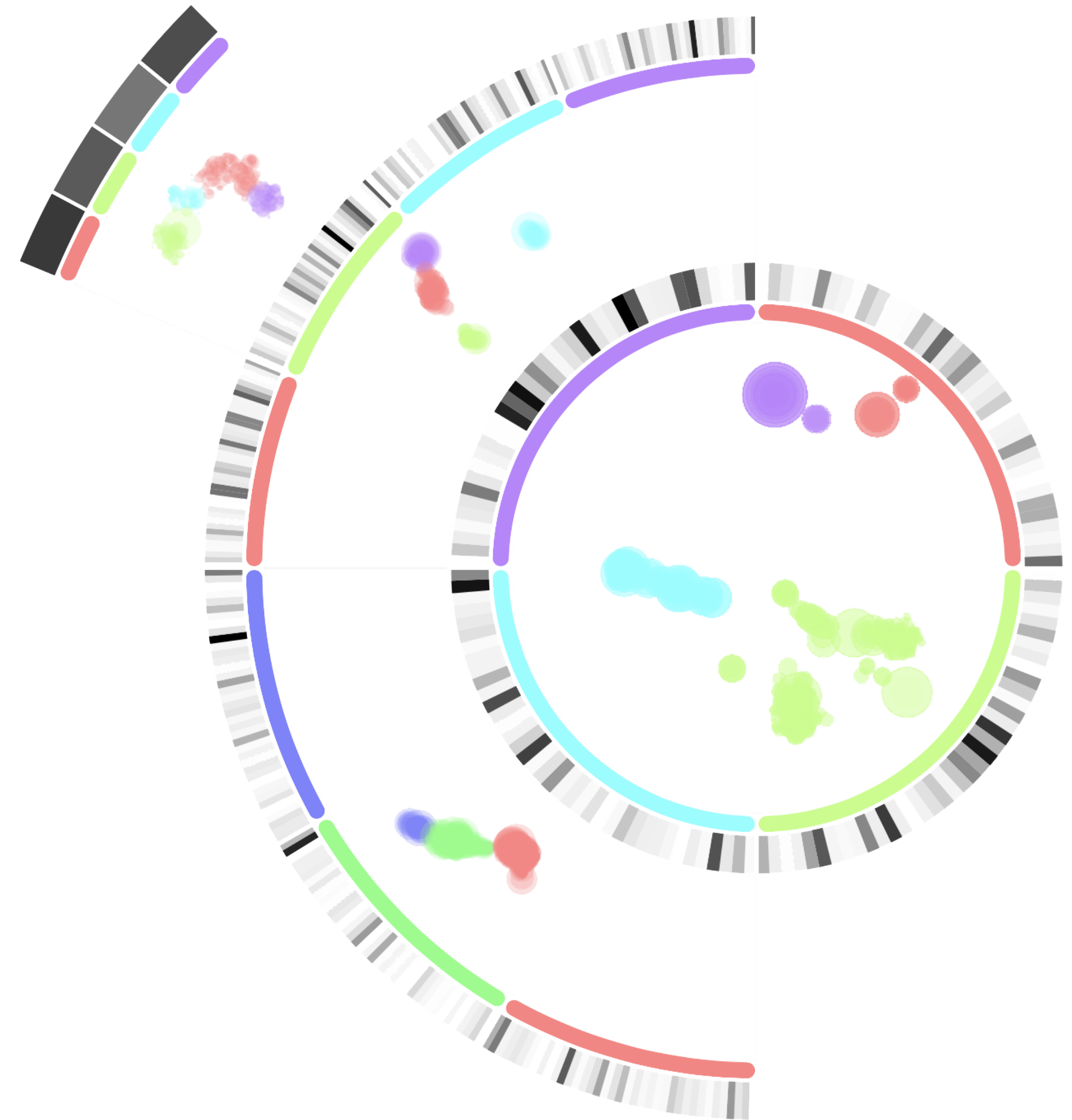
*“It is very useful to know whether
a cell cluster still contains variability,
and if drilling is necessary.”*

CyteGuide - vis for hierarchy exploration

- Eases navigation
- Increases confidence

Applicable to other hierarchies:

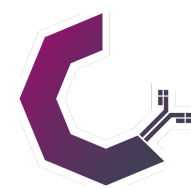
- Embedded visualizations
- Abstract feature space
- Limited number of features (of interest)



Acknowledgements:

STW grant 12720 VAnPIRe

LKEB & IHB @ Leiden University Medical Centre



[cyteguide.]cytosplore.org



lcbc.nl



graphics.tudelft.nl