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

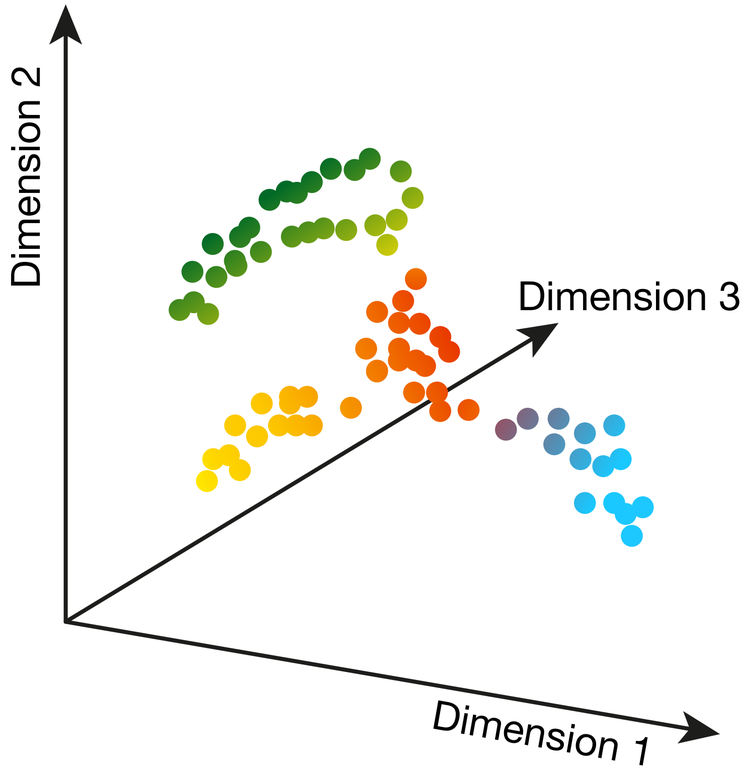
CyteGuide: Visual Guidance for Hierarchical Single-Cell Analysis

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

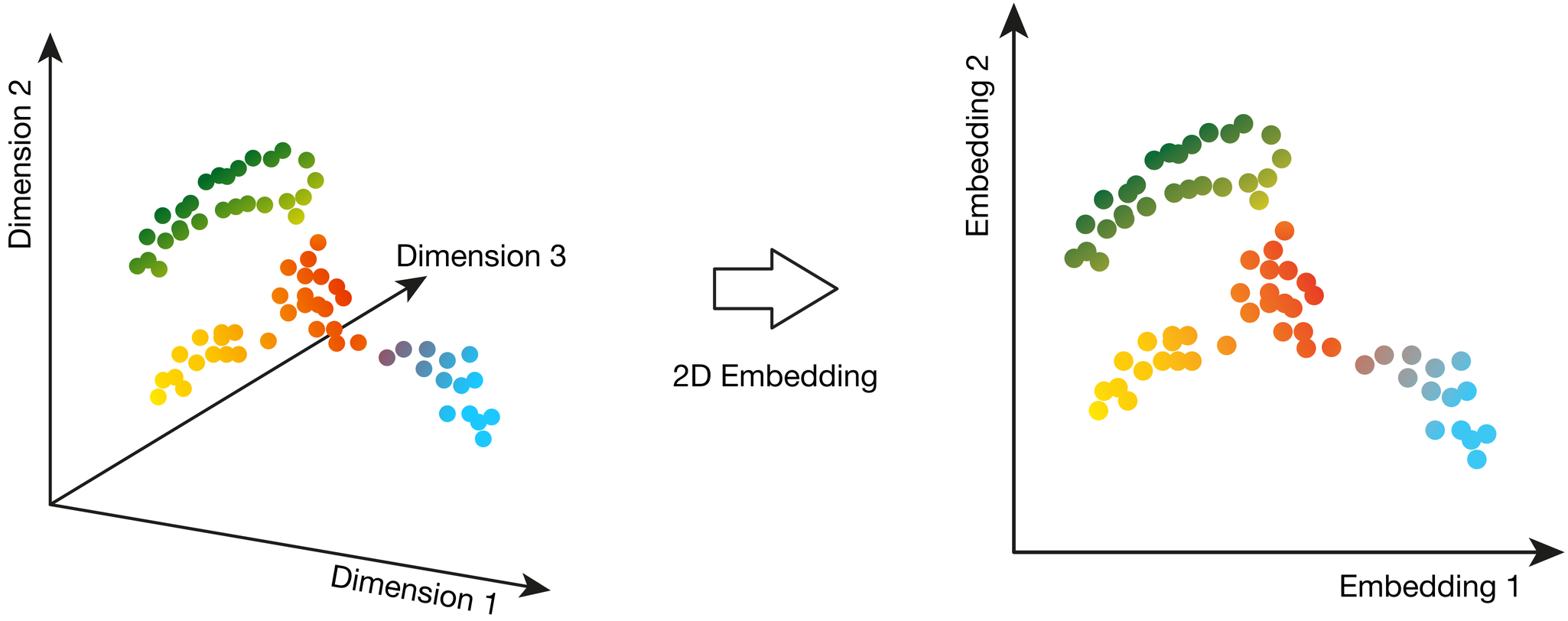
♥ TU Delft
♠ Leiden University MC


VIS 2017
LSTMVis: A Tool for Visual Analysis of Hidden State Dynamics in Recurrent Neural Networks
Hendrik Strobelt, Sebastian Gehrmann, Hanspeter Pfister, and Alexander M. Rush
Skeleton-based Scagnostics
José Matute, Alexandru C. Telea, and Lars Linsen
Visual Diagnosis of Tree Boosting Methods
Shixia Liu, Jiannan Xiao, Junlin Liu, Xiting Wang, Jing Wu, Jun Zhu
ACTIVIS: Visual Exploration of Industry-Scale Deep Neural Network Models
Minsuk Kahng, Pierre Y. Andrews, Aditya Kalro, and Duen Horng (Polo) Chau
DeepEyes: Progressive Visual Analytics for Designing Deep Neural Networks
Nicola Pezzotti, Thomas Höllt, Jan van Gemert, Boudewijn P.F. Lelieveldt, Elmar Eisemann, Anna Vilanova
VIGOR: Interactive Visual Exploration of Graph Query
Robert Pienta, Fred Hohman, Alex Endert, Acar Tamersoy, Kevin Roundy, Chris Gates, Shamkant Navathe, Duen Horng Chau
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
Visual Exploration of Semantic Relationships in Neural Word Embeddings
Shusen Liu, Peer-Timo Bremer, Jayaraman J. Thiagarajan, Vivek Srikumar, Bei Wang, Yarden Livnat and Valerio Pascucci
LDSScanner: Exploratory Analysis of Low-Dimensional Structures in High-Dimensional Datasets
Jiazhi Xia, Fenjin Ye, Wei Chen*, Yusi Wang, Weifeng Chen, Yuxin Ma, and Anthony K.H. Tung
SkyLens: Visual Analysis of Skyline on Multi-dimensional Da
Xun Zhao, Yanhong Wu, Weiwei Cui, Xinnan Du, Yuan Chen, Yong Wang, Dik Lun Lee, and Huamin Qu
CyteGuide: Visual Guidance for Hierarchical Single-Cell Analysis
Thomas Höllt, Nicola Pezzotti, Vincent van Unen, Frits Koning, Boudewijn P.F. Lelieveldt, and Anna Vilanova
HiPiler: Visual Exploration of Large Genome Interaction Matrices with Interactive Small Multiples
Fritz Lekschas, Benjamin Bach, Peter Kerpedjiev, Nils Gehlenborg, and Hanspeter Pfister
What Would a Graph Look Like in This Layout? A Machine Learning Approach to Large Graph Visualization
Oh-Hyun Kwon, Student Member, IEEE, Tarik Crnovrsanin, and Kwan-Liu Ma, Fellow, IEEE
Towards a Systematic Combination of Dimension Reduction and Clustering in Visual Analytics
John Wenskovitch, Student Member, IEEE, Ian Crandell, Naren Ramakrishnan, Member, IEEE, Leanna House, Scotland Leman, Chris North
Comparing Visual-Interactive
An Experim
Jürgen Bernard, Marco Hutter, Matthias Zepp
Visual Supervision of Unsupervised Clustering
Bum Chul Kwon, Ben Eysenbach, Janu Verma, Ng, Christopher deFilippi, Walter F. Stewart, and Adam Perer
CGV

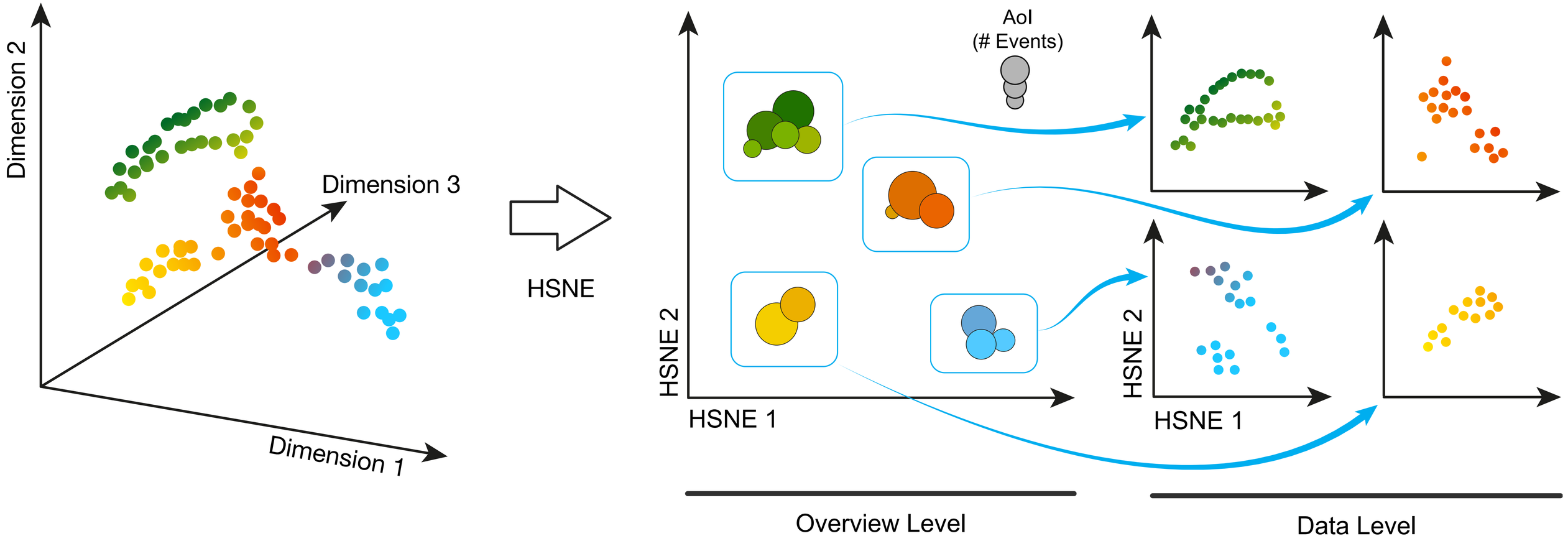
Dimensionality Reduction



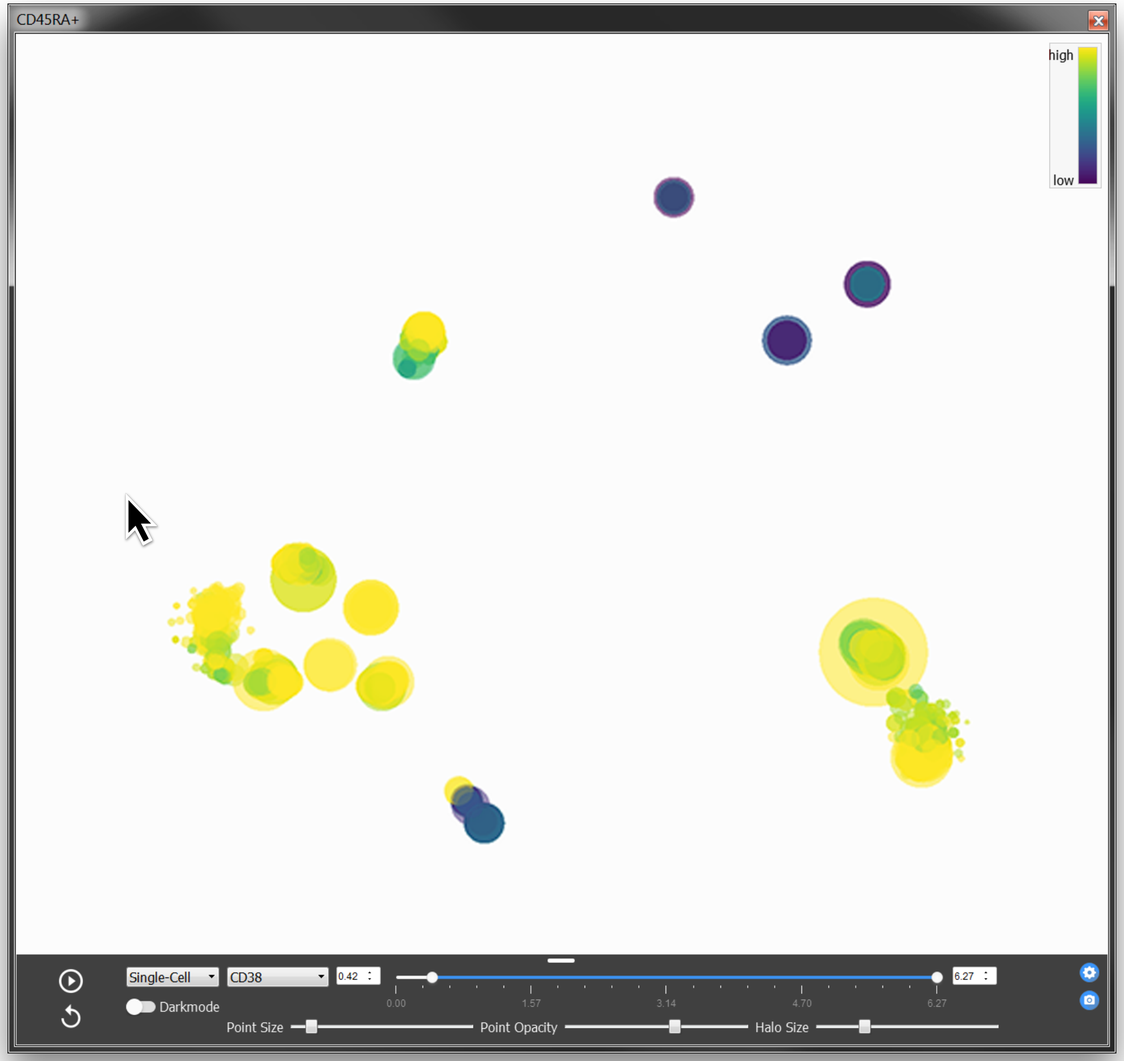
Dimensionality Reduction

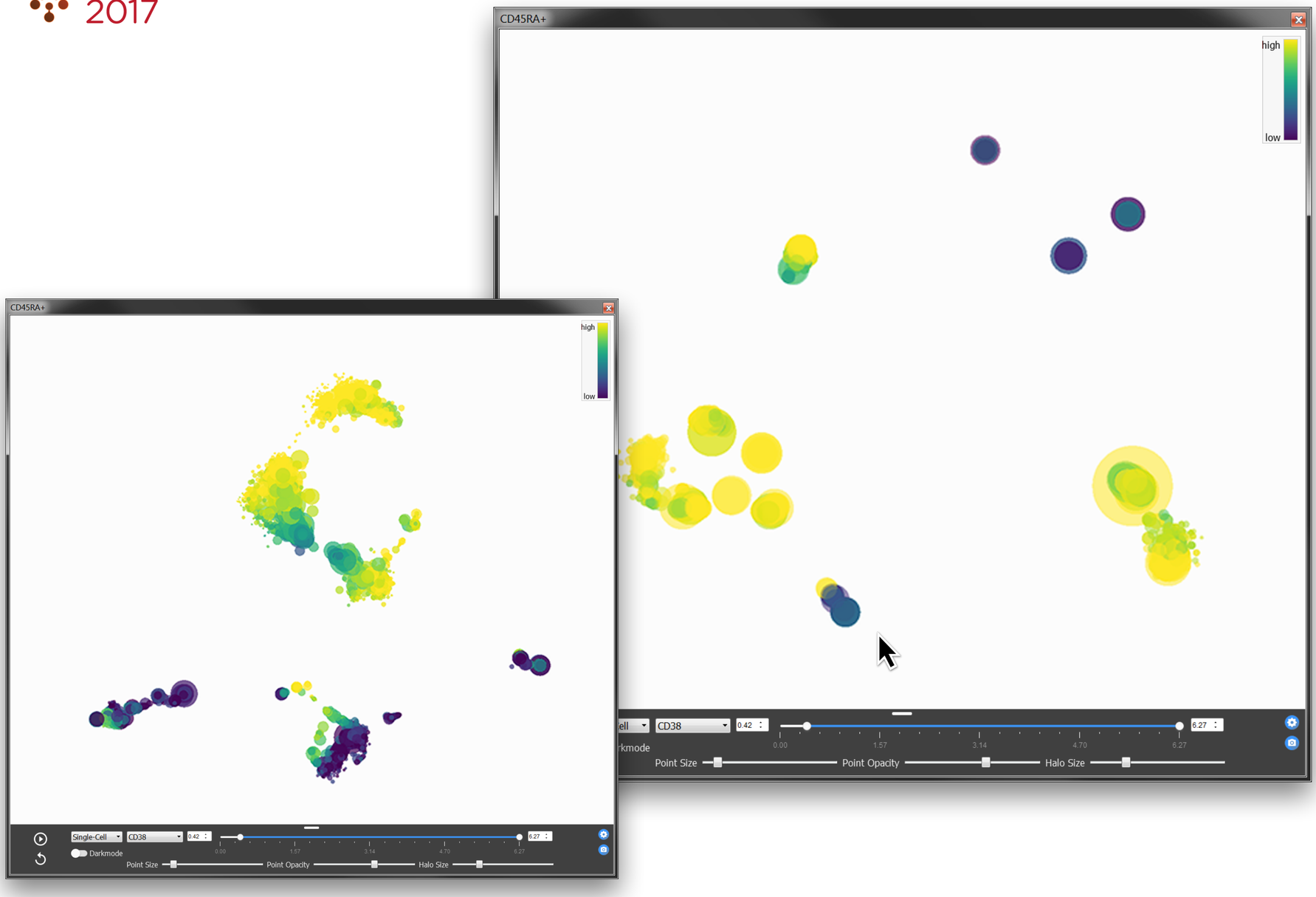


Hierarchical SNE

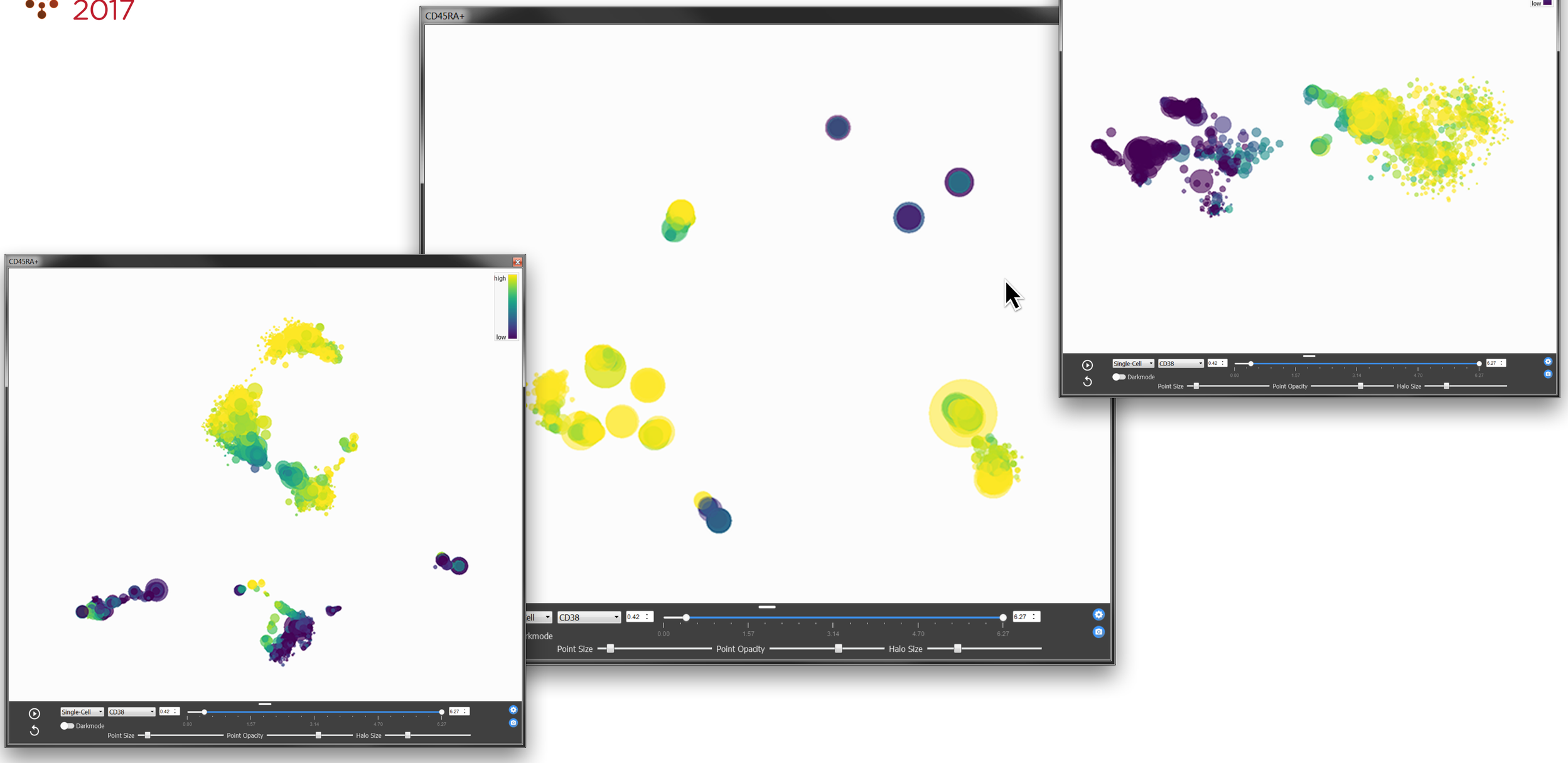


Pezzotti et al., EuroVis, 2016 || v. Unen, Höllt, Pezzotti et al., biorXiv 2017

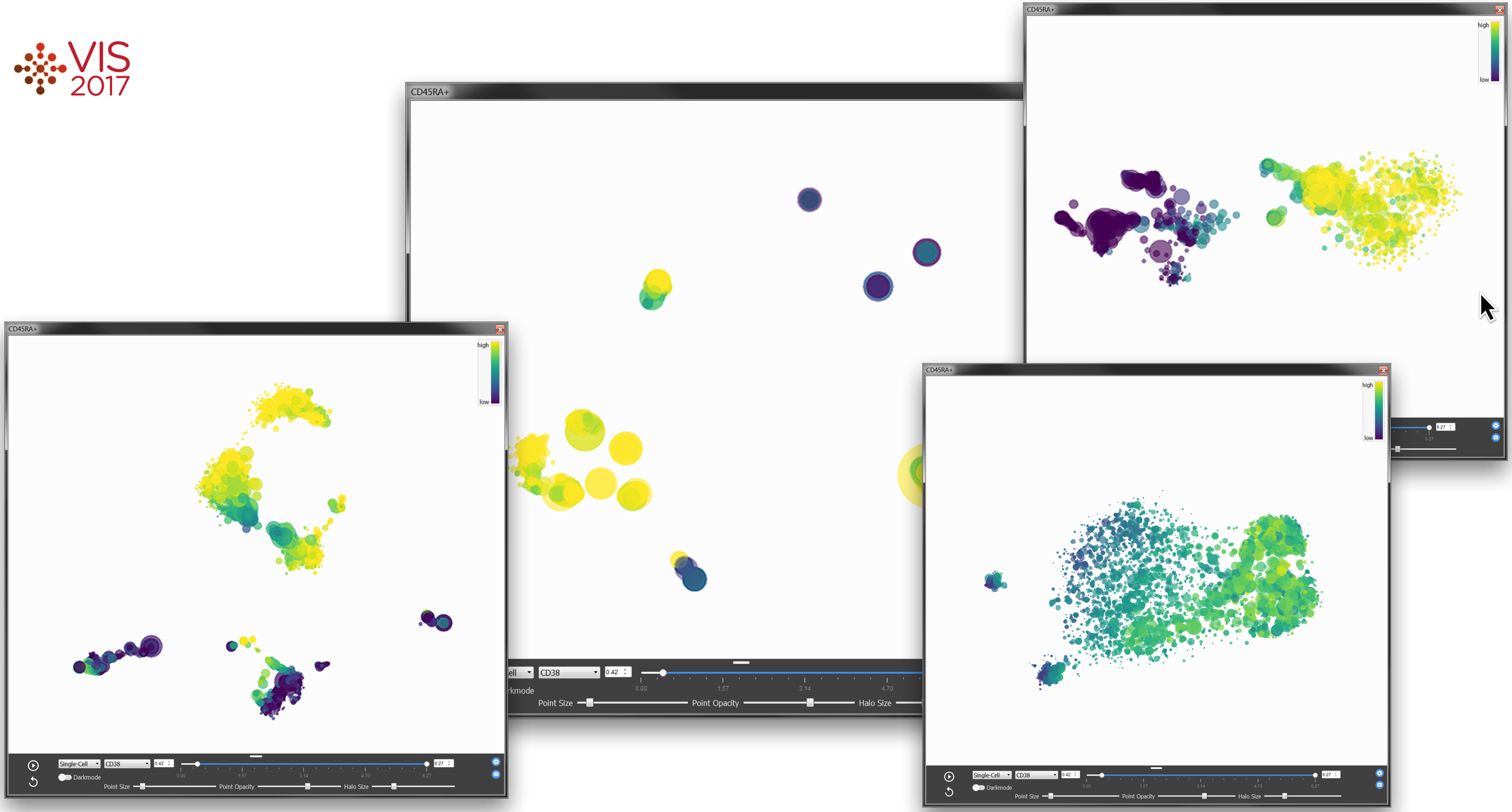

CD45RA+
high
low
Single-Cell
CD38
0.42
6.27
0.00
1.57
3.14
4.70
6.27
Darkmode
Point Size
Point Opacity
Halo Size


CD45RA+
high
low
Single-Cell
CD38
0.42
6.27
Darkmode
0.00
1.57
3.14
4.70
6.27
Point Size
Point Opacity
Halo Size

VIS 2017


CD45RA+
high
low
Single-Cell
CD38
0.42
6.27
0.00
1.57
3.14
4.70
6.27
Darkmode
Point Size
Point Opacity
Halo Size

VIS 2017


CD45RA+
high
low
Single-Cell
CD38
0.42
6.27
0.00
1.57
3.14
4.70
Darkmode
Point Size
Point Opacity
Halo Size

CyteGuide

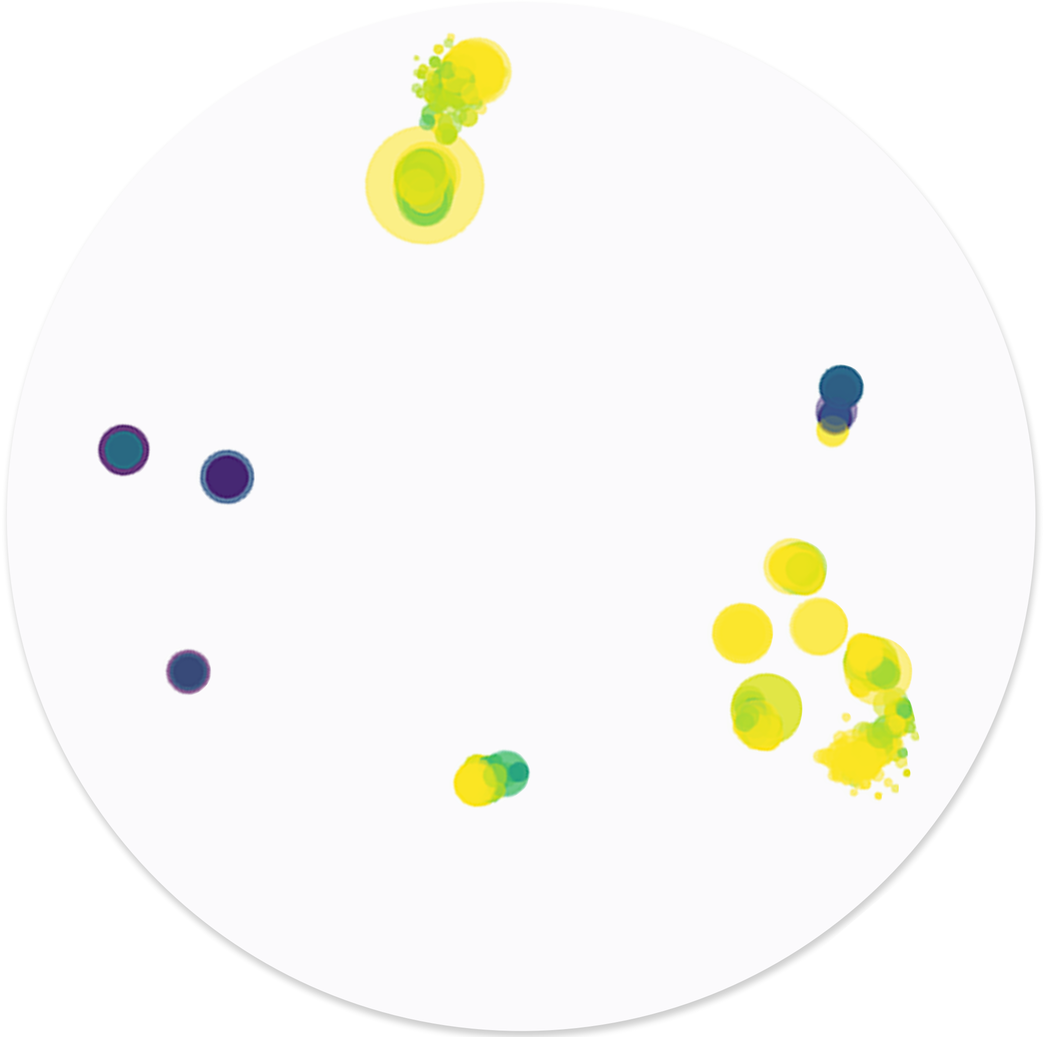
Mass Cytometry

- 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

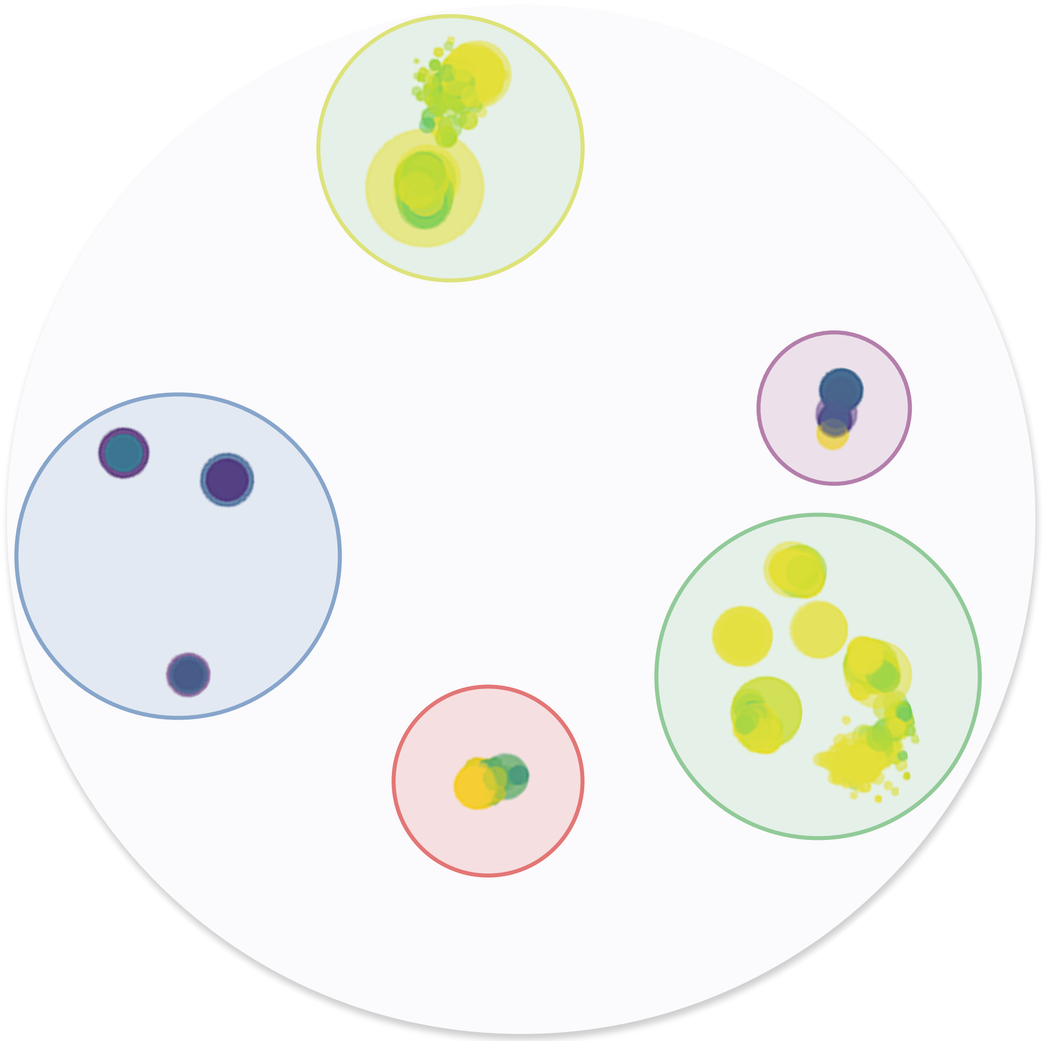
Data Abstraction

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



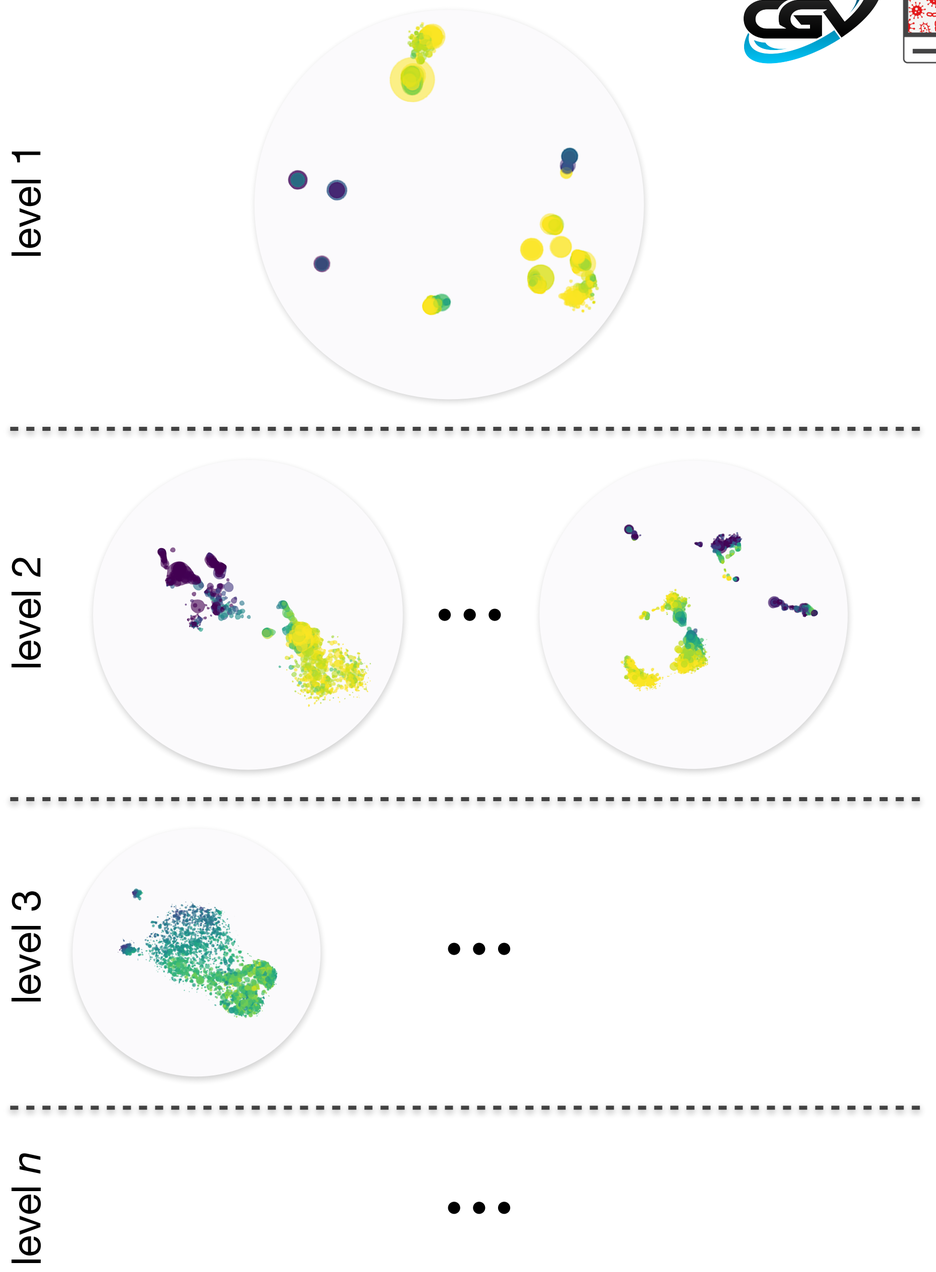
Data Abstraction

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



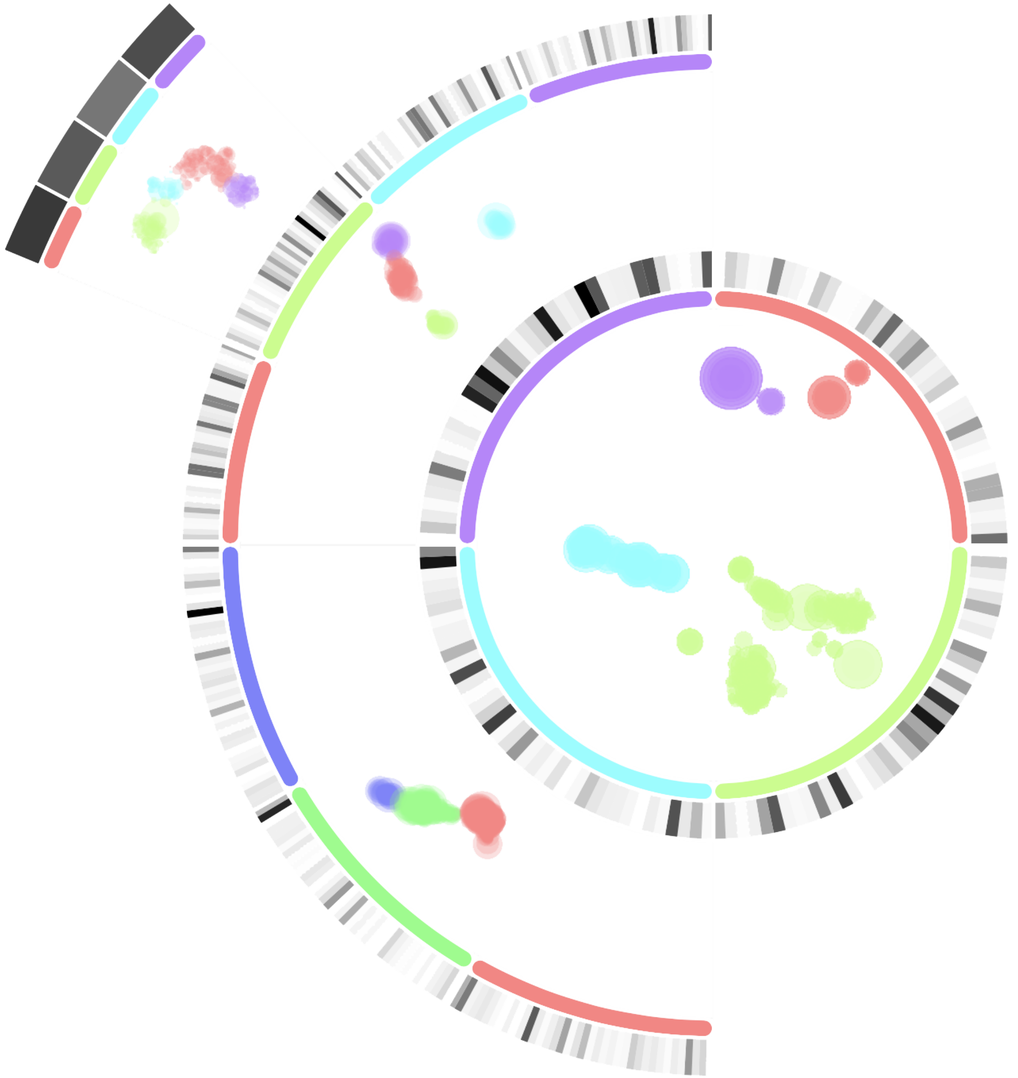
Data Abstraction

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



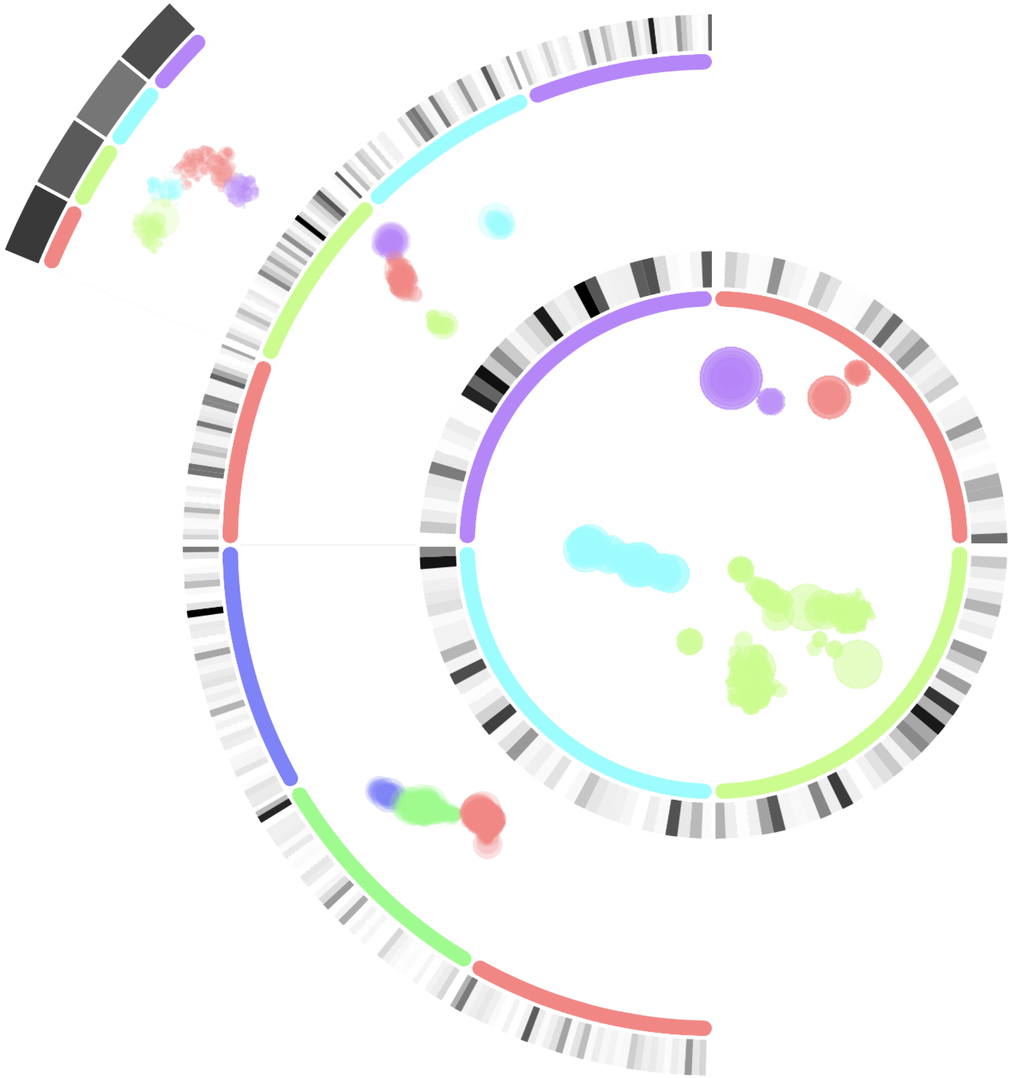
Goal

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



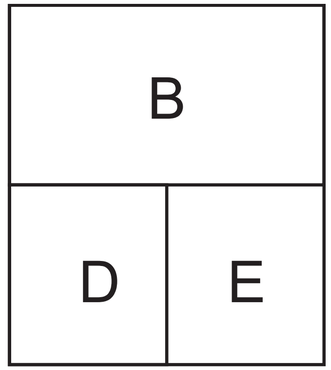
Requirements

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

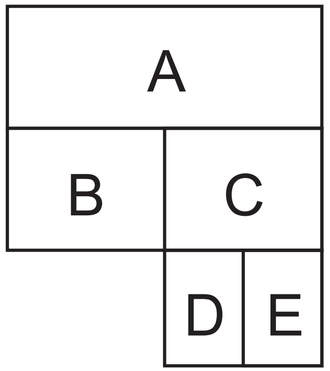


Solutions

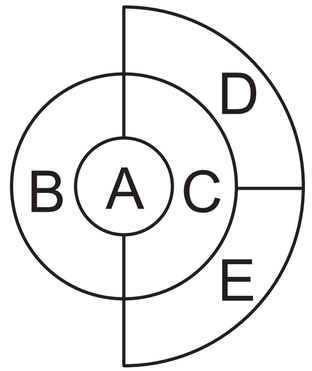
- Hierarchy
  - Up to 5 levels
  - Up to 5 branches per node
  - Embedded vis in nodes



Treemap



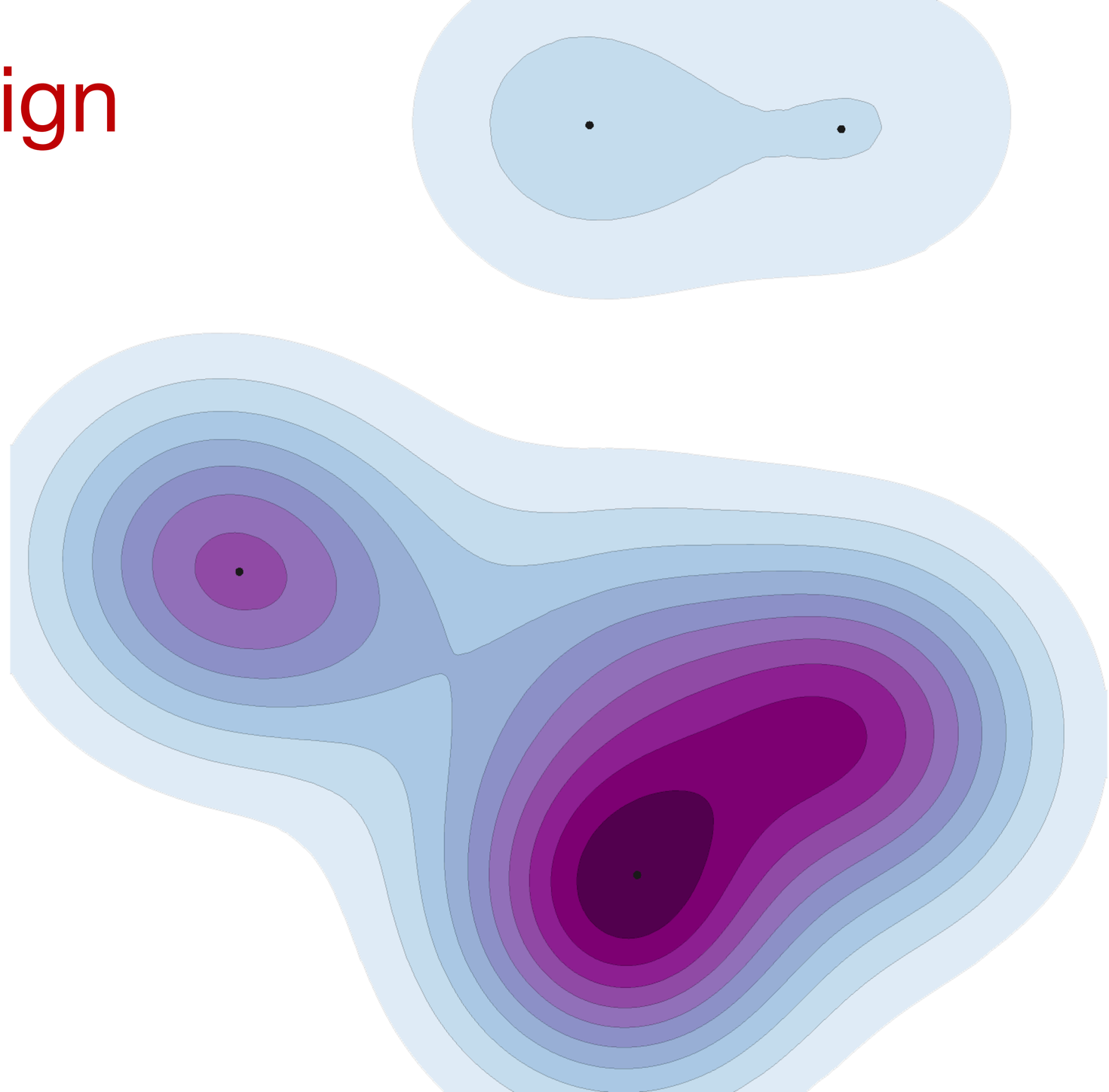
Icicles



Sunburst

Design

Design



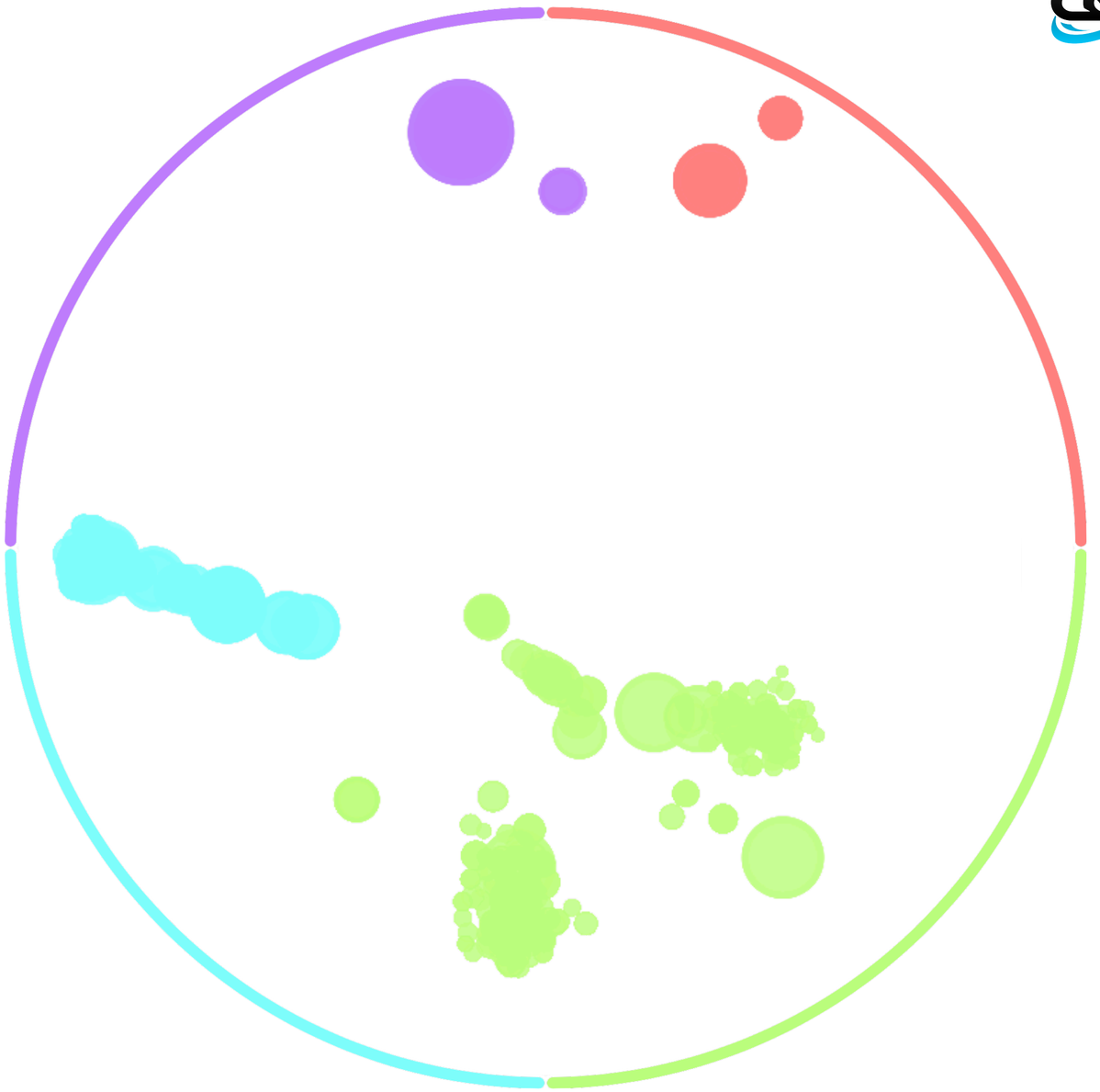
Design

clusters

Design

clusters

arc segment

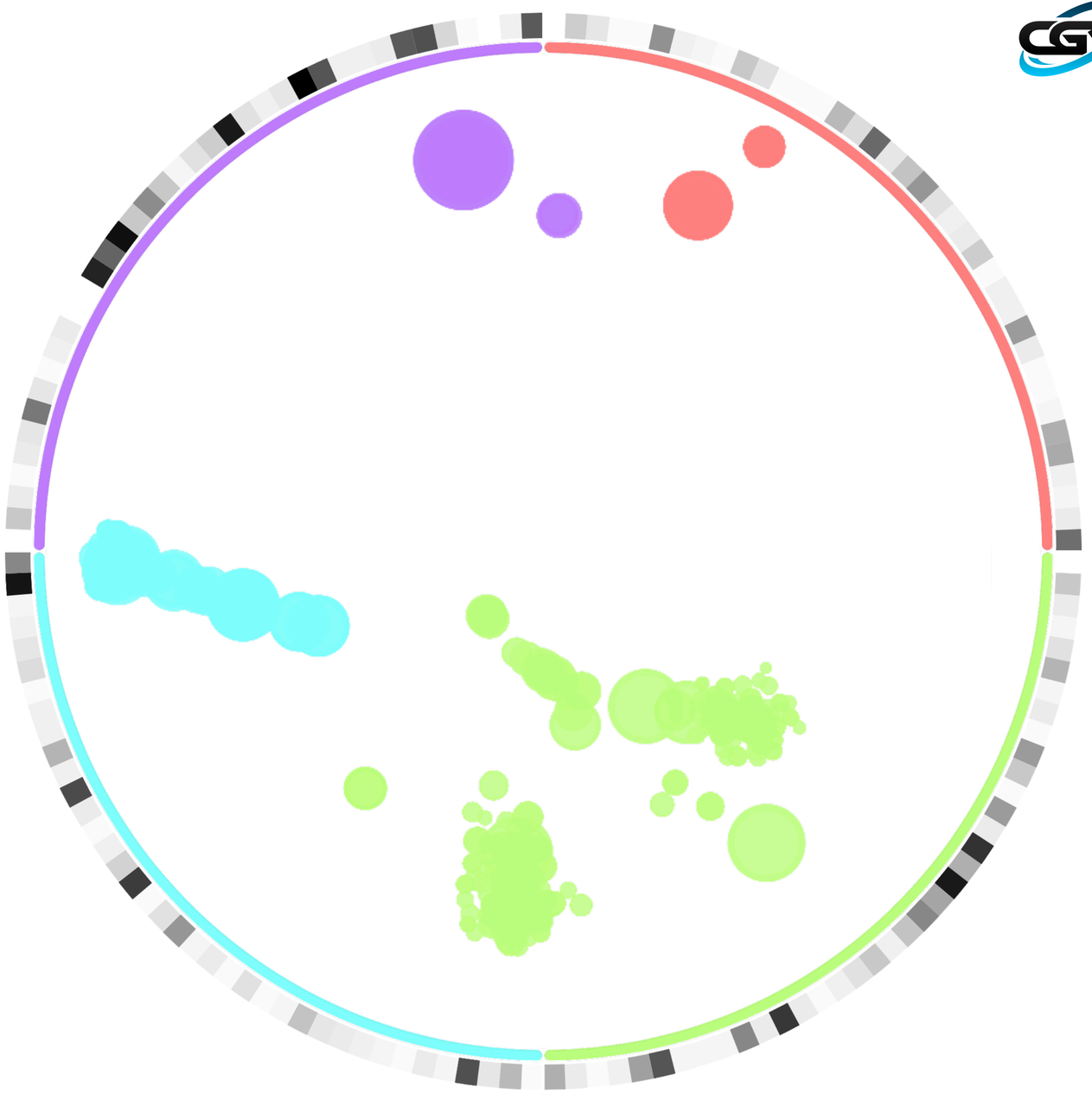


Design

clusters

arc segment

marker heatmap



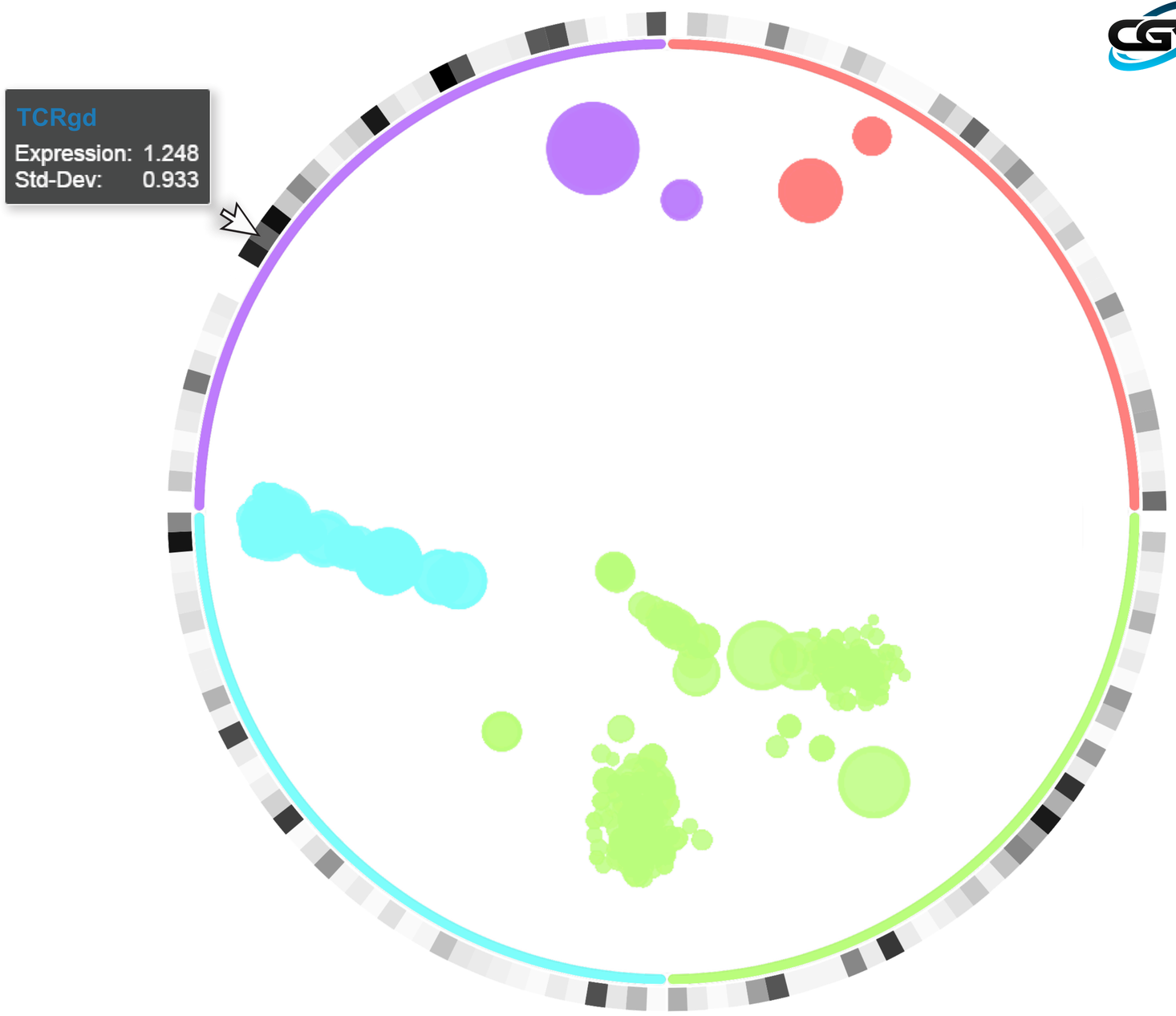
std-dev

Design

clusters

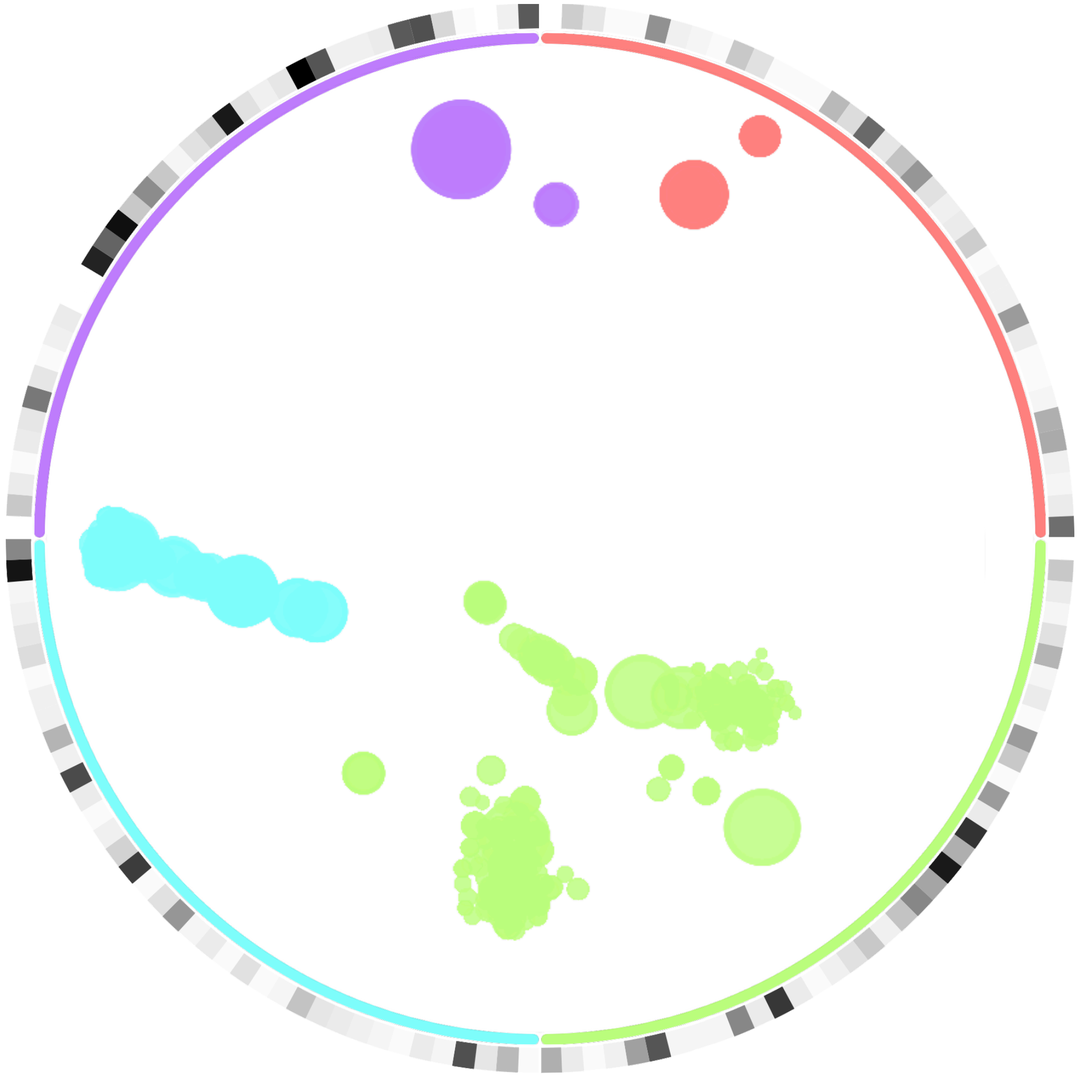
arc segment

marker heatmap

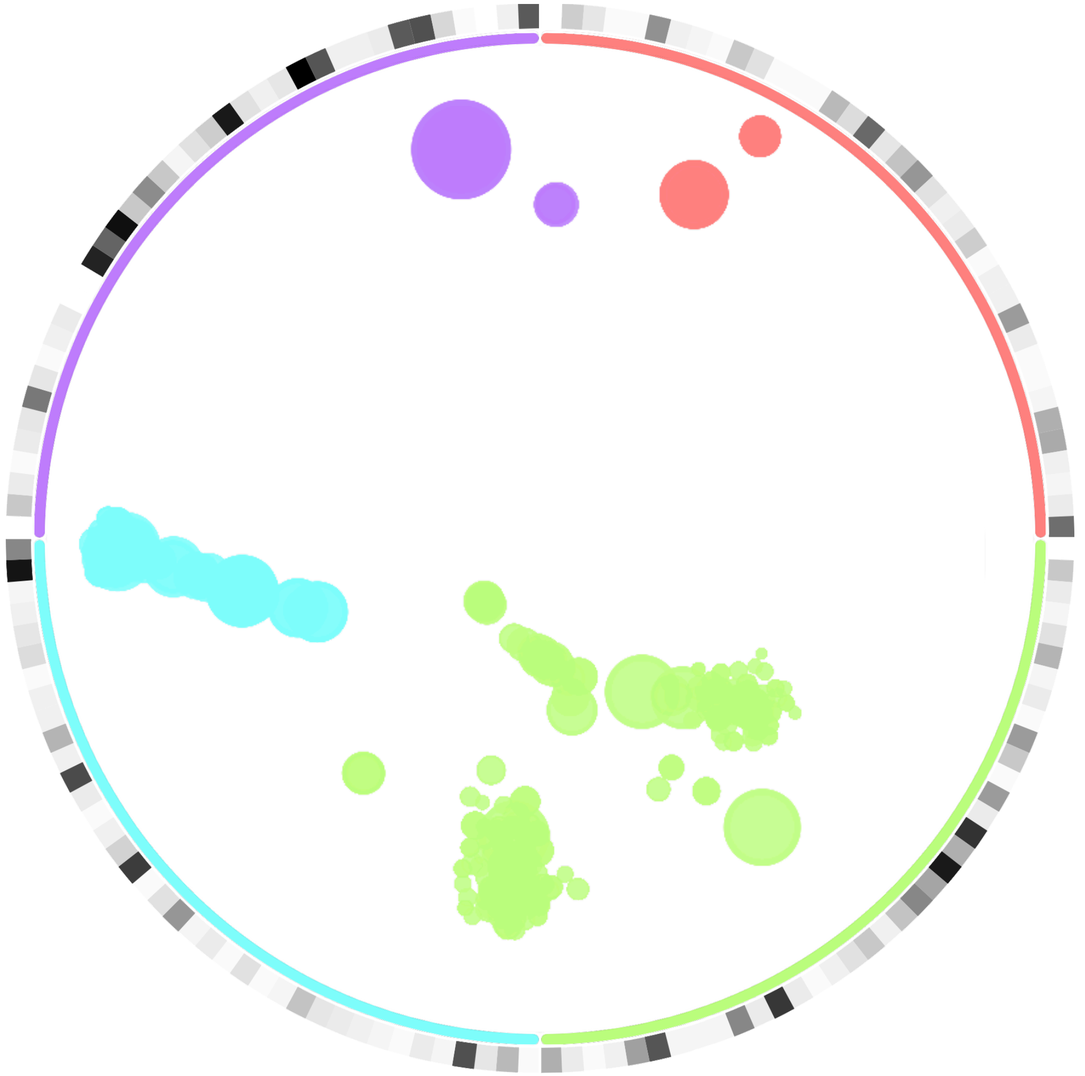


std-dev

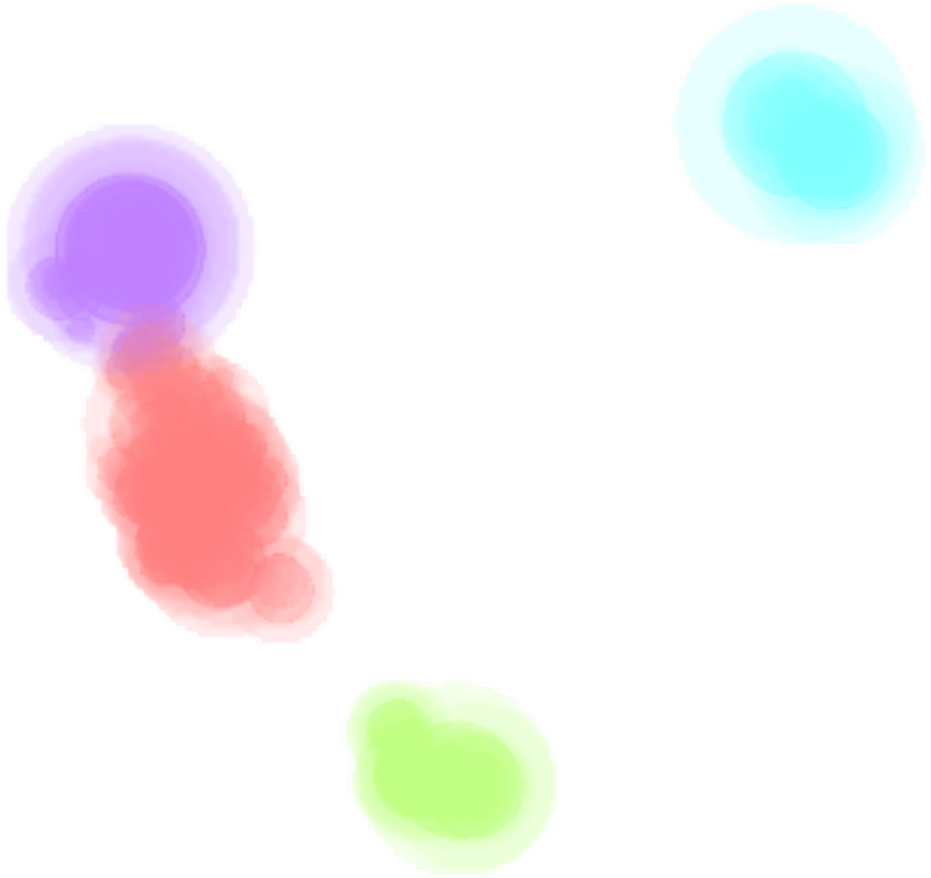
Design

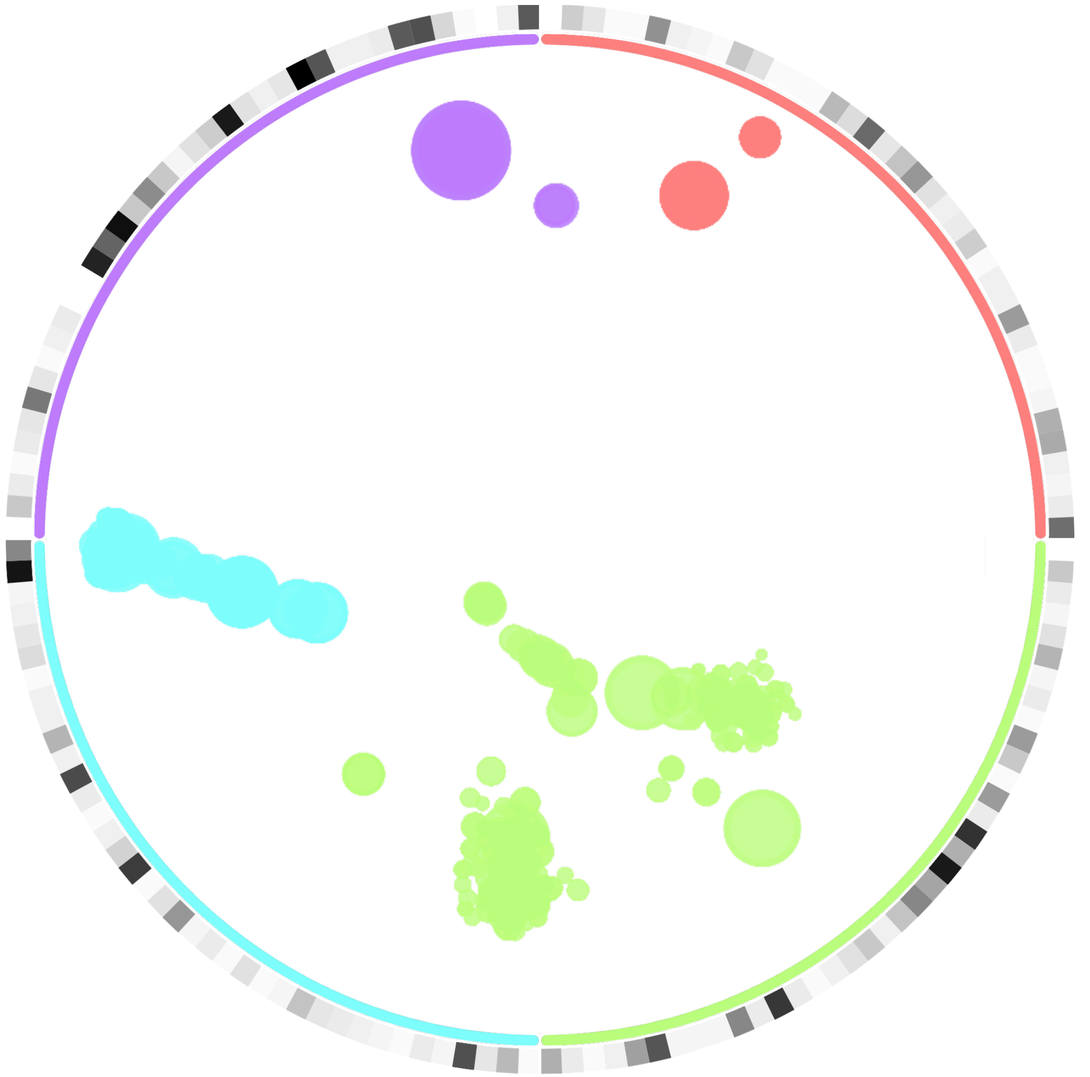


Design



Design



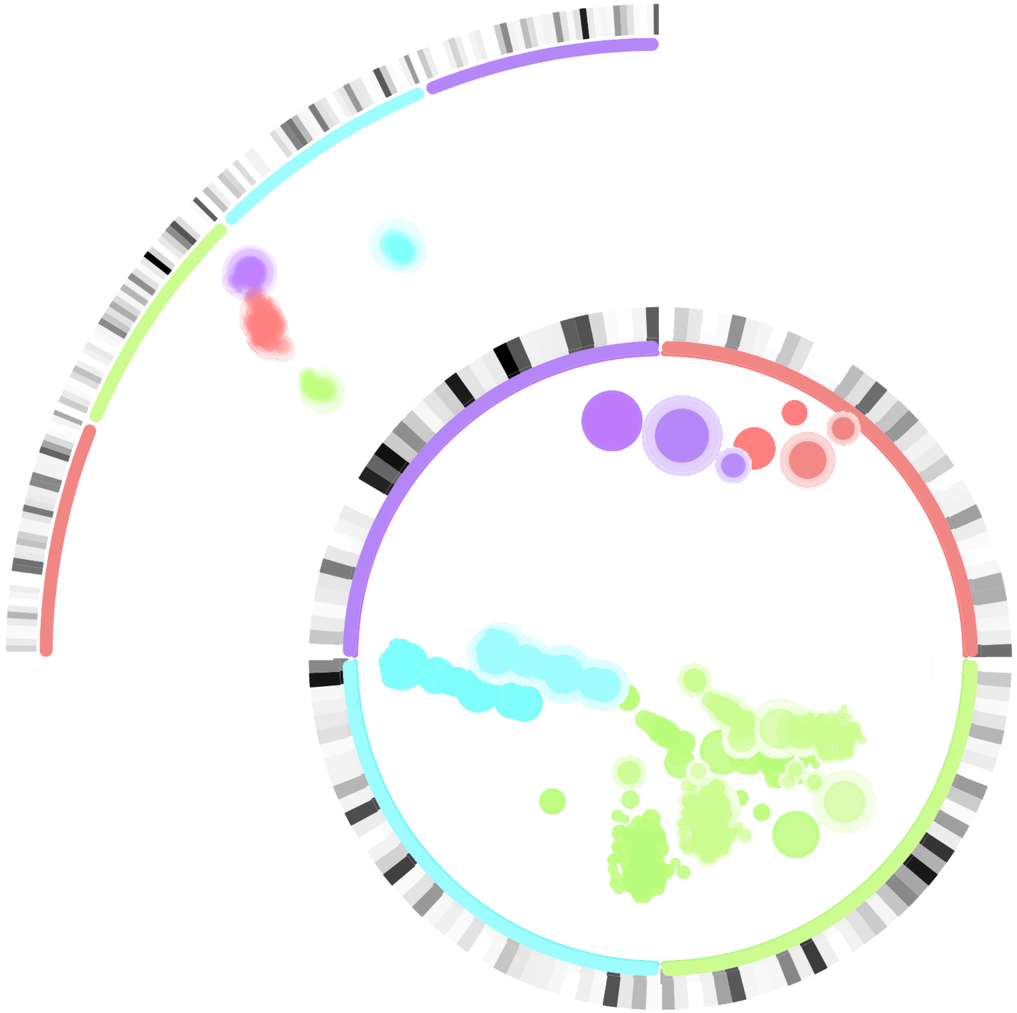


Design

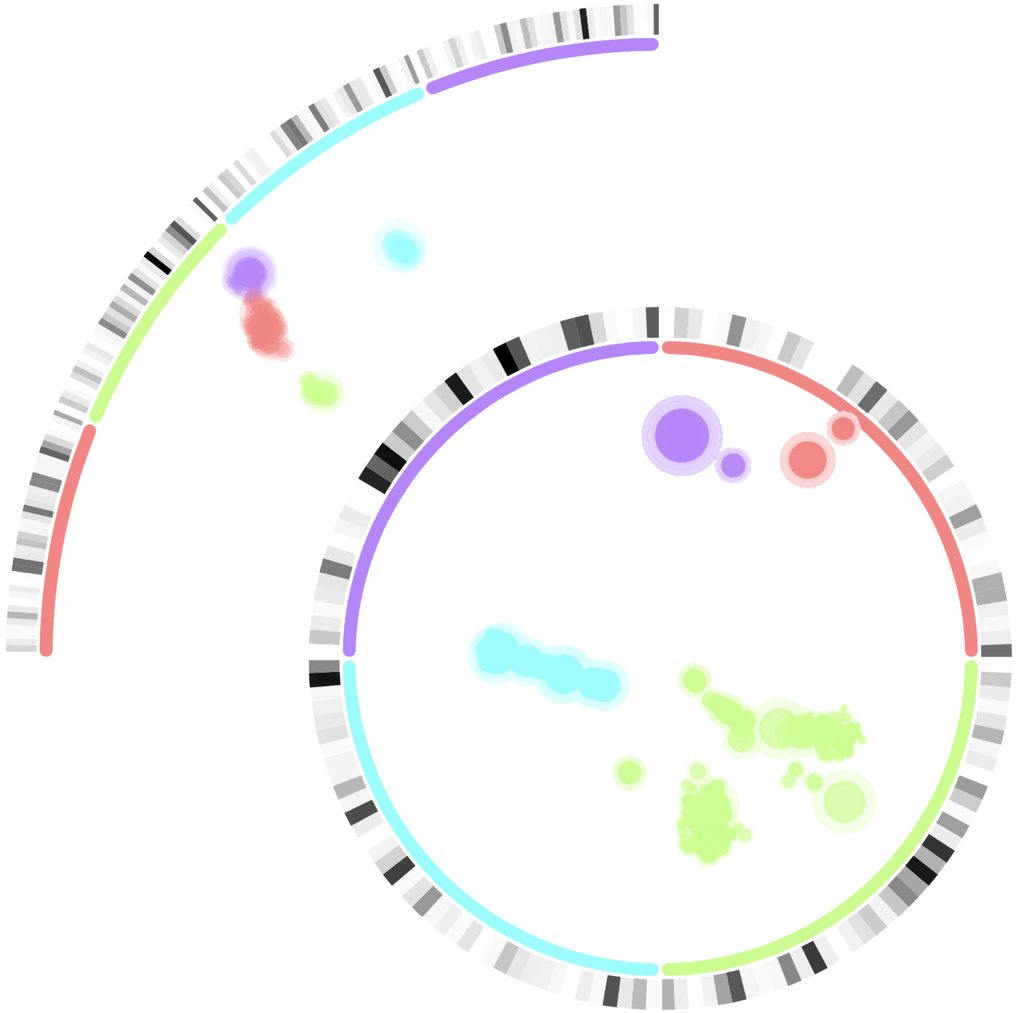
clusters

arc segment

marker heatmap



Design



clusters

arc segment

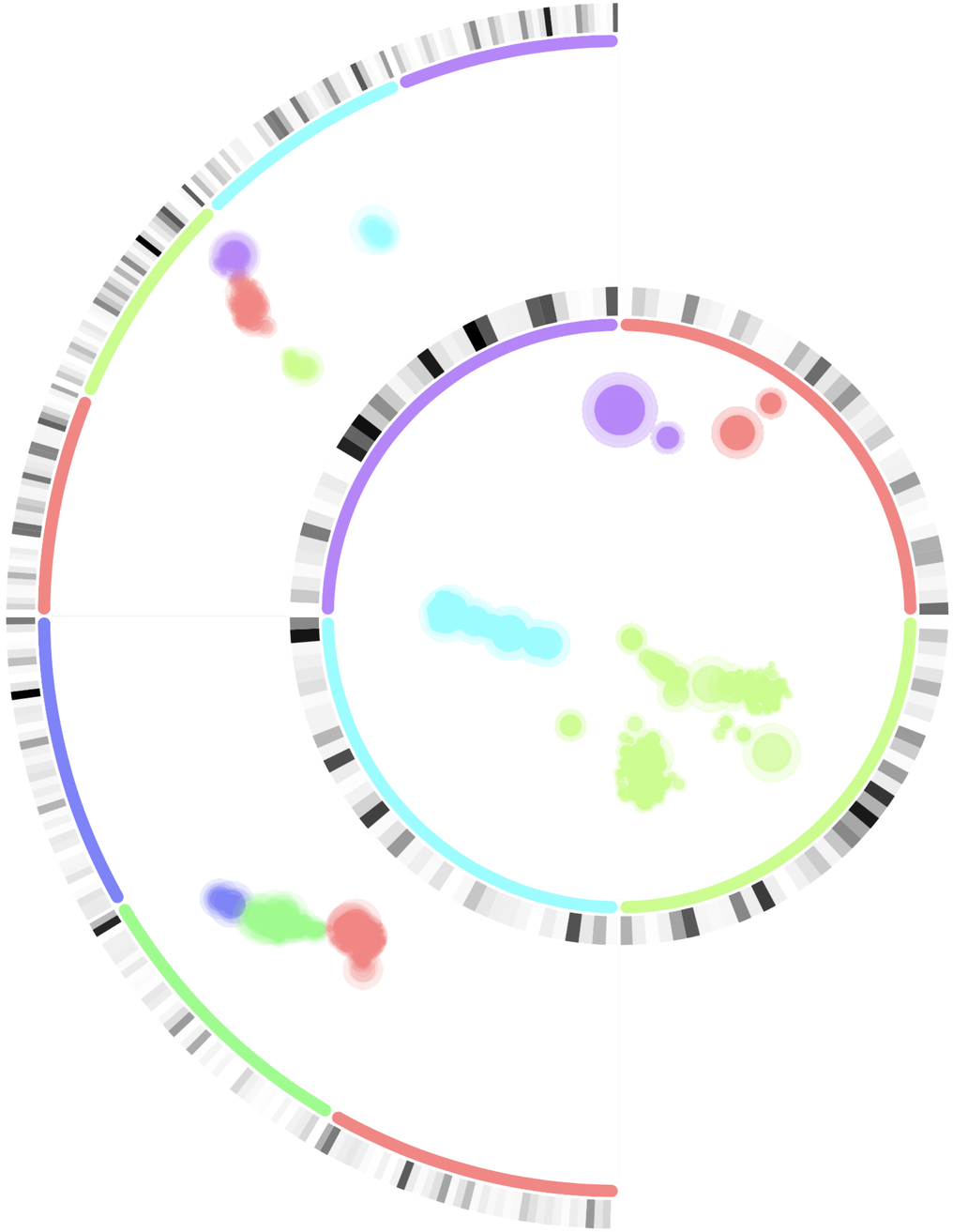
marker heatmap

Design

clusters

arc segment

marker heatmap

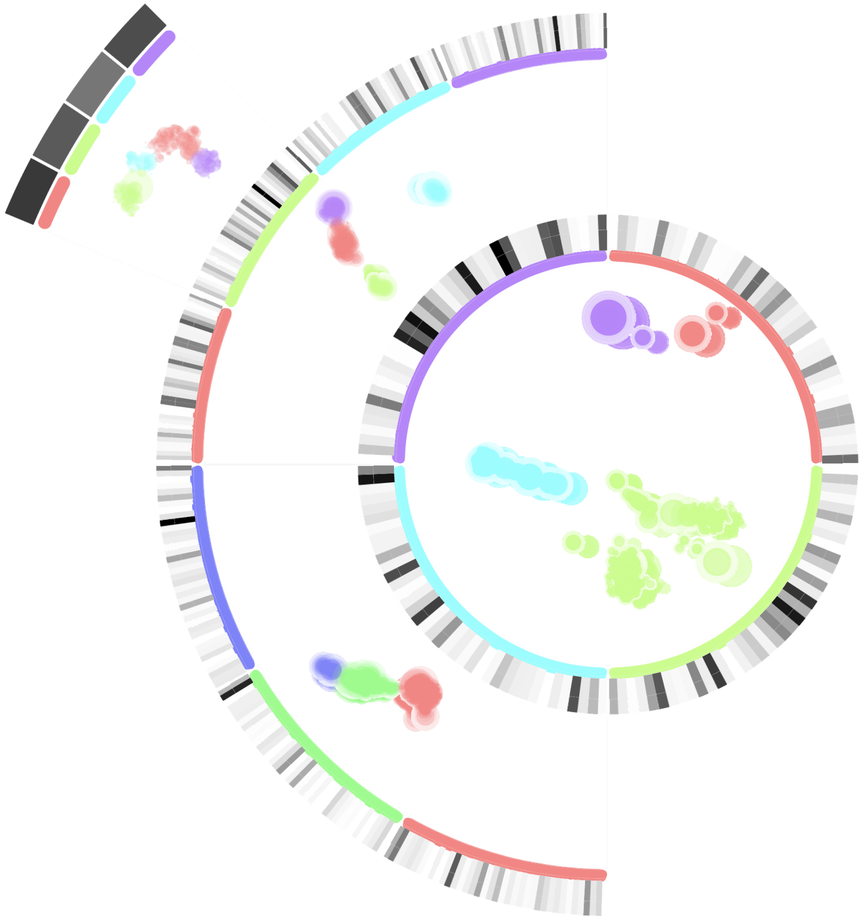


Design

clusters

arc segment

marker heatmap

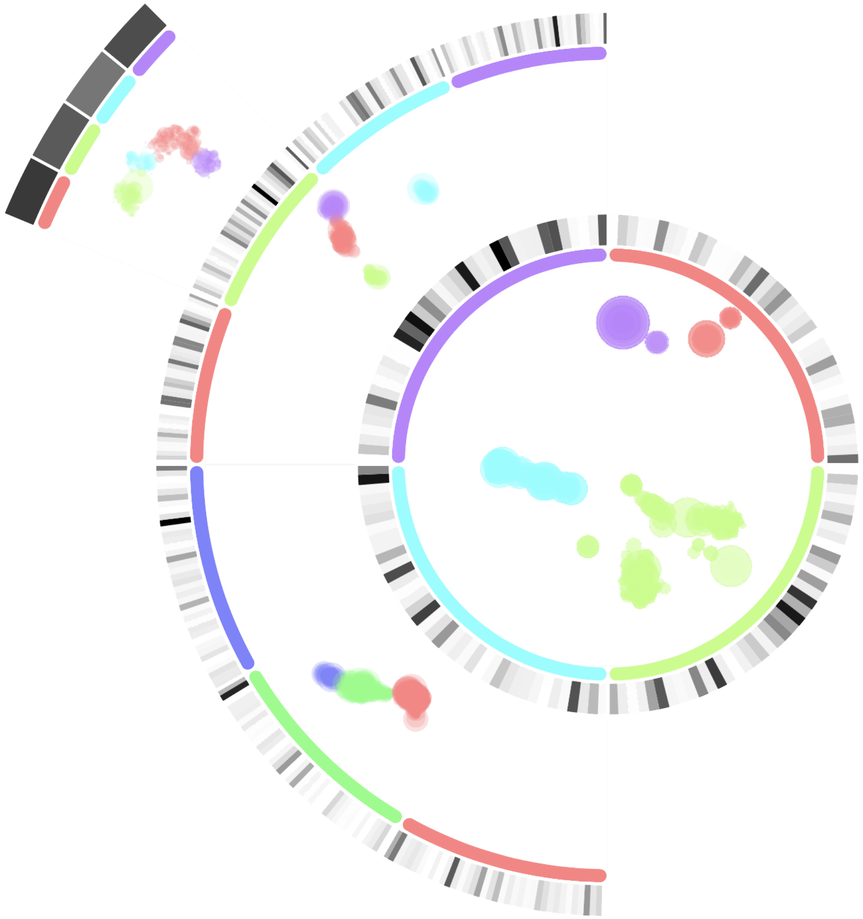


Design

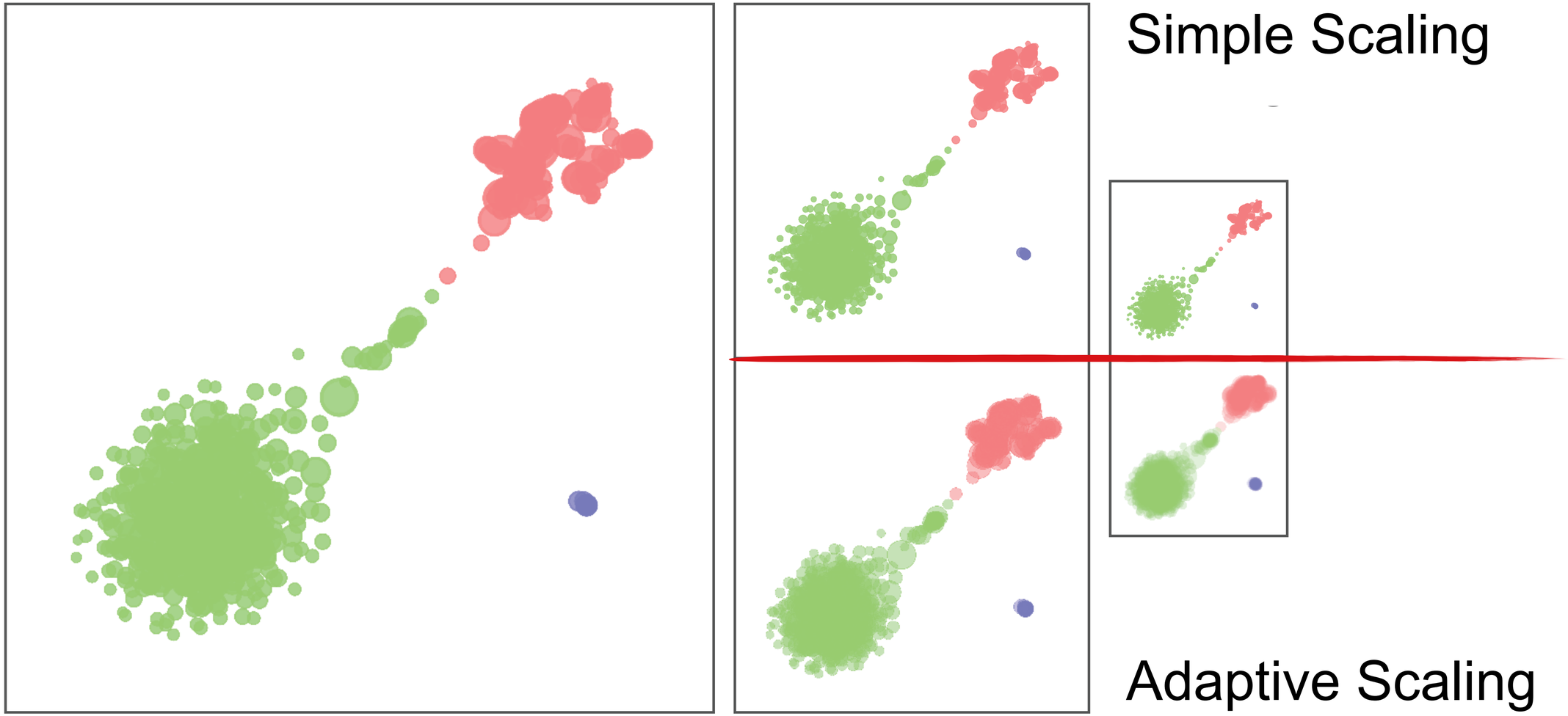
clusters

arc segment

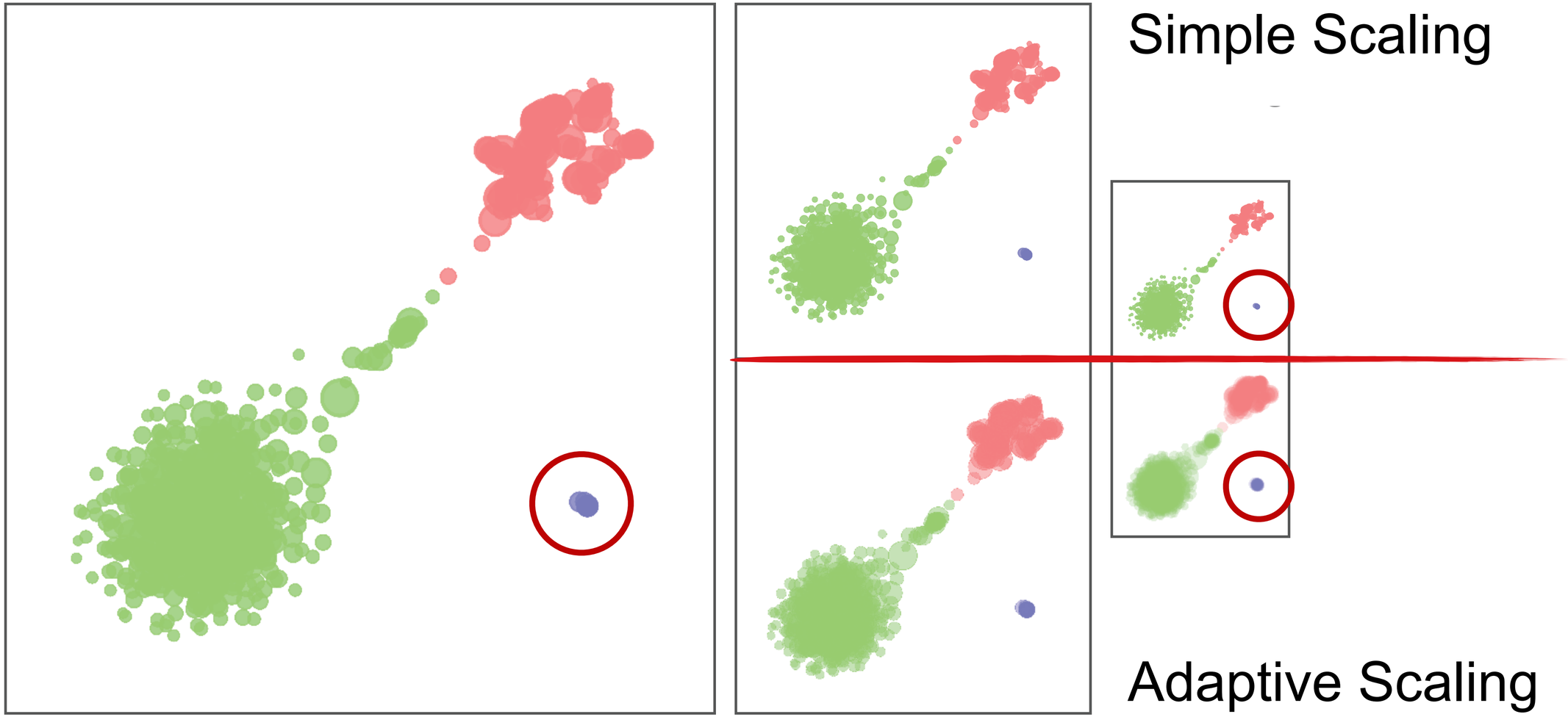
marker heatmap



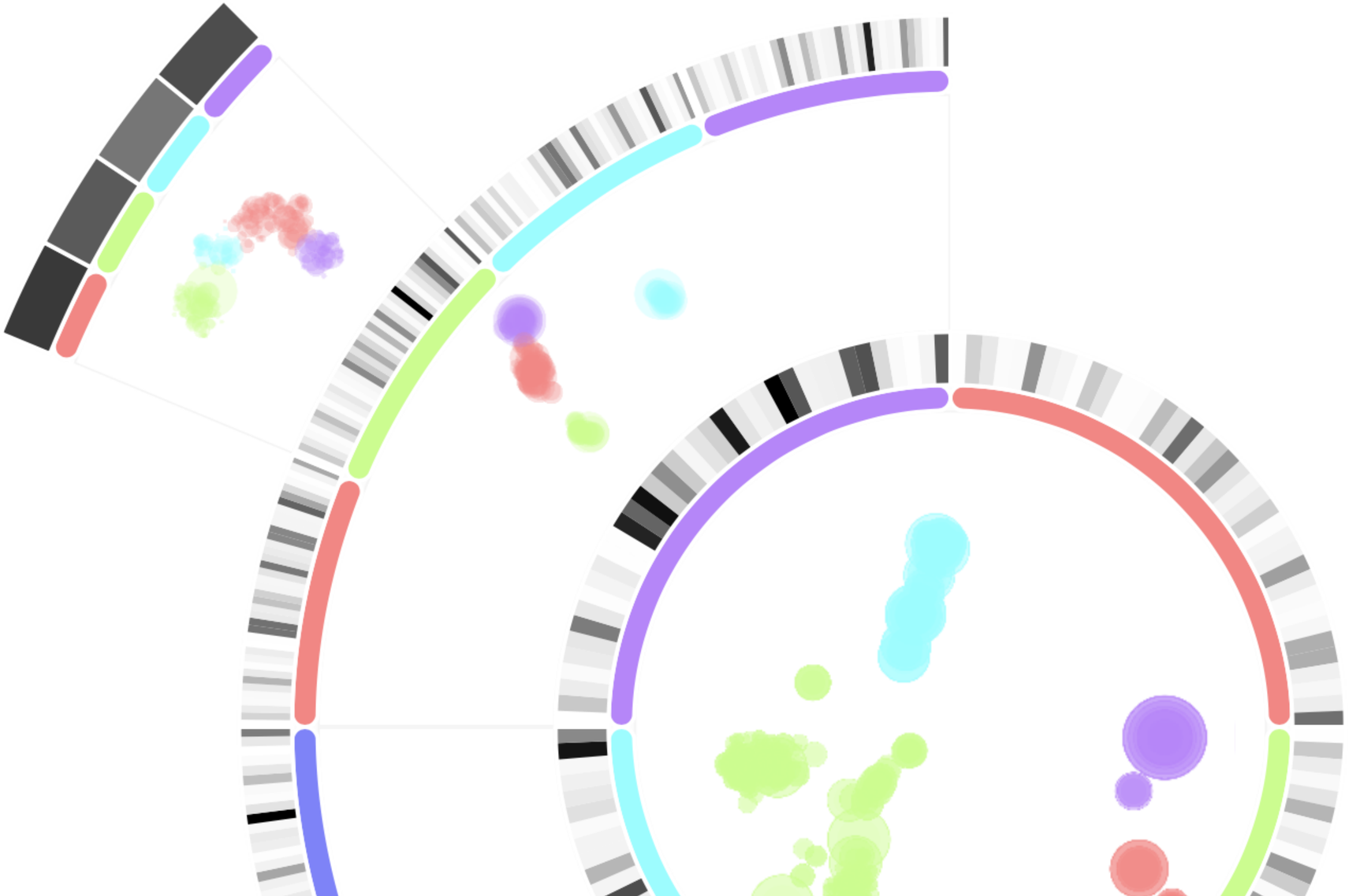
Adaptive Design - Scatterplots



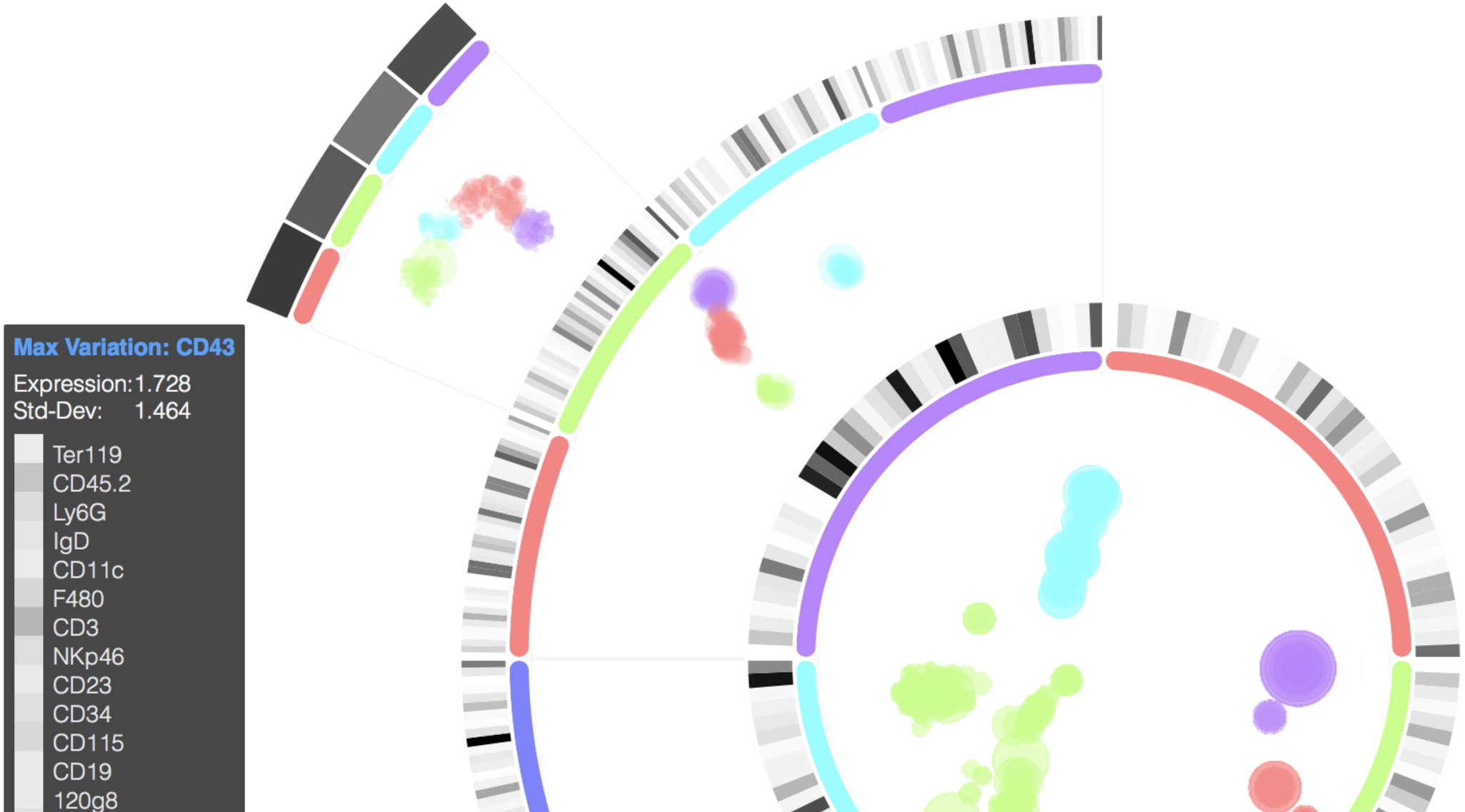
Adaptive Design - Scatterplots



Adaptive Design - Heatmaps



Adaptive Design - Heatmaps



Cytosplore

Settings

HSNE Analysis

Input Data

Active Markers

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

Inverse Markers

Load Selection | Set Default

Advanced Parameters

Parameters

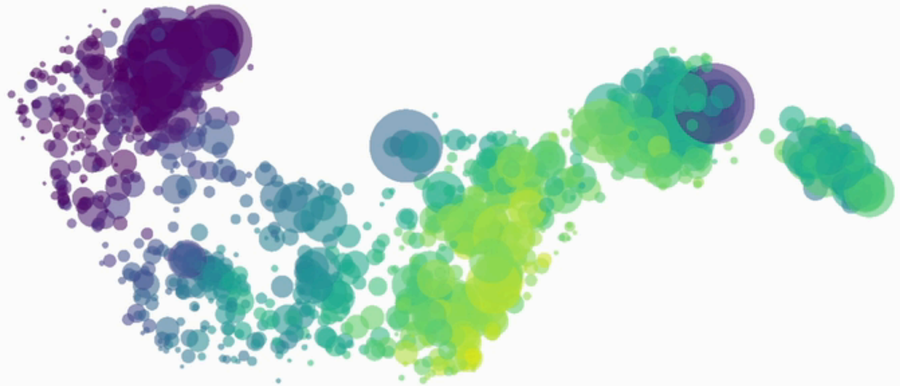
Number of Scales 4

Compute HSNE

CD4+

high

low



Single-Cell | Sca1 | -0.18 | 7.03

-0.18 1.62 3.43 5.23 7.03

Darkmode

Point Size | Point Opacity | Halo Size

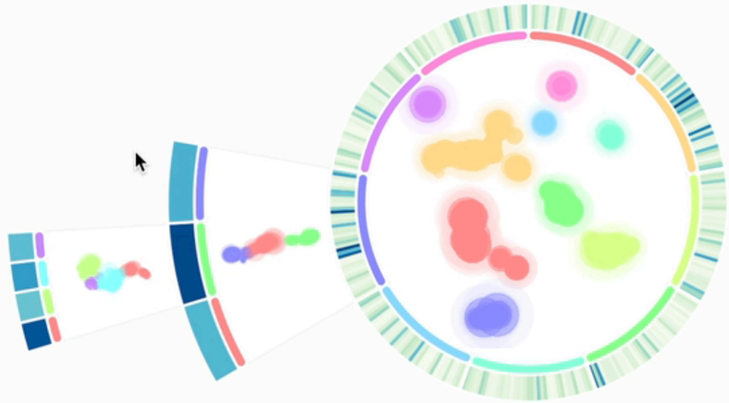
HSNE Analysis - Overview Level | T-Cells | CD4+

Hierarchy View

HSNE Analysis - Overview Level

1.92

0.12



Cytosplore

Settings

HSNE Analysis

Input Data

Active Markers

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

Inverse Markers

Load Selection | Set Default

Advanced Parameters

Parameters

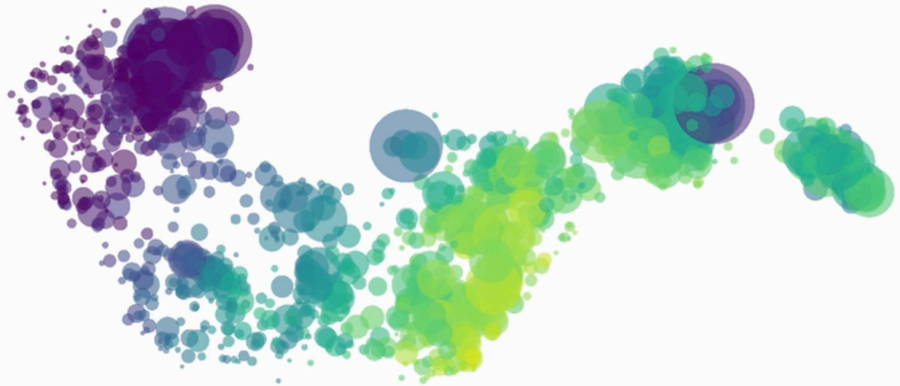
Number of Scales 4

Compute HSNE

CD4+

high

low



Single-Cell | Sca1 | -0.18 | 7.03

-0.18 1.62 3.43 5.23 7.03

Darkmode

Point Size | Point Opacity | Halo Size

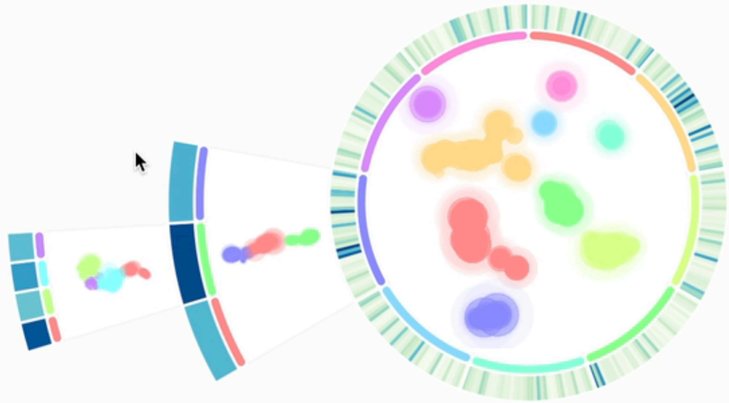
HSNE Analysis - Overview Level | T-Cells | CD4+

Hierarchy View

HSNE Analysis - Overview Level

1.92

0.12



Cytosplore

Settings

HSNE Analysis

Input Data

Active Markers

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

Inverse Markers

Load Selection | Set Default

Advanced Parameters

Parameters

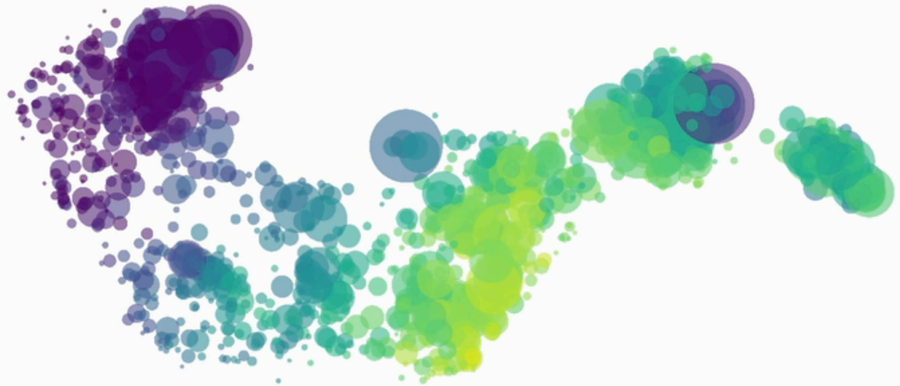
Number of Scales 4

Compute HSNE

CD4+

high

low



Single-Cell | Sca1 | -0.18 | 7.03

-0.18 1.62 3.43 5.23 7.03

Darkmode

Point Size | Point Opacity | Halo Size

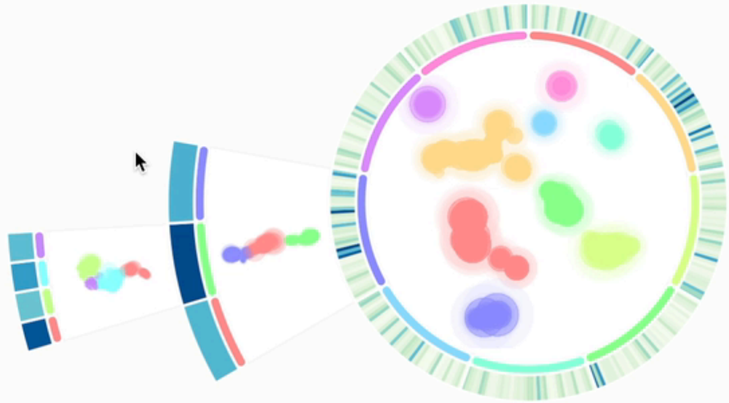
HSNE Analysis - Overview Level | T-Cells | CD4+

Hierarchy View

HSNE Analysis - Overview Level

1.92

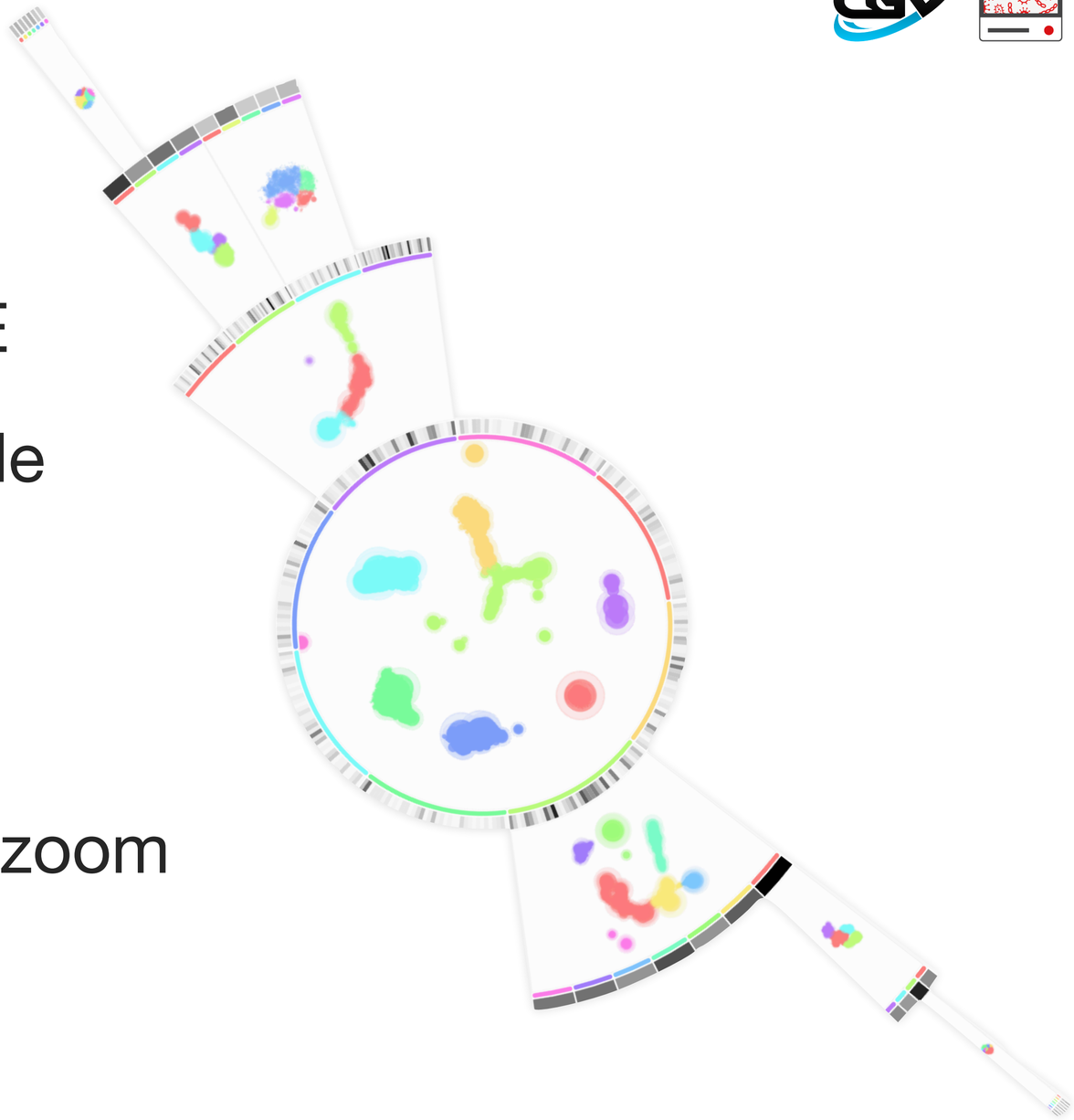
0.12



Evaluation

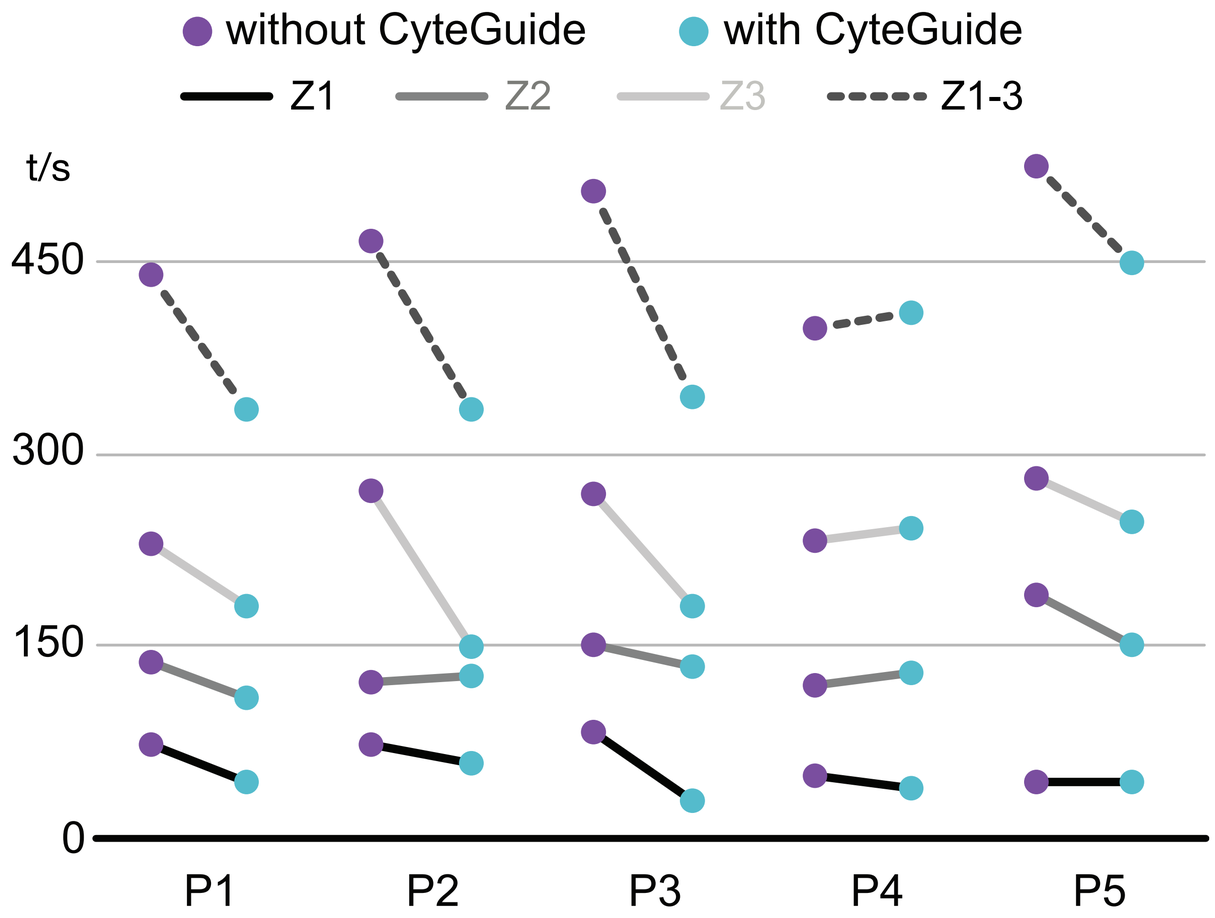
Evaluation

- 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



Performance

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



Performance

- 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.”

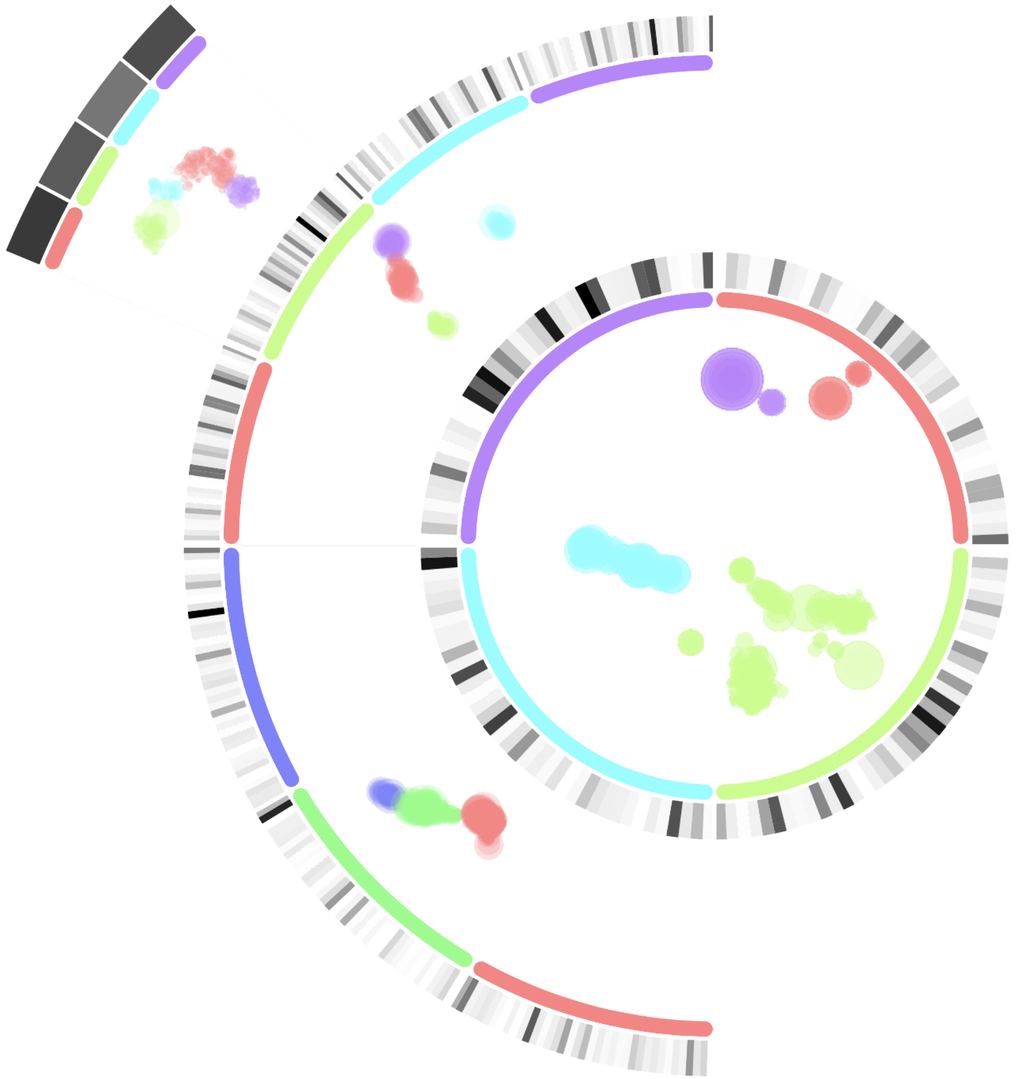
Conclusion

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

VIS 2017