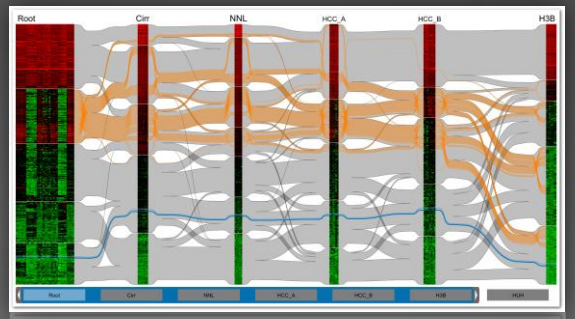
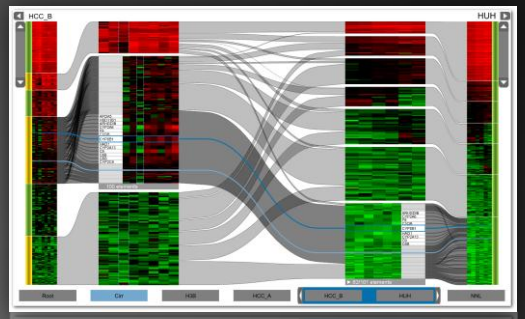




Comparative Analysis of Multidimensional, Quantitative Data

The Caleydo Matchmaker

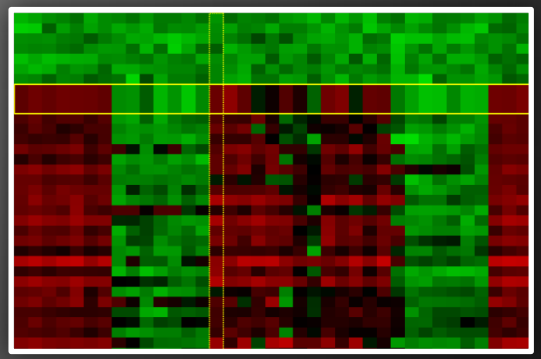
Alexander Lex
lex@icg.tugraz.at
Marc Streit, Christian Partl, Karl Kashofer, Dieter Schmalstieg
Graz University of Technology, Austria

Graz University of Technology

InfoVis 10

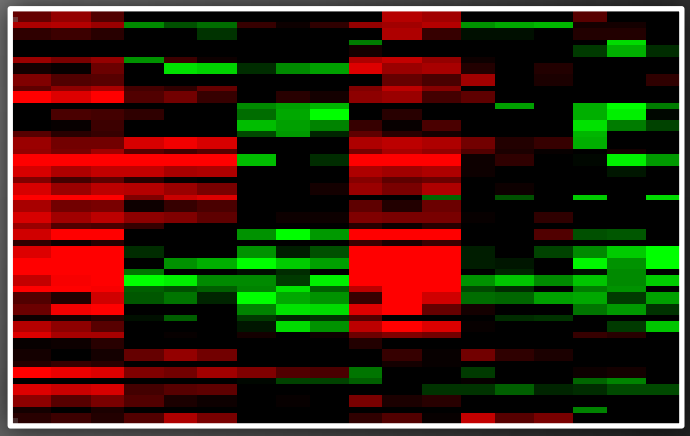
Problem Domain

- Multidimensional data often consists of homogeneous groups of dimensions
  - Examples from biology:
    replicates, time-series, comparison of strains
- Common task for multidimensional data
  - Compare those groups
- Traditional approach:
  - Filter, cluster all,
    visualize with Heat Maps!



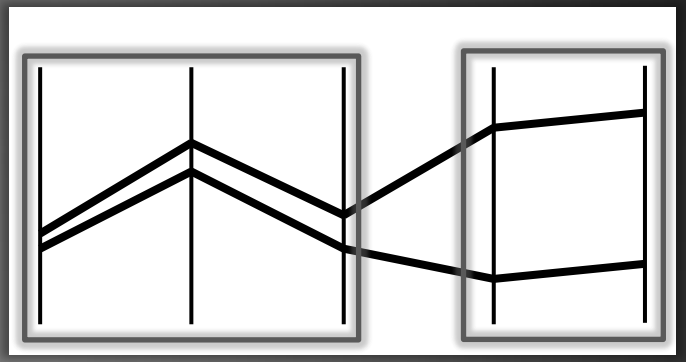
Clustering Inhomogeneous Data

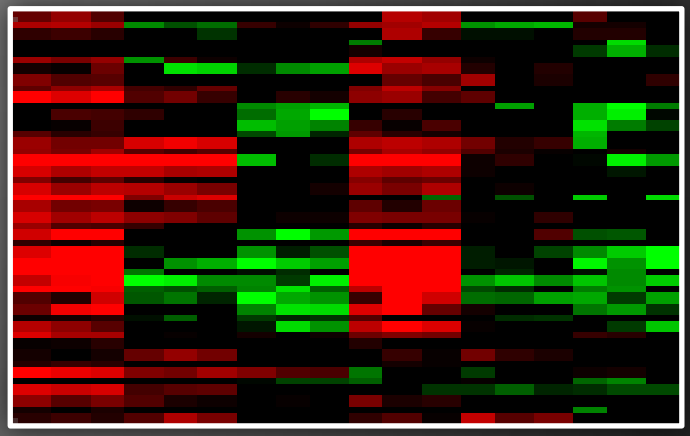
- The Problem
  - Clustering all dimensions at once obscures relations in homogeneous groups
- The Solution
  - Divide & conquer!



Clustering Inhomogeneous Data

- The Problem
  - Clustering all dimensions at once obscures relations in homogeneous groups
- The Solution
  - Divide & conquer!





Comparing Clustering Algorithms

- Choice of algorithm, parameters and distance measures are important
- No good quality metrics for clustering algorithms
  - Visual assessment is best solution

METHOD

InfoVis IO

The Process

1,1	1,2	1,3	1,4	1,5	1,6	1,7
2,1	2,2	2,3	2,4	2,5	2,6	2,7
3,1	3,2	3,3	3,4	3,5	3,6	3,7
4,1	4,2	4,3	4,4	4,5	4,6	4,7
5,1	5,2	5,3	5,4	5,5	5,6	5,7
6,1	6,2	6,3	6,4	6,5	6,6	6,7

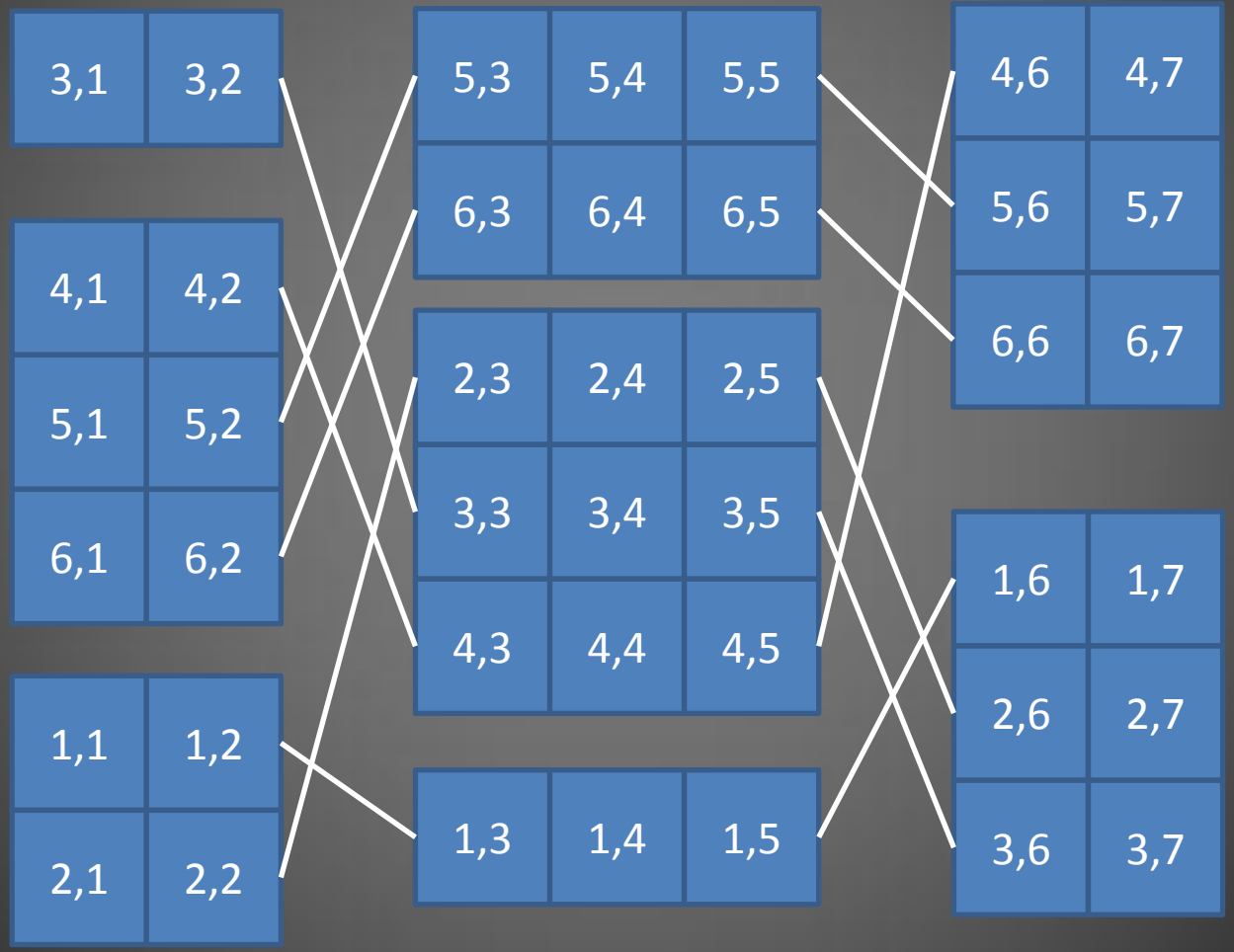
The Process

1,1	1,2
2,1	2,2
3,1	3,2
4,1	4,2
5,1	5,2
6,1	6,2

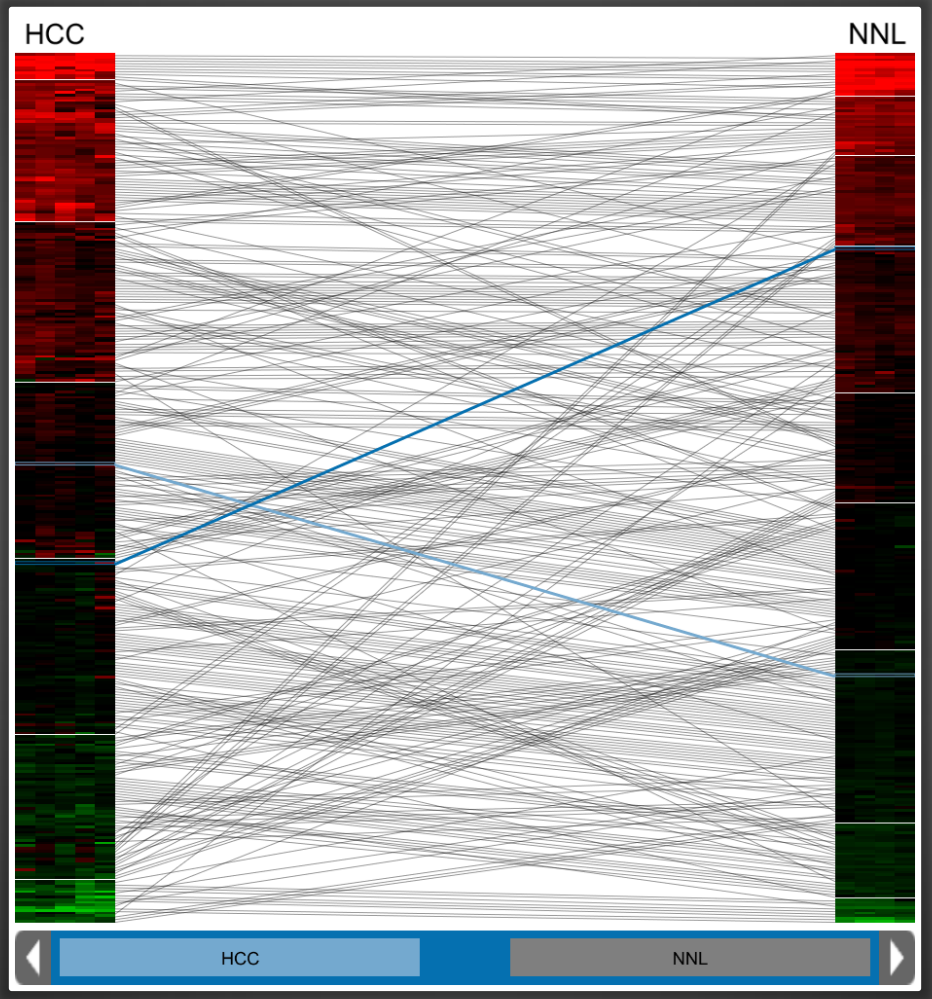
1,3	1,4	1,5
2,3	2,4	2,5
3,3	3,4	3,5
4,3	4,4	4,5
5,3	5,4	5,5
6,3	6,4	6,5

1,6	1,7
2,6	2,7
3,6	3,7
4,6	4,7
5,6	5,7
6,6	6,7

The Process


3,1
3,2
4,1
4,2
5,1
5,2
6,1
6,2
1,1
1,2
2,1
2,2
5,3
5,4
5,5
6,3
6,4
6,5
2,3
2,4
2,5
3,3
3,4
3,5
4,3
4,4
4,5
1,3
1,4
1,5
4,6
4,7
5,6
5,7
6,6
6,7
1,6
1,7
2,6
2,7
3,6
3,7

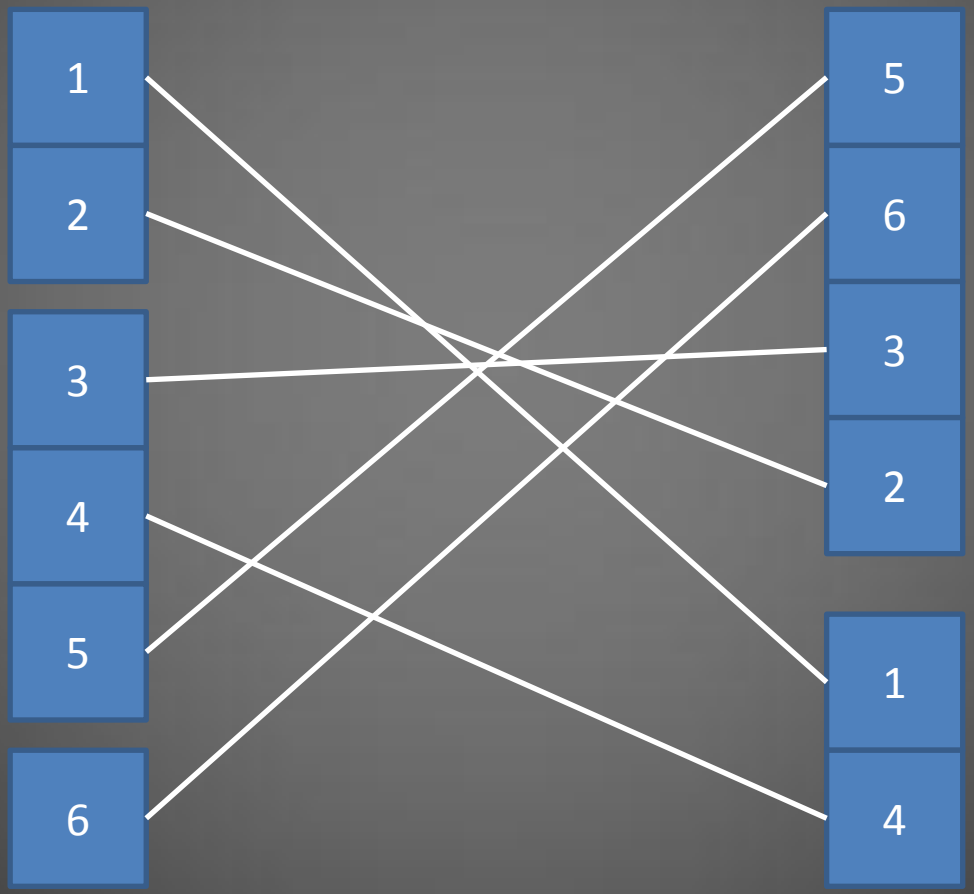
Immediate Result



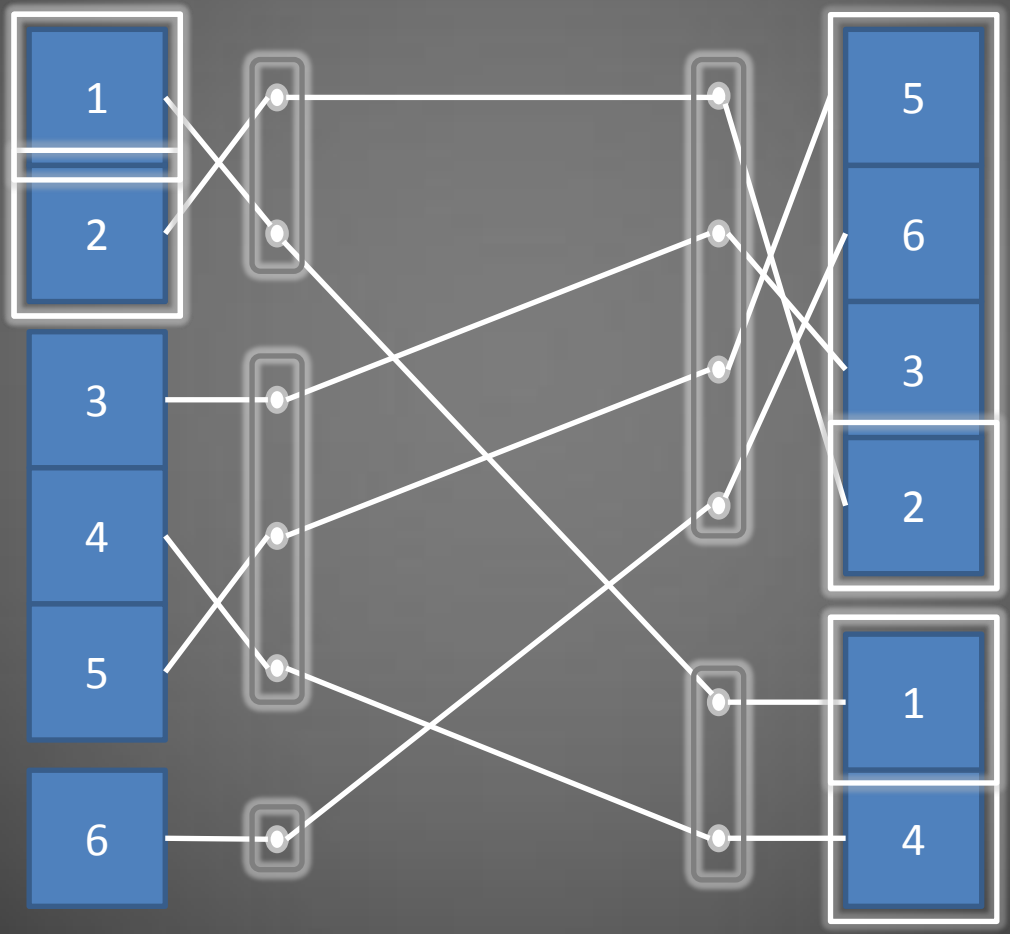
Alexander Lex

InfoVis 10

Connecting Records



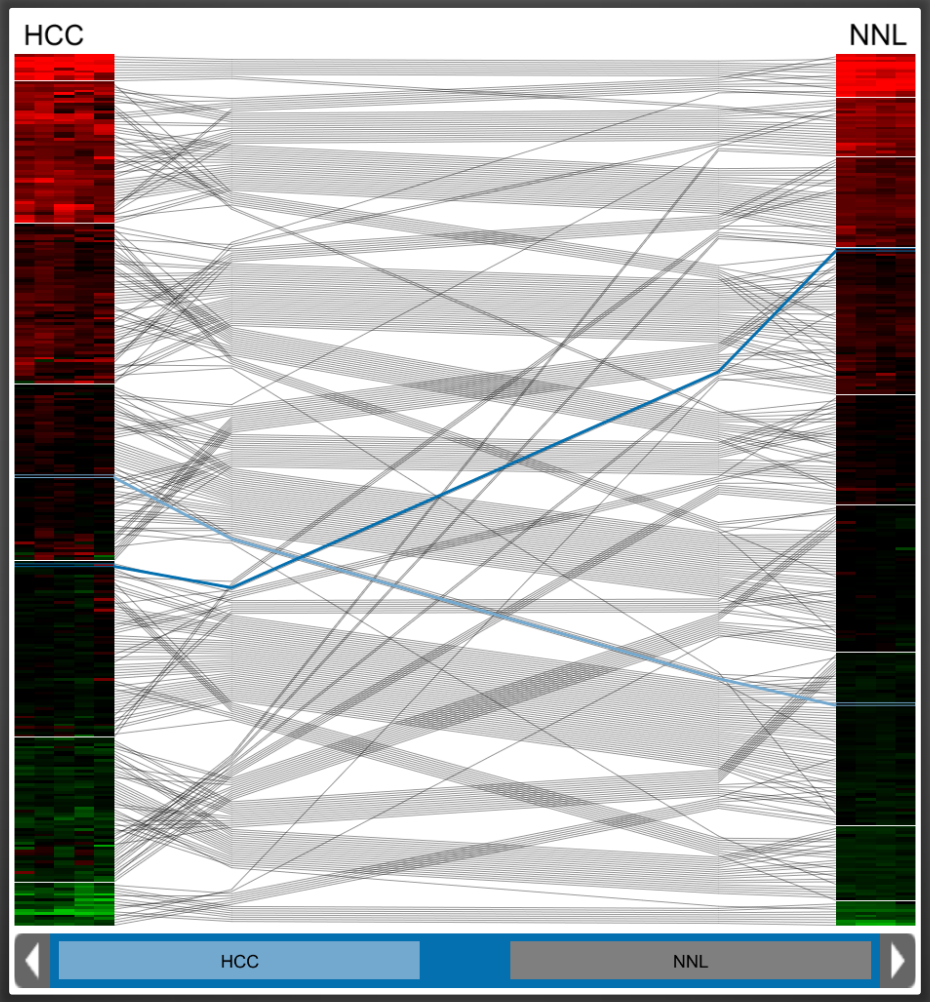
Cluster-driven Edge Bundling



Alexander Lex

InfoVis IO

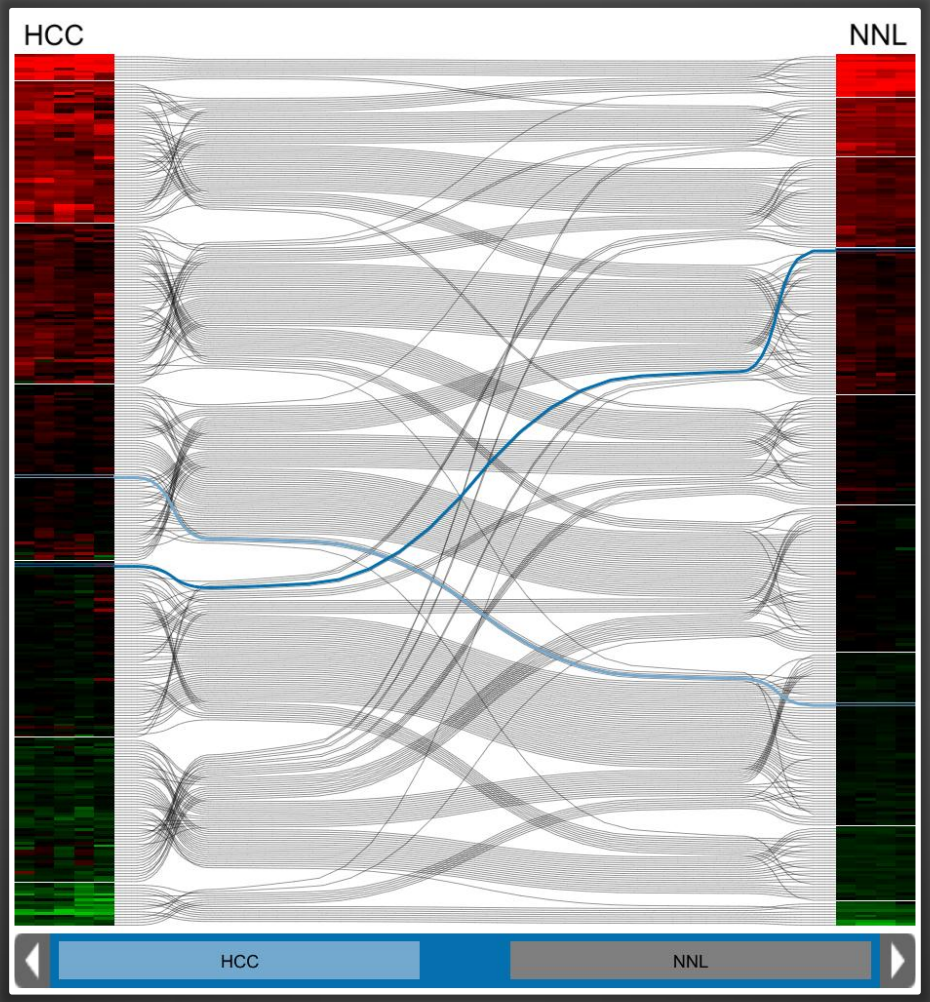
Using Bundling Points



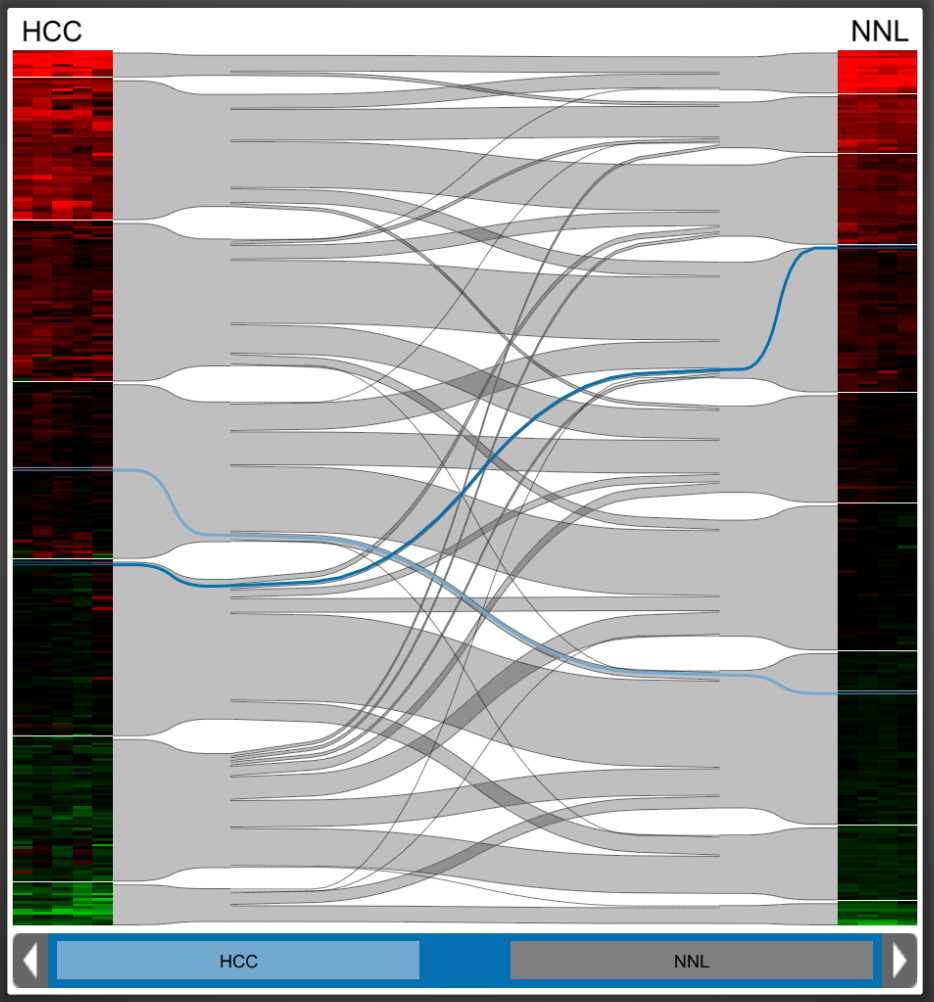
Alexander Lex

InfoVis IO

Using Splines



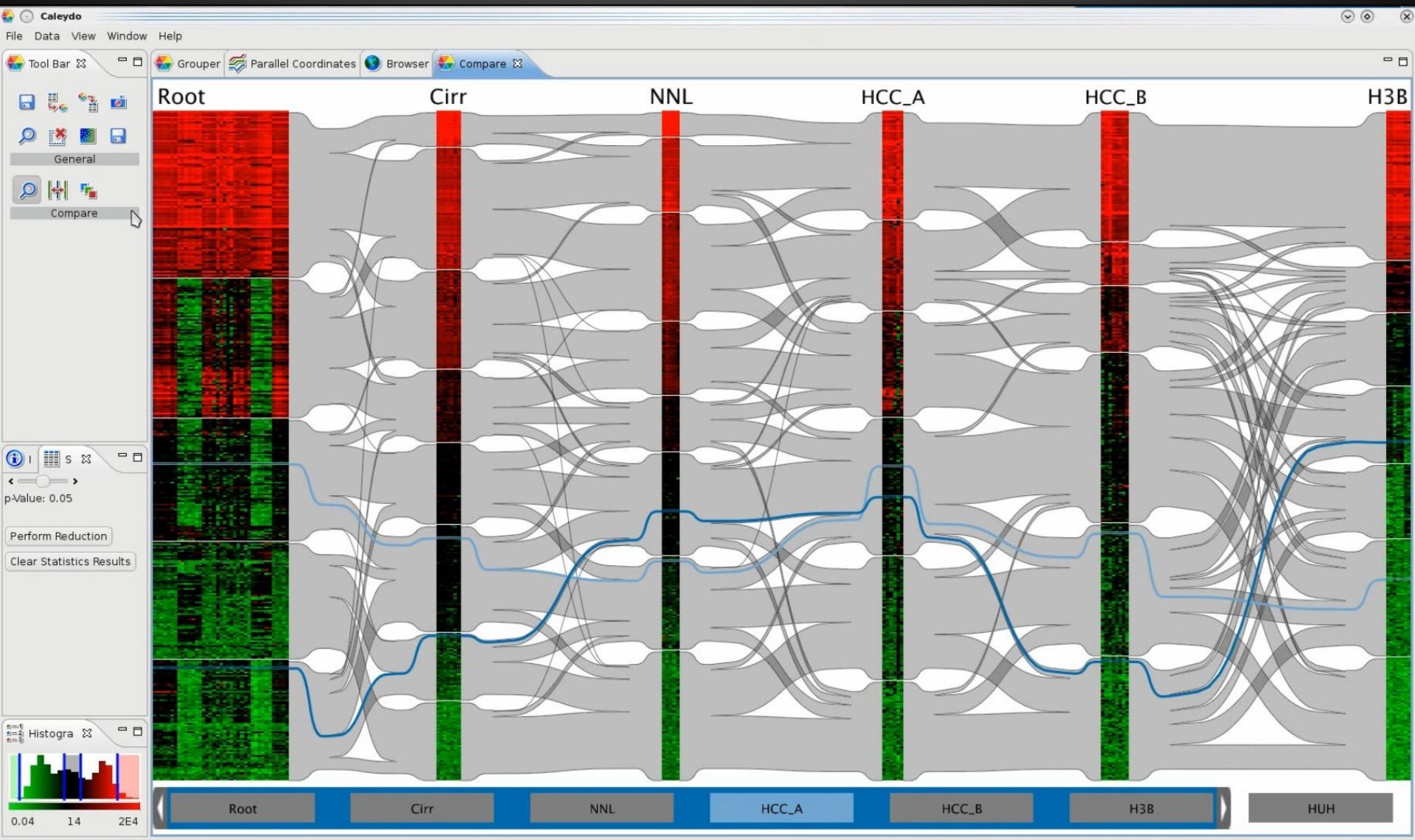
Using Pipes



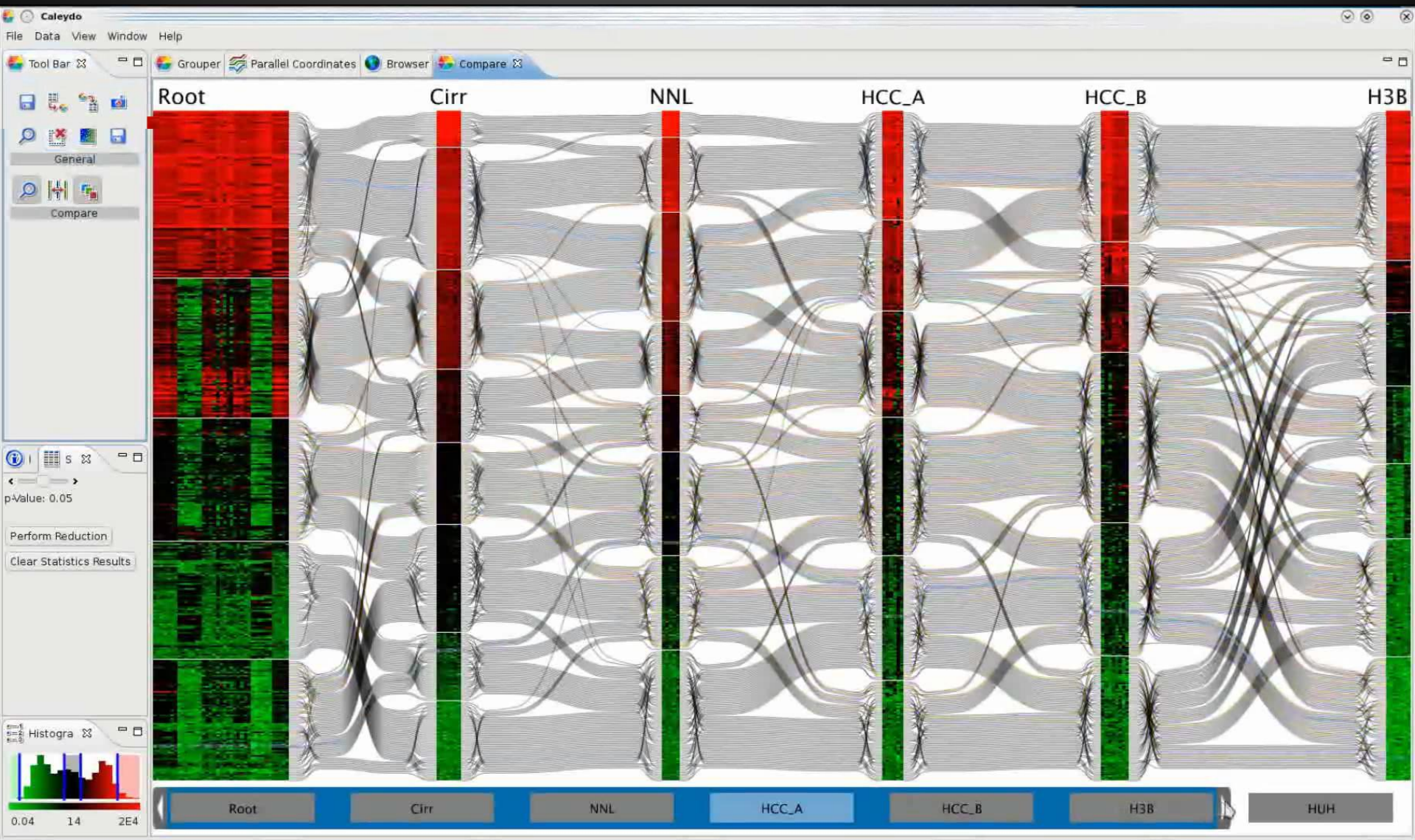
Alexander Lex

InfoVis IO

Overview


Caleydo
File Data View Window Help
Tool Bar
Grouper
Parallel Coordinates
Browser
Compare
General
Compare
p-Value: 0.05
Perform Reduction
Clear Statistics Results
Histogra
0.04 14 2E4
Root
Cirr
NNL
HCC_A
HCC_B
H3B
Root
Cirr
NNL
HCC_A
HCC_B
H3B
HUH

Details-on-Demand

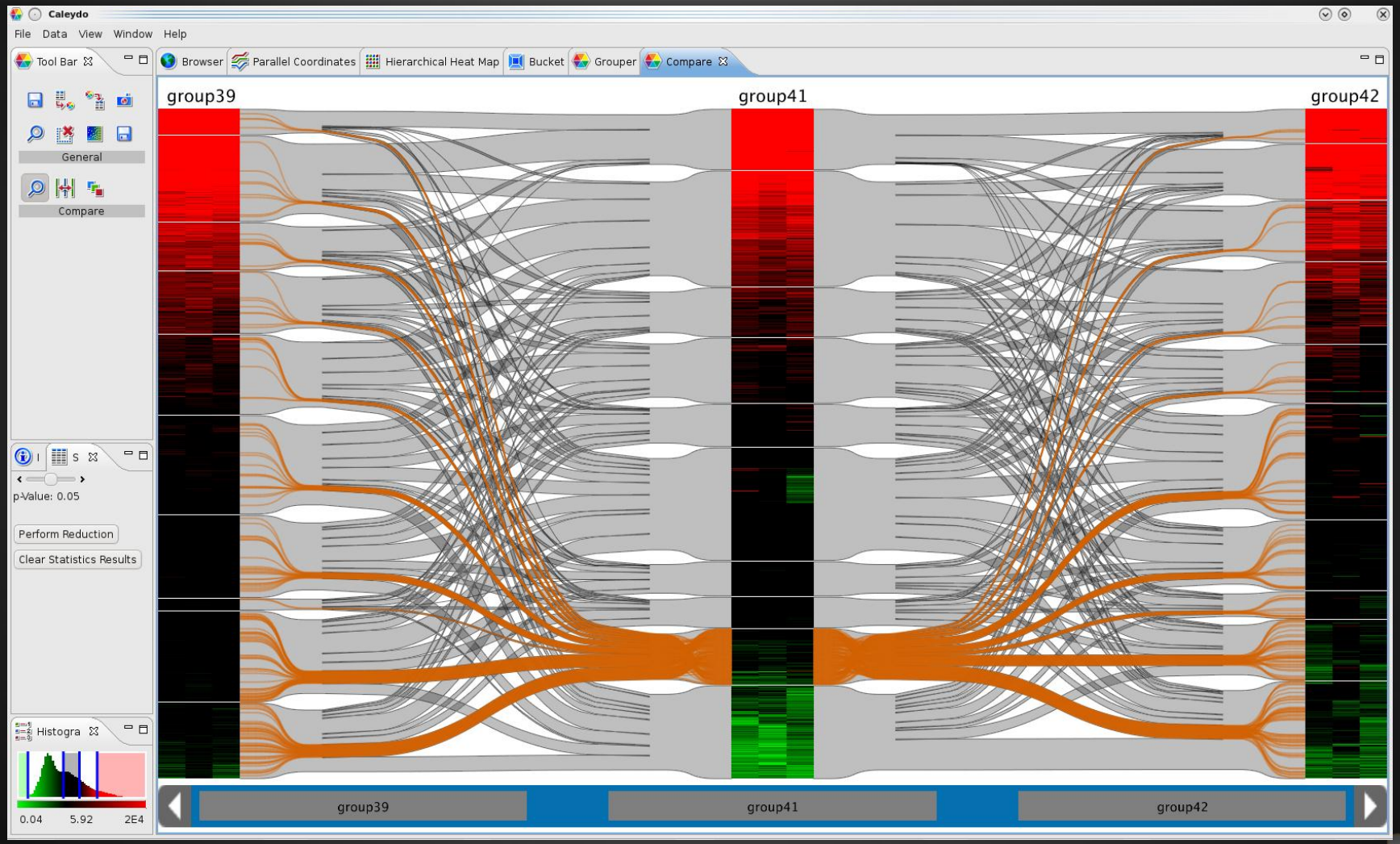


SCALABILITY AND IMPLEMENTATION

Scalability

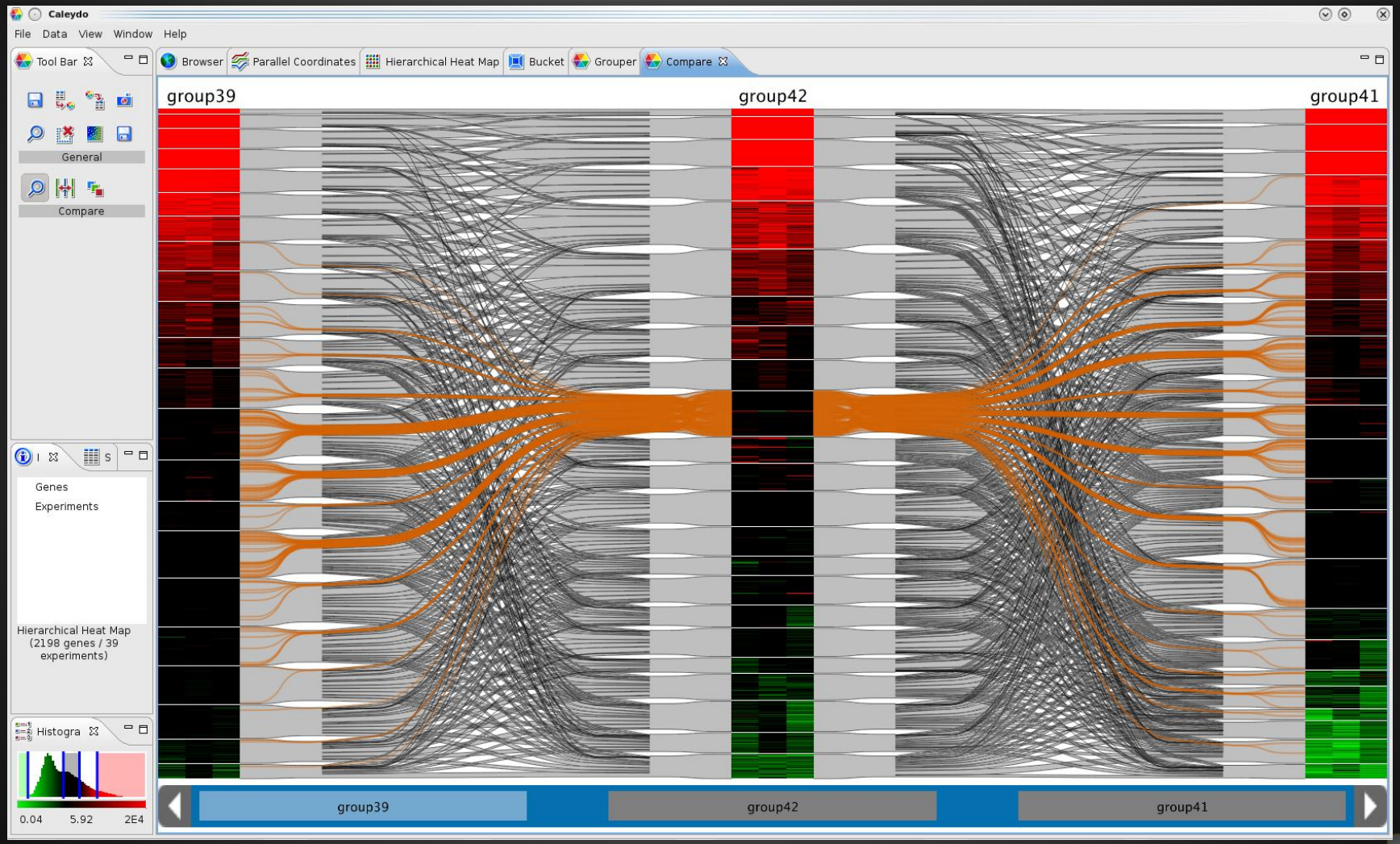
- Most sensitive to number of clusters
- For Overview
  - 20 clusters for < 2000 records
  - 10 clusters for < 3000 records
- 6 groups of dimensions simultaneously
- 100 or more dimensions

2000 Records, 10 Clusters

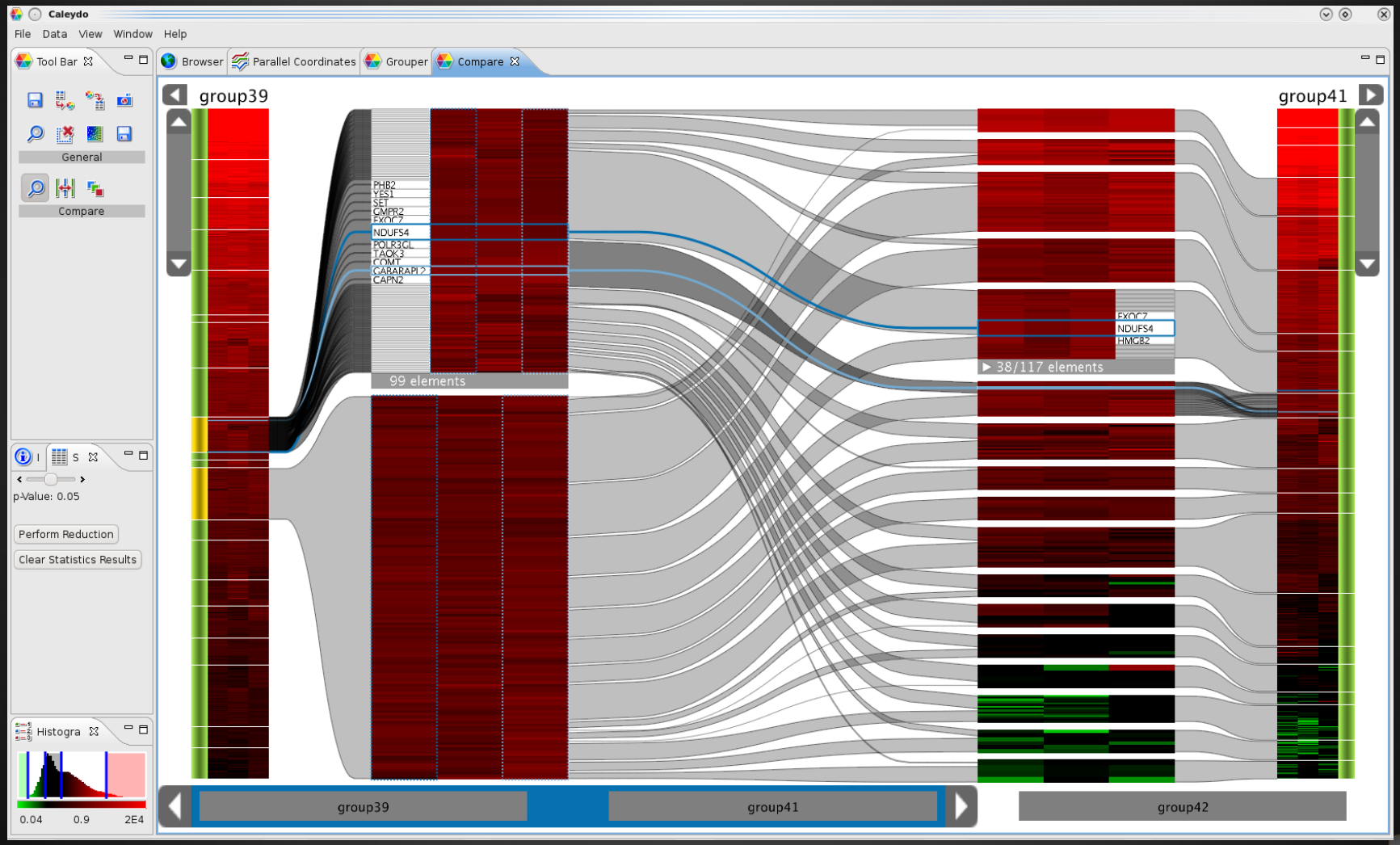


Alexander Lex

2000 Records, 20 Clusters



2000 Records, 20 Clusters in Detail



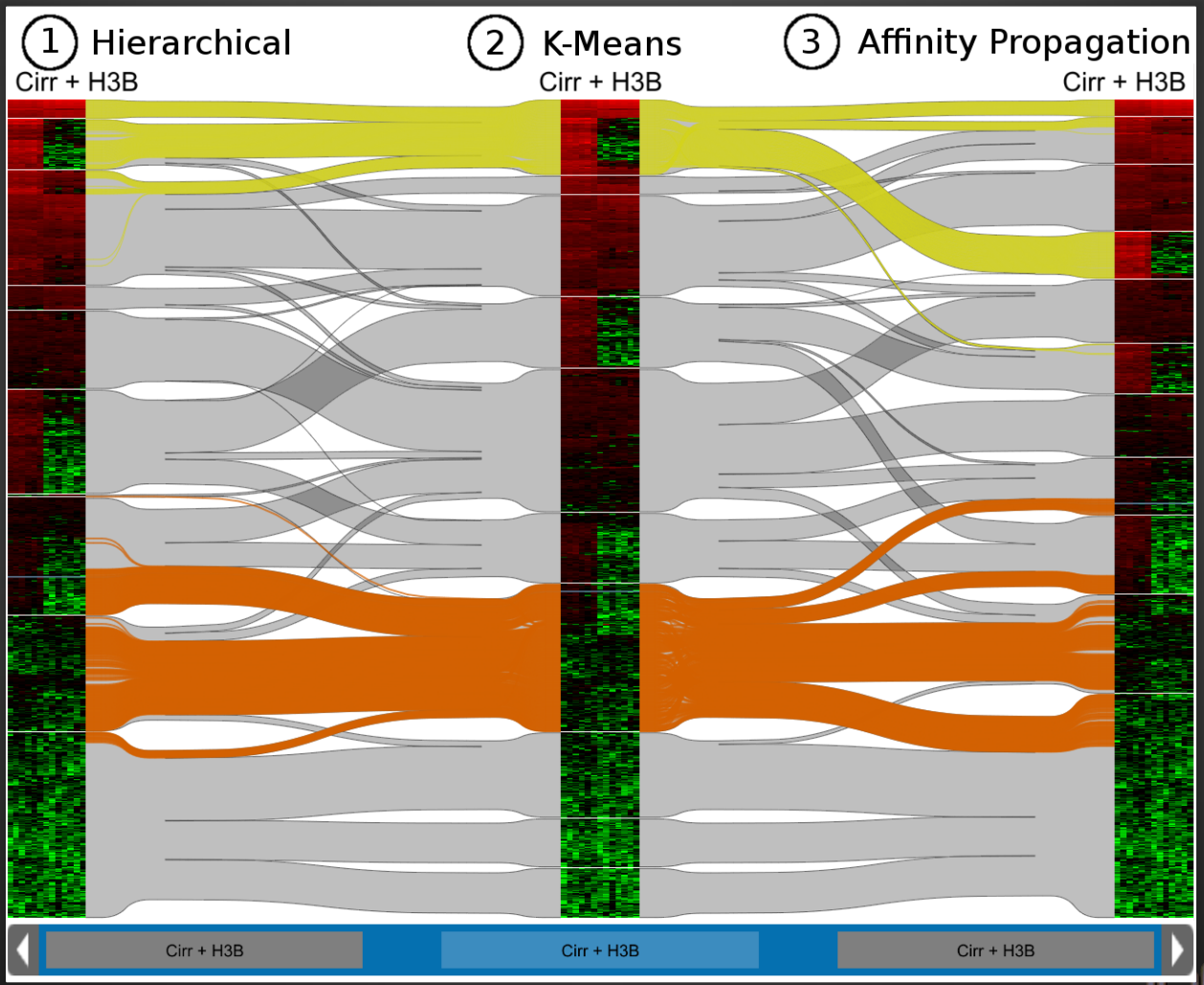
Alexander Lex

Implementation

- Based on the Caleydo Bioinformatics Visualization Framework
- Java, OpenGL
- Clustering either native implementation, R or Weka

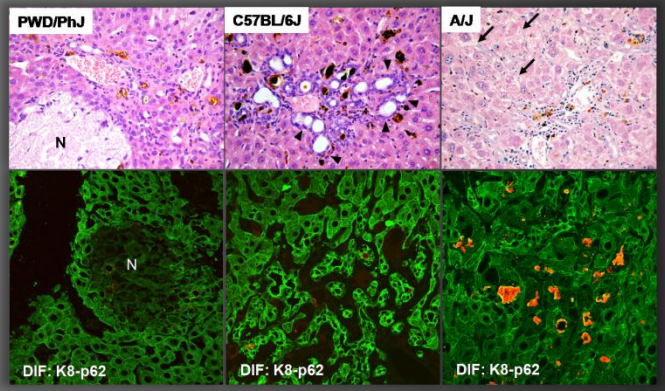
APPLICATION EXAMPLES

Example: Cluster Algorithm Comparison



Example: Comparative analysis of mice strains under intoxication

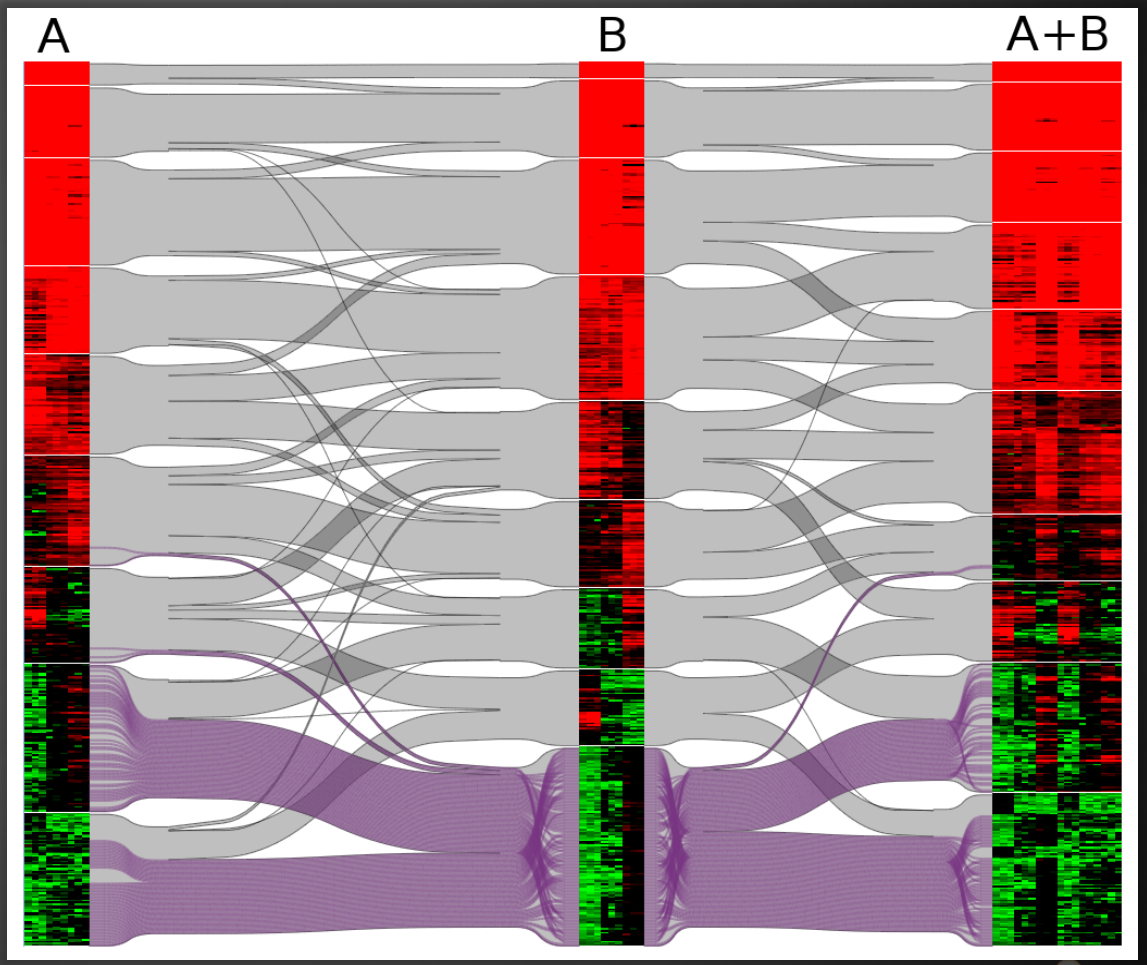
- Goal: find genetic reasons for liver cirrhosis
- Different mouse strains
  - Some show symptoms of cirrhosis after being fed poison for 8 weeks
  - Others don‘t



Low Responders *(< 5% MB- containing hepatocytes)*	Medium Responders *(5-30% MB- containing hepatocytes)*	High Responders *(>30% MB- containing hepatocytes)*
PWD/PhJ	C57BL/6J	OF1 Swiss Albino
	FVB/N	C3H/He
	129P2	A/J

Example: Comparative analysis of mice strains under intoxication

- 3 replicates
- 3 time points
  - 0 days
  - 7 days
  - 8 weeks
- 2 strains: A and B
- Grouping
  - Group A
  - Group B
  - Group A+B



CONCLUSION

Conclusion

- Technique for comparing groups of dimensions in multidimensional datasets
- General approach
  1. Define interesting subspaces
  2. Use clustering and heat maps to visualize subspace
  3. Use curves or pipes to re-introduce connections

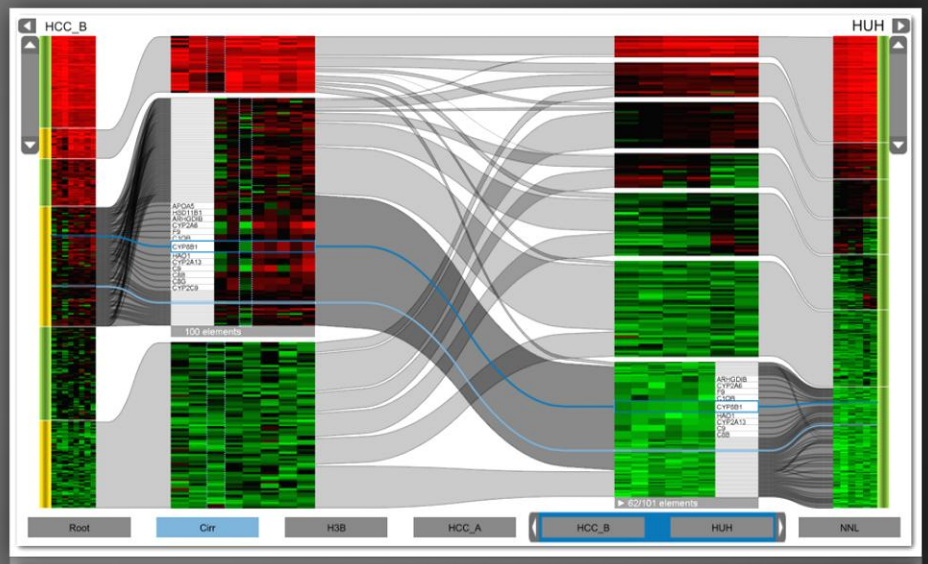
Thanks

- To our reviewers for their valuable feedback
- To Helmut Doleisch and Bernhard Schlegl
- Our funding agencies: FFG and FWF
- For your attention!

Caleydo Matchmaker

http://www.caleydo.org/matchmaker

Alexander Lex, Marc Streit, Christian Partl, Karl Kashofer, Dieter Schmalstieg



CALEYDO

InfoVis 10