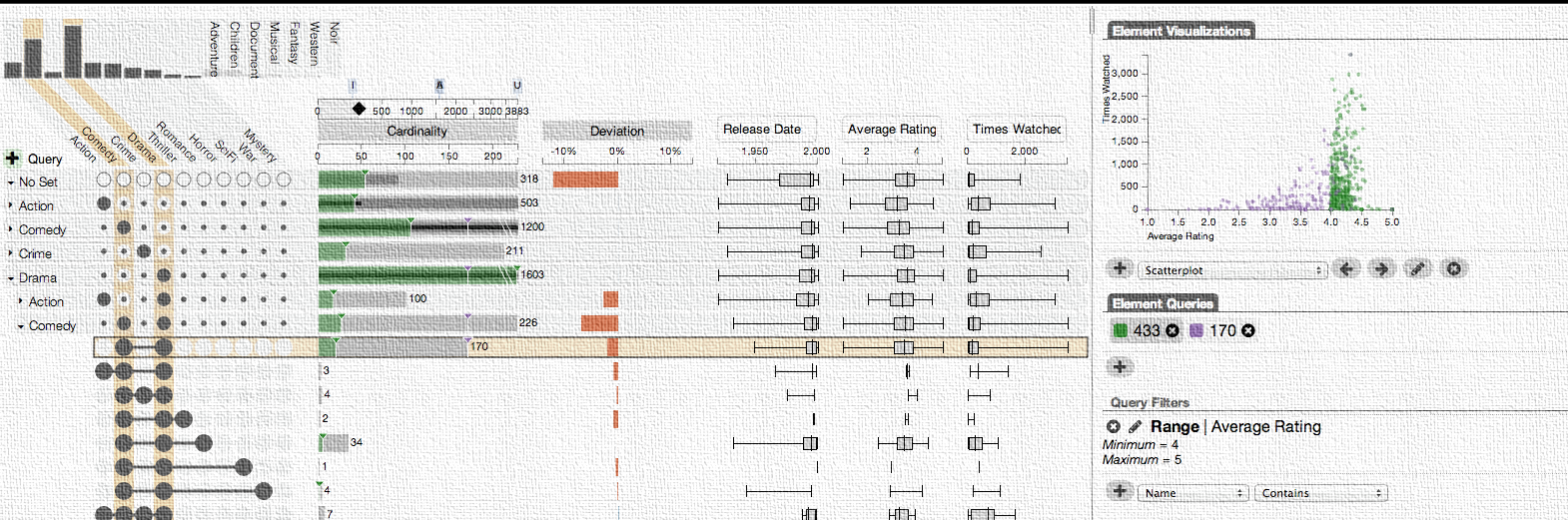




Enabling Scientific Discovery through Interactive Visual Data Analysis



visualization

pictures

The purpose of computing is insight, not numbers.

- Richard Wesley Hamming
- Card, Mackinlay, Shneiderman

Banana ***M. acuminata***

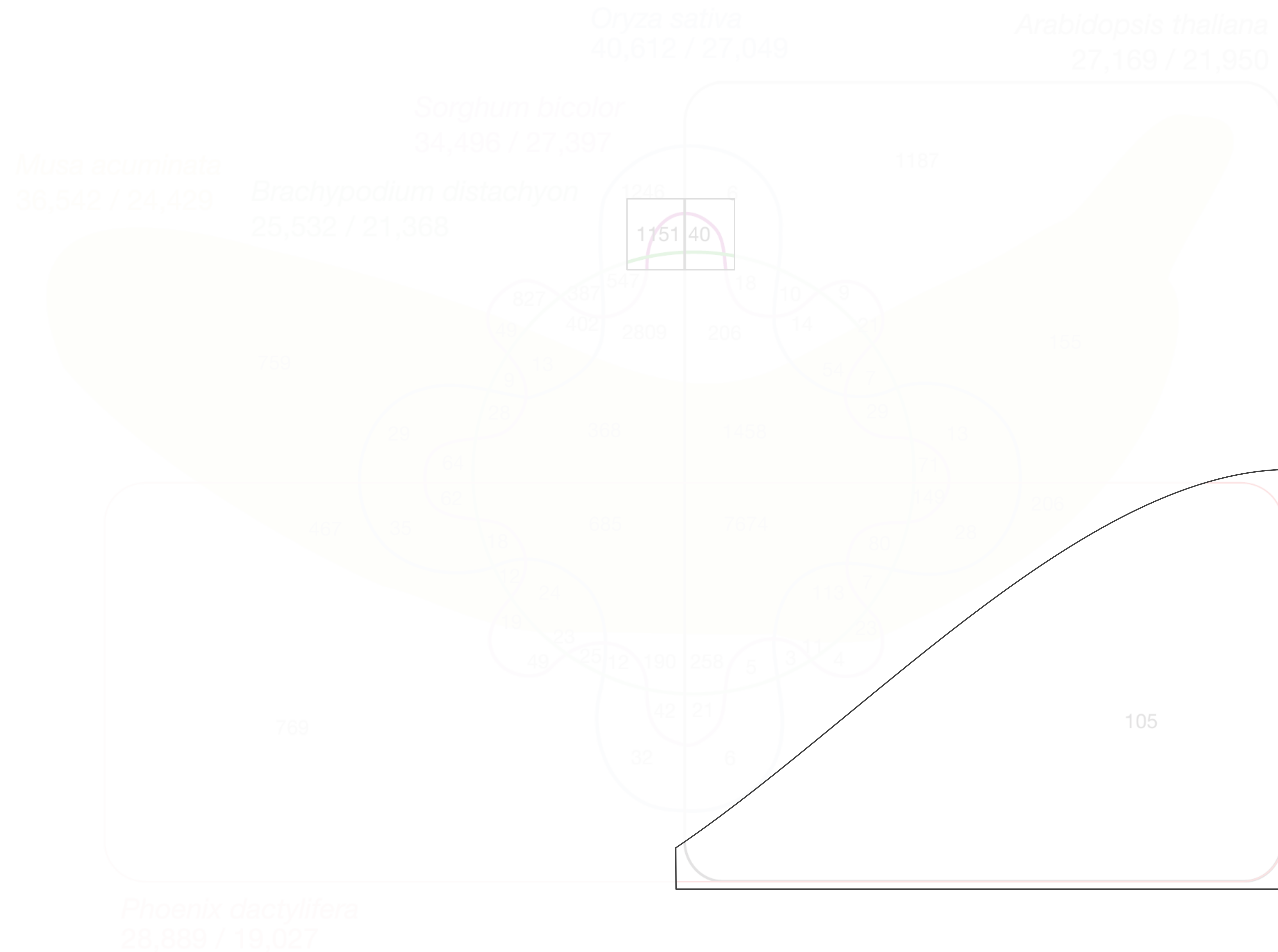
Date ***P. dactylifera***

Cress ***Arabidopsis thaliana***

Rice ***Oryza sativa***

Sorghum ***Sorghum bicolor***

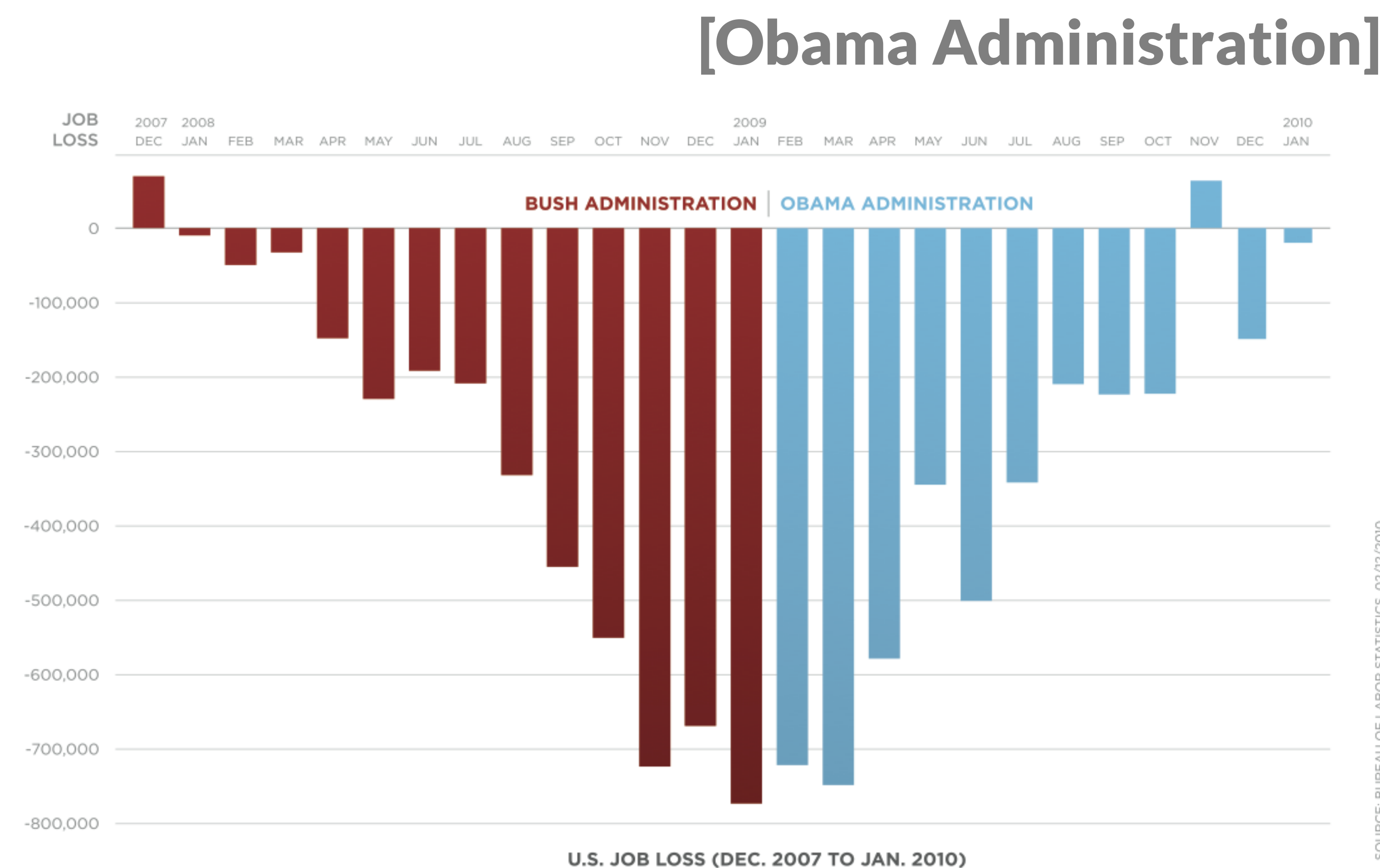
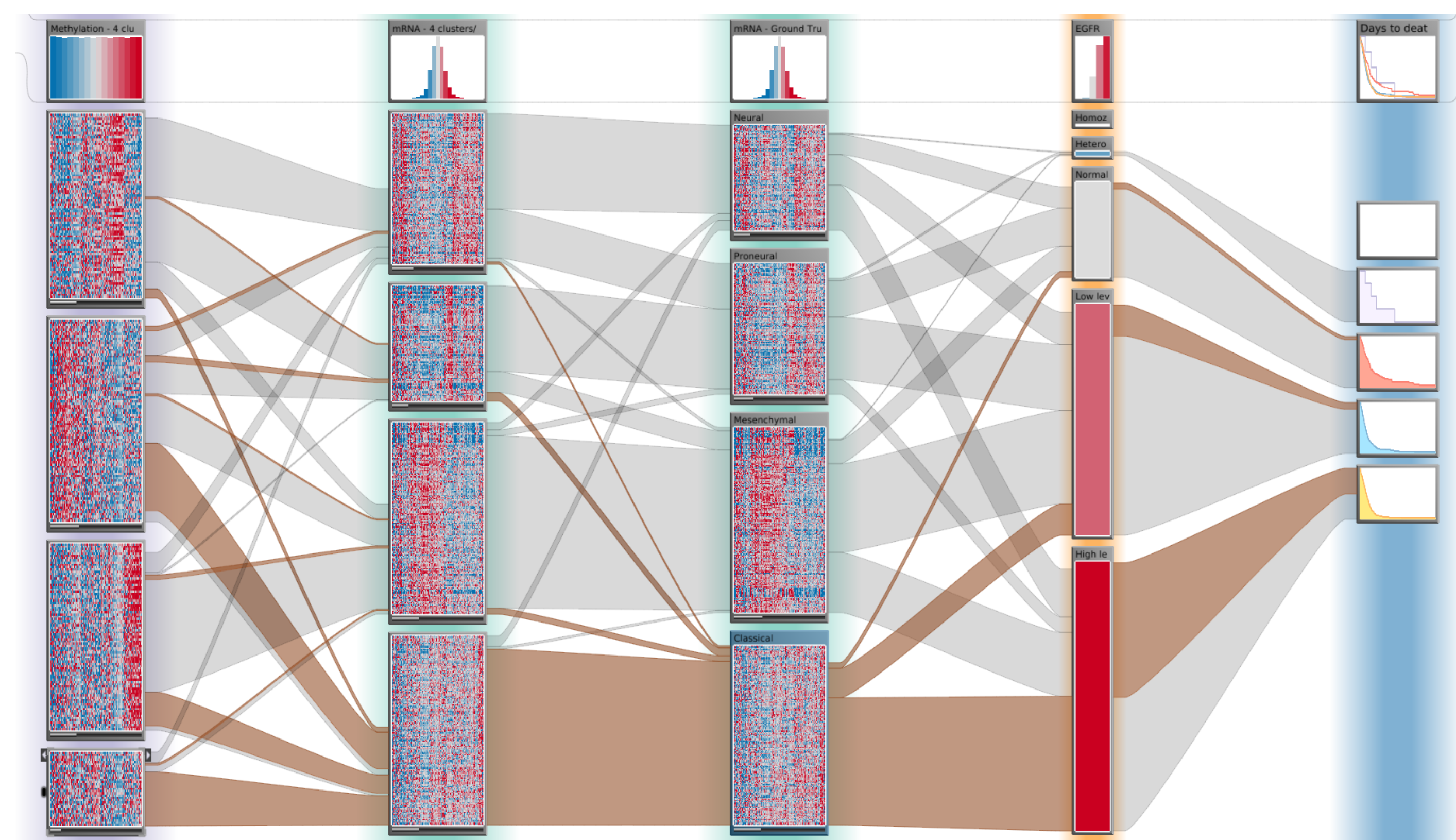
Brome ***Brachypodium distachyon***



Good Data Visualization

- ... makes data **accessible**
- ... combines strengths of
humans and computers
- ... enables **insight**
- ... **communicates**

Purpose of Visualization



Open Exploration

Communication

Research Approach

Technique Driven

Solving a general visual analysis problem

Broad application

General questions

How to visualize many sets?

How to make rankings interactive?

Visualization Method

Problem Driven

Solving a domain problem of collaborator

Complex, targeted systems

Help answer important questions

What characterizes a cancer subtype?

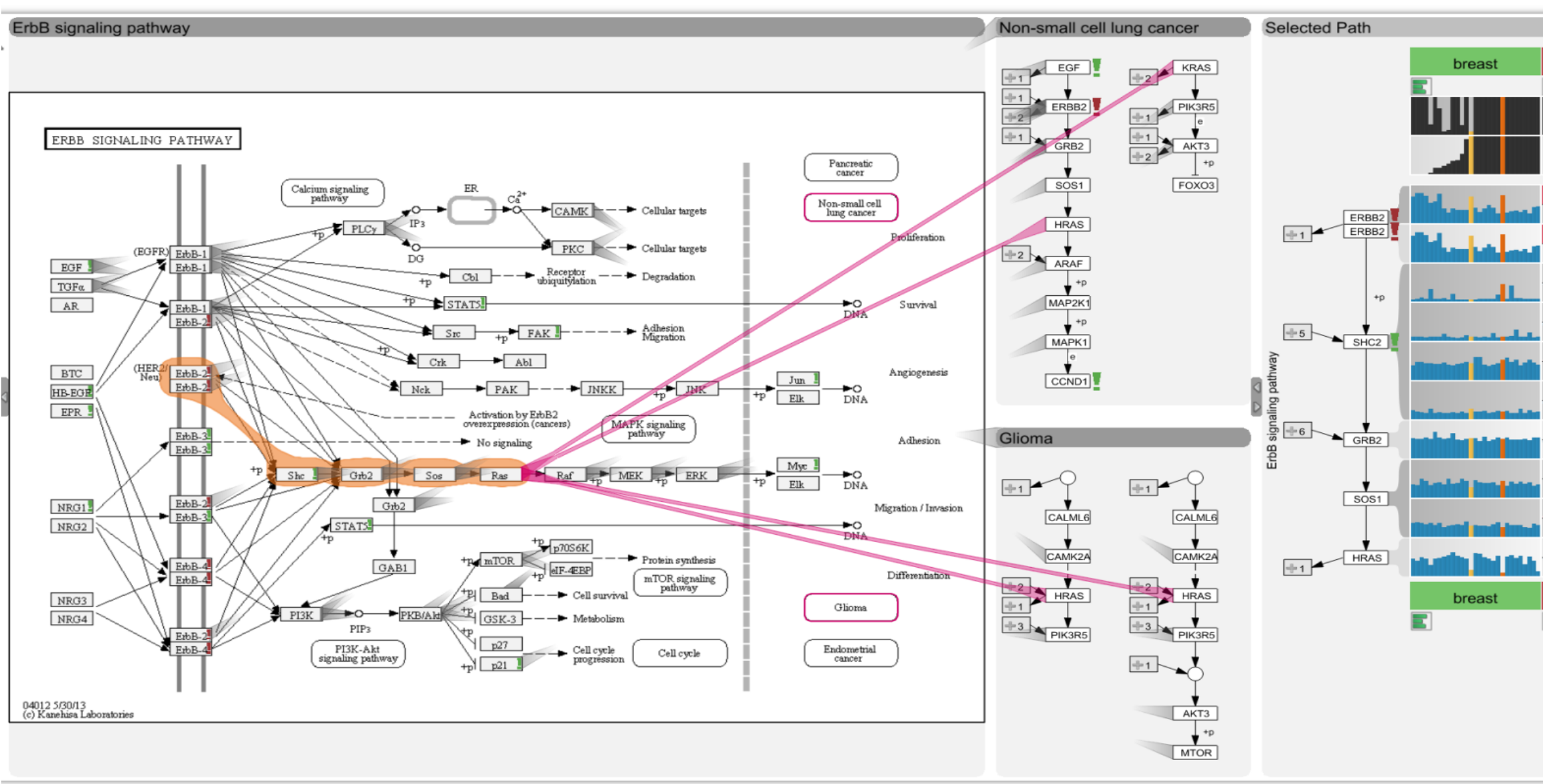
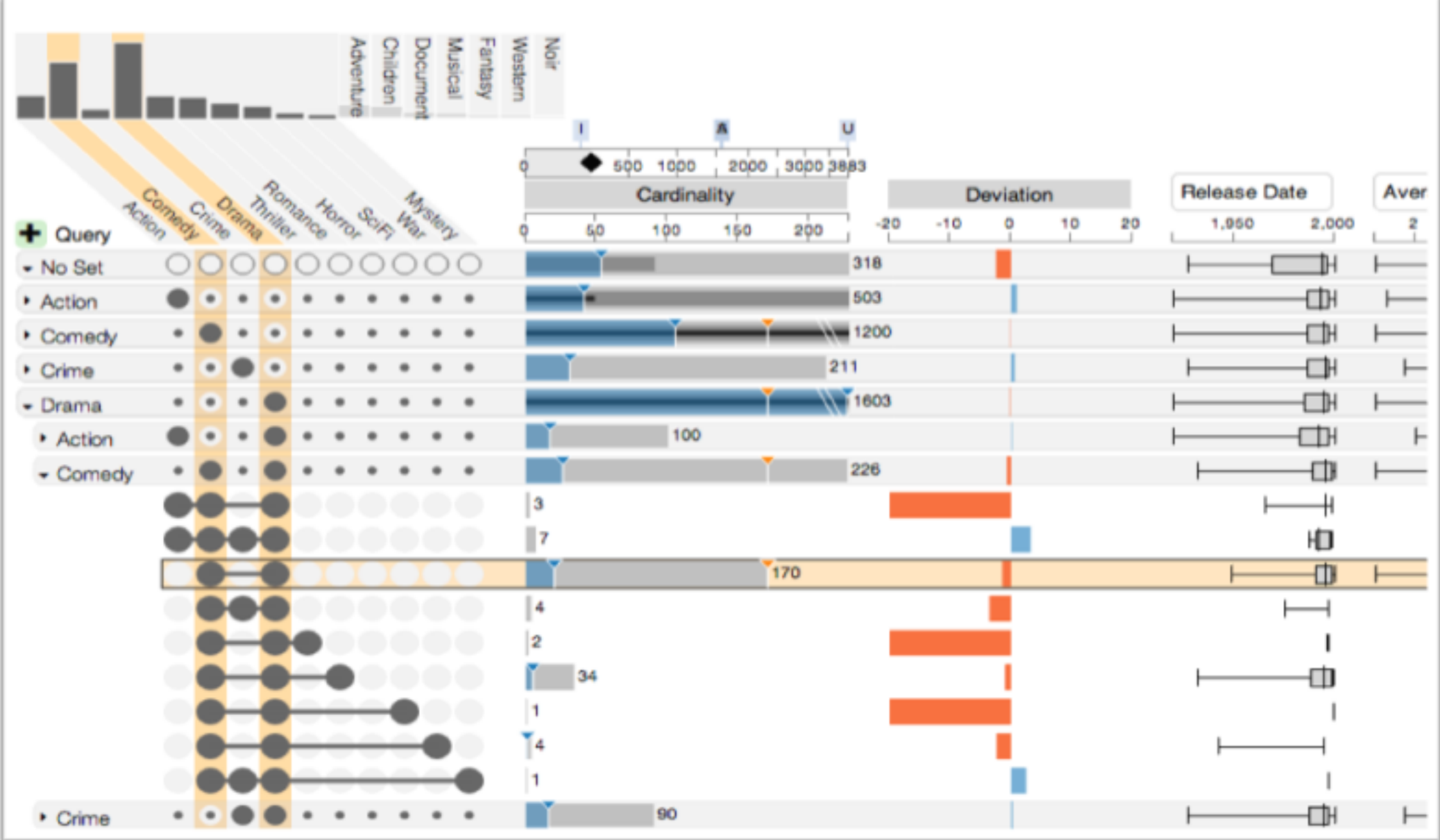
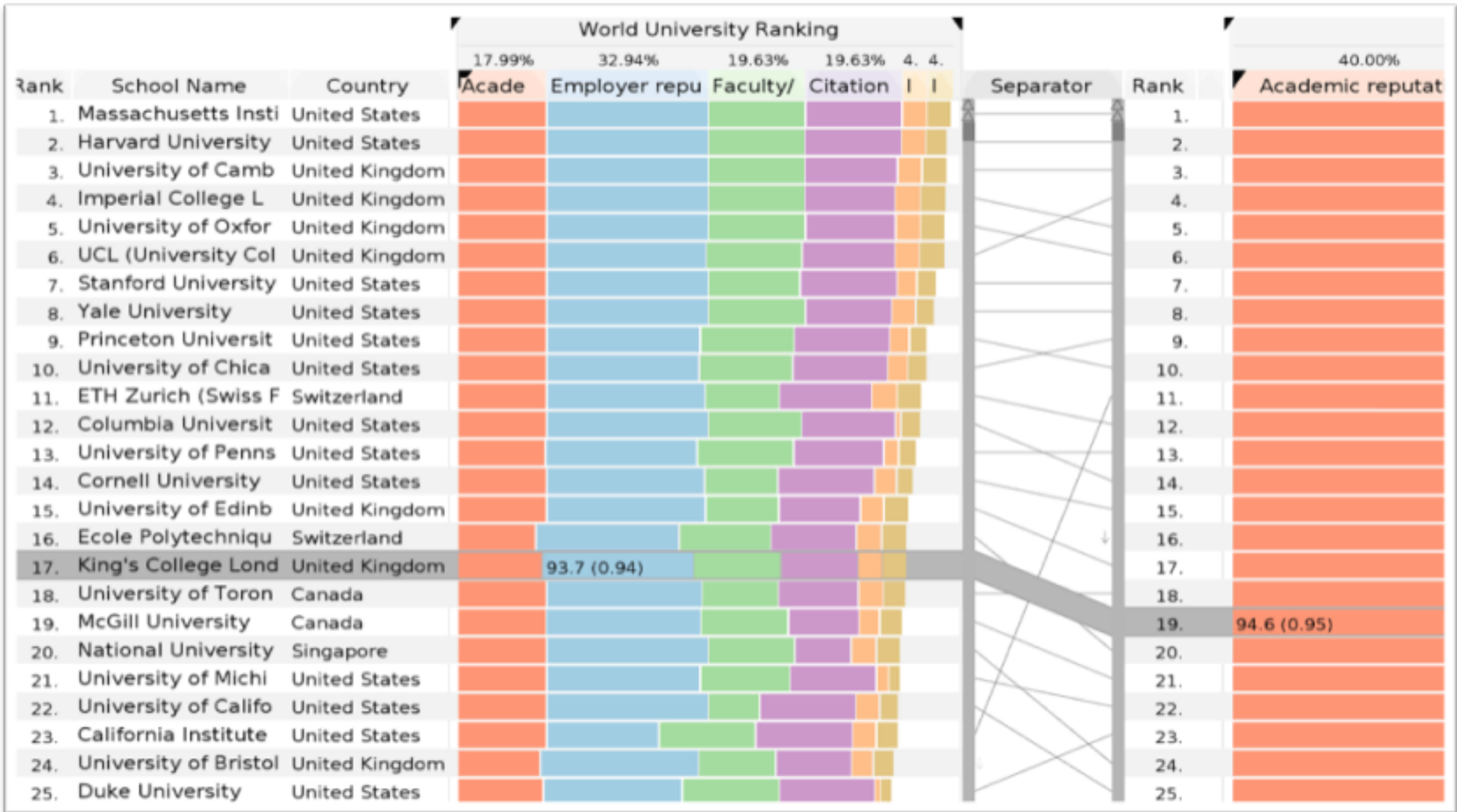
Which drug works for these patients?

Visualization Application

Content

Techniques

Domain Solution



Multi-Attribute Rankings

Set Visualization

Multivariate Networks Pathway Analysis

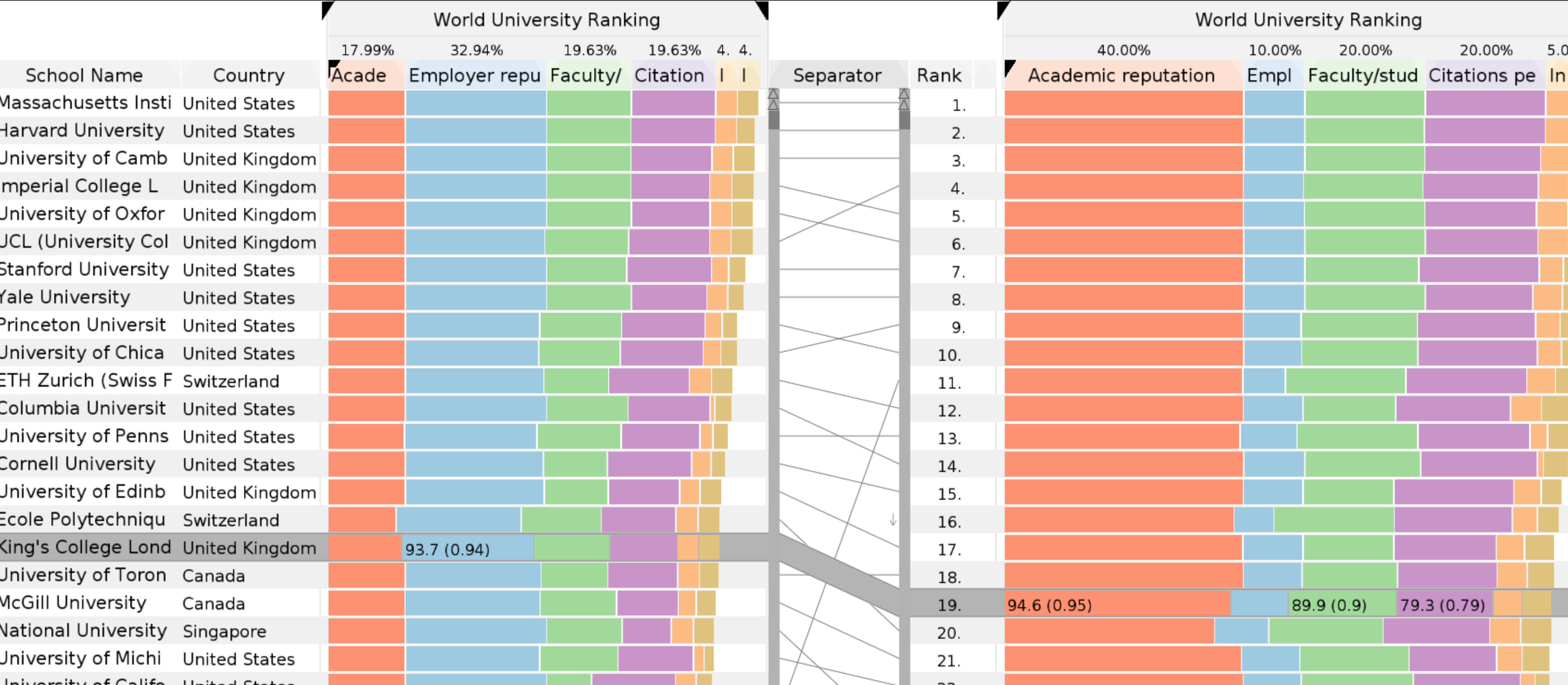
Lay Users

Expert Analysts

[InfoVis '13]

Best Paper Award

Rankings





Rankings are popular

Goal

Intuitive

Interactive

Multi-Attribute

Ranking Visualization

To Create

Refine

Explore

Rank	University	Score
1.	MIT, USA	
2.	Harvard, U	
3.	Princeton,	
4.	Cambridge	
5.	Oxford, UK	4.0

Support Multiple Attributes

Score = f(A, B, C)

Rank	University	A	B	C
1.	MIT, USA			
2.	Harvard, USA			
3.	Princeton, US			
4.	Cambridge, U			
5.	Oxford, UK			

Combiner functions: $f(A,B,C)$

(Weighted) sum

$$\text{Score} = w_a A + w_b B + w_c C$$

→ Serial

Maximum

$$\text{Score} = \max(A, B, C)$$

→ Parallel

Product

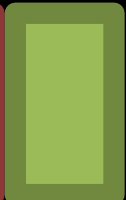
Nesting

...

} → Complex
Combiners

Serial Combiner (as Stacked Bar)



Rank	University	A	B	C
1.	MIT, USA			
2.	Harvard, USA			
3.	Princeton, USA			
4.	Cambridge, UK			
5.	Oxford, UK			

Serial Combiner (as Stacked Bar)



Rank	University	A	B	C
1.	MIT, USA			
2.	Harvard, USA			
3.	Princeton, USA			
4.	Cambridge, UK			
5.	Oxford, UK			

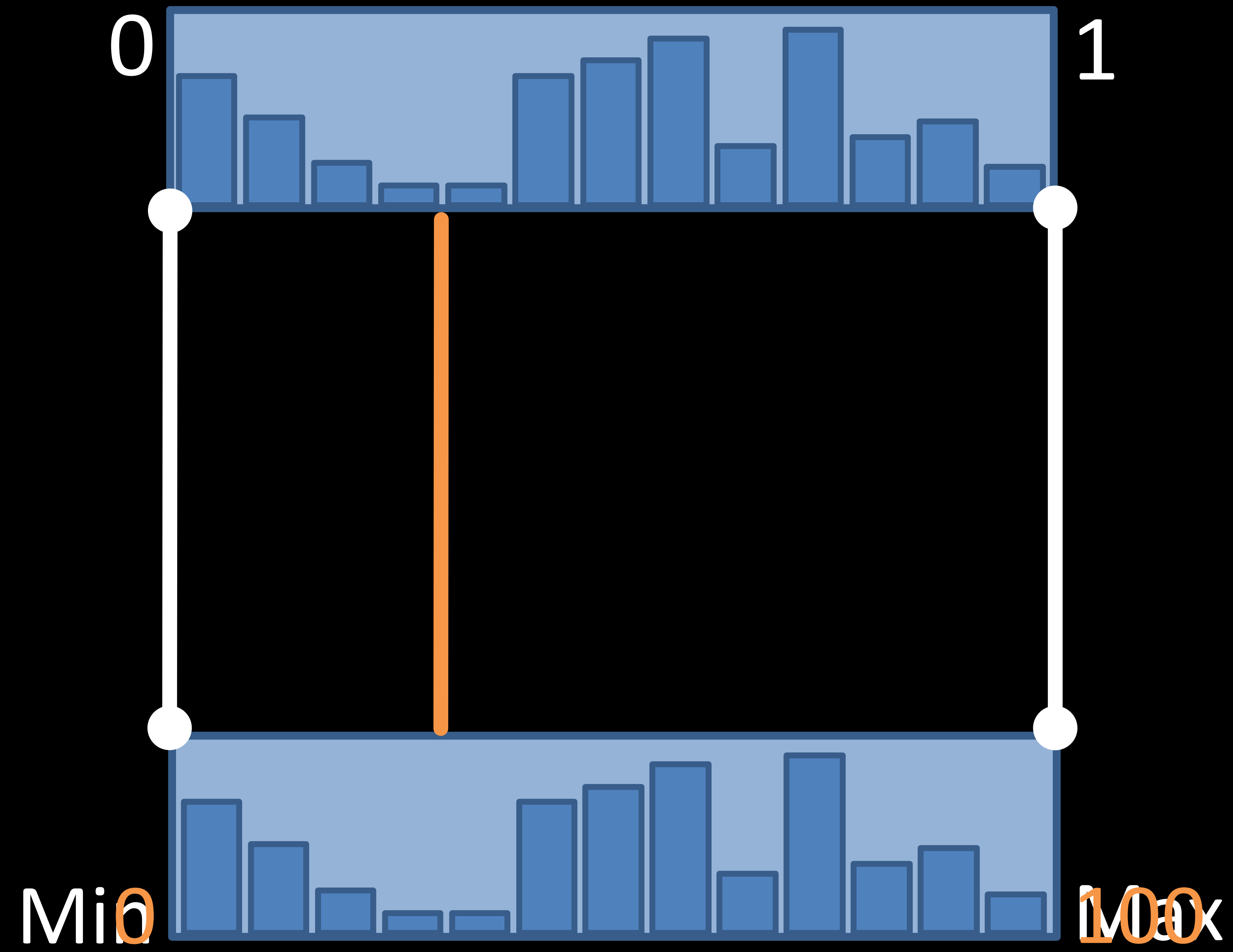
Serial Combiner (as Stacked Bar)

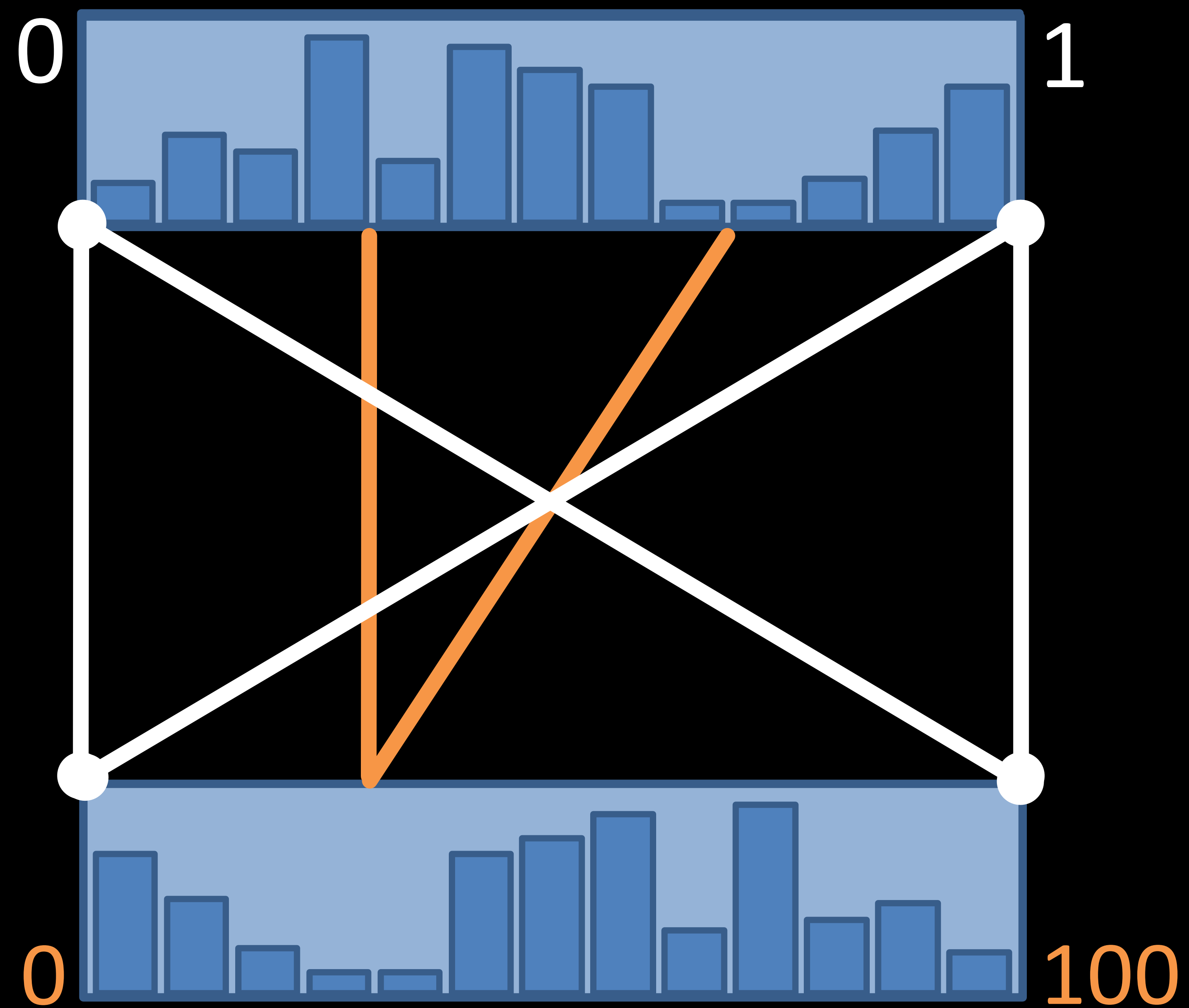
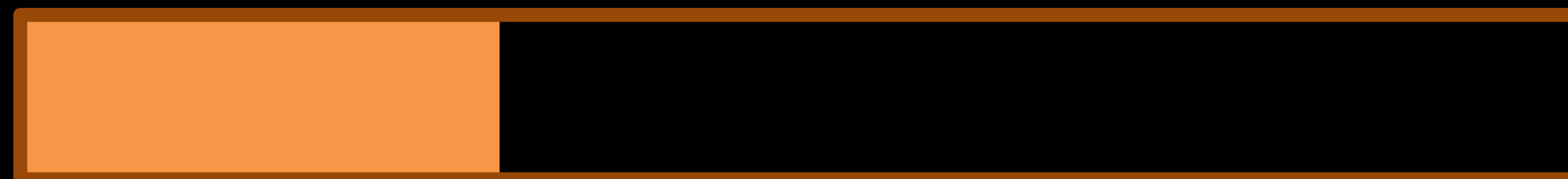


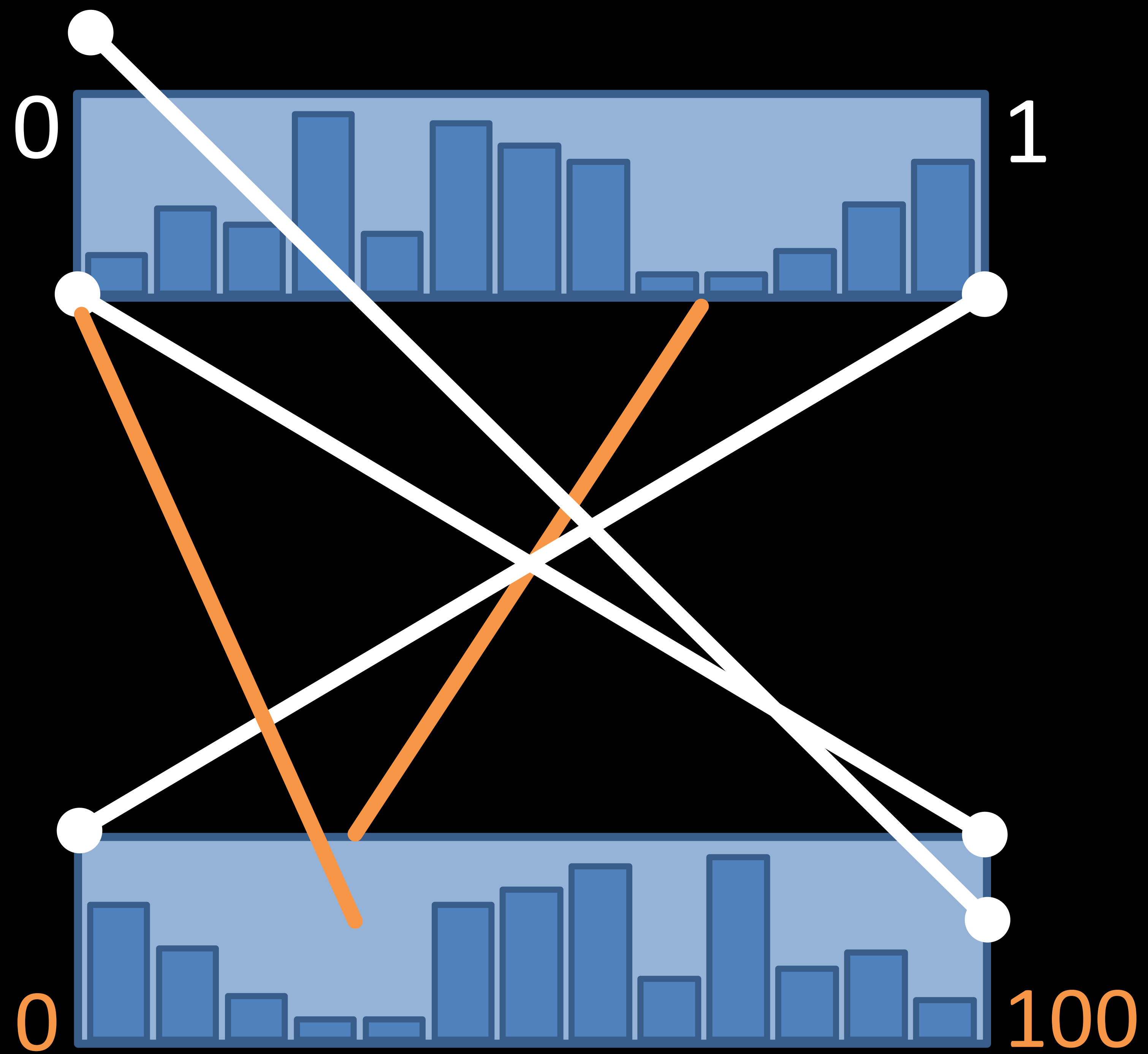
Rank	University	A	B	C
1.	MIT, USA			
2.	Harvard, USA			
3.	Princeton, USA			
4.	Cambridge, UK			
5.	Oxford, UK			

Rank	School Name	Country	Faculty/student ratio	Employer reputation	Citations per faculty
			Filter: <None>	Filter: 2 out of 72	
1.	American University	United States			
2.	Arizona State University	United States			
3.	Aston University	United Kingdom			
4.	Birkbeck College, University of L	United Kingdom			
5.	Boston College	United States			
6.	Boston University	United States			
7.	Brandeis University	United States			
8.	Brown University	United States			
9.	Brunel University	United Kingdom			
10.	California Institute of Technology	United States			
11.	Cardiff University	United Kingdom			
12.	Case Western Reserve University	United States			
13.	City University London	United Kingdom			
14.	College of William & Mary	United States			
15.	Colorado State University	United States			
16.	Columbia University	United States			
17.	Cornell University	United States			
18.	Cranfield University	United Kingdom			
19.	Dartmouth College	United States			
20.	Drexel University	United States			
21.	Duke University	United States			
22.	Durham University	United Kingdom			

Flexible Mapping of Attributes to Scores

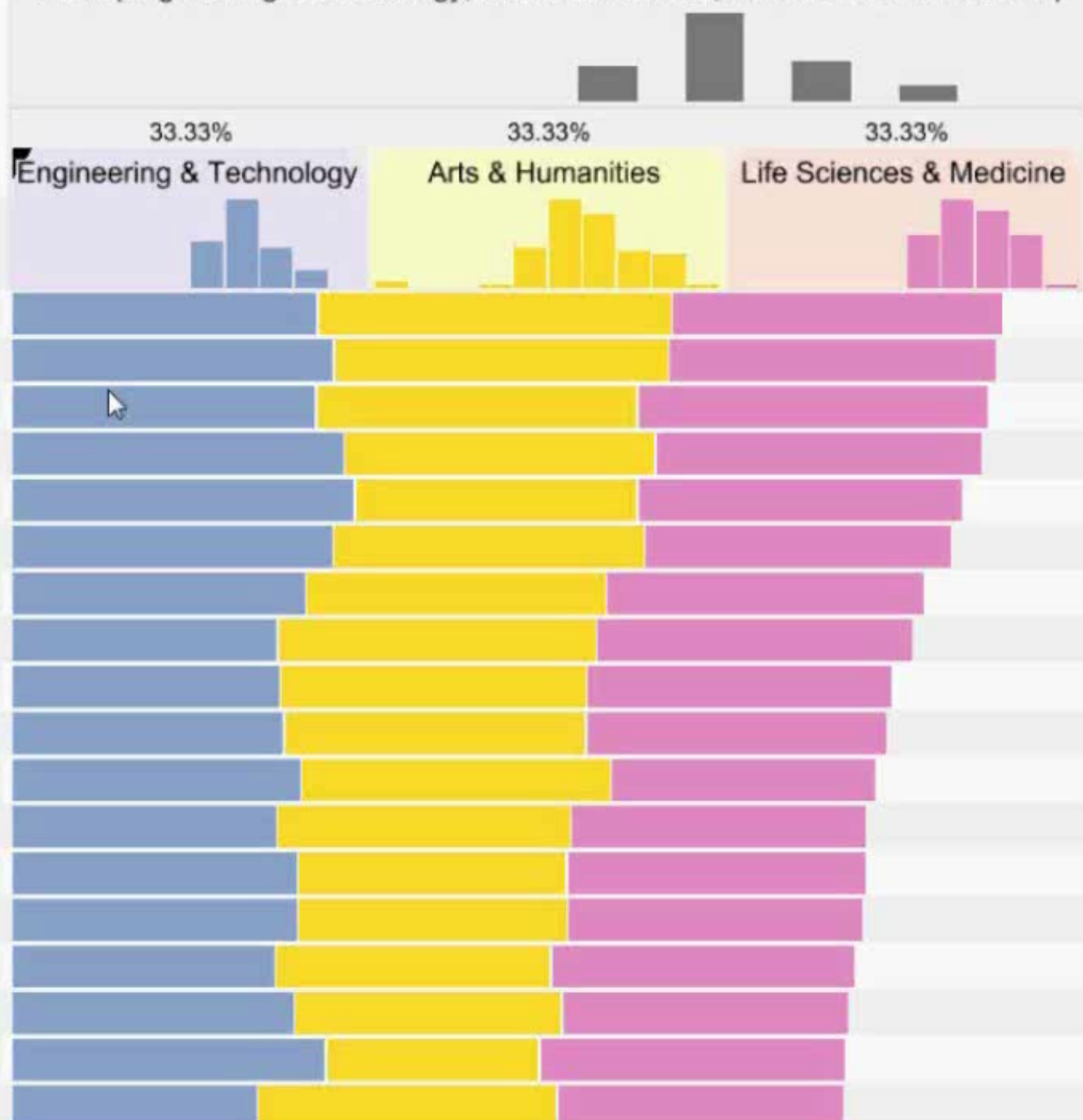






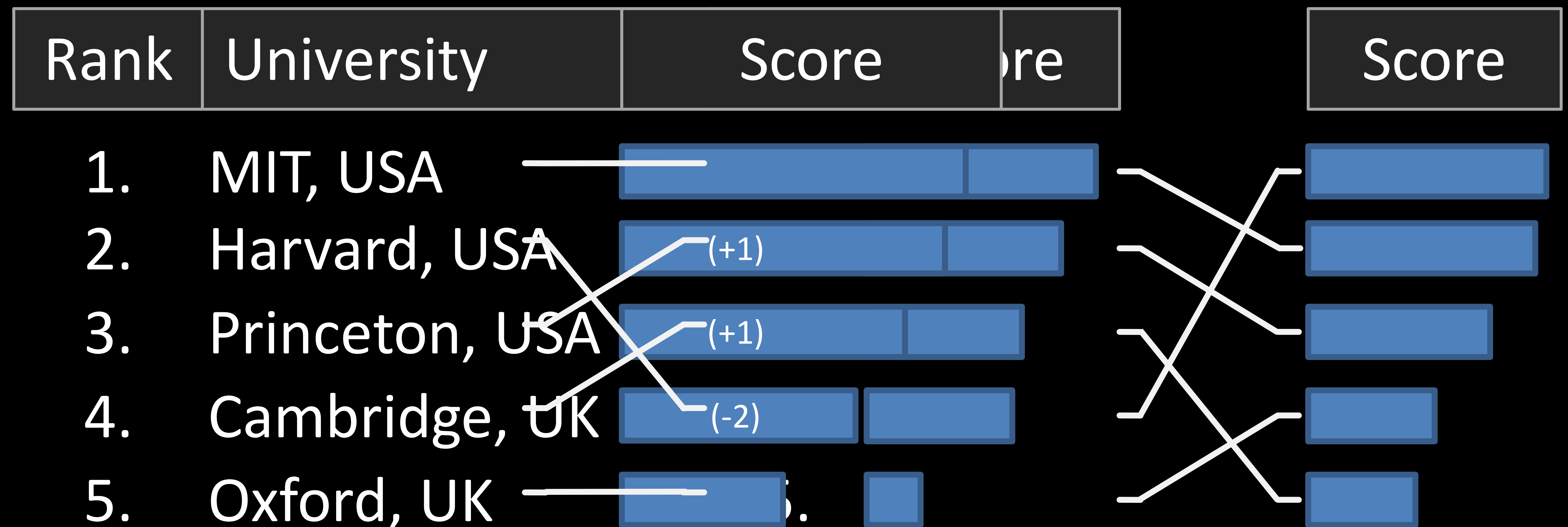
Ran	School Name	Country
Filter: <None>		Filter: 2 out of 43
1.	University of Oxford	United Kingdom
2.	University of Cambridge	United Kingdom
3.	Harvard University	United States
4.	Stanford University	United States
5.	Massachusetts Institute of Technology (MIT)	United States
6.	University of California, Berkeley (UCB)	United States
7.	University of California, Los Angeles (UCL)	United States
8.	Yale University	United States
9.	UCL (University College London)	United Kingdom
10.	Columbia University	United States
11.	Princeton University	United States
12.	University of Edinburgh	United Kingdom
13.	University of Michigan	United States
14.	Cornell University	United States
15.	University of Pennsylvania	United States
16.	The University of Manchester	United Kingdom
17.	Imperial College London	United Kingdom
18.	University of Chicago	United States

SUM (Engineering & Technology, Arts & Humanities, Life Sciences & Medicine)

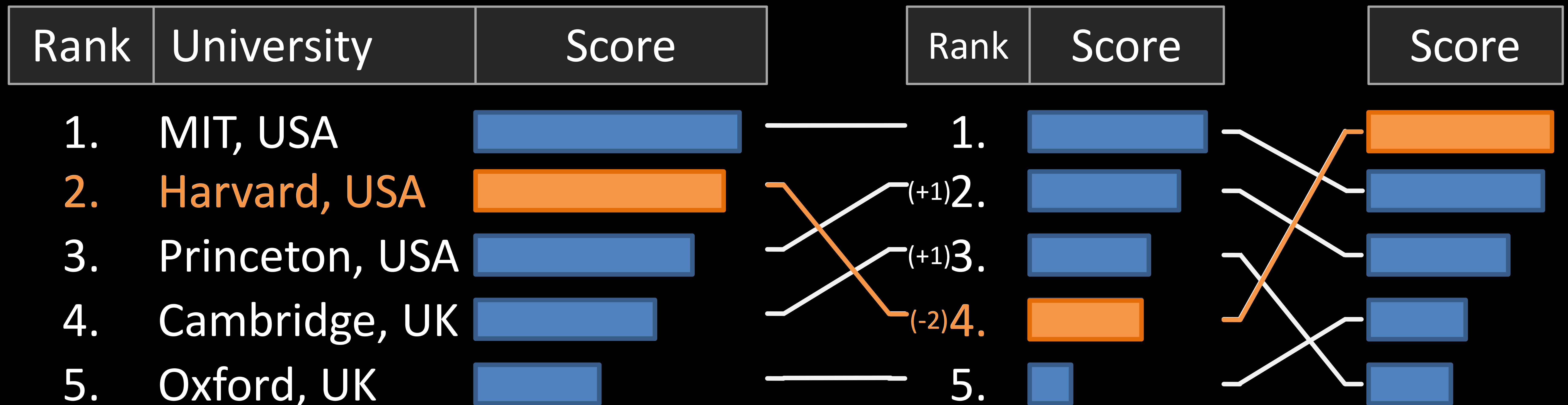


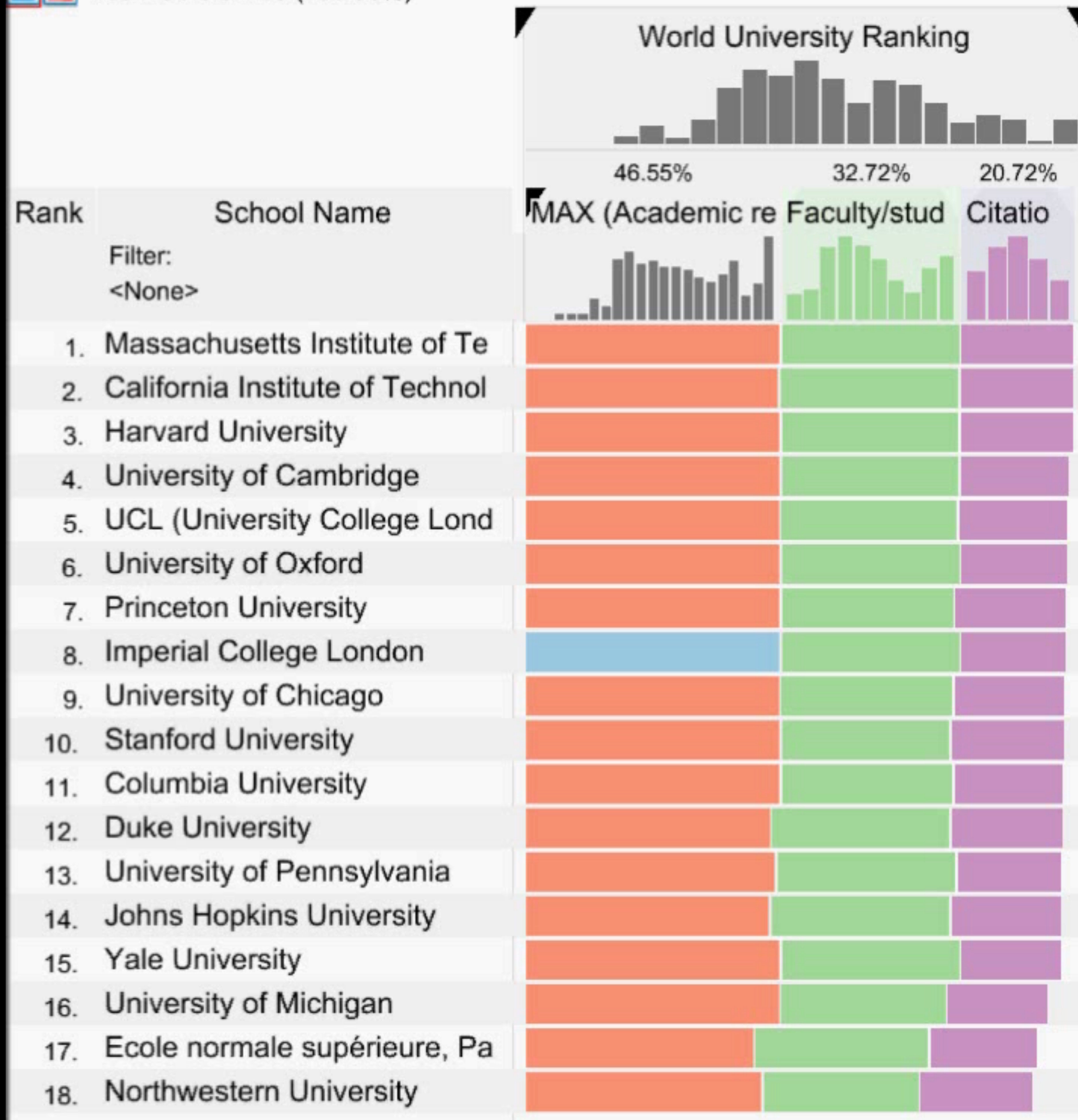
Compare Rankings

Bump Charts



Bump Charts

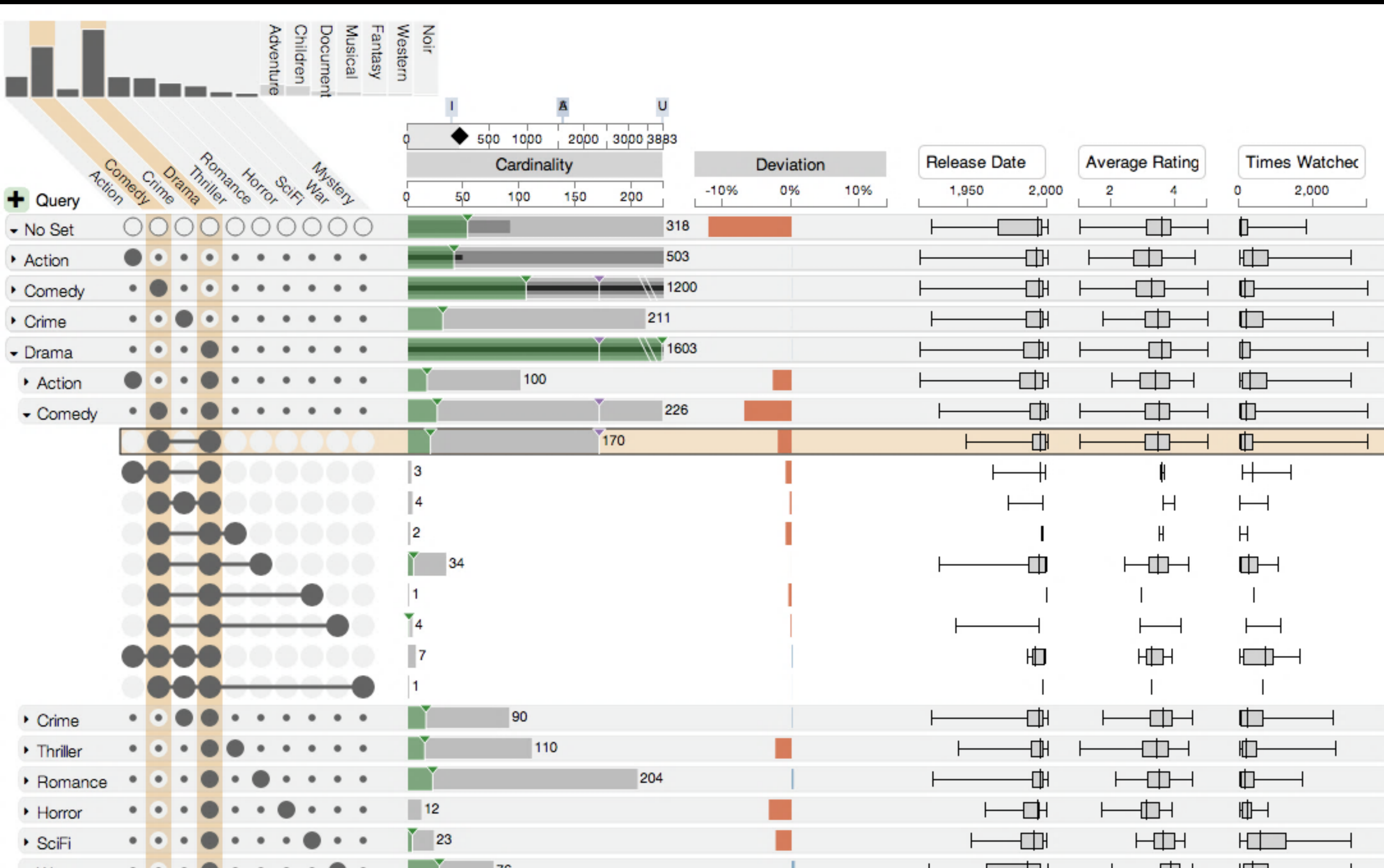




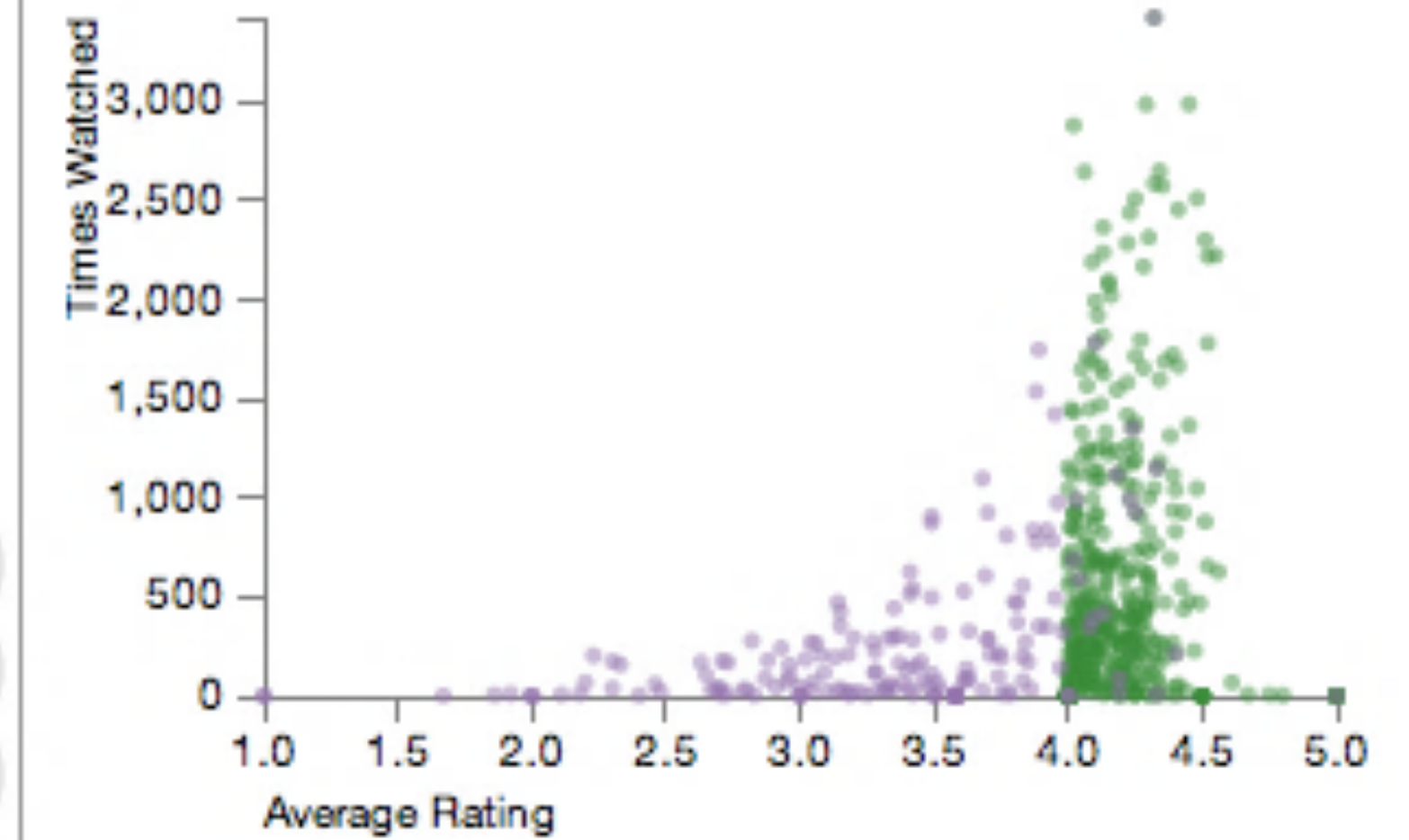
<http://lineup.caleydo.org>
<http://caleydo.github.io/lineup.js/>

[InfoVis'14]

UpSet Visualizing Intersecting Sets



Element Visualizations



+ Scatterplot

Element Queries

433 170

+

Query Filters

✖ Range | Average Rating

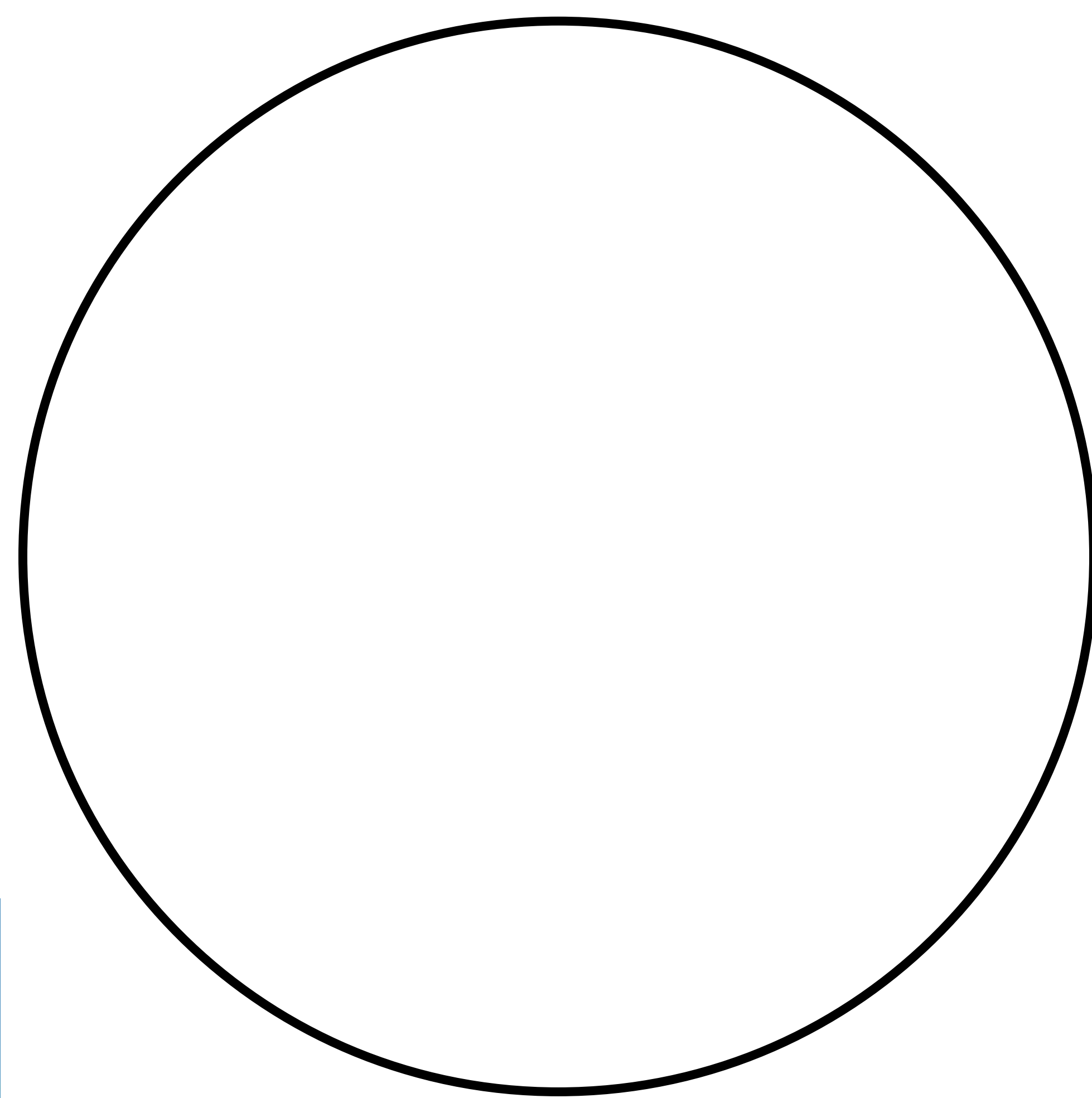
Minimum = 4

Maximum = 5

+ Name Contains

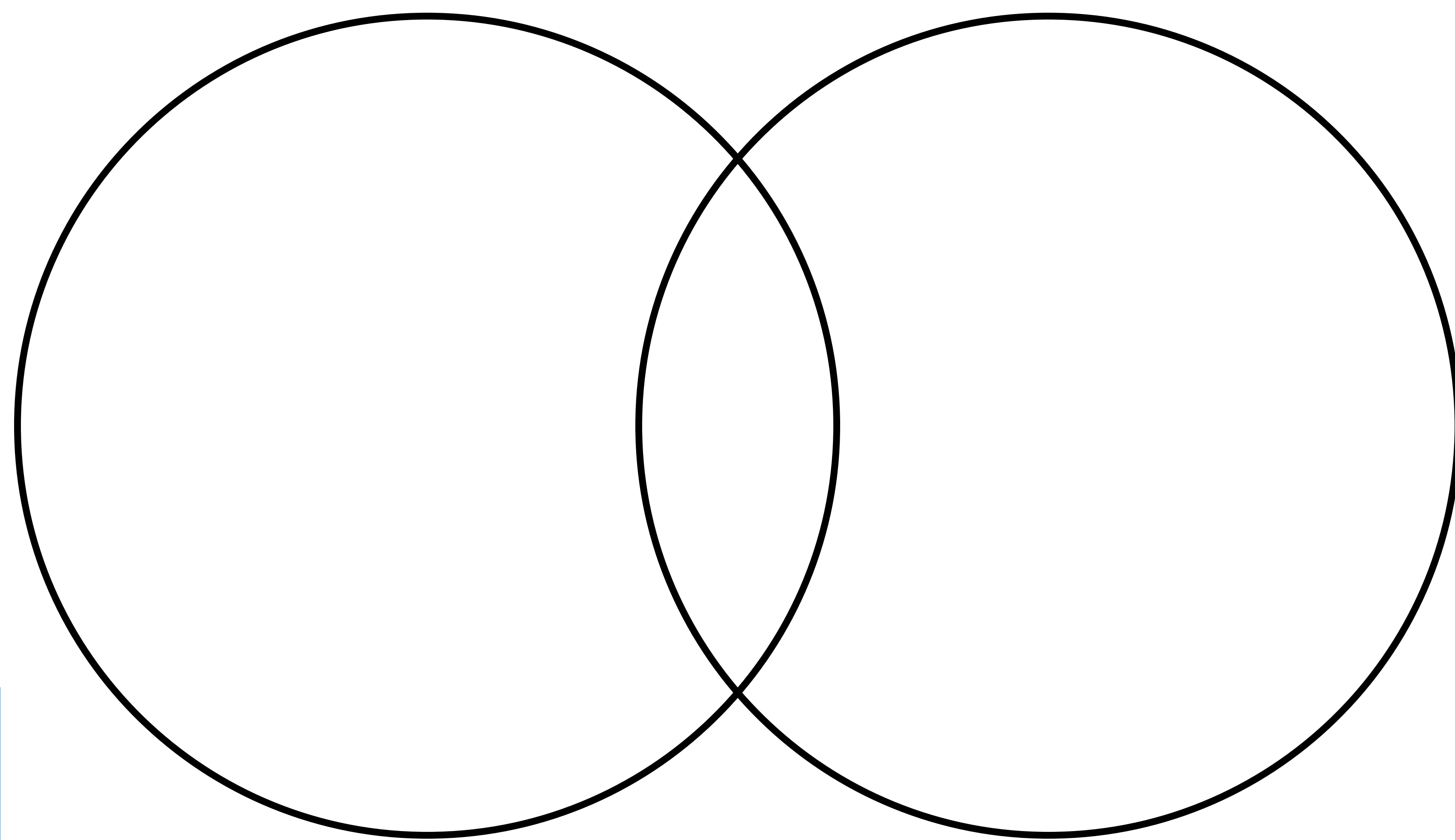
Query Results

Name	Release Date	Average Rating	Times Watched	Set Count
Toy Story (1995)	1995	4.15	2077	2
Sense and Sensibility (1995)	1995	4.03	835	2
Persuasion (1995)	1995	4.06	179	1
City of Lost Children, The (1995)	1995	4.06	403	2



School





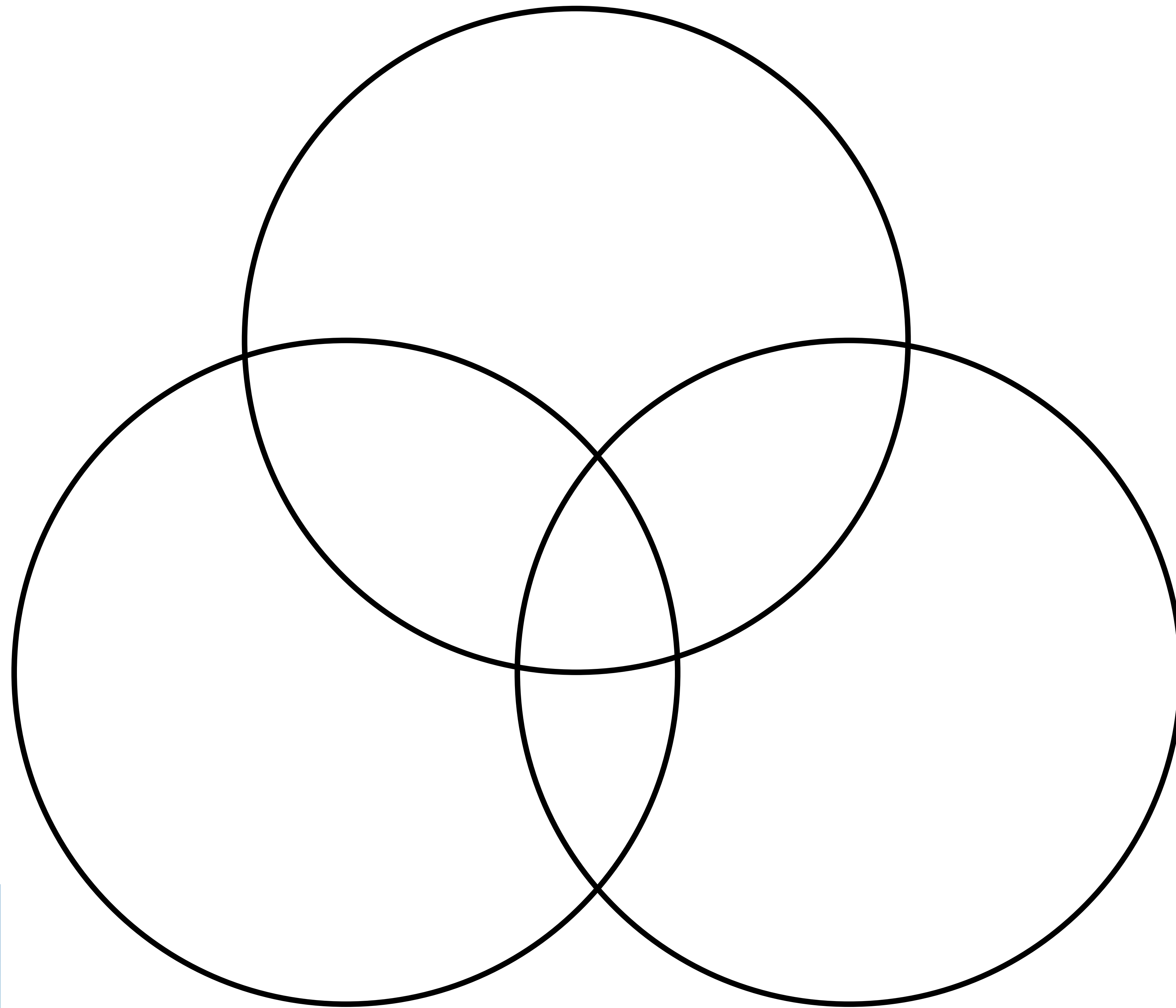
School

Power Plant





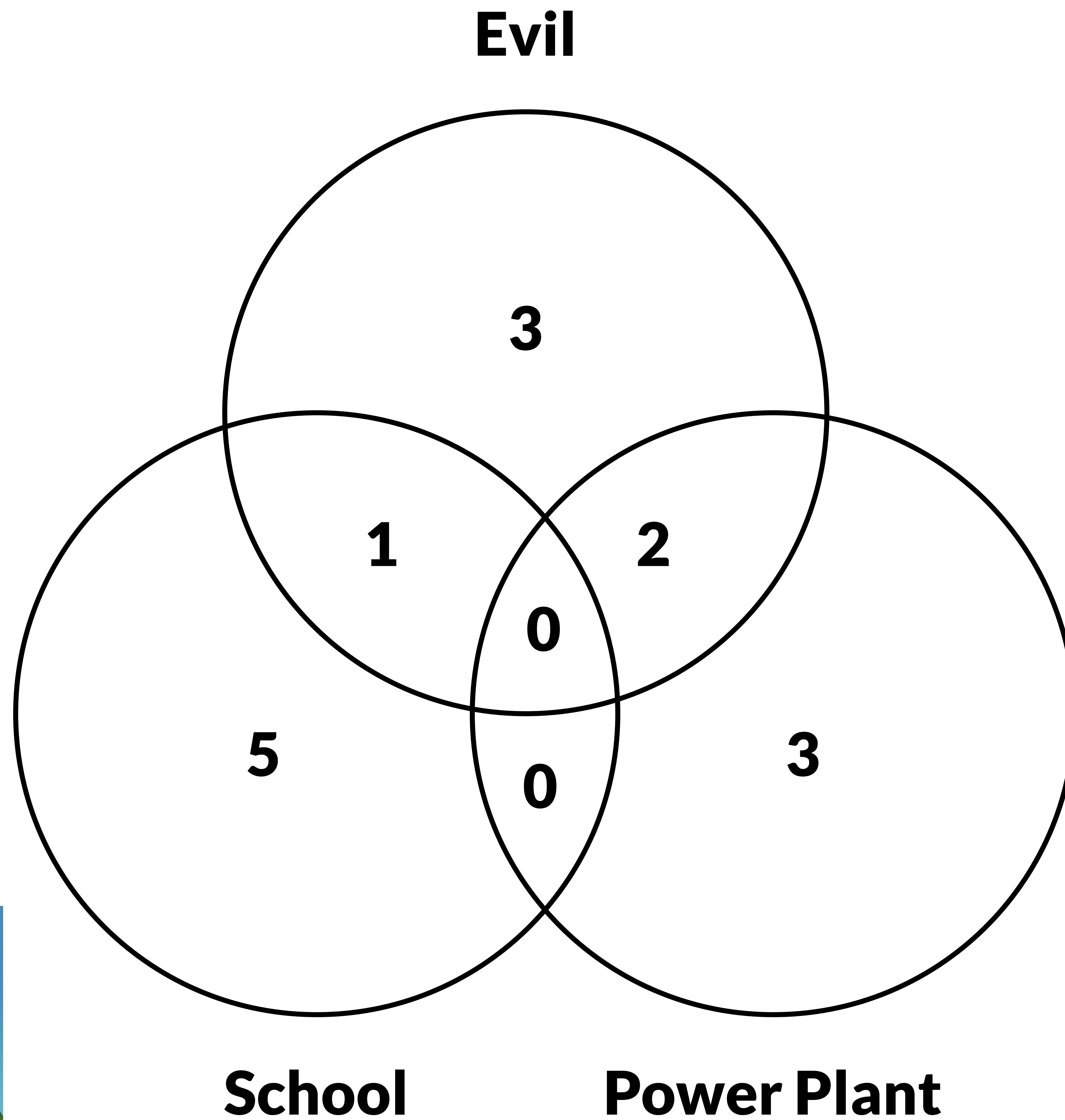
Evil



School

Power Plant

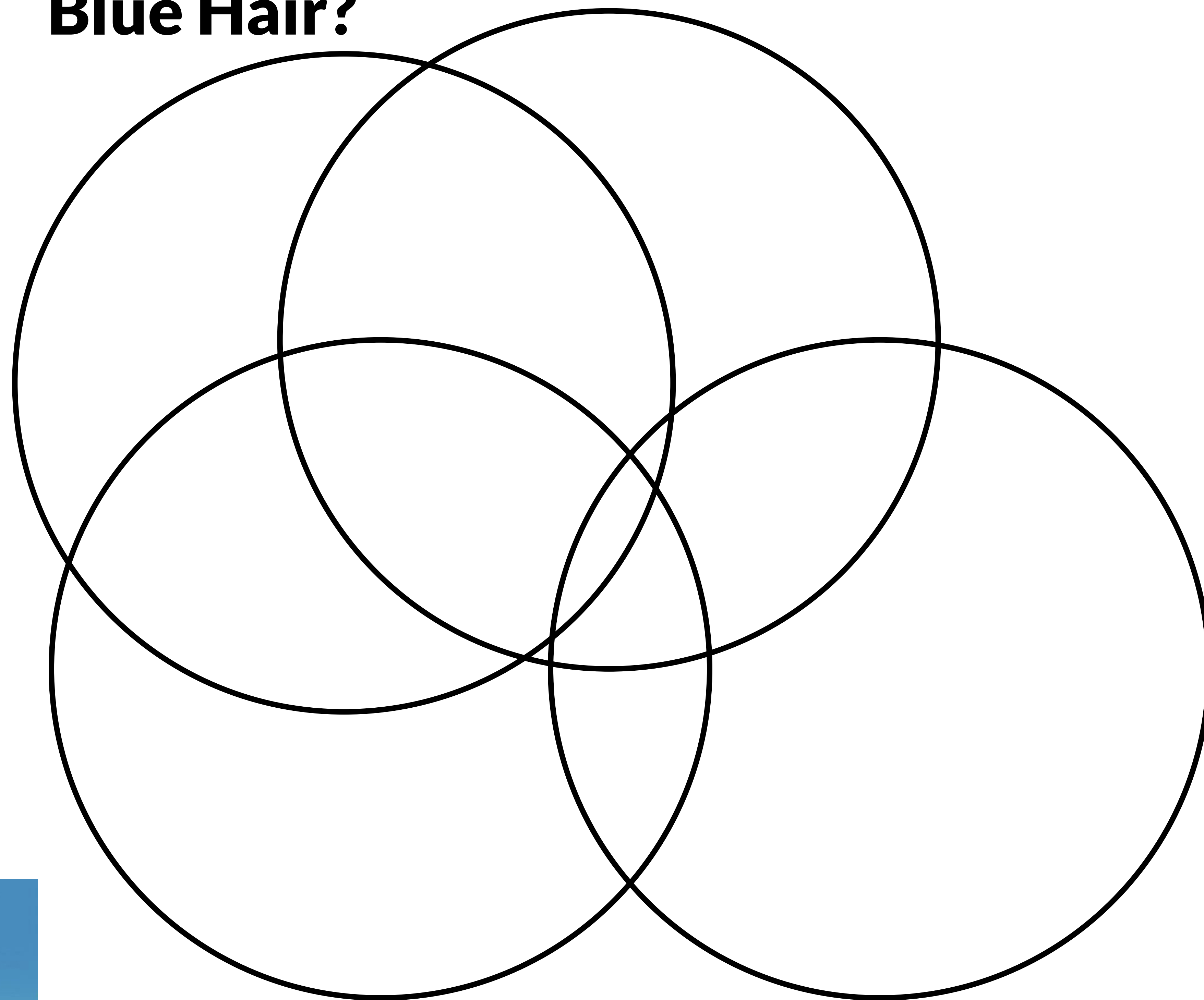






Blue Hair?

Evil



School

Power Plant

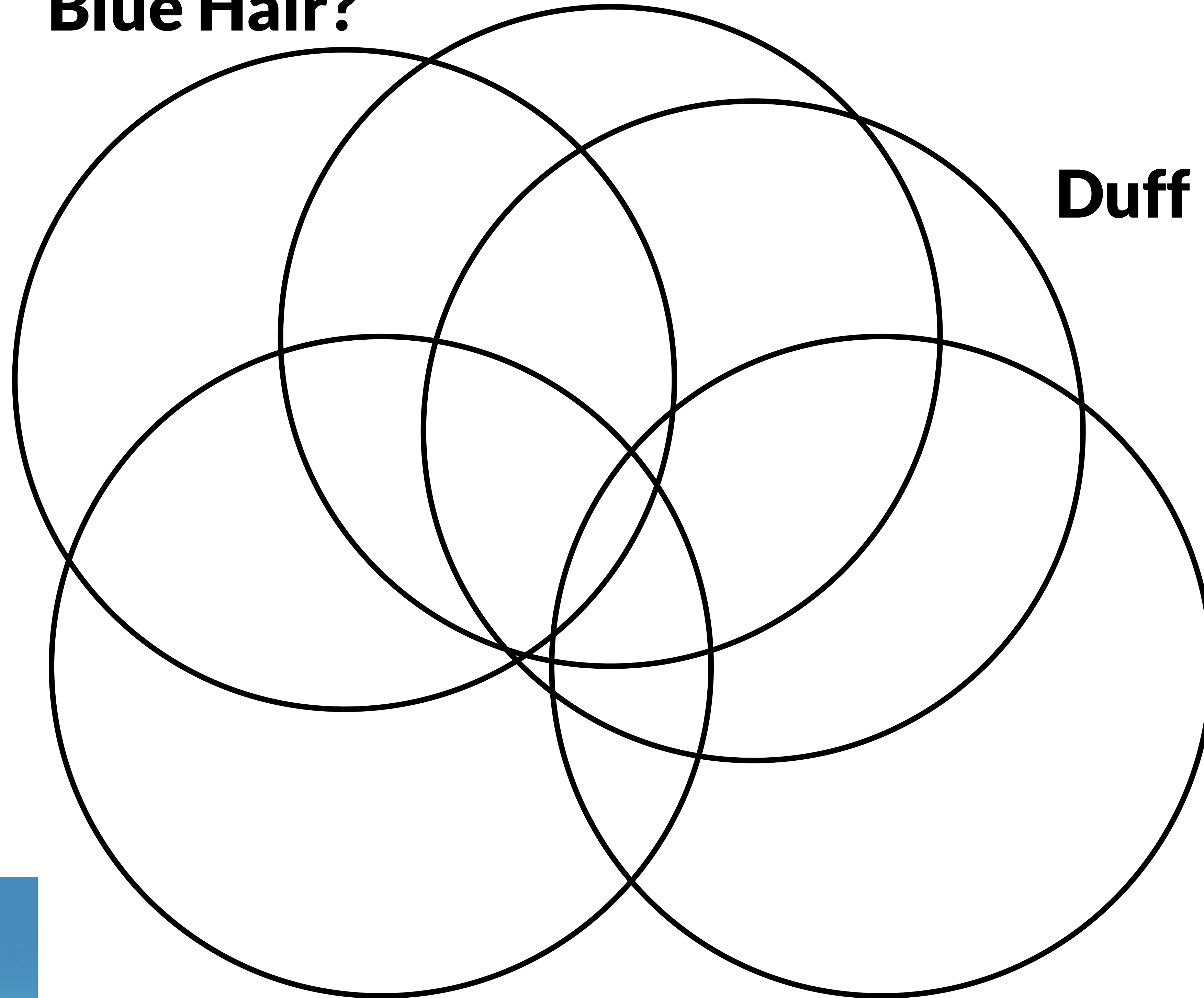




Blue Hair?

Evil

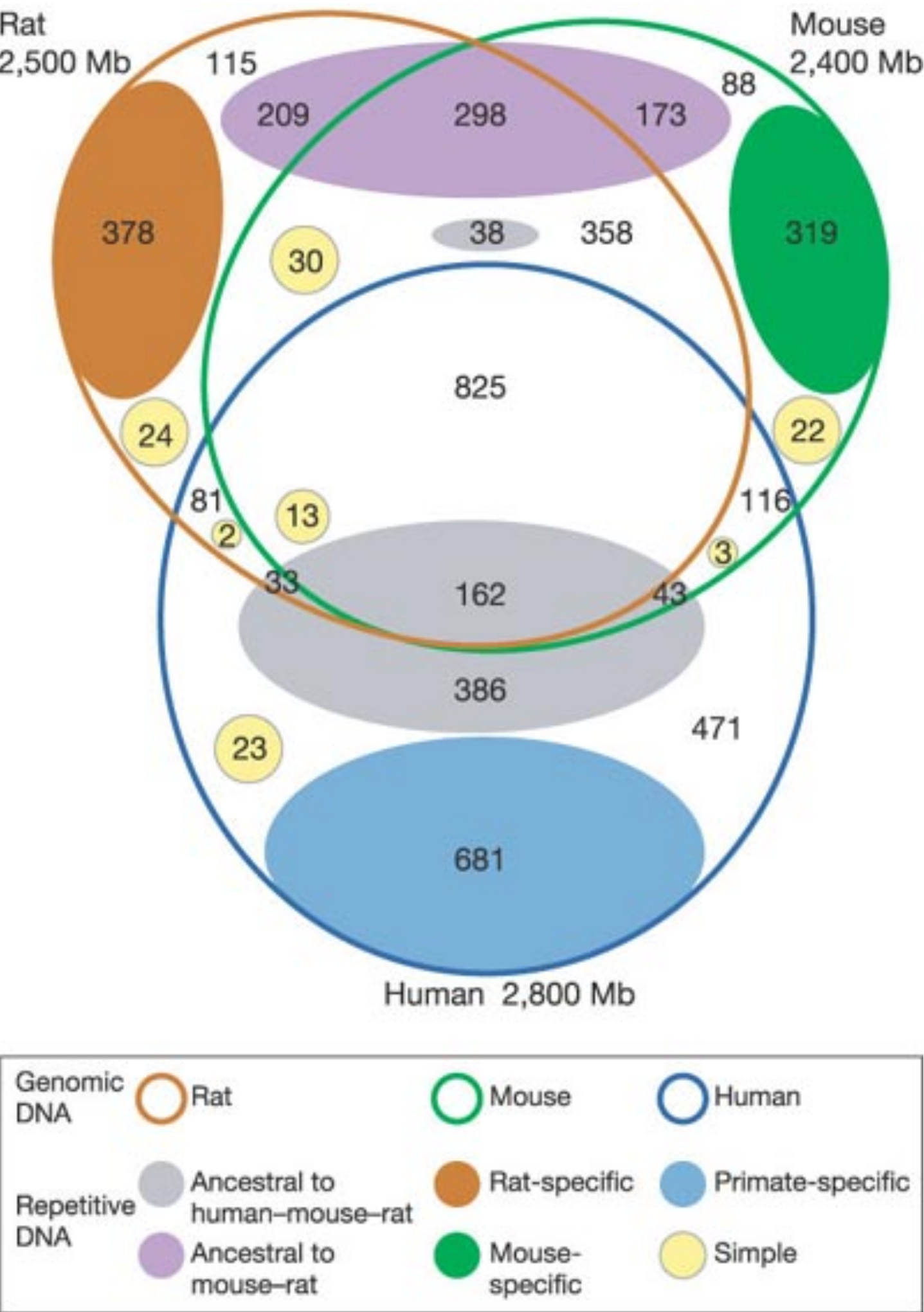
Duff Beer?



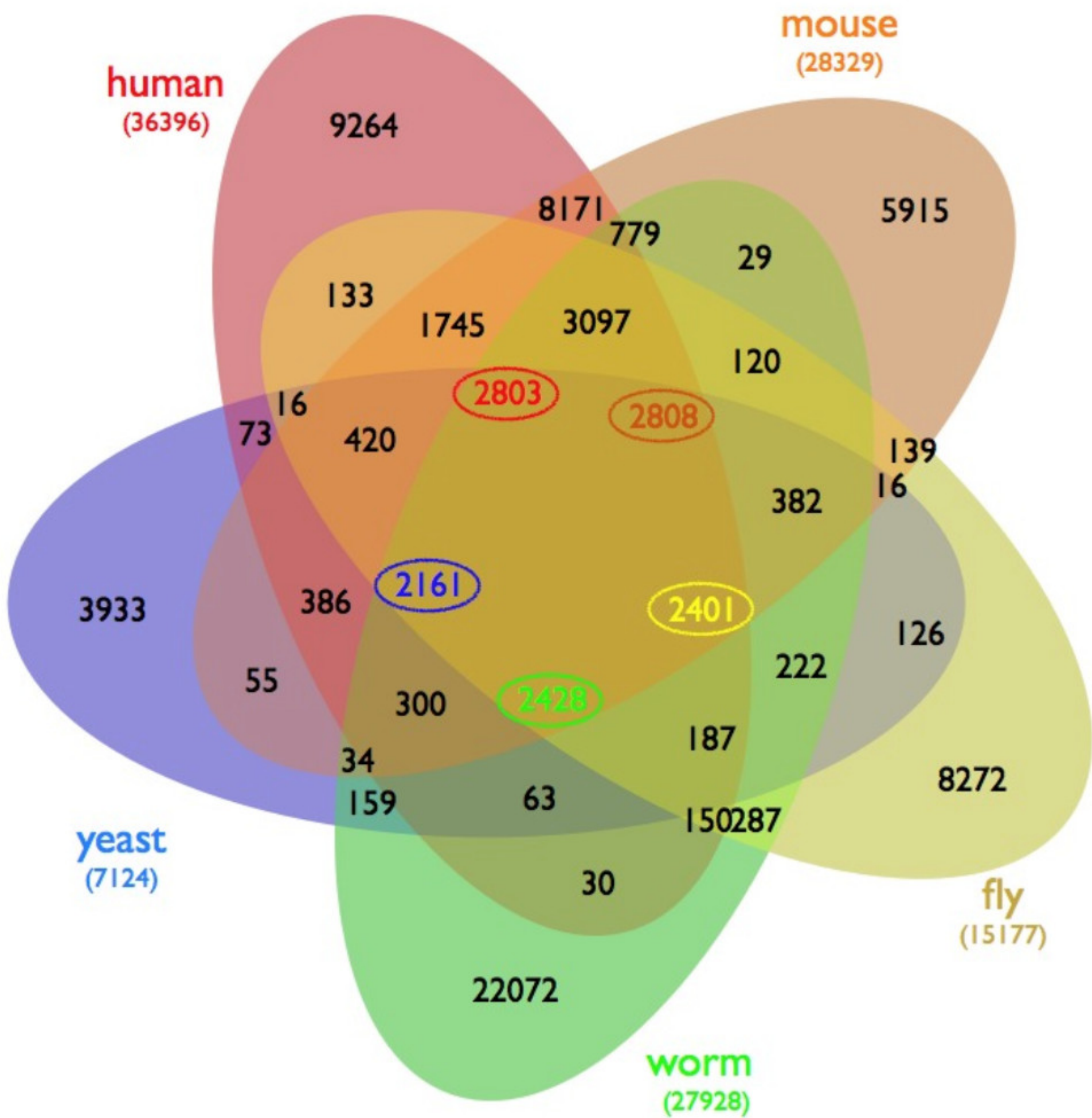
School

Power Plant

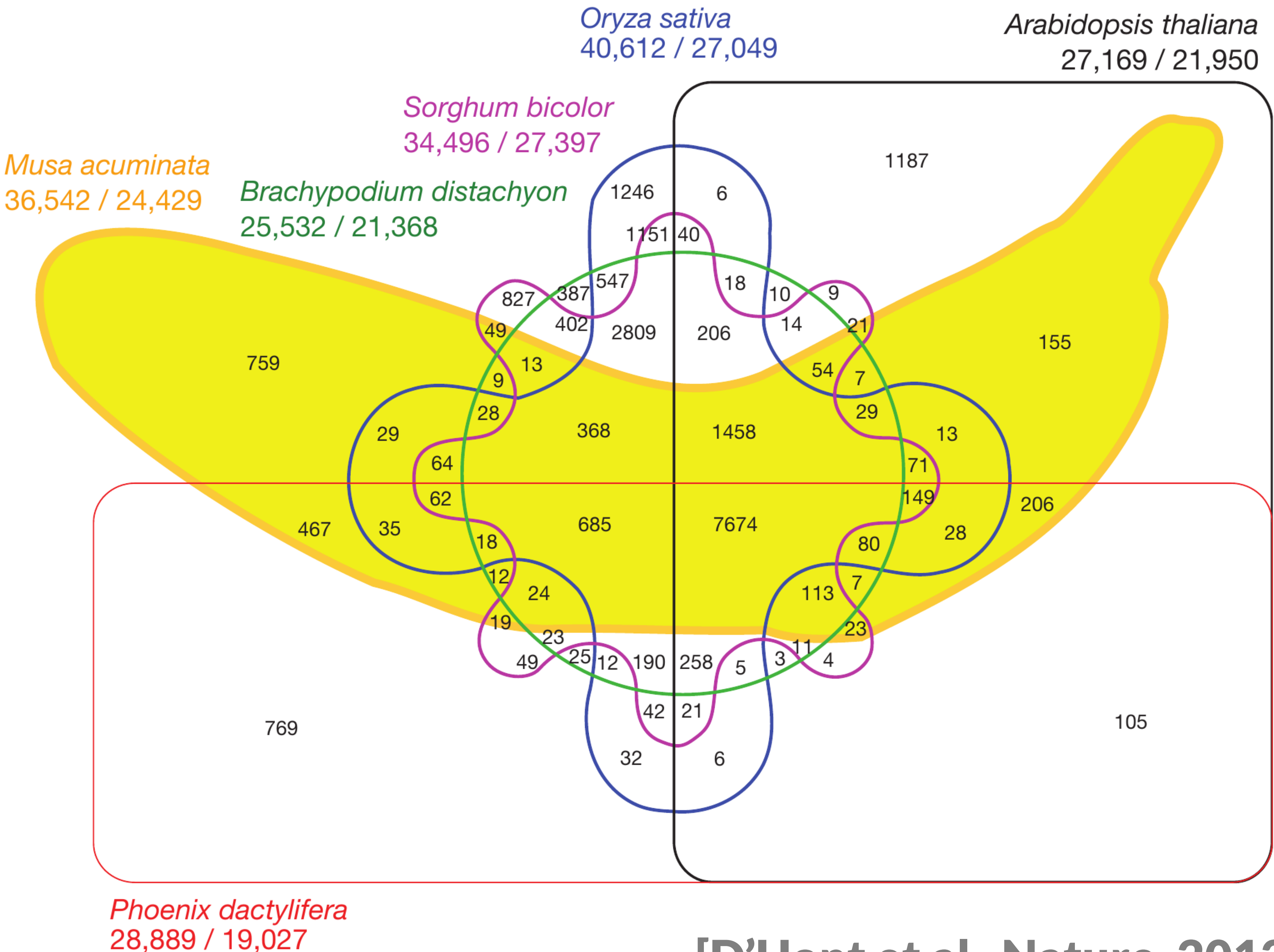




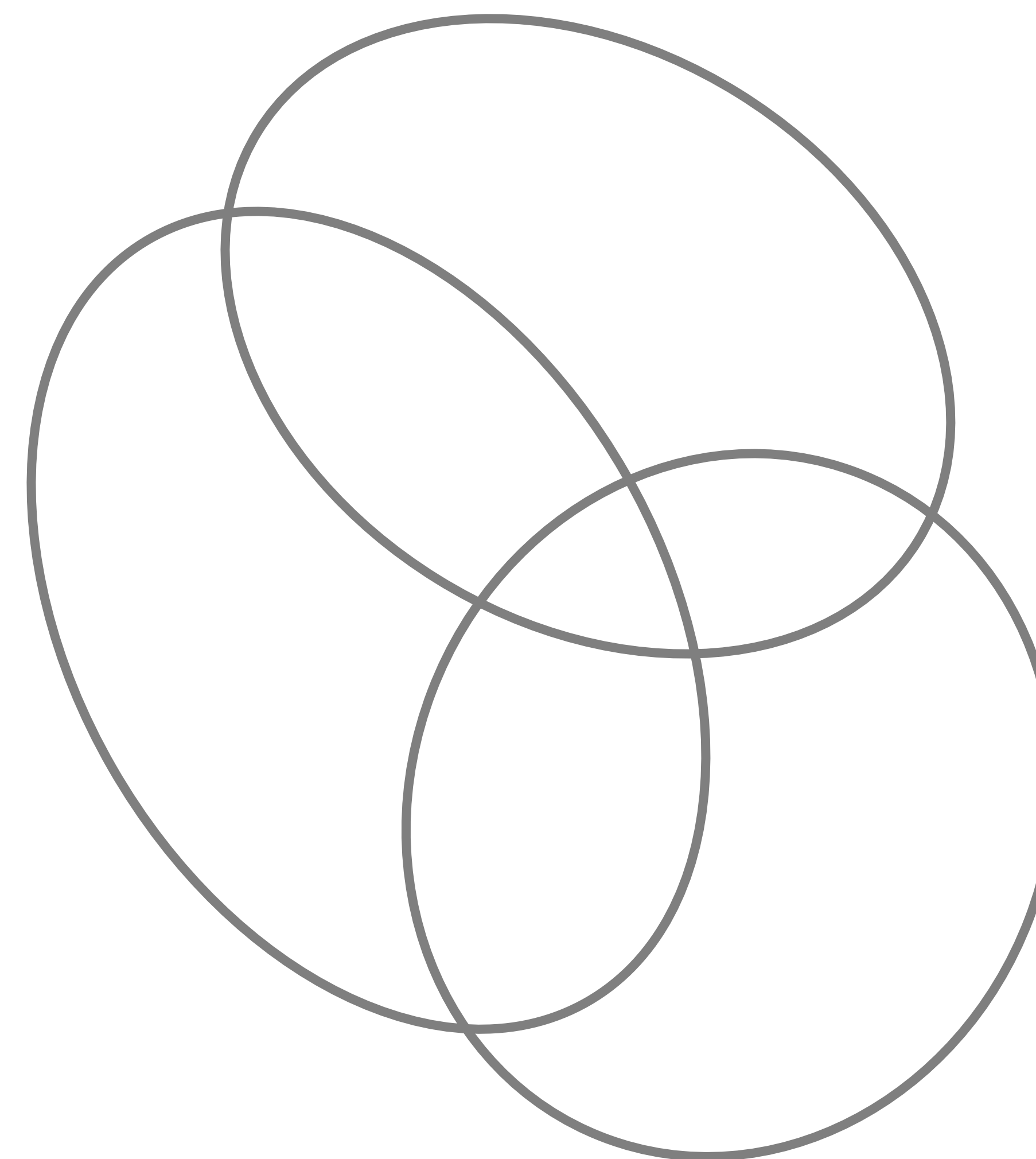
[Gibbs et al., Nature, 2004]



[Wiles et al., BMC Systems Biology]



[D'Hont et al., Nature, 2012]

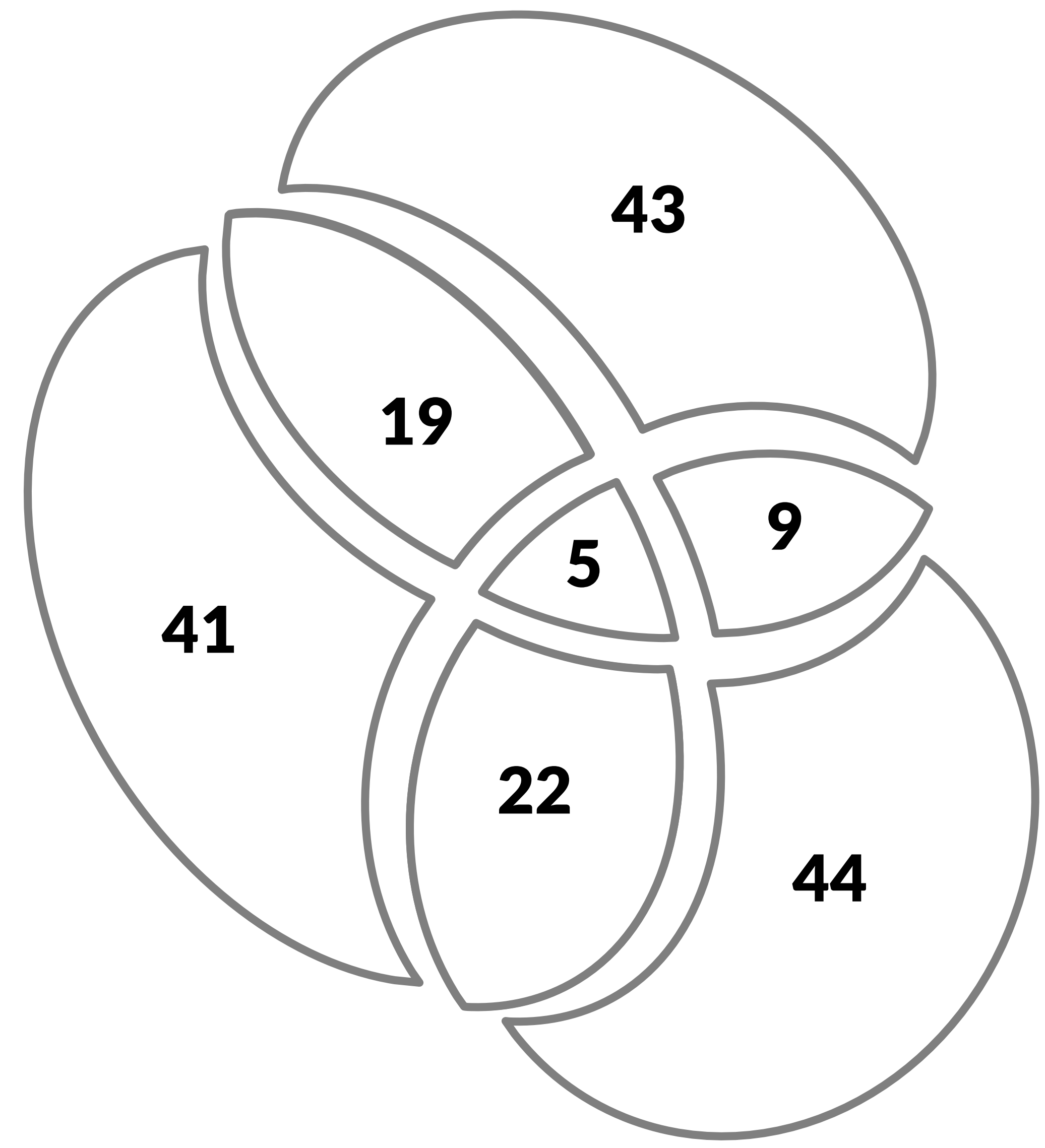


22

?

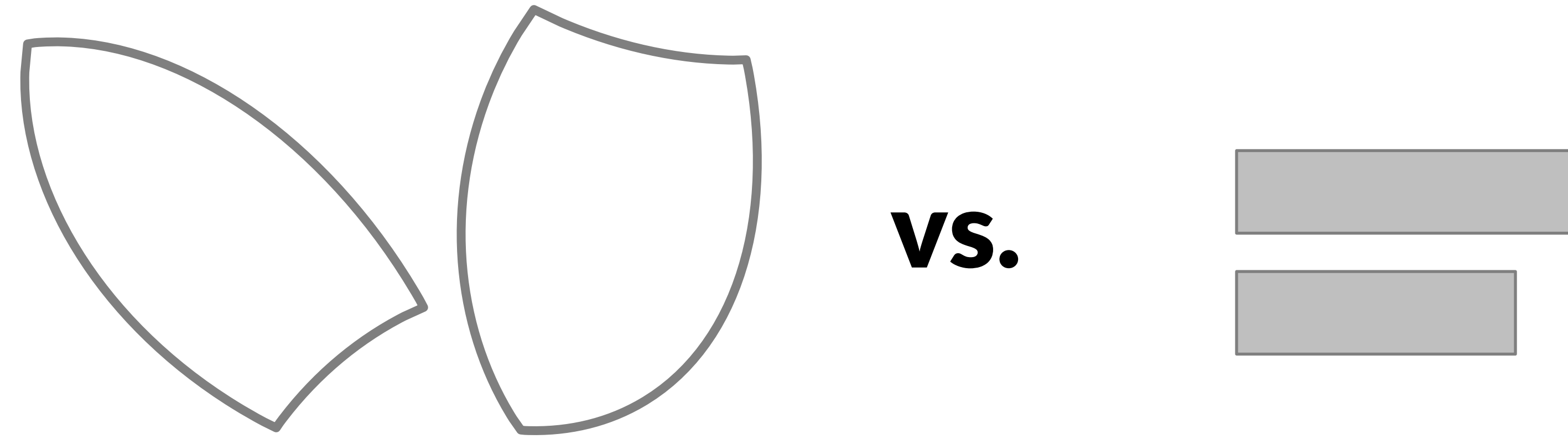
^v • ^v

19

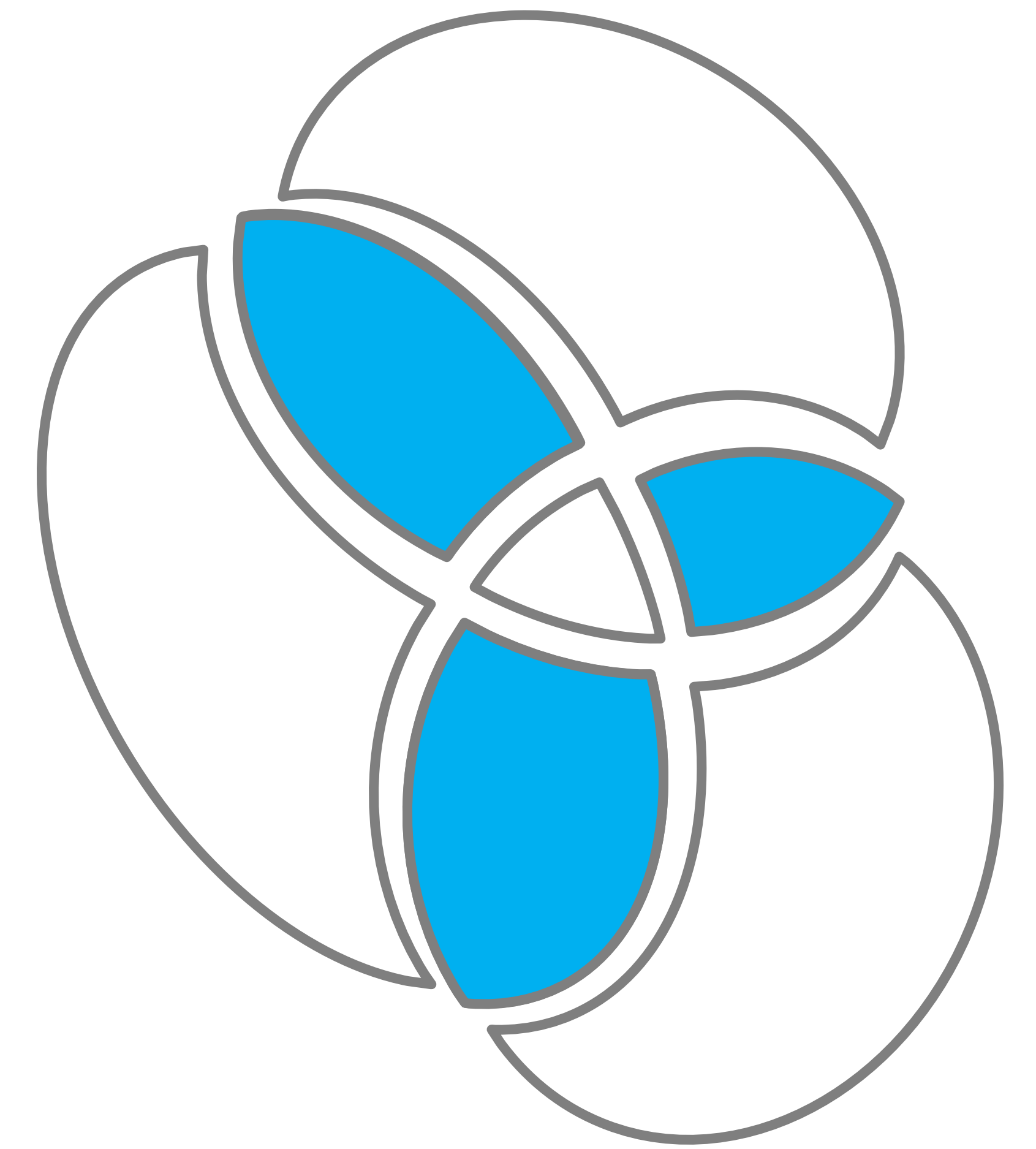


Set Vis Goals

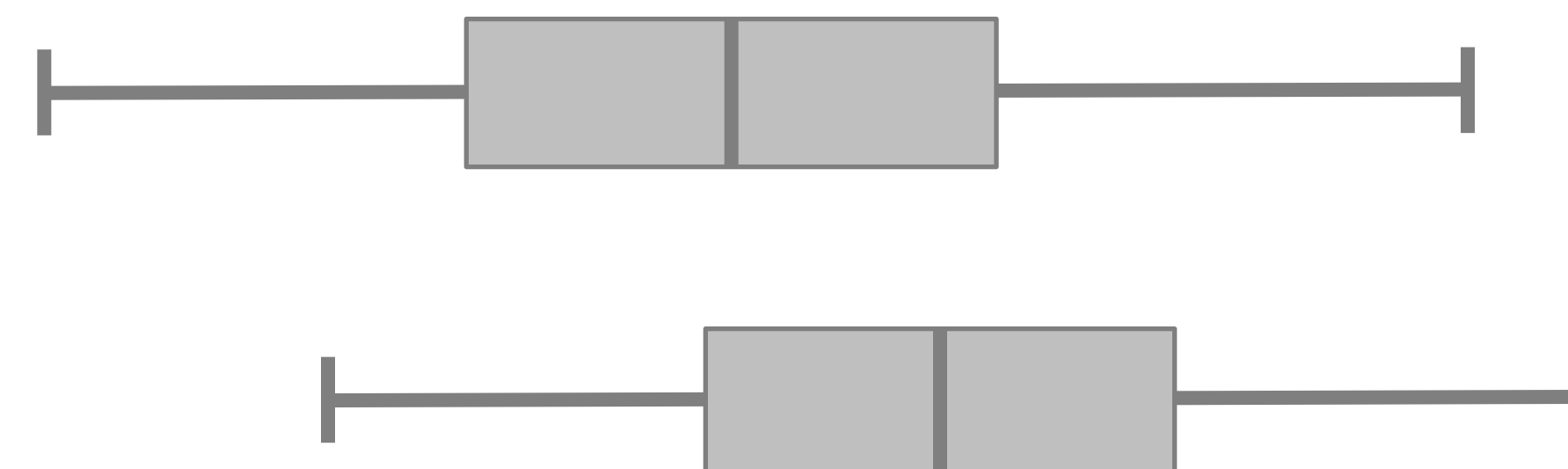
1. Efficient visual encoding



2. Creating complex slices of a dataset



3. Visualize attributes



[Movie Lens Dataset]

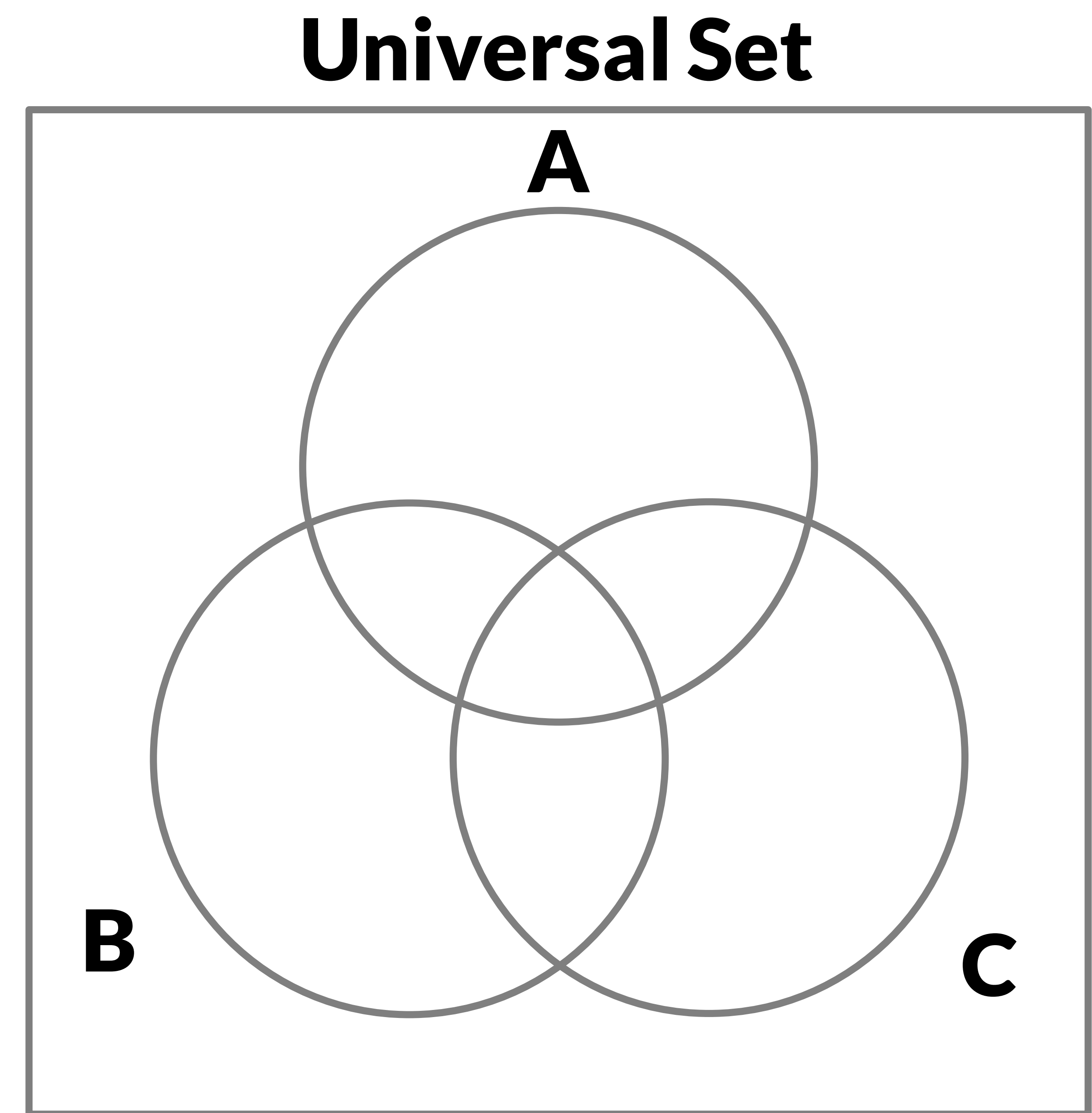
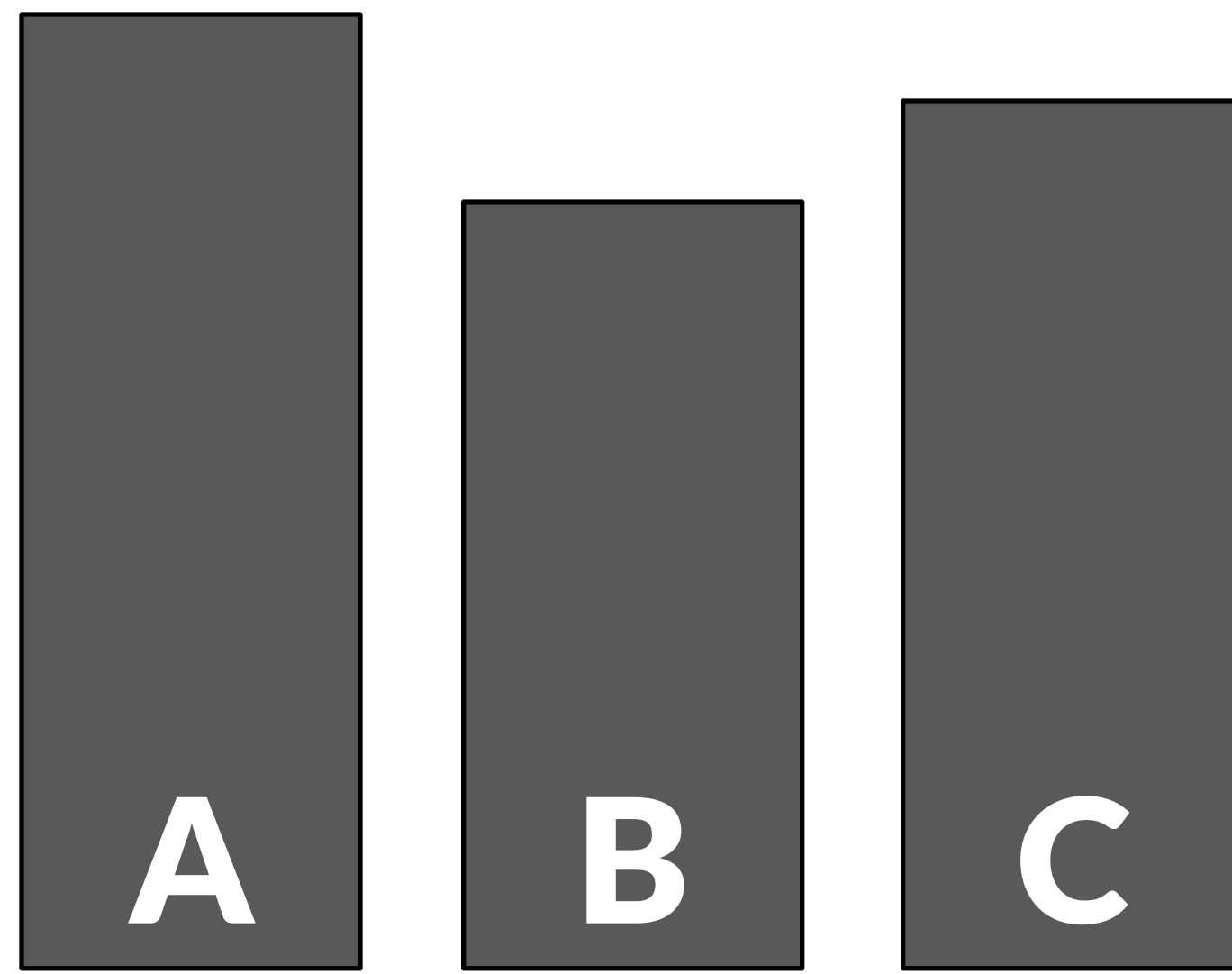
Attribute Details

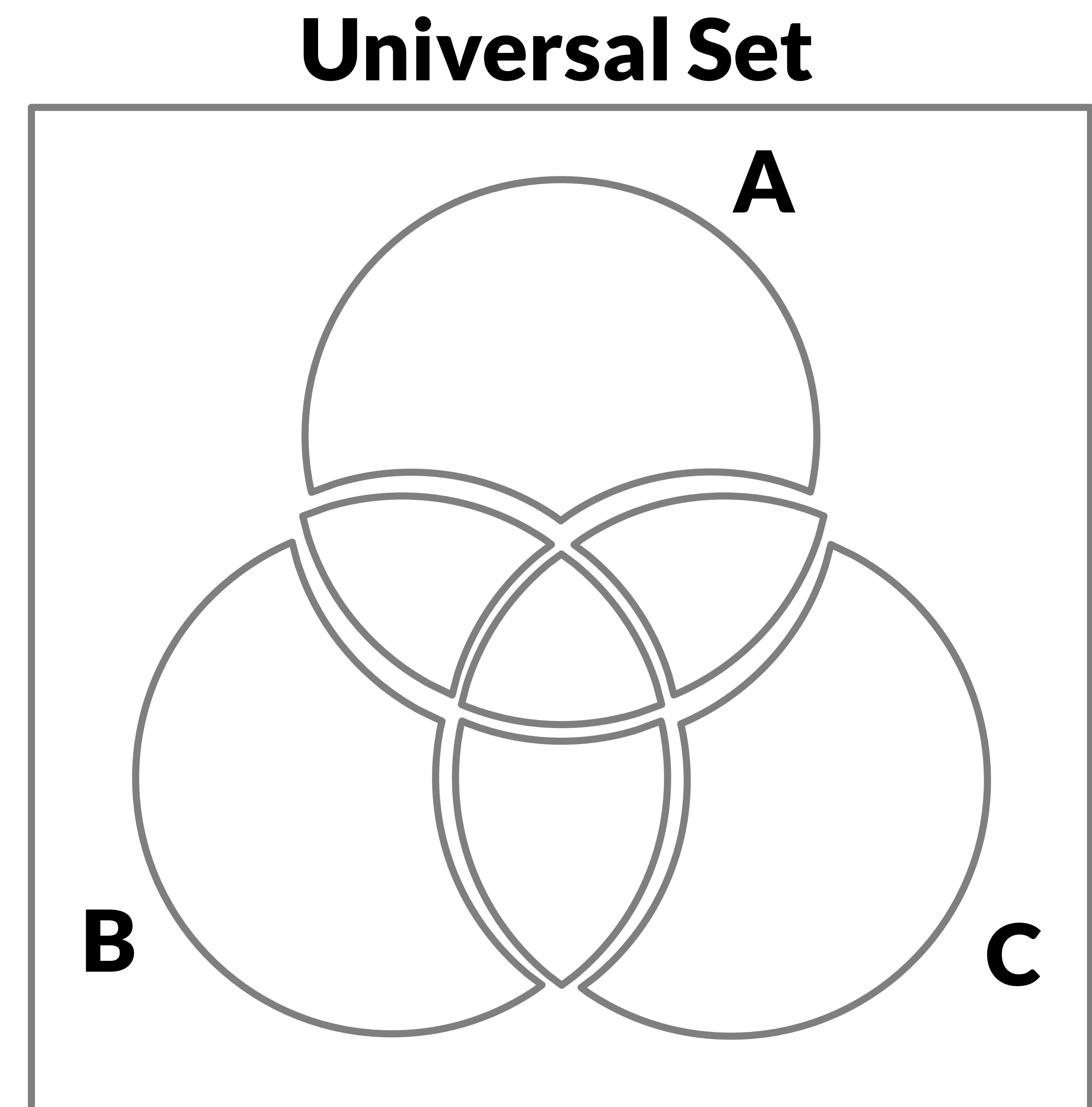
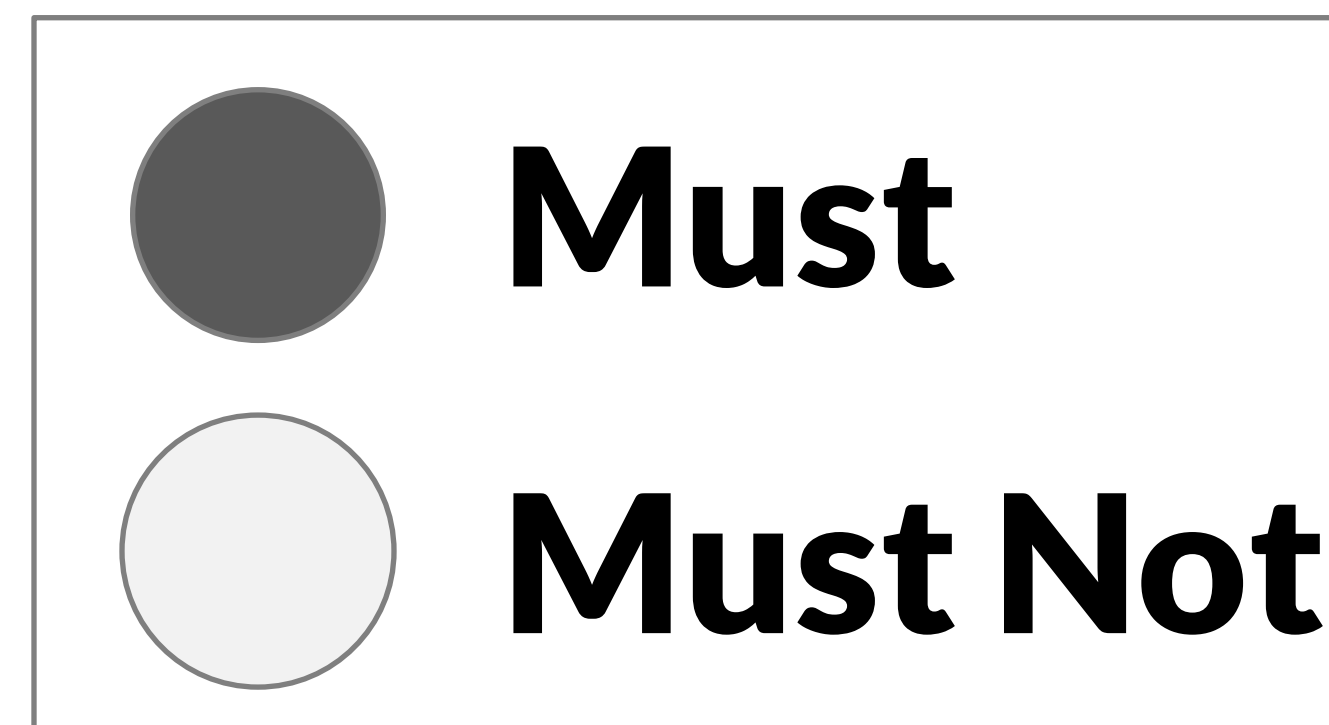
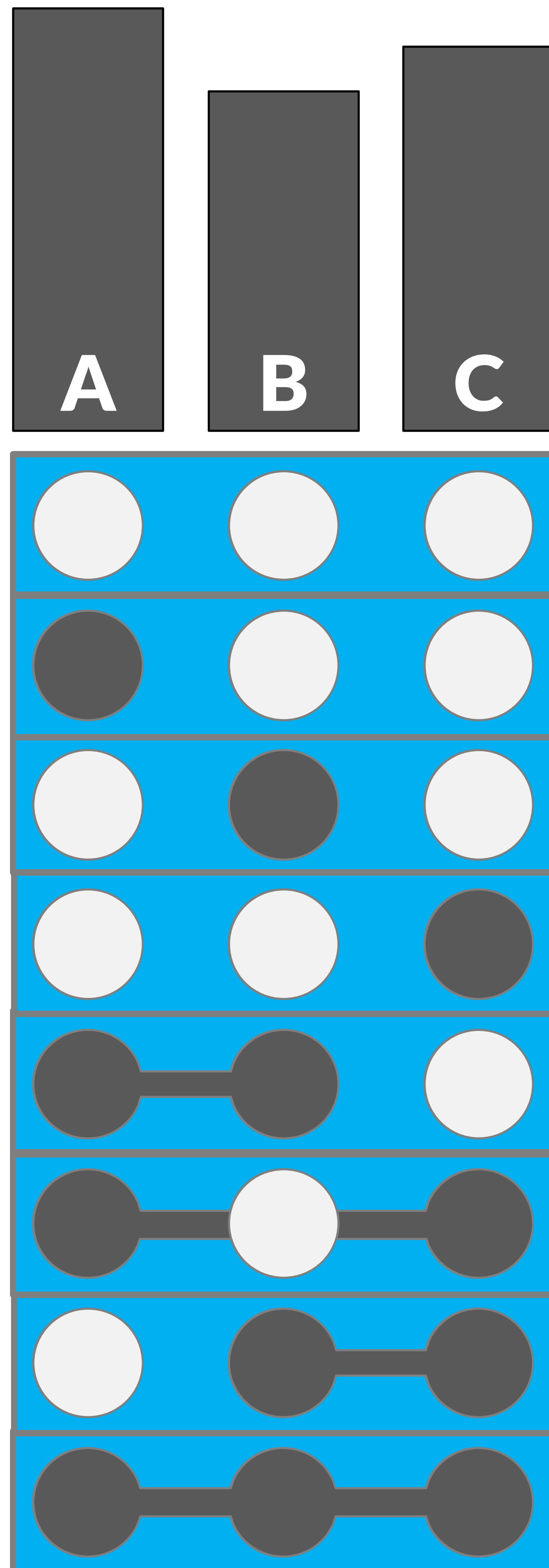
Visualizing Intersections

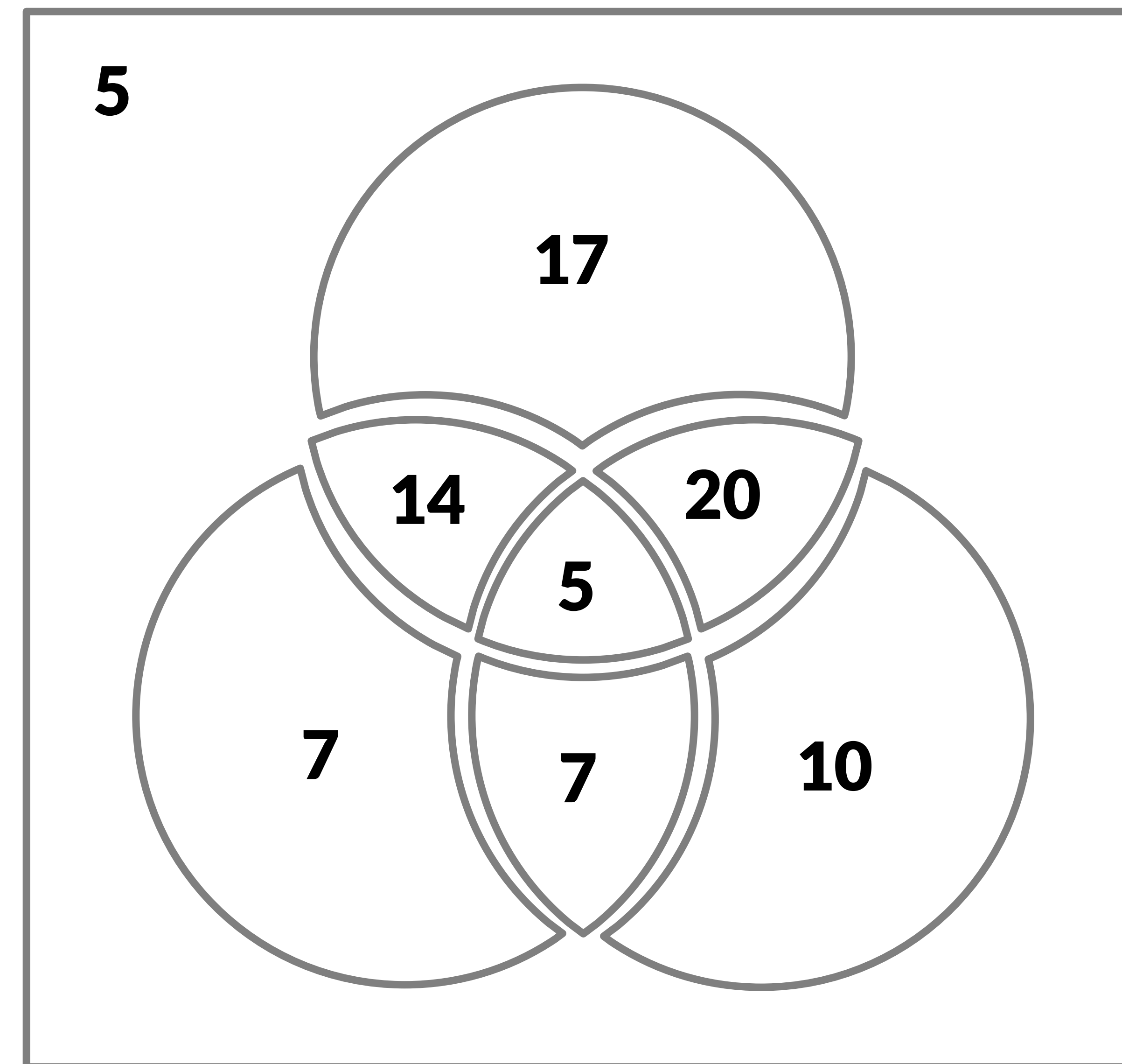
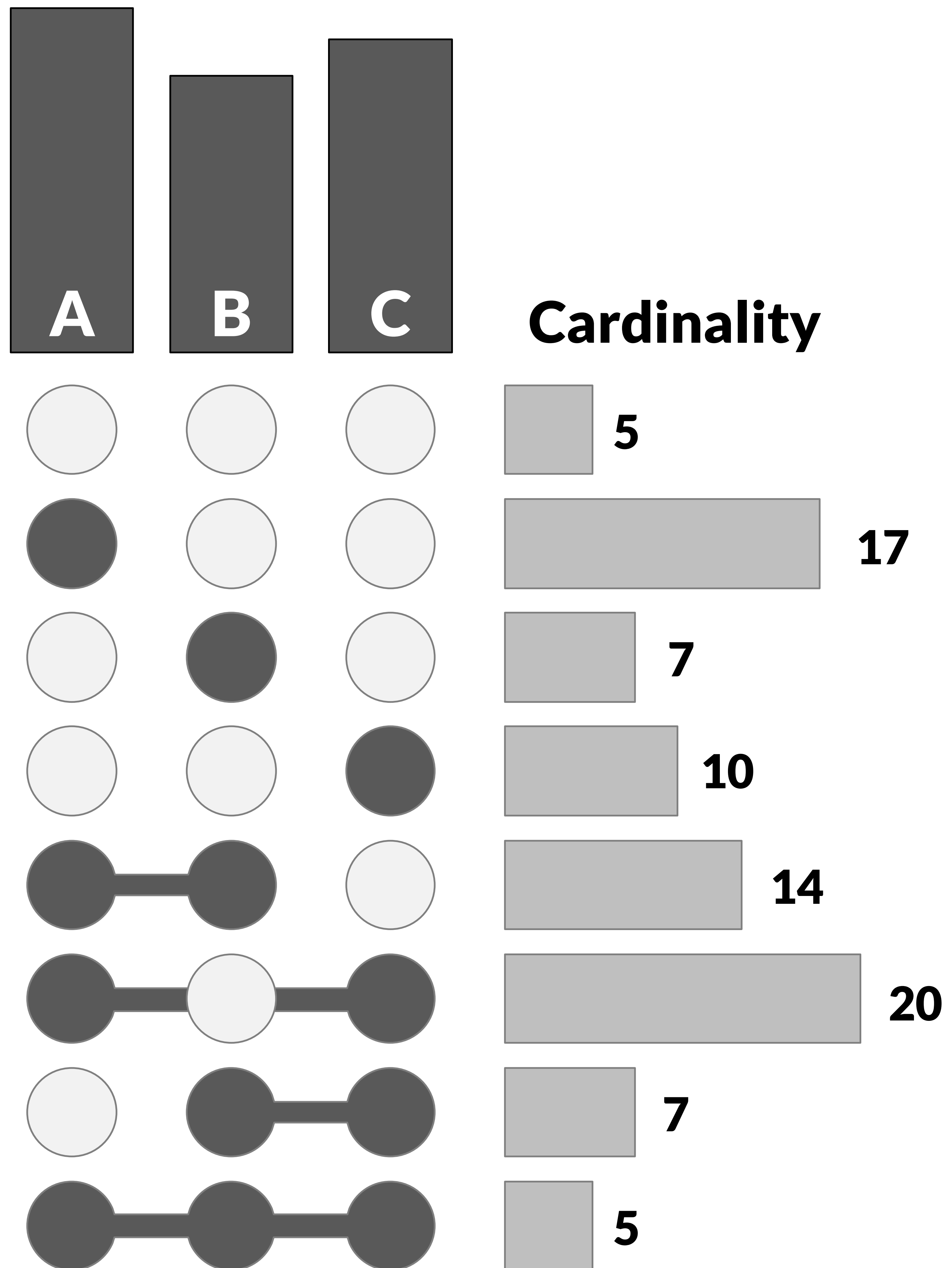
Visualizing Properties

Element List & Queries

Visualizing Intersections

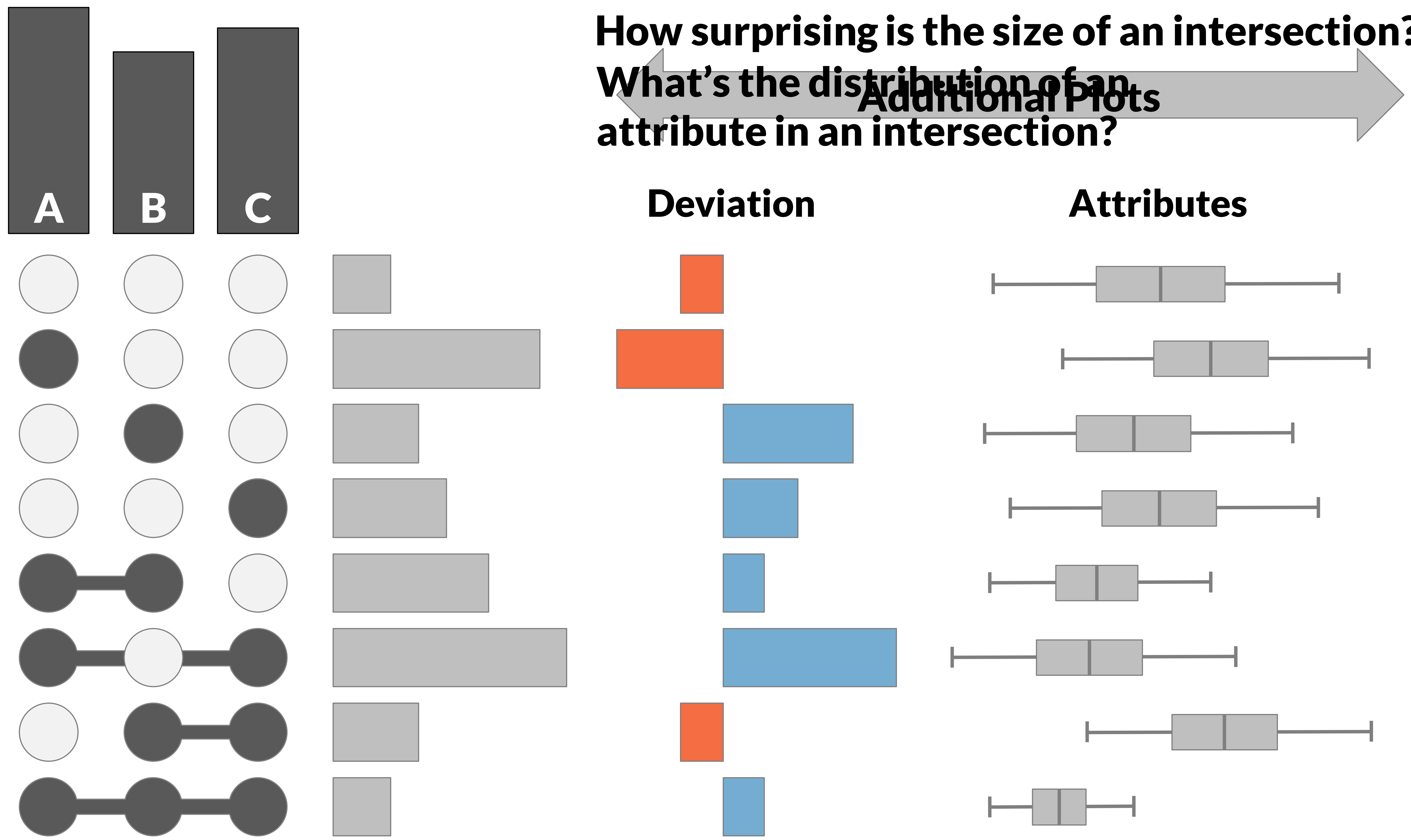






Plotting Attributes

How surprising is the size of an intersection?
 What's the distribution of an attribute in an intersection?



First, aggregate by
Don't Aggregate ▾

Then, aggregate by
Don't Aggregate ▾

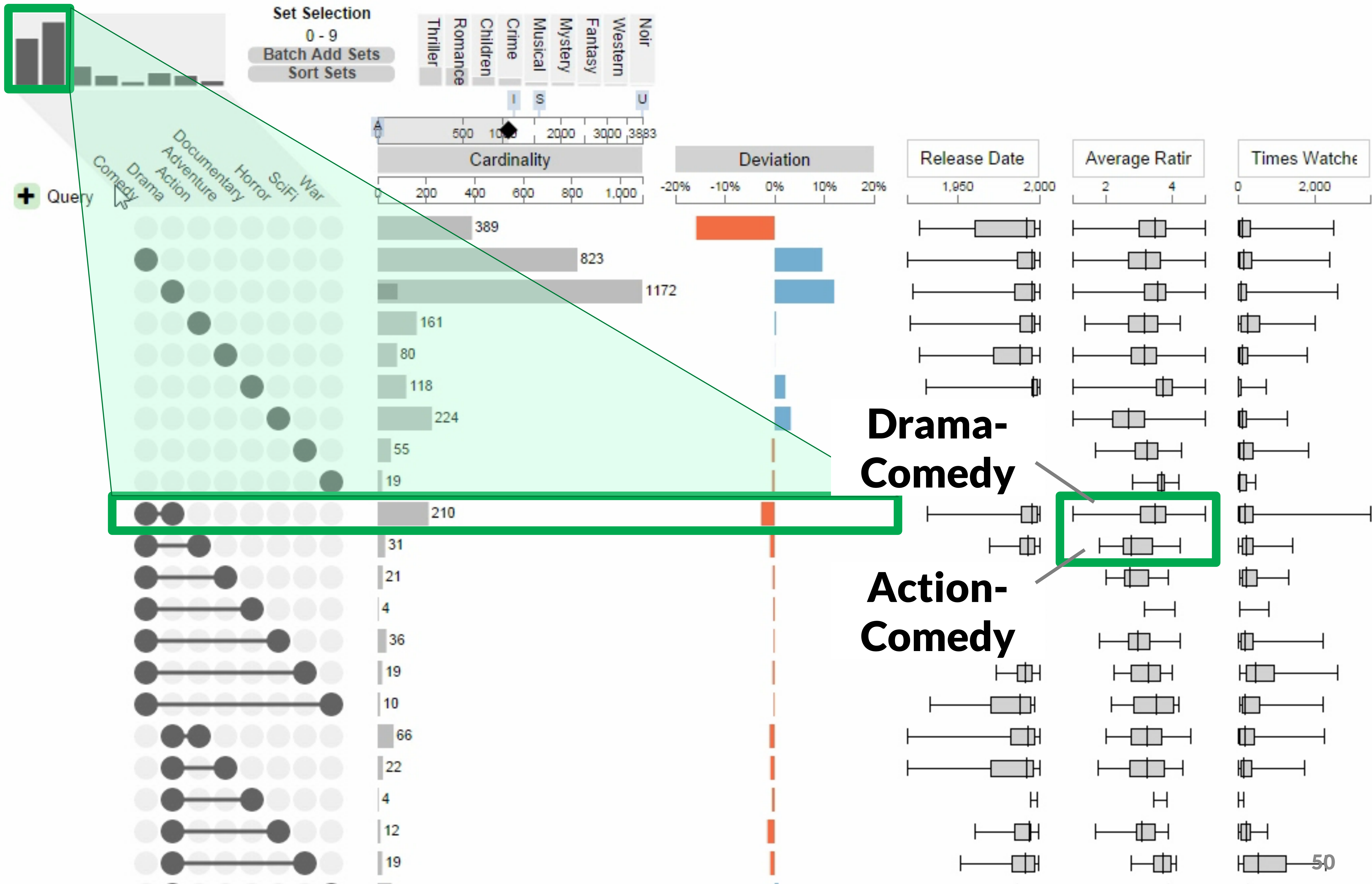
Sort by
☒ Degree
☐ Cardinality
☐ Deviation

Aggregates
Collapse All
Expand All

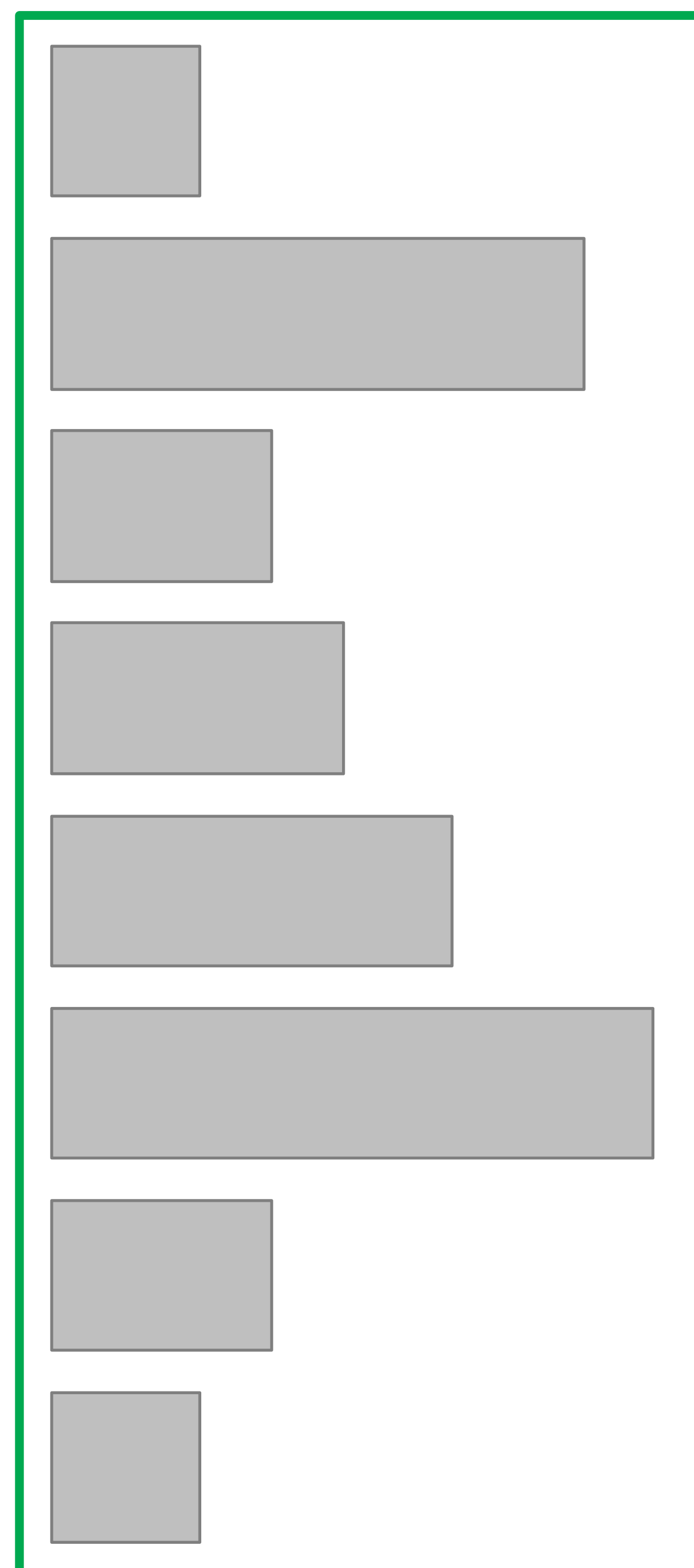
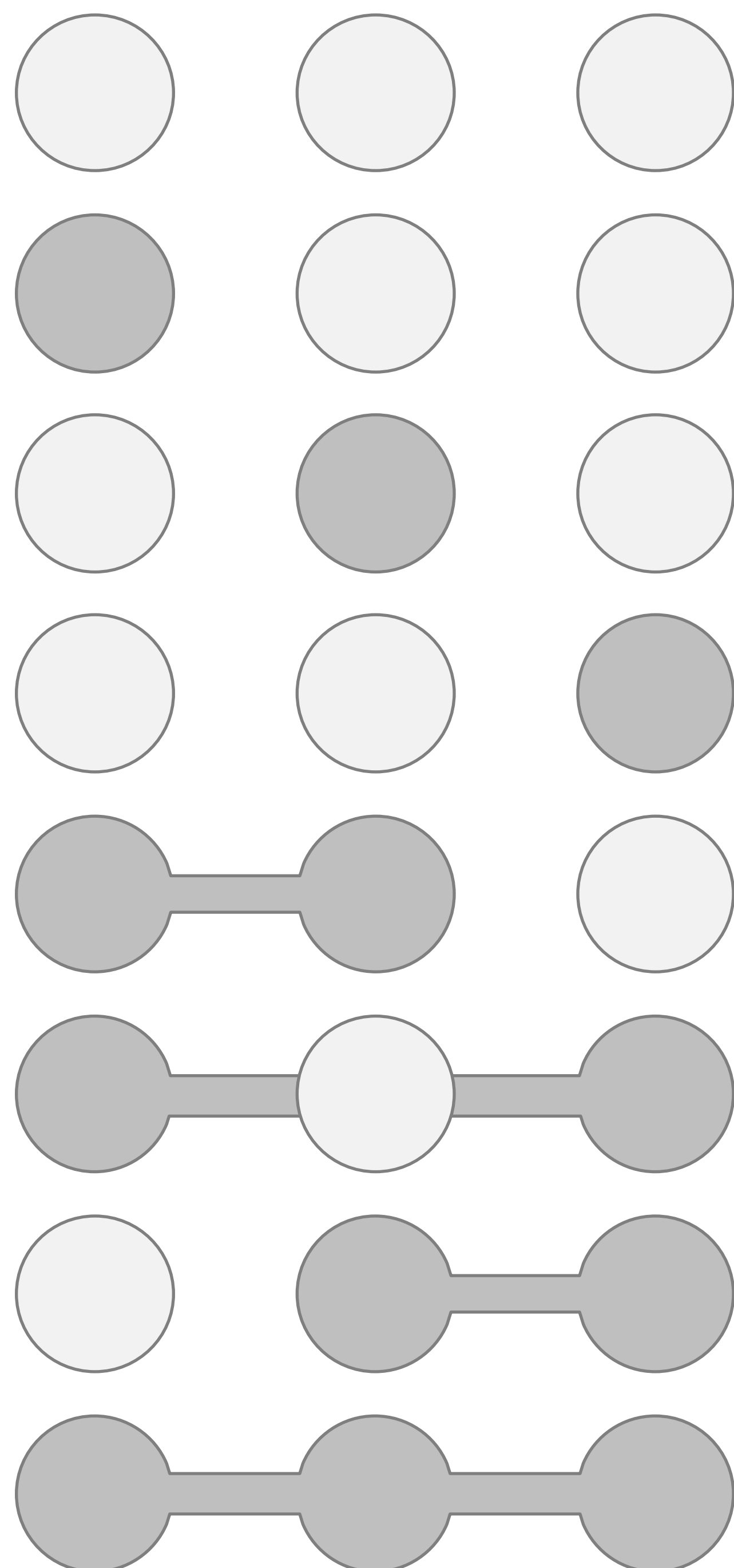
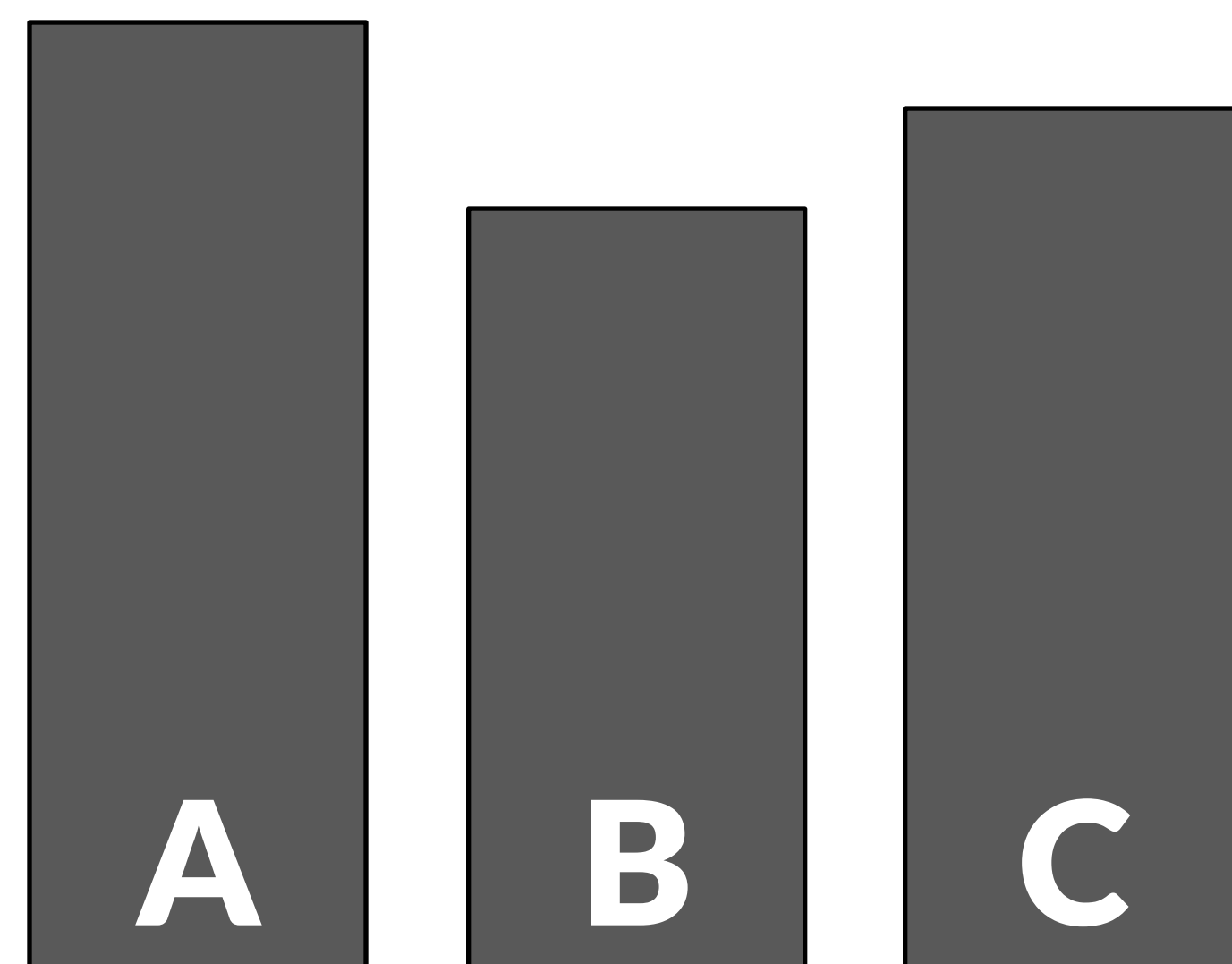
Row Height
Large ▾

Data
Min Degree:
0
Max Degree:
5
☐ Hide Empty Intersections

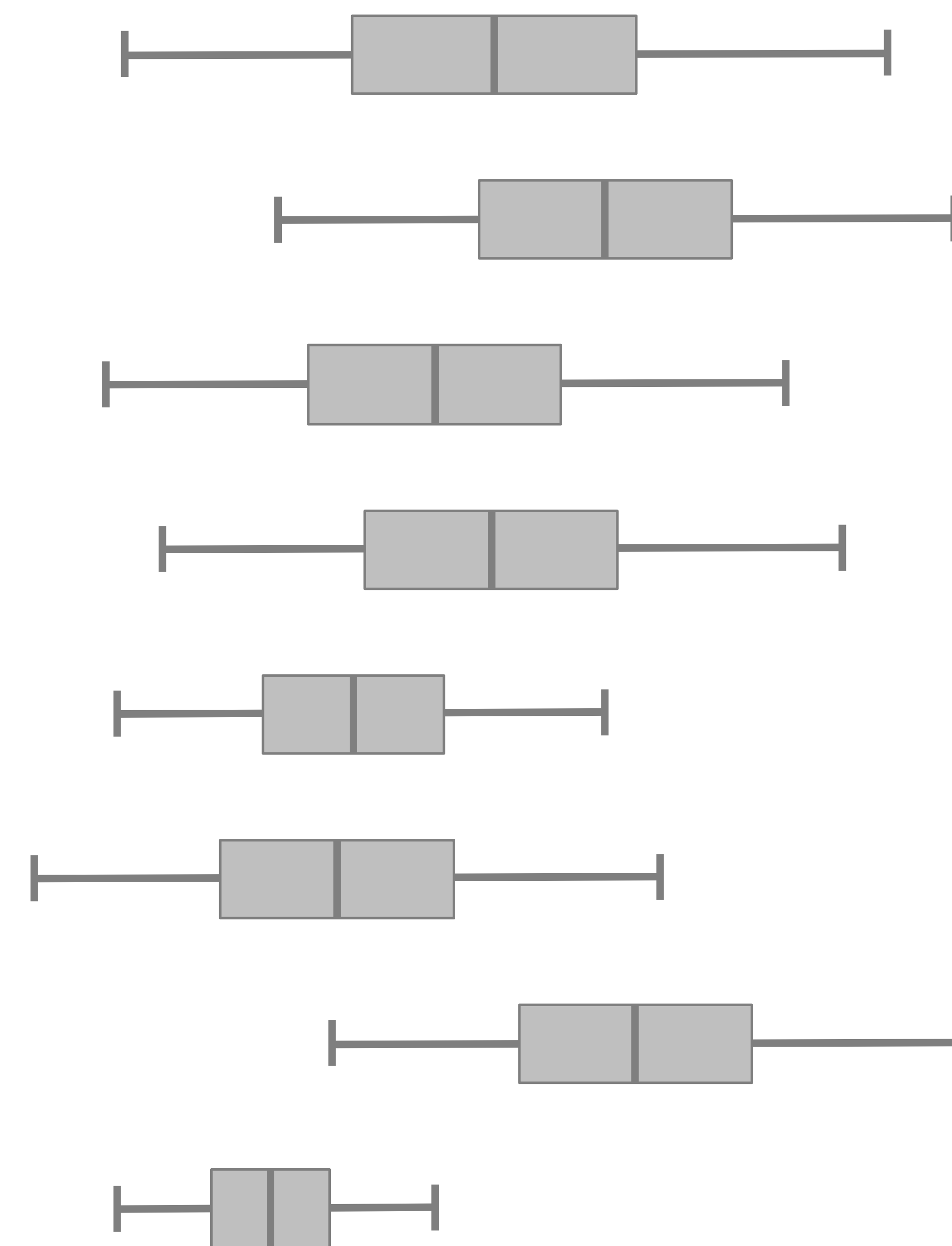
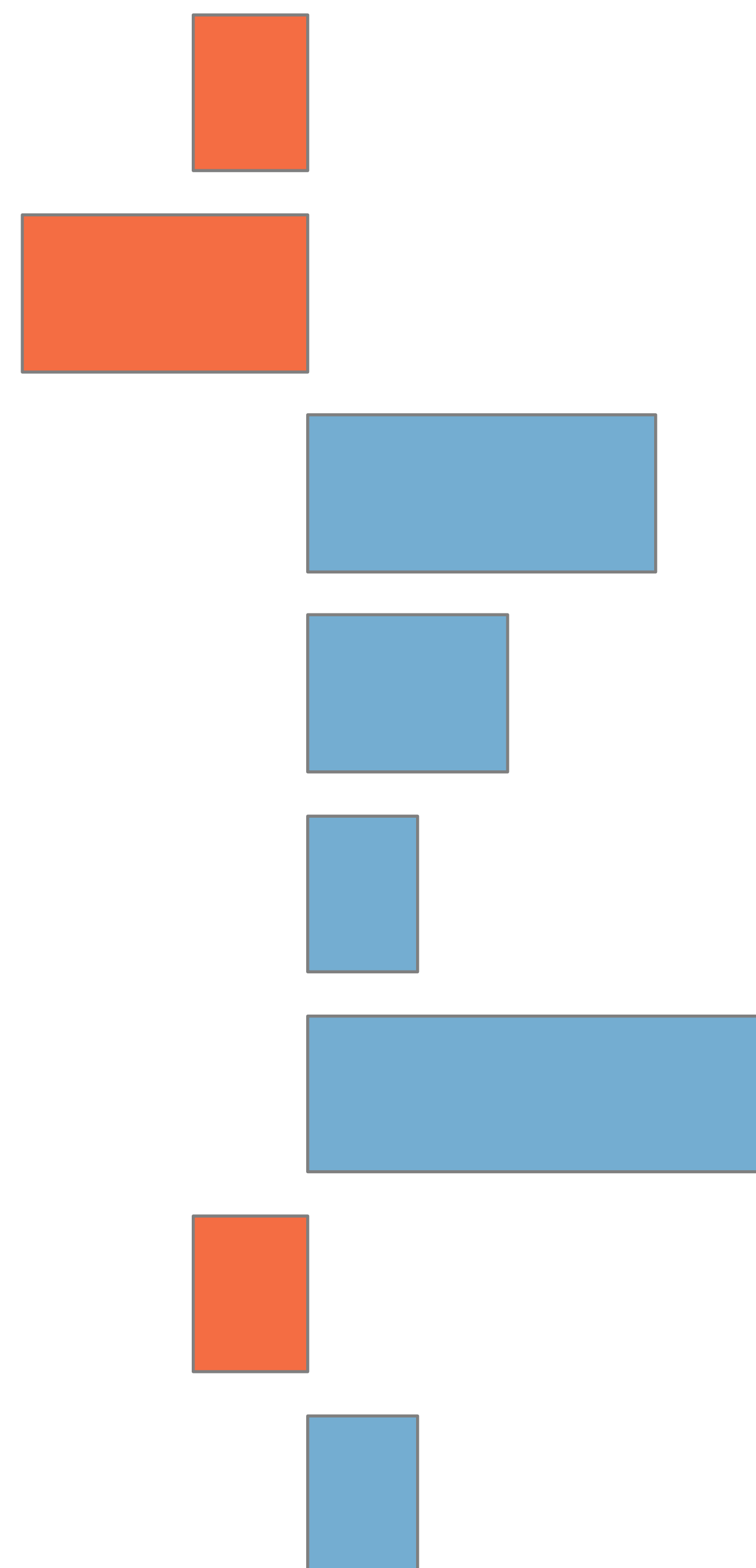
Dataset Information
Name: Movies
Genres
Sets: 17
Attributes: 6
Elements: 3883
Author: grouplens
Description:
MovieLens ratings
dataset, curated and
filtered by Alsallakh.
Source:
<http://grouplens.org/d..>

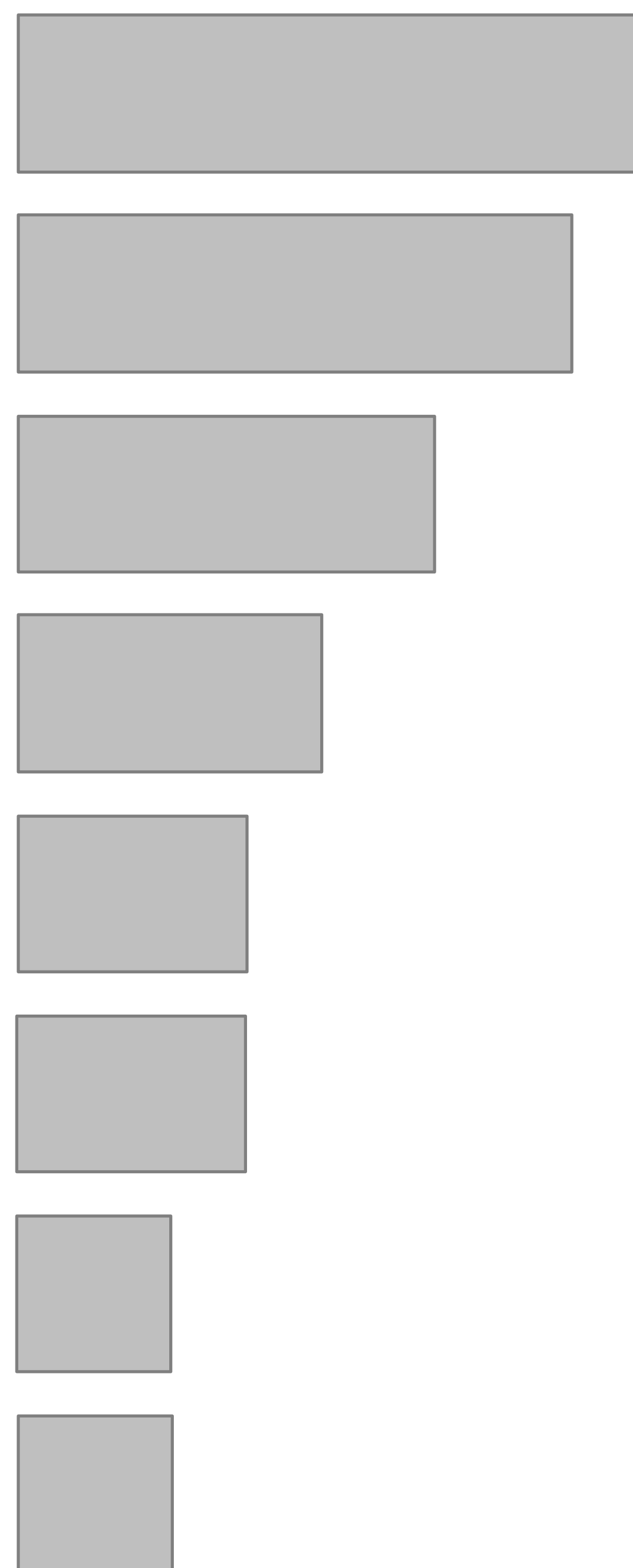
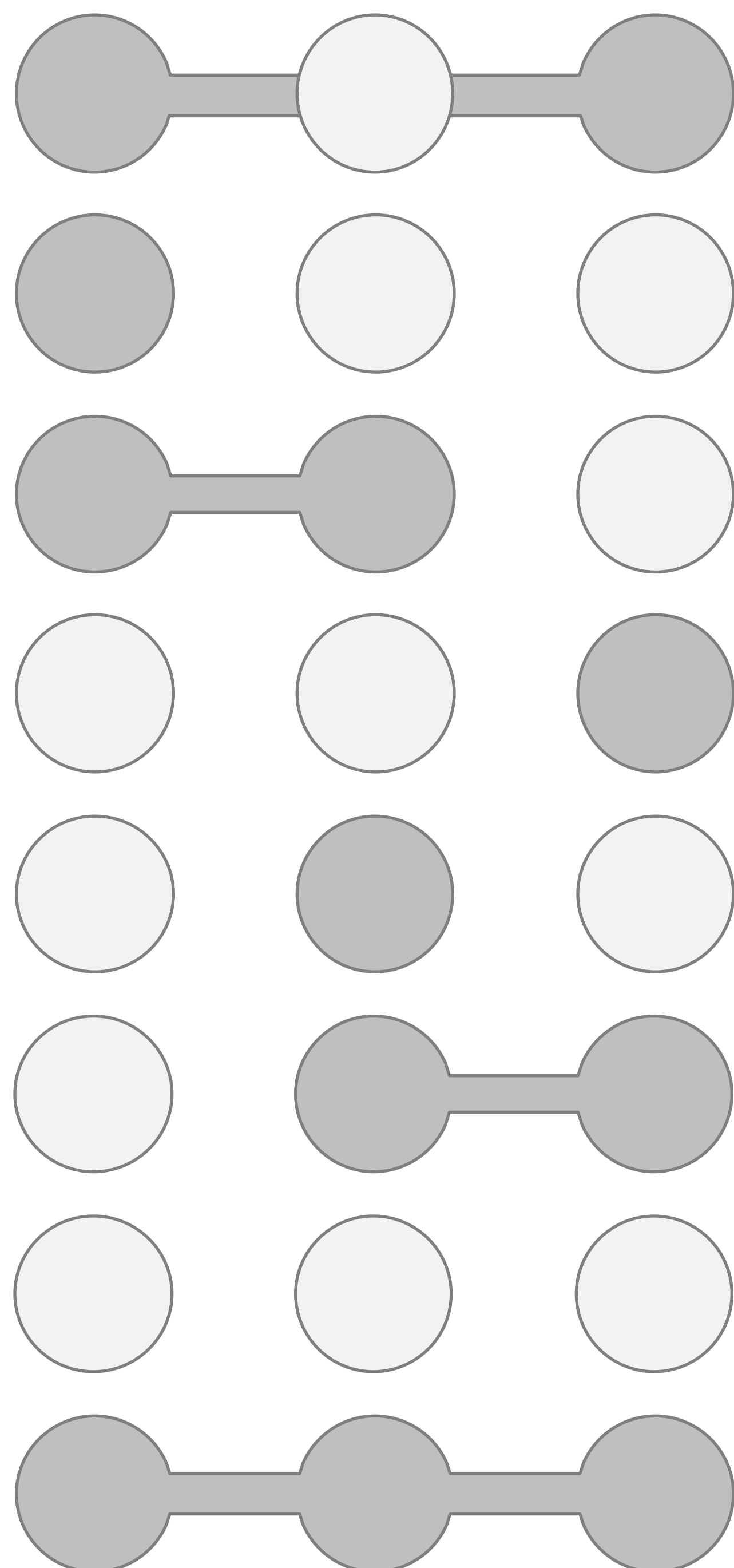
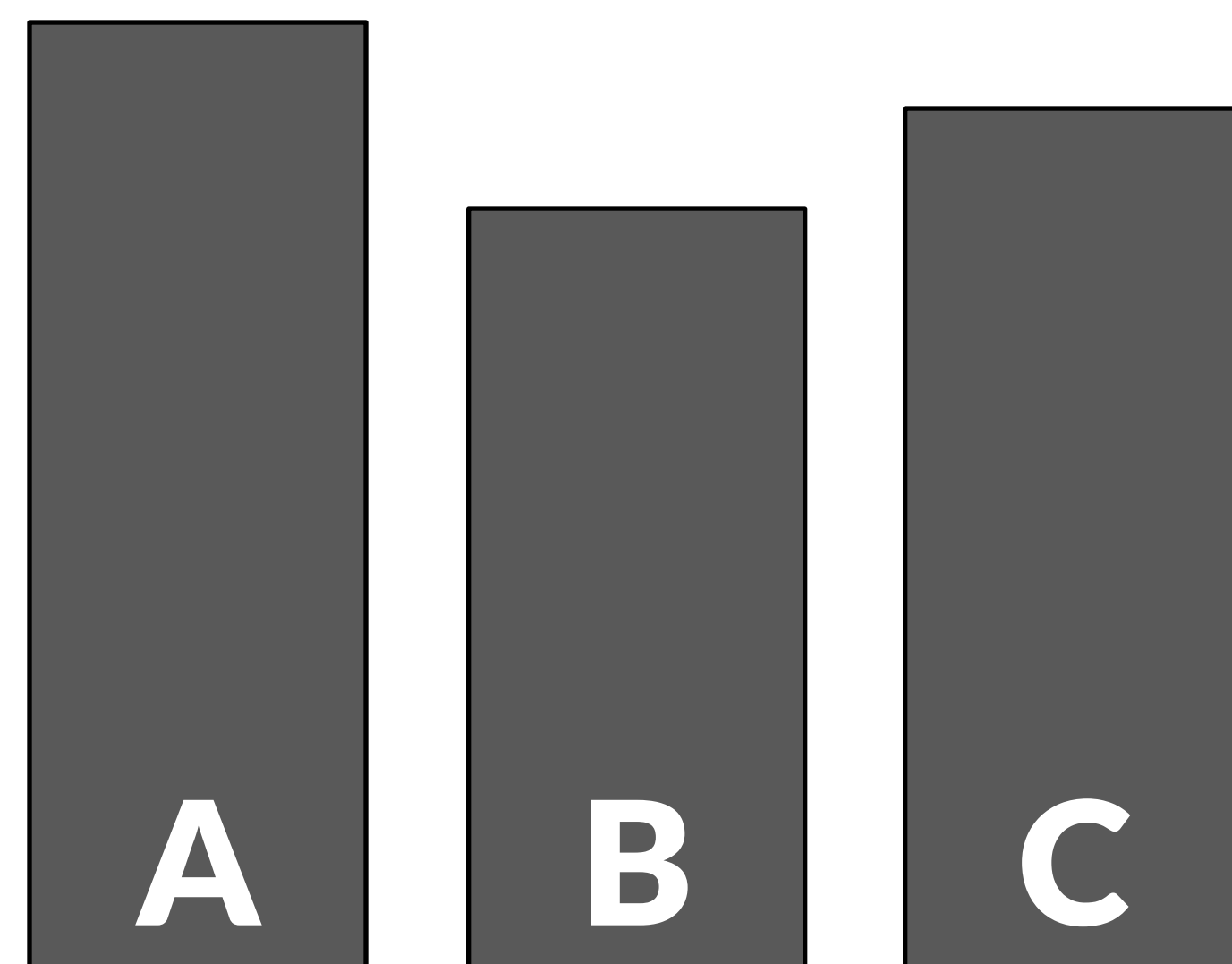


Sorting

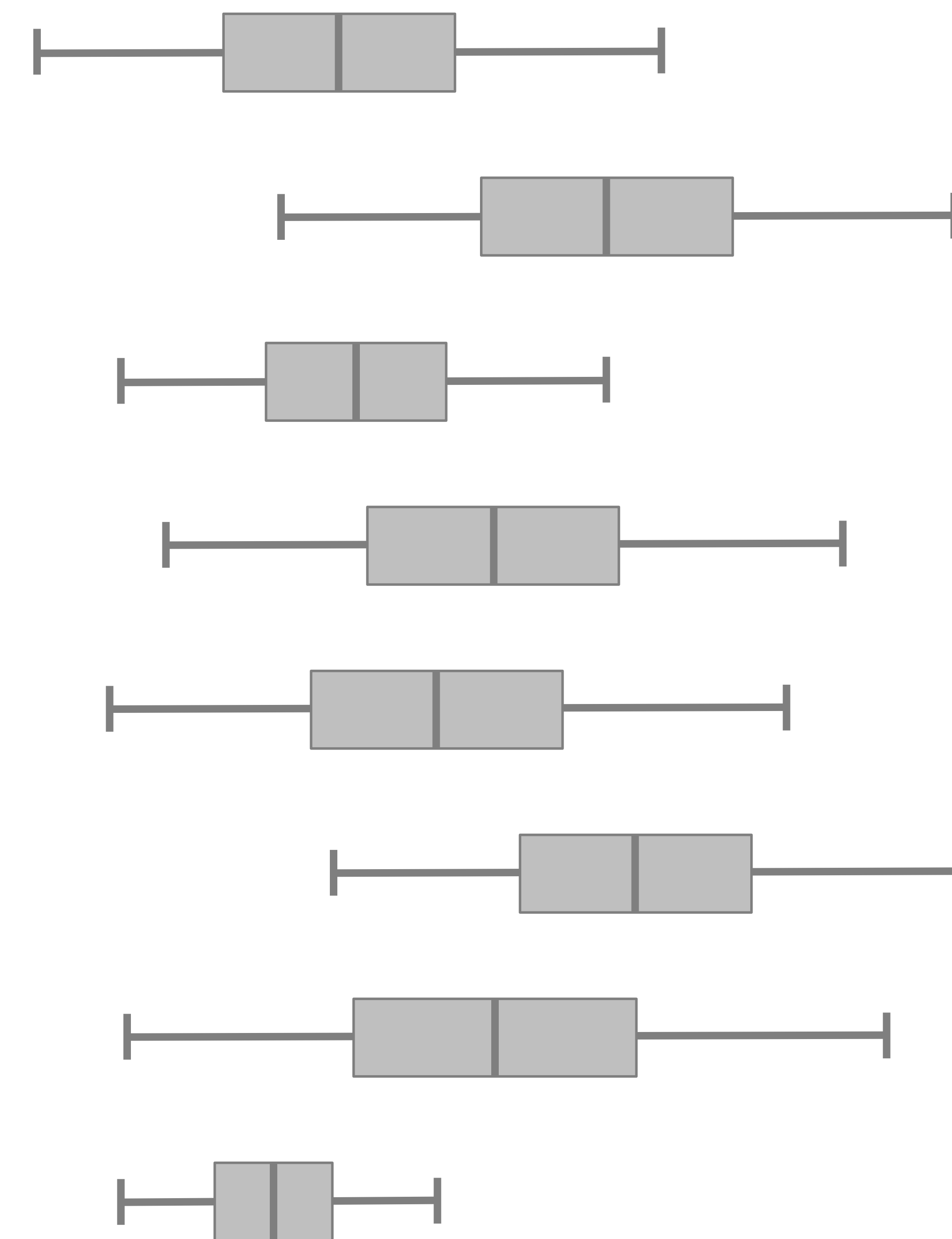
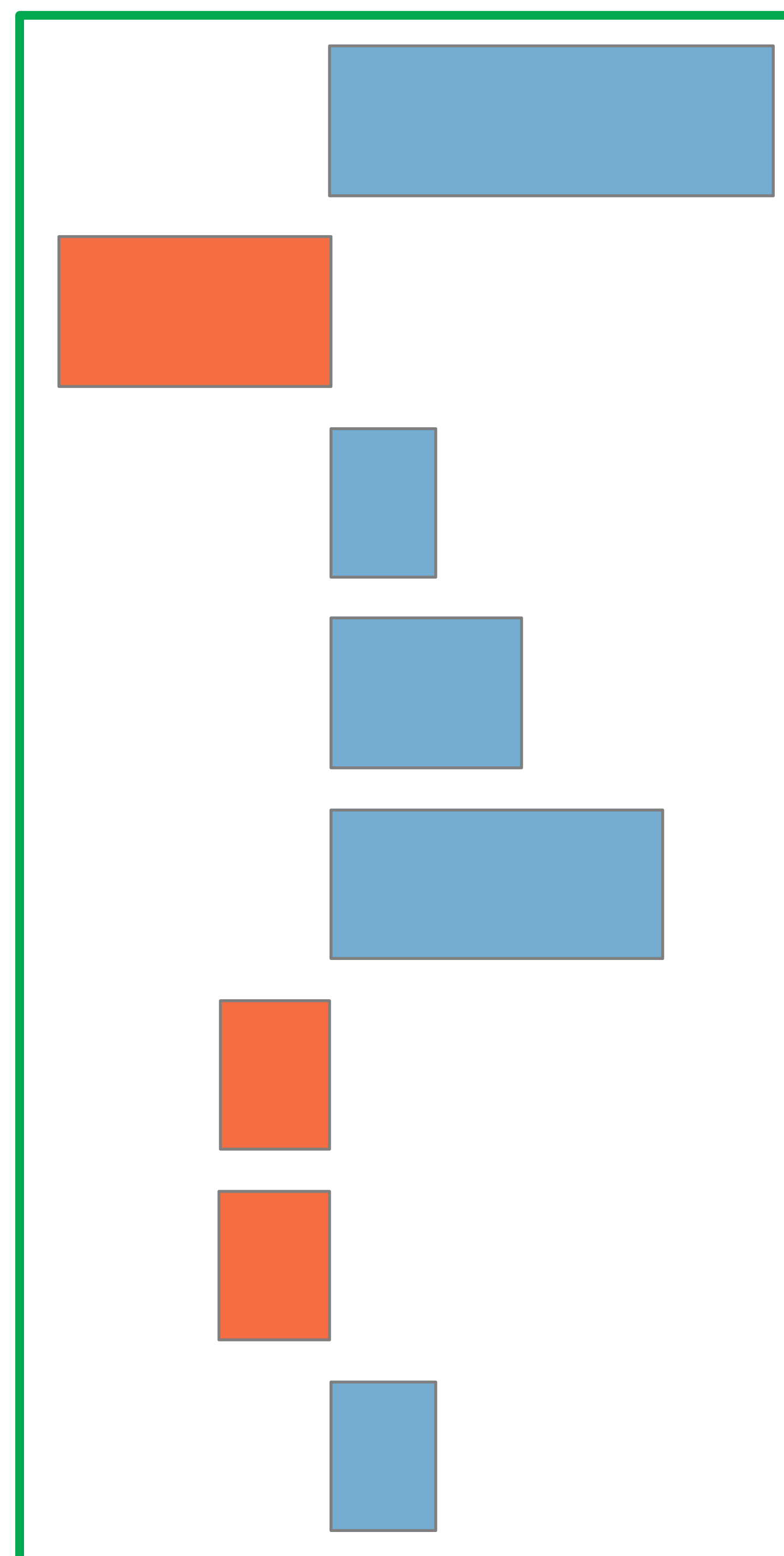


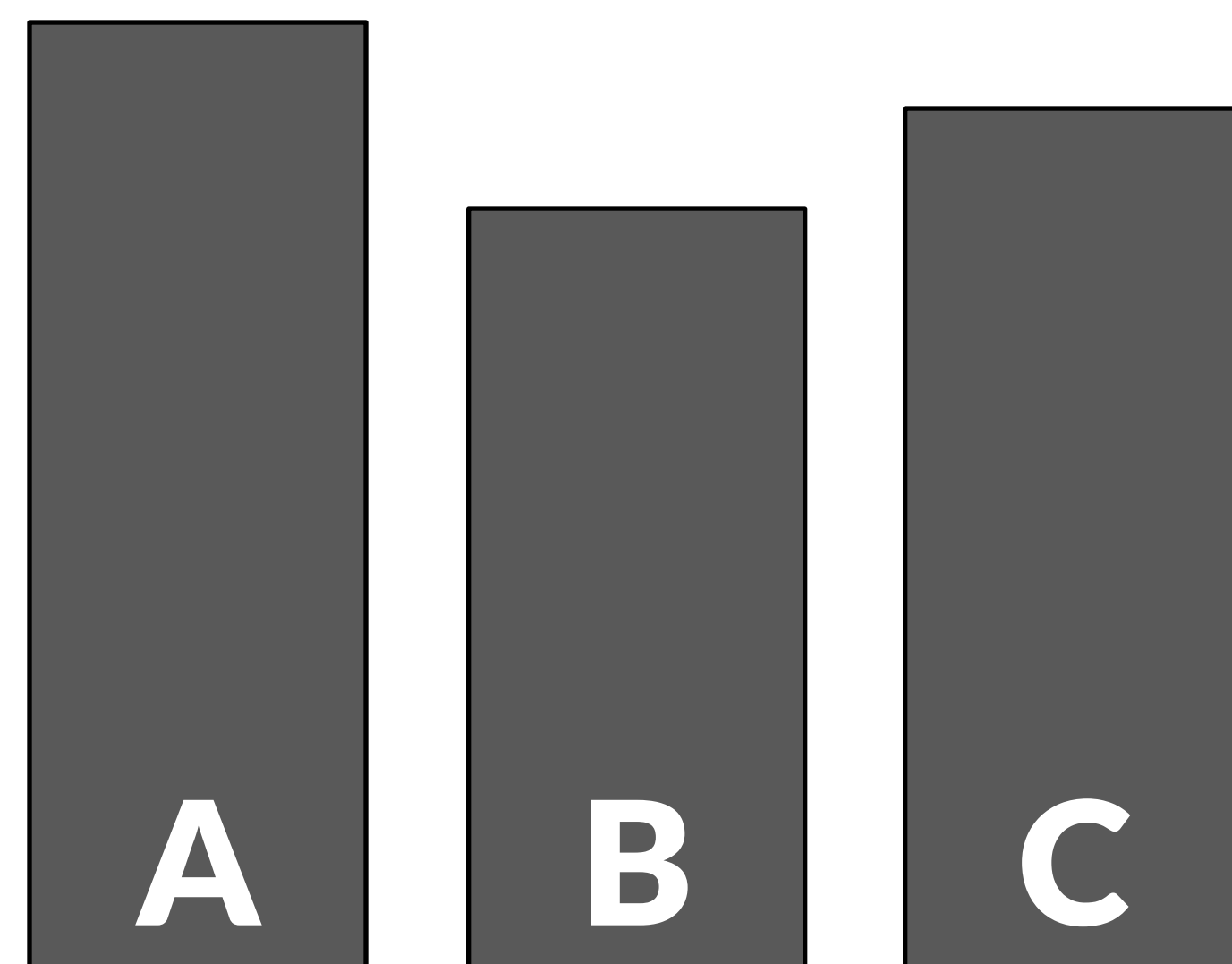
Which is the biggest intersection?
Sort By: Cardinality



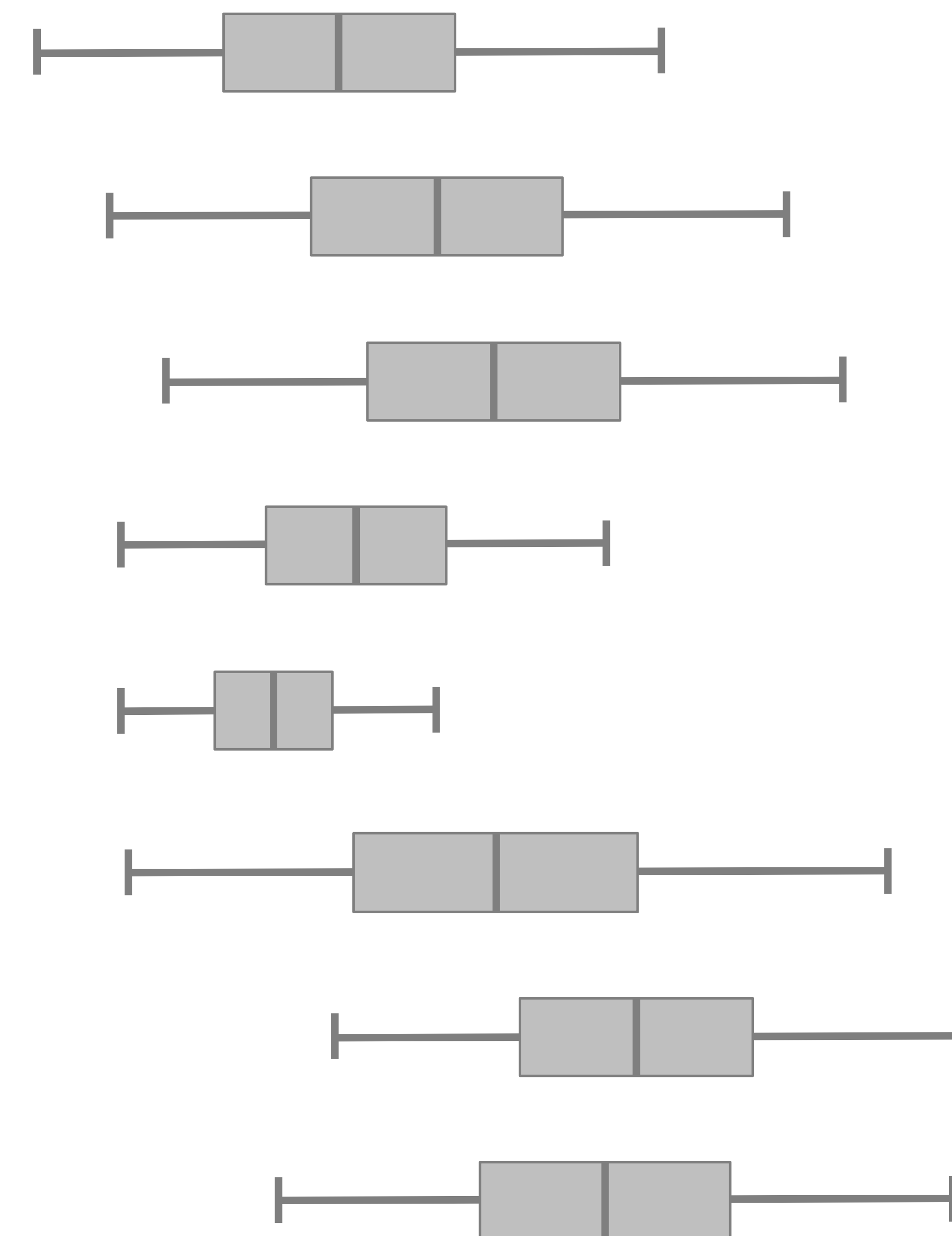
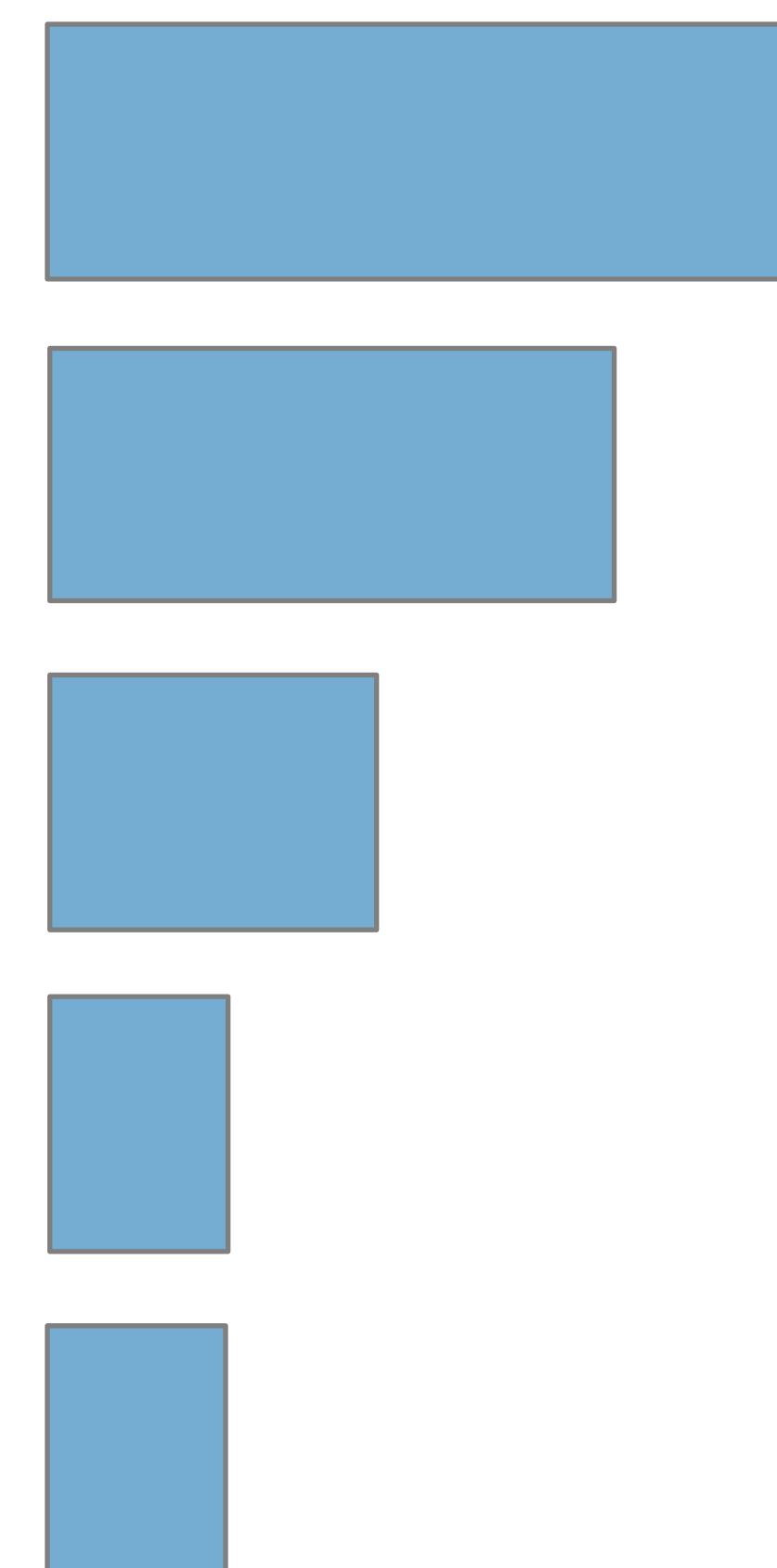
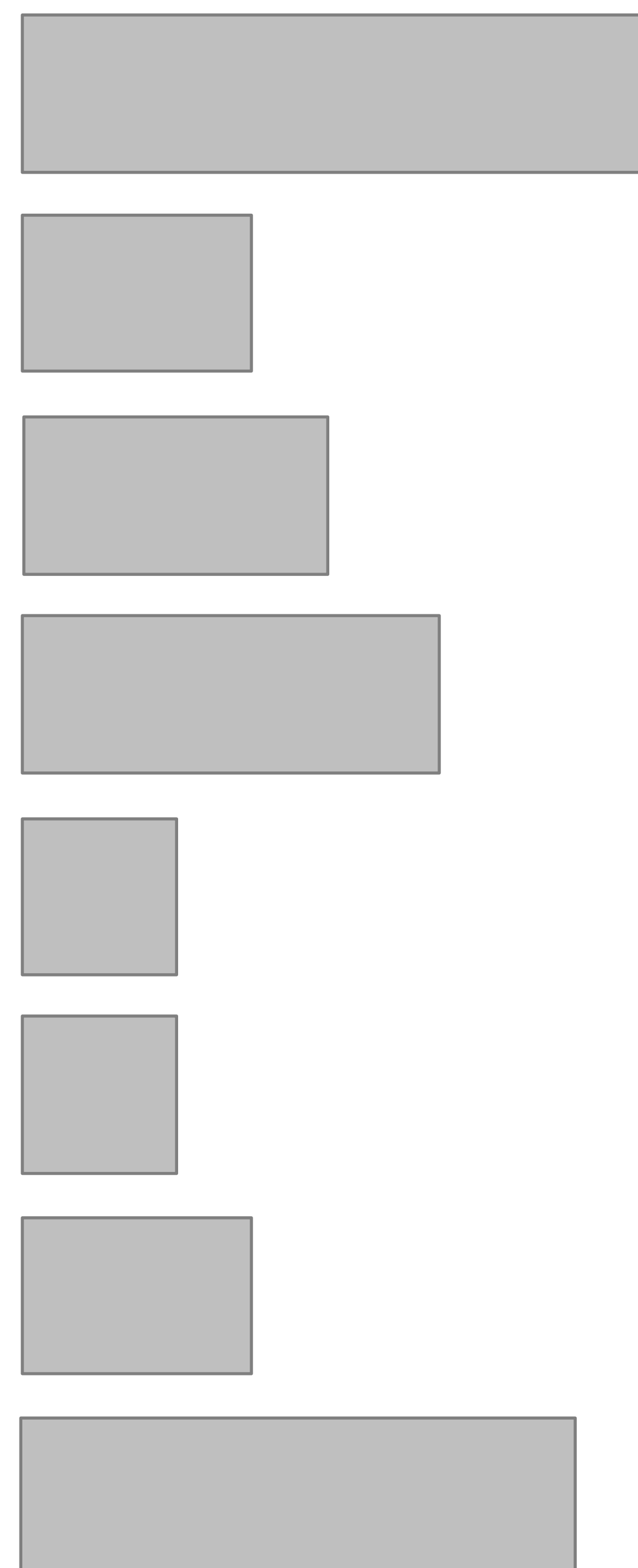
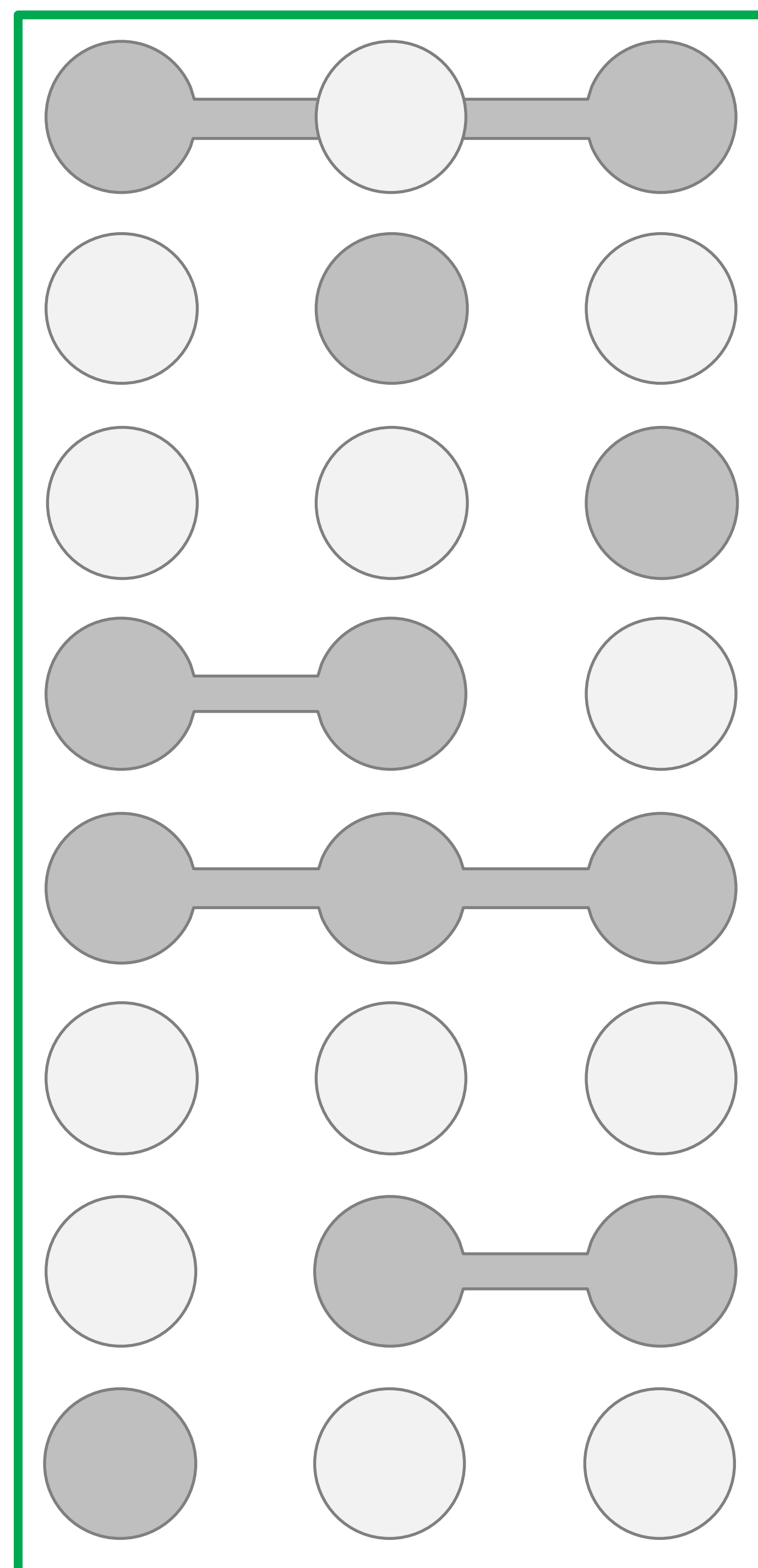


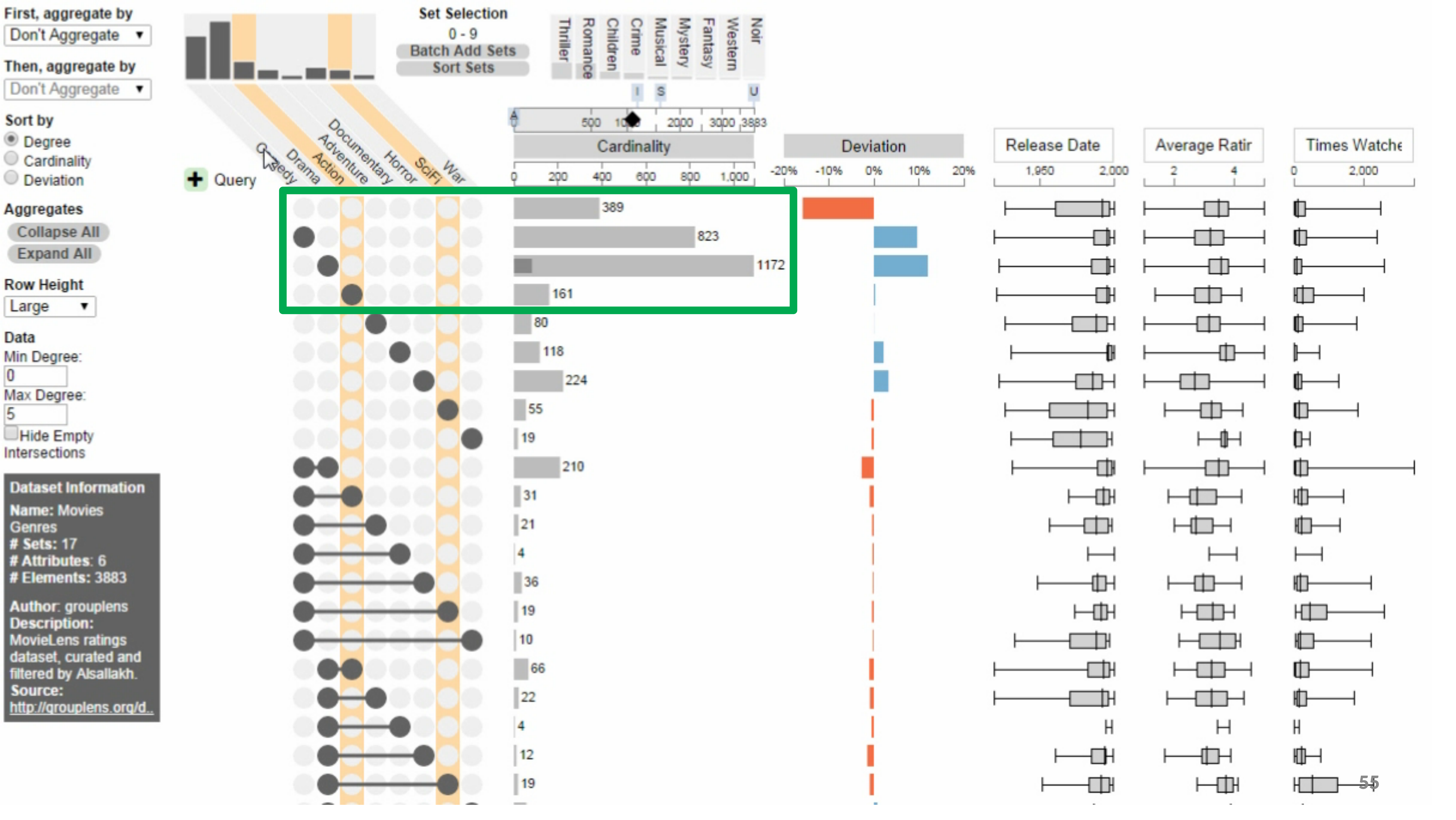
Which is the most 'surprising' intersection?
Sort By: Deviation



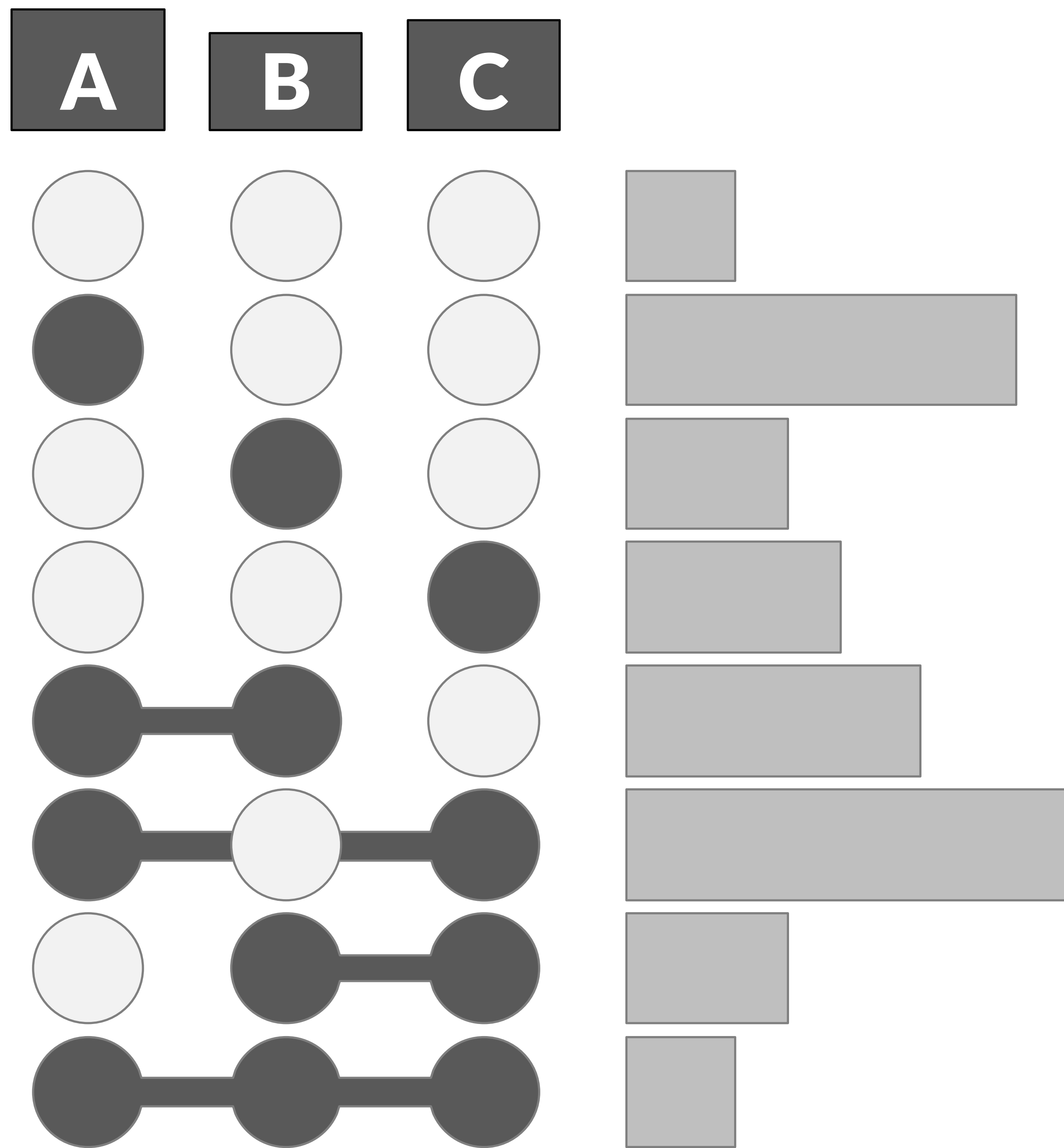


What are the properties of the intersections involving 'A'?
Sort By: Set

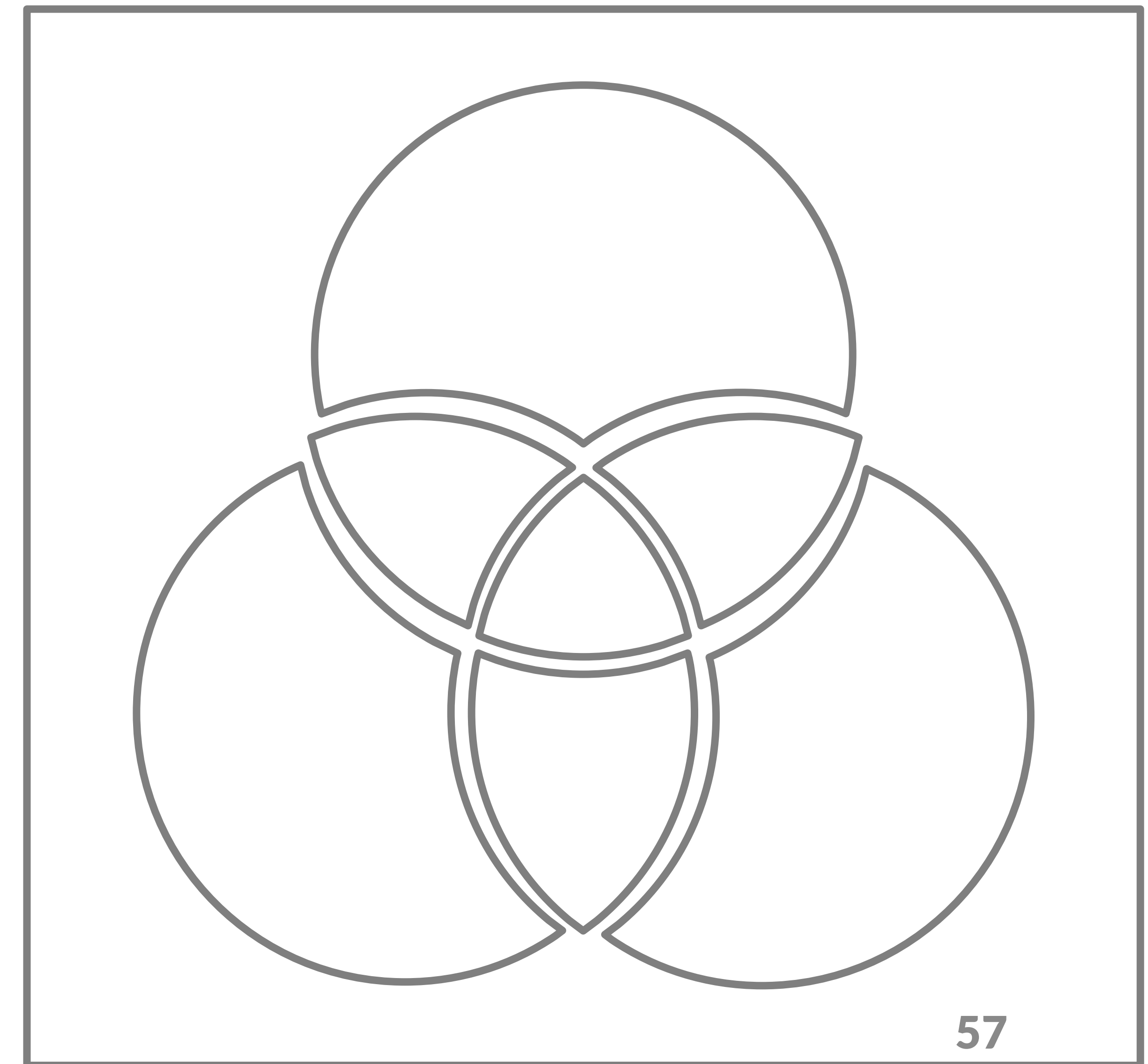




Aggregation

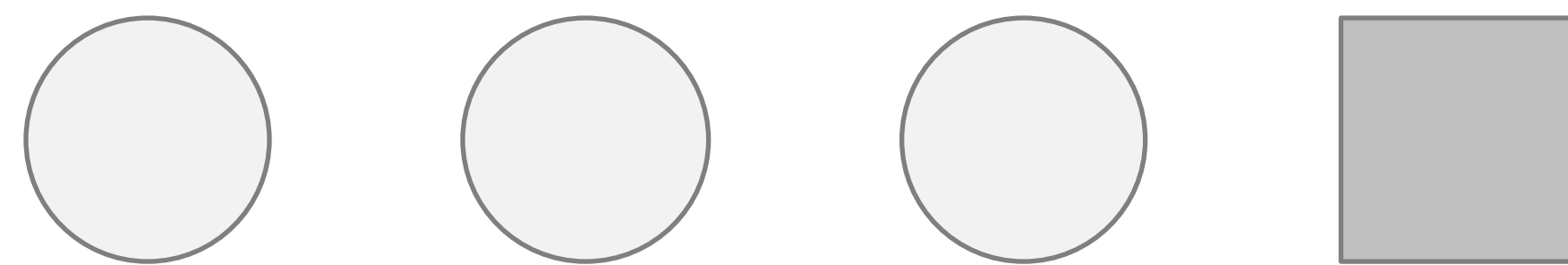


Are many items shared between two sets?
Aggregate By: Degree

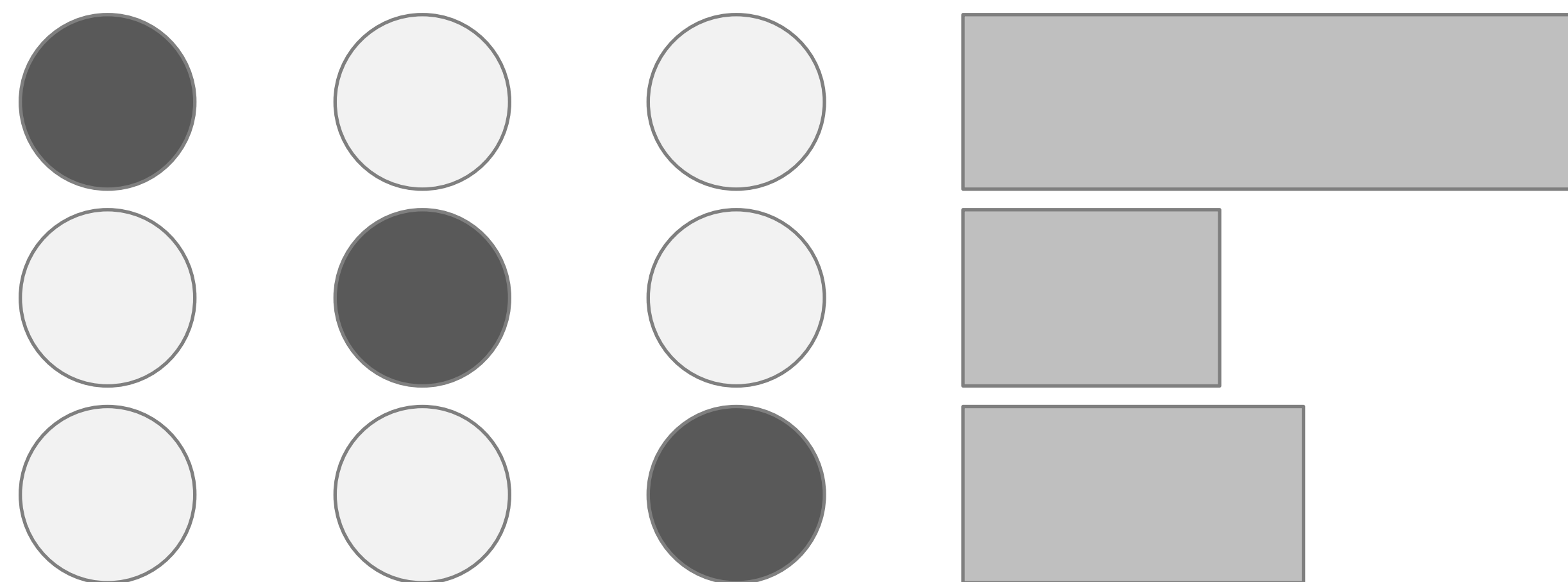


A **B** **C**

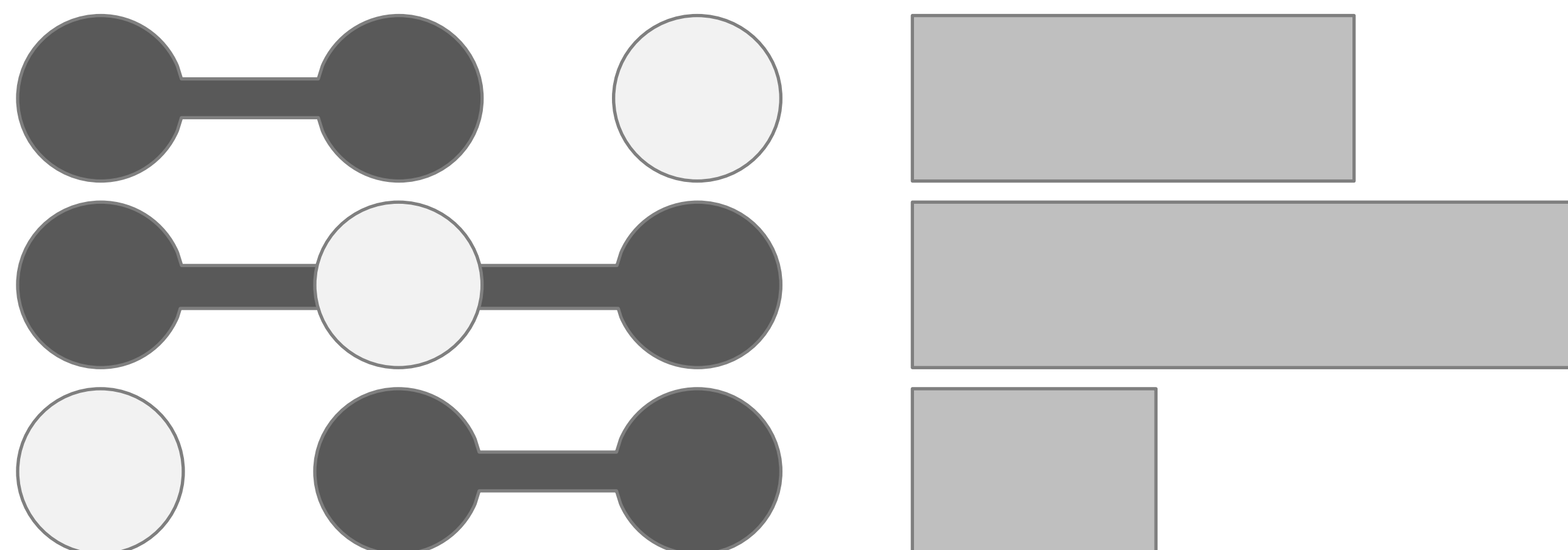
Degree 0



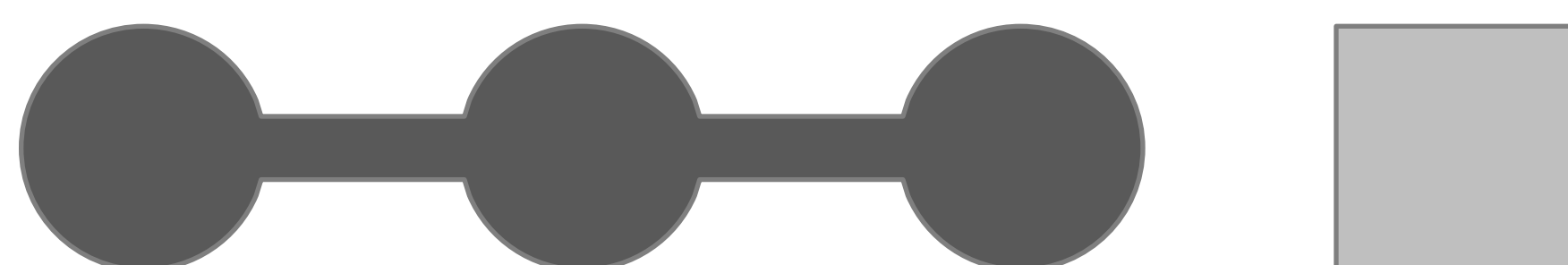
Degree 1



Degree 2

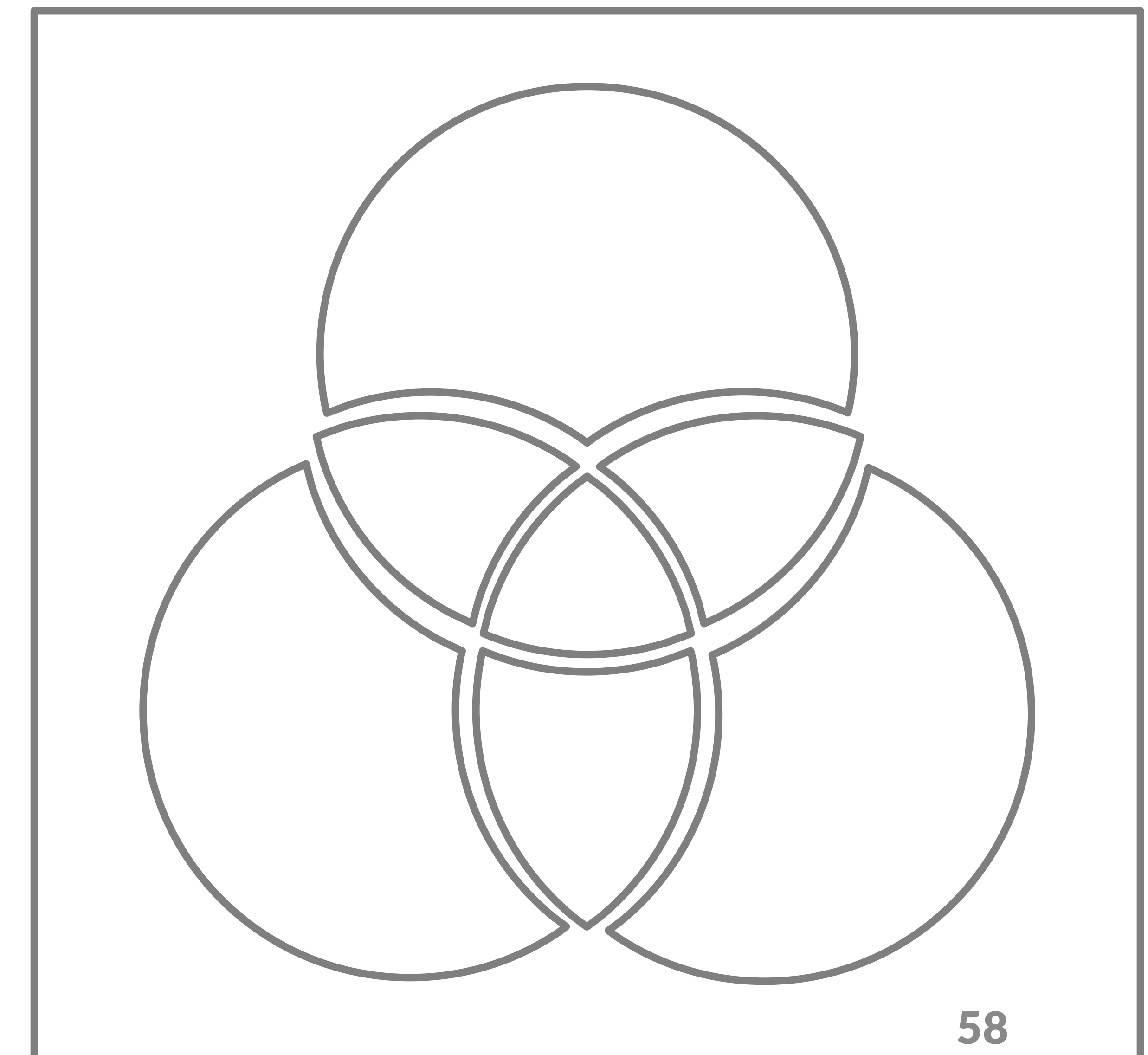


Degree 3



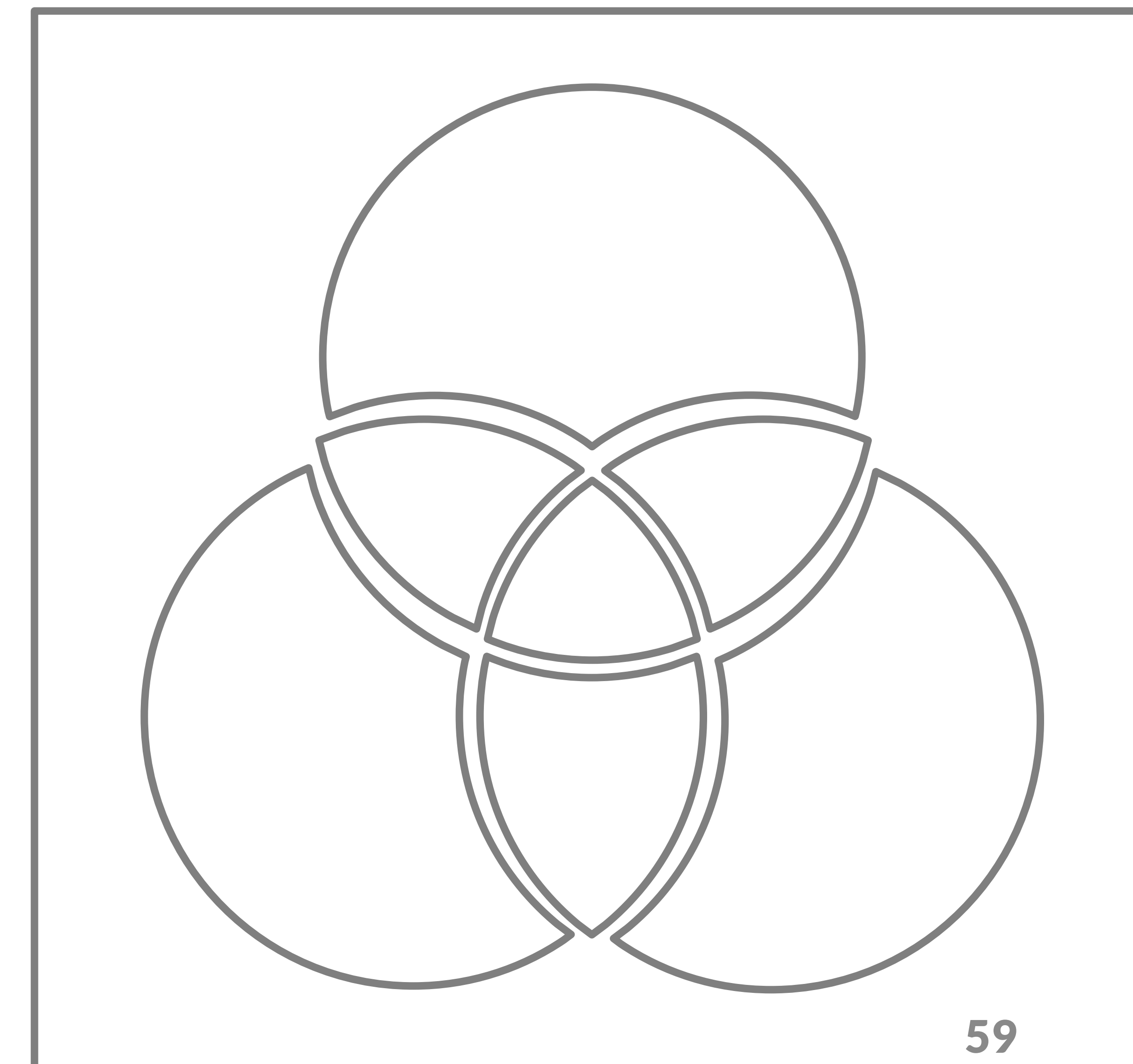
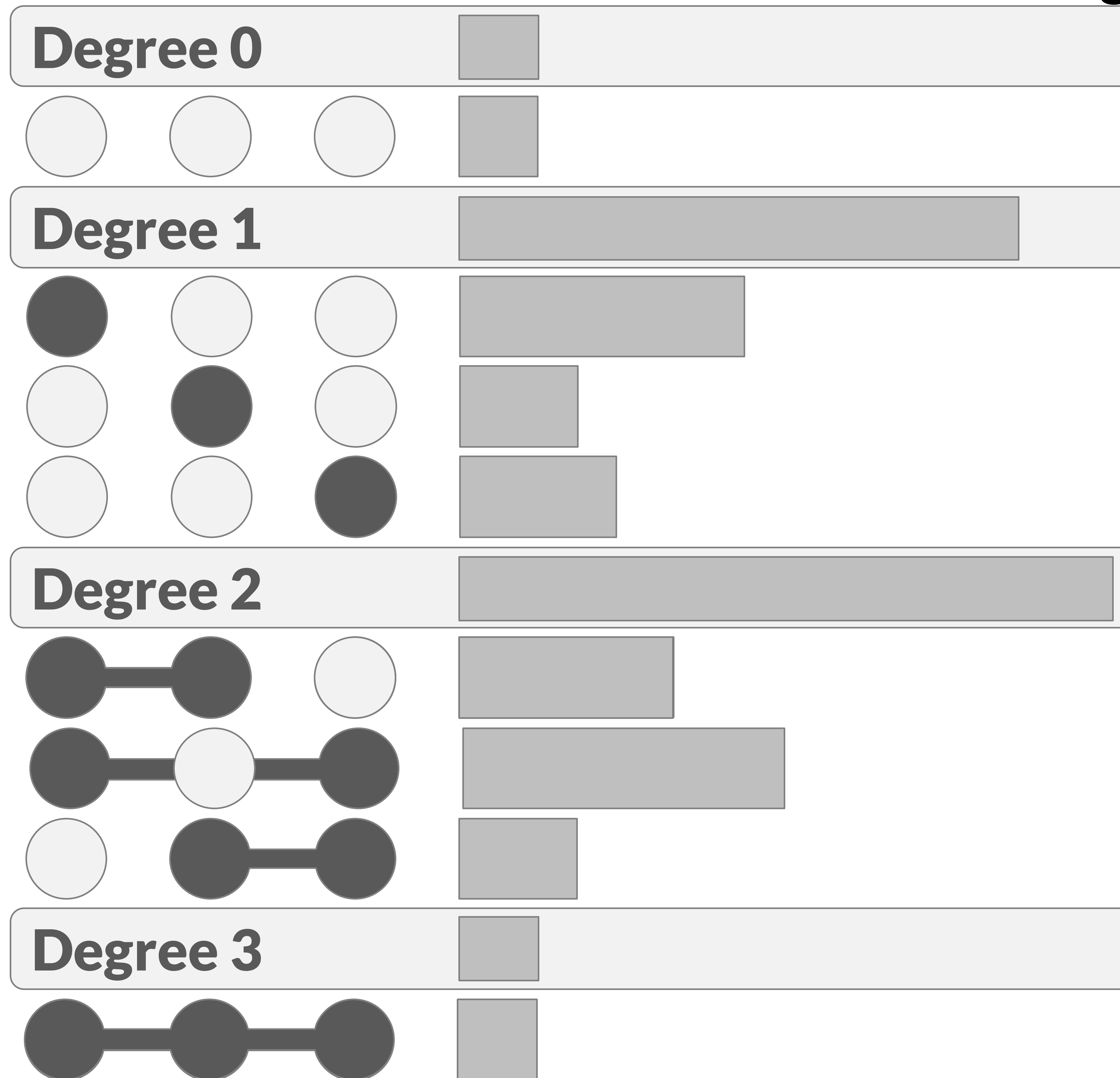
Are many items shared between two sets?
Aggregate By: Degree

Sum of children

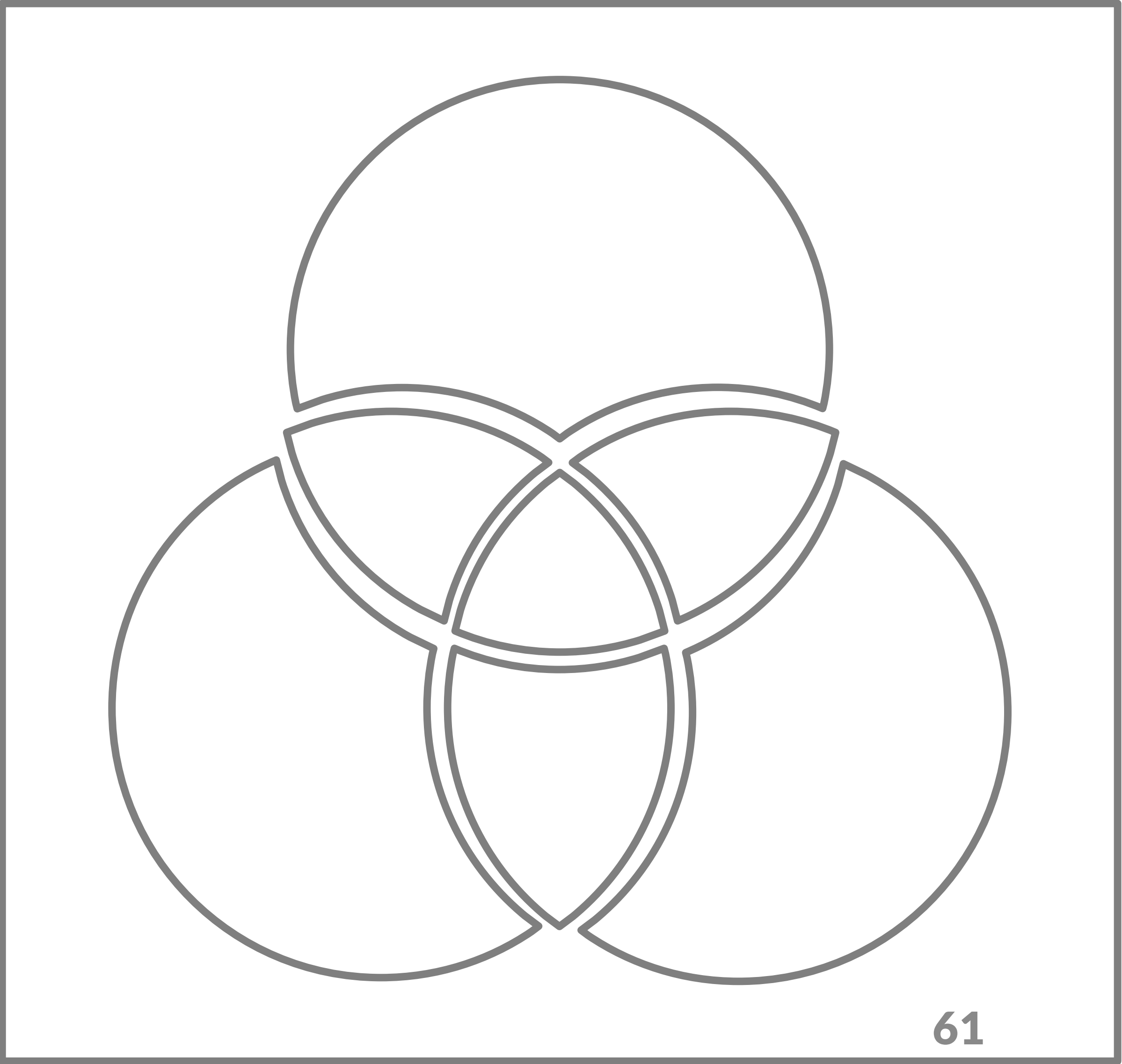
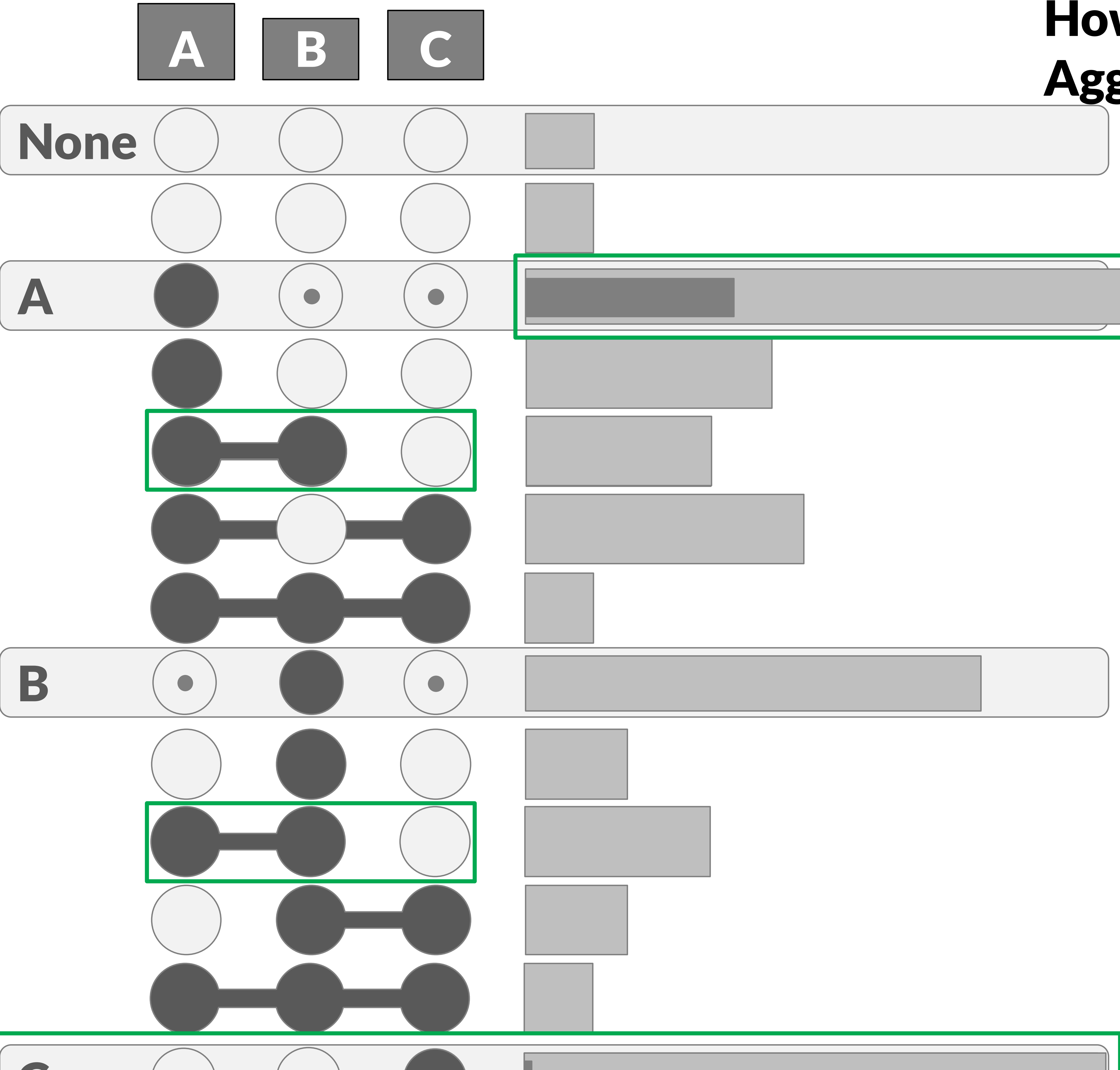


A **B** **C**

How are the elements of 'B' distributed?
Aggregate By: Set



How are the elements of 'B' distributed?
Aggregate By: Set



First, aggregate by
Degree

Then, aggregate by
Don't Aggregate

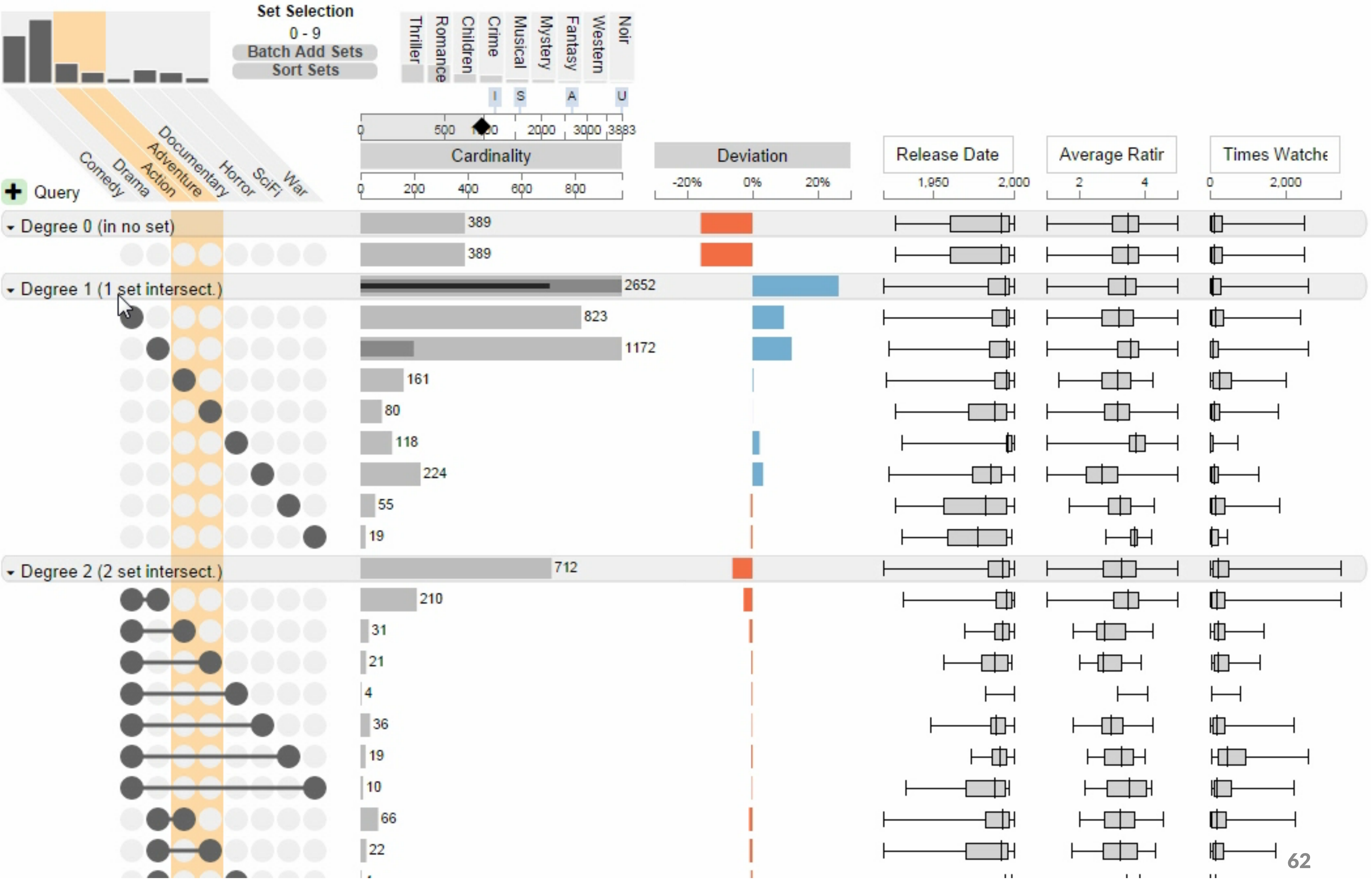
Sort by
☒ Degree
☐ Cardinality
☐ Deviation

Aggregates

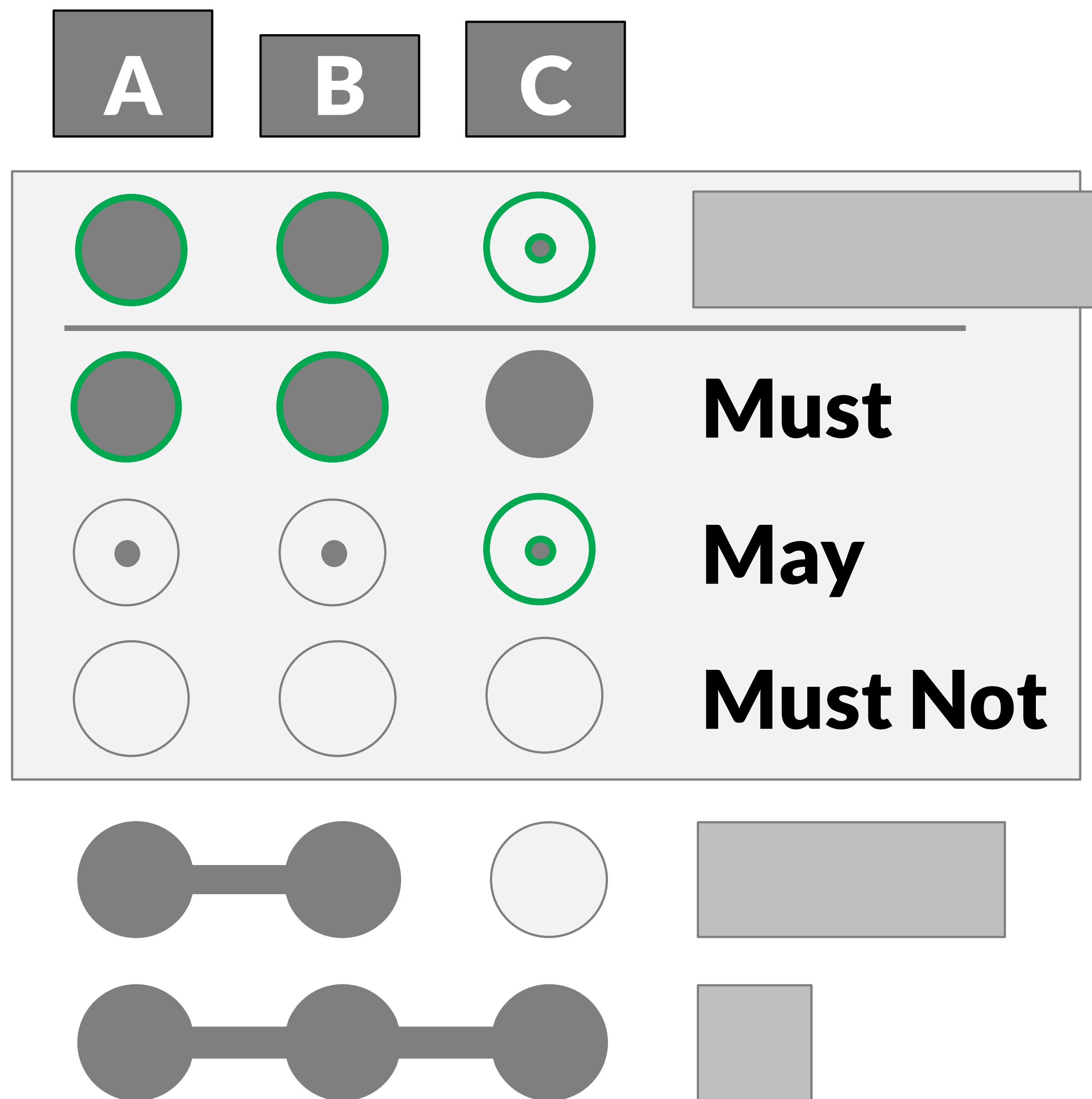
Row Height
Large

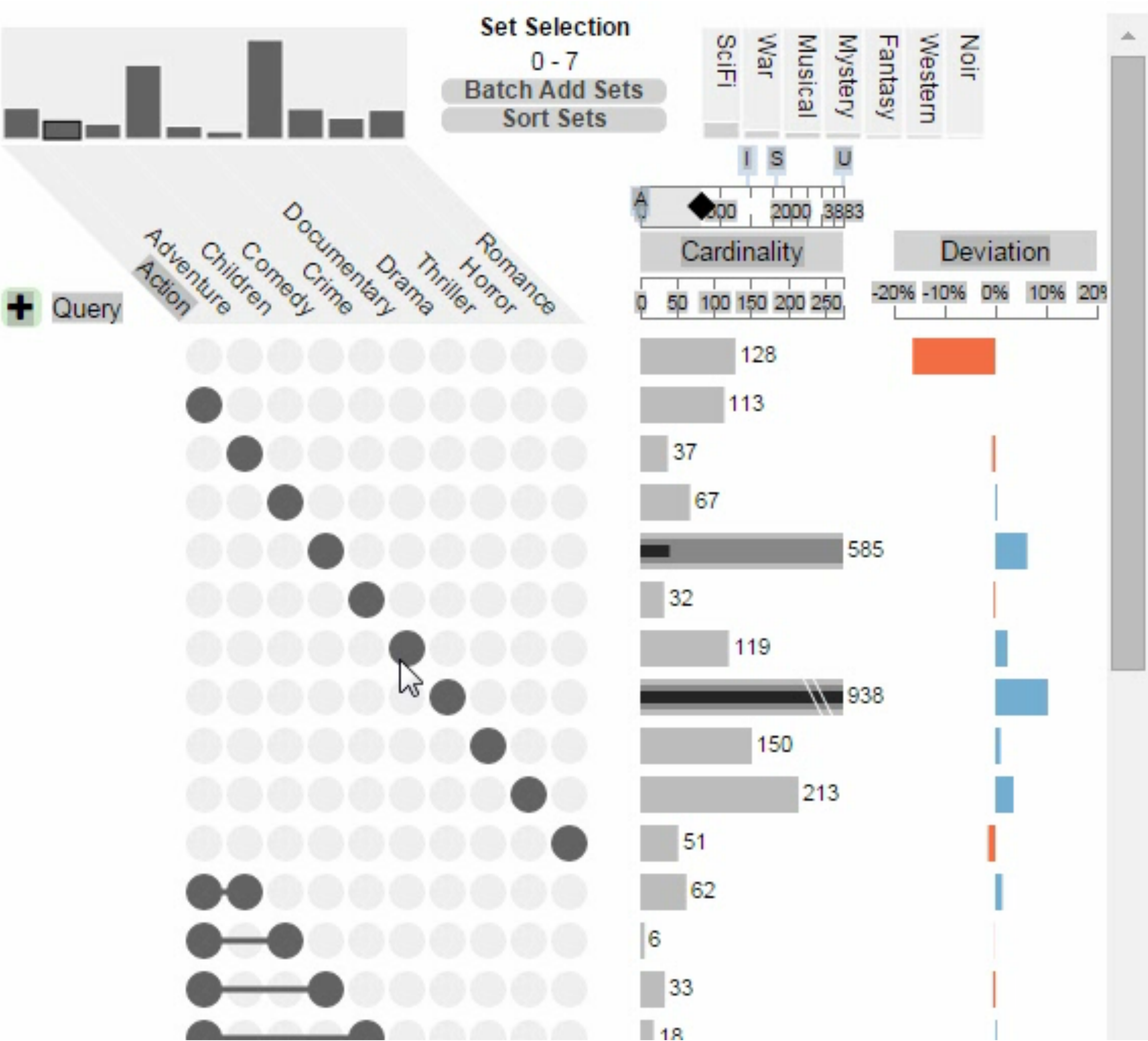
Data
Min Degree:
0
Max Degree:
5
☒ Hide Empty
Intersections

Dataset Information
Name: Movies
Genres
Sets: 17
Attributes: 6
Elements: 3883
Author: grouplens
Description:
MovieLens ratings
dataset, curated and
filtered by Alsallakh.
Source:
<http://grouplens.org/d..>



Queries





Element Visualizations

No visualizations configured. Click + button to add a new visualization.

+

Scatterplot

▼

←

→

Element Queries

No queries. Click + button to add a new query.

+

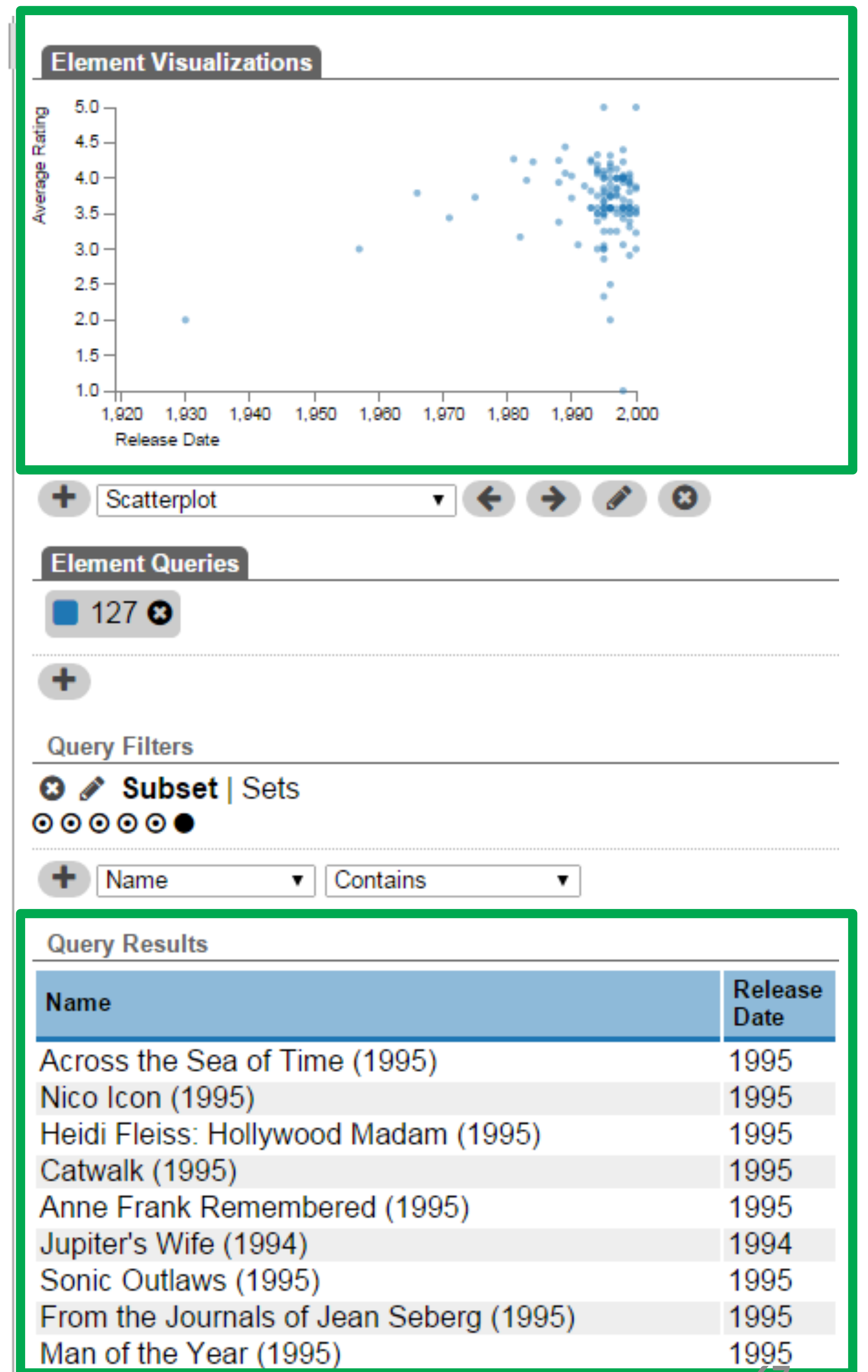
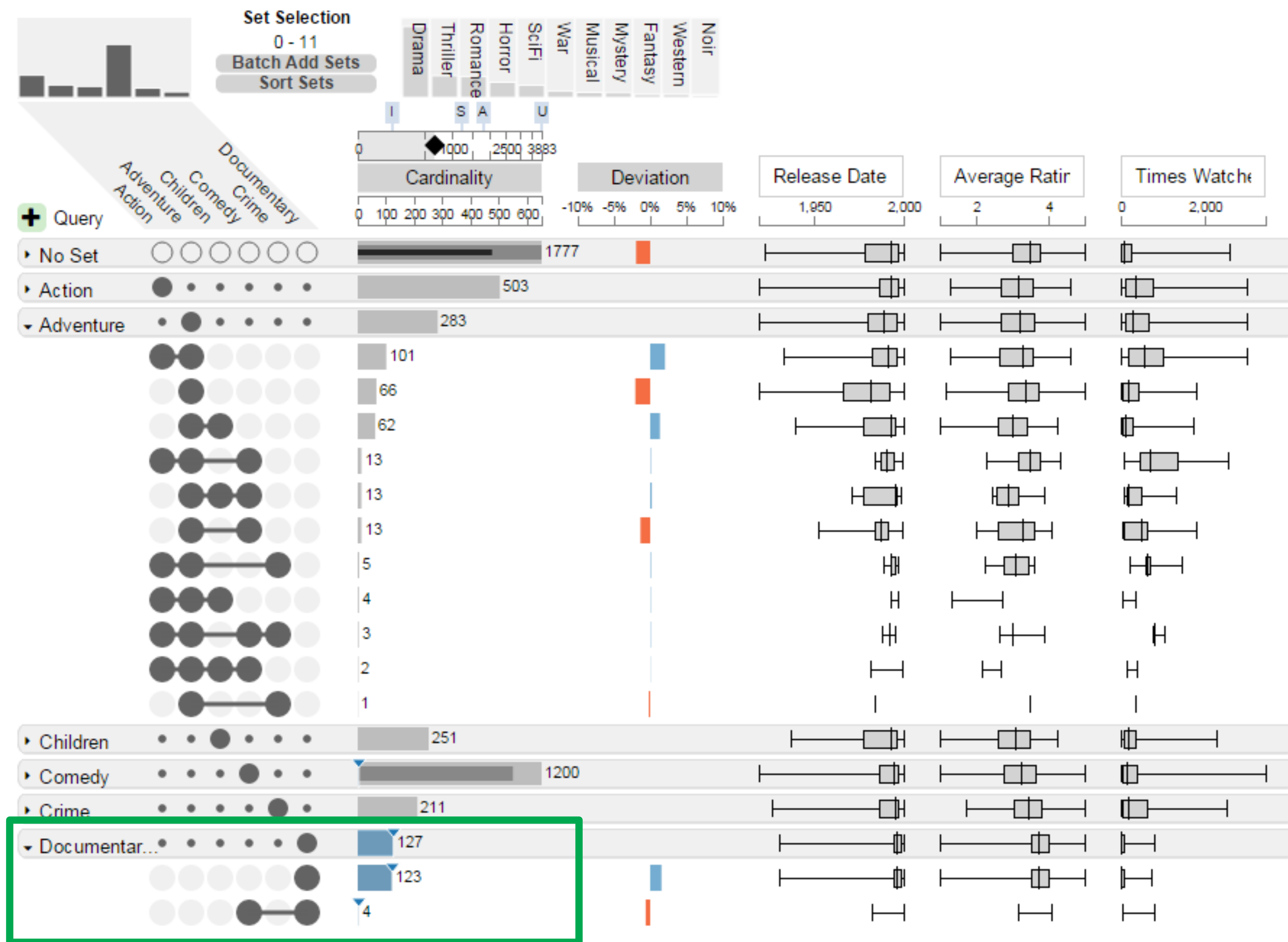
Query Filters

No active query.

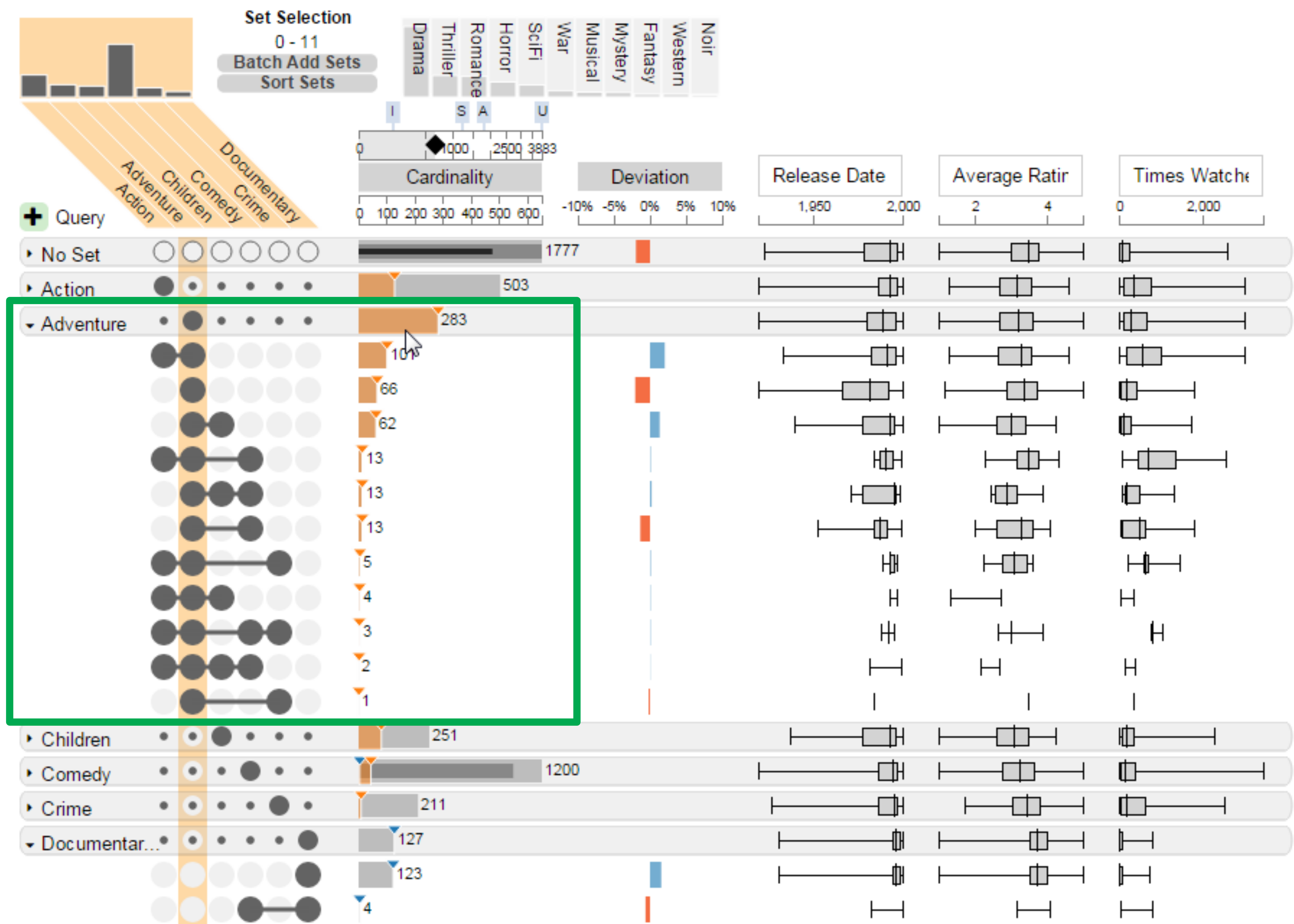
Query Results

No active query.

Elements & Attributes



How do documentaries compare to adventure movies?



Scatterplot

Element Queries

127 283

Query Filters

Subset | Sets

Name Contains

Query Results

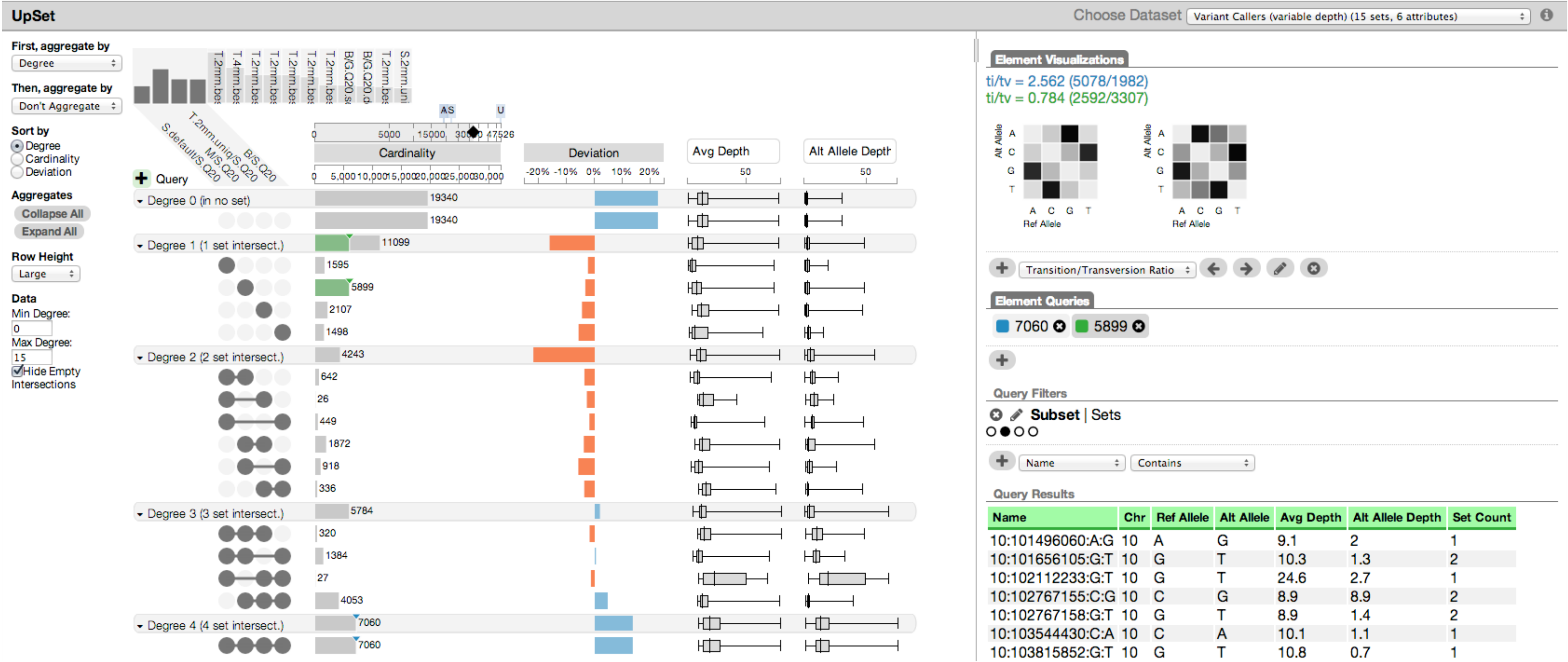
Name	Release Date
Jumanji (1995)	1995
Tom and Huck (1995)	1995
GoldenEye (1995)	1995
Cutthroat Island (1995)	1995
City of Lost Children, The (1995)	1995
Wings of Courage (1995)	1995
Mortal Kombat (1995)	1995
Kids of the Round Table (1995)	1995
Indian in the Cupboard, The (1995)	1995
White Squall (1996)	1996
Muppet Treasure Island (1996)	1996

How do documentaries compare to adventure movies?

Applications

Applications

Genetics
Economics
Pharmacology
Social Networks



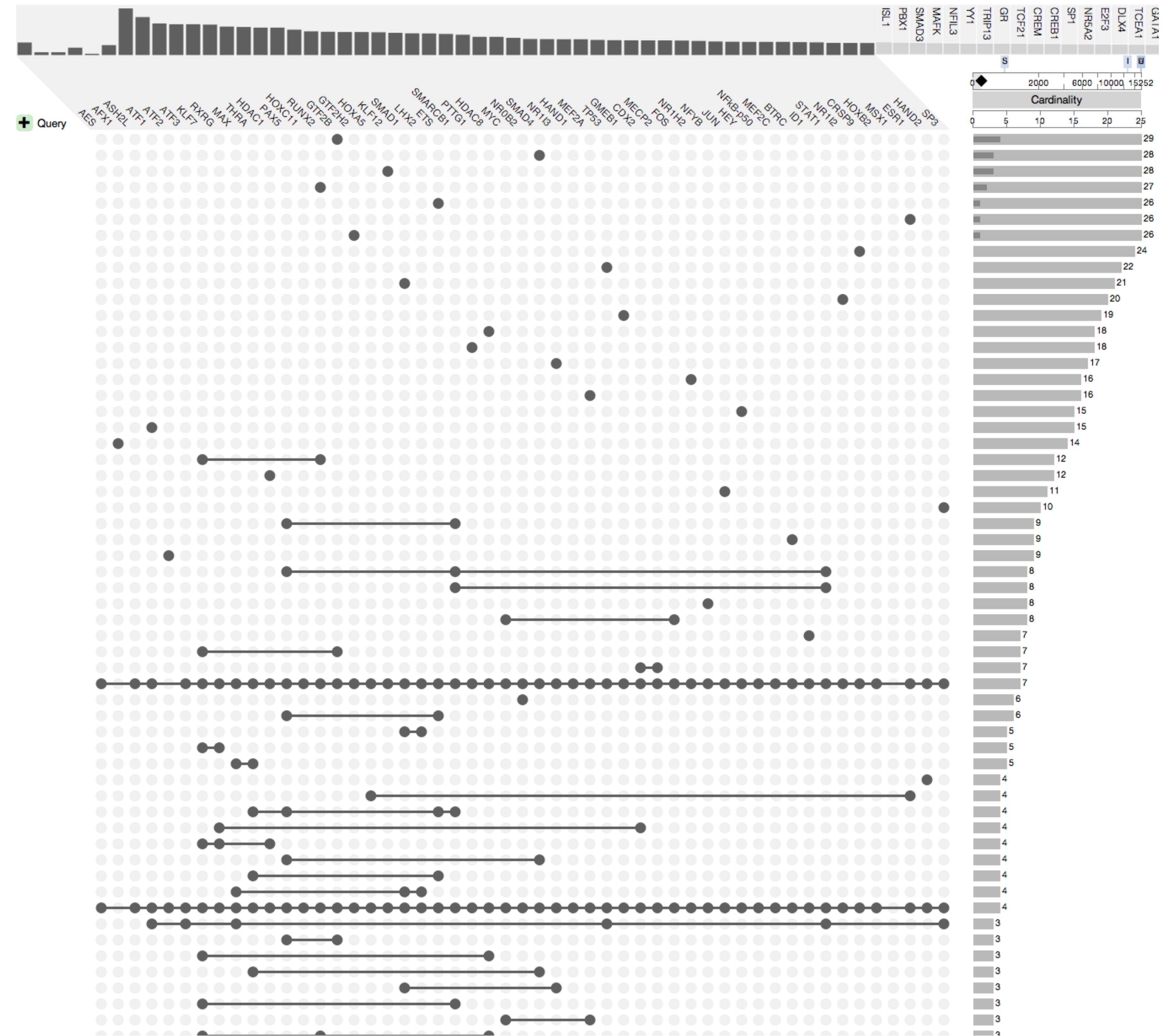
Scalability

Comfortable: ~15 sets

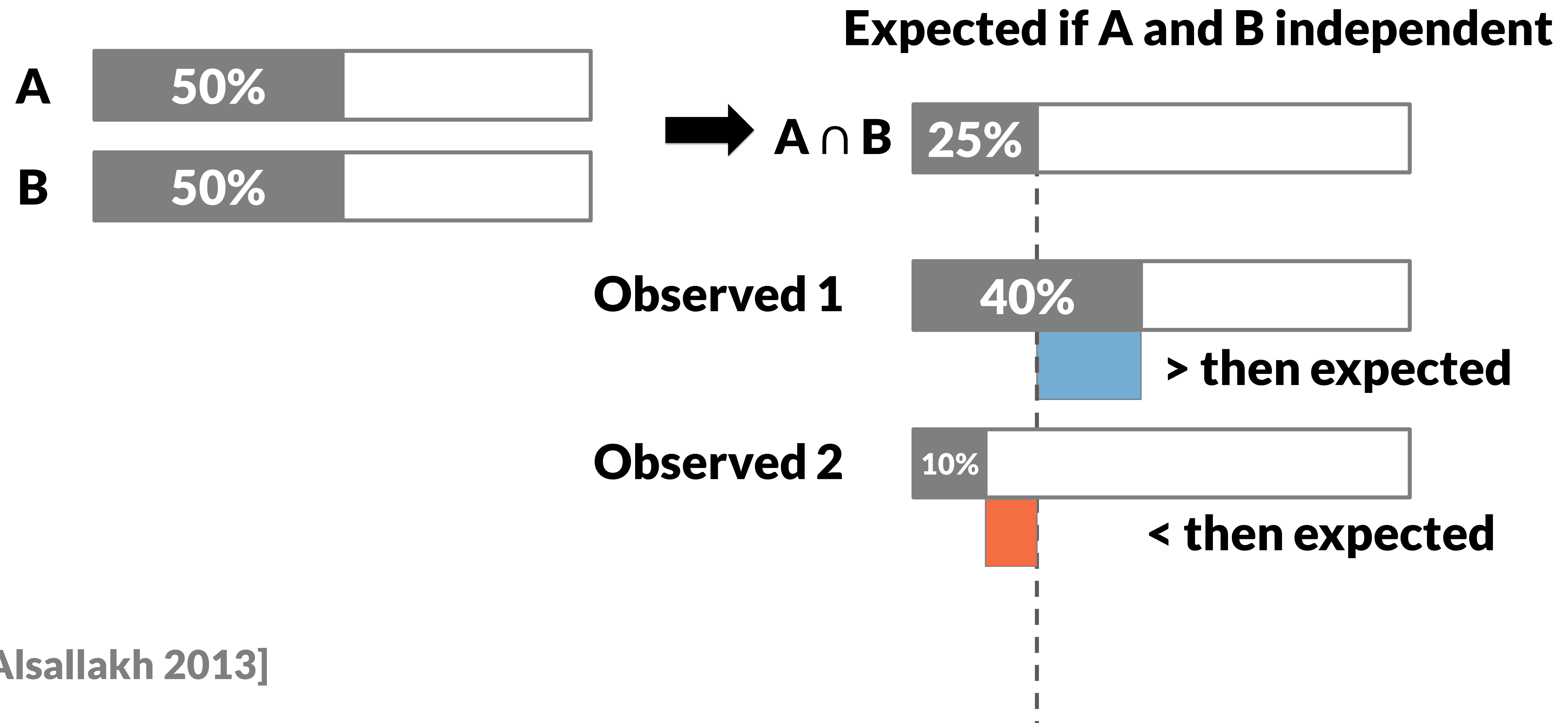
Possible: ~40 sets

**Scales with the number of
non-empty intersections**

Most datasets are sparse



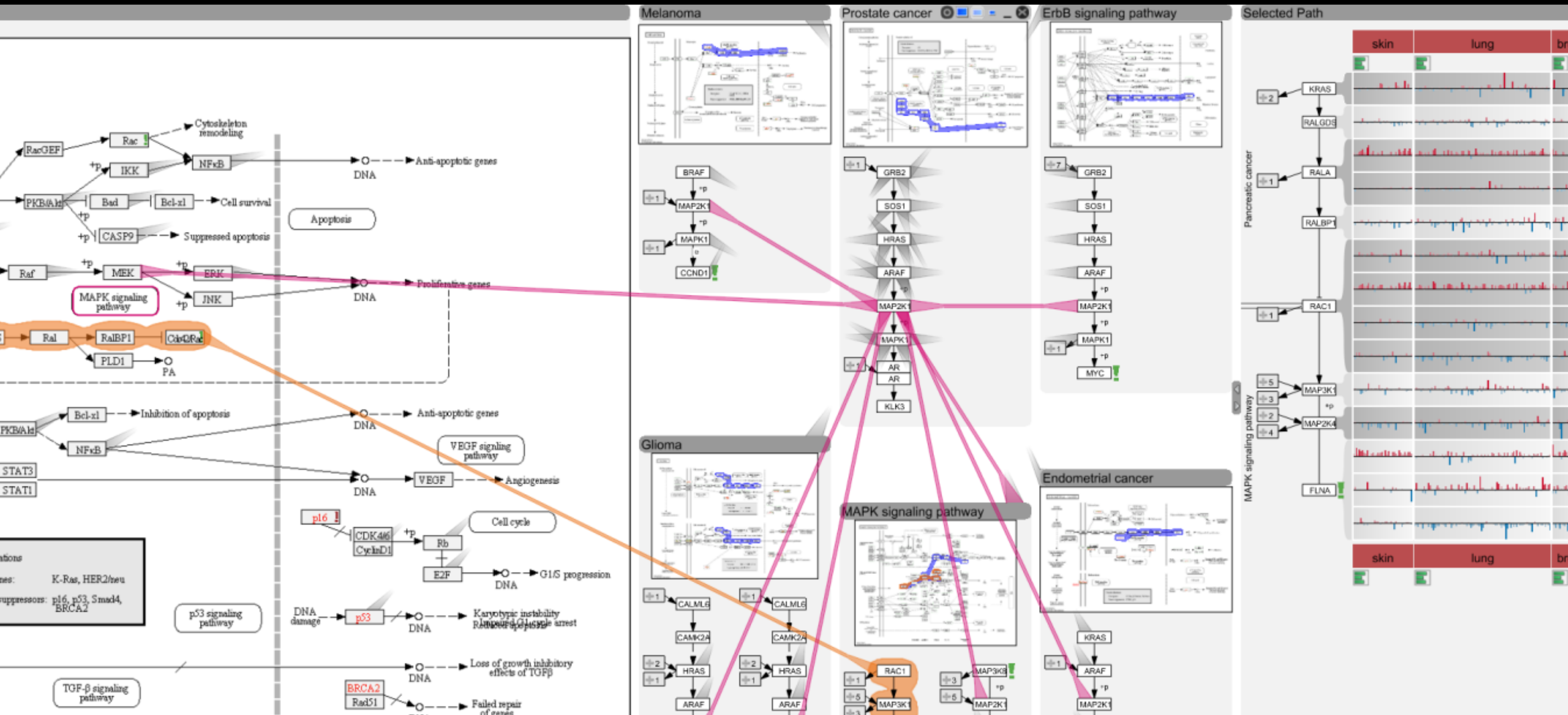
Deviation Measure



<http://vcg.github.io/upset>

[PartI, BioVis '12]
Best Paper Award
[PartI, BMC Bioinf. '13]
[Lex, InfoVis '13]

Pathways



Publications

Alexander Lex, Christian Partl, Denis Kalkofen, Marc Streit, Anne Mai Wasserman, Samuel Gratzl, Dieter Schmalstieg and Hanspeter Pfister

Entourage: Visualizing Relationships between Biological Pathways using Contextual Subsets

IEEE TVCG (InfoVis '13), 2013

Christian Partl, Alexander Lex, Marc Streit, Denis Kalkofen, Karl Kashofer and Dieter Schmalstieg
enRoute: Dynamic Path Extraction from Biological Pathway Maps for Exploring Heterogeneous Experimental Datasets

BMC Bioinformatics, 2013

Christian Partl, Alexander Lex, Marc Streit, Denis Kalkofen, Karl Kashofer, Dieter Schmalstieg
enRoute: Dynamic Path Extraction from Biological Pathway Maps for In-Depth Experimental Data Analysis

BioVis, 2012

Best Paper Award



KEGG PATHWAY Database

Wiring diagrams of molecular interactions, reactions, and relations

KEGG2 PATHWAY BRITE MODULE KO GENOME GENES LIGAND DISEASE DRUG DBGET

Select prefix

map

Organism

Enter keywords

Go

Help

[New maps | Update history]

Pathway Maps

KEGG PATHWAY is a collection of manually drawn pathway maps representing our knowledge on the molecular interaction and reaction networks for:

1. Metabolism

Global/overview Carbohydrate Energy Lipid Nucleotide Amino acid Other amino Glycan
Cofactor/vitamin Terpenoid/PK Other secondary metabolite Xenobiotics Chemical structure

2. Genetic Information Processing

3. Environmental Information Processing

4. Cellular Processes

5. Organismal Systems

6. Human Diseases

and also on the structure relationships (KEGG drug structure maps) in:

7. Drug Development

Pathway Mapping

KEGG PATHWAY mapping is the process to map molecular datasets, especially large-scale datasets in genomics, transcriptomics, proteomics, and metabolomics, to the KEGG pathway maps for biological interpretation of higher-level systemic functions.

- Search Pathway - basic pathway mapping tool
- Search&Color Pathway - advanced pathway mapping tool
- Color Pathway - selected pathway map coloring tool

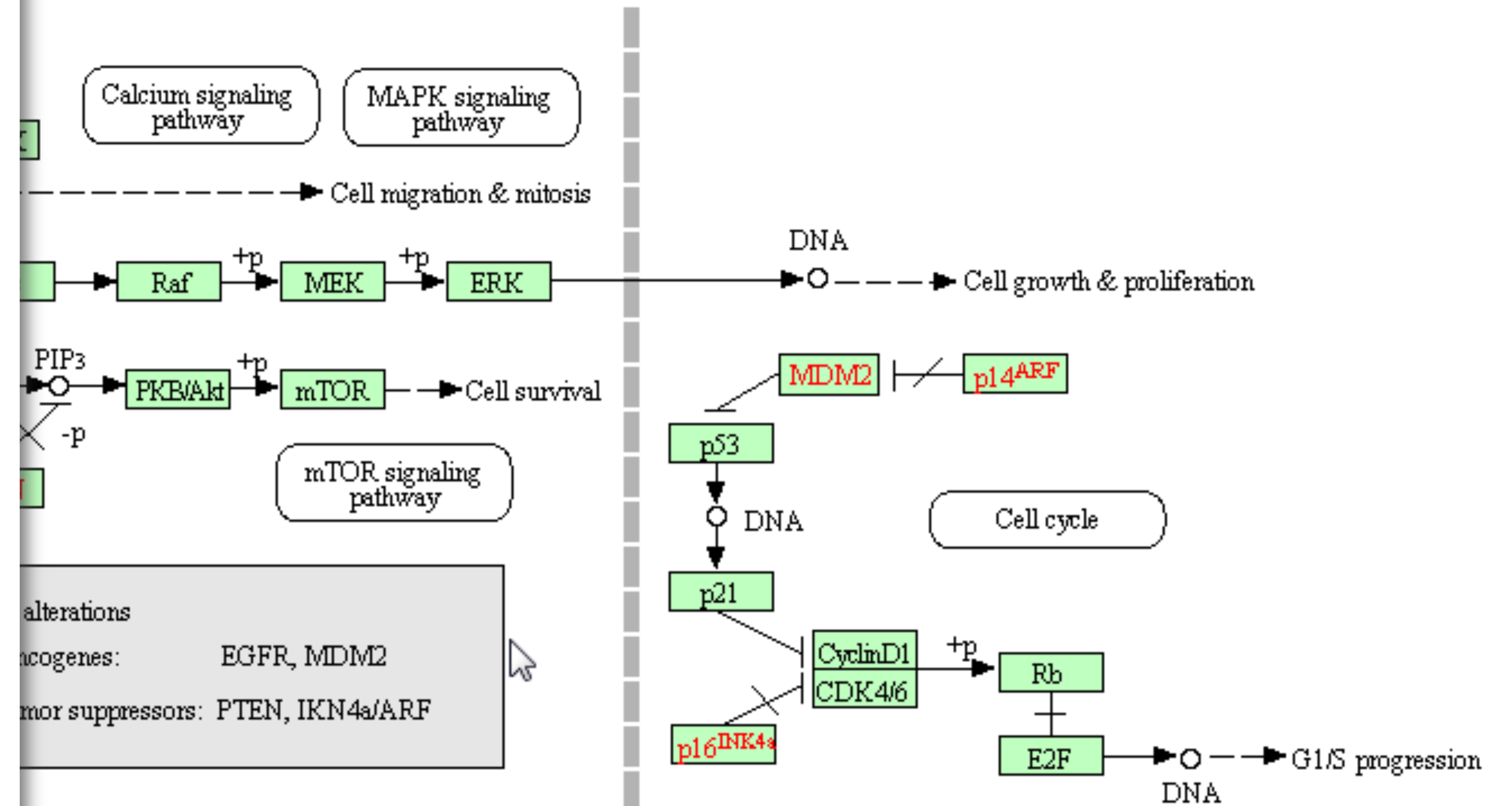
1. Metabolism

1.0 Global and overview maps

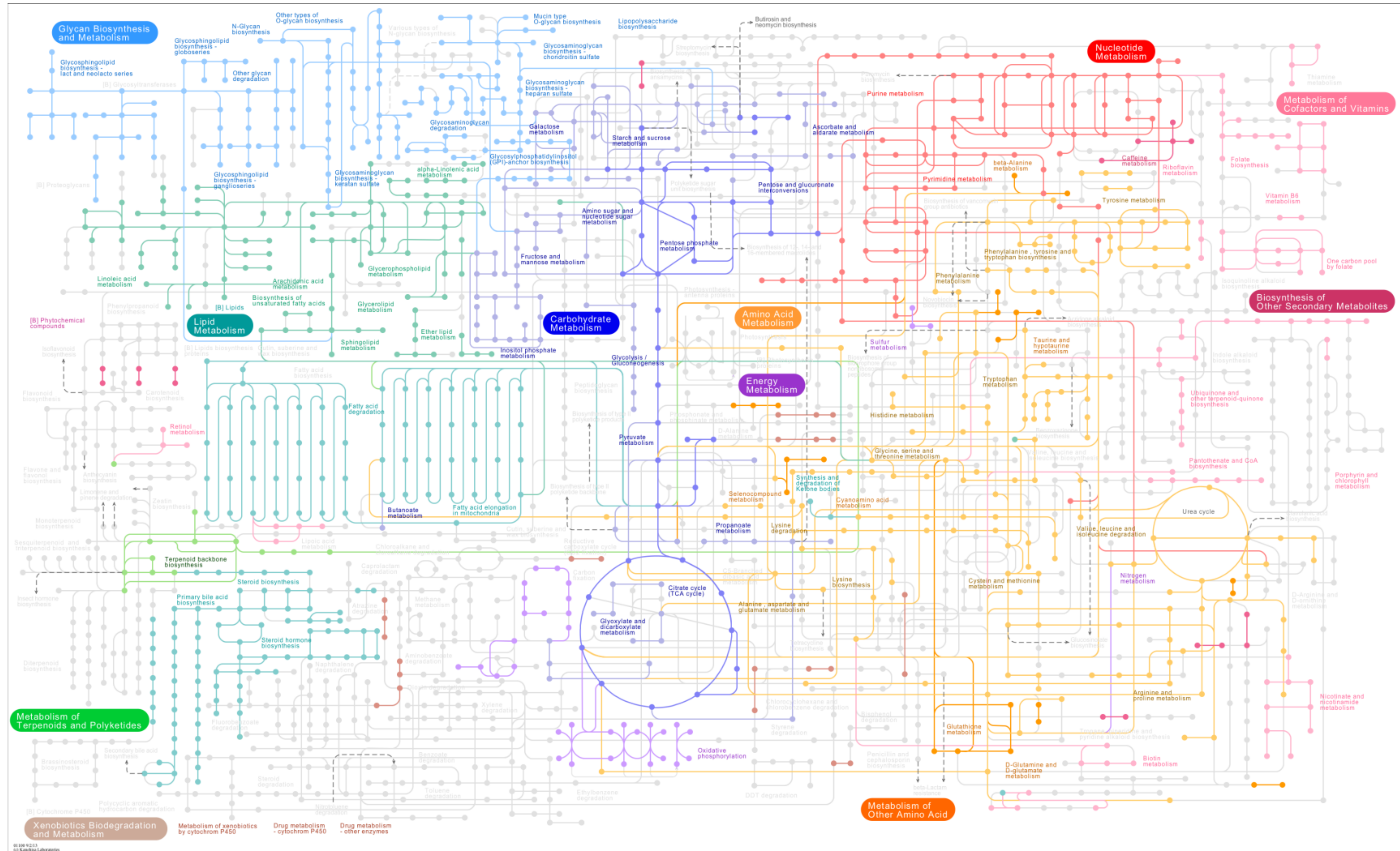
Metabolic pathways	[KEGG Atlas]	KEGG modules
Biosynthesis of secondary metabolites	[KEGG Atlas]	KEGG reaction modules
Microbial metabolism in diverse environments	[KEGG Atlas]	
Carbon metabolism	[KEGG Atlas]	
2-Oxocarboxylic acid metabolism	[KEGG Atlas]	
Fatty acid metabolism	[KEGG Atlas]	
Biosynthesis of amino acids	[KEGG Atlas]	
Degradation of aromatic compounds	[KEGG Atlas]	

1.1 Carbohydrate metabolism

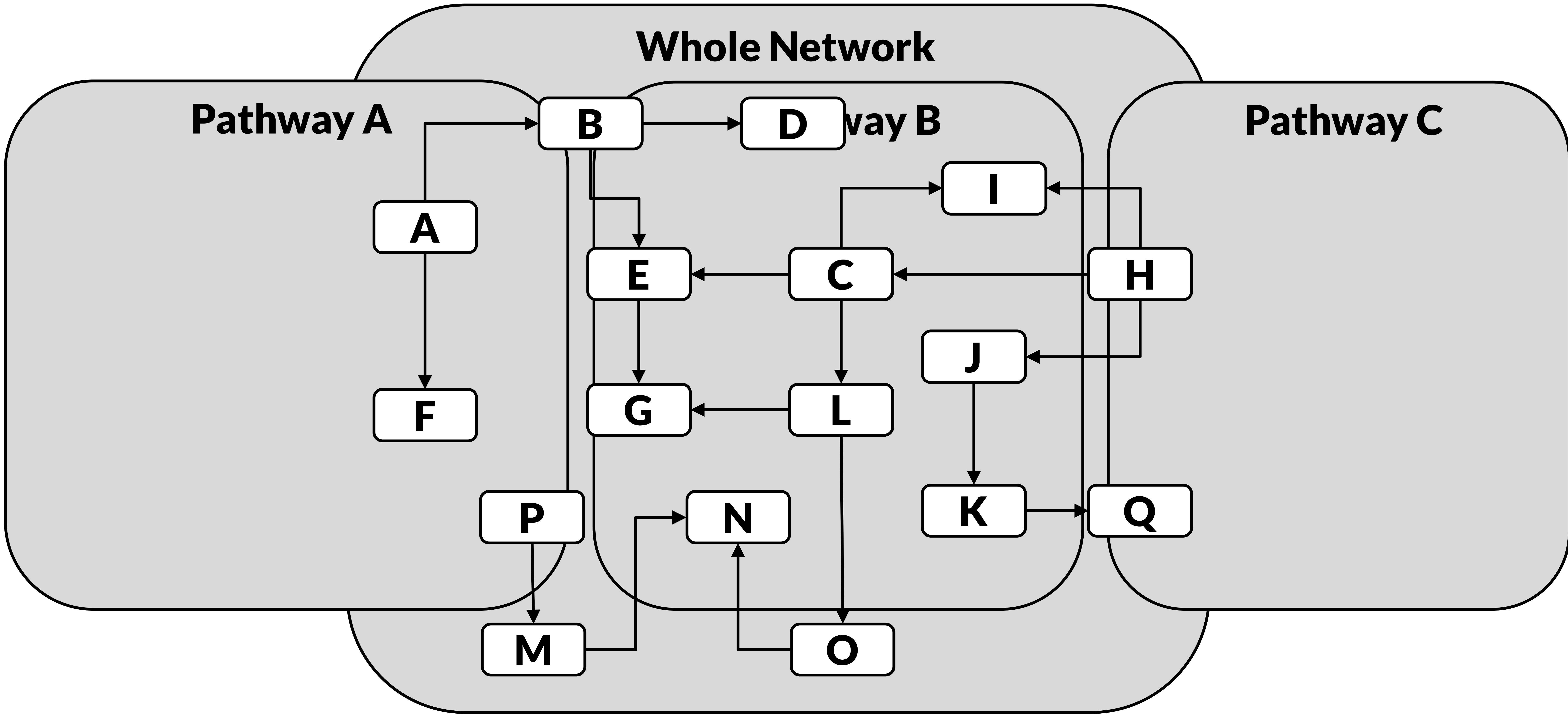
Glycolysis / Gluconeogenesis	Enzymes
------------------------------	---------



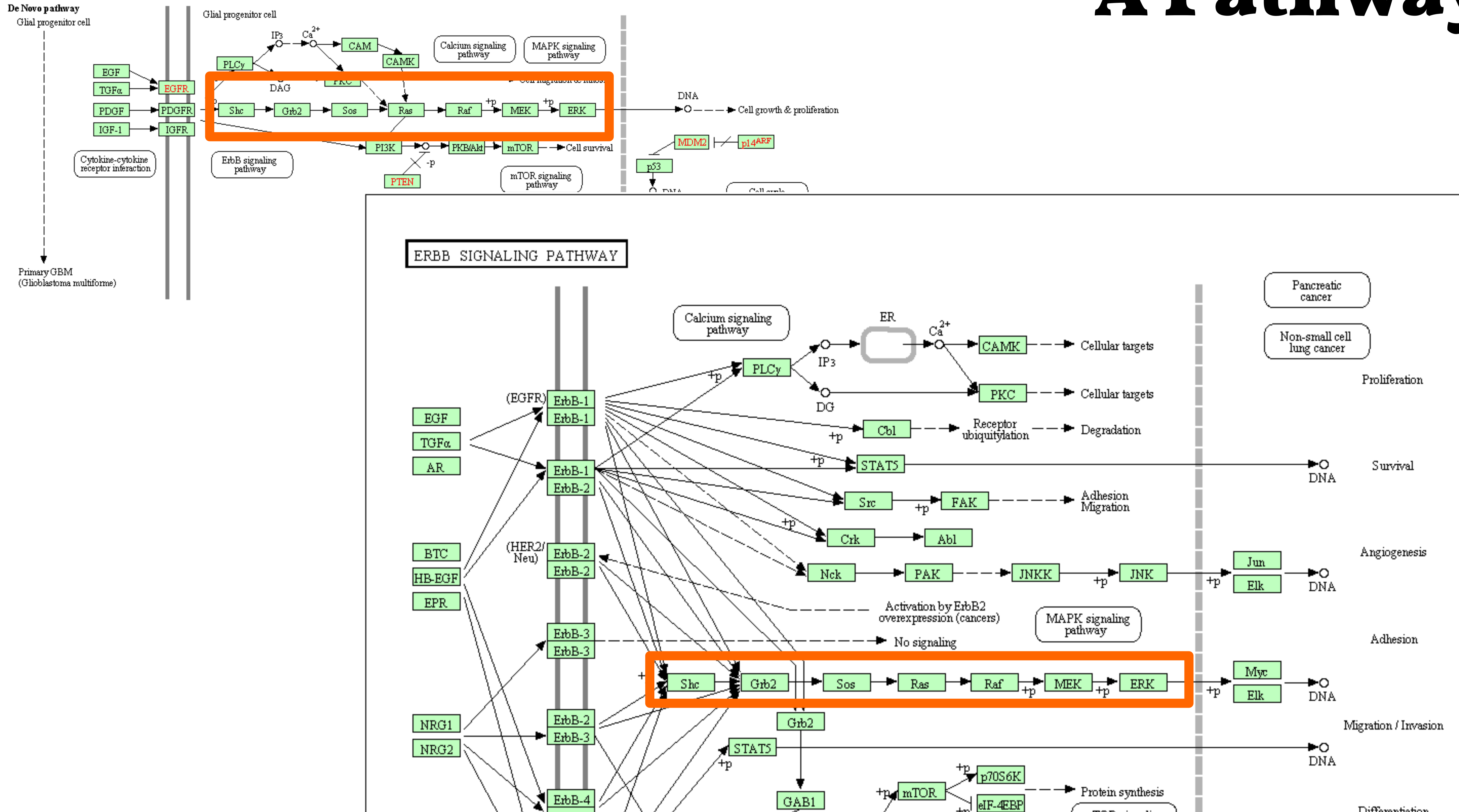
The bigger picture



Background



A Pathway

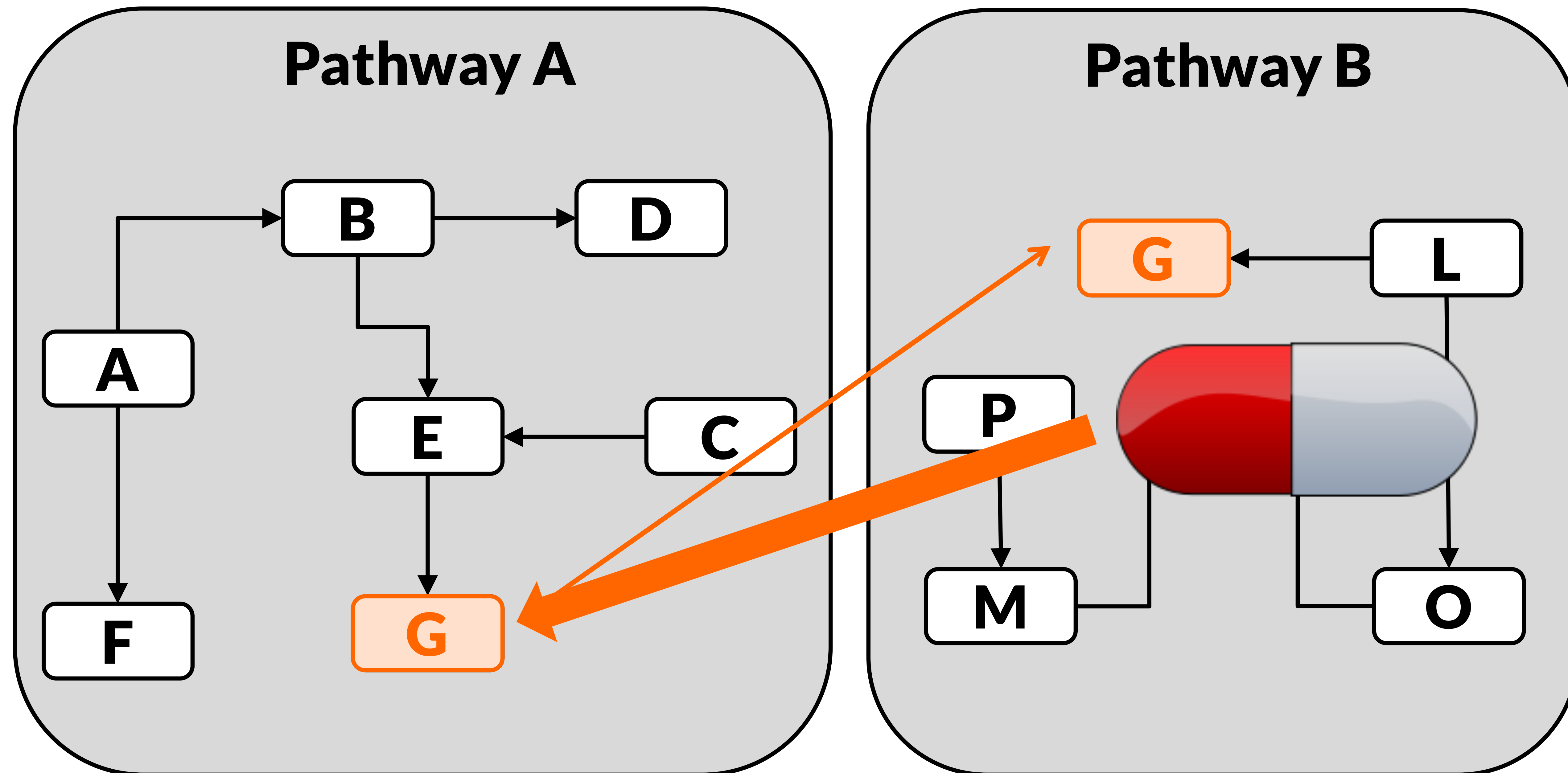


Two Problems - Two Solutions

1. Large Partitioned Network

2. Many Node Attributes

Large Partitioned Network

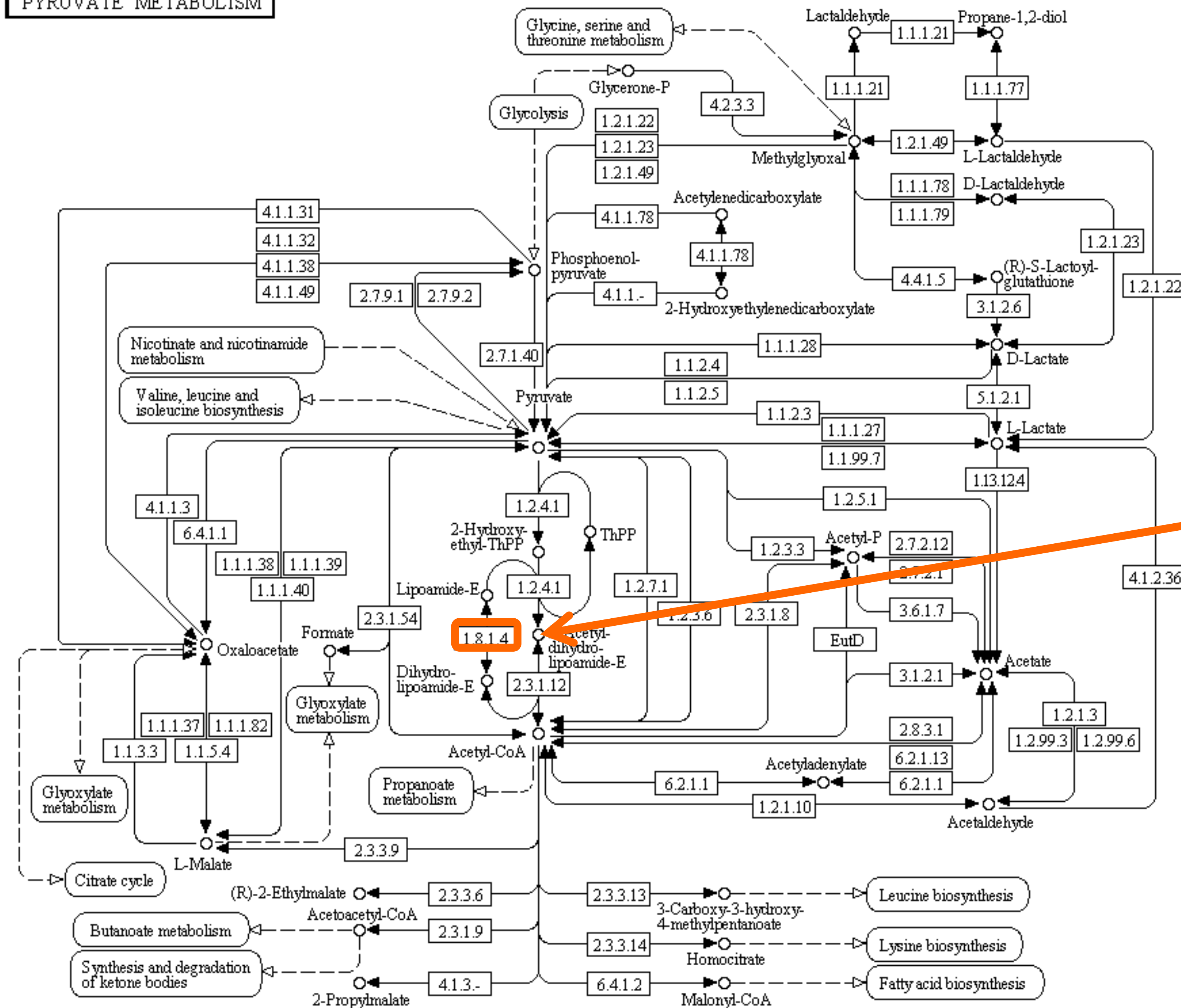


Drug side-effects

Drug repositioning

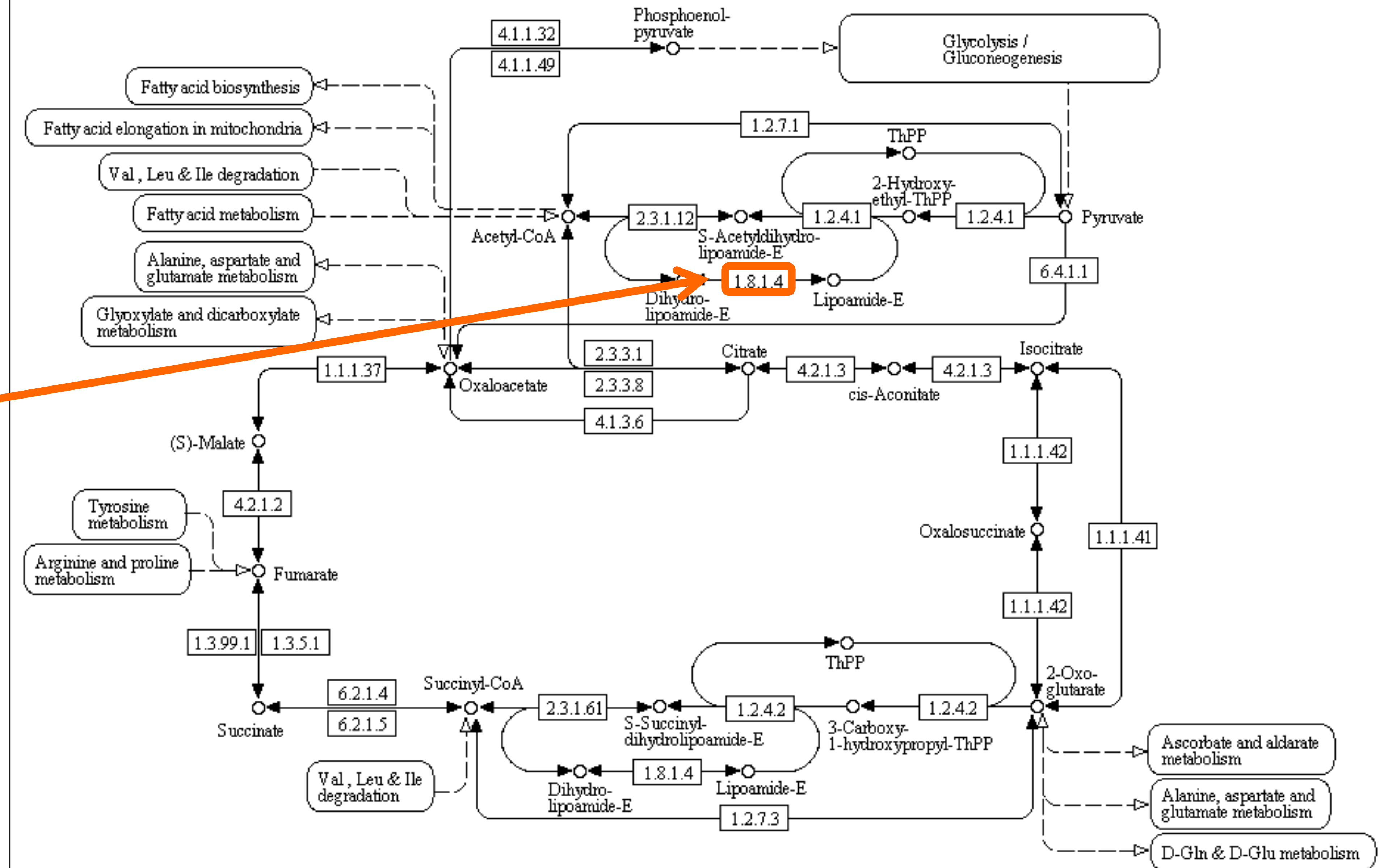
Large Partitioned Network

PYRUVATE METABOLISM



00620 6/6/12
(c) Kanehisa Laboratories

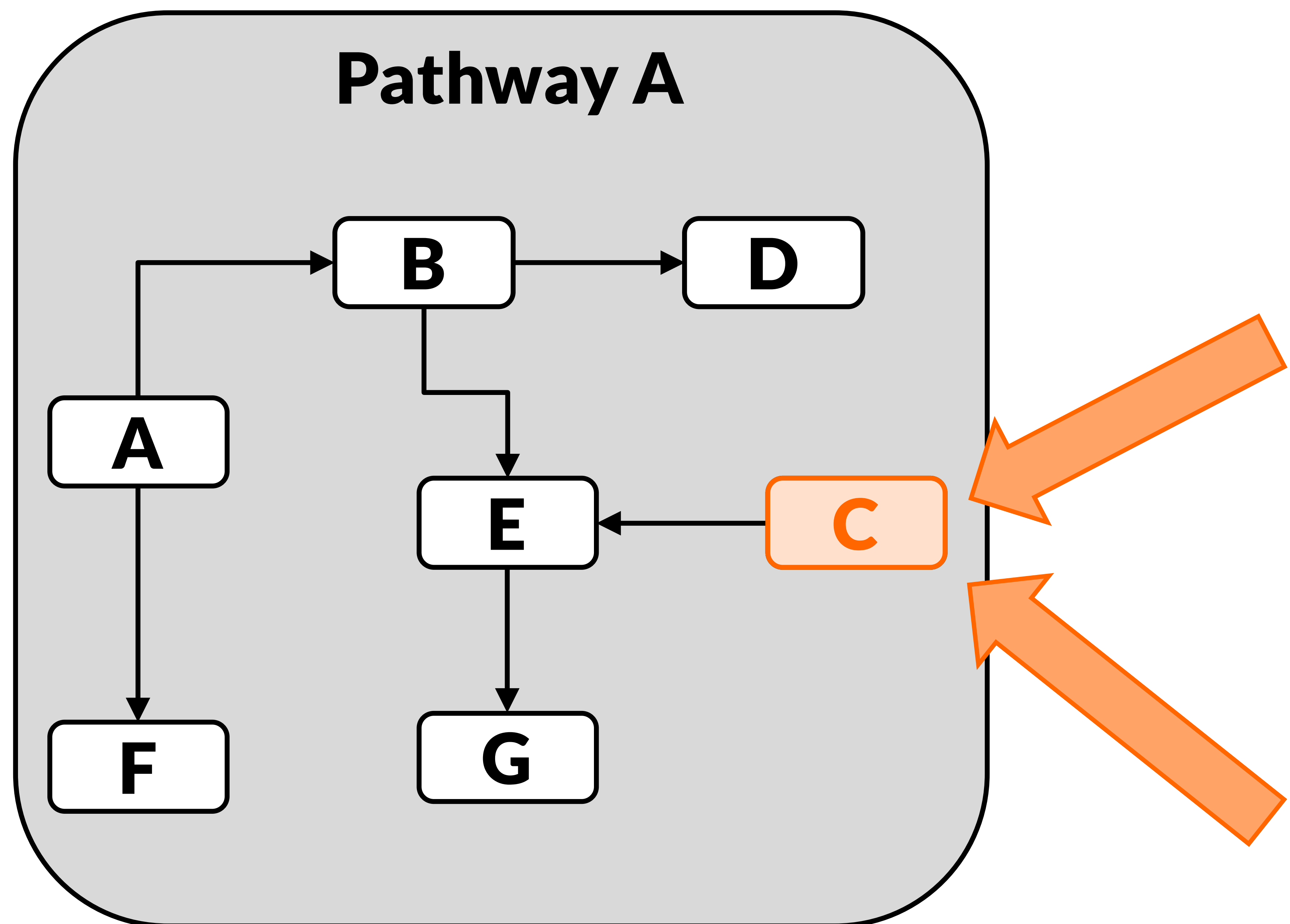
CITRATE CYCLE (TCA CYCLE)



00020 5/31/12
(c) Kanehisa Laboratories

How to visualize pathway relationships?

Many Node Attributes



Node	Sample 1	Sample 2	Sample 3	...
A	0.55	0.95	0.83	...
B	0.12	0.42	0.16	...
C	0.33	0.65	0.38	...
...	...	...	...	...

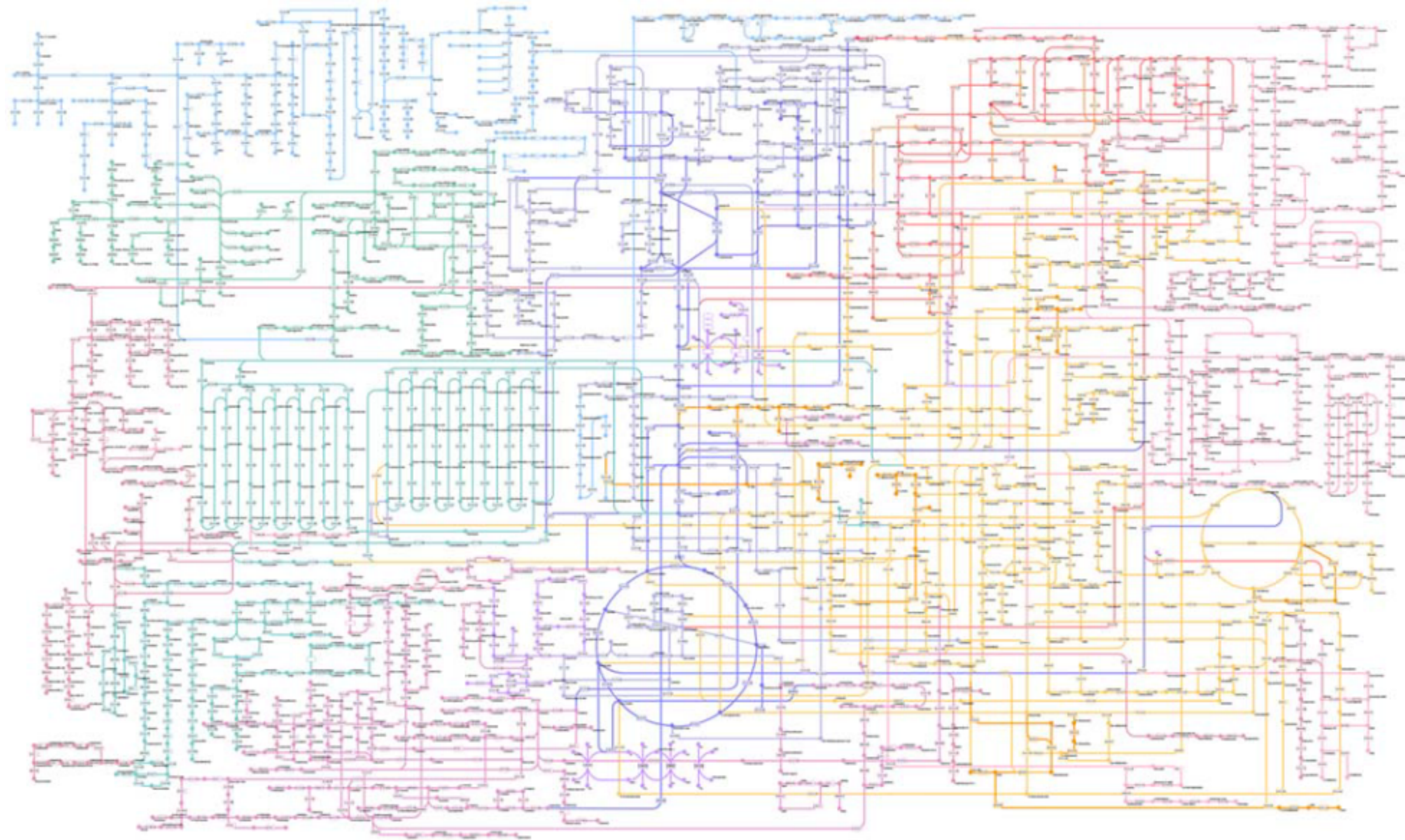
Node	Sample 1	Sample 2	Sample 3	...
A	low	low	very high	...
B	normal	low	high	...
C	high	very low	normal	...
...	...	...	...	...

How to visualize experimental data on pathways?

Visualizing Large Partitioned Networks

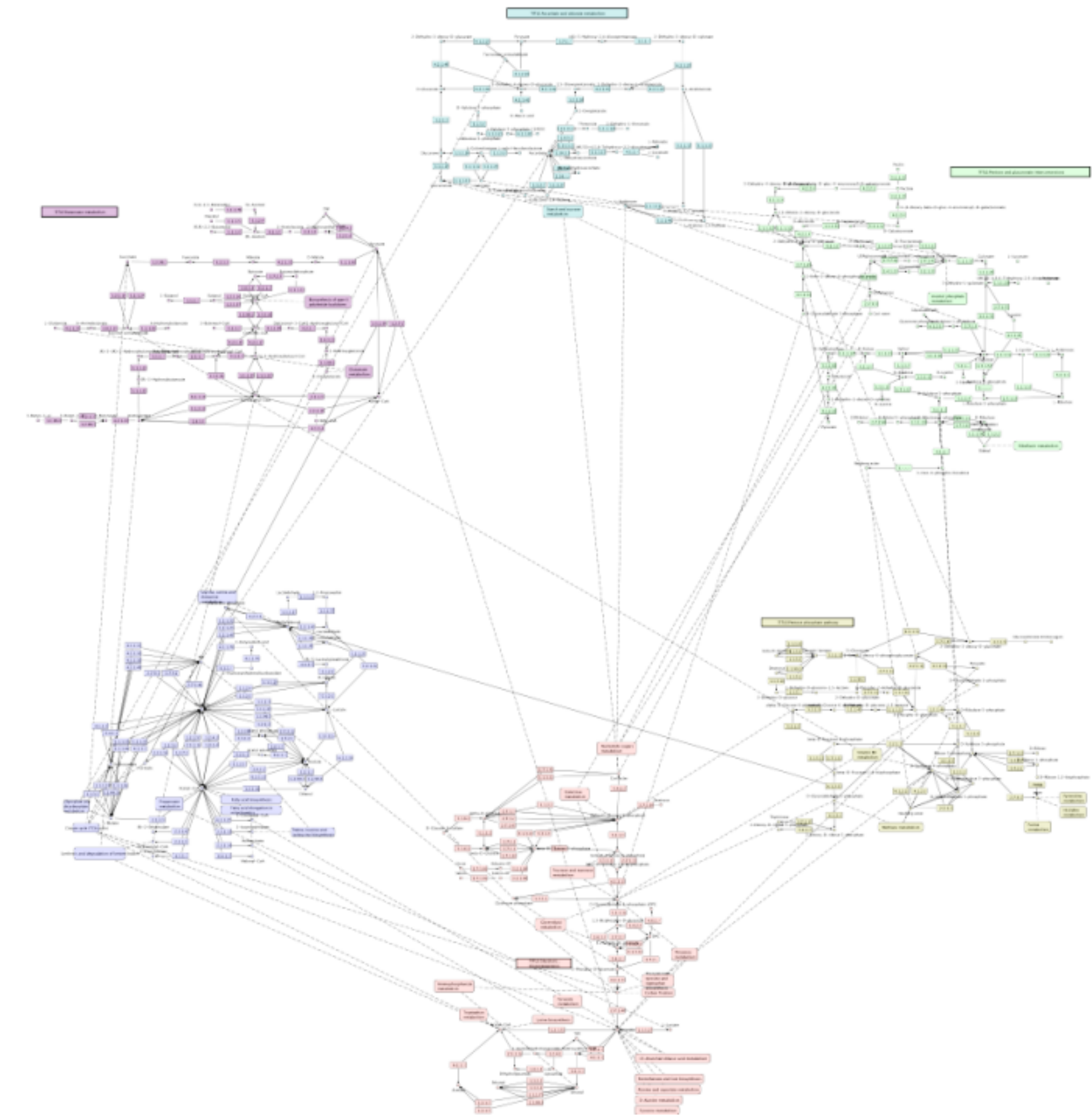
Approaches

Whole Network



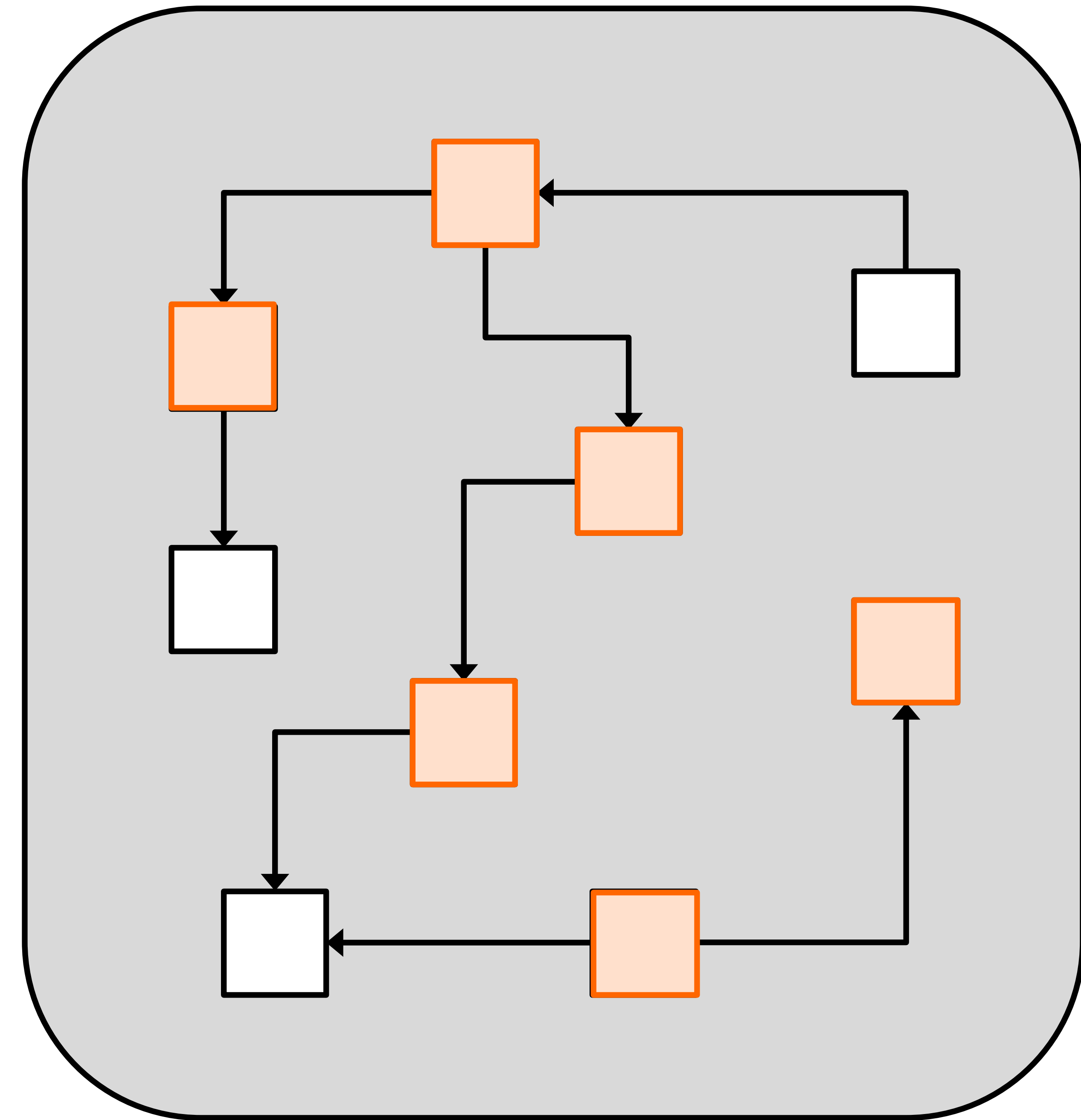
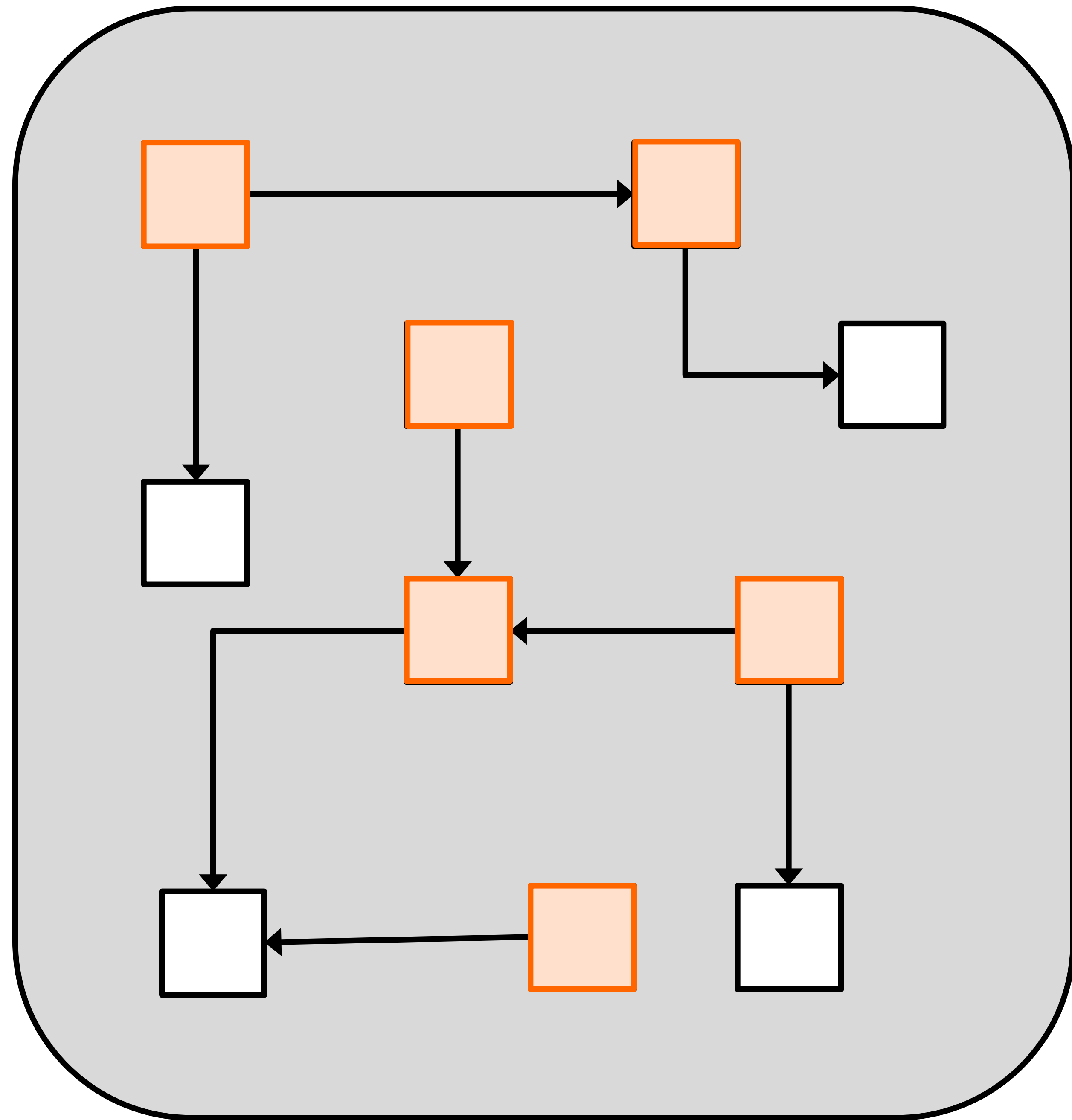
[Kono2009]

Connected Pathways

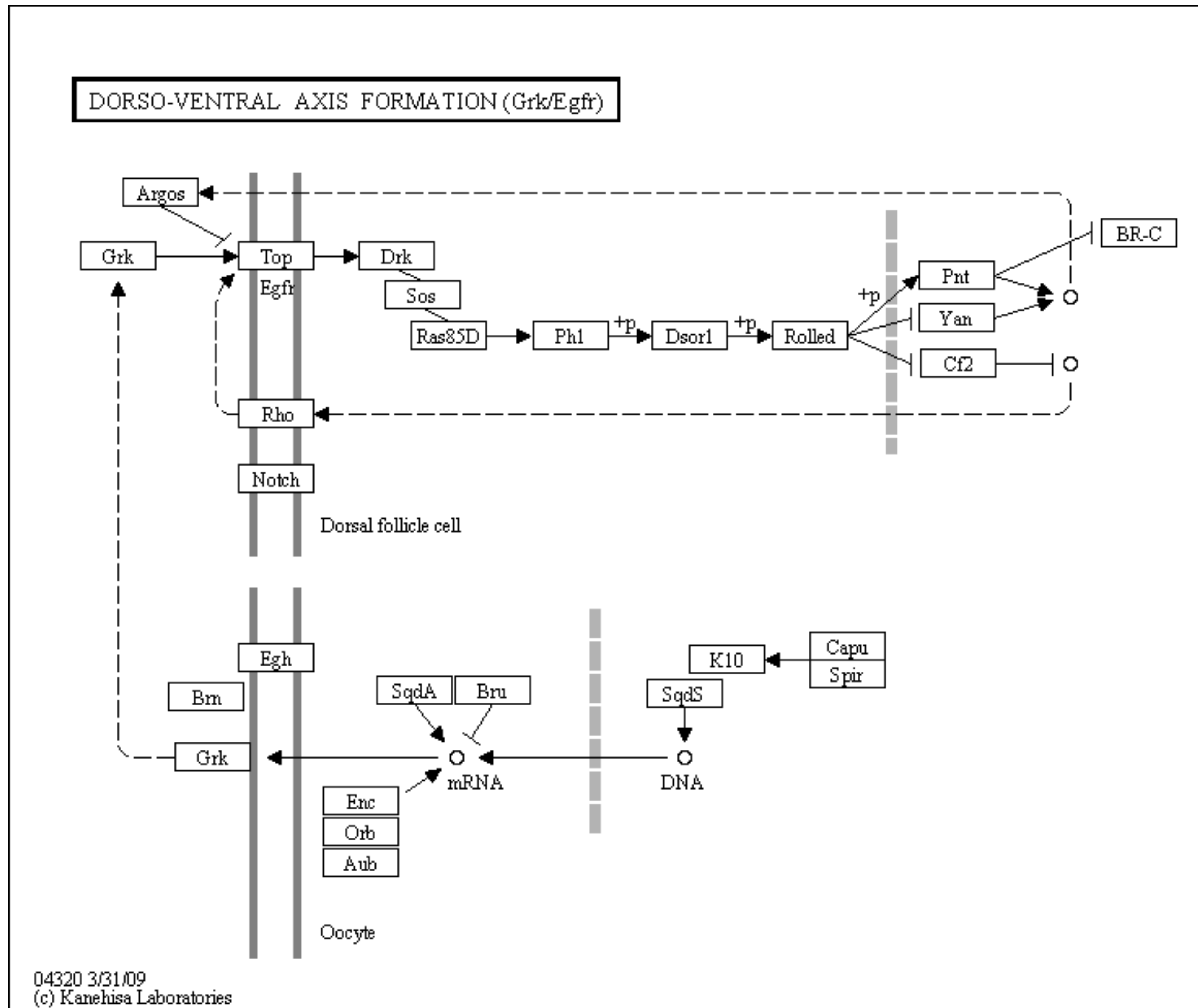


[Klukas2006]

Step 1: Finding Related Pathways

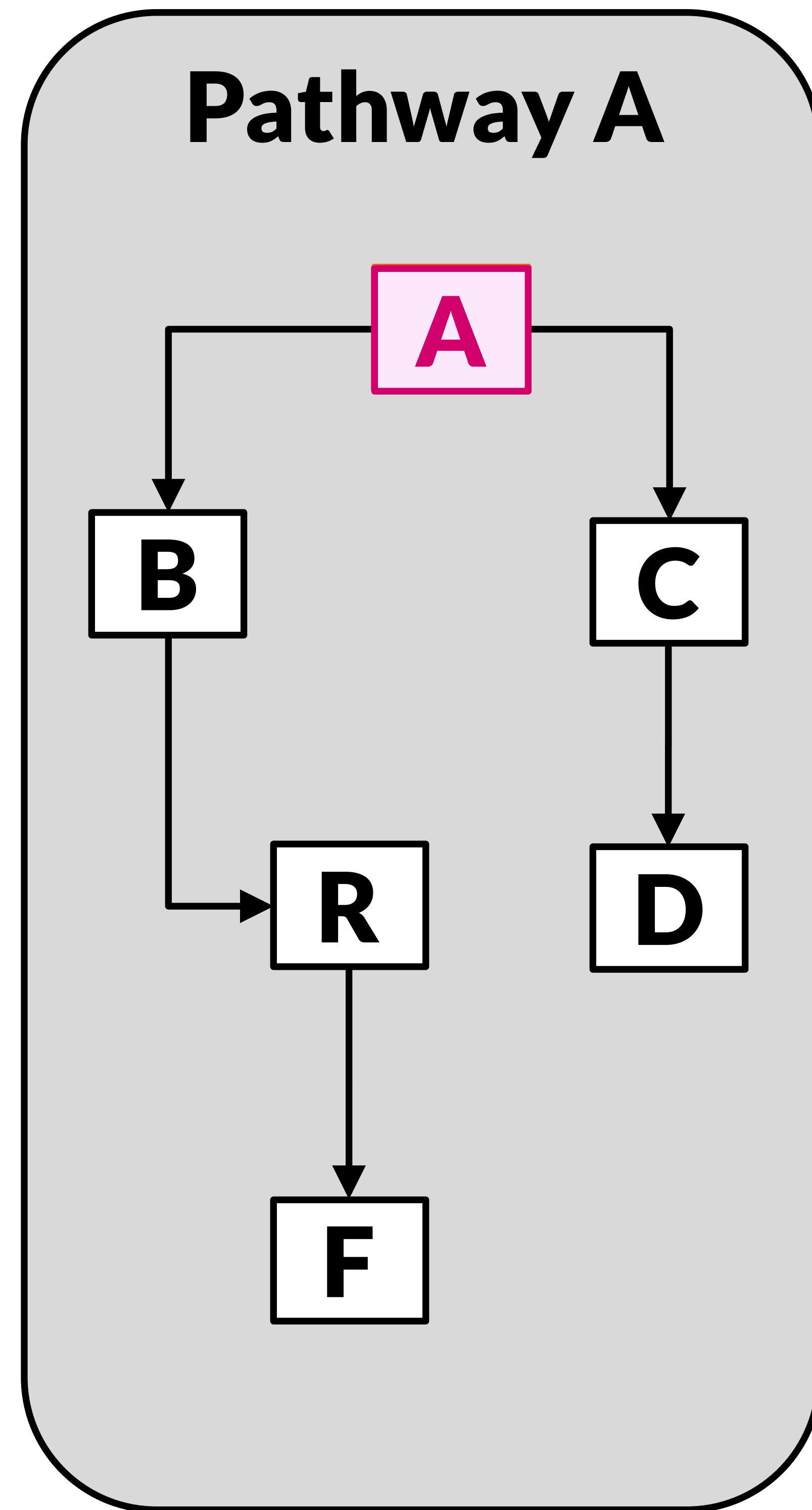


Finding Related Pathways

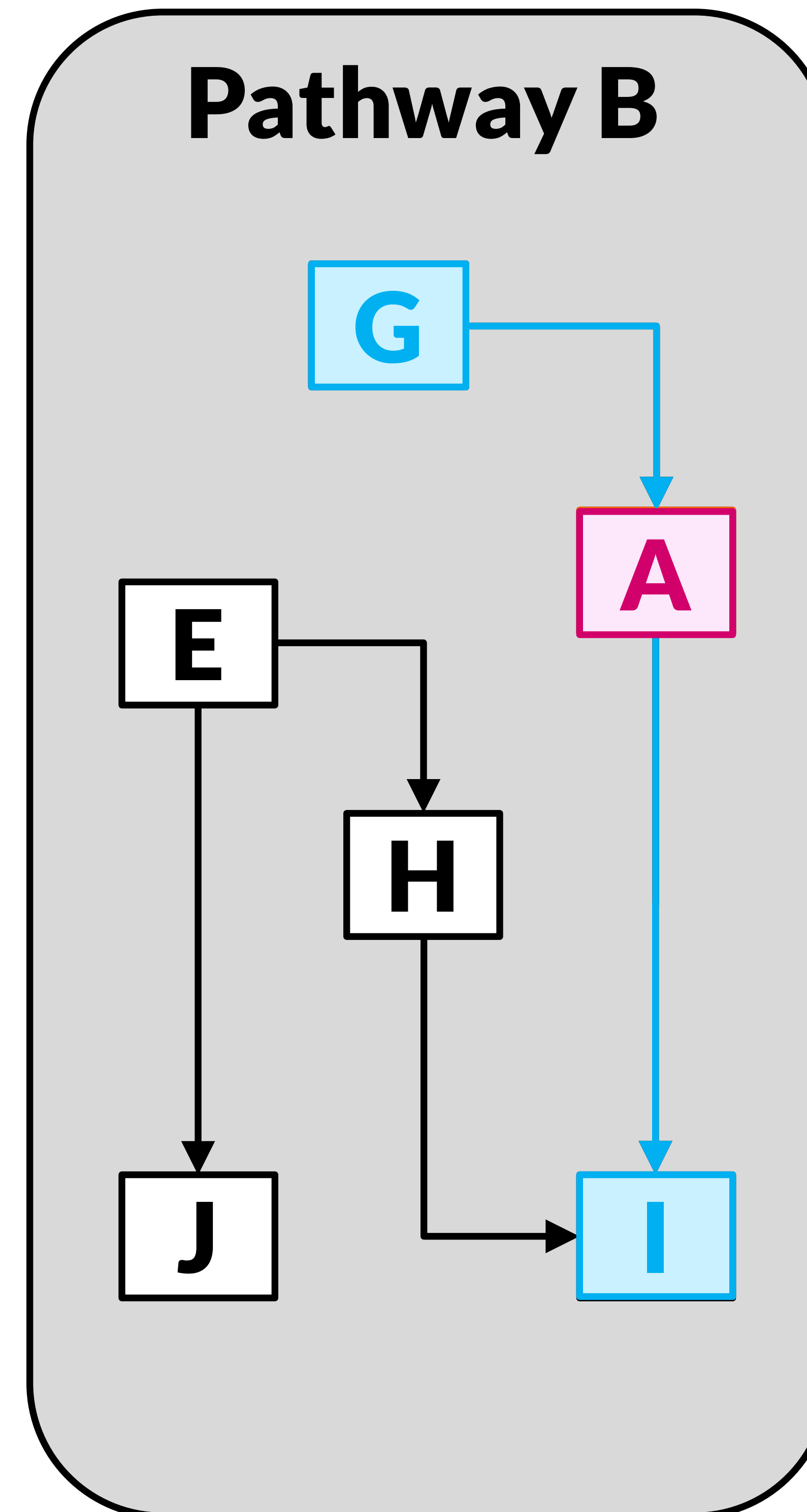


Pathways	
Pathway	Commo
Nifedipine Activity	
Endometrial cancer	
Thyroid cancer	
Melanoma	
Acute myeloid leukem	
Signaling of Hepatocy	
Non-small cell lung c	
Renal cell carcinoma	
Bladder cancer	
Gap junction	
GnRH signaling path	
Glioma	
IL-9 signaling pathwa	
Progesterone-mediate	
Estrogen signaling pa	
EPO Receptor Signali	
Prostate cancer	
IL-5 signaling pathwa	
Chronic myeloid leuk	
ErbB signaling pathw	
Fc epsilon RI signali	
Osteopontin Signaling	
Serotonin Receptor 2	
ErbB signaling pathw	
IL-3 Signaling Pathwa	
B cell receptor signal	
Prion diseases	
Prolactin signaling pa	
Focal adhesion	

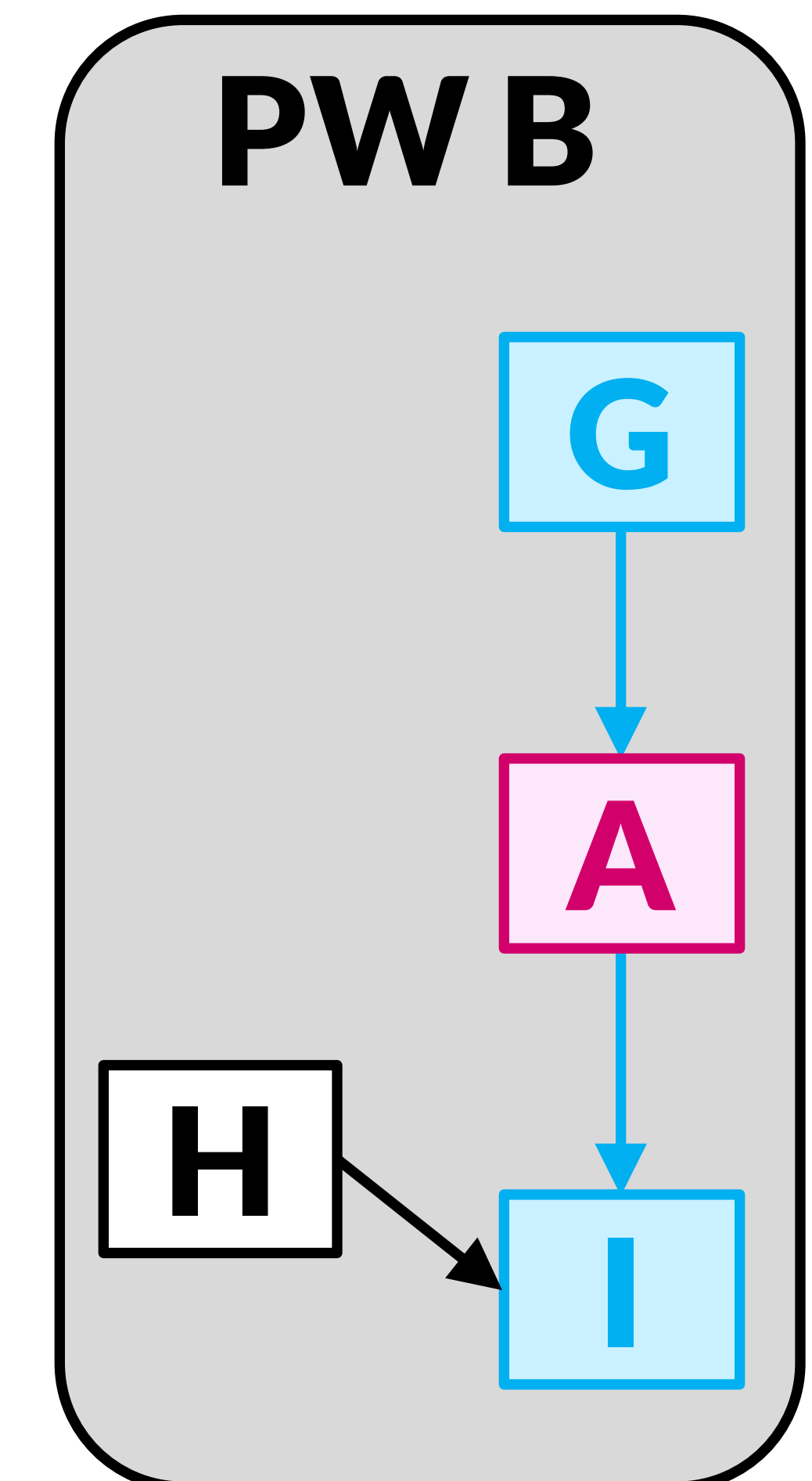
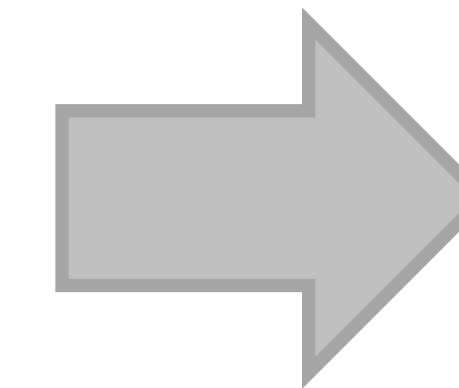
Step 2: Managing Display Space



Focus Pathway

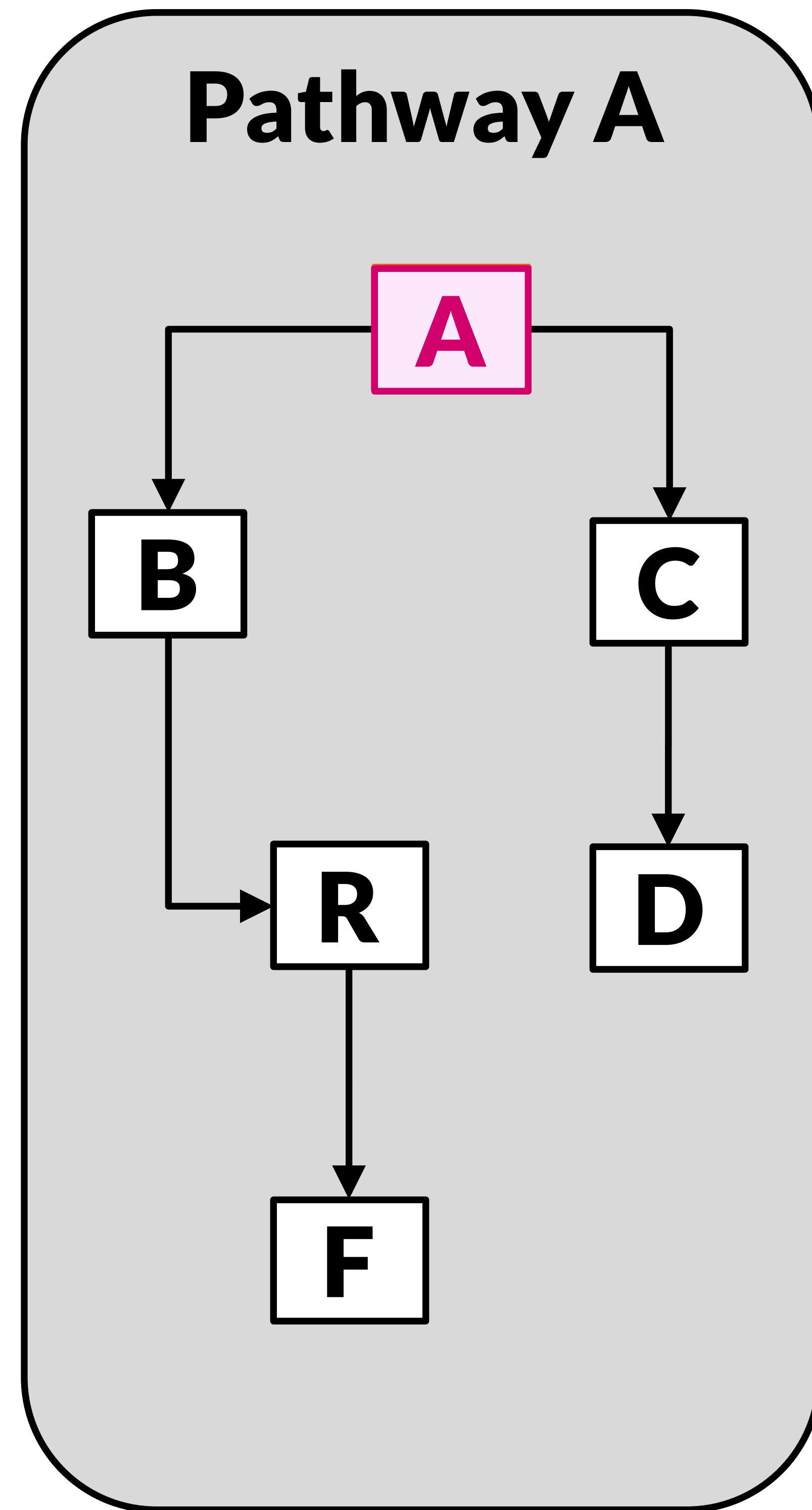


Context Pathway

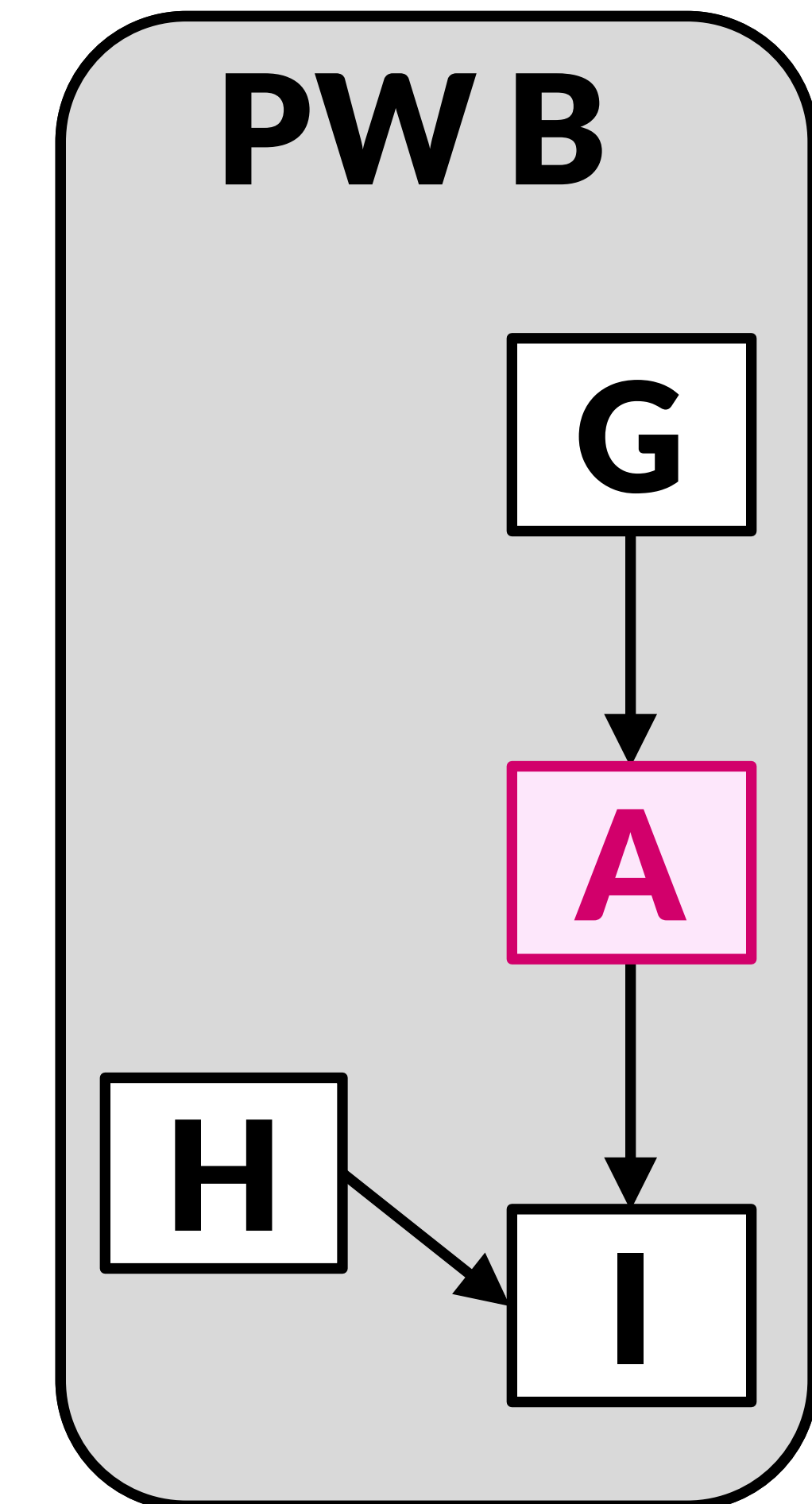
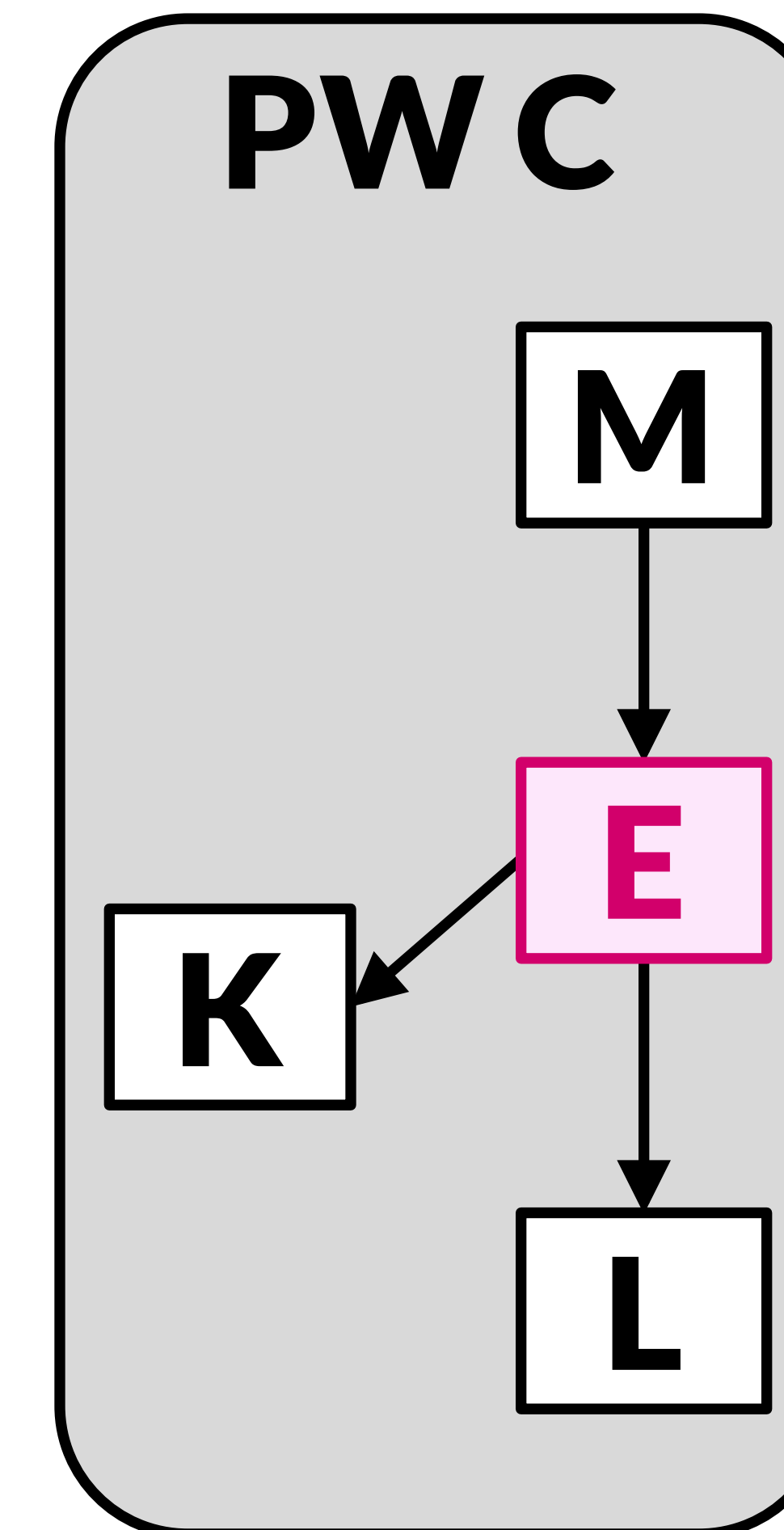


Contextual Subset

Step 2: Managing Display Space



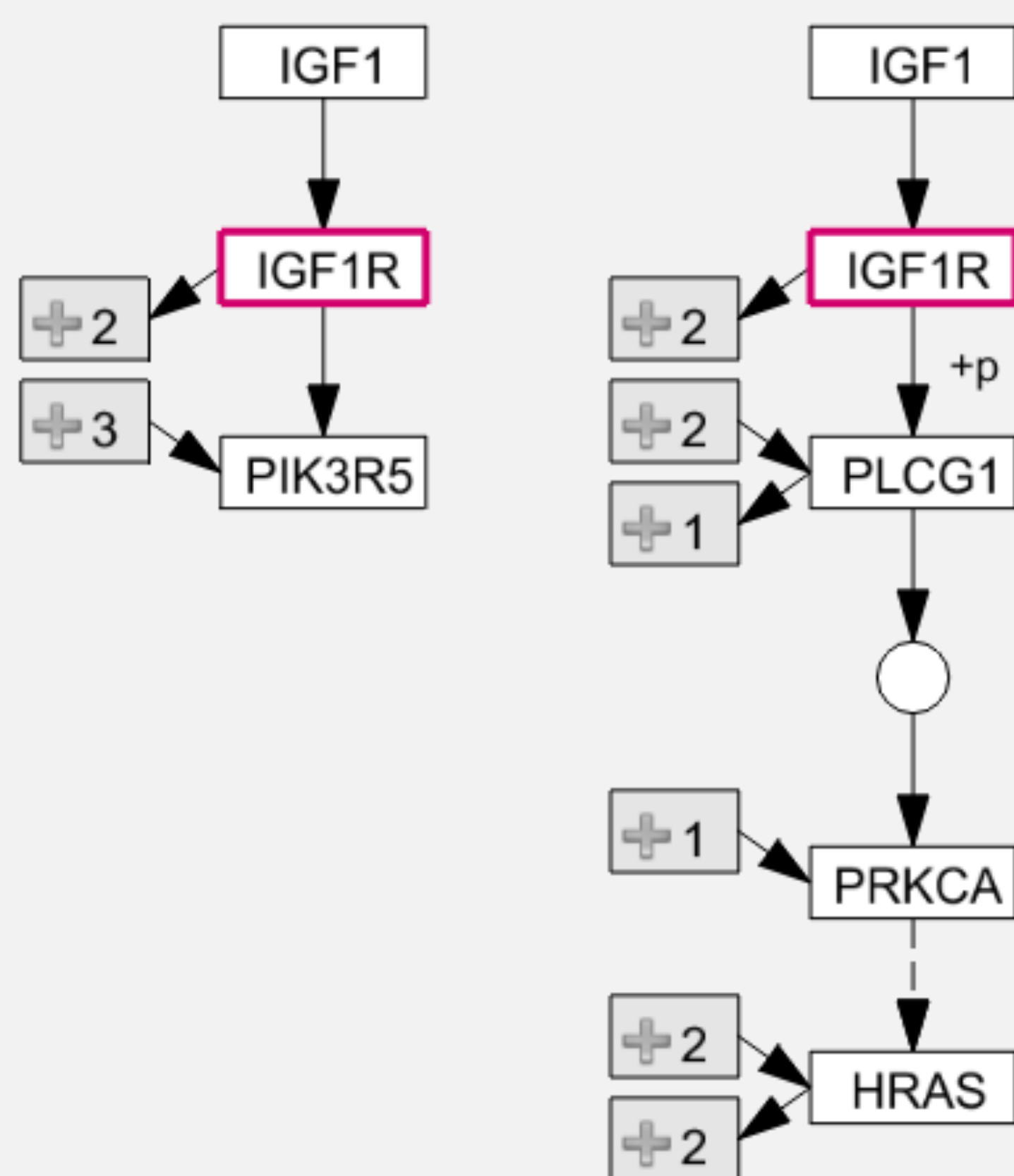
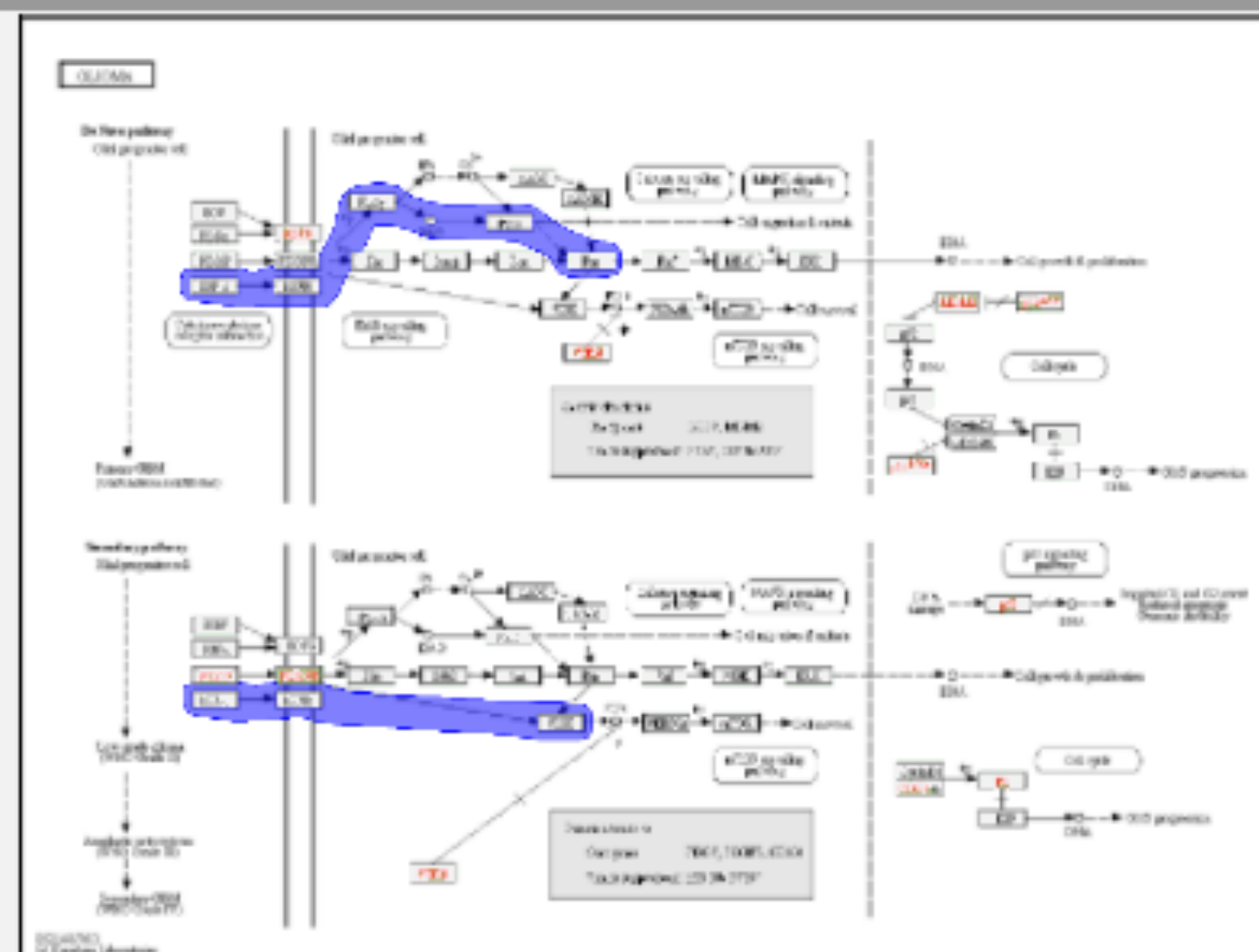
Focus Pathway



Context Pathways

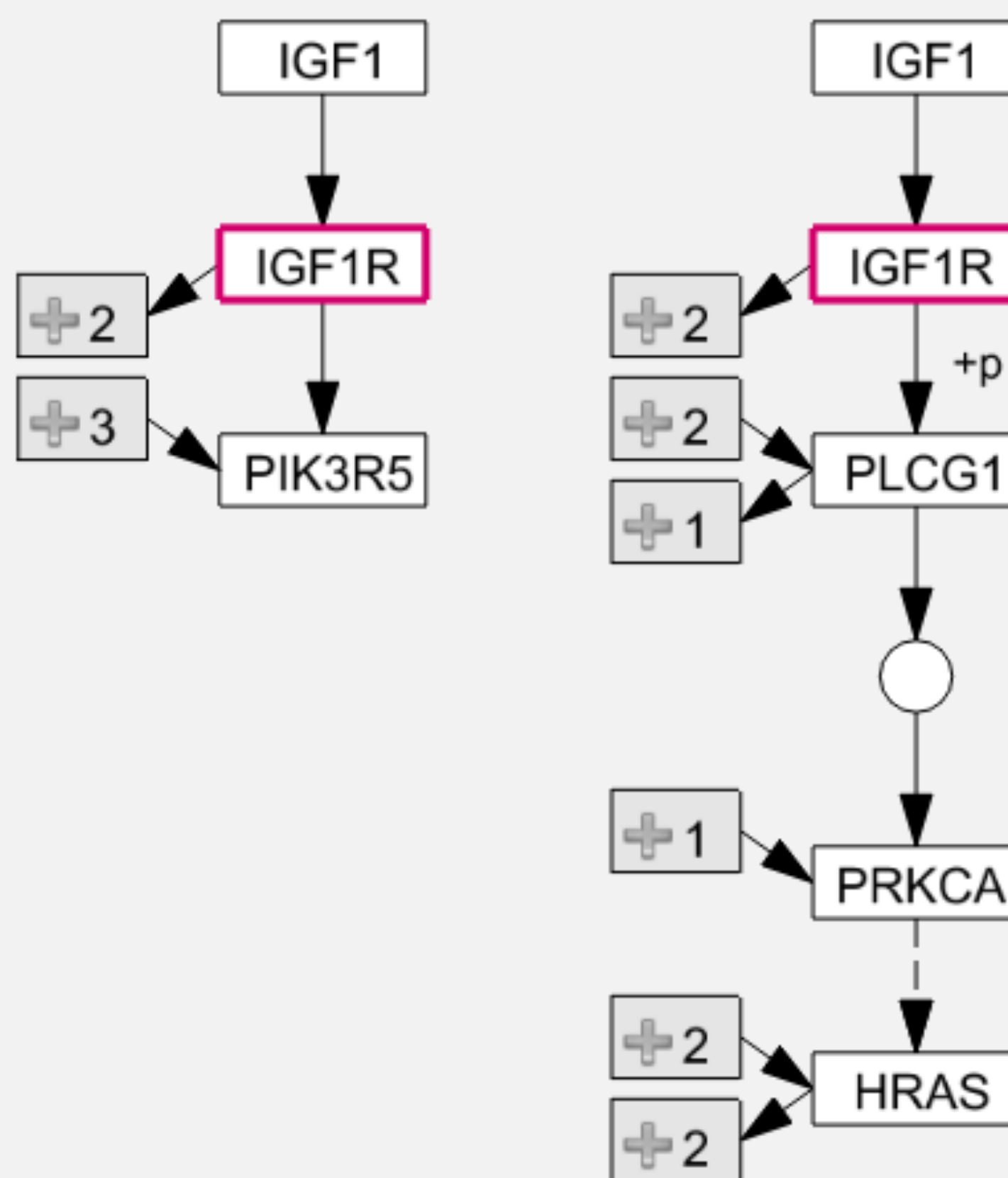
Levels of Detail

Glioma



High

Glioma



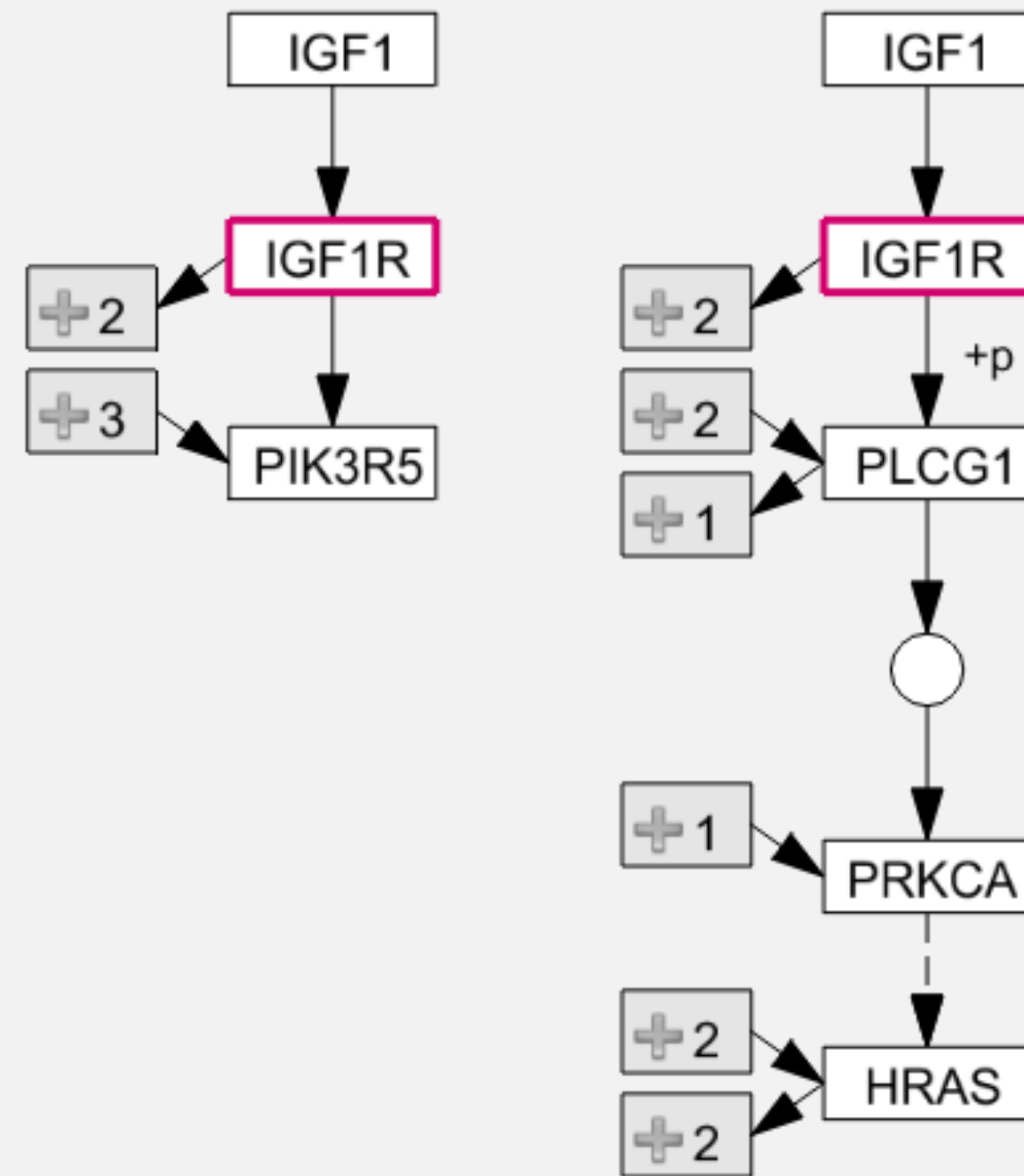
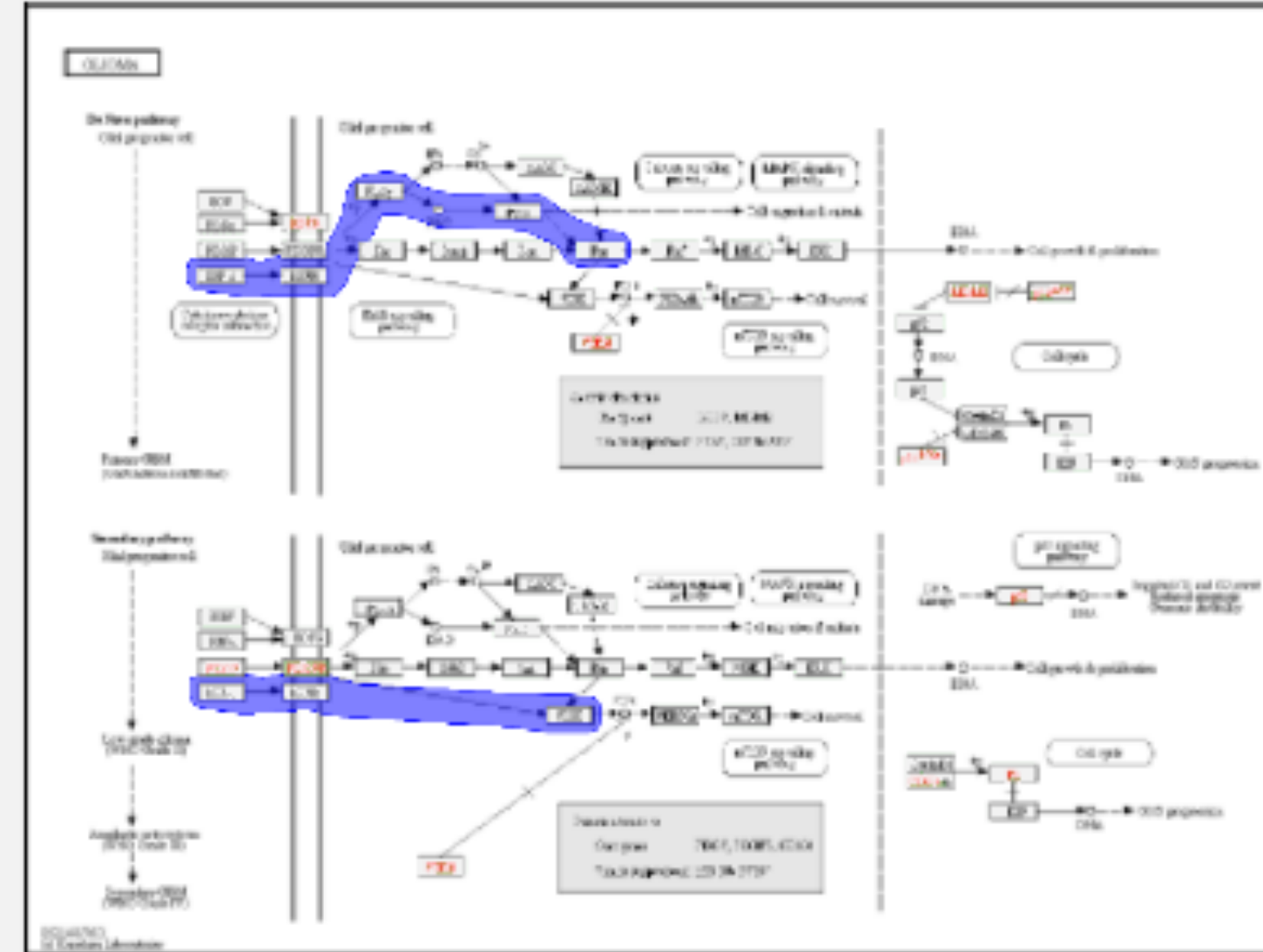
Medium

Glioma

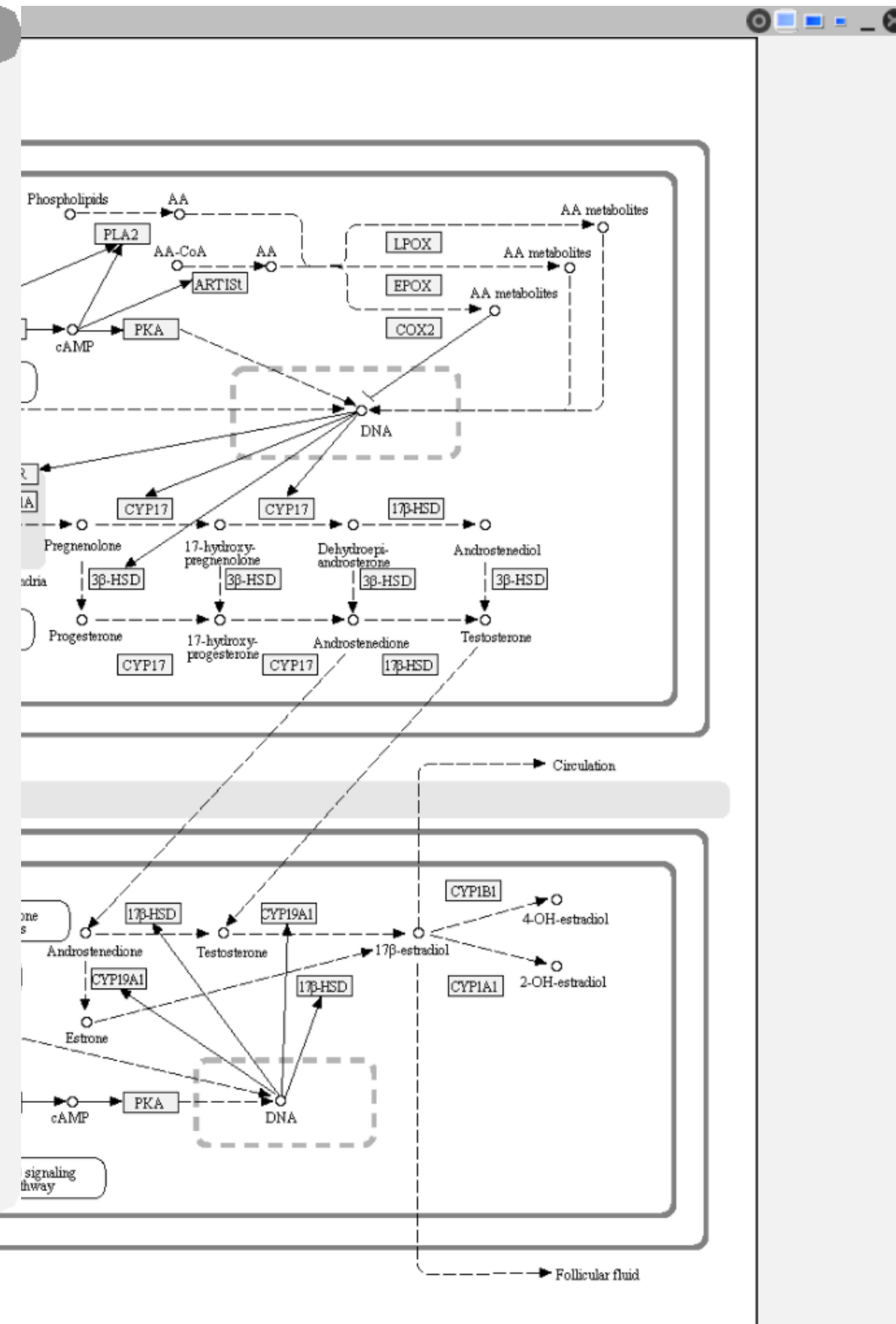
Low

Layout

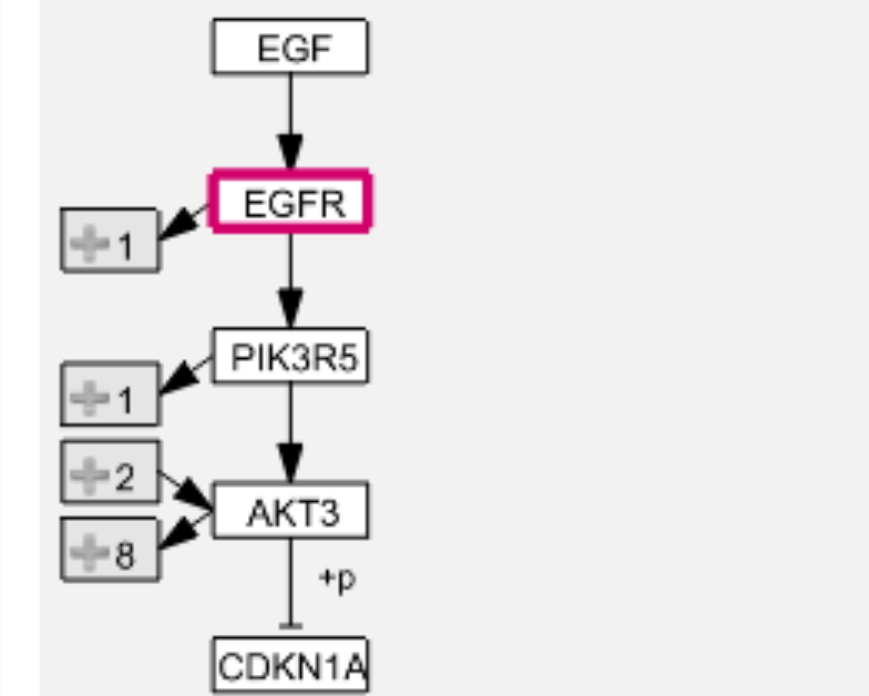
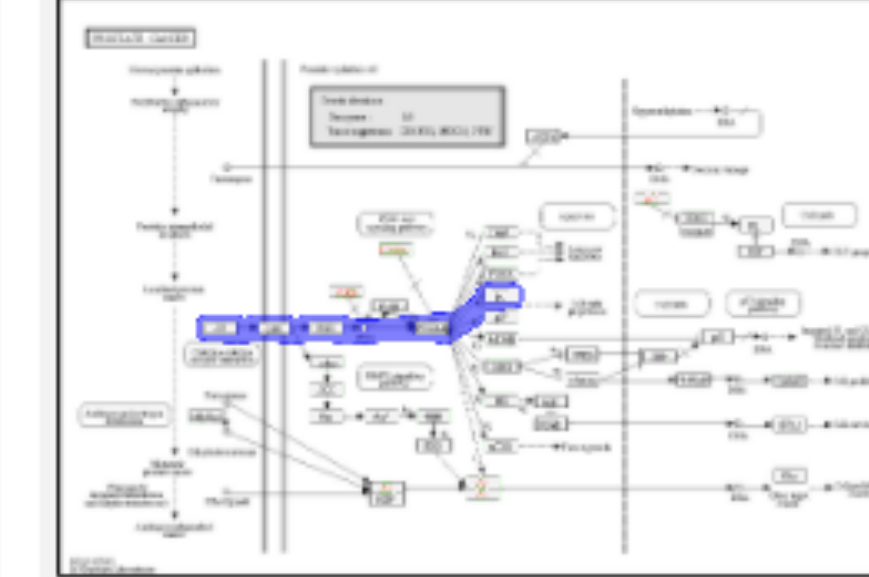
Glioma



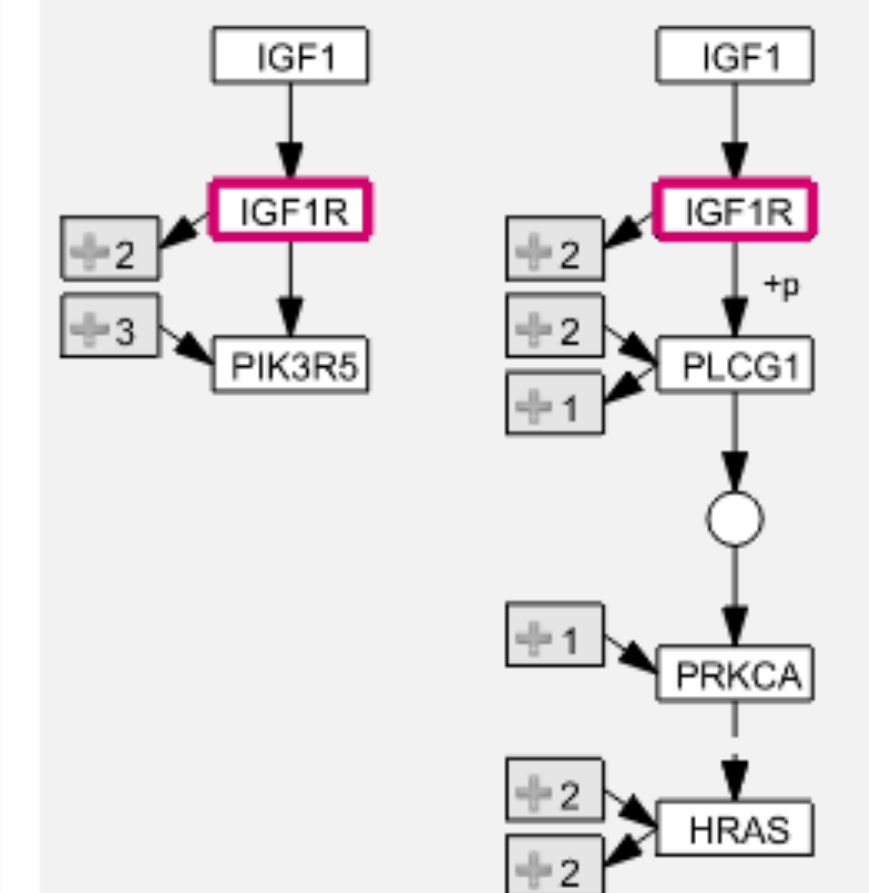
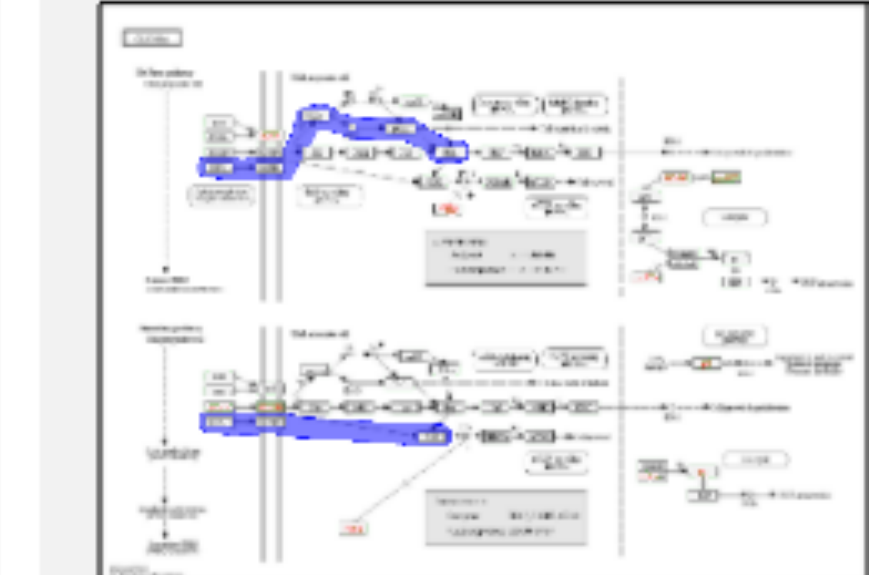
04913 4/22/13
(c) Kanehisa Laboratories



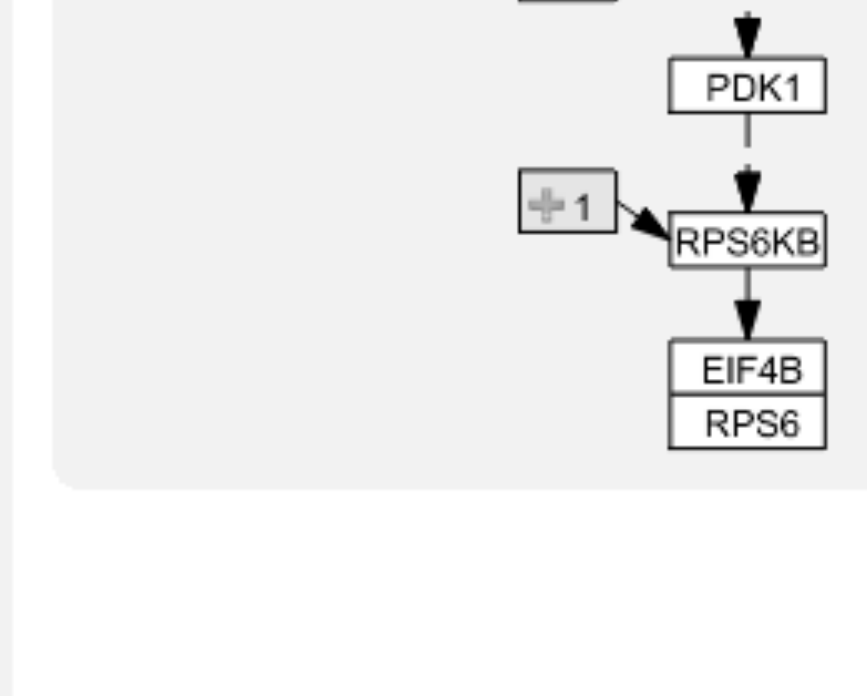
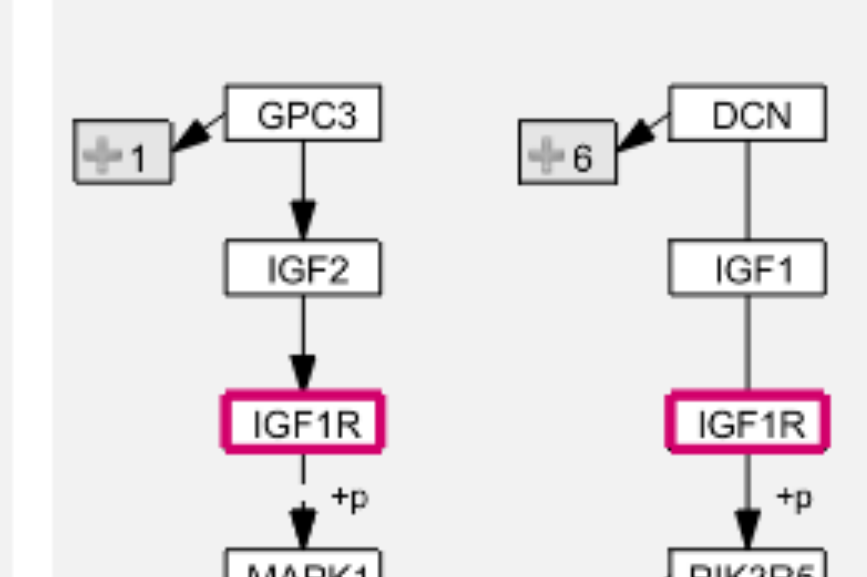
Prostate cancer



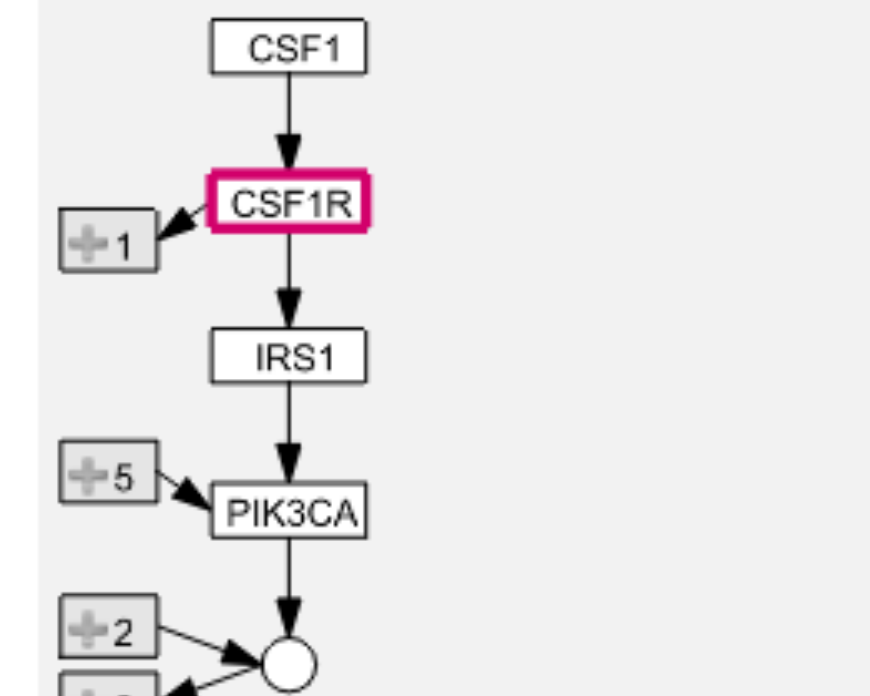
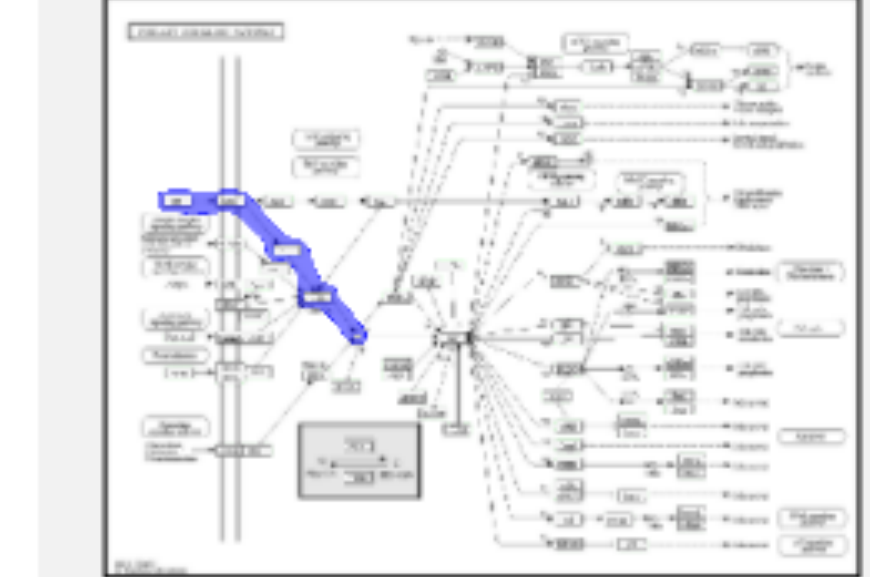
Glioma



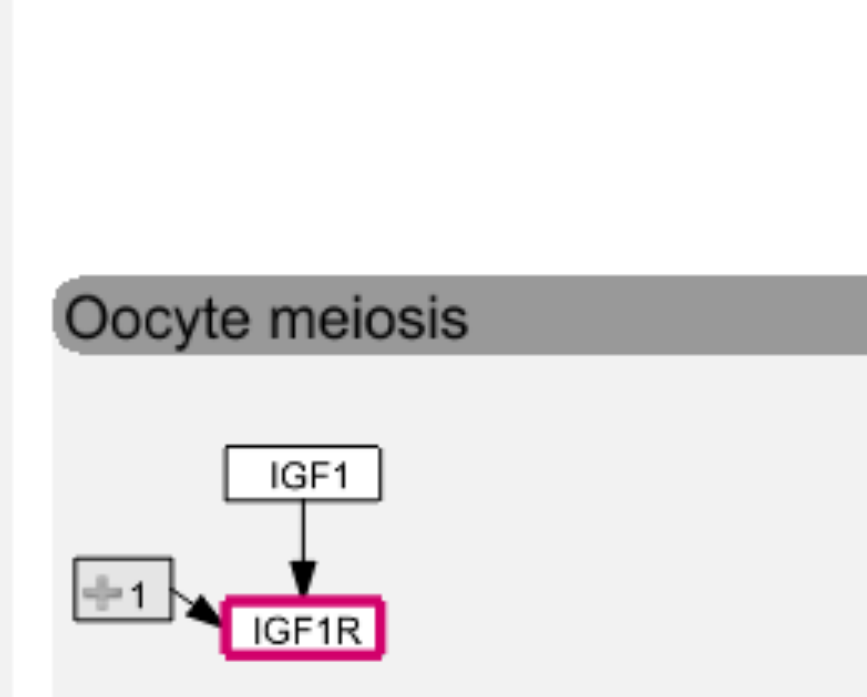
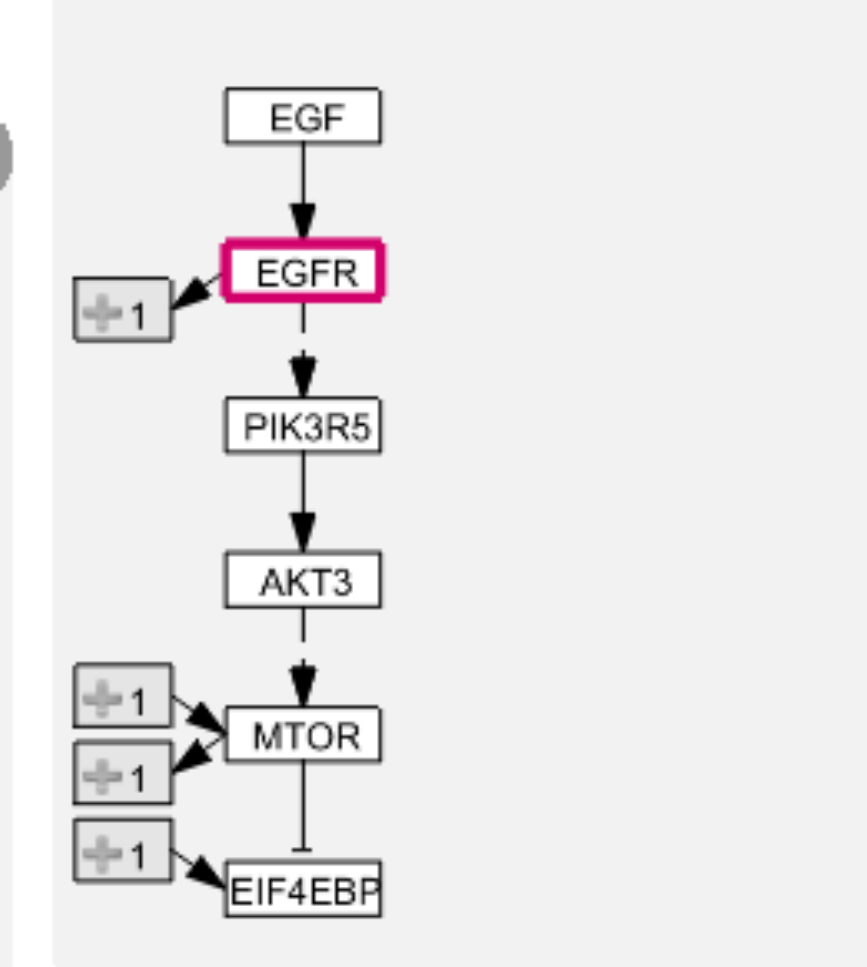
Proteoglycans in cancer



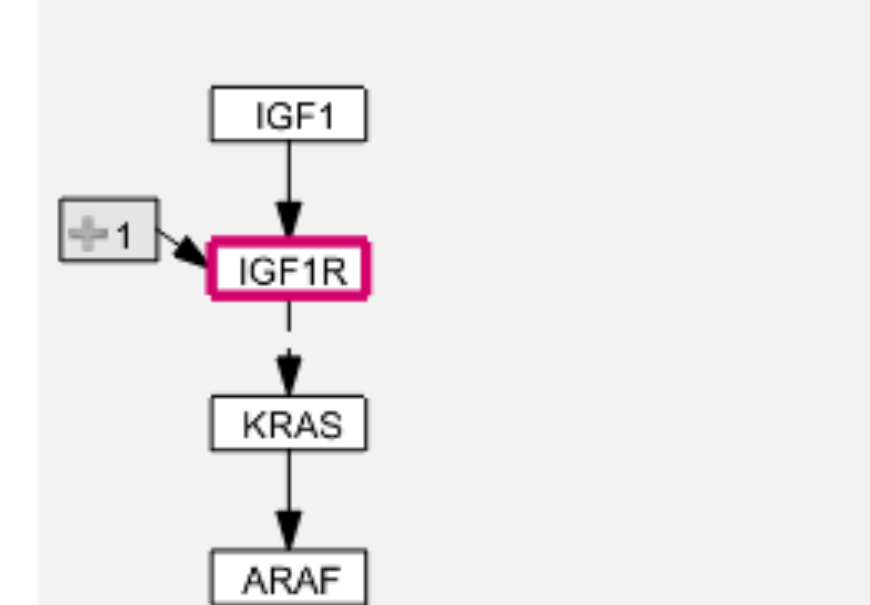
PI3K-Akt signaling pathway



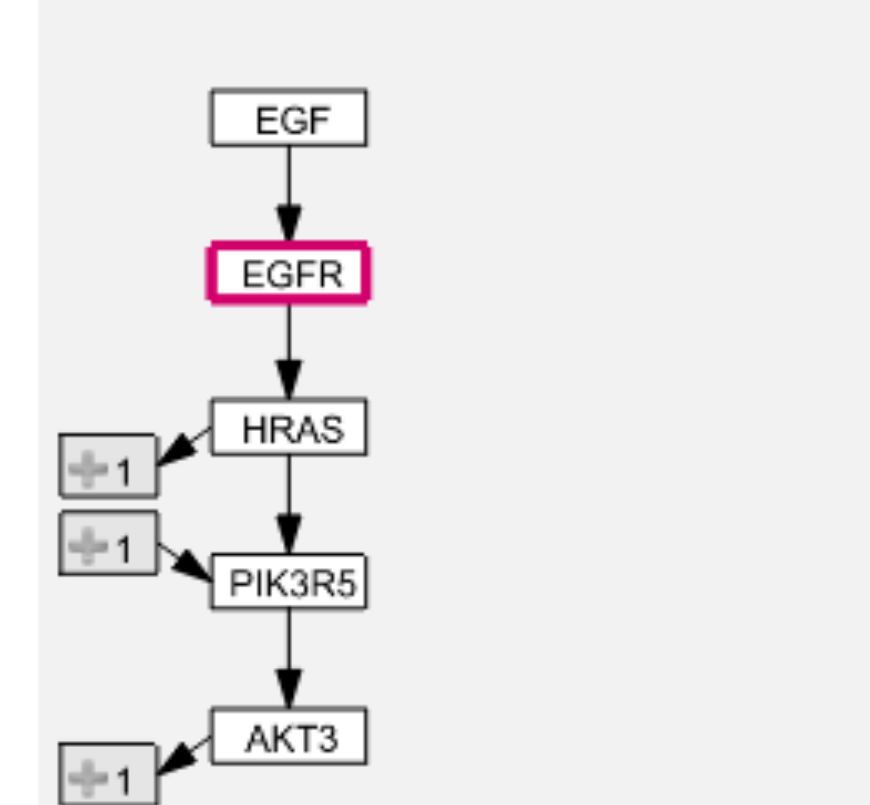
HIF-1 signaling pathway



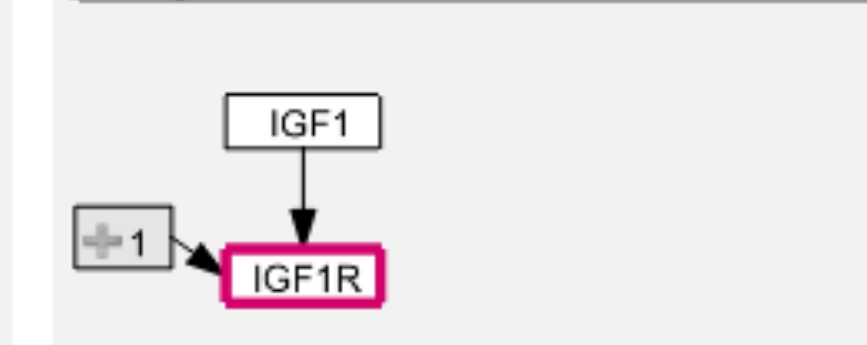
Progesterone-mediated oocyte



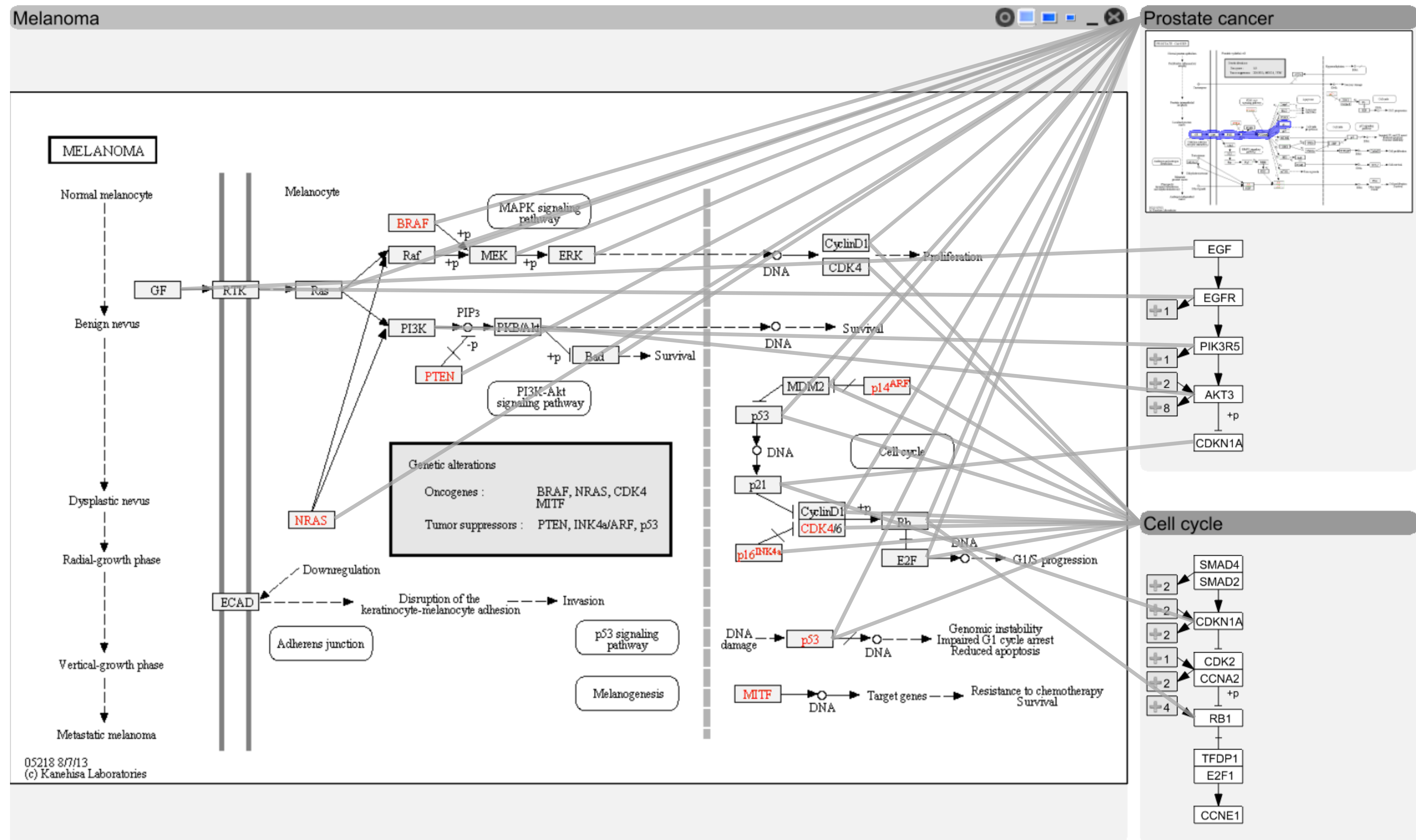
Melanoma



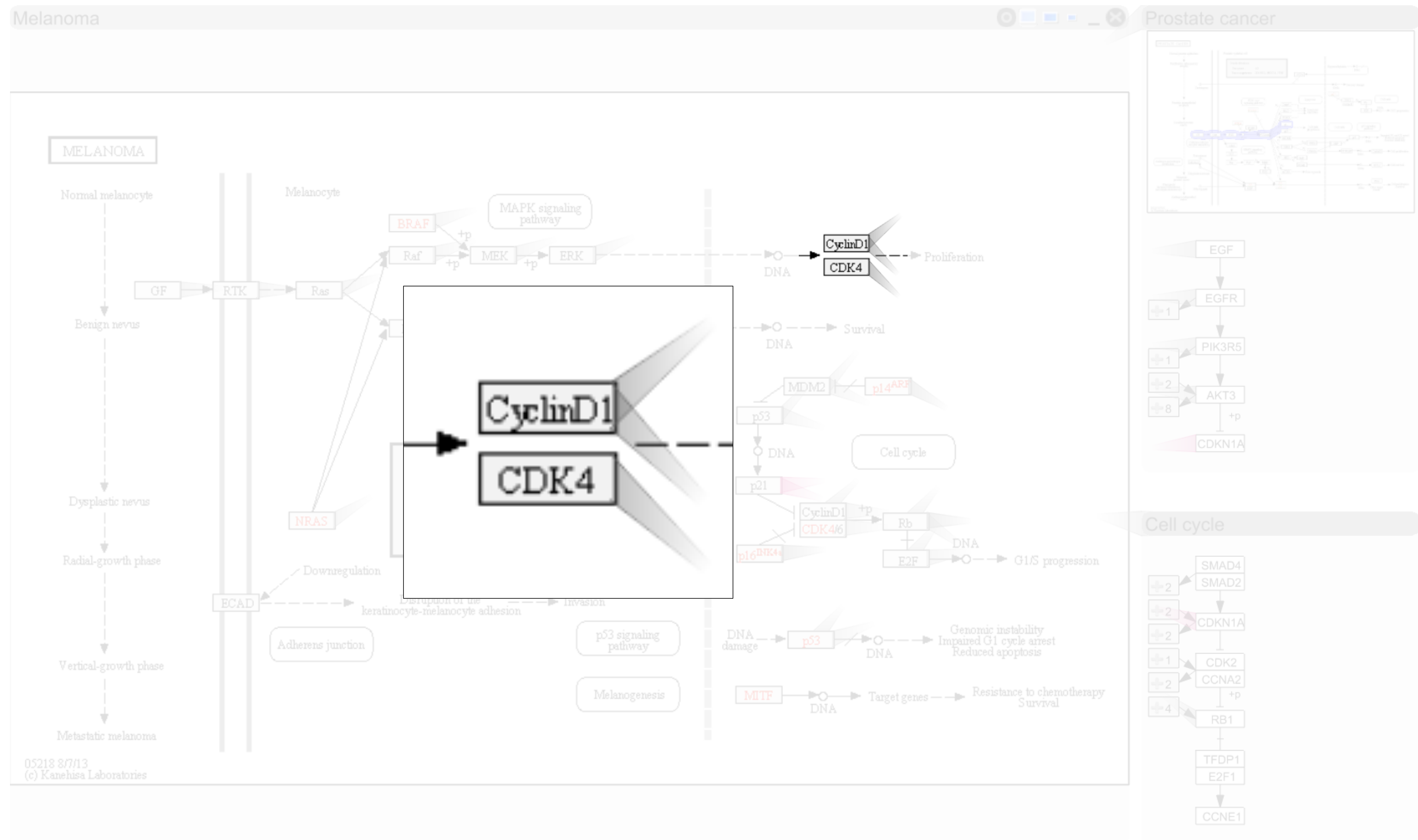
Oocyte meiosis



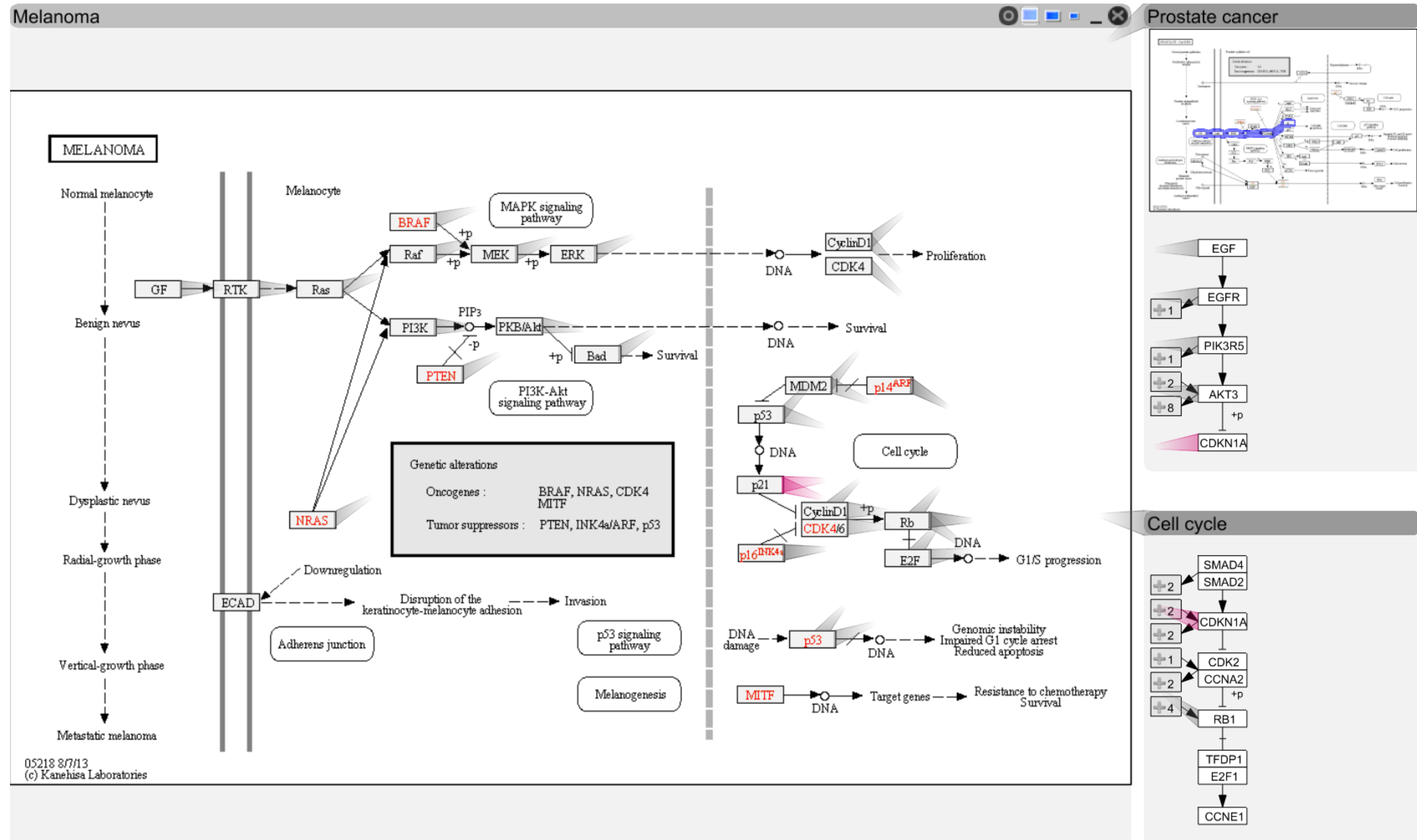
Step 3: Visualizing Relationships



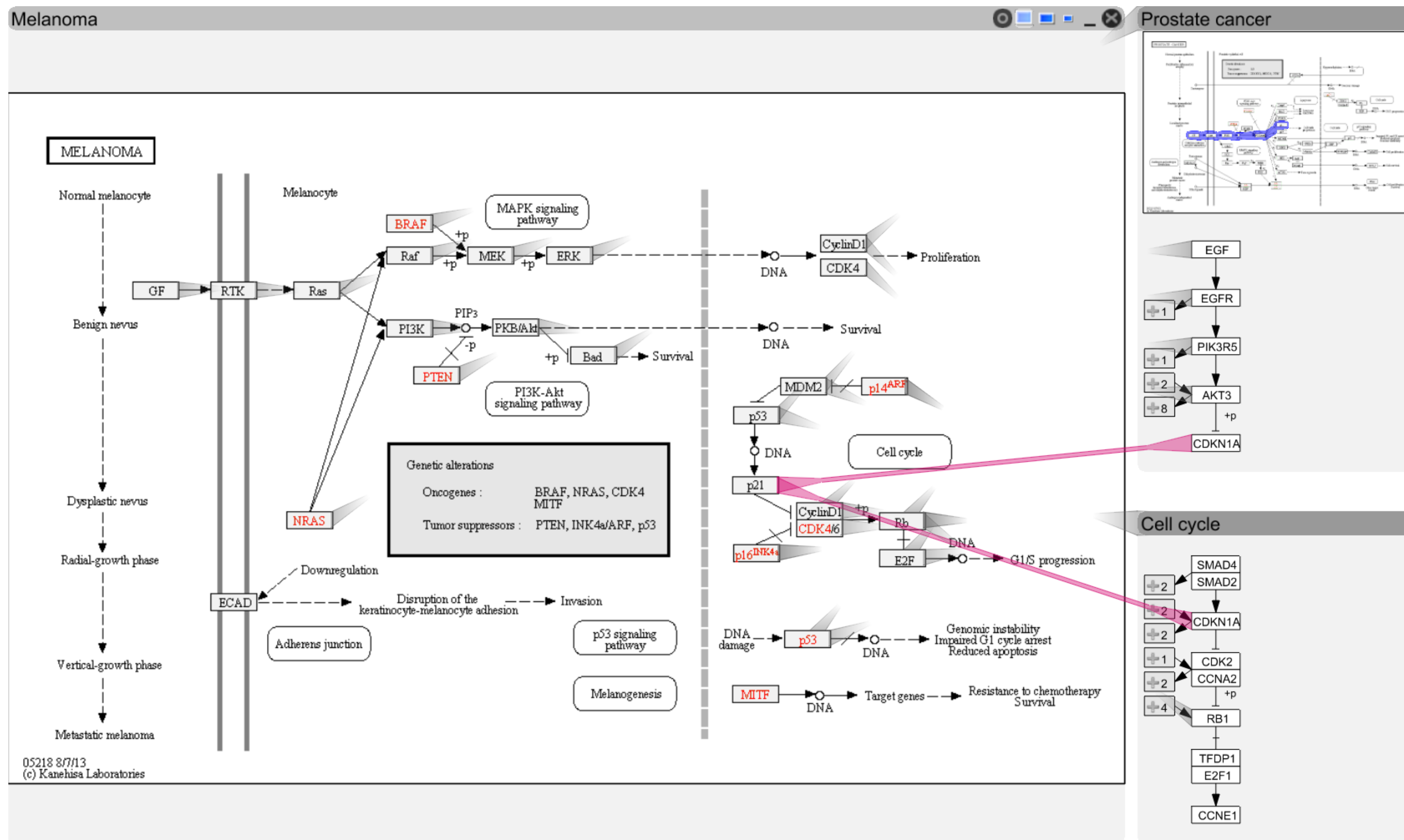
Step 3: Visualizing Relationships



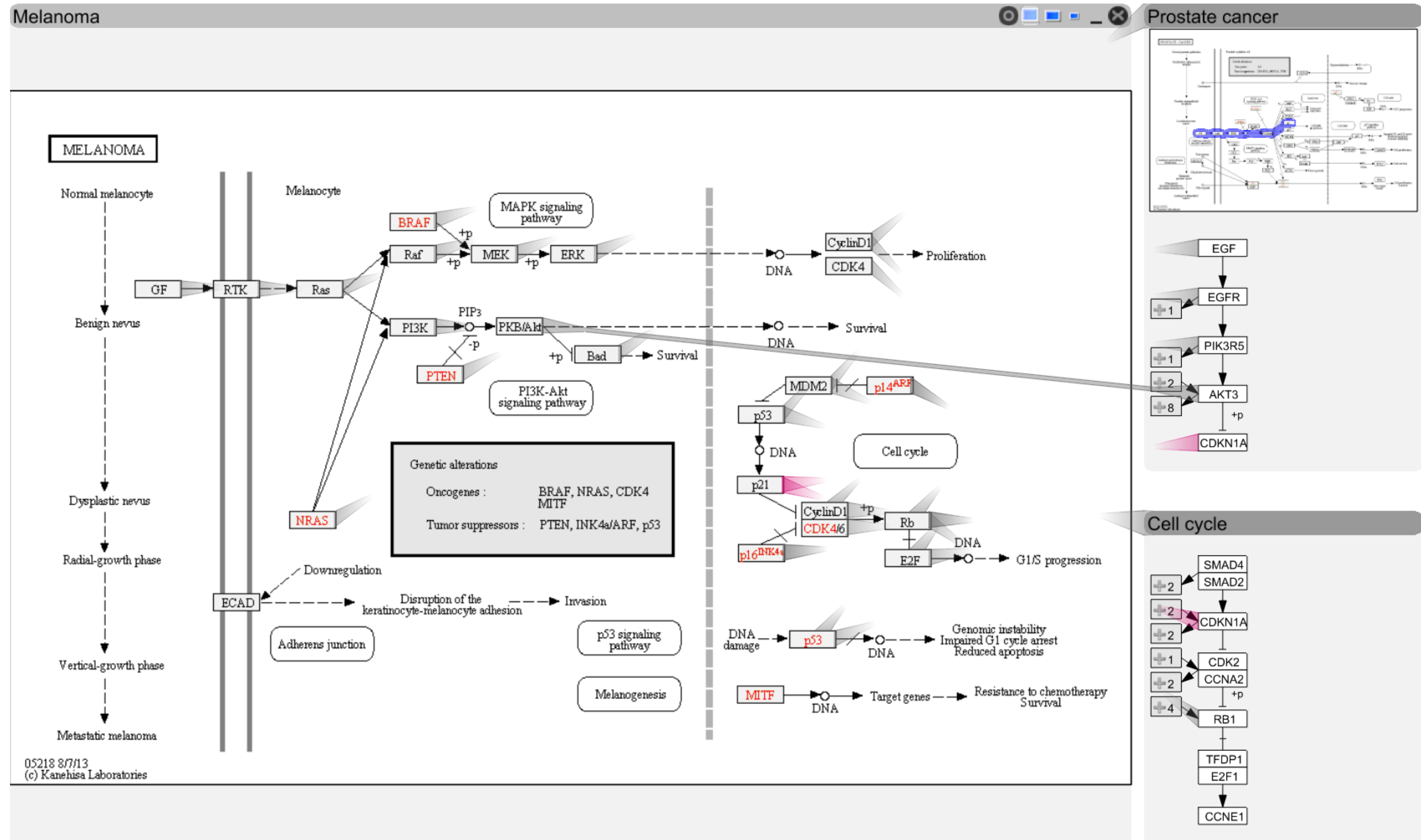
Step 3: Visualizing Relationships



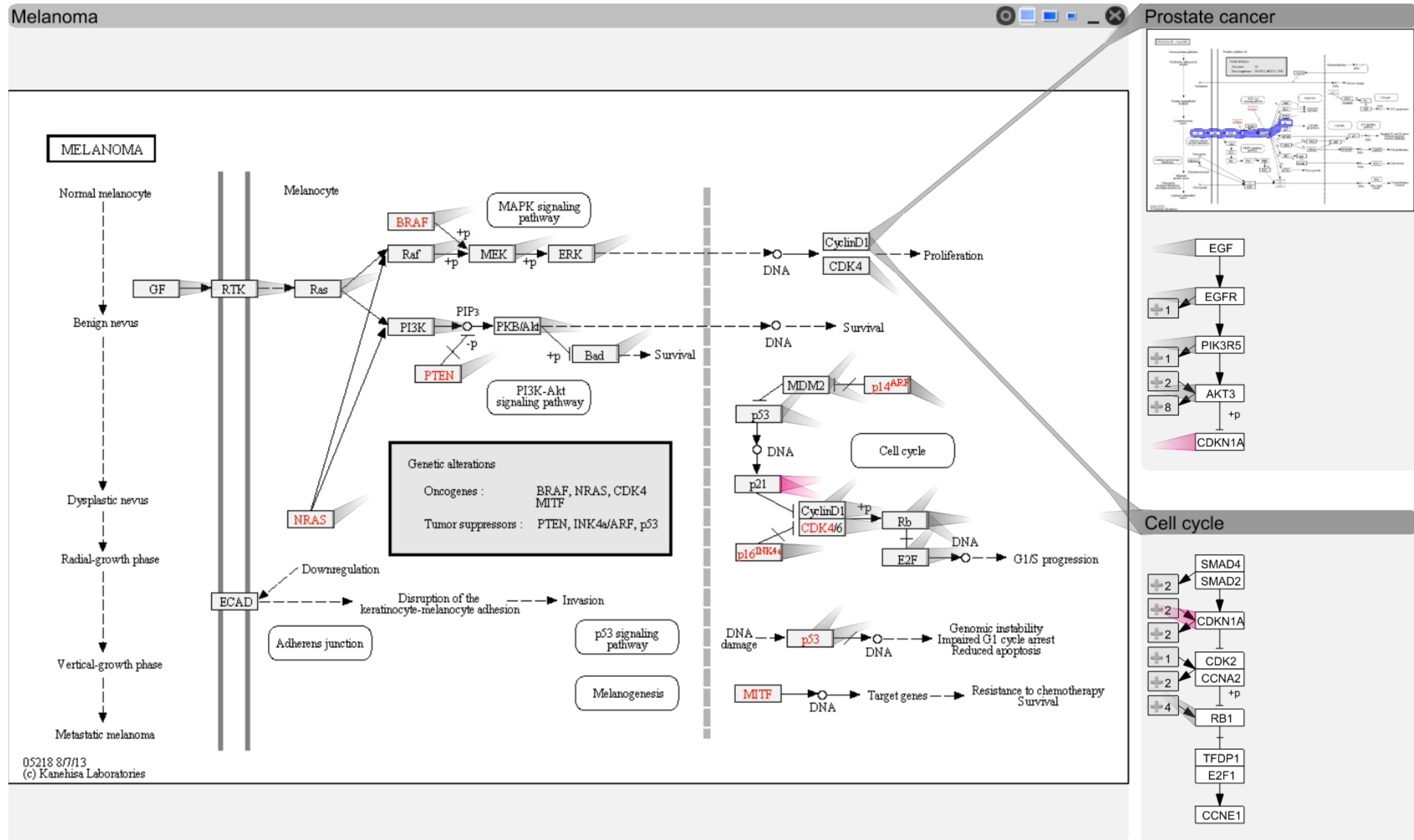
Step 3: Visualizing Relationships



Step 3: Visualizing Relationships



Step 3: Visualizing Relationships



FileDataWindowViewHelp

Entourage

Pathways

Pathway

Filter:
<None>

1 C donor

2-Oxocarboxylic acid

ABC transporters

ABC-family proteins

ACE Inhibitor Pathwa

Acetylcholine Synthes

Acute myeloid leukem

Adherens junction

Adipocyte TarBase

Adipocytokine signali

Adipogenesis

Advanced glycosylatio

Aflatoxin B1 metaboli

African trypanosomias

AGE/RAGE pathway

AhR pathway

Alanine and aspartate

Alanine, aspartate an

Alcoholism

Aldosterone-regulated

Allograft rejection

Allograft rejection

Alpha 6 Beta 4 signal

alpha-Linolenic acid

Alzheimer's disease

Alzheimers Disease

amino acid conjugatio

amino acid conjugatio

Amino sugar and nucl

Aminoacyl-tRNA bios

Amoebiasis

Amphetamine addicti

AMPK signaling

Amyotrophic lateral sc

Androgen receptor si

Angiogenesis

Angiogenesis

angiogenesis overvie

Antigen processing an

APC/C-mediated degra

Apoptosis

Apoptosis

Apoptosis Meta Path

Apoptosis Modulation

Apoptosis Modulation

Apoptosis, anoikis an

Selected Path

Visualizing Many Node Attributes

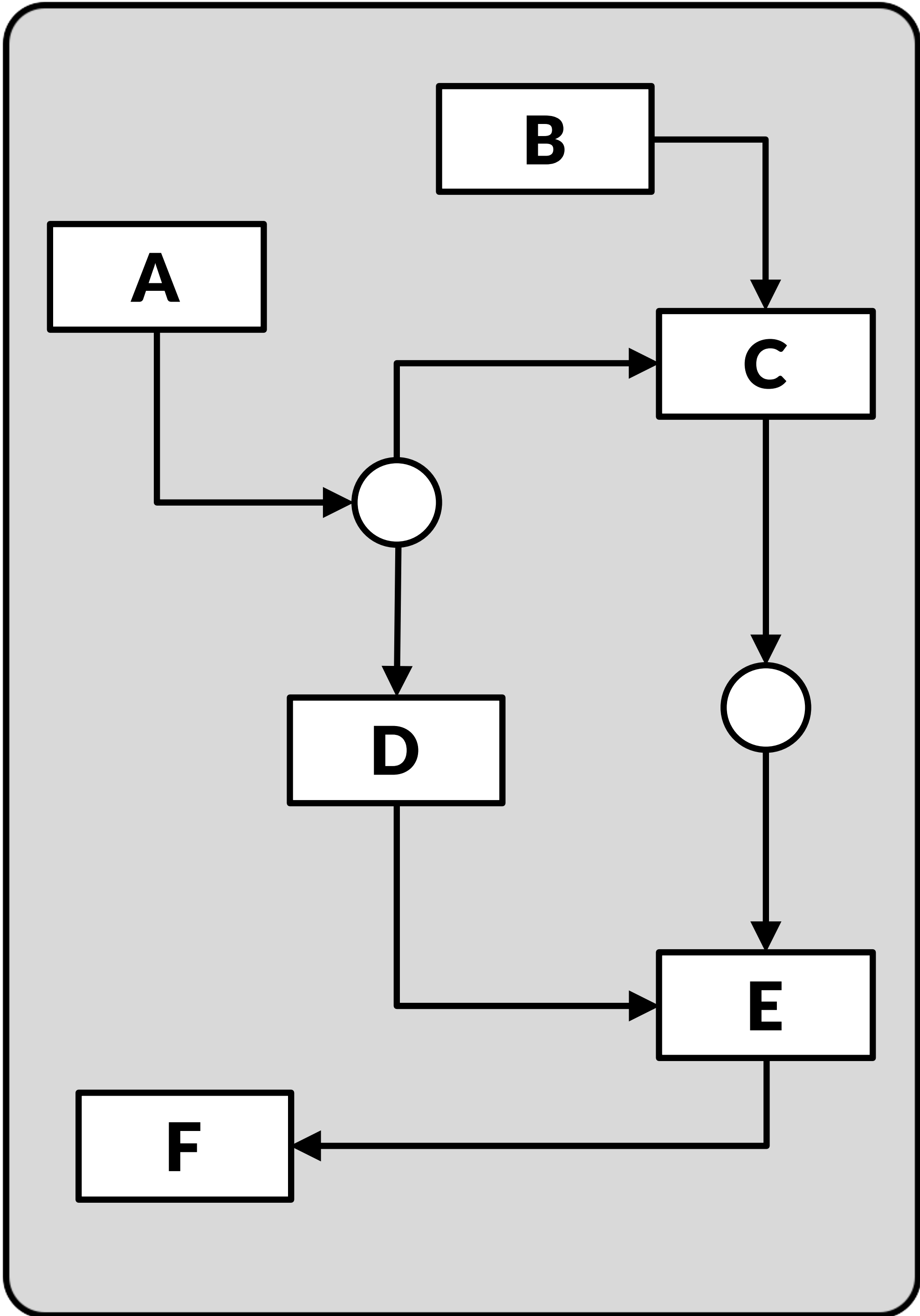
Experi- mental Data and Pathways

[PartI, BioVis '12]

Cannot account for **variation** found in
real-world data

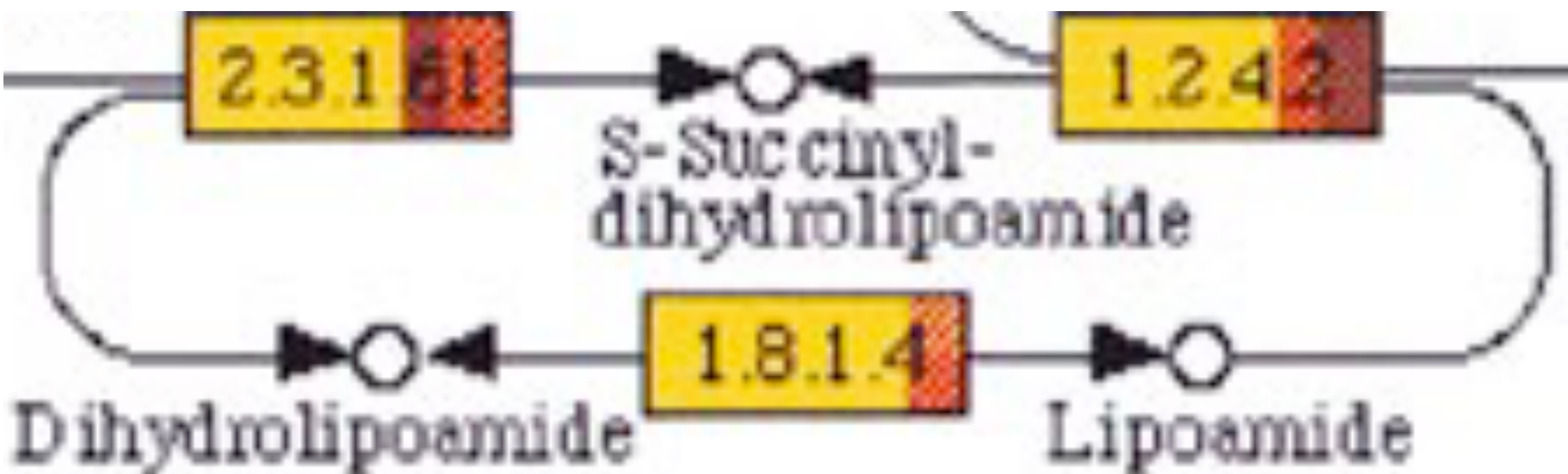
Branches can be **(in)activated** due to
mutation,
changed gene expression,
modulation due to drug treatment,
etc.

Good Old Color Coding



A	-3.4	4.2	5.1	4.2
B	2.8	1.8	1.3	1.1
C	3.1	-2.2	2.4	2.2
D	-3	-2.8	1.6	1.0
E	0.5	0.3	-1.1	1.3
F	0.3	0.3	1.8	-0.3

[Lindroos2002]



Challenge: Data Scale & Heterogeneity

Large **number of experiments**

Large datasets have more than 500 experiments

Multiple **groups/conditions**

Different **types** of data

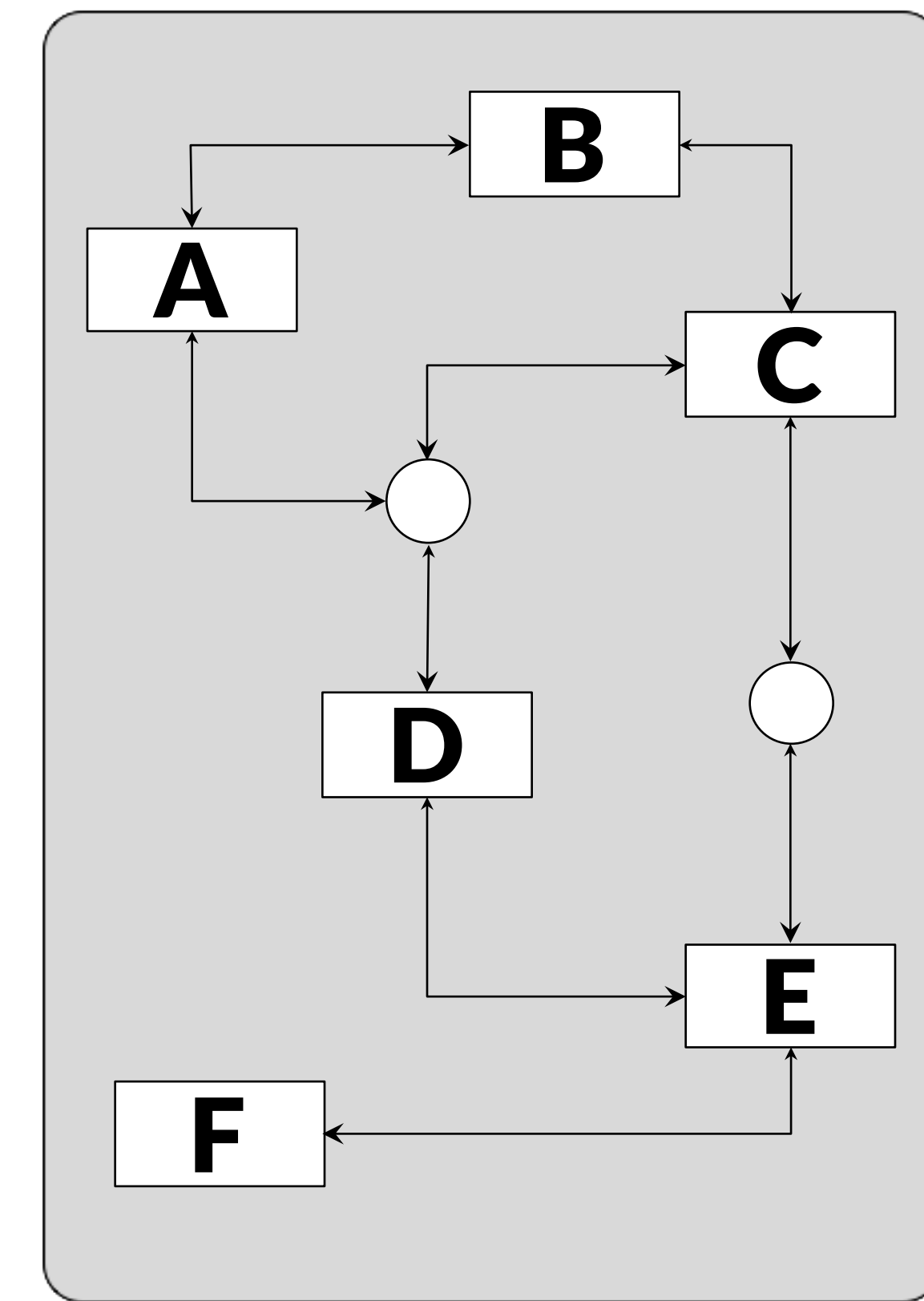
Challenge: Supporting Multiple Tasks

Two central tasks:

Explore **topology** of pathway

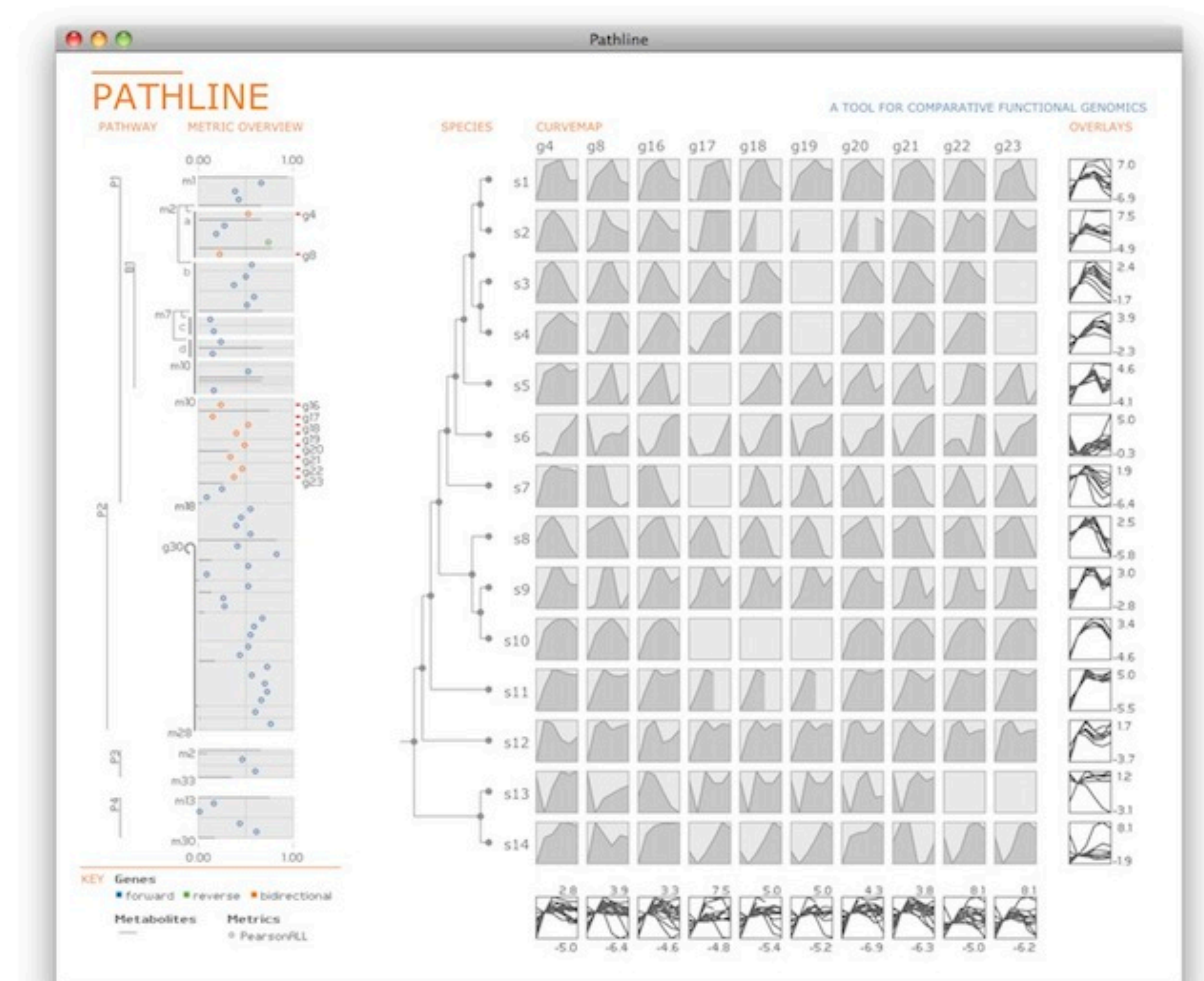
Explore the **attributes** of the nodes
(experimental data)

Need to support both!



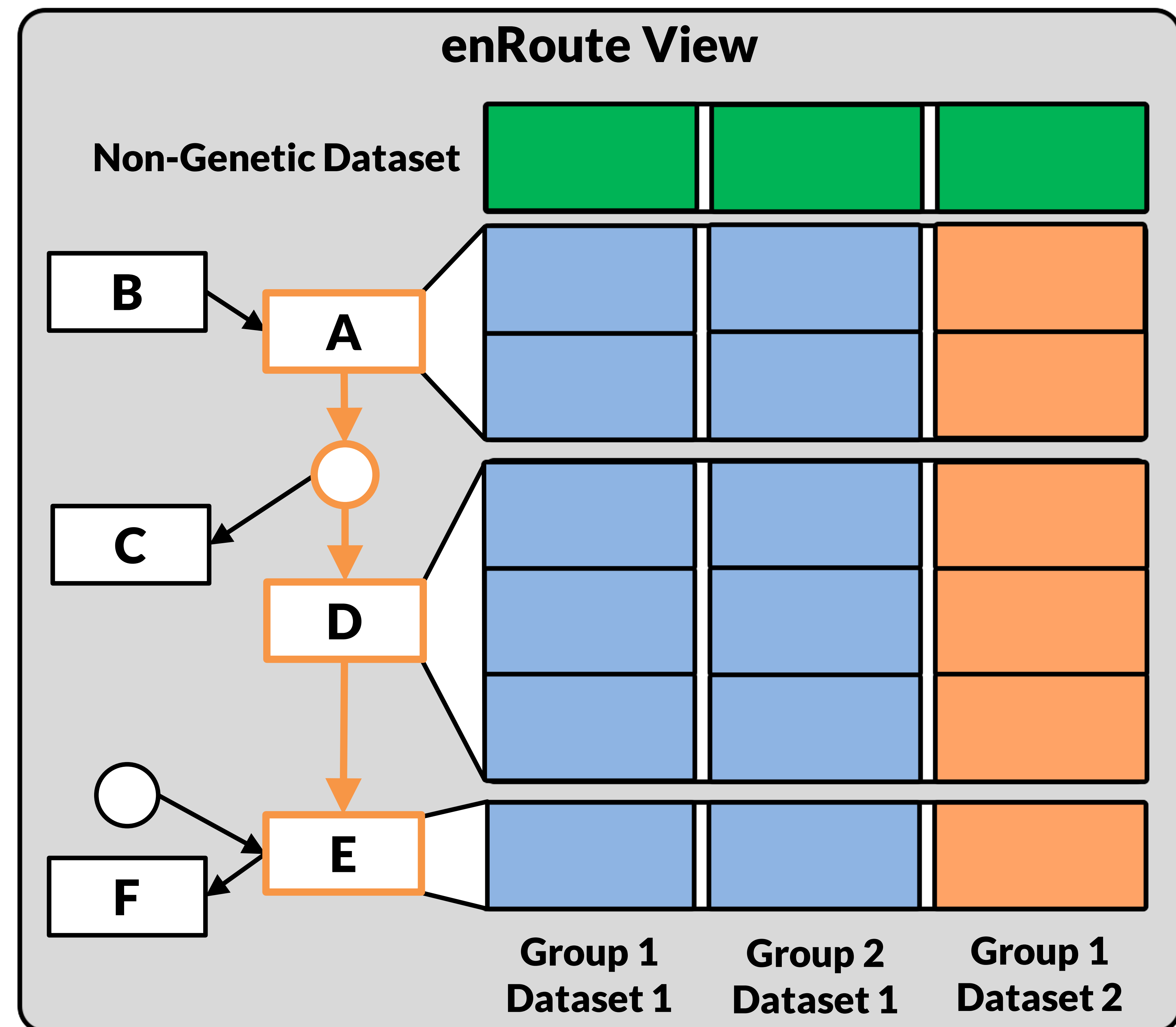
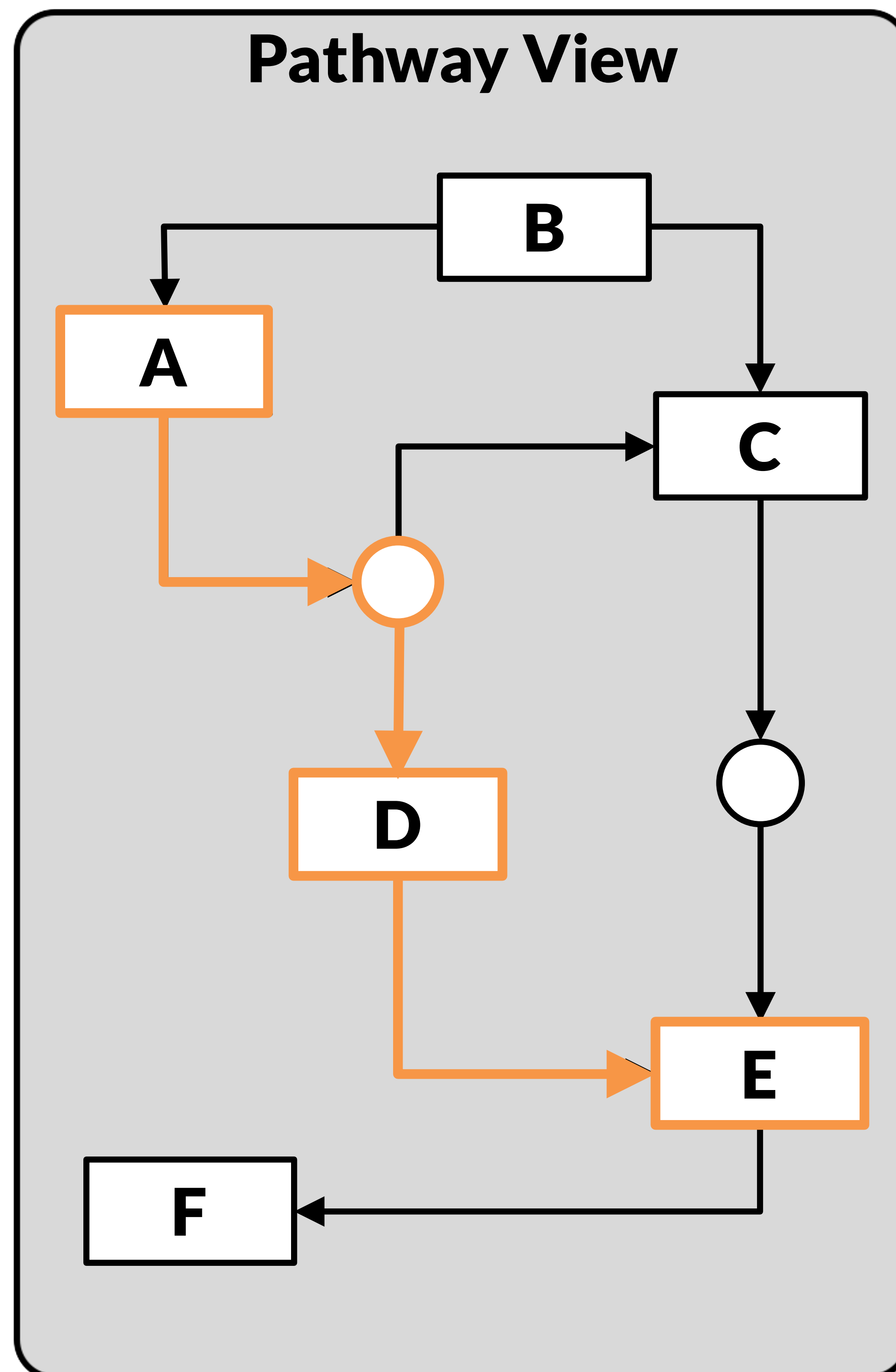
	Sample 1	Sample 2	Sample 3
Gene 1	1	1.1	0.4
Gene 2	2	0.5	1.2
Gene 3	1.4	0.2	0.5
Gene 4	0.3	0.5	0.7

Inspiration

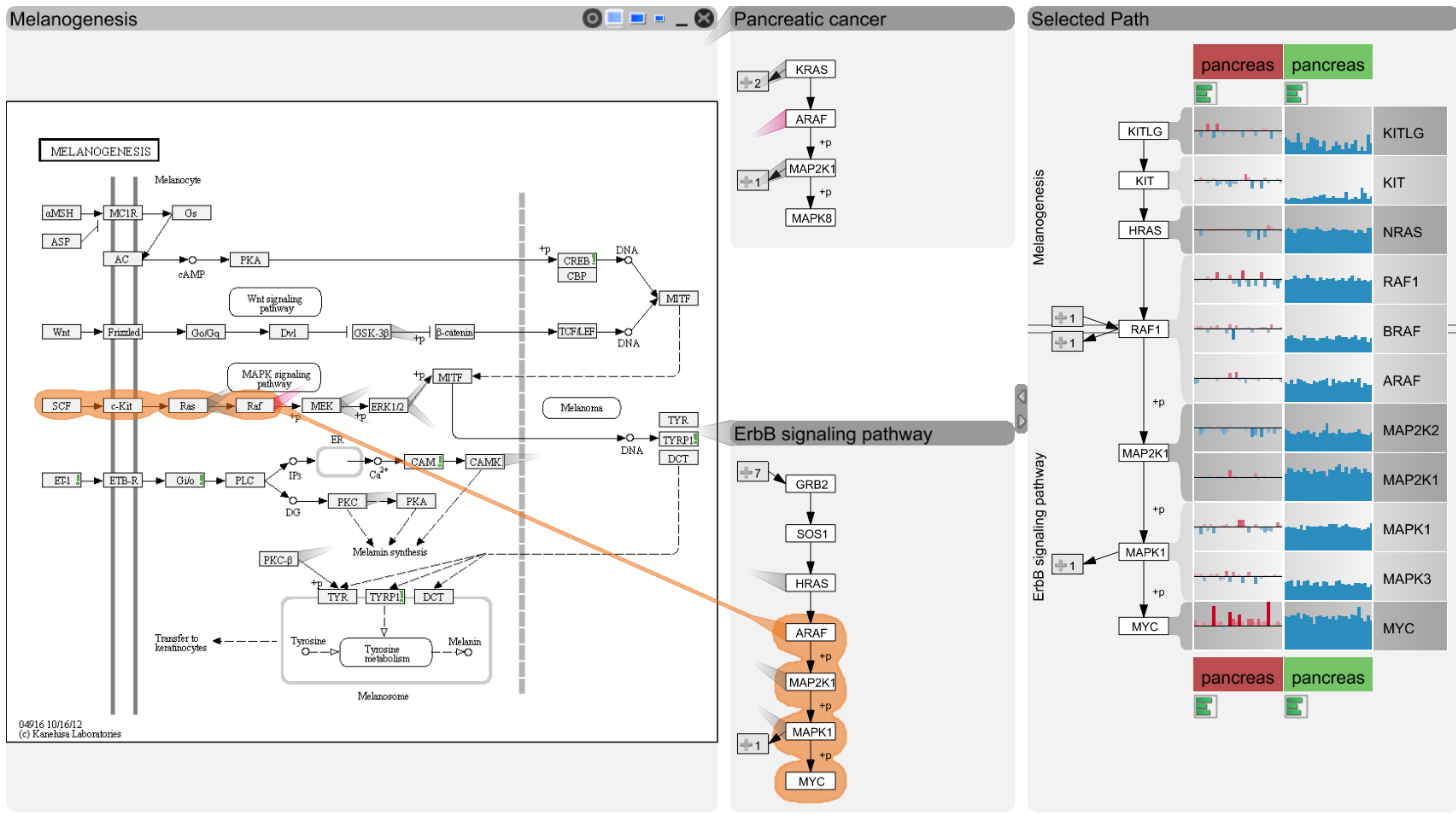


[Meyer 2010]

Concept



enRoute



FileDataWindowViewHelp

Entourage

Pathways

Pathway

Filter:
<None>

1 C donor

2-Oxocarboxylic acid

ABC transporters

ABC-family proteins

ACE Inhibitor Pathwa

Acetylcholine Synthes

Acute myeloid leukem

Adherens junction

Adipocyte TarBase

Adipocytokine signali

Adipogenesis

Advanced glycosylatio

Aflatoxin B1 metaboli

African trypanosomias

AGE/RAGE pathway

AhR pathway

Alanine and aspartate

Alanine, aspartate an

Alcoholism

Aldosterone-regulated

Allograft rejection

Allograft rejection

Alpha 6 Beta 4 signal

alpha-Linolenic acid

Alzheimer's disease

Alzheimers Disease

amino acid conjugatio

amino acid conjugatio

Amino sugar and nucl

Aminoacyl-tRNA bios

Amoebiasis

Amphetamine addicti

AMPK signaling

Amyotrophic lateral sc

Androgen receptor si

Angiogenesis

Angiogenesis

angiogenesis overvie

Antigen processing an

APC/C-mediated degra

Apoptosis

Apoptosis

Apoptosis Meta Path

Apoptosis Modulation

Apoptosis Modulation

Apoptosis, anoikis an

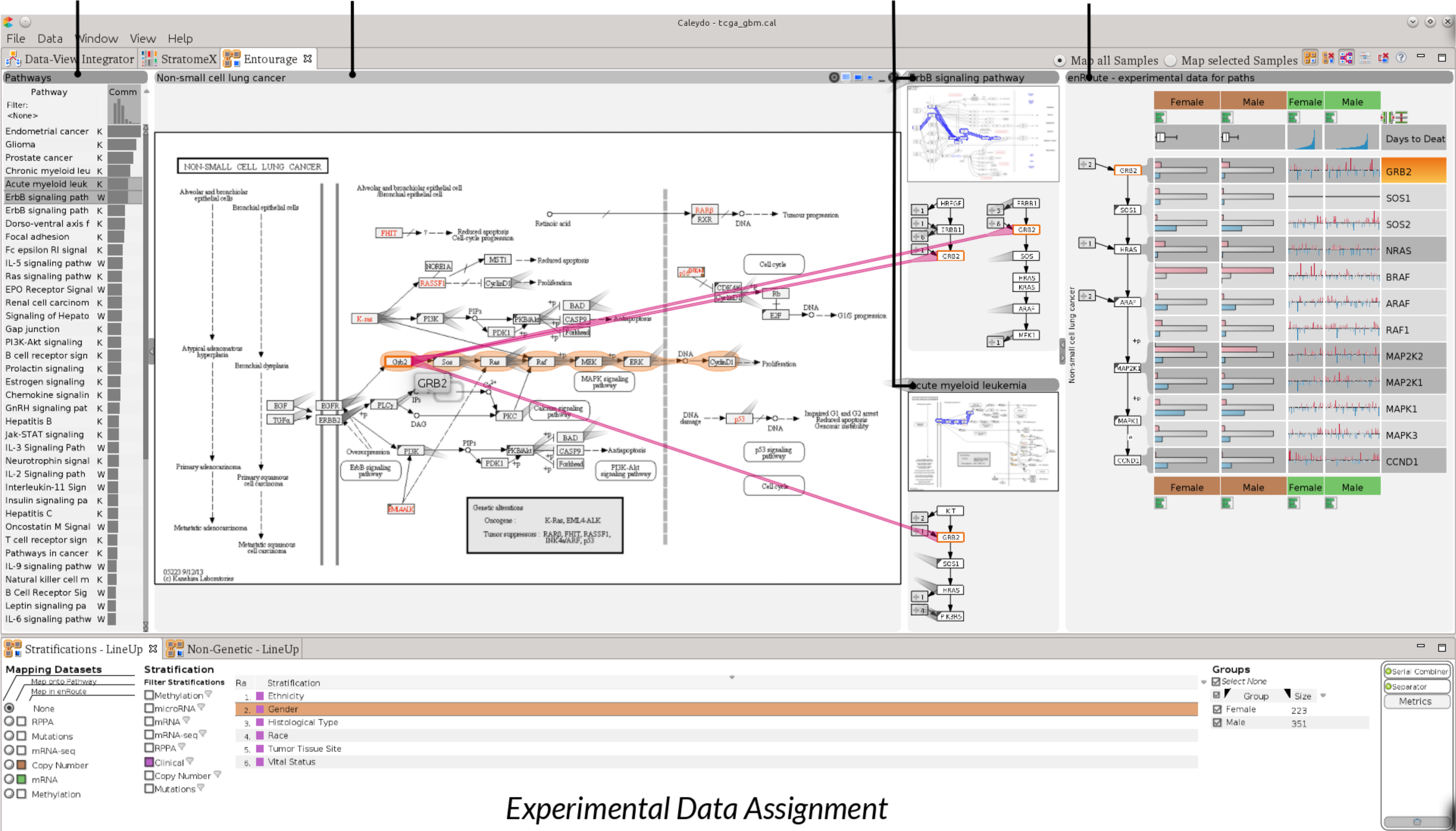
Selected Path

Pathway List

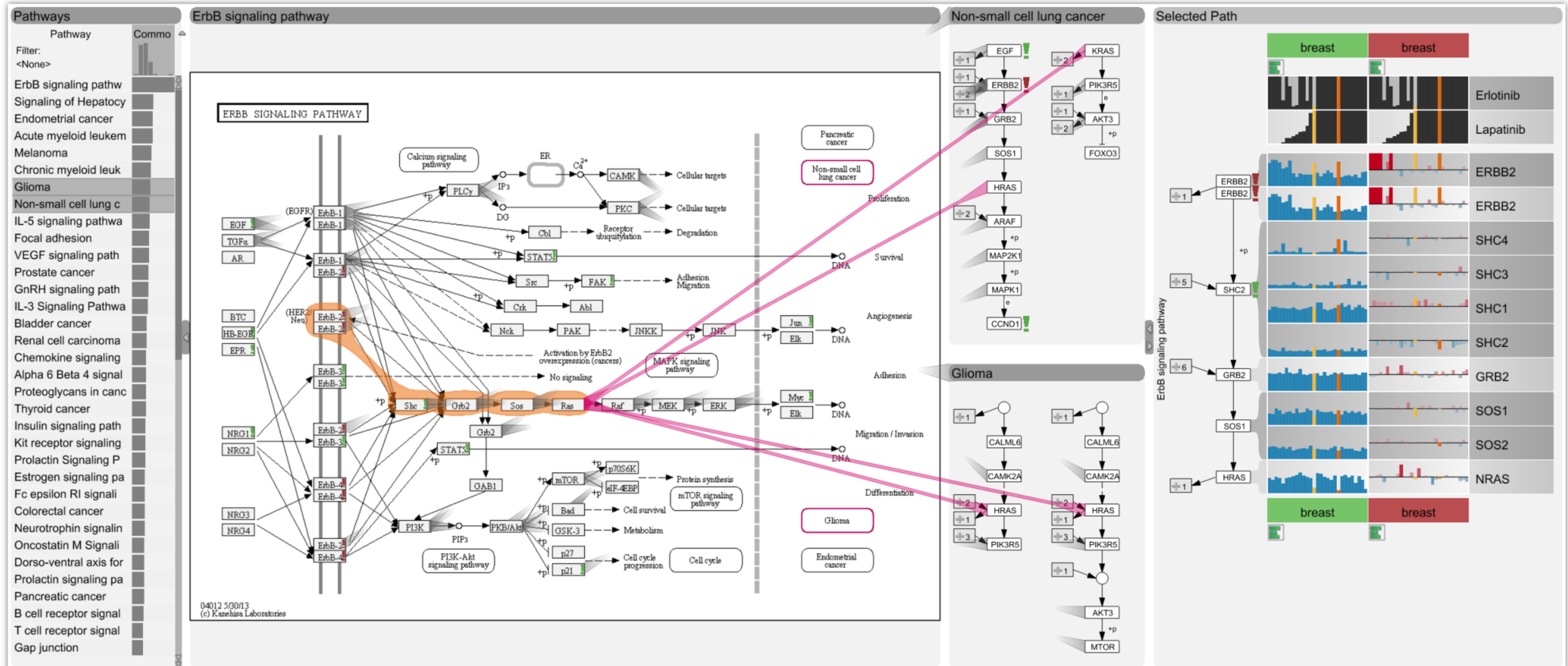
Focus Pathway

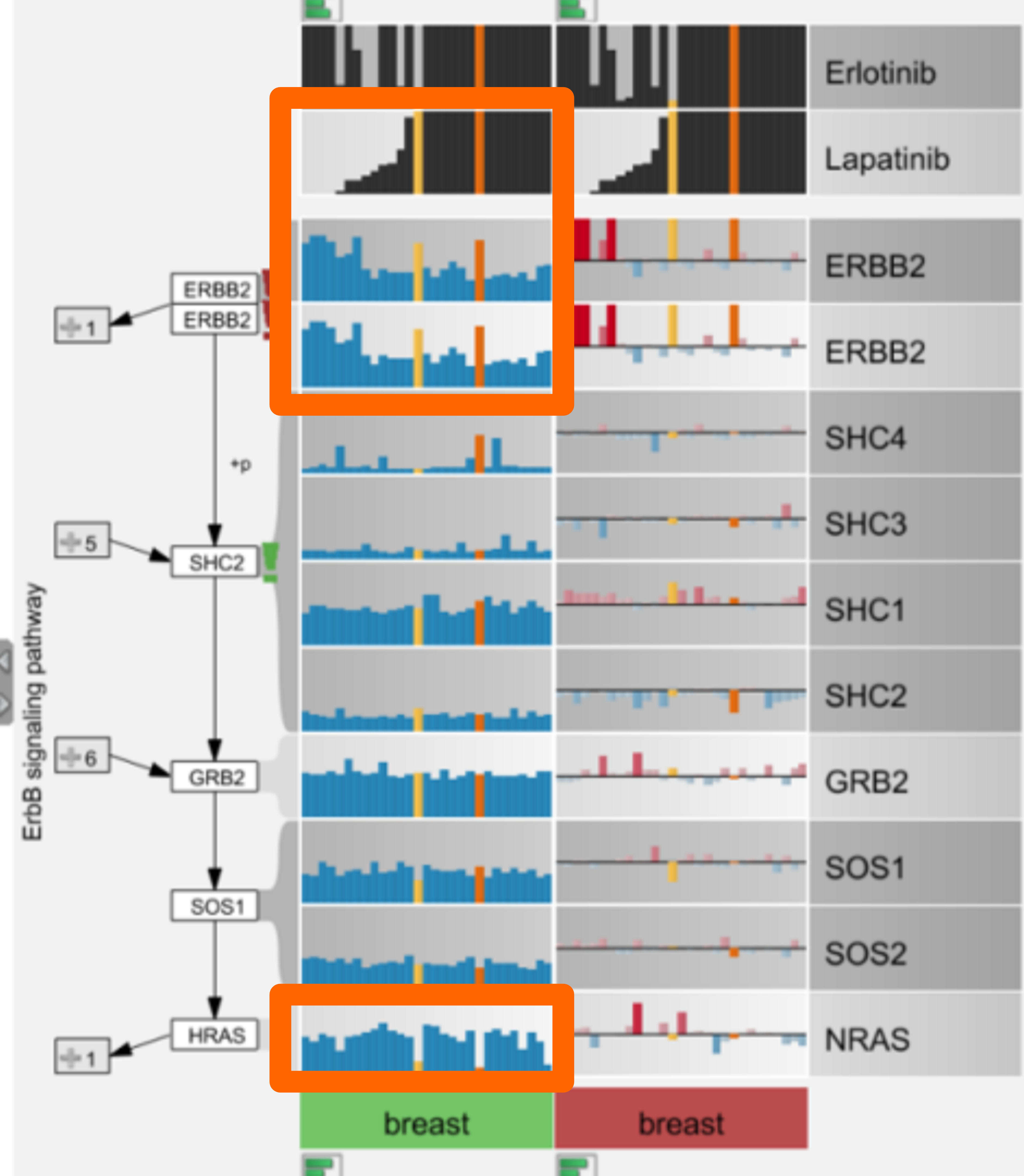
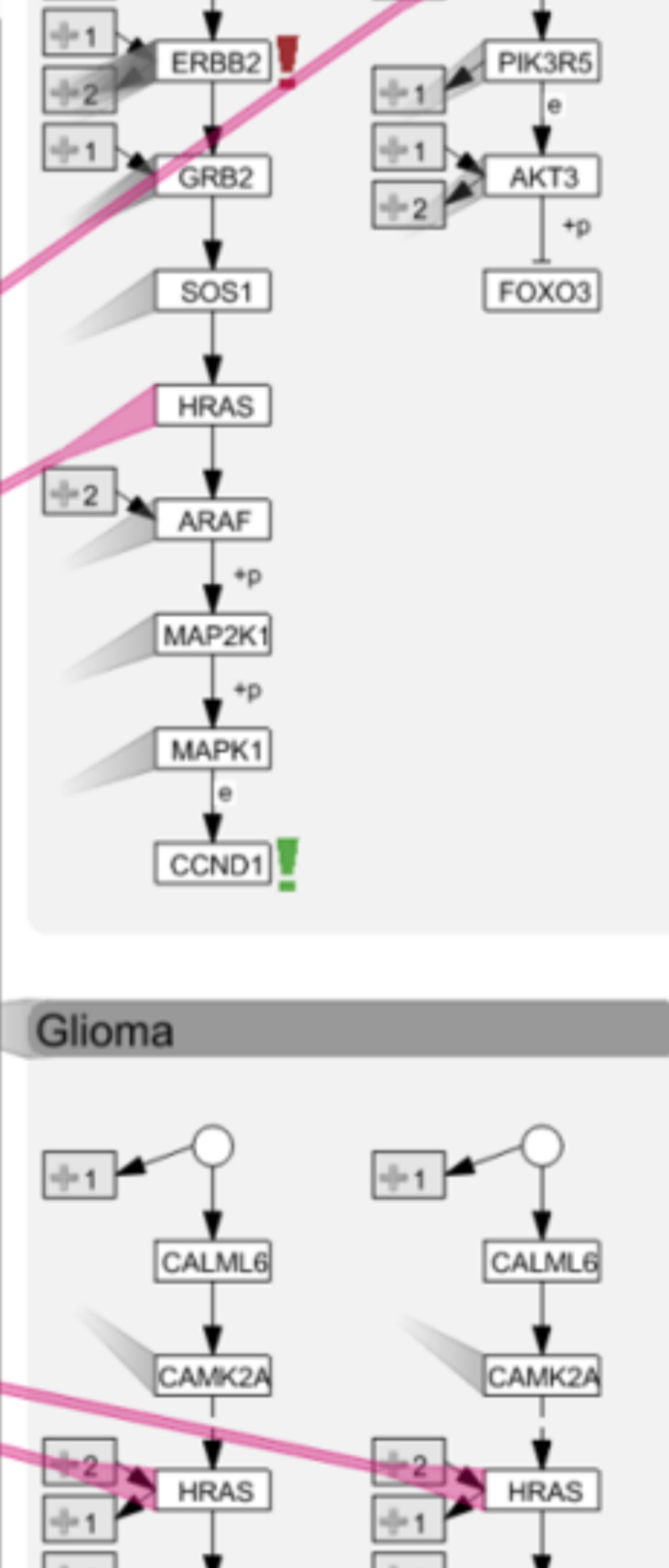
Context Pathways

enRoute View



Case Study: CCLE Data

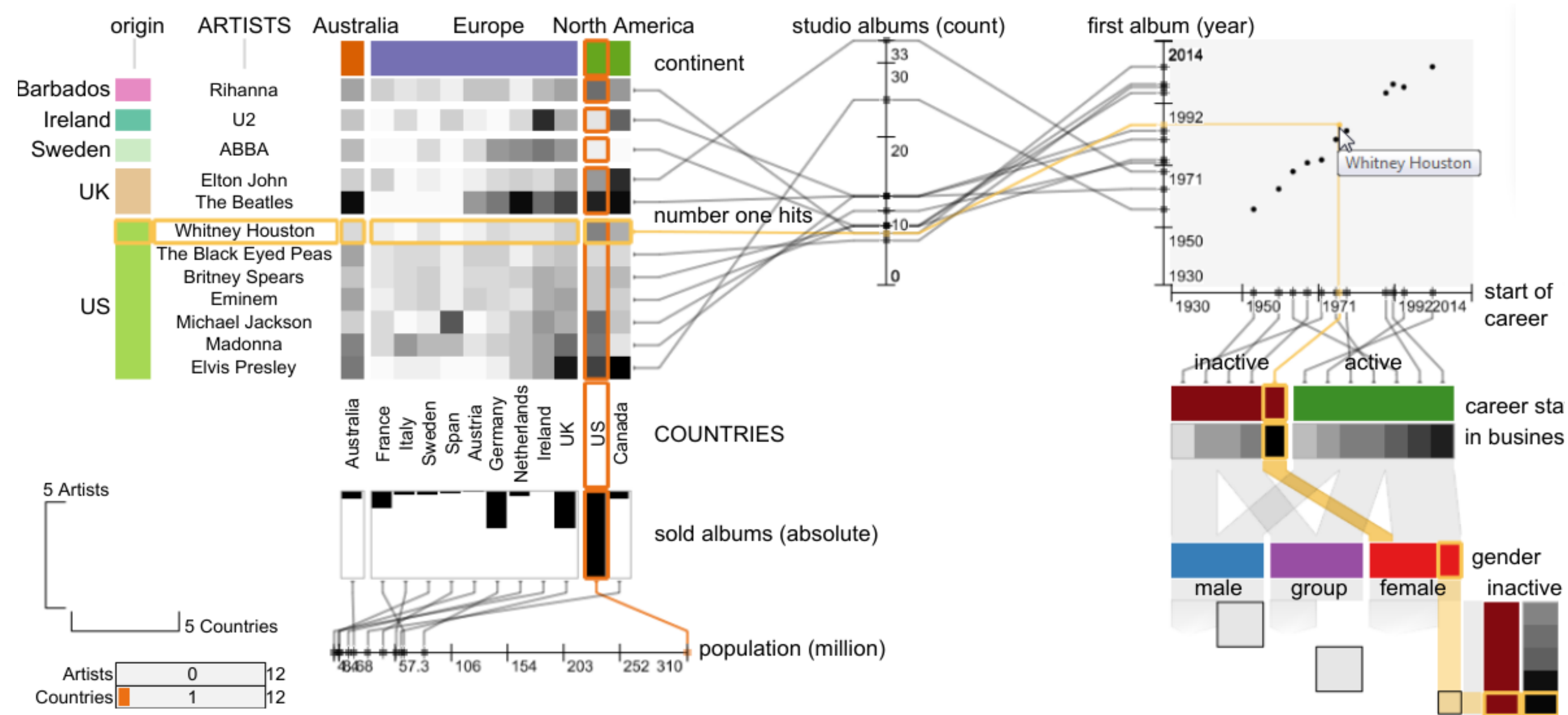




<http://entourage.caleydo.org>

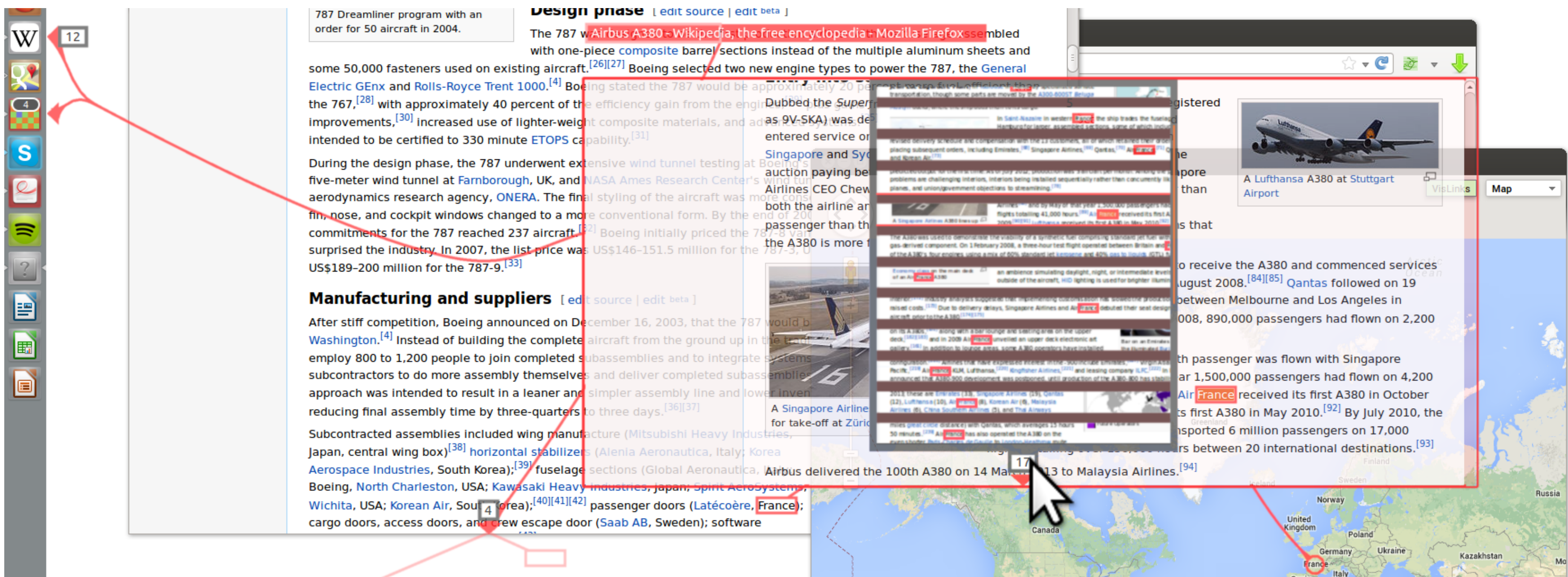
Other Techniques

Creating Integrated Visualizations from Subsets



IEEE TVCG (InfoVis '14)
HONORABLE MENTION AWARD

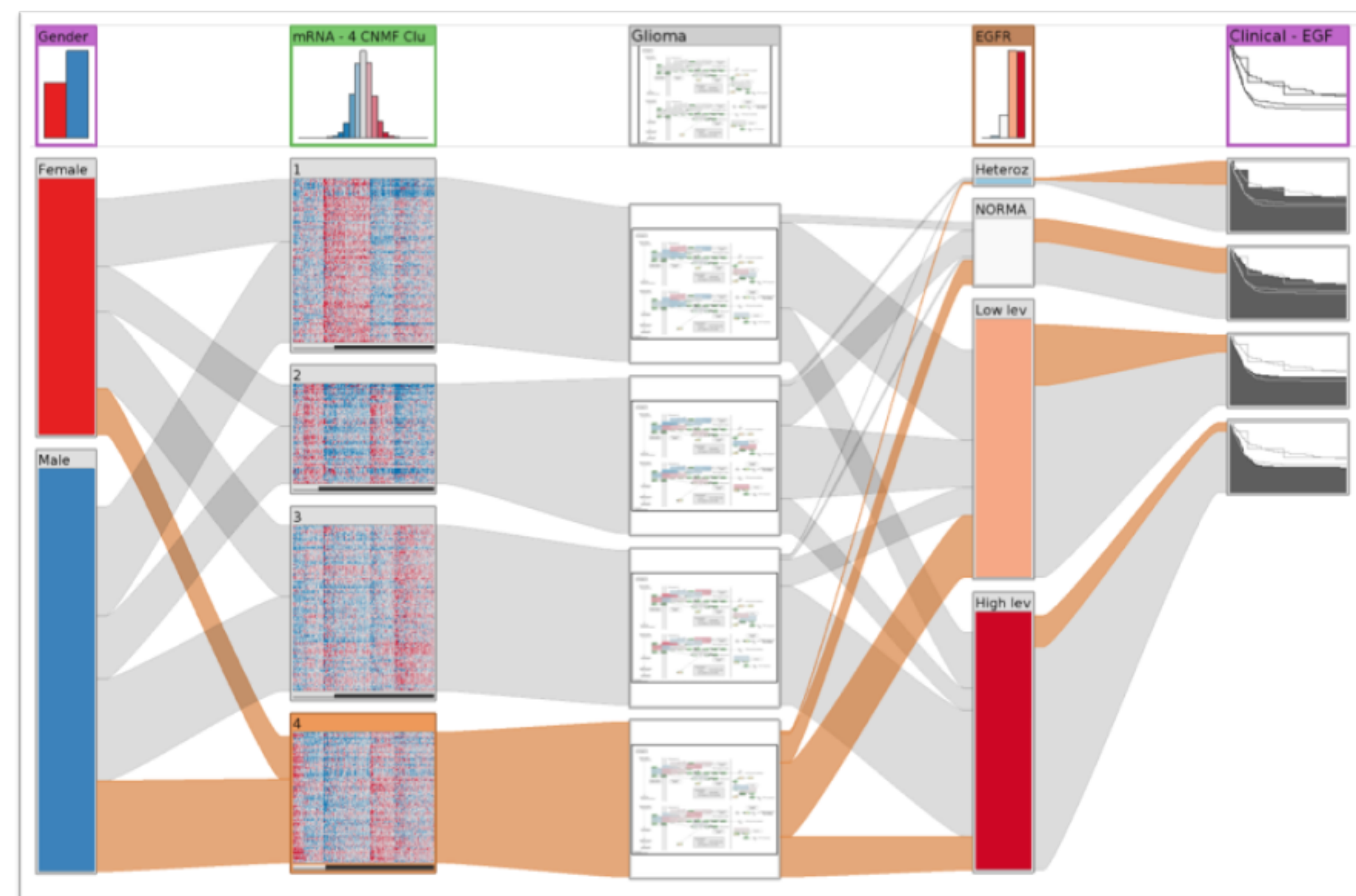
Visualizing Hidden Content



ACM CHI'14
HONORABLE MENTION AWARD

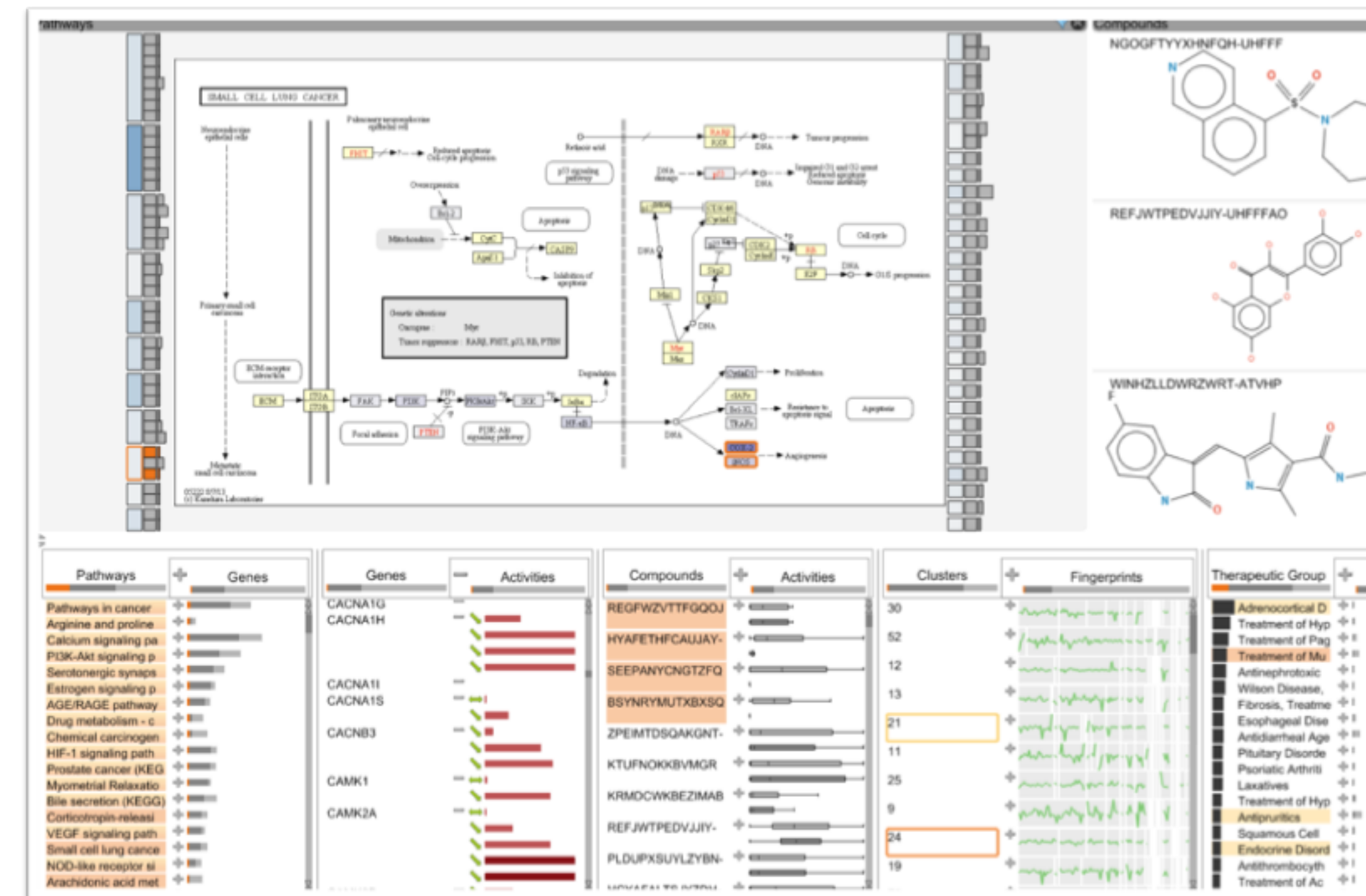
Other Domain Problem Projects

Heterogeneous Datasets
Cancer Subtype Analysis



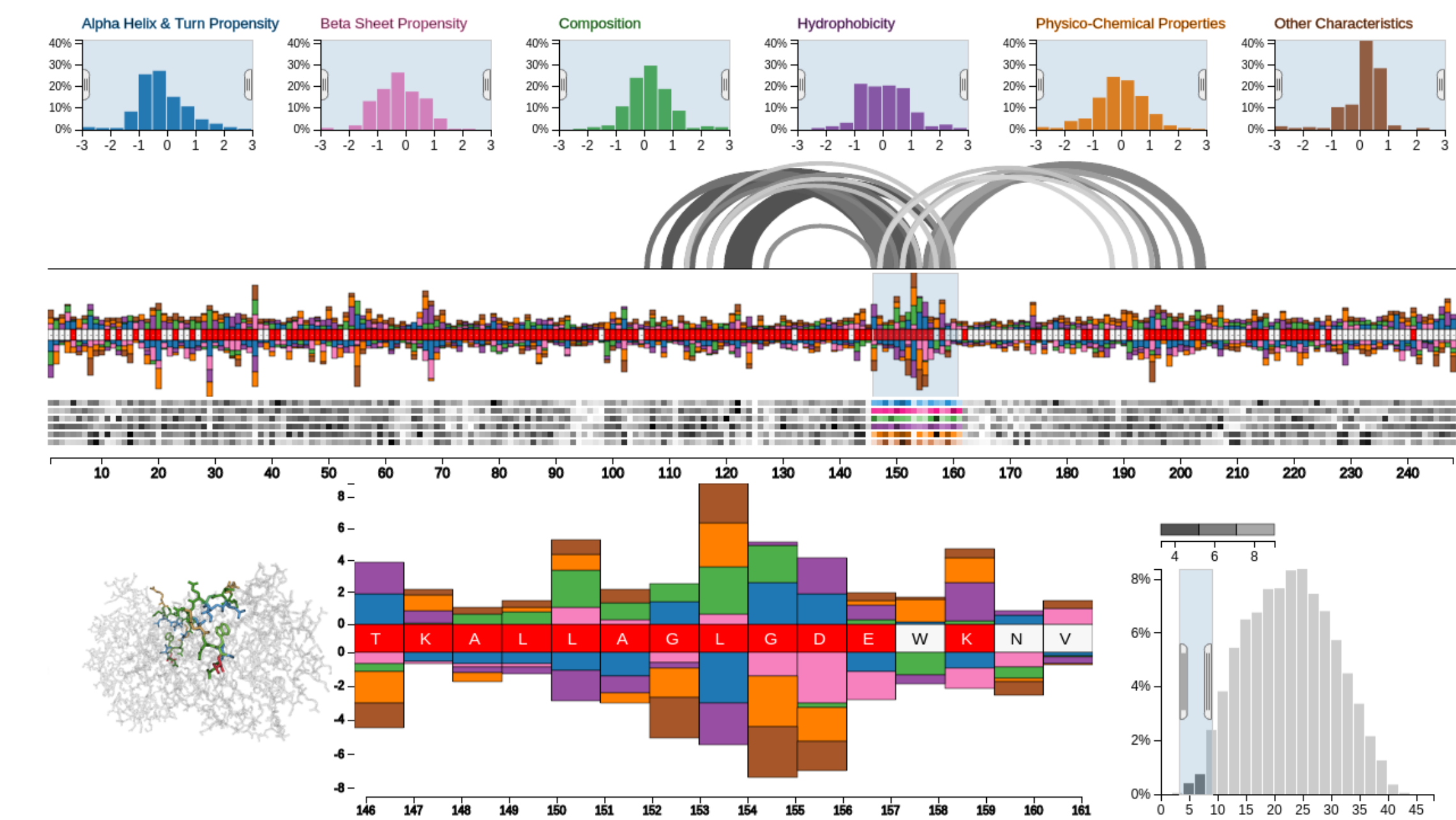
Nature Methods, 2014
CG & A 2014
EuroVis 2012
3rd Best Paper

Faceted Browsing of Graphs
Drug Discovery



IEEE TVCG (VAST '14)

Protein Sequence
Analysis



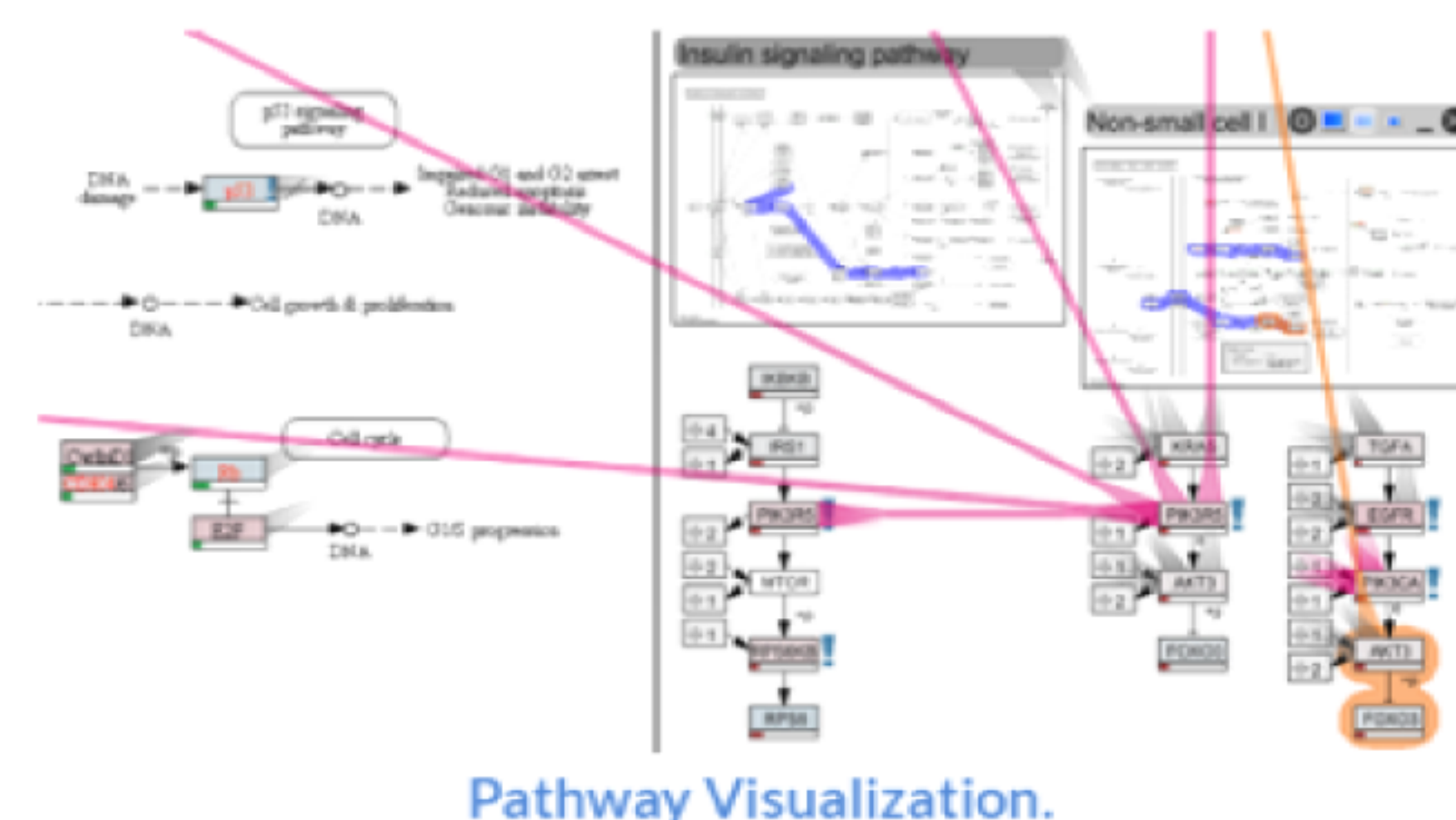
BMC Proceedings 2014

Home

About

Caleydo is an open source visual analysis framework targeted at biomolecular data. The biggest strength of Caleydo is the visualization of interdependencies between multiple datasets. Caleydo can load tabular data and groupings/clustering. You can explore relationships between multiple groupings, between different datasets and see how your data maps onto pathways. Caleydo has been successfully used to analyze mRNA, miRNA, methylation, copy number variation, mutation status and clinical data as well as other dataset types.

All published parts of the software are available as open source. Developers and researchers that want to use or enhance the current tools are invited to contribute. See our [Github page](#) for further details.



Partners

-  [Graz University of Technology](#)
Graz, Austria
-  [Johannes Kepler University Linz](#)
Linz, Austria
-  [Harvard University](#)
Cambridge, MA, USA
-  [Harvard Medical School](#)
Boston, MA, USA

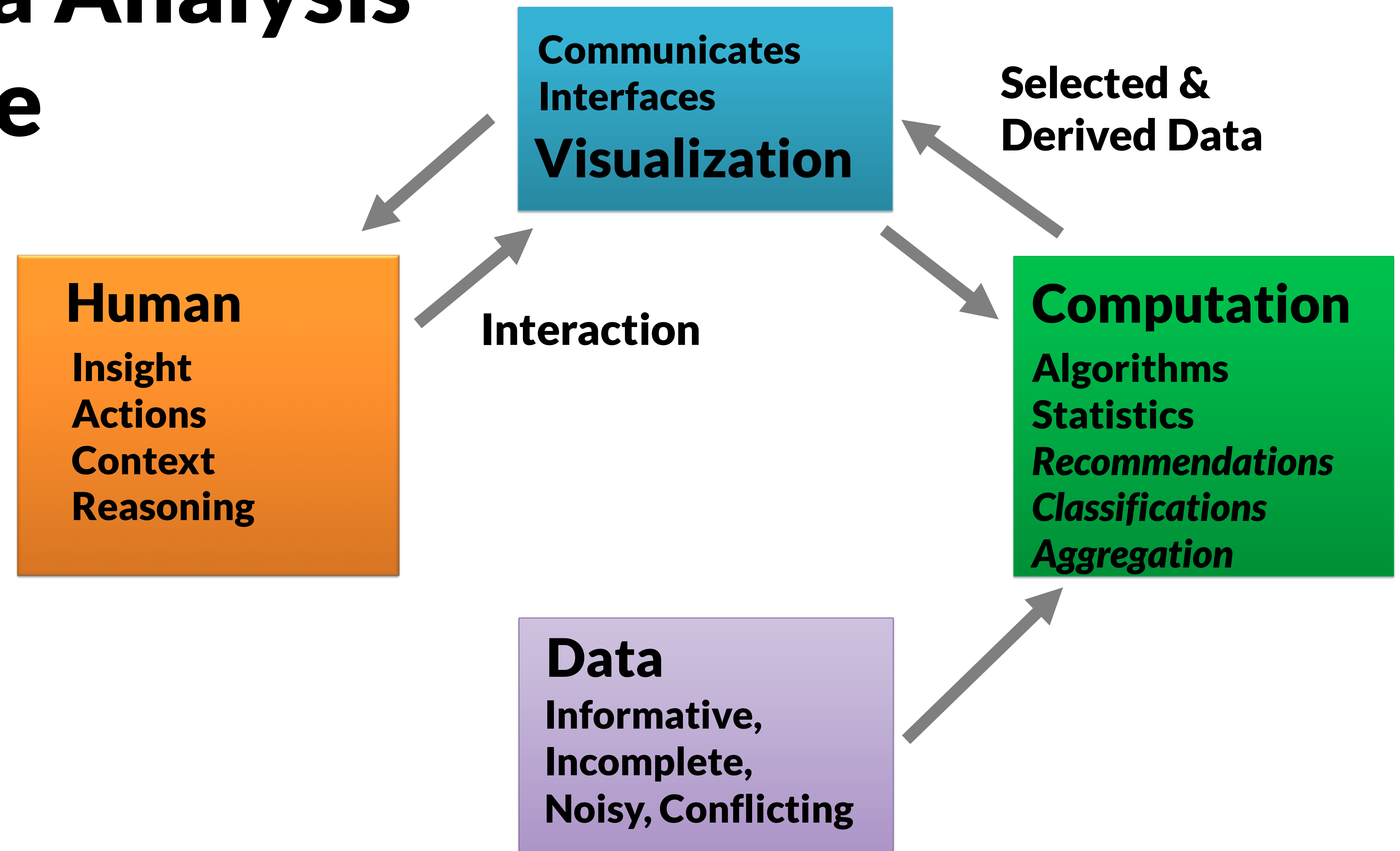
Projects

-  [StratomeX](#)
Genomic Stratification of Diseases
-  [Entourage & enRoute](#)
Experimental Data in Pathways
-  [LineUp](#)
Multi-Attribute Rankings
-  [ConTour](#)
Relational Data in Drug Discovery
-  [Domino](#)
Exploring Subsets of Tabular Datasets
-  [Furby](#)
Bicluster Exploration
-  [Visual Linking & Hidden Content](#)
Showing Relationships

RESEARCH CHALLENGES

Interacting with Data

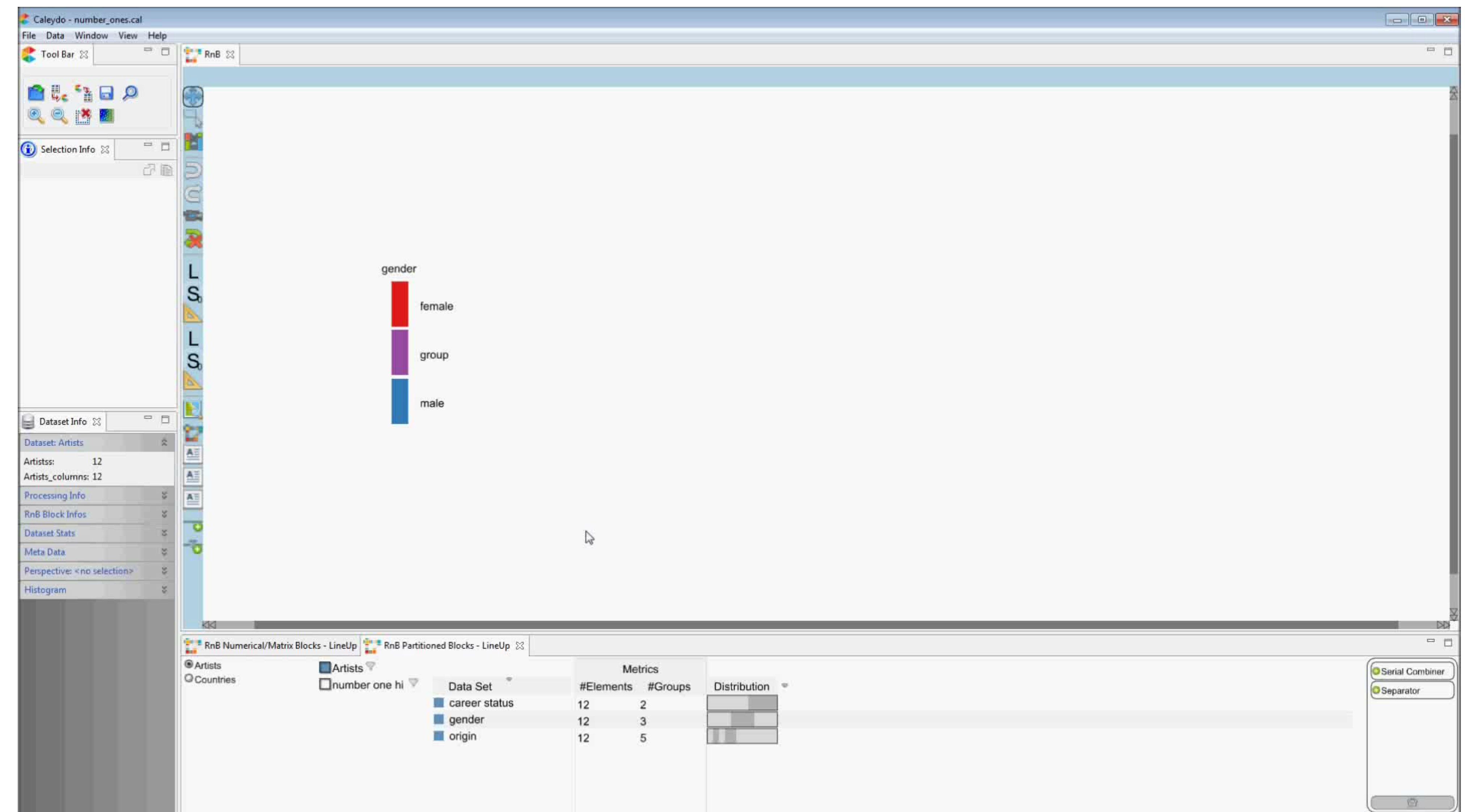
**The Future of Data Analysis
is (also) Interactive**



Can we make creating interactive vis systems easier?

Right now: prerogative of skilled software developers

Goal: democratize the creation of visualizations



How can we scale to increasingly big datasets?

No local copies of datasets

No in-memory processing

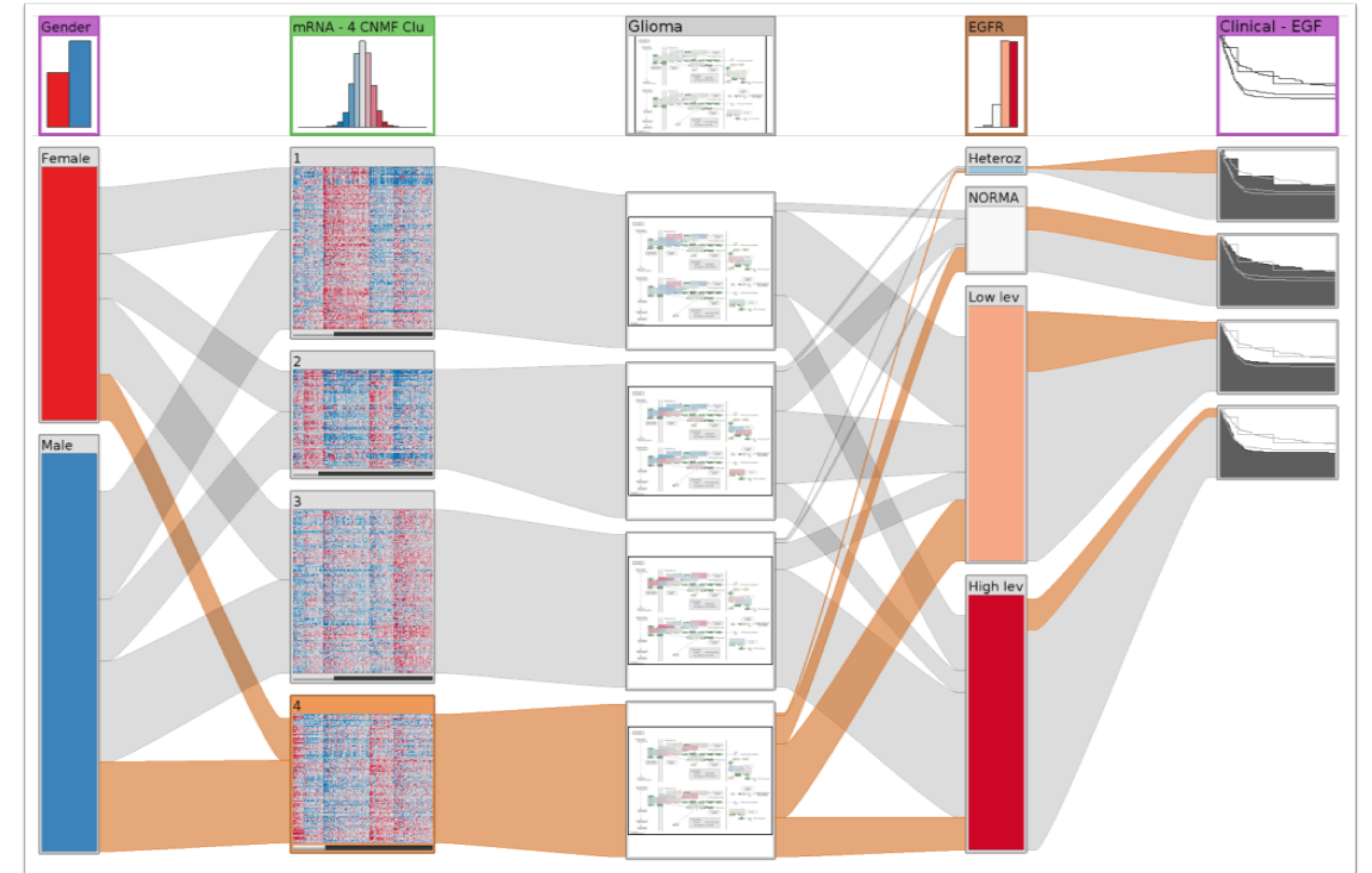
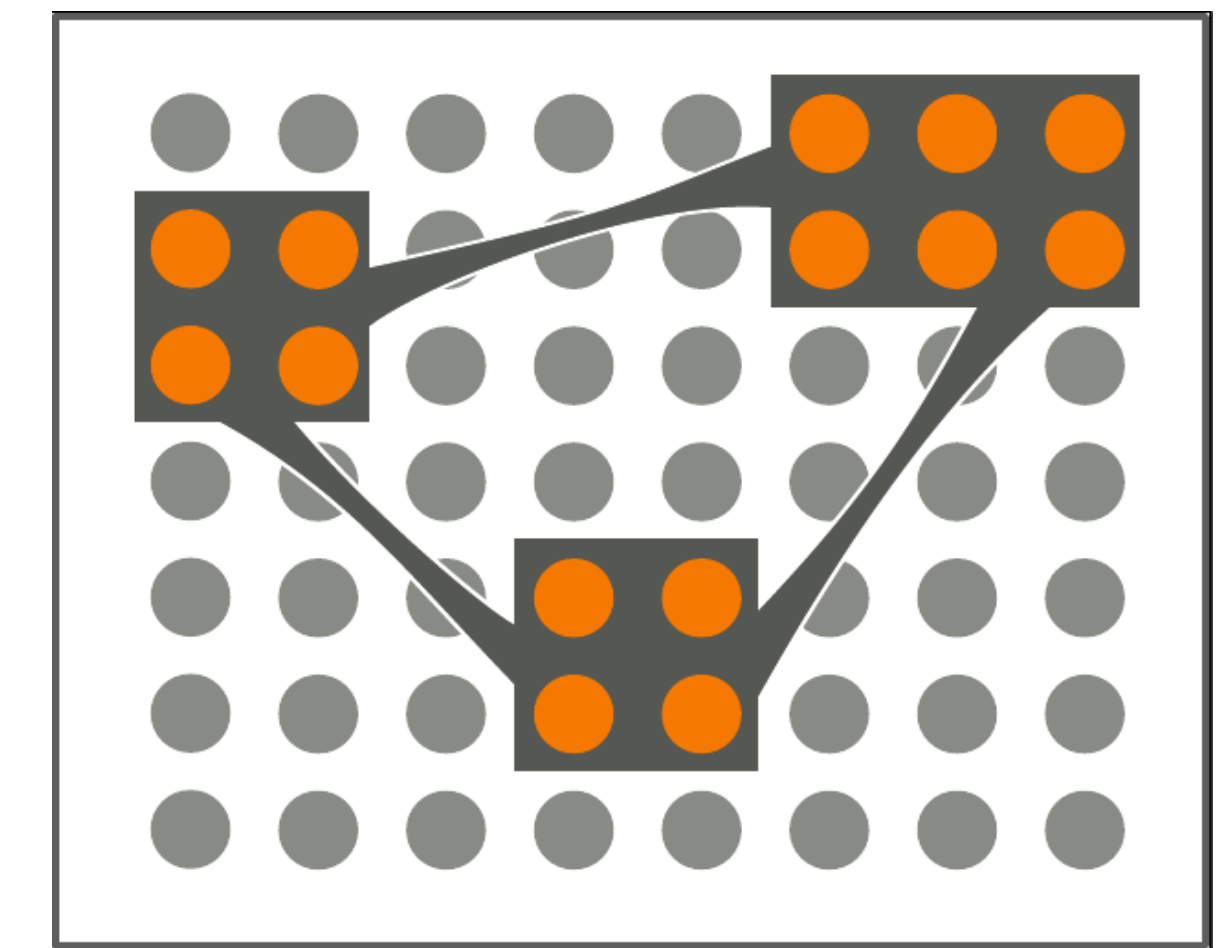
....

Approach:

Automatic slicing into subsets

Query or explore ranked subsets

Pull in relationships for subset

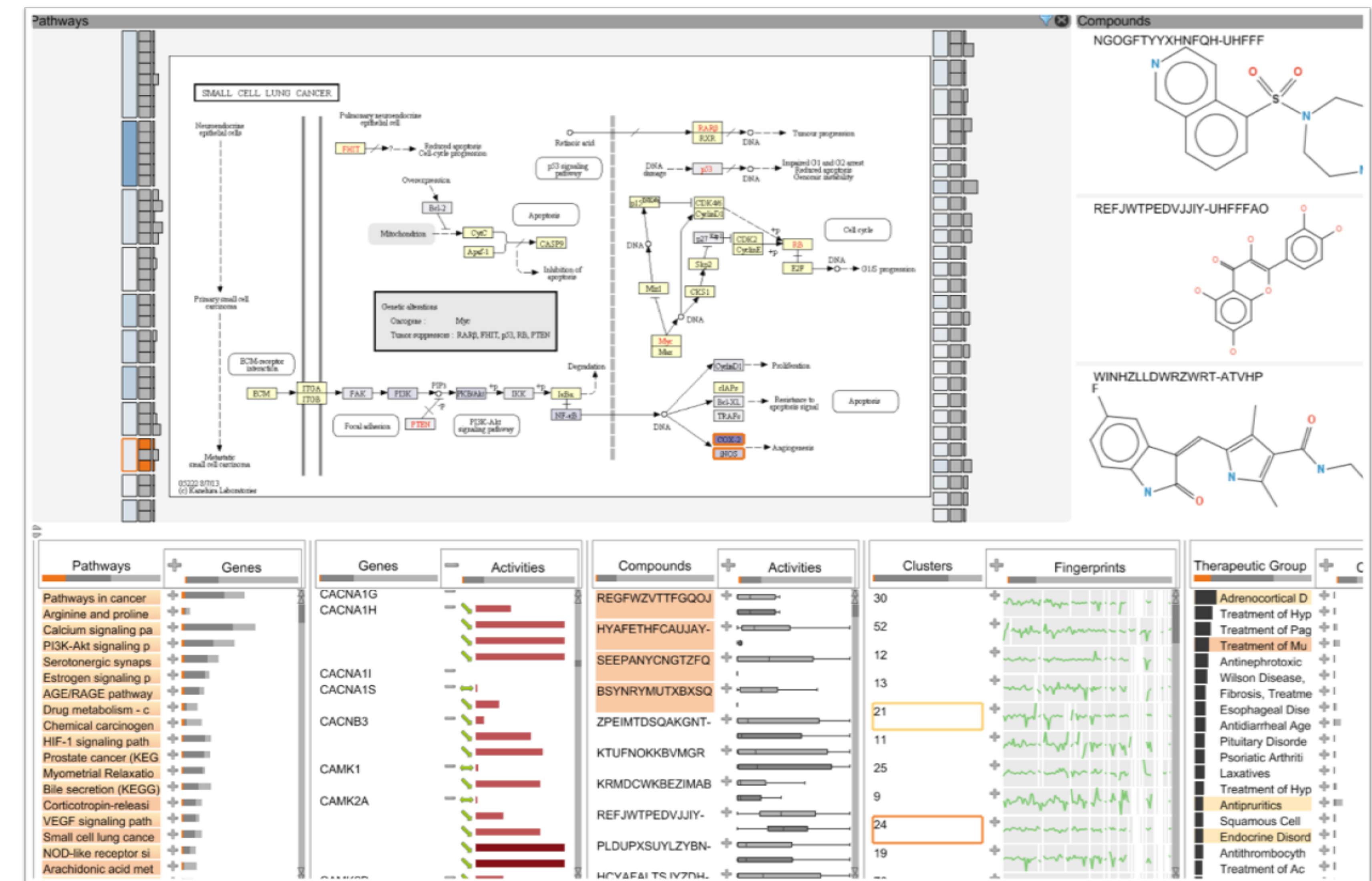


Solve Important Domain Problems

Some problems too specific for
general methods

High impact when solving them

Inspiration for research



Alexander Lex
@alexander_lex
<http://alexander-lex.com>

University of Utah
Scientific Computing Institute
December 3, 2014



HARVARD
School of Engineering
and Applied Sciences

Enabling Scientific Discovery through Interactive Visual Data Analysis



Credits:

Marc Streit, Nils Gehlenborg,
Christian Partl, Samuel Gratzl,
Markus Steinberger, Manuela
Waldner, Hendrik Strobelt, Romain
Vuillemot, Dieter Schmalstieg, Denis
Kalkofen, Mark Borowsky, Anne Mai
Wasserman, Hanspeter Pfister

