Pixel-Adaptive Visual Comparison Between Many Phylogenetic Trees

Tamara Munzner
Department of Computer Science
University of British Columbia

Asilomar Microcomputer Workshop #50
25 Apr 2024

http://www.cs.ubc.ca/~tmm/talks.html#amw24

DESIGNING for PEOPLE

CAIDA

UBC Data Science Institute

@tamara@vis.social

@tamaramunzner

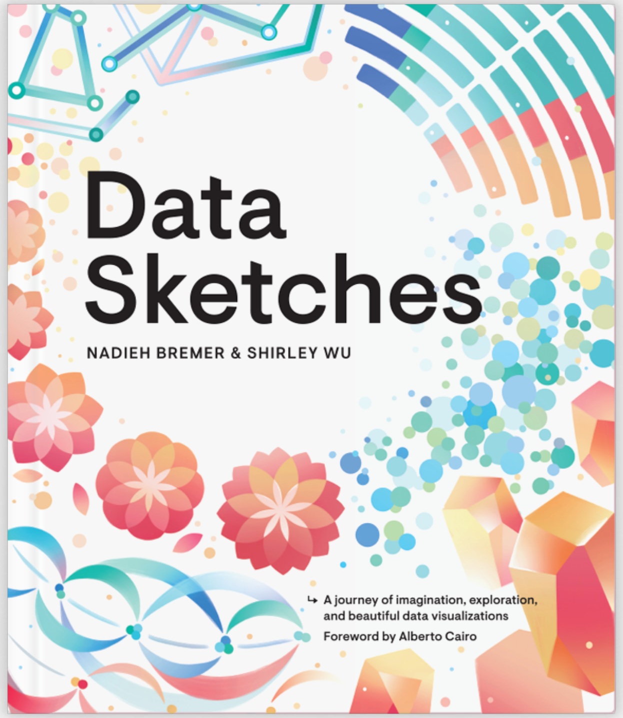
Hi again, 15 years later!

- still doing data visualization
  - yet more papers / projects / videos / software
  - I edit a book series

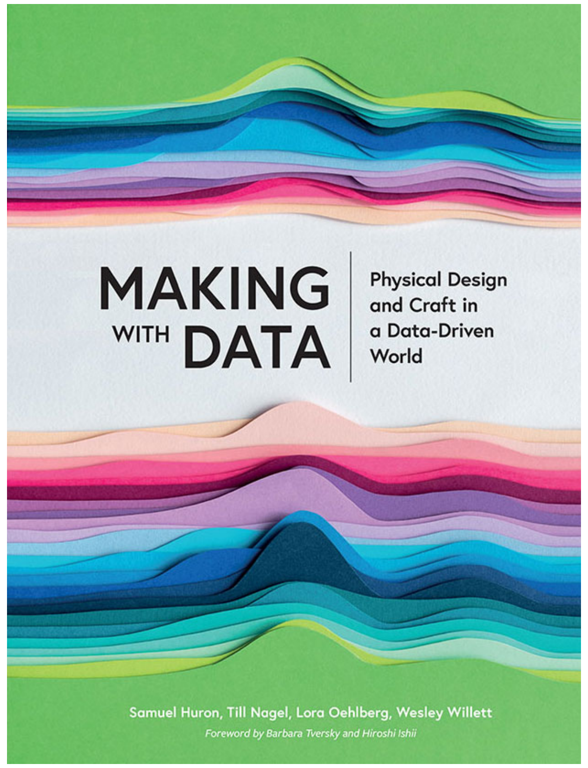
http://www.cs.ubc.ca/~tmm/talks.html#amw24

Visualization book series highlights

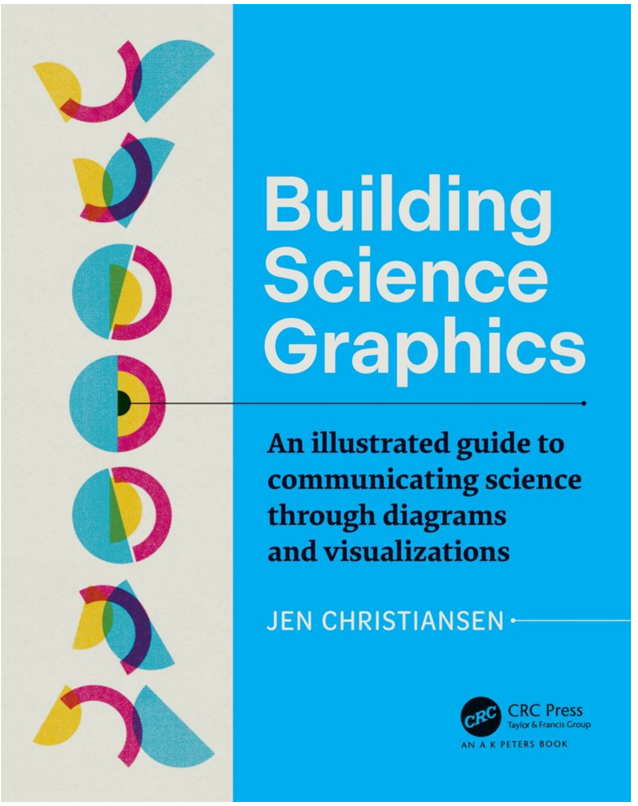
- Data Sketches, by Nadieh Bremer & Shirley Wu



- Making with Data, by multitudes



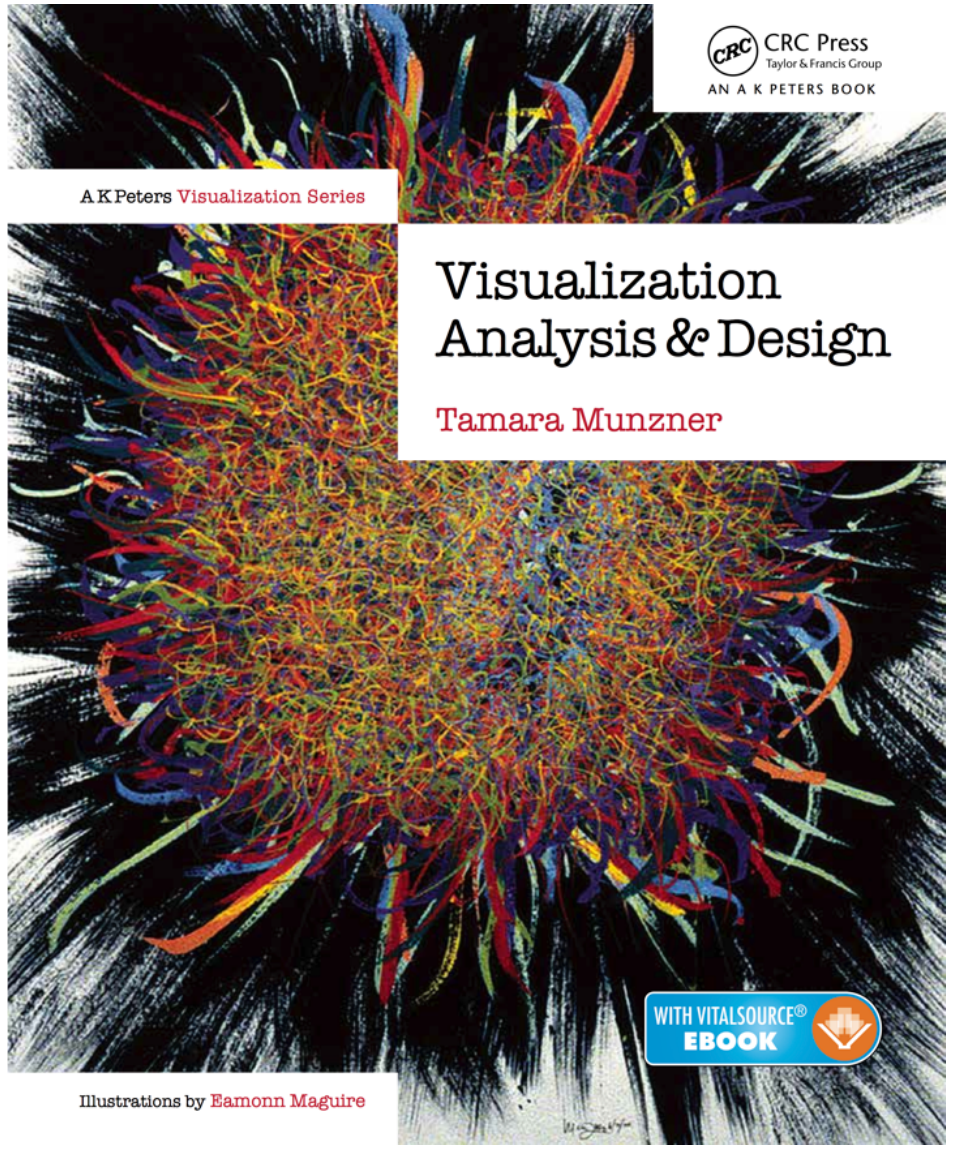
- Building Science Graphics, by Jen Christiansen



https://www.routledge.com/AK-Peters-Visualization-Series/book-series/CRCVIS

Hi again, 15 years later!

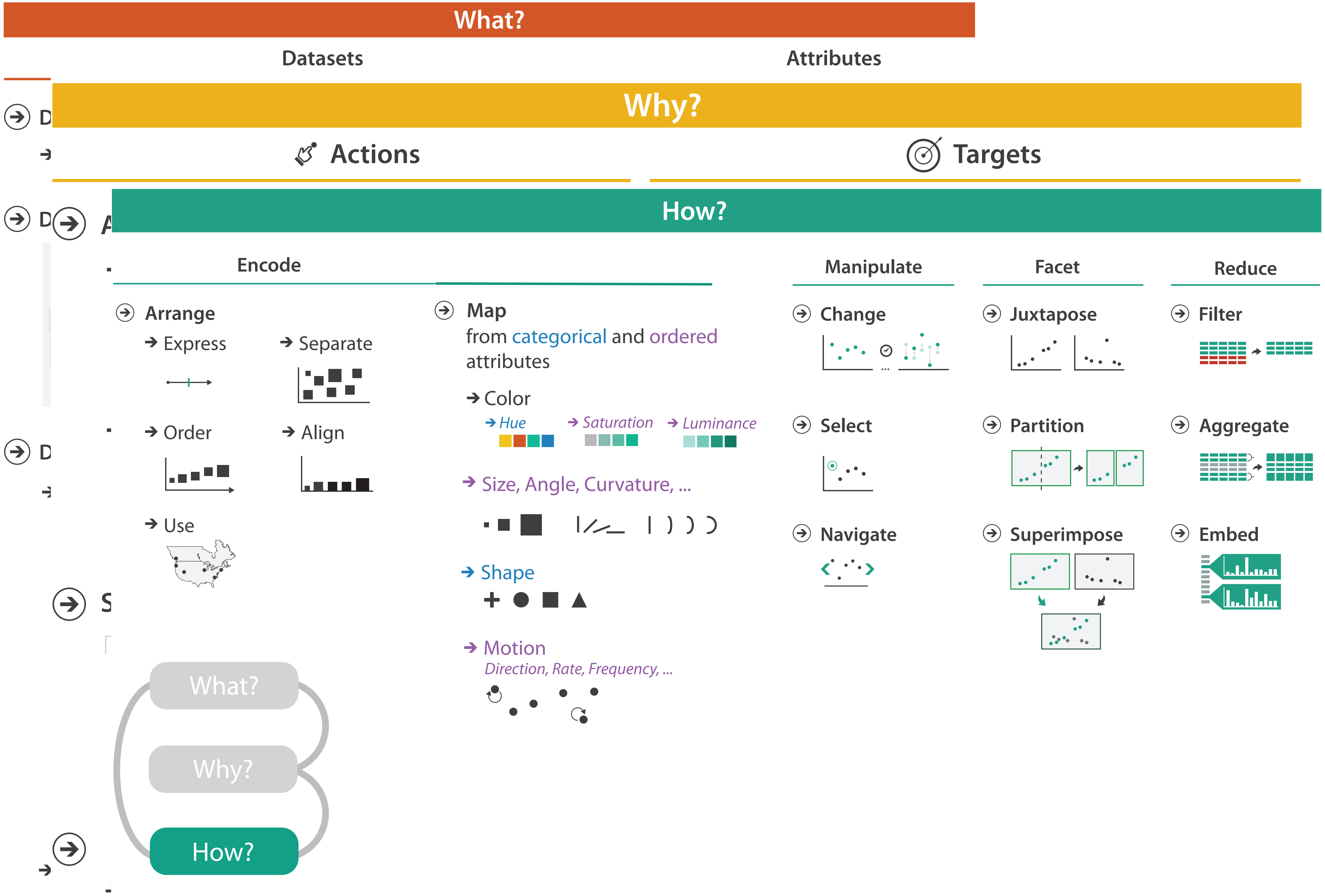
- still doing data visualization
  - yet more papers / projects / videos / software
  - I edit a book series
  - I wrote a book

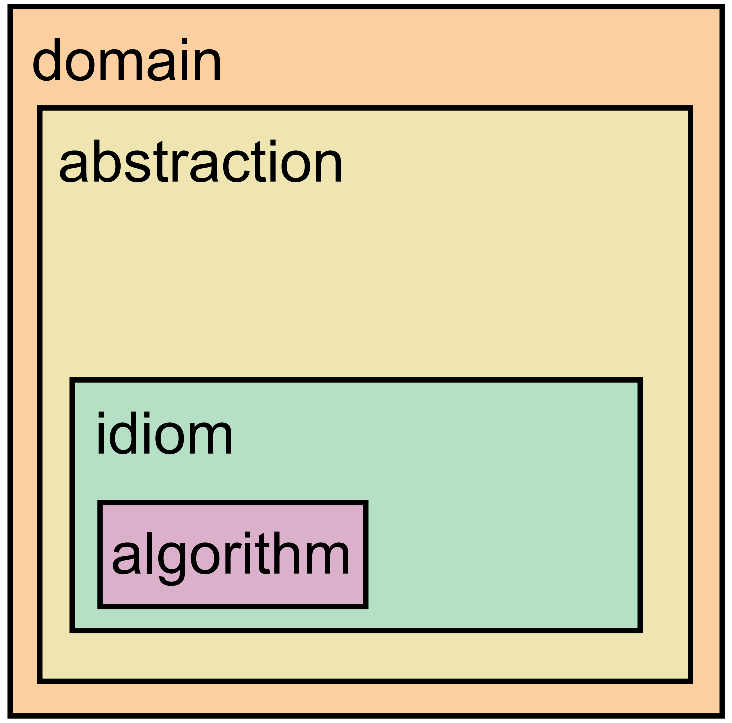


Visualization Analysis and Design. Munzner. *CRC Press, AK Peters Visualization Series, 2014.*

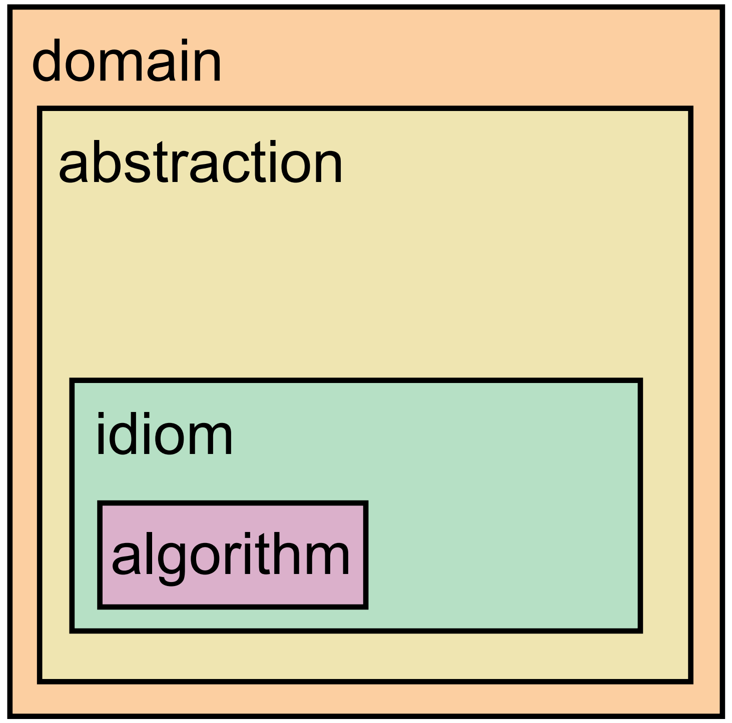
http://www.cs.ubc.ca/~tmm/talks.html#amw24

Visualization Analysis & Design book





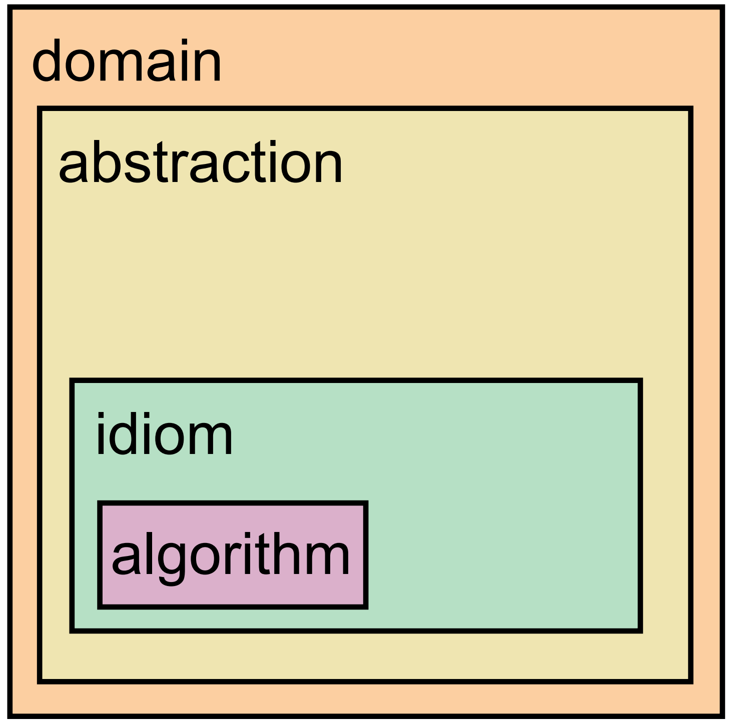
Nested model: Four levels of visualization concerns



[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Nested model: Four levels of visualization concerns

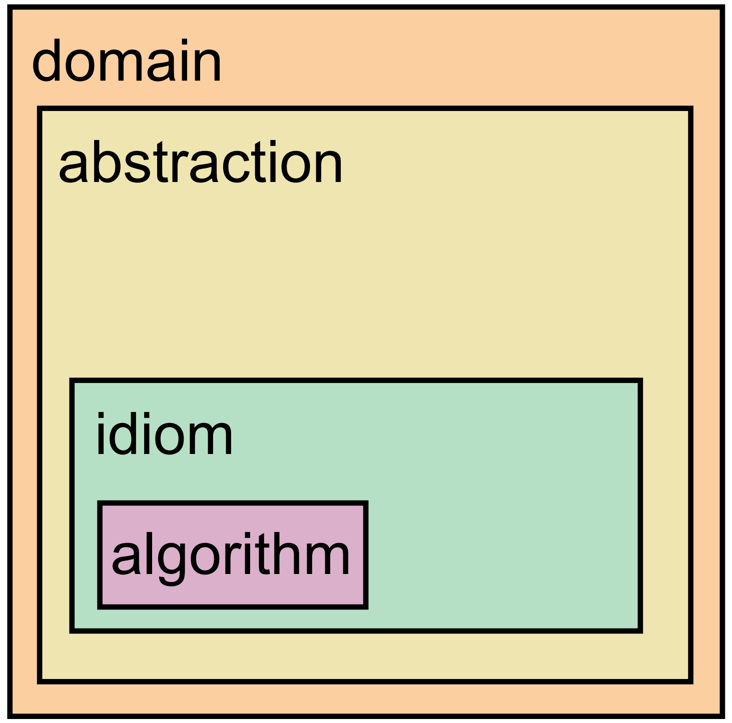
- *domain* situation
  - **who** are the target users?



[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Nested model: Four levels of visualization concerns

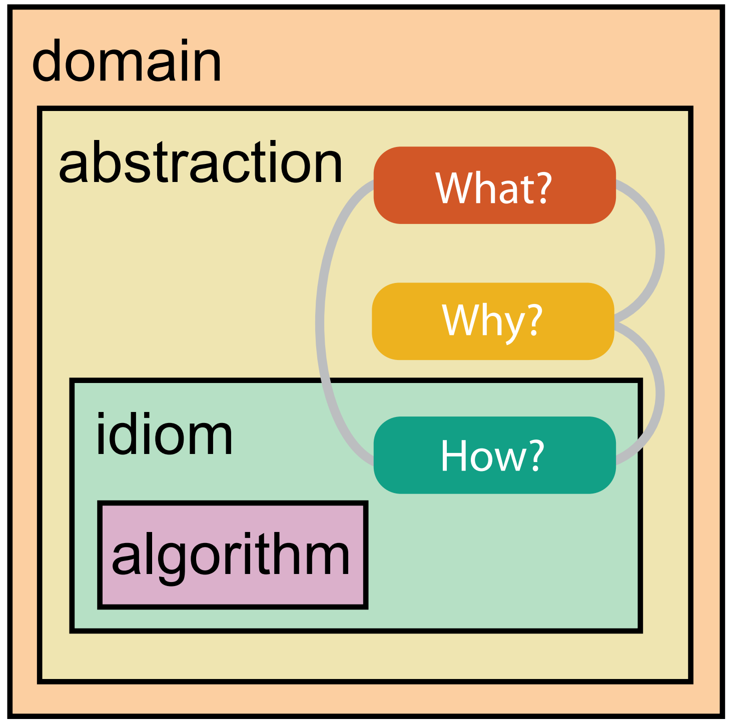
- *domain* situation
  - **who** are the target users?
- *abstraction*
  - translate from specifics of domain to vocabulary of vis



[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Nested model: Four levels of visualization concerns

- *domain* situation
  - **who** are the target users?
- *abstraction*
  - translate from specifics of domain to vocabulary of vis
  - **what** is shown? **data abstraction**

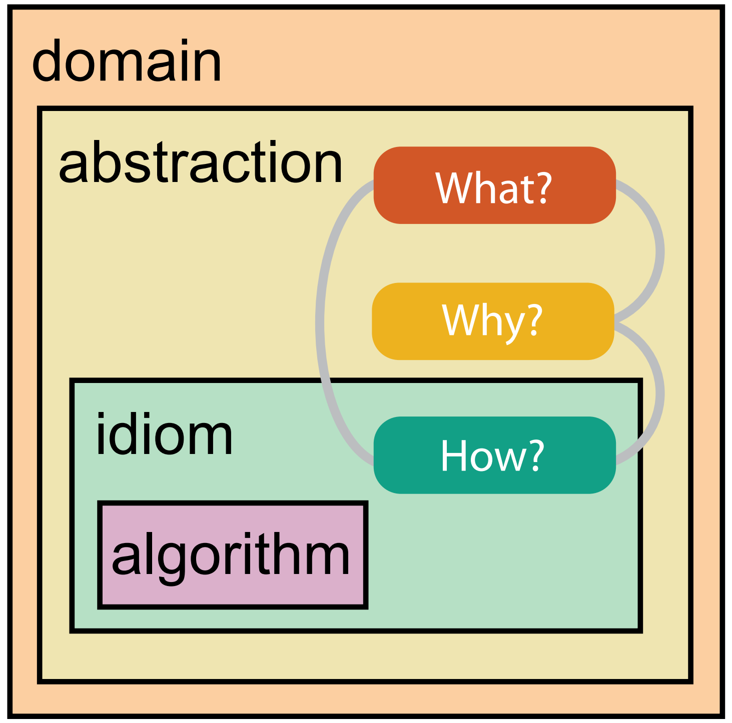


[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

[A Multi-Level Typology of Abstract Visualization Tasks Brehmer and Munzner. IEEE TVCG 19(12):2376-2385, 2013 (Proc. InfoVis 2013).]

Nested model: Four levels of visualization concerns

- *domain* situation
  - **who** are the target users?
- *abstraction*
  - translate from specifics of domain to vocabulary of vis
  - **what** is shown? **data abstraction**
    - often don't just draw what you're given: transform to new form

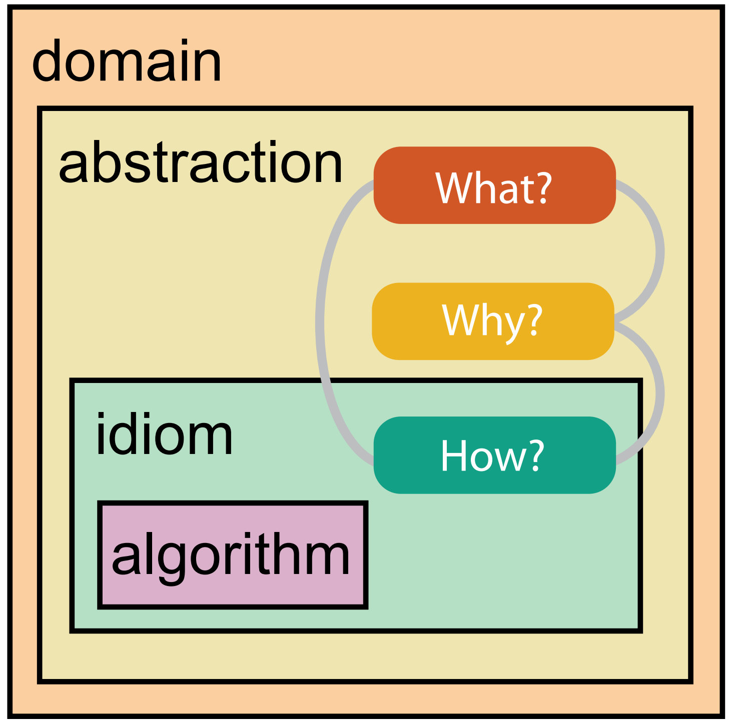


[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

[A Multi-Level Typology of Abstract Visualization Tasks Brehmer and Munzner. IEEE TVCG 19(12):2376-2385, 2013 (Proc. InfoVis 2013).]

Nested model: Four levels of visualization concerns

- *domain* situation
  - **who** are the target users?
- *abstraction*
  - translate from specifics of domain to vocabulary of vis
  - **what** is shown? **data abstraction**
    - often don't just draw what you're given: transform to new form
  - **why** is the user looking at it? **task abstraction**

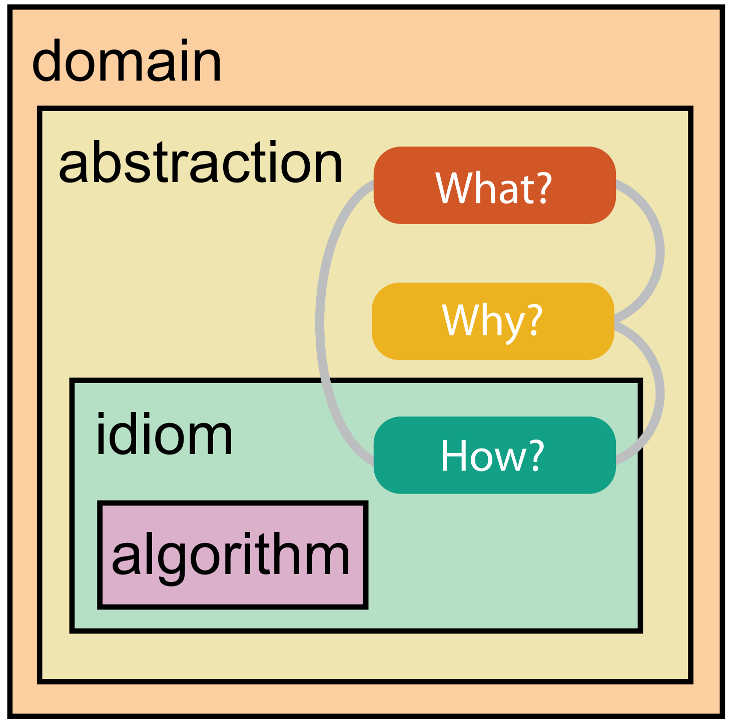


[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

[A Multi-Level Typology of Abstract Visualization Tasks Brehmer and Munzner. IEEE TVCG 19(12):2376-2385, 2013 (Proc. InfoVis 2013).]

Nested model: Four levels of visualization concerns

- *domain* situation
  - **who** are the target users?
- *abstraction*
  - translate from specifics of domain to vocabulary of vis
  - **what** is shown? **data abstraction**
    - often don’t just draw what you’re given: transform to new form
  - **why** is the user looking at it? **task abstraction**
- *idiom*
  - **how** is it shown?

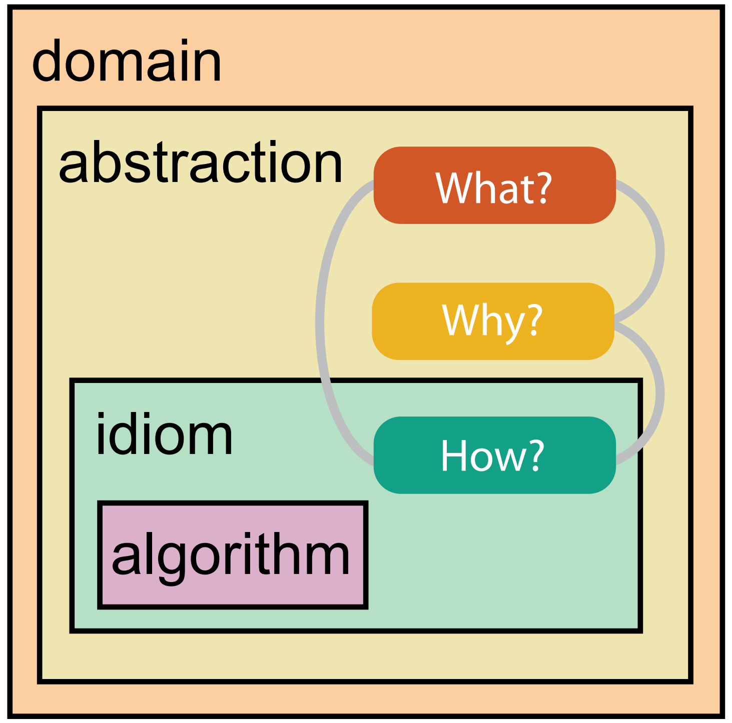


[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

[A Multi-Level Typology of Abstract Visualization Tasks Brehmer and Munzner. IEEE TVCG 19(12):2376-2385, 2013 (Proc. InfoVis 2013).]

Nested model: Four levels of visualization concerns

- *domain* situation
  - **who** are the target users?
- *abstraction*
  - translate from specifics of domain to vocabulary of vis
  - **what** is shown? **data abstraction**
    - often don't just draw what you're given: transform to new form
  - **why** is the user looking at it? **task abstraction**
- *idiom*
  - **how** is it shown?
    - **visual encoding idiom**: how to draw

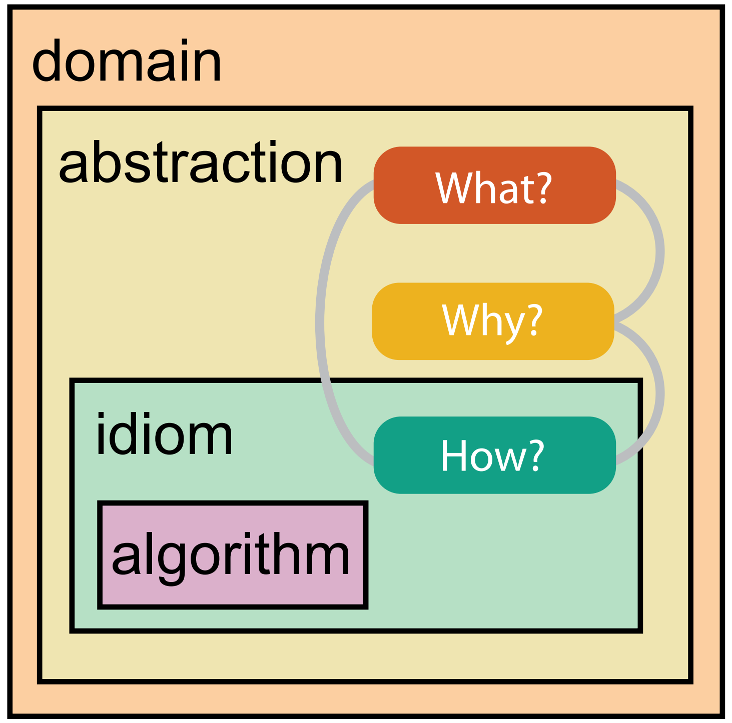


[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

[A Multi-Level Typology of Abstract Visualization Tasks Brehmer and Munzner. IEEE TVCG 19(12):2376-2385, 2013 (Proc. InfoVis 2013).]

Nested model: Four levels of visualization concerns

- *domain* situation
  - **who** are the target users?
- *abstraction*
  - translate from specifics of domain to vocabulary of vis
  - **what** is shown? **data abstraction**
    - often don't just draw what you're given: transform to new form
  - **why** is the user looking at it? **task abstraction**
- *idiom*
  - **how** is it shown?
    - **visual encoding idiom**: how to draw
    - **interaction idiom**: how to manipulate

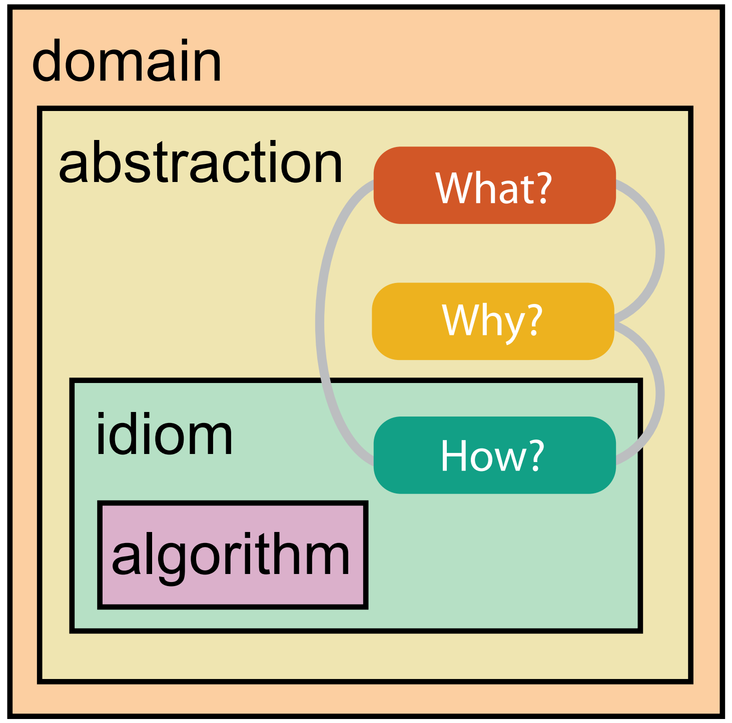


[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

[A Multi-Level Typology of Abstract Visualization Tasks Brehmer and Munzner. IEEE TVCG 19(12):2376-2385, 2013 (Proc. InfoVis 2013).]

Nested model: Four levels of visualization concerns

- *domain* situation
  - **who** are the target users?
- *abstraction*
  - translate from specifics of domain to vocabulary of vis
  - **what** is shown? **data abstraction**
    - often don't just draw what you're given: transform to new form
  - **why** is the user looking at it? **task abstraction**
- *idiom*
  - **how** is it shown?
    - **visual encoding idiom**: how to draw
    - **interaction idiom**: how to manipulate
- *algorithm*
  - efficient computation



[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

[A Multi-Level Typology of Abstract Visualization Tasks Brehmer and Munzner. IEEE TVCG 19(12):2376-2385, 2013 (Proc. InfoVis 2013).]

Why is validation difficult?

- different threats to validity at each level

[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Why is validation difficult?

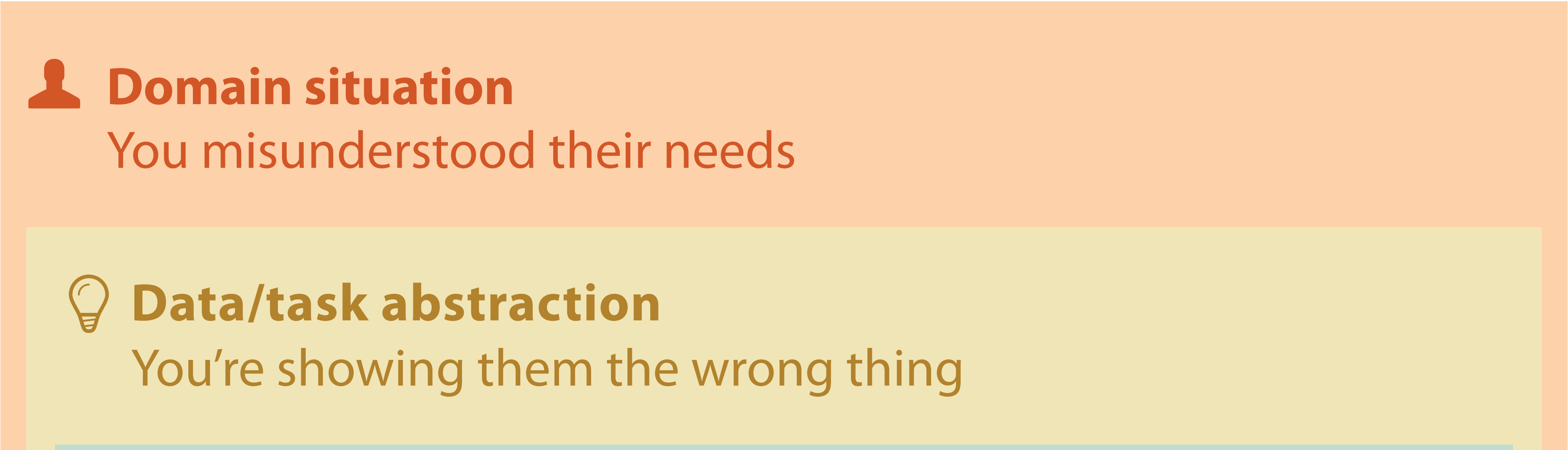
- different threats to validity at each level

Domain situation
You misunderstood their needs

[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Why is validation difficult?

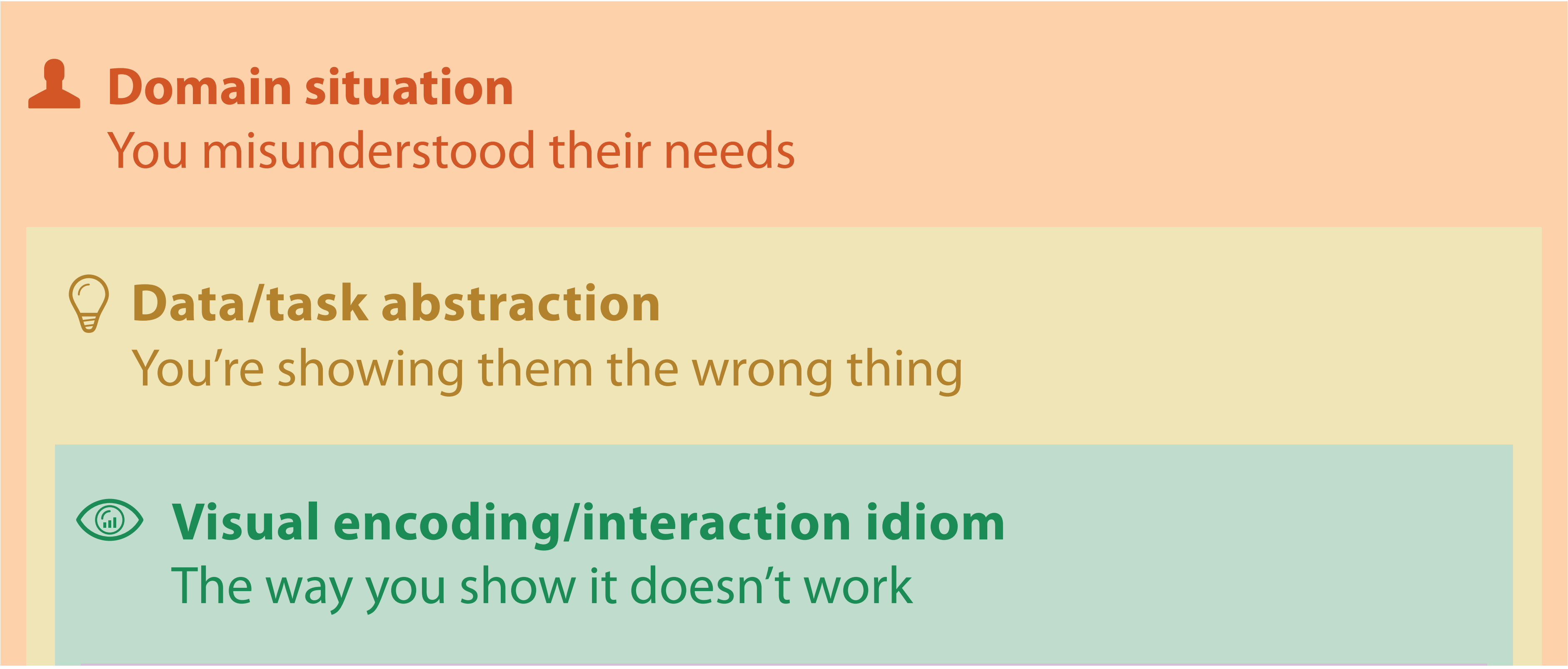
- different threats to validity at each level



[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Why is validation difficult?

- different threats to validity at each level



[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Why is validation difficult?

- different threats to validity at each level

[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Validation solution: Use methods from appropriate fields at each level

[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Validation solution: Use methods from appropriate fields at each level

computer science

Algorithm
Measure system time/memory
Analyze computational complexity

[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Validation solution: Use methods from appropriate fields at each level

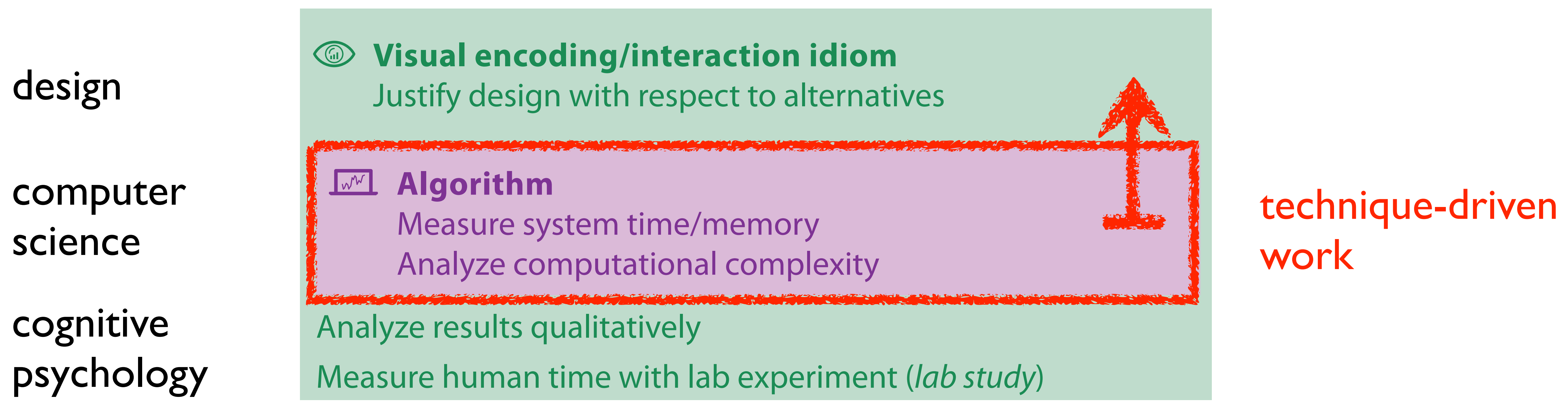
computer
science

Algorithm
Measure system time/memory
Analyze computational complexity

technique-driven
work

[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Validation solution: Use methods from appropriate fields at each level



[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Validation solution: Use methods from appropriate fields at each level

anthropology/ ethnography

Domain situation
Observe target users using existing tools

Data/task abstraction

design

Visual encoding/interaction idiom
Justify design with respect to alternatives

computer science

Algorithm
Measure system time/memory
Analyze computational complexity

technique-driven work

cognitive psychology

Analyze results qualitatively

Measure human time with lab experiment (*lab study*)

anthropology/ ethnography

Observe target users after deployment (*field study*)

Measure adoption

[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Validation solution: Use methods from appropriate fields at each level

- avoid mismatches between level and validation

anthropology/ ethnography

Domain situation
Observe target users using existing tools

Data/task abstraction

design

Visual encoding/interaction idiom
Justify design with respect to alternatives

computer science

Algorithm
Measure system time/memory
Analyze computational complexity

technique-driven work

cognitive psychology

Analyze results qualitatively

Measure human time with lab experiment (*lab study*)

anthropology/ ethnography

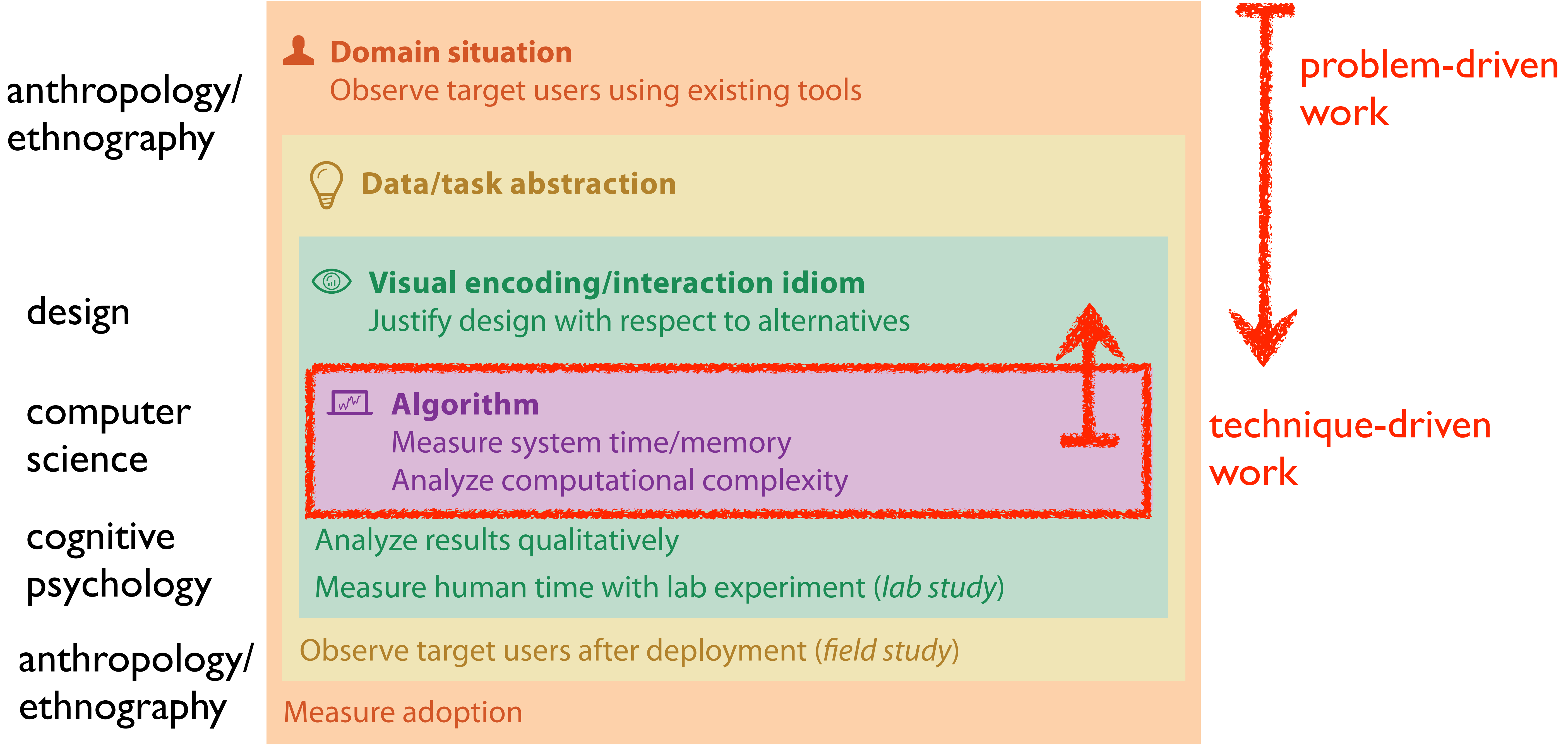
Observe target users after deployment (*field study*)

Measure adoption

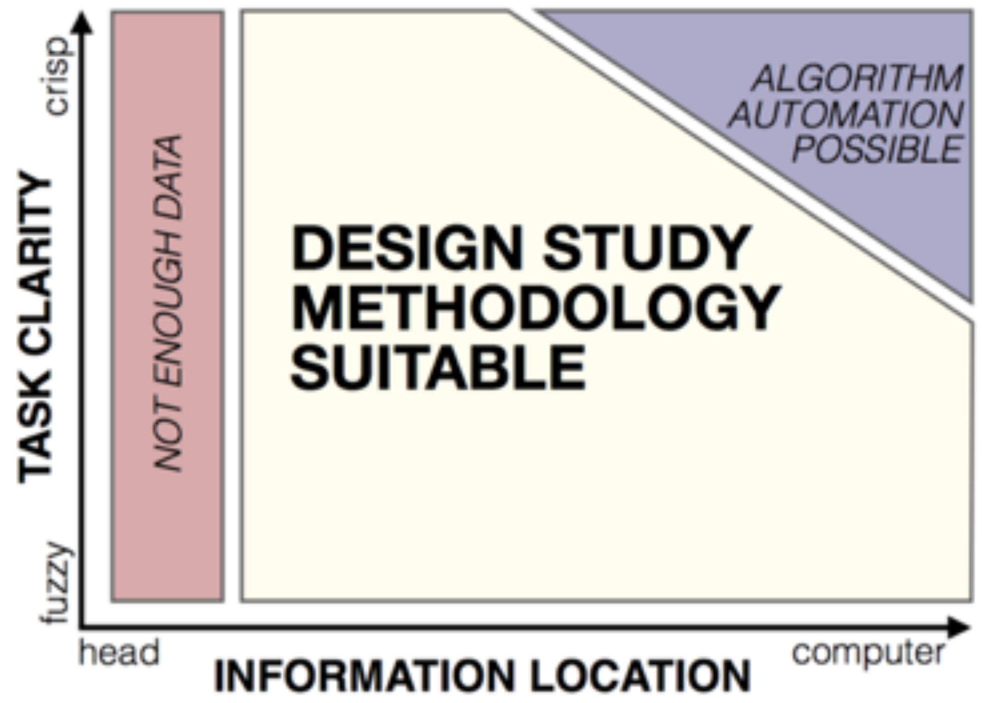
[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Validation solution: Use methods from appropriate fields at each level

- avoid mismatches between level and validation



[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]



Michael Sedlmair

Miriah Meyer

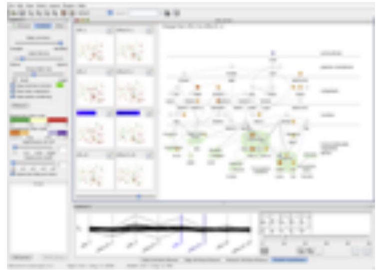
Design Study Methodology

Reflections from the Trenches and from the Stacks

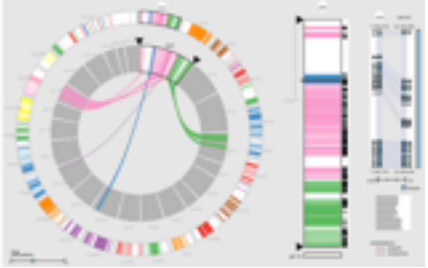
http://www.cs.ubc.ca/labs/imager/tr/2012/dsm/

Design Study Methodology: Reflections from the Trenches and from the Stacks.
Sedlmair, Meyer, Munzner. IEEE Trans. Visualization and Computer Graphics 18(12): 2431-2440, 2012 (Proc. InfoVis 2012).

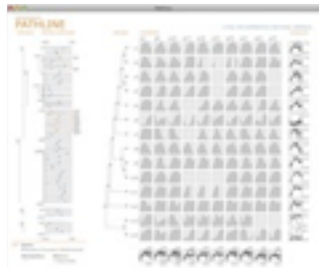
Lessons learned from the trenches: 20+ between us



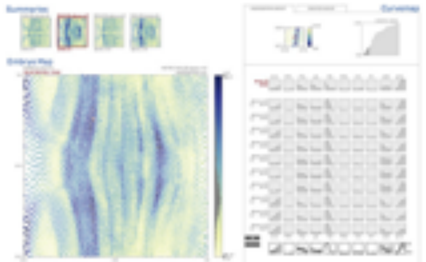
Cerebral
genomics



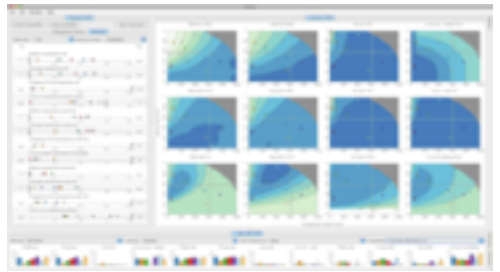
MizBee
genomics



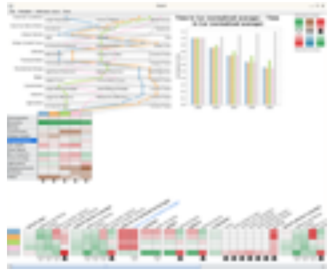
Pathline
genomics



MulteeSum
genomics



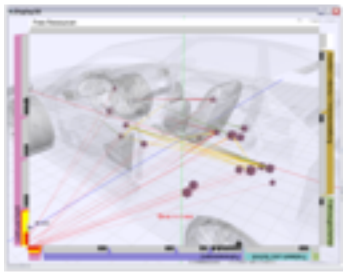
Vismon
fisheries management



QuestVis
sustainability

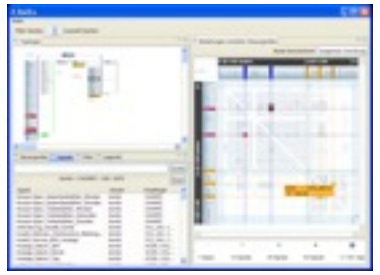
WiKeVis
in-car networks

MostVis
in-car networks

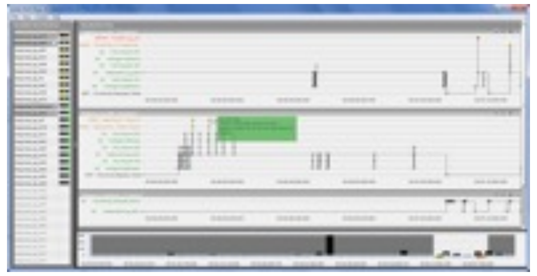


Car-X-Ray
in-car networks

ProgSpy2010
in-car networks



RelEx
in-car networks

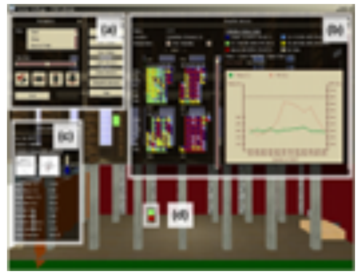


Cardiogram
in-car networks

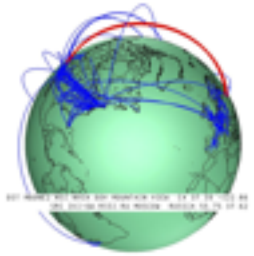
AutobahnVis
in-car networks

VisTra
in-car networks

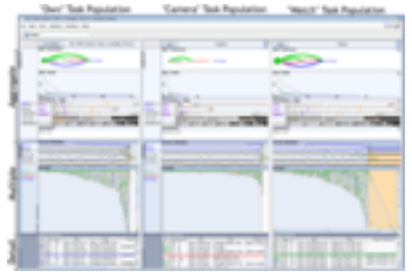
Constellation
linguistics



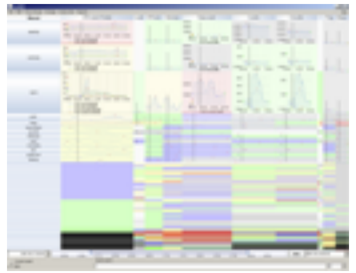
LibVis
cultural heritage



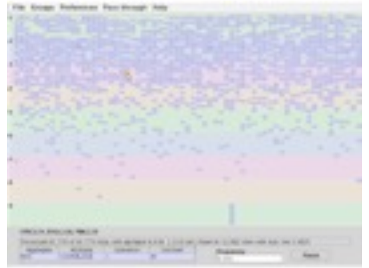
Caidants
multicast



SessionViewer
web log analysis



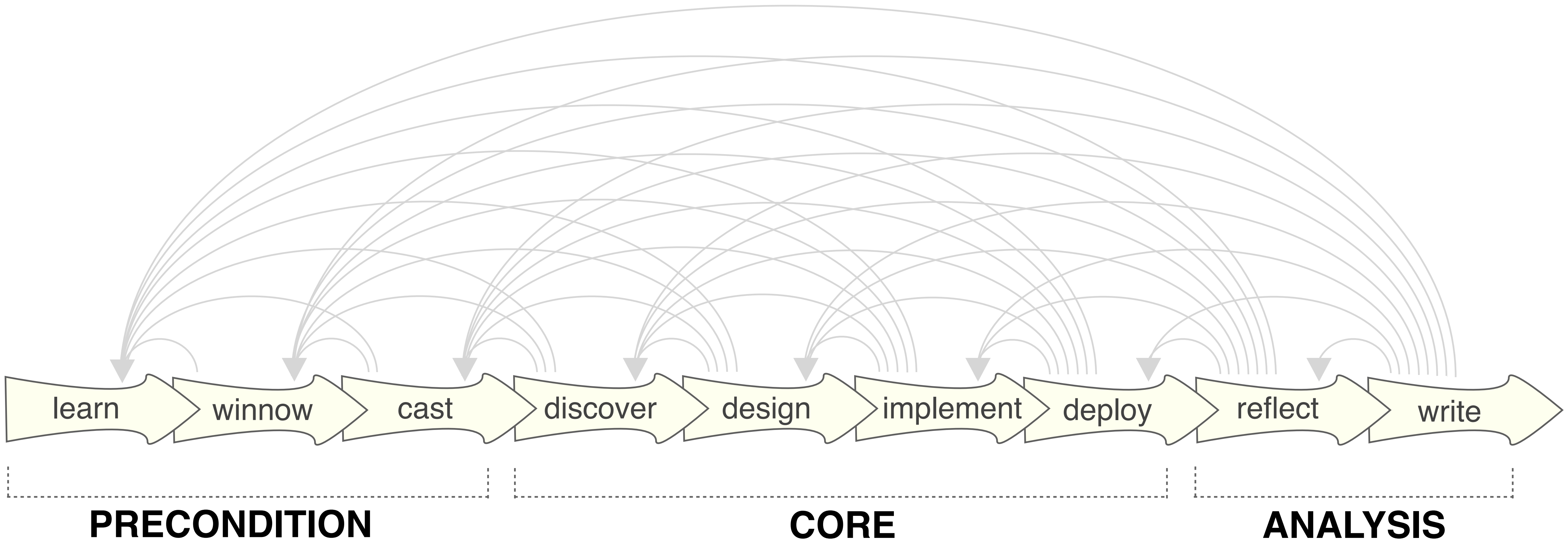
LiveRAC
server hosting



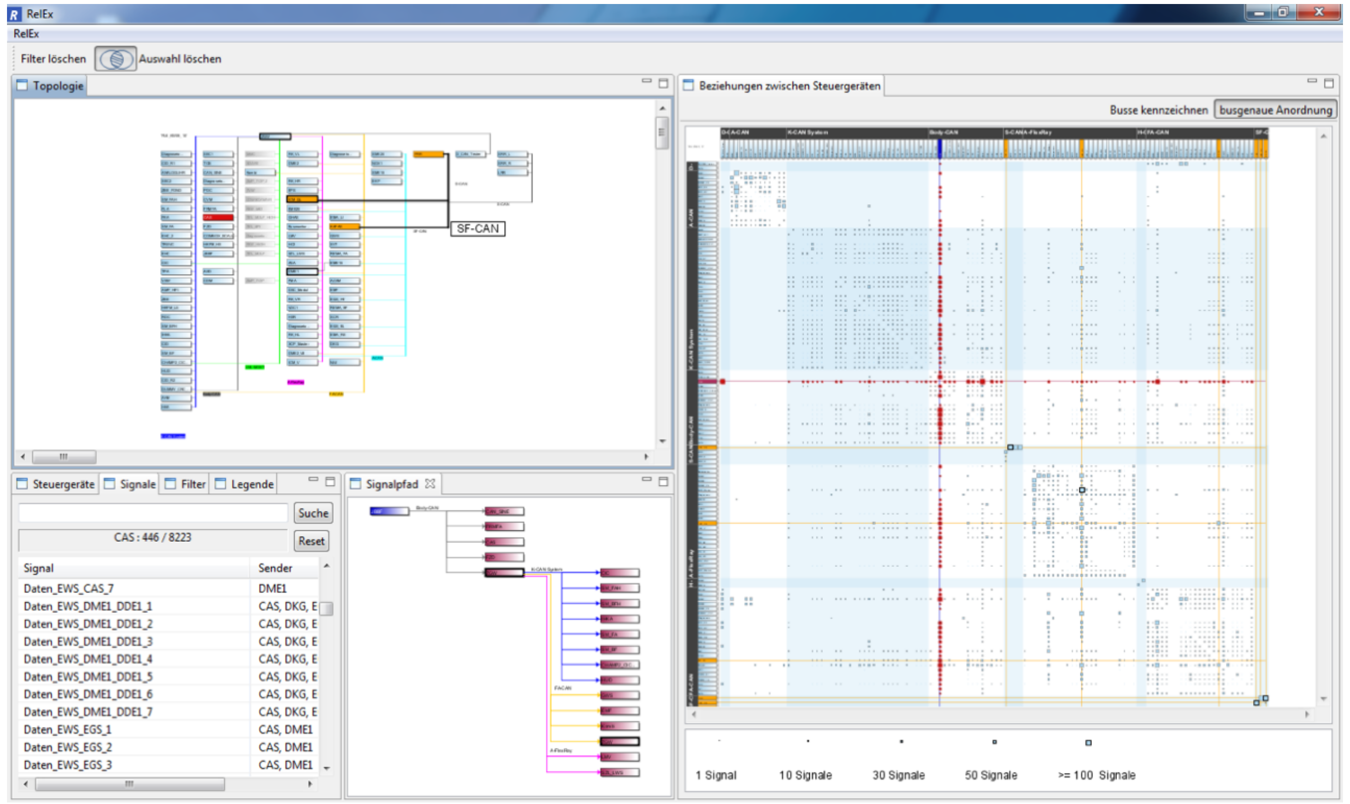
PowerSetViewer
data mining

9-stage framework

iterative



[Design Study Methodology: Reflections from the Trenches and the Stacks. Sedlmair, Meyer & Munzner. IEEE TVCG 18(12): 2431-2440, 2012 (Proc. InfoVis 2012).]



RelEx

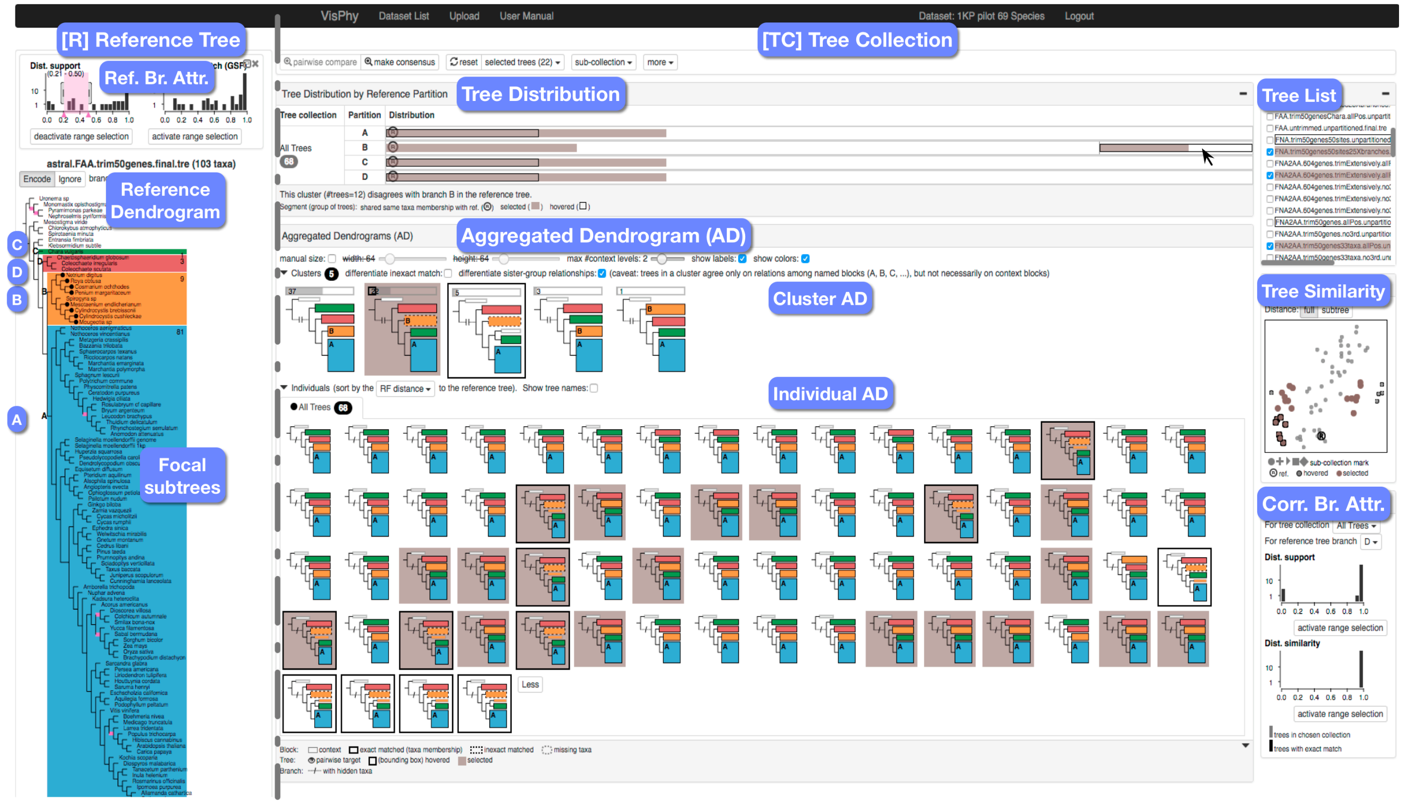
Visualization for Actively Changing Overlay Network Specifications

joint work with:

Michael Sedlmair, Annika Frank, Andreas Butz

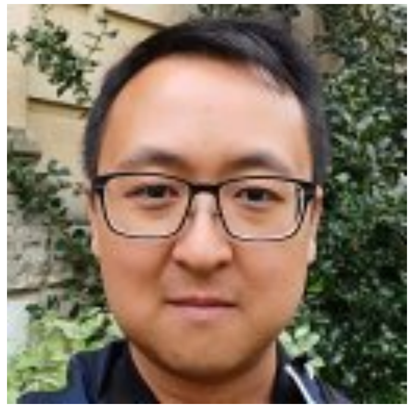
http://www.cs.ubc.ca/labs/imager/tr/2012/relex/

RelEx: Visualization for Actively Changing Overlay Network Specifications.
Sedlmair, Frank, Butz, Munzner. IEEE TVCG 18(12): 2729-2738, 2012 (Proc. InfoVis 2012).



Zipeng Liu

Shing Hei Zhan



Aggregated Dendrograms

for Visual Comparison Between Many Phylogenetic Trees

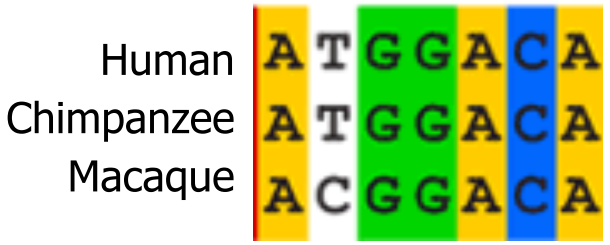
http://www.cs.ubc.ca/labs/imager/tr/2019/adview

Aggregated Dendrograms for Visual Comparison Between Many Phylogenetic Trees.

Liu, Zhan, Munzner. IEEE Trans. Visualization and Computer Graphics (TVCG) 26(9):2732-2747, 2019.

Phylogenetic tree

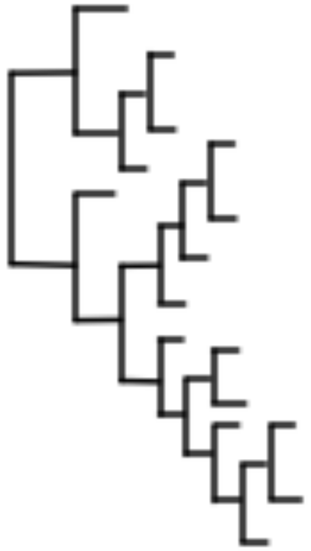
Evolutionary relationships of organisms



Genetic information

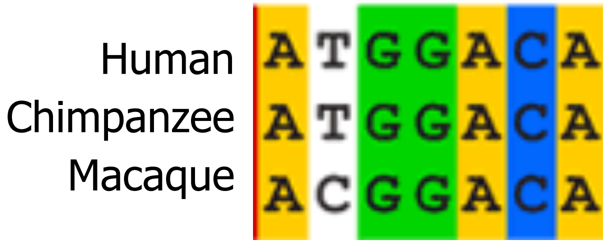
Computational workflow

Phylogenetic tree



Many phylogenetic trees

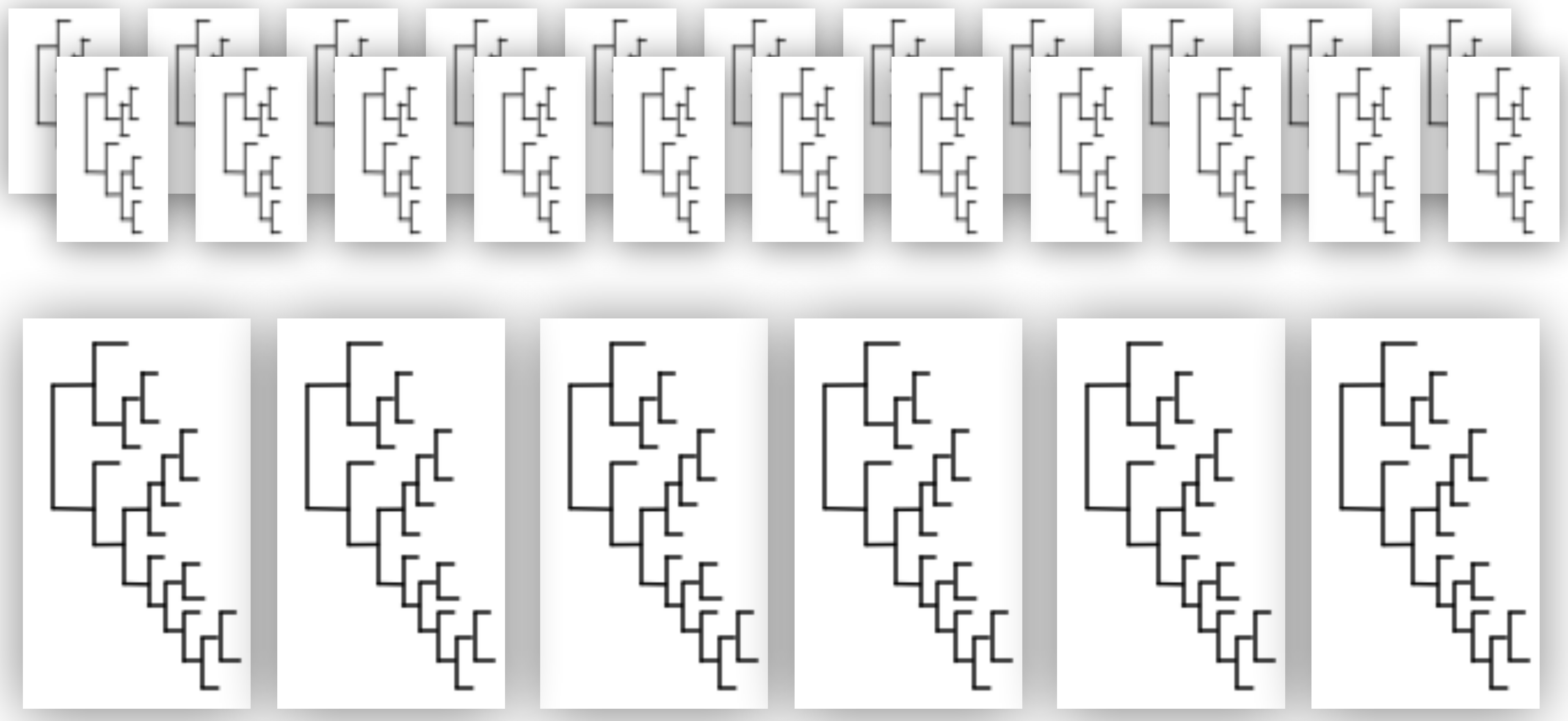
- Understand relationships between genes and species trees
- Explore trees generated with different methods and data



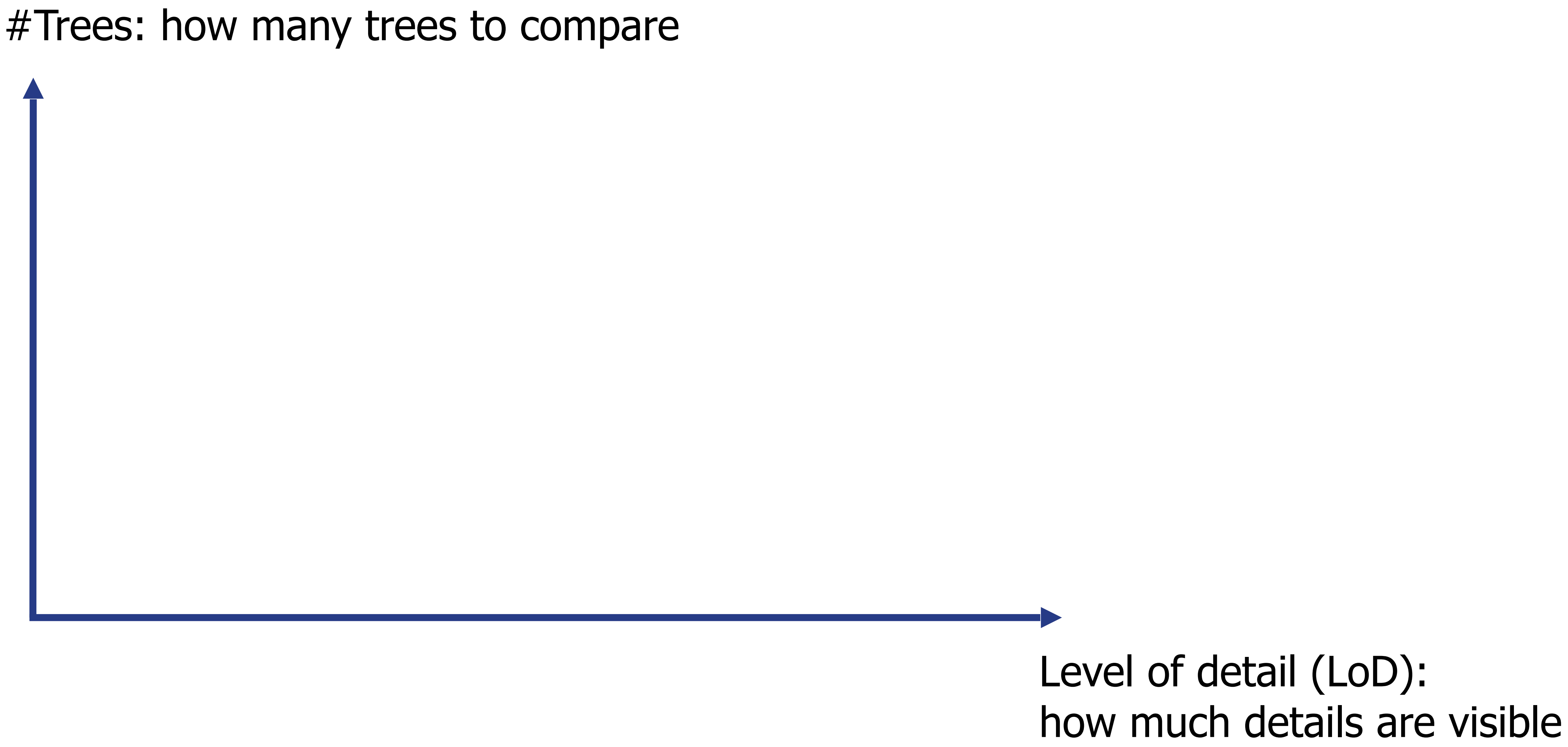
Genetic information

Computational workflow

Phylogenetic tree



Scalability of existing tree comparison systems



Scalability of existing tree comparison systems

#Trees: how many trees to compare

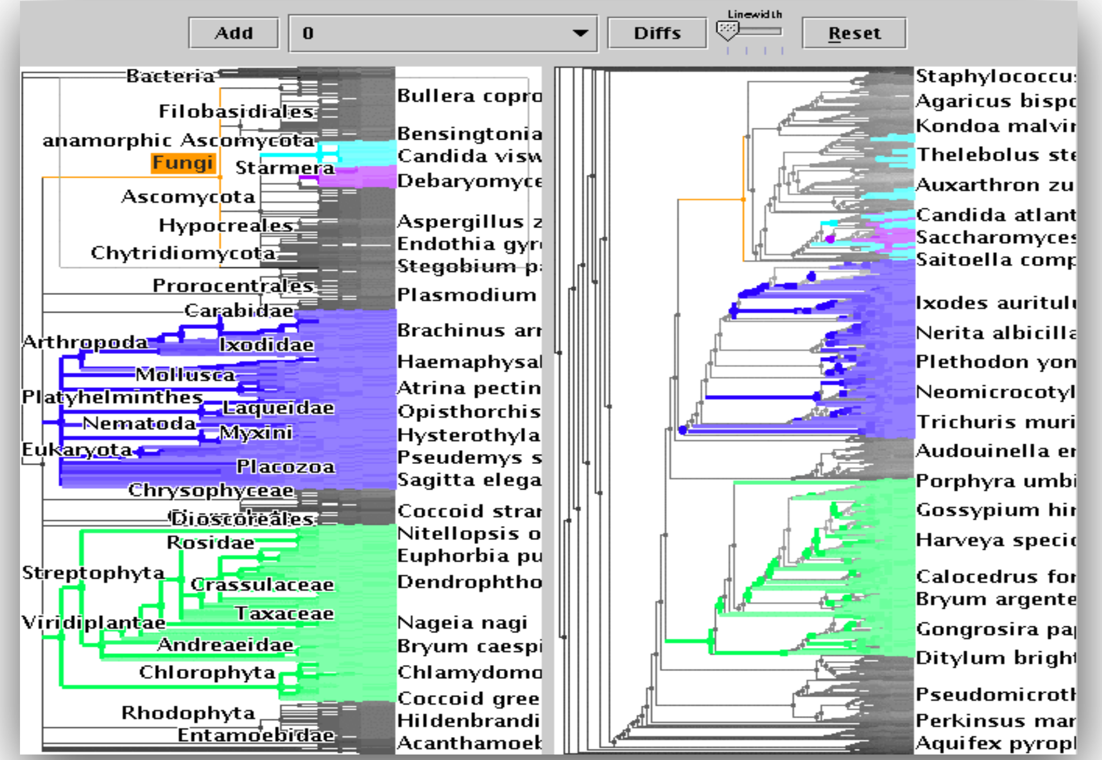
Pairs

Few in full

Simplified structure

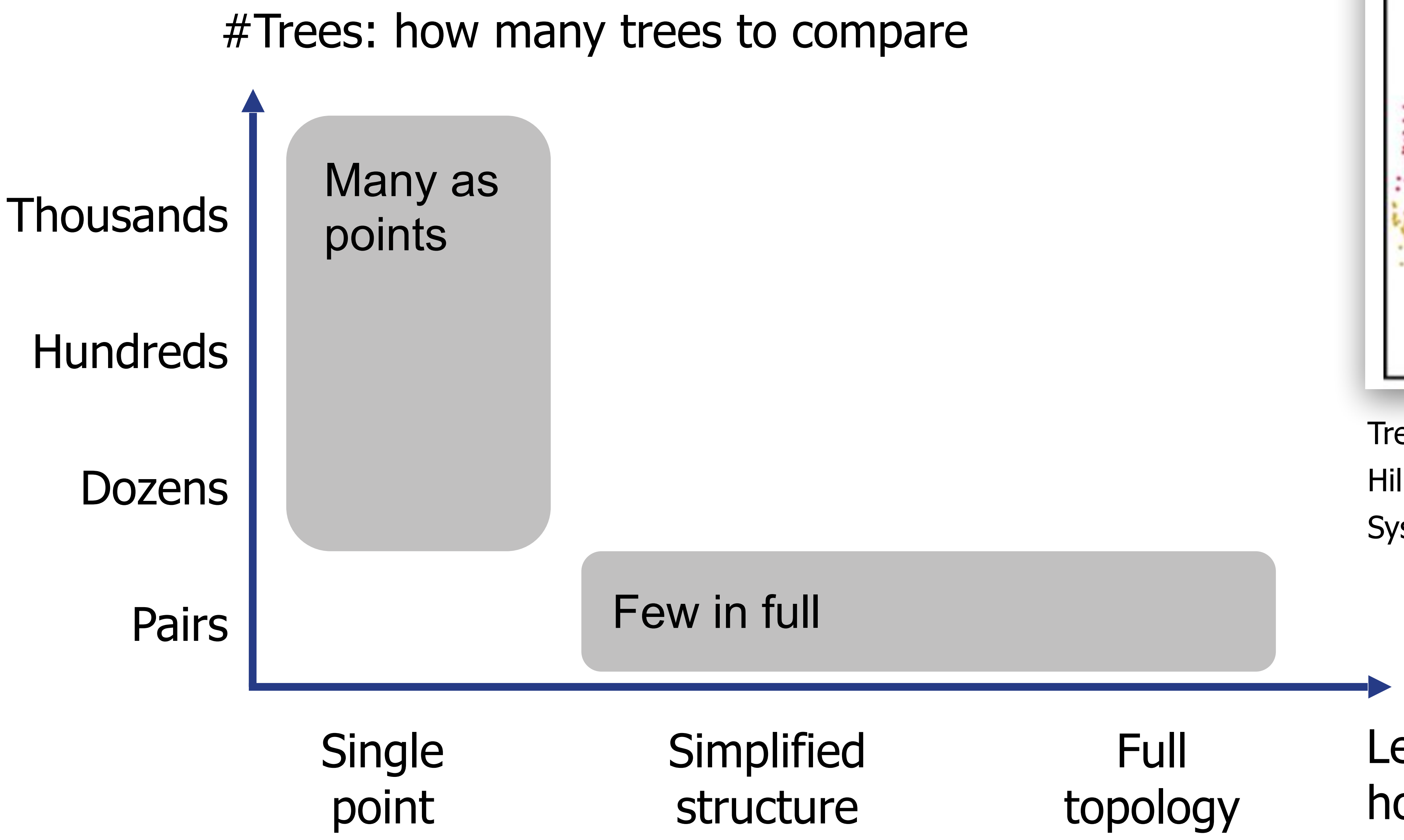
Full topology

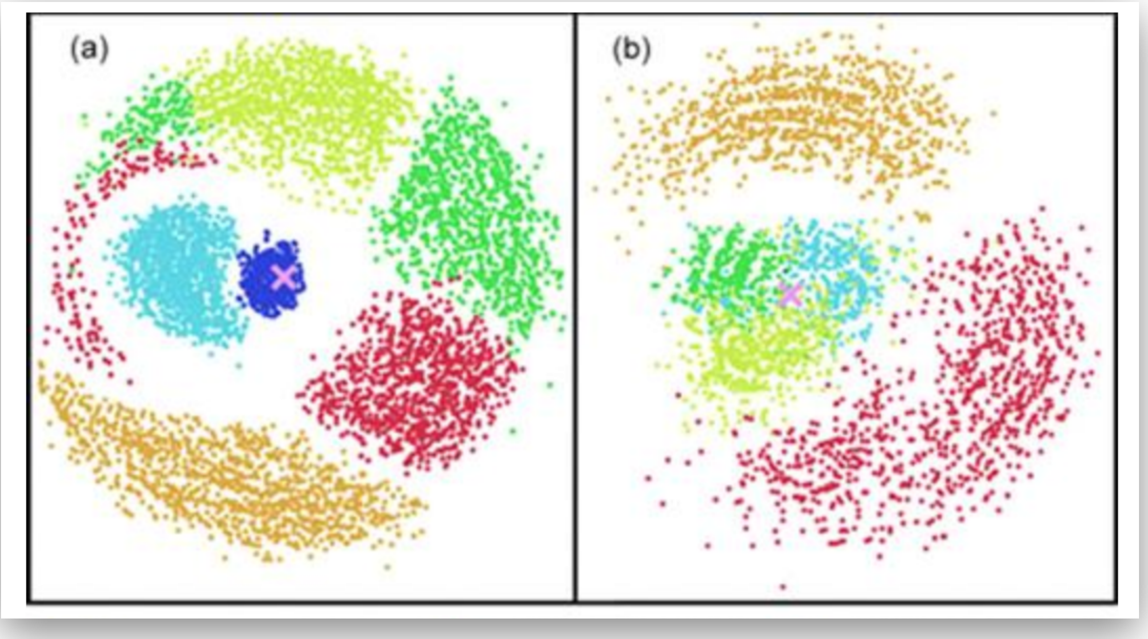
Level of detail (LoD):
how much details are visible



TreeJuxtaposer.
Munzner, Guimbretière, Zhang, Zhou.
SIGGRAPH 2003

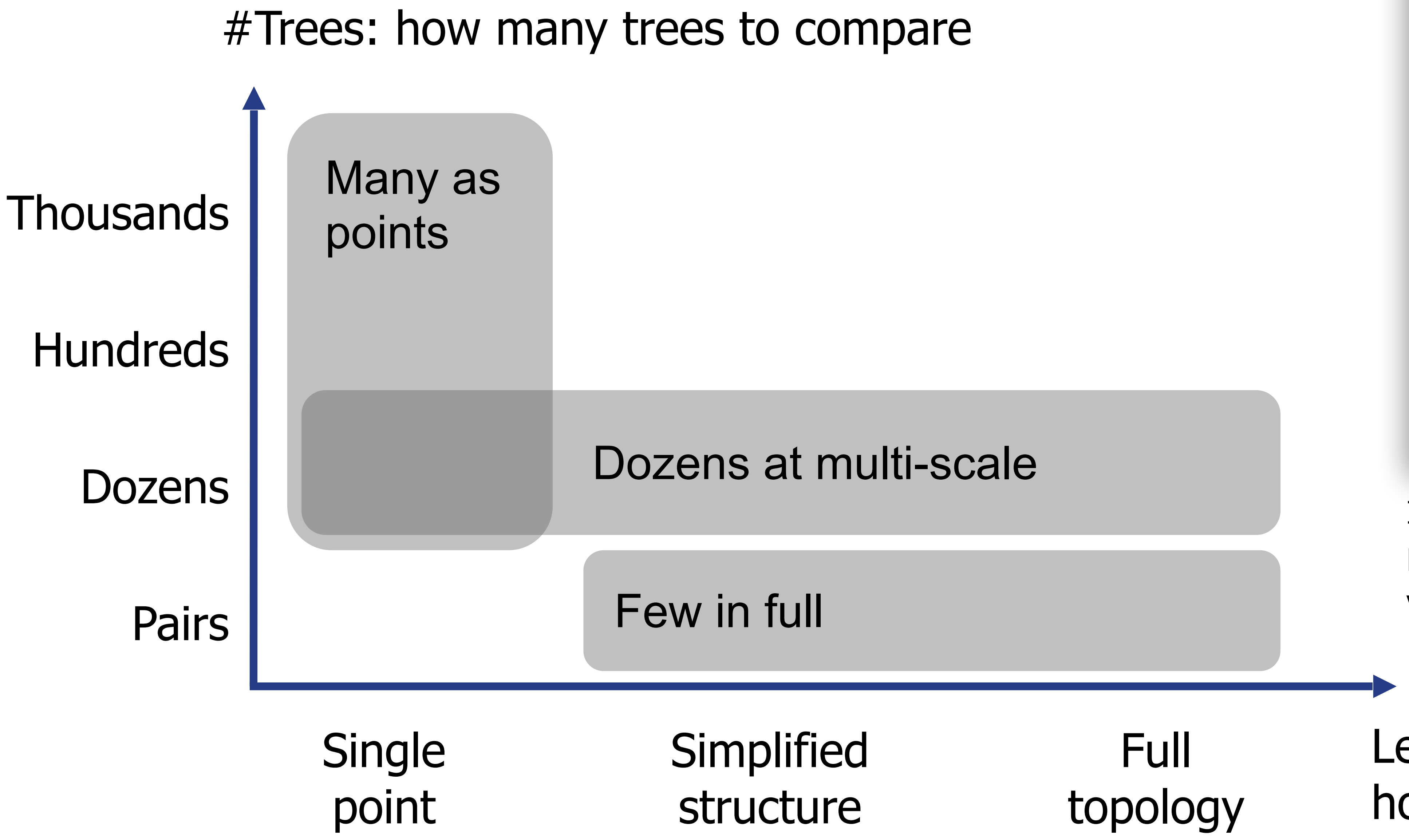
Scalability of existing tree comparison systems

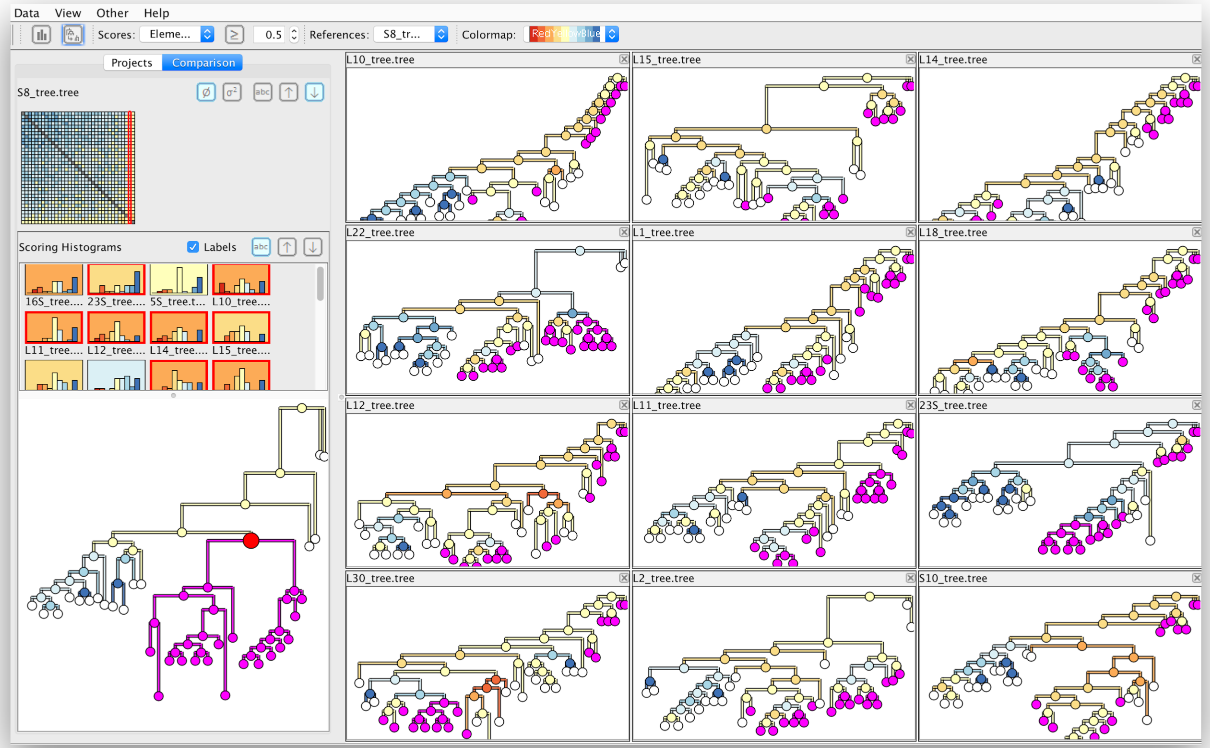




Tree space.
Hillis, Health, John.
Systematic Biology 2005.

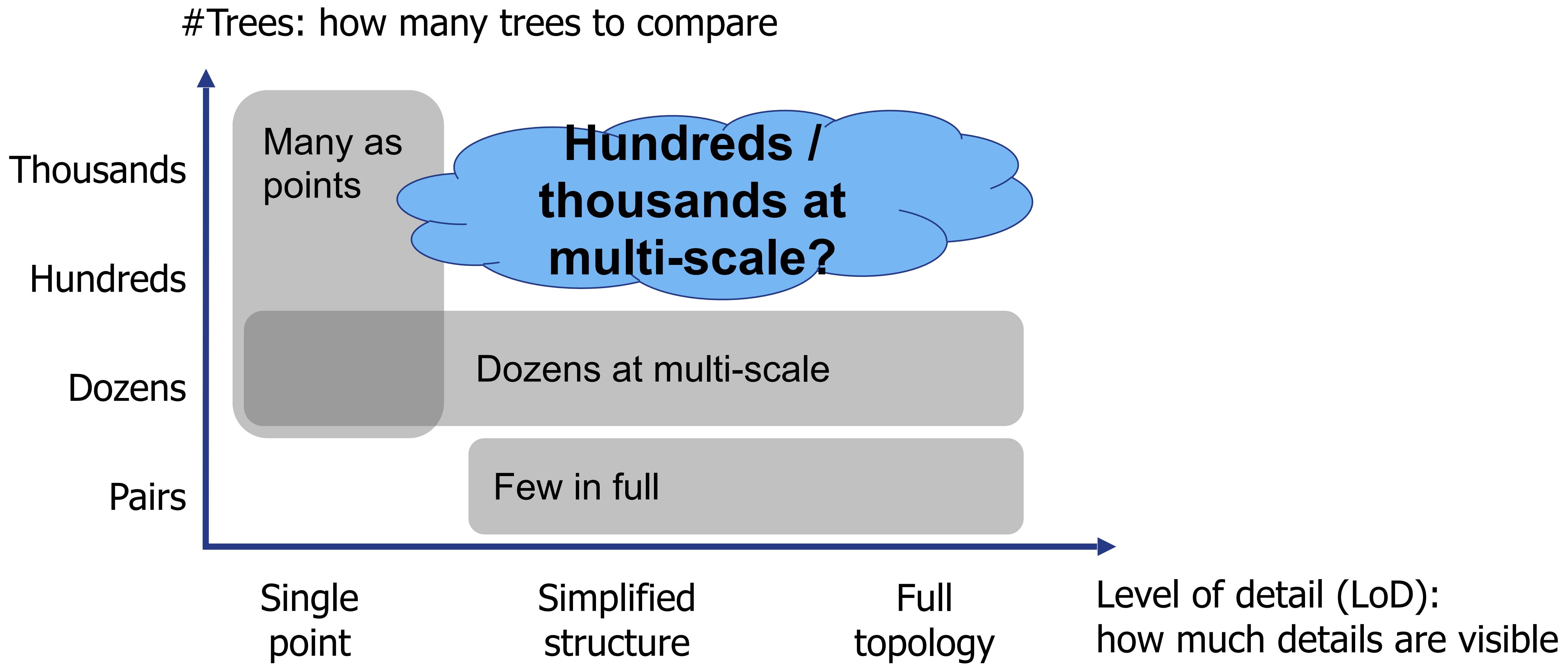
Scalability of existing tree comparison systems





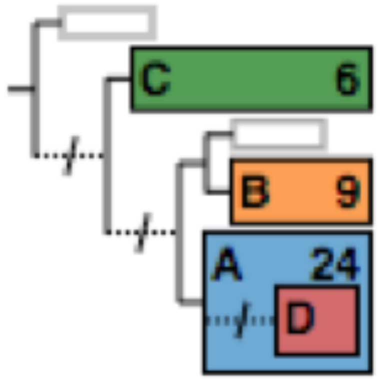
Interactive visual comparison of multiple trees.
Bremm, Landesberger, Heß, Schreck, Weil, Hamacher.
VAST 2011.

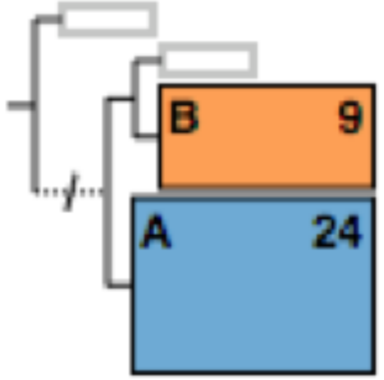
Comparing many phylogenetic trees

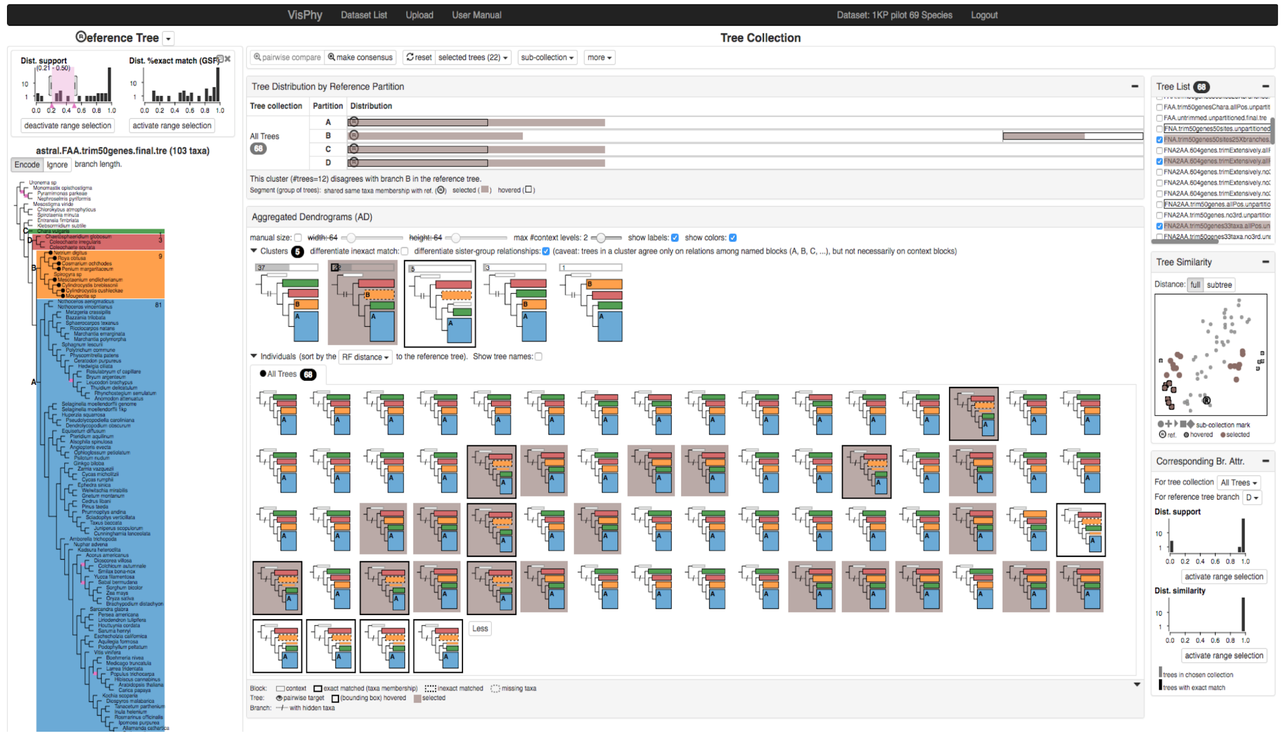


Contributions at abstraction, idiom, & algorithm levels

- data and task **abstractions** for comparison of phylogenetic trees
- new visual encoding **idiom**: Aggregated Dendrogram
  - compact tree representation that focuses on selected subtrees
  - **algorithm** that adapts to available screen space
- interactive multi-view tool: ADView
  - covers multiple levels of details for tree comparison

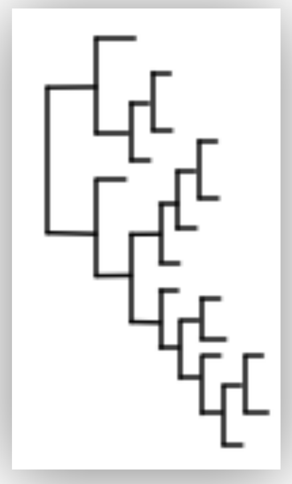


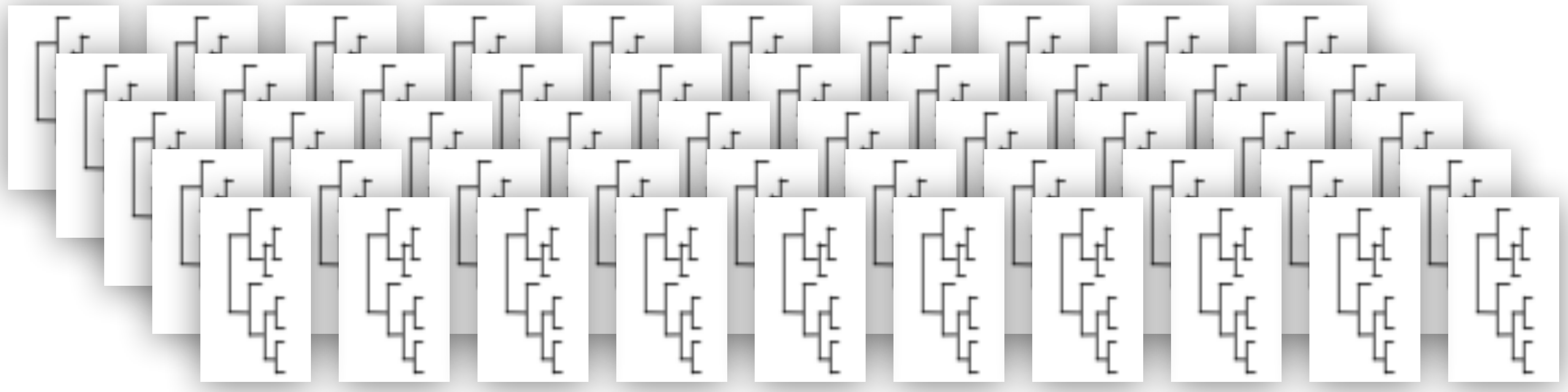




Data abstraction: Trees

Reference tree vs. Tree collection

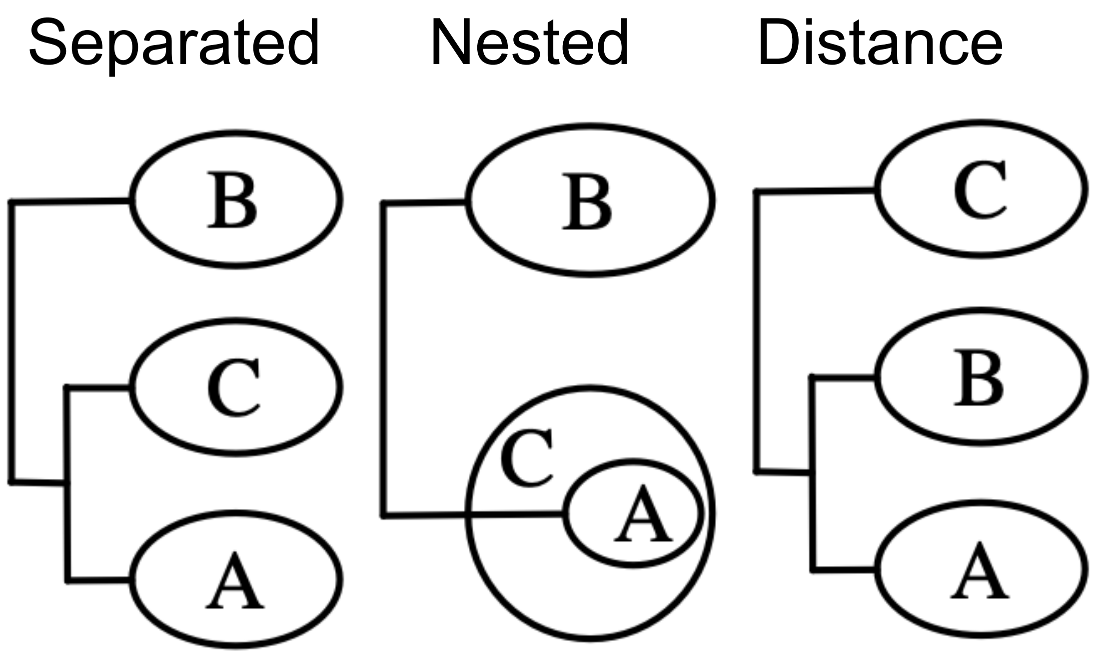




Task abstraction

Task abstraction

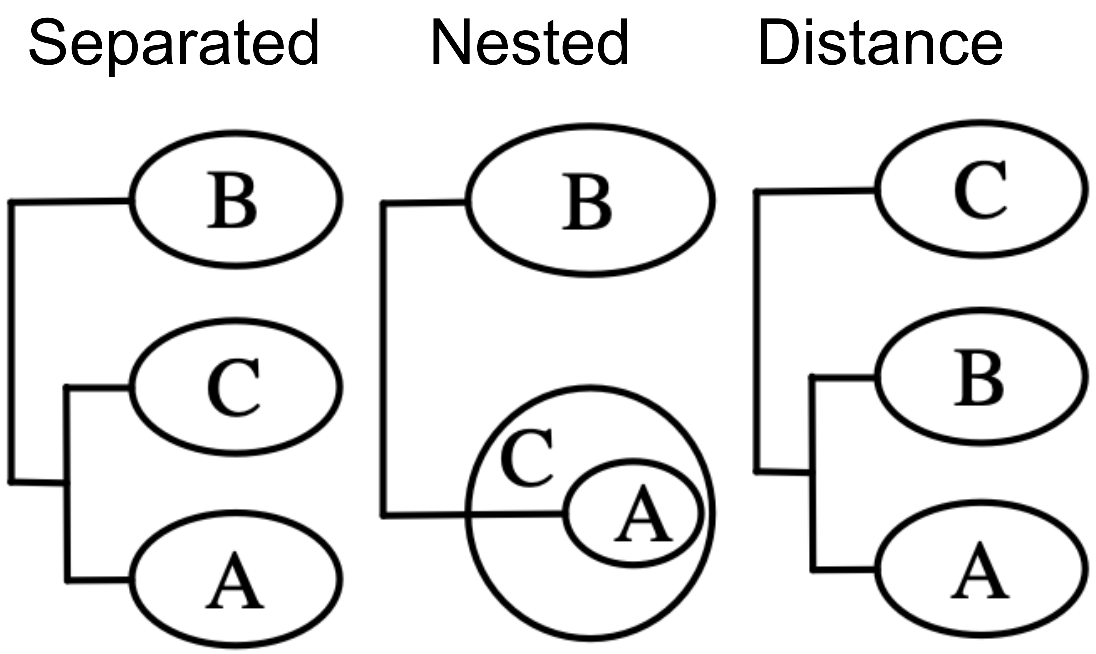
Topological relationships & distance
between subtrees / leaf nodes

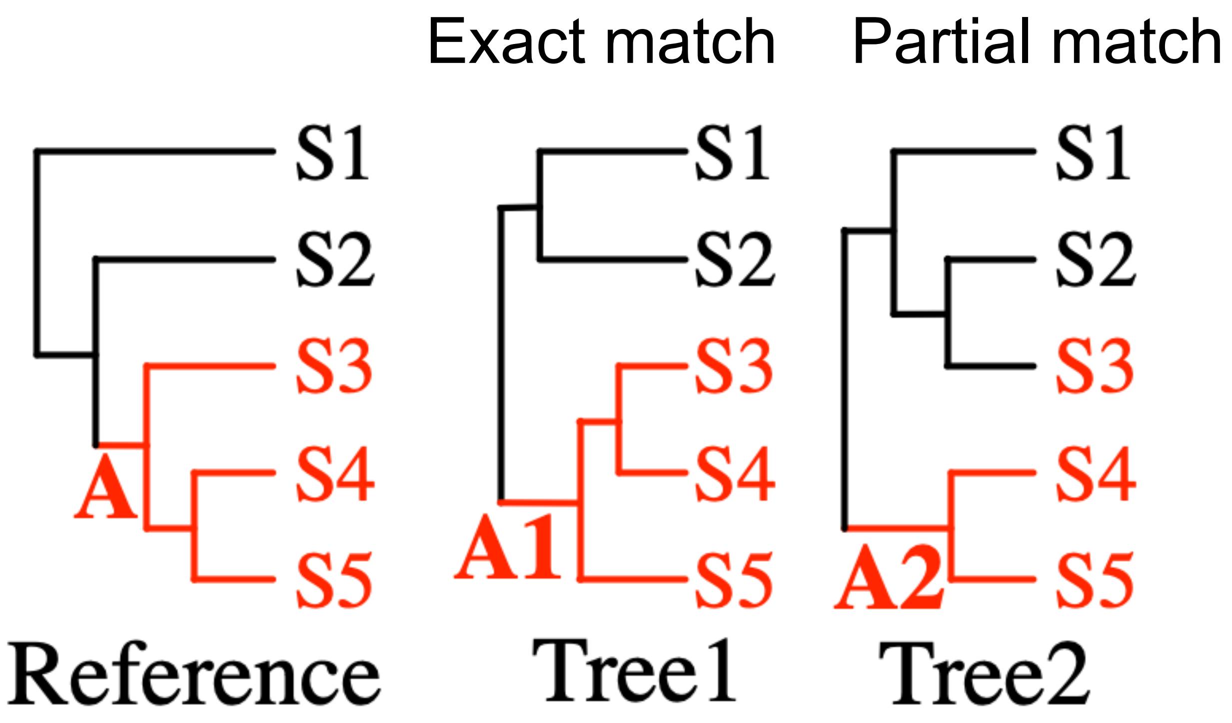


Task abstraction

Topological relationships & distance between subtrees / leaf nodes

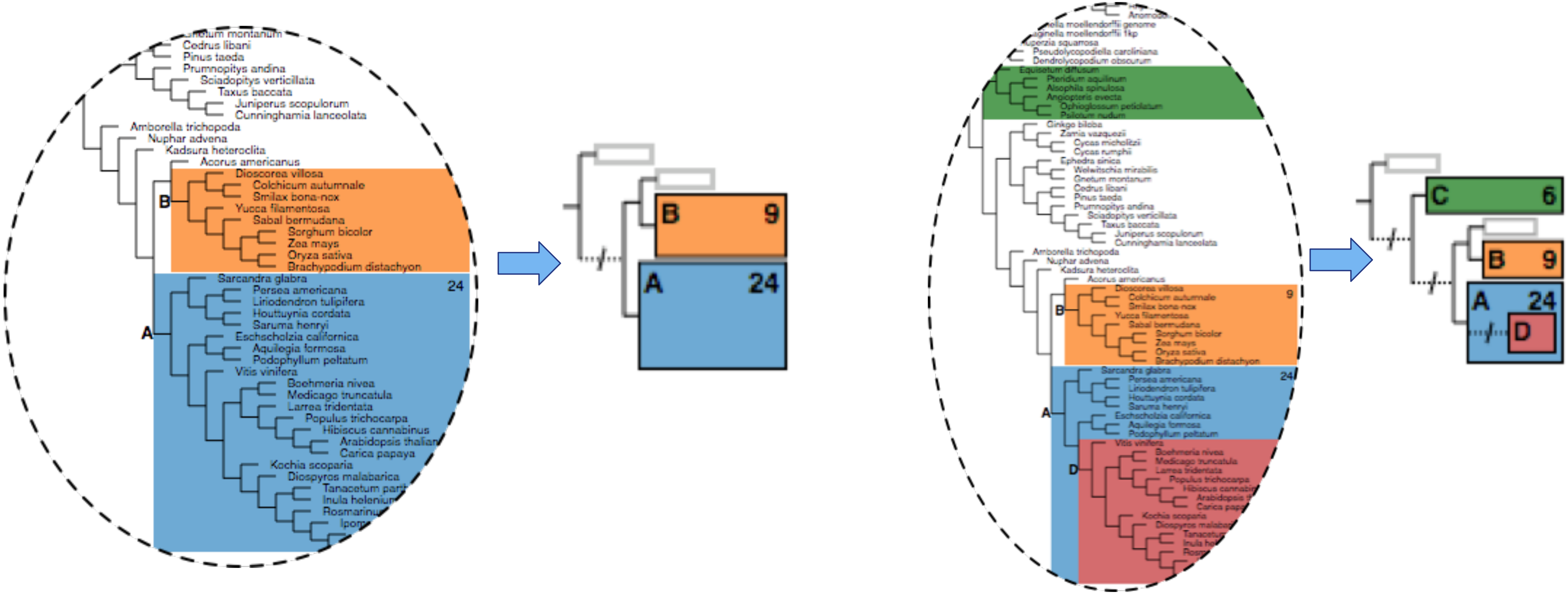
Leaf node memberships compared to reference tree





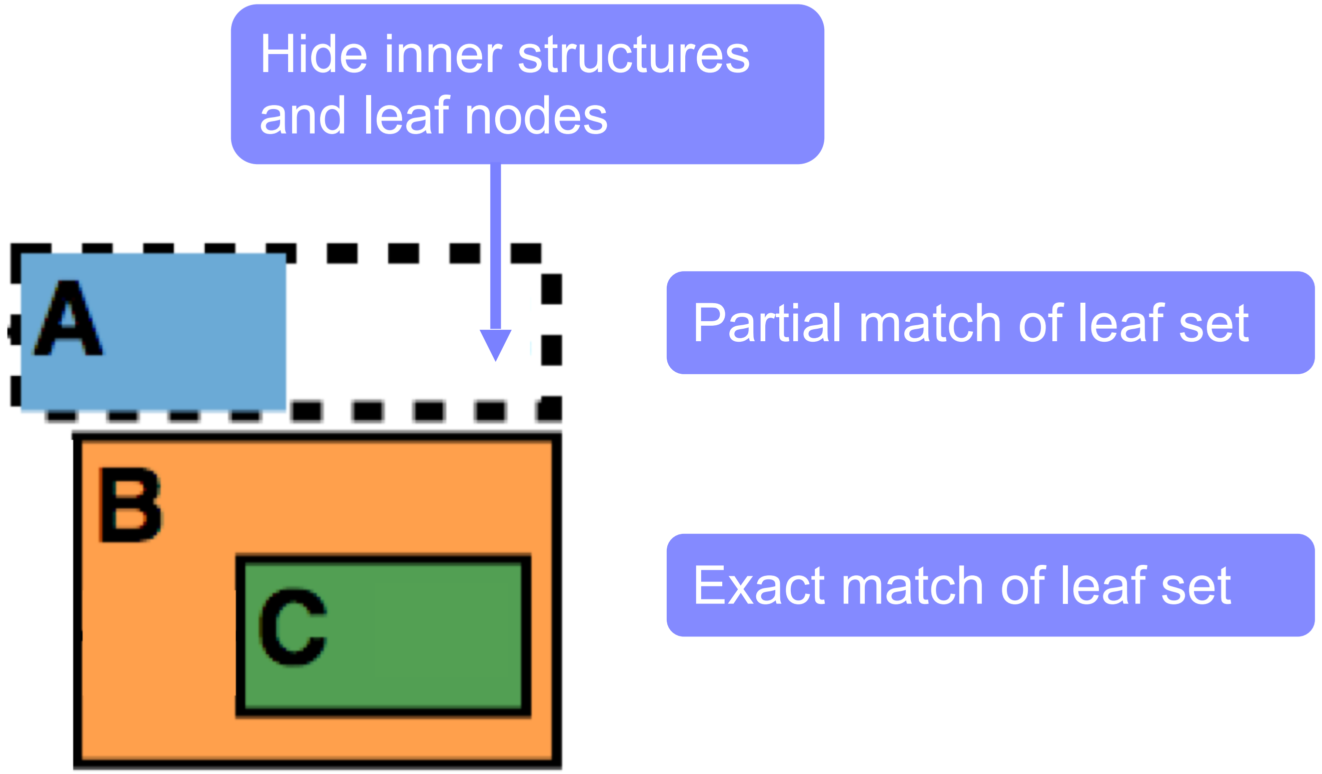
Aggregated Dendrogram: Intuition

Use glyphs to compress a tree according to user selections



Visual design: focus + context

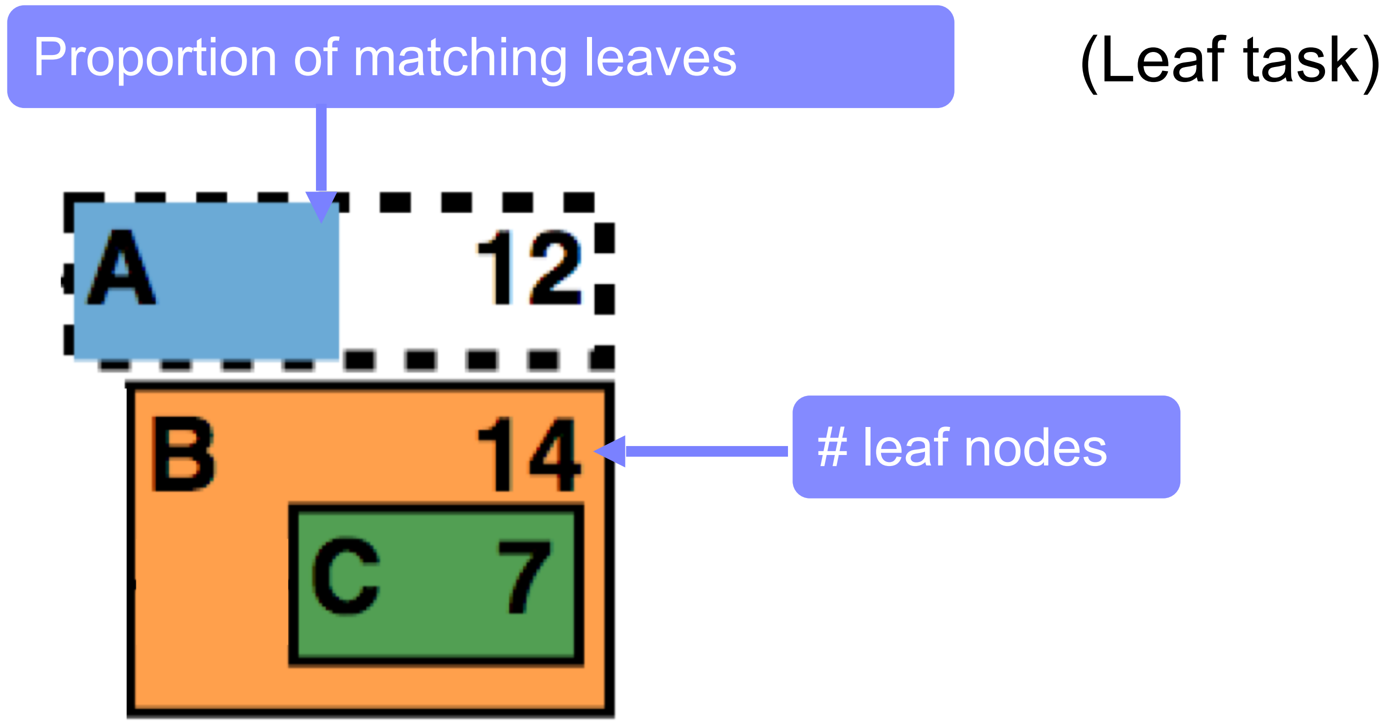
- focus
  - selected subtrees



(Leaf task)

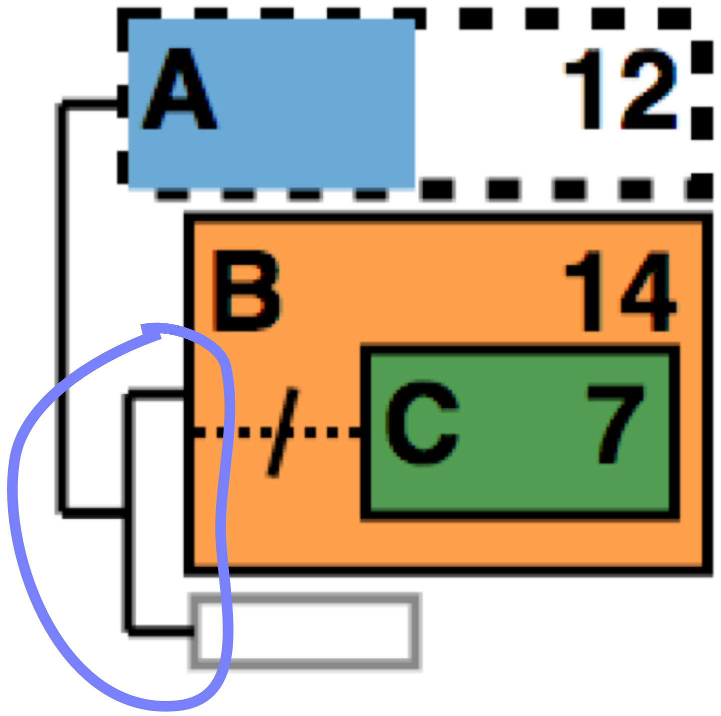
Visual design: focus + context

- focus
  - selected subtrees



Visual design: focus + context

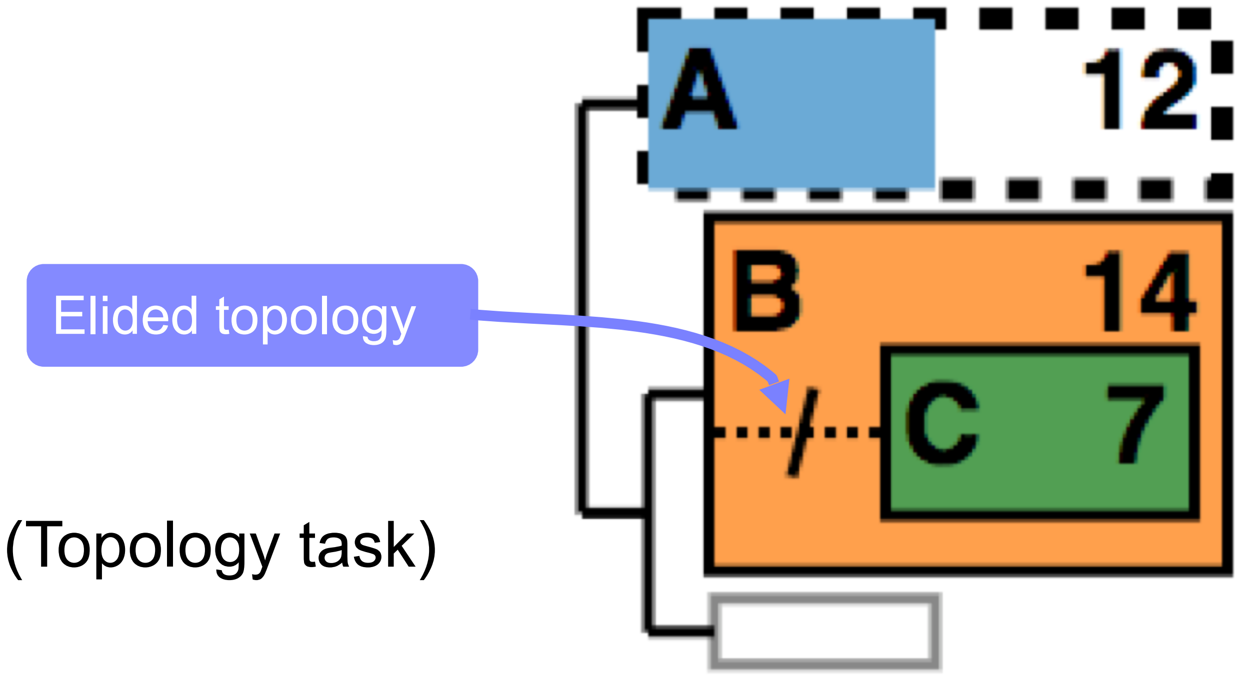
- focus
  - selected subtrees
  - topological relationships between them



(Topology task)

Visual design: focus + context

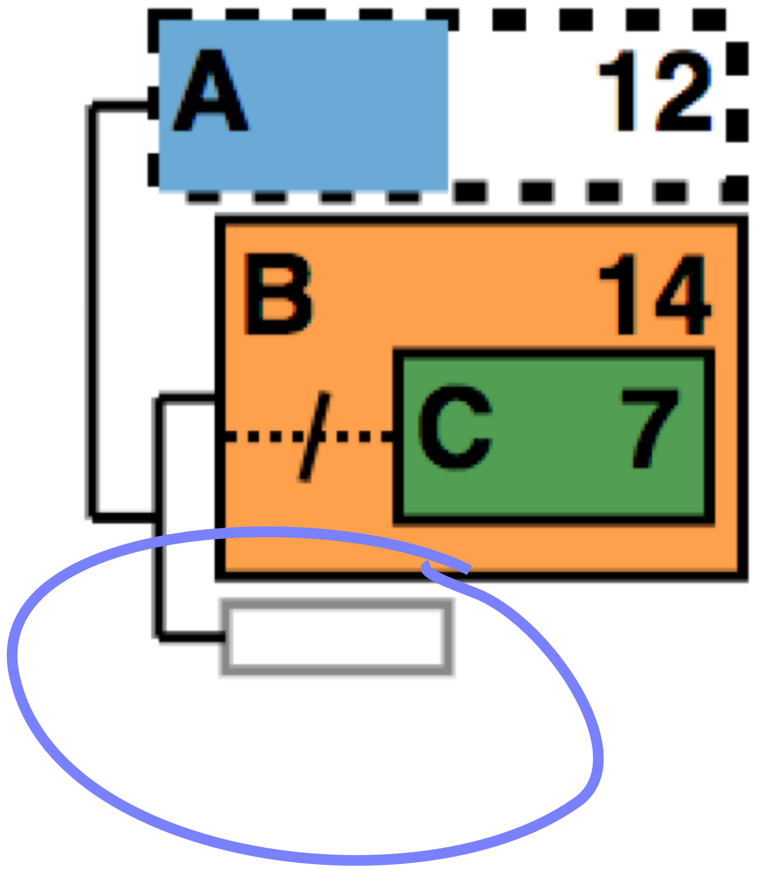
- focus
  - selected subtrees
  - topological relationships between them



(Topology task)

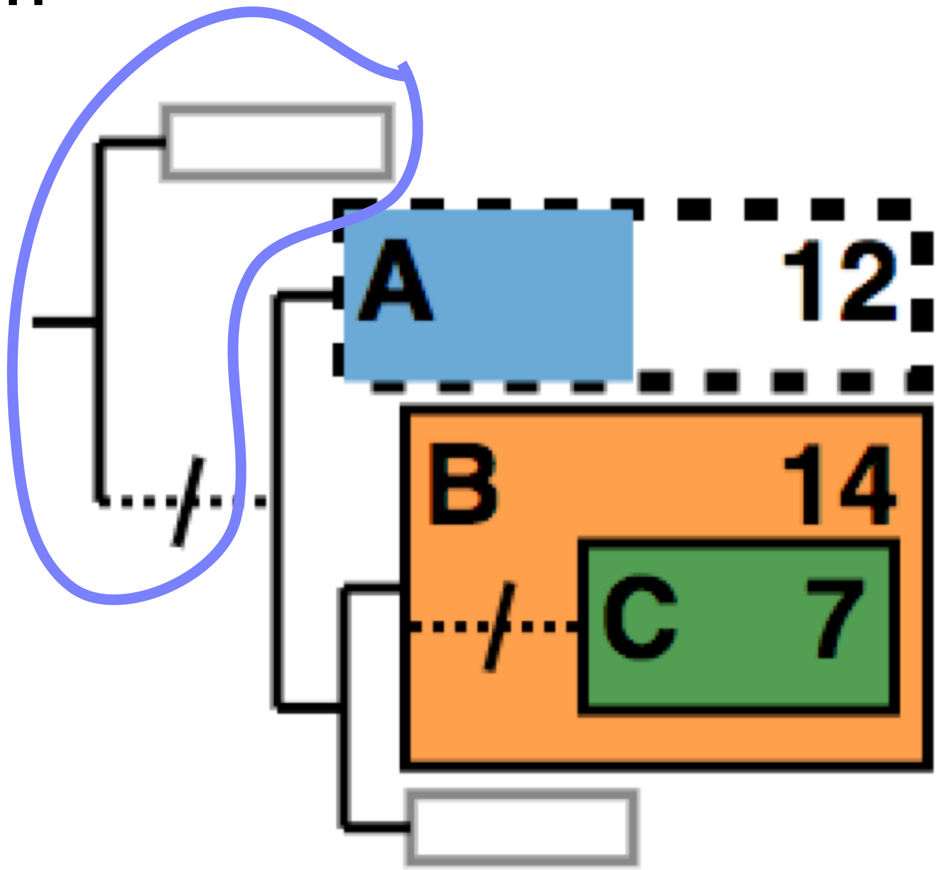
Visual design: focus + context

- focus
  - selected subtrees
  - topological relationships between them
- context
  - neighboring subtrees



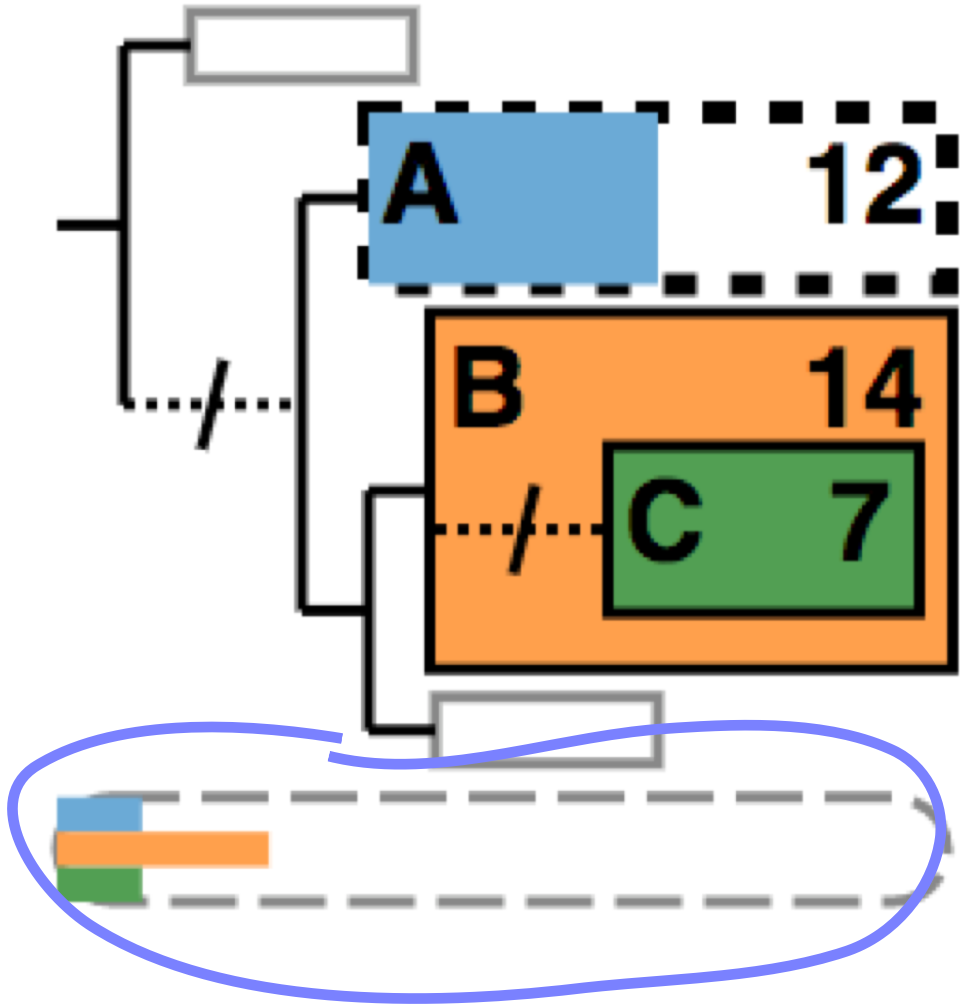
Visual design: focus + context

- focus
  - selected subtrees
  - topological relationships between them
- context
  - neighboring subtrees
  - upstream topology and root



Visual design: focus + context

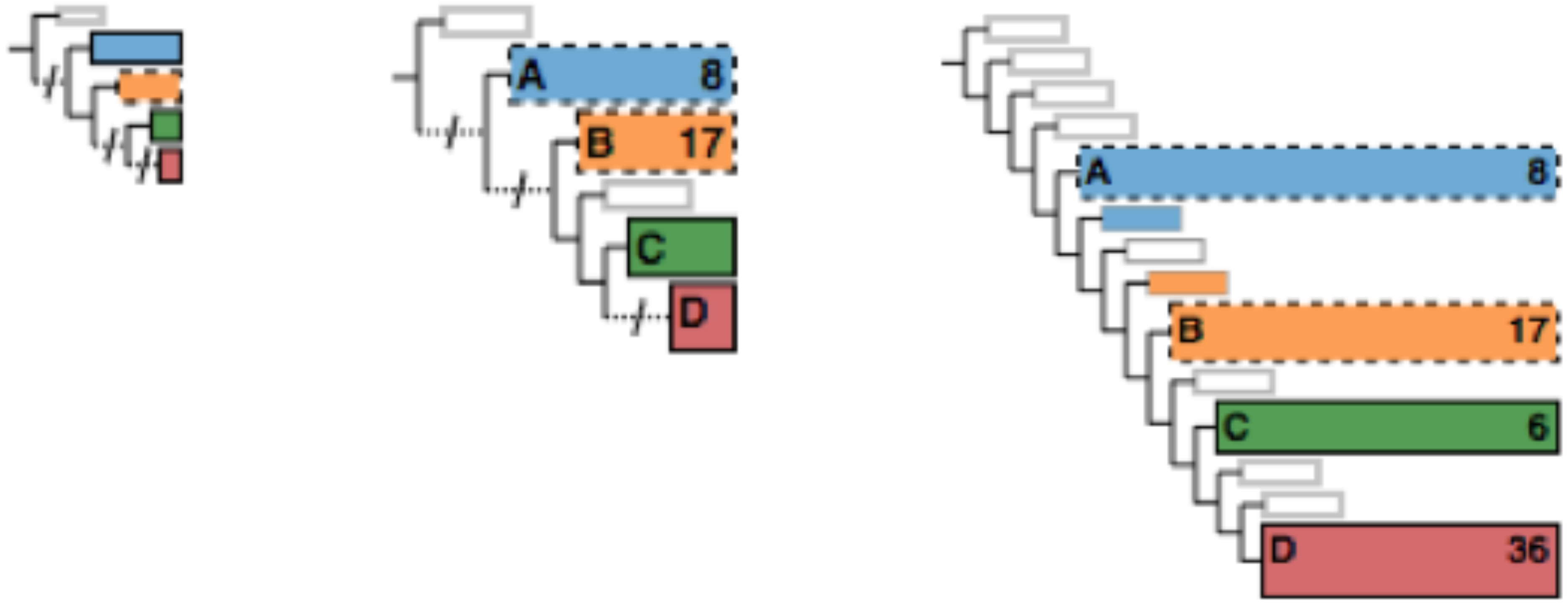
- focus
  - selected subtrees
  - topological relationships between them
- context
  - neighboring subtrees
  - upstream topology and root
  - missing leaf nodes



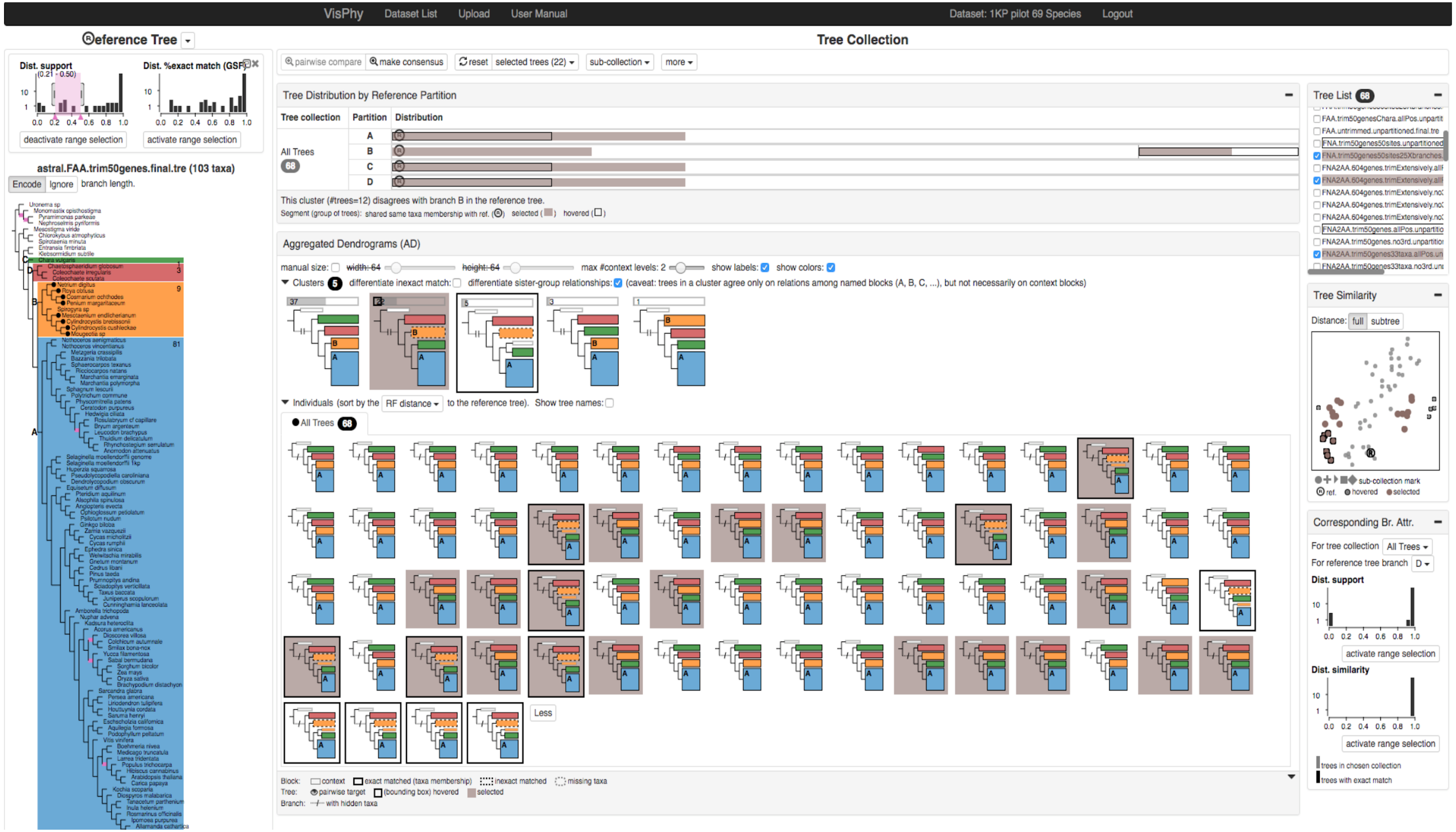
Visual design: algorithm adapts to space

- show more info when space permitted
  - labels
  - # leaf nodes
  - neighboring blocks

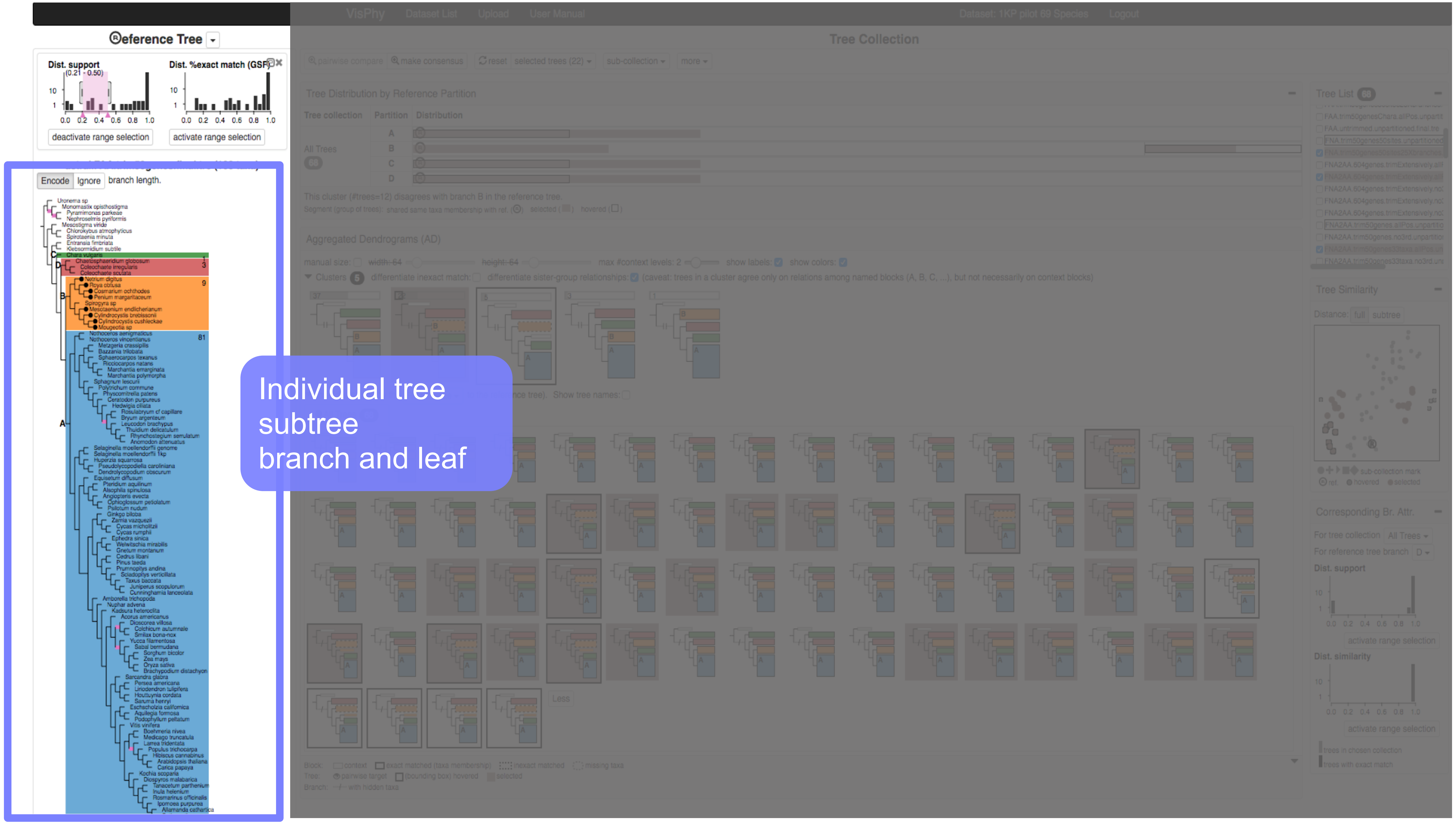
40x40 px **80x80 px** **160x160 px**



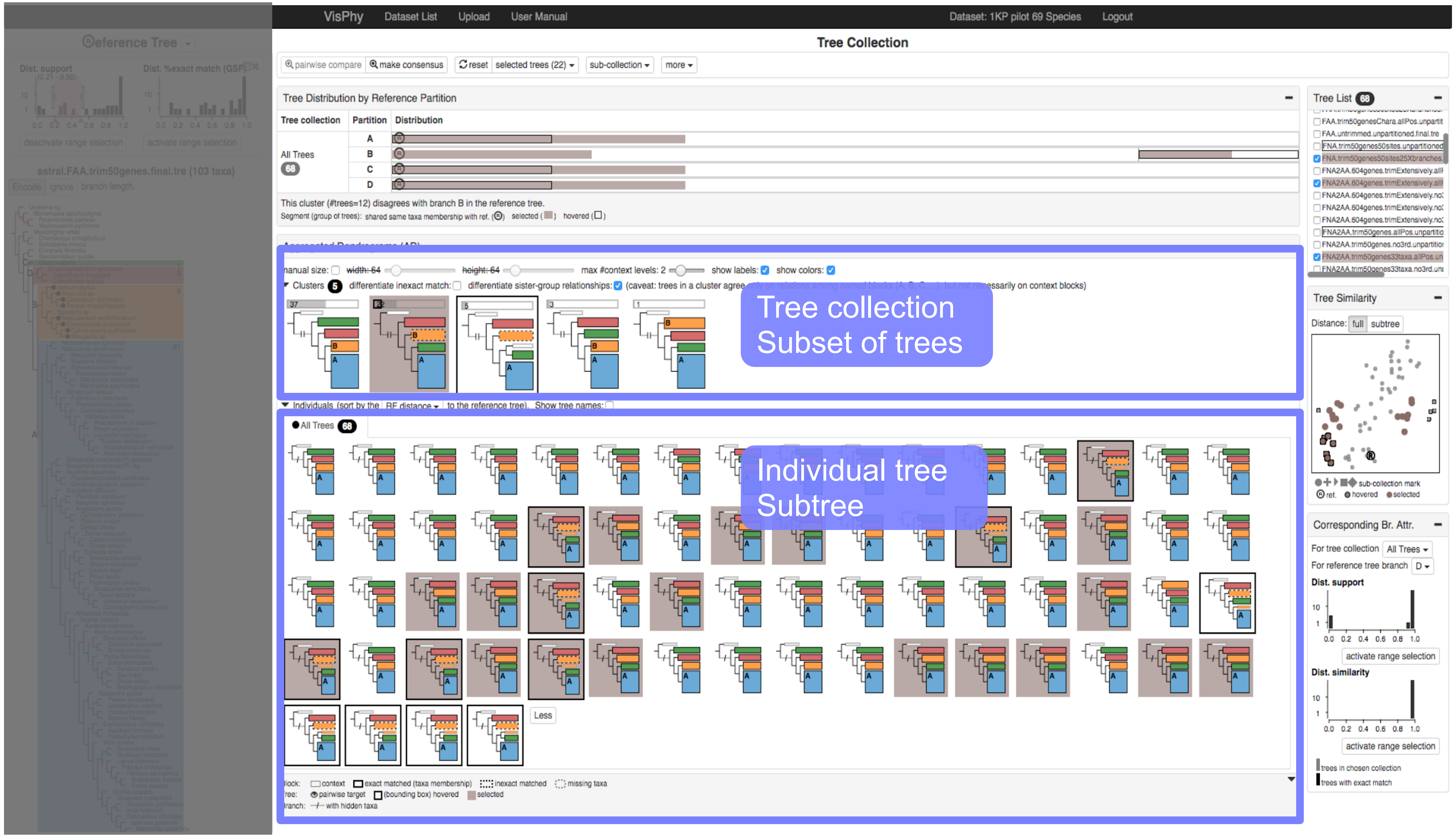
ADView interface: Multi-level structure across views



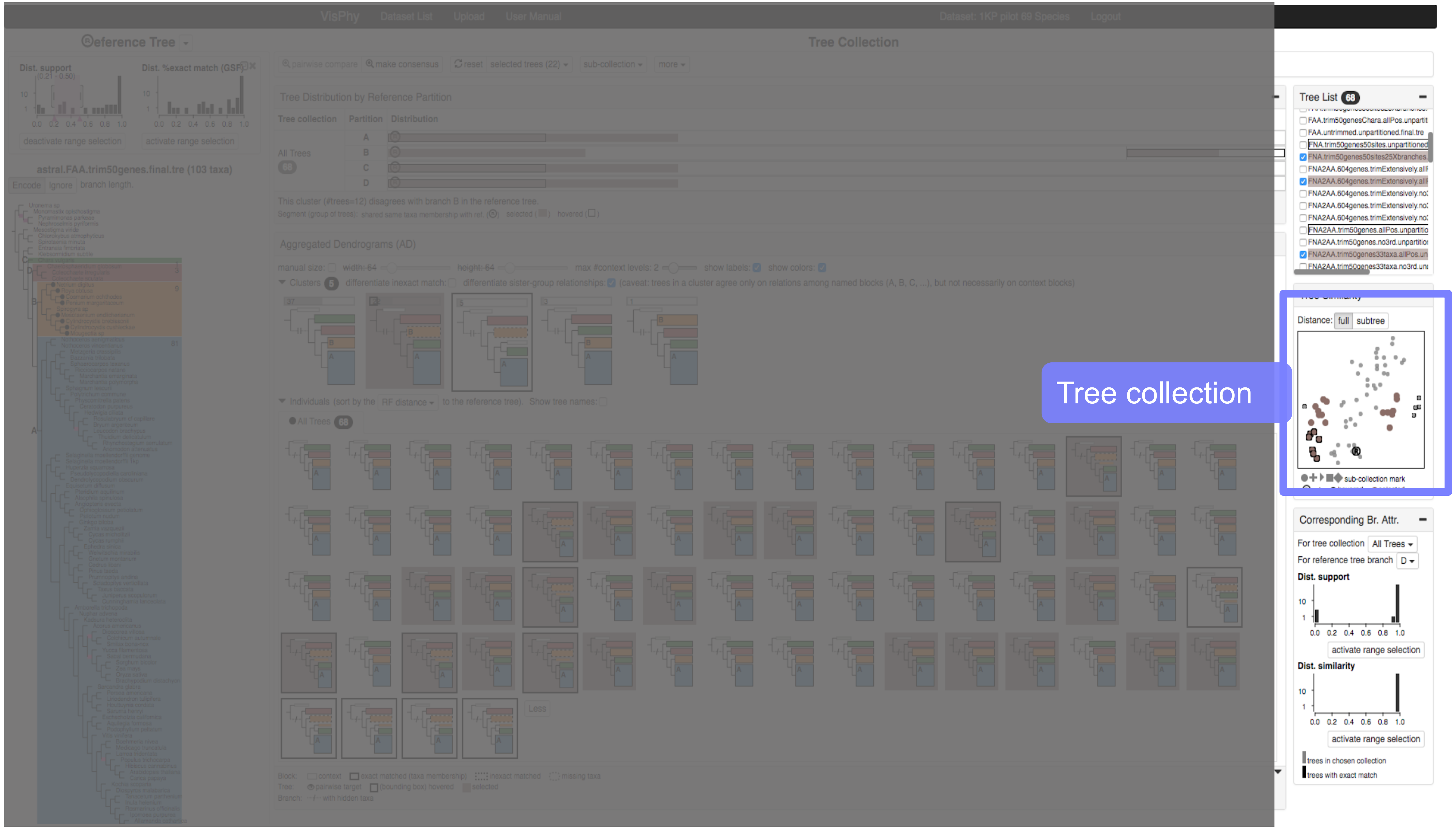
Interface walkthrough: reference tree



Interface walkthrough: individual & cluster ADs



Interface walkthrough: treespace



Validation with many biologists

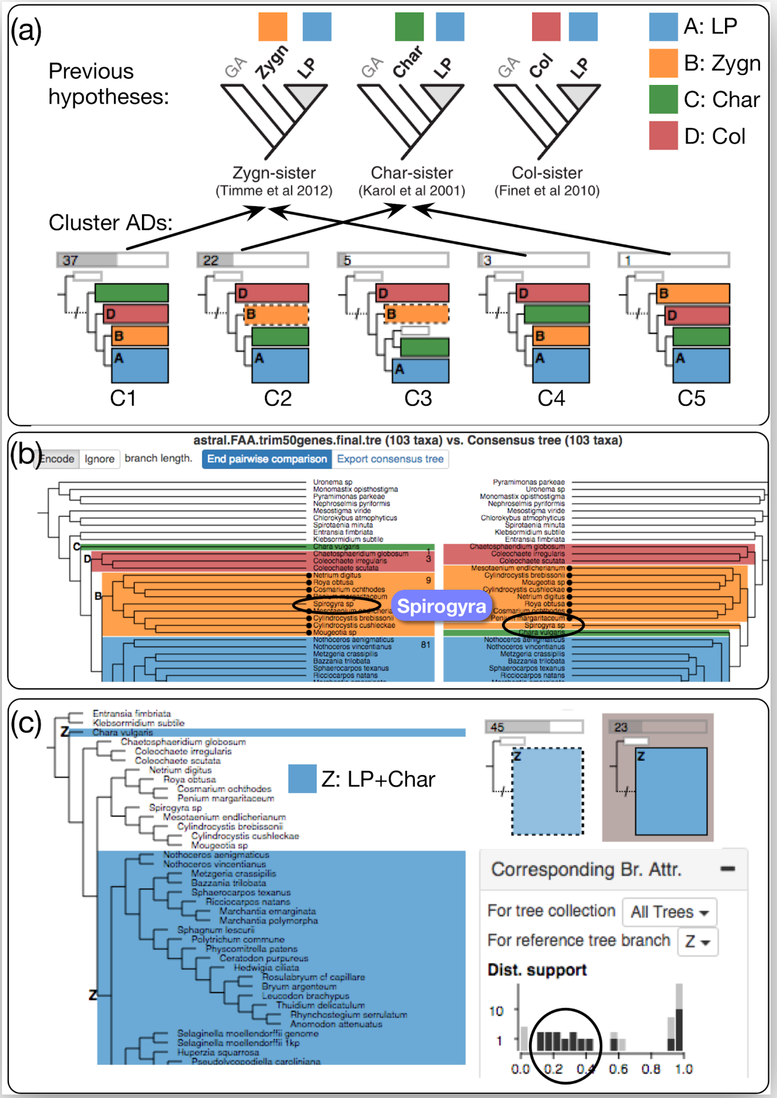
- worked closely with a biology PhD student (second author)

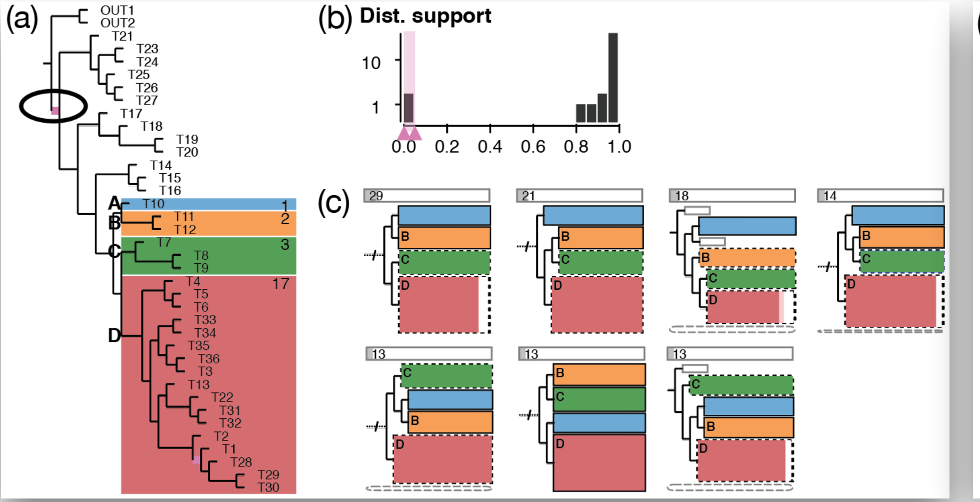
Validation with many biologists

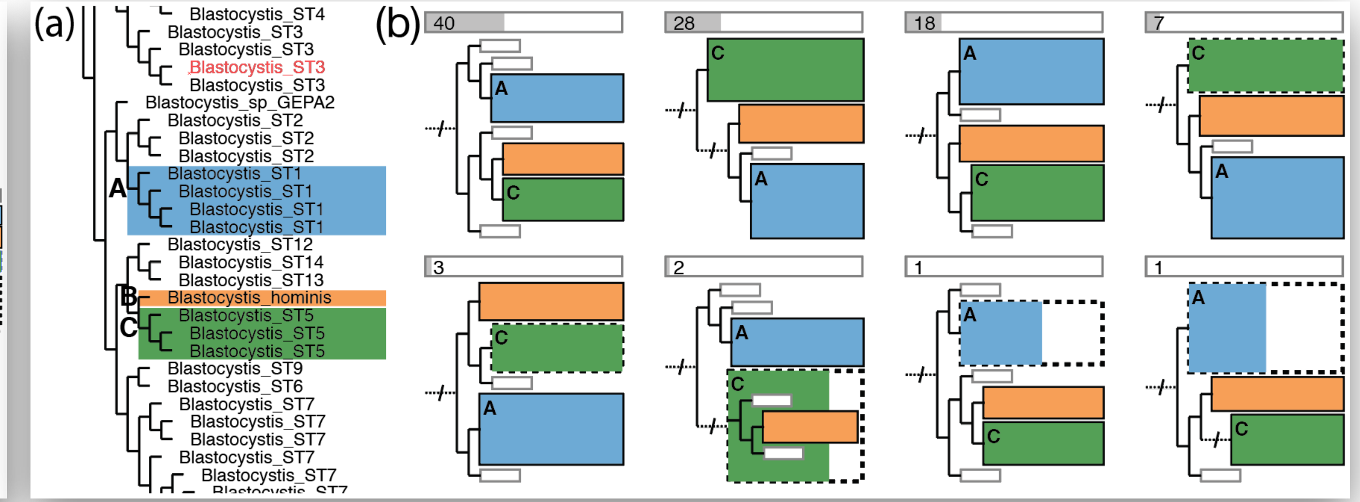
- worked closely with a biology PhD student (second author)
- demos, interviews and discussions
  - 10 biologists at different times throughout project

Validation with many biologists

- worked closely with a biology PhD student (second author)
- demos, interviews and discussions
  - 10 biologists at different times throughout project
- user study sessions
  - 5 biologists, using their own datasets

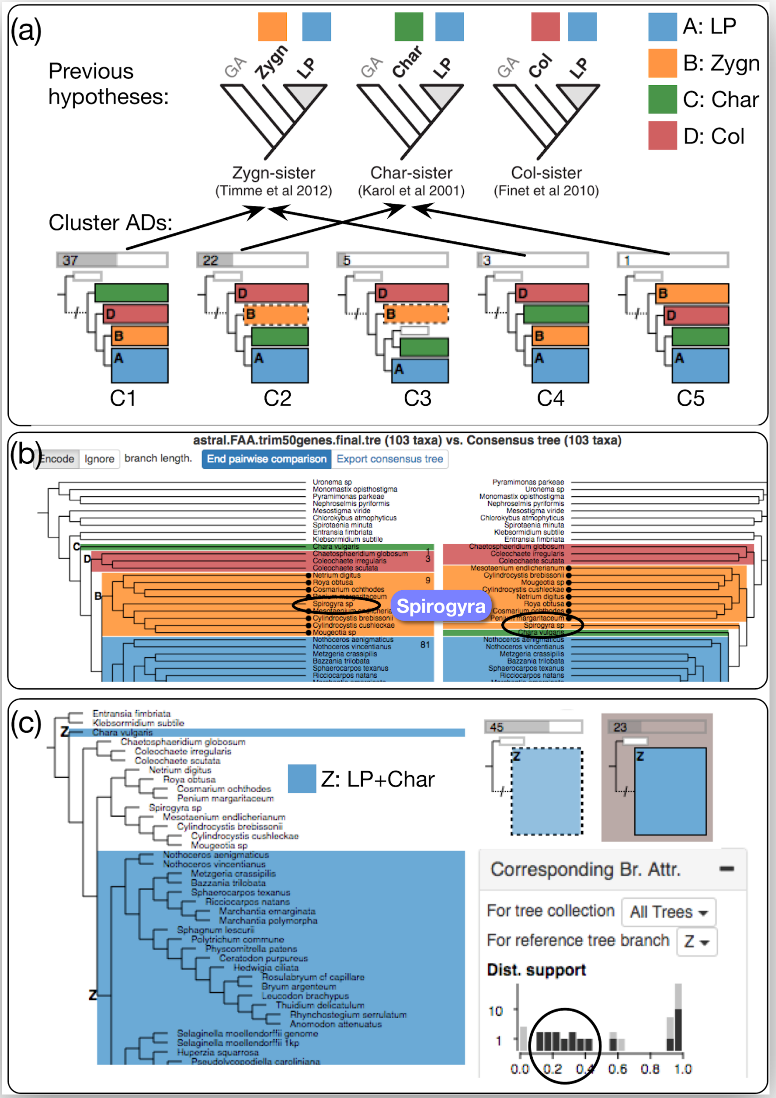


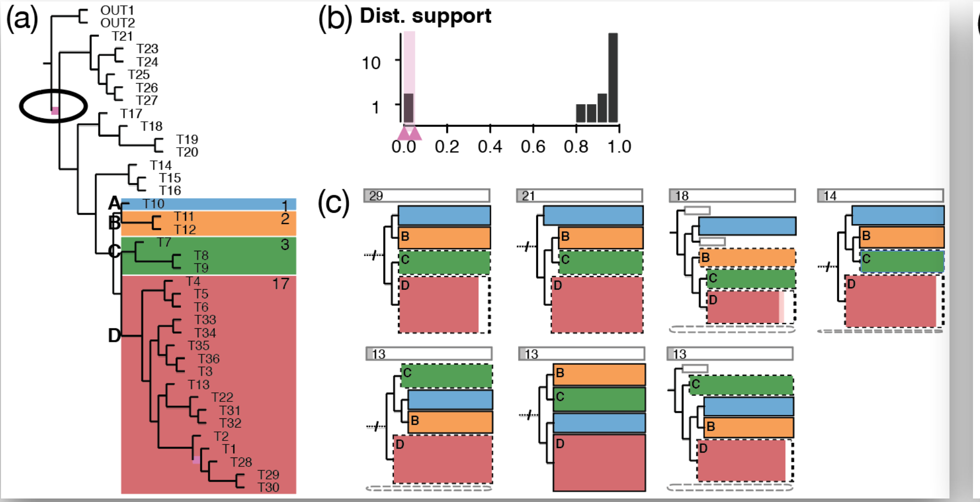


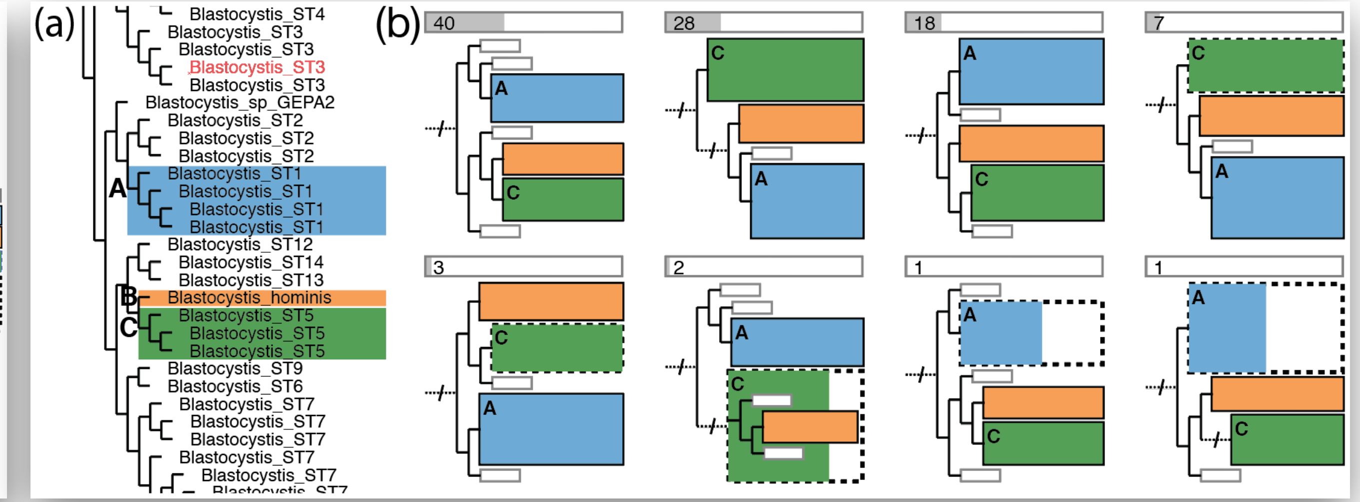


Validation with many biologists

- worked closely with a biology PhD student (second author)
- demos, interviews and discussions
  - 10 biologists at different times throughout project
- user study sessions
  - 5 biologists, using their own datasets
- biologists confirmed
  - validity of data and task abstractions
  - utility of ADView





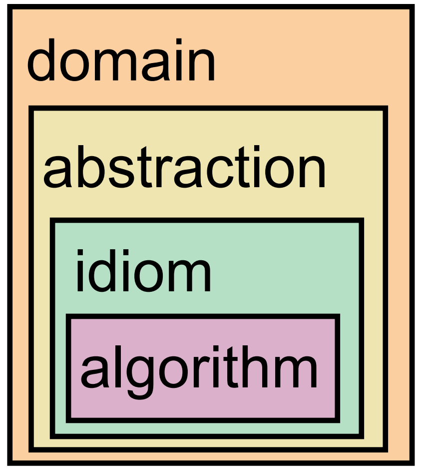


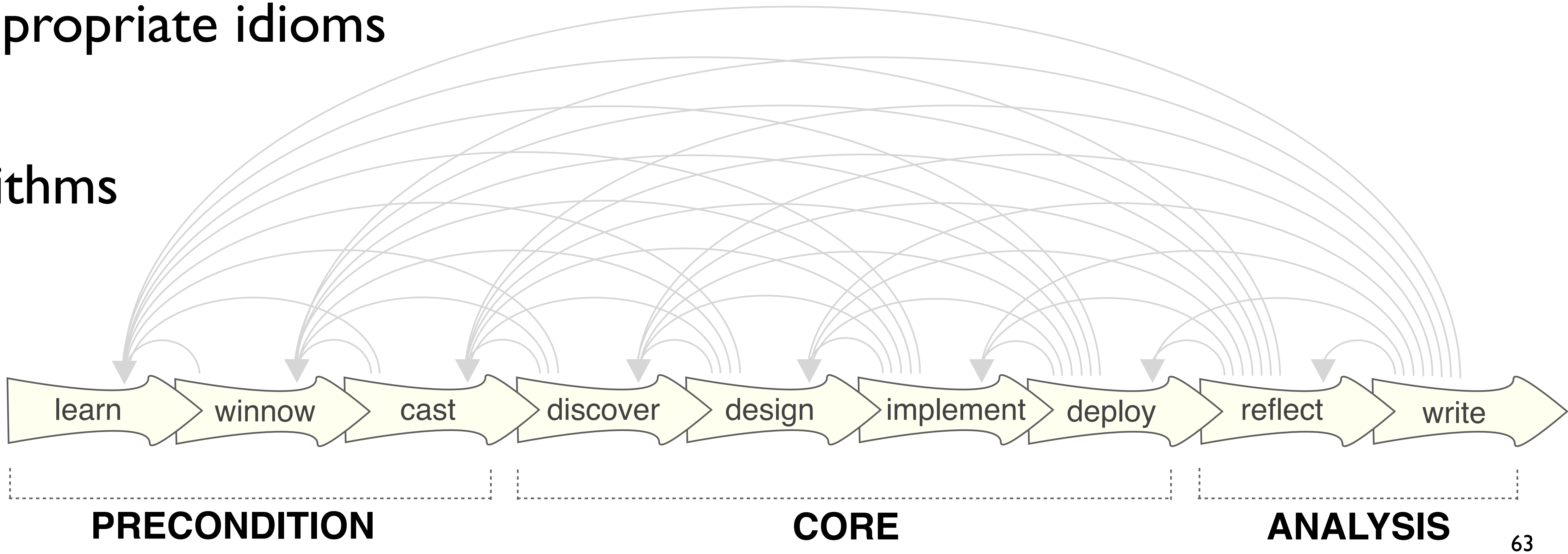
Video

https://www.youtube.com/watch?v=2SLcz7KNLJw

Problem-driven visualization with design study methodology

- work through all four levels of nested model
  - investigate domain
  - identify abstractions
    - crucial -- & difficult -- iterative process
  - select or create appropriate idioms
  - develop new algorithms
    - if need be



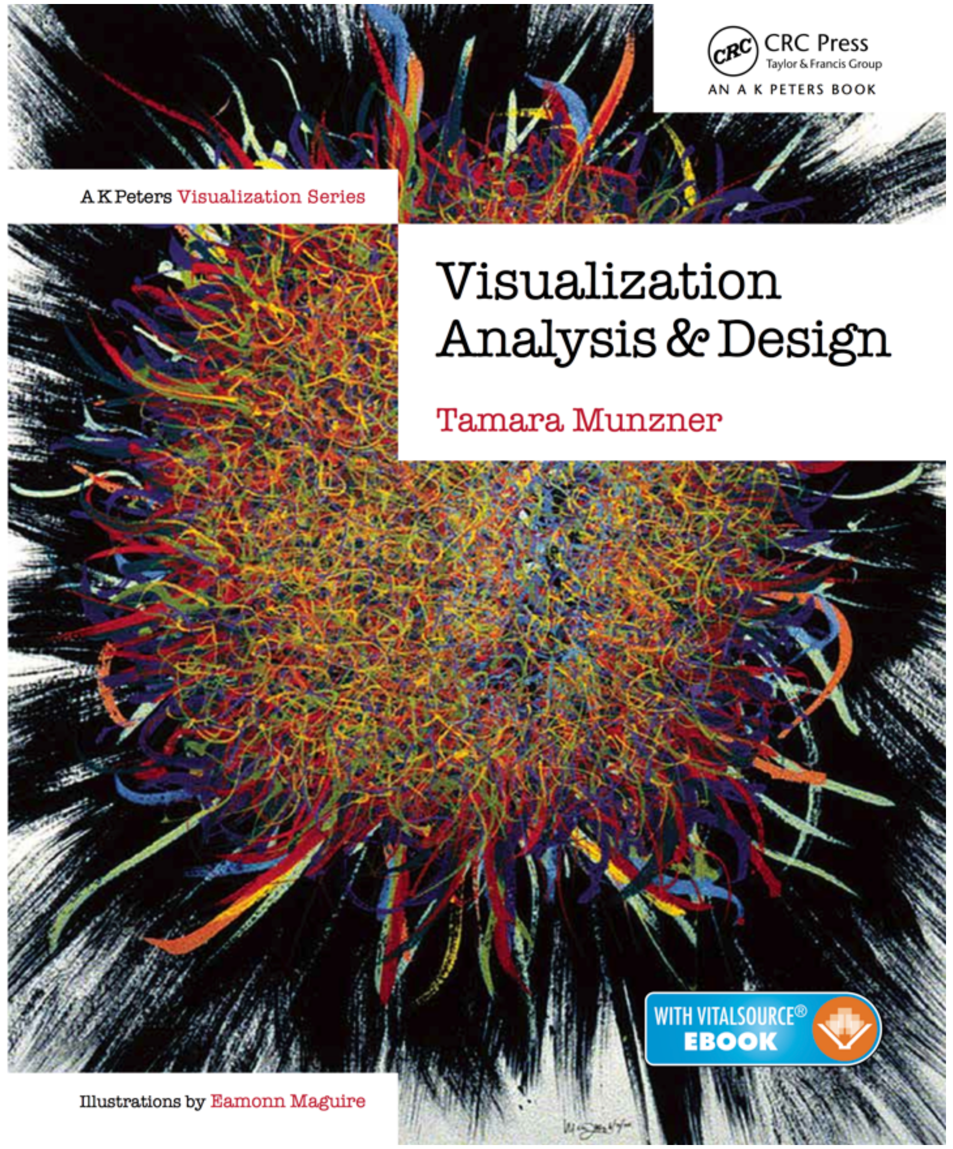


More information

- this talk
  http://www.cs.ubc.ca/~tmm/talks.html#amw24

- book
  http://www.cs.ubc.ca/~tmm/vadbook
  (hardcopy on demo/stuff table)

- full courses, papers, videos, software, talks
  http://www.cs.ubc.ca/group/infovis
  http://www.cs.ubc.ca/~tmm



Visualization Analysis and Design. Munzner. *CRC Press, AK Peters Visualization Series, 2014.*

@tamara@vis.social

@tamaramunzner