Plant Cell Lineage Specification based on the Interactive Visualization of Hybrid 3D and 2D Data and with the Support of Machine Learning

Tobias Isenberg

Inria INVENTEURS DU MONDE NUMÉRIQUE

université PARIS-SACLAY

LISN LABORATOIRE INTERDISCIPLINAIRE DES SCIENCES DU NUMÉRIQUE

NAVISCOPE

INRAE science for people, life & earth

Collaboration

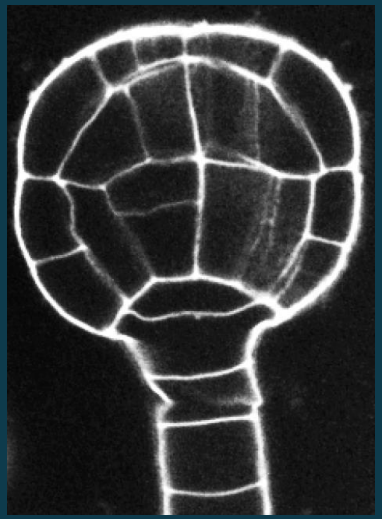
Jiayi Hong

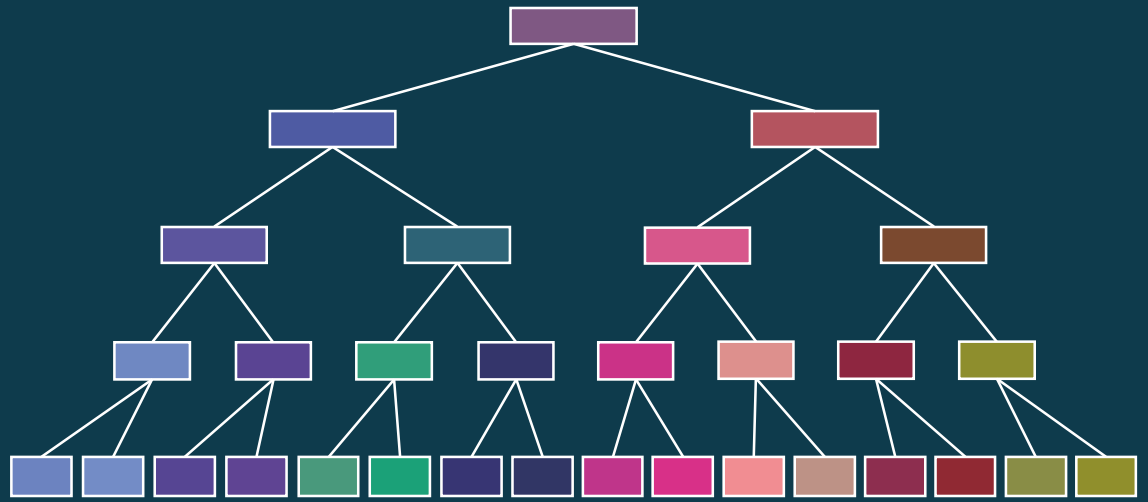
Inria
INVENTEURS DU MONDE NUMÉRIQUE

Alain Trubuil

INRAE
science for people, life & earth

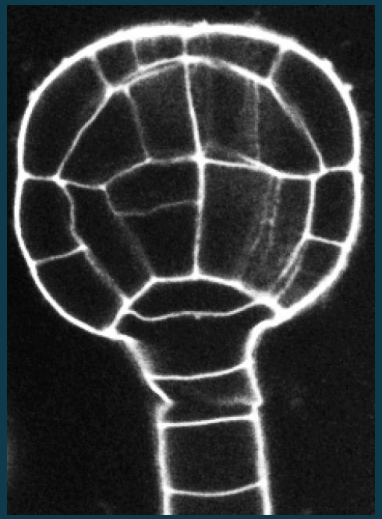
Plant Cell Division

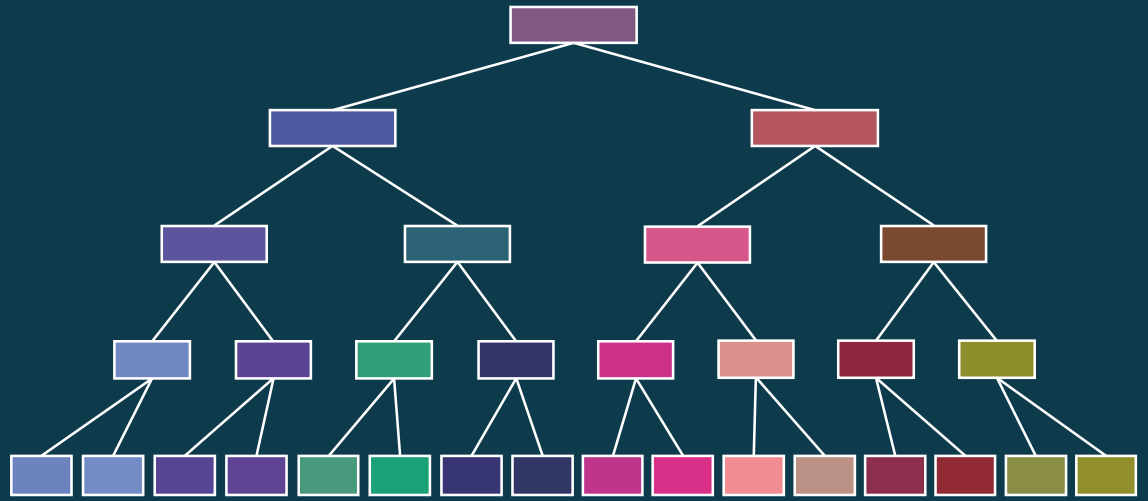




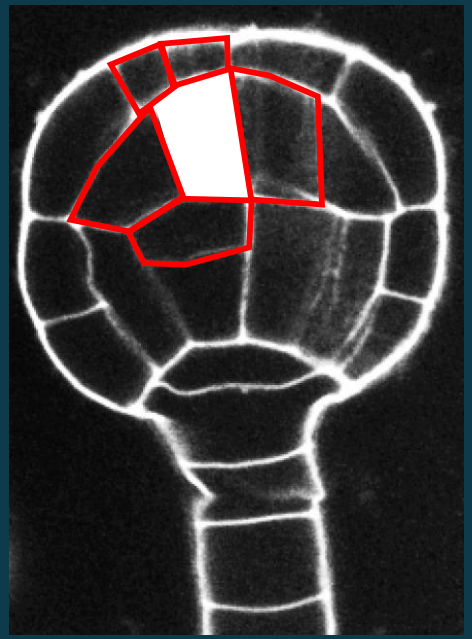
Plant Cell Lineage

For a specific embryo, what do biologists do to build the hierarchy?



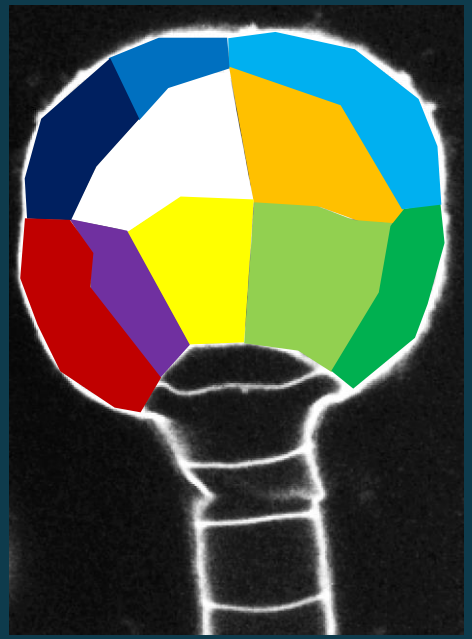


How to do cell lineage?



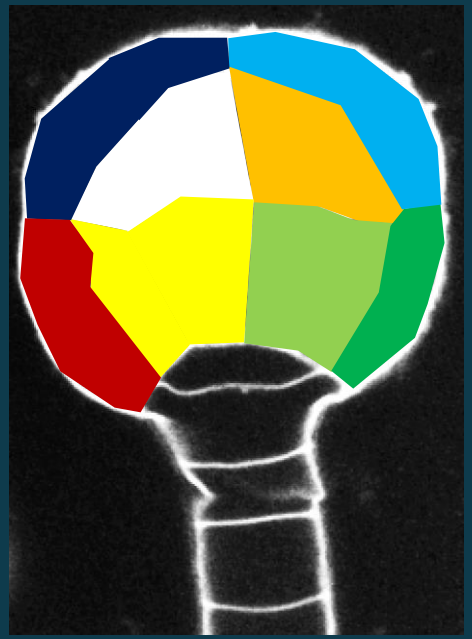
Biologists need to find the right sister cell for every cell in an embryo.

How to do cell lineage?



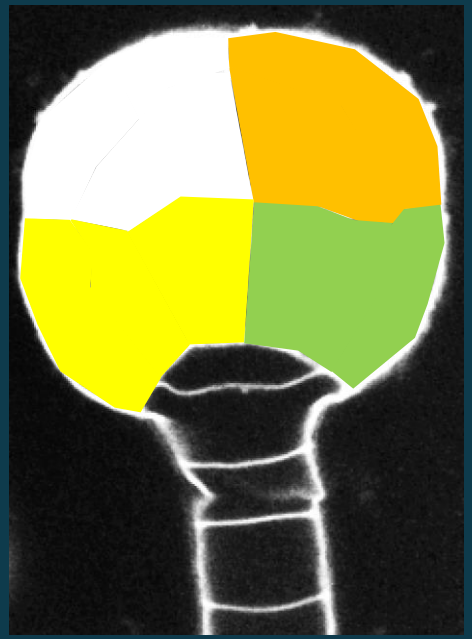
Once decided, they would merge cells and continue assigning the remaining cells.

How to do cell lineage?



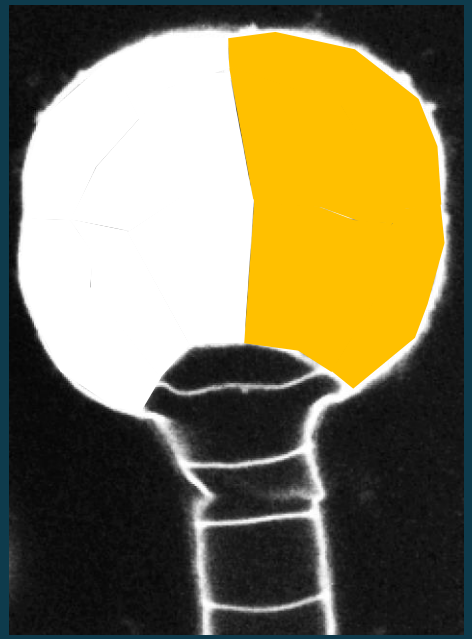
Biologists will continue this process to the new generation until there is only one cell left.

How to do cell lineage?



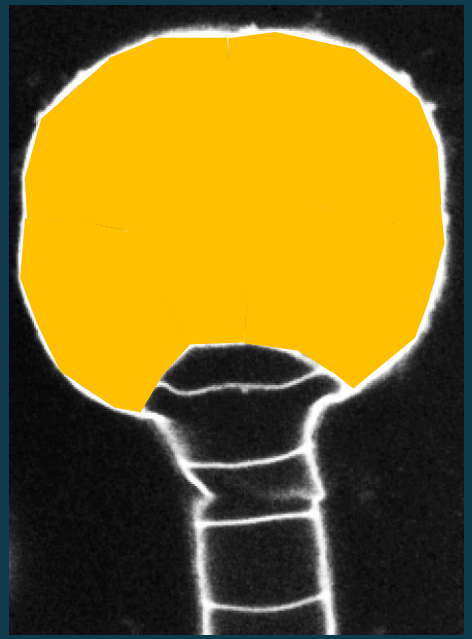
Biologists will continue this process to the new generation until there is only one cell left.

How to do cell lineage?



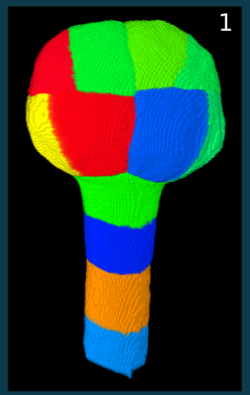
Biologists will continue this process to the new generation until there is only one cell left.

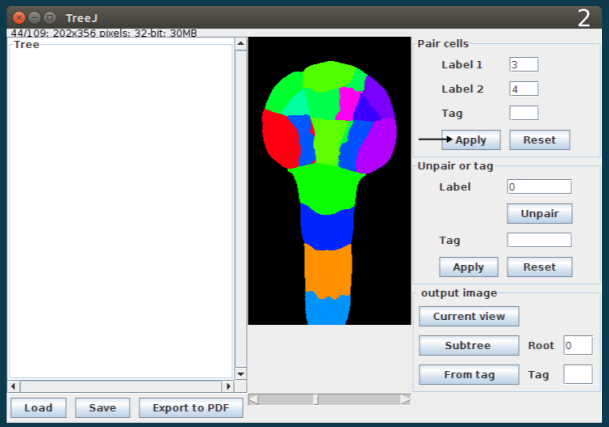
How to do cell lineage?

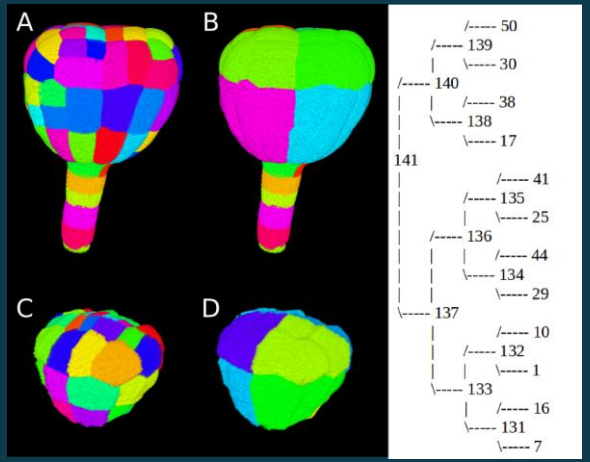


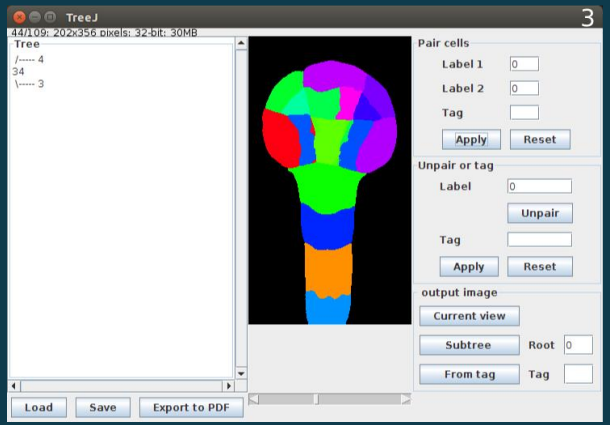
Biologists will continue this process to the new generation until there is only one cell left.

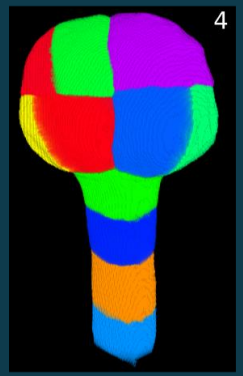
Traditional Tools











Traditional Workflow

Get segmented 2D slices

Check every cell

Manually find the sister for each cell

Write down the hierarchy on paper

Traditional Workflow

Get segmented 2D slices

Visualization

Interaction

Check every cell

Manually find the sister for each cell

Visualization

Interaction

ML

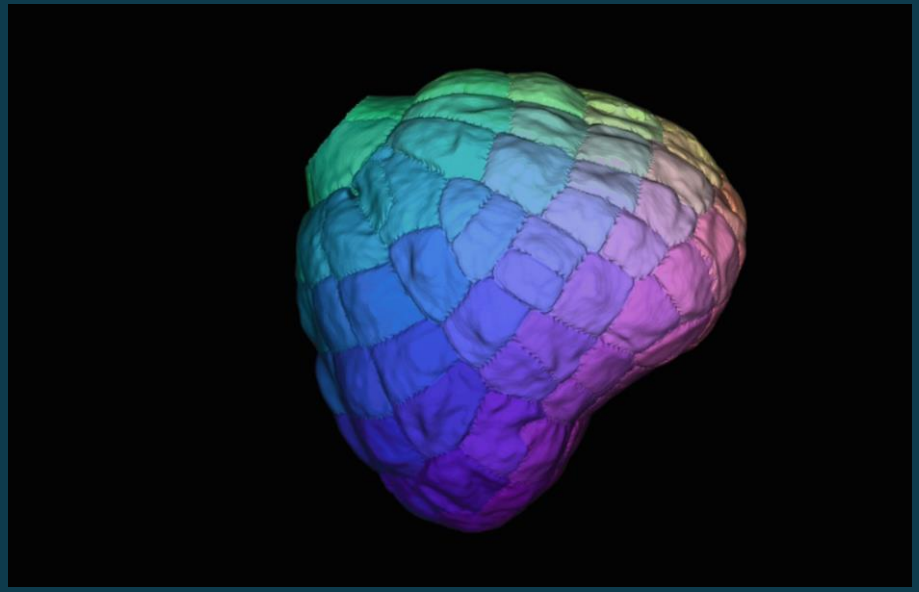
Visualization

Write down the hierarchy on paper

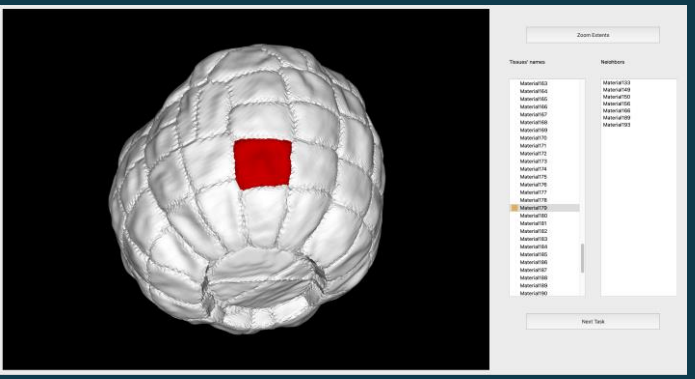
Problem with Plant Cell Embryos

Cells are densely packed together.

Interaction

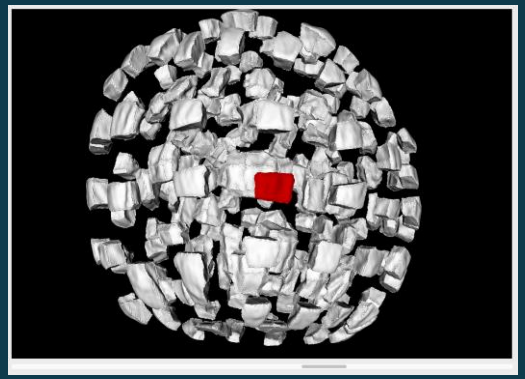


Exploration



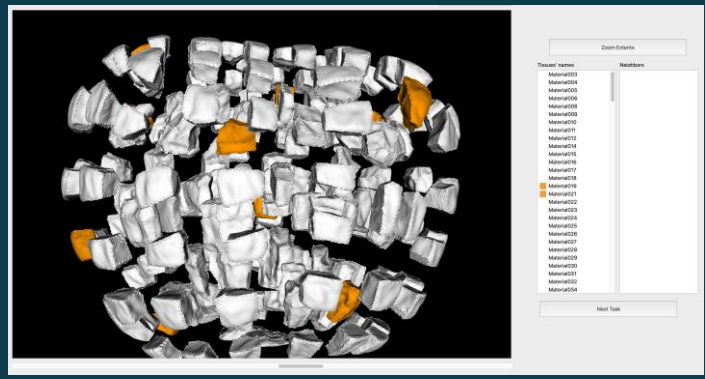
1D

Select from the list



3D

Select from the explosion view



1D + 3D

Select from both views

Study Results

- The explosion technique makes it possible to access objects in such a tightly packed 3D environment.
- Combination (List + Explosion) seems to combine these advantages of the single technique.

Traditional Workflow

Get segmented 2D slices

Visualization

Interaction

Check every cell

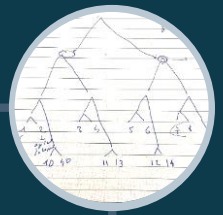
Manually find the sister for each cell

Visualization

Interaction

ML

Visualization



Write down the hierarchy on paper

Datasets

≡

Color Modes

By District | By Normalized Shared Area | By Model Confidence | Randomized

Explosion Extent

Peeling

Interactions

Confirm | Wrong Children | New Sister | Supporting Cell | Predict One Level

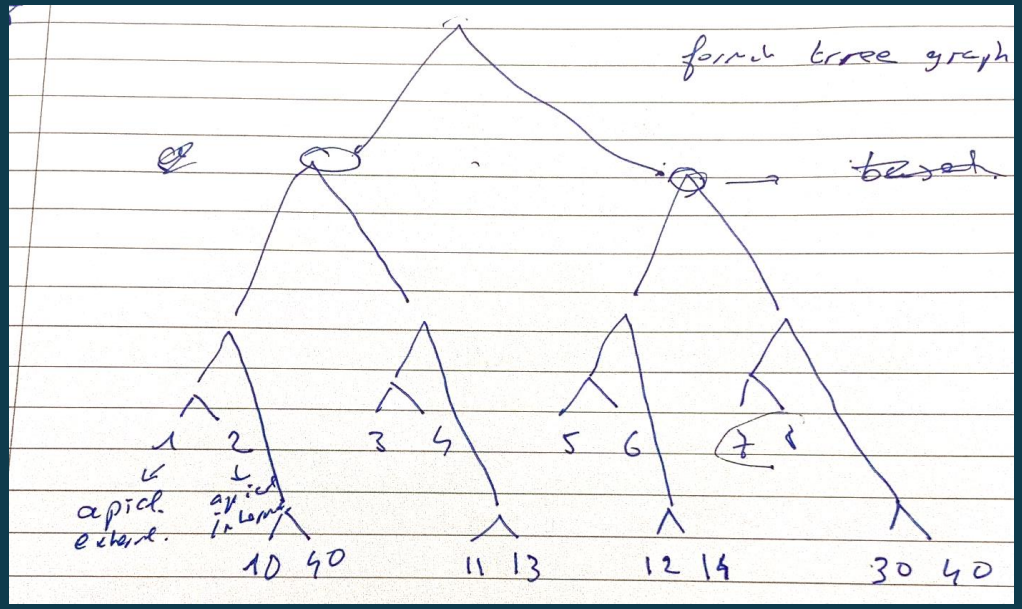
Selections:

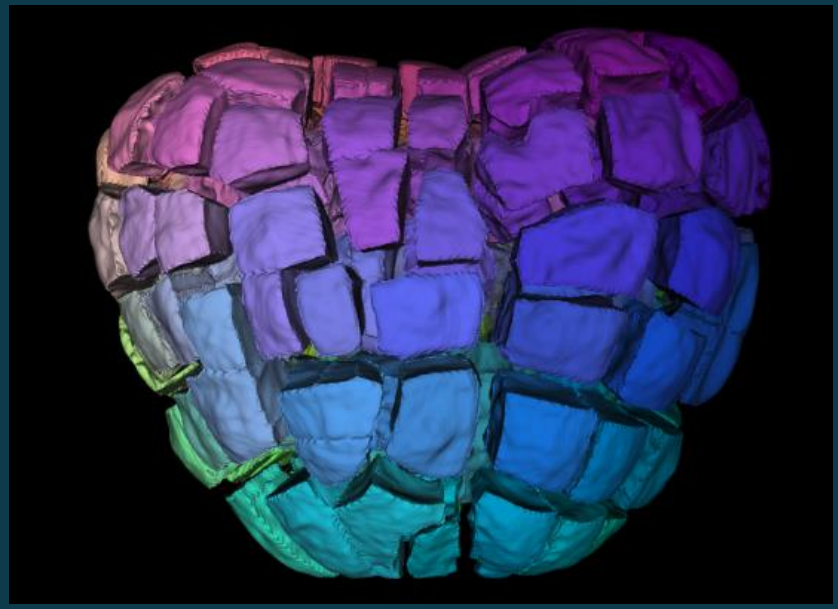
Target & Sister

Interactions

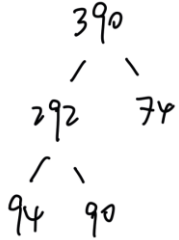
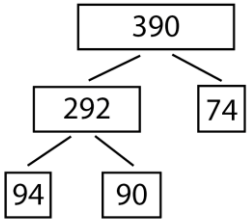
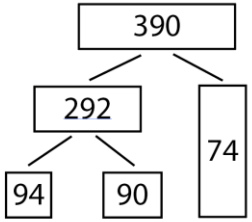
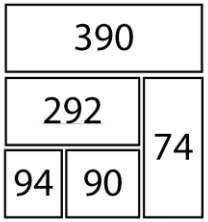
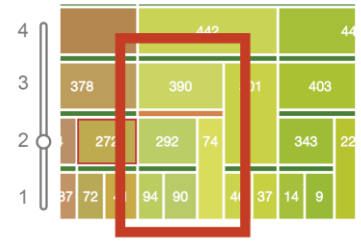
Confirm | Wrong Children | New Sister | Supporting Cell | Predict One Level

Traditional Hand-written Tree

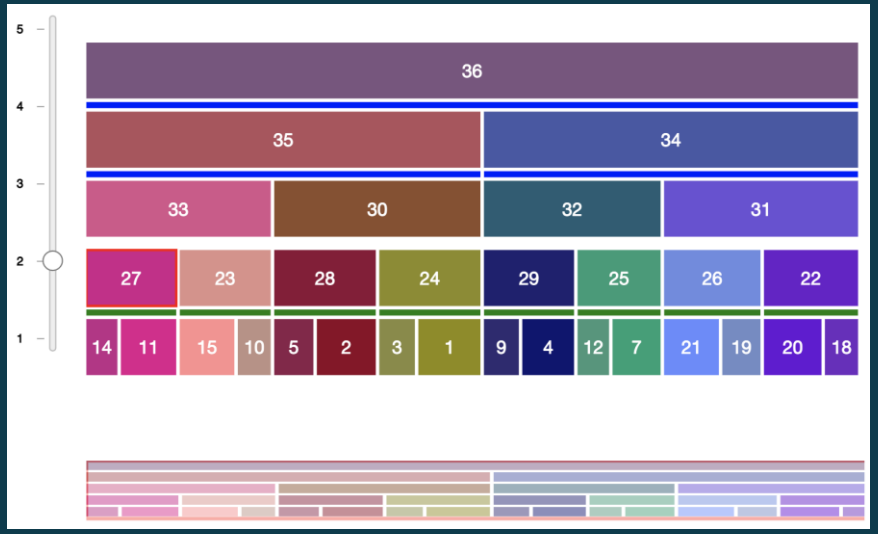


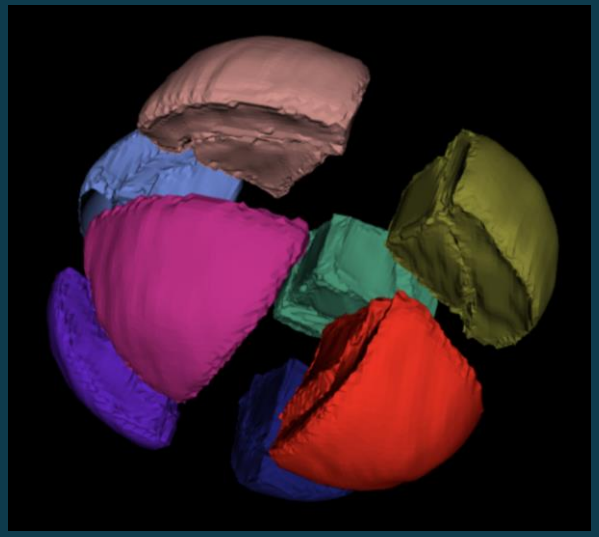


2D Abstract Hierarchy

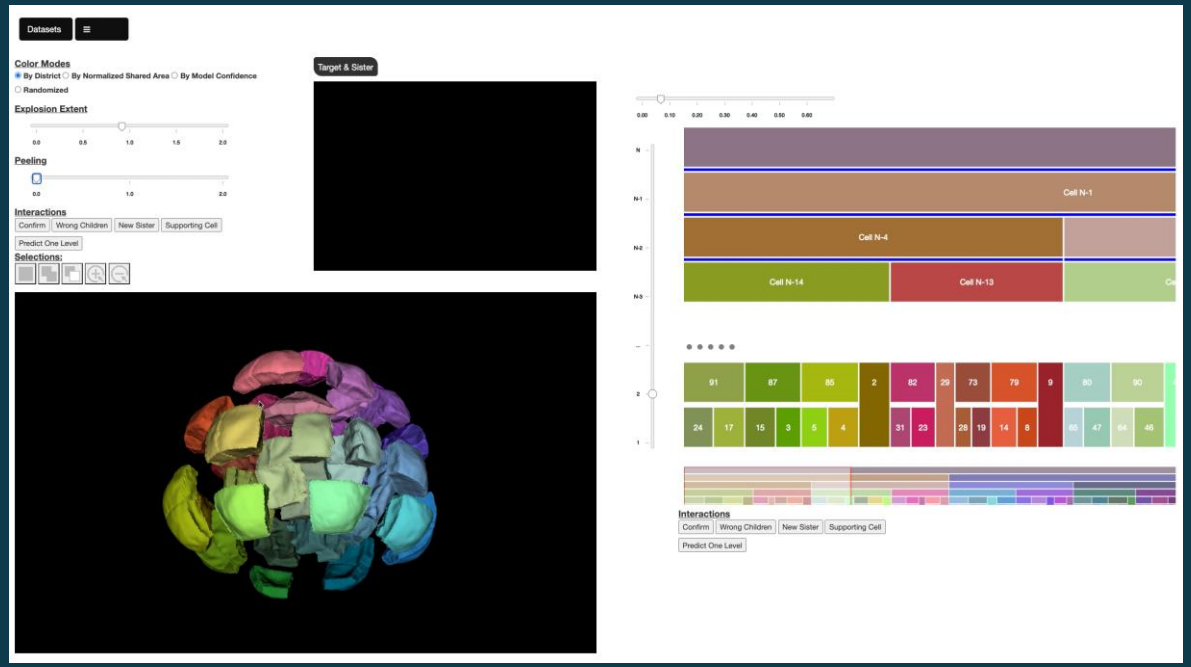
Hand writing records	Encode volume with width	Encode division time stage with height	Remove Links	Final Design
				

Visually and Interactively **Connected**

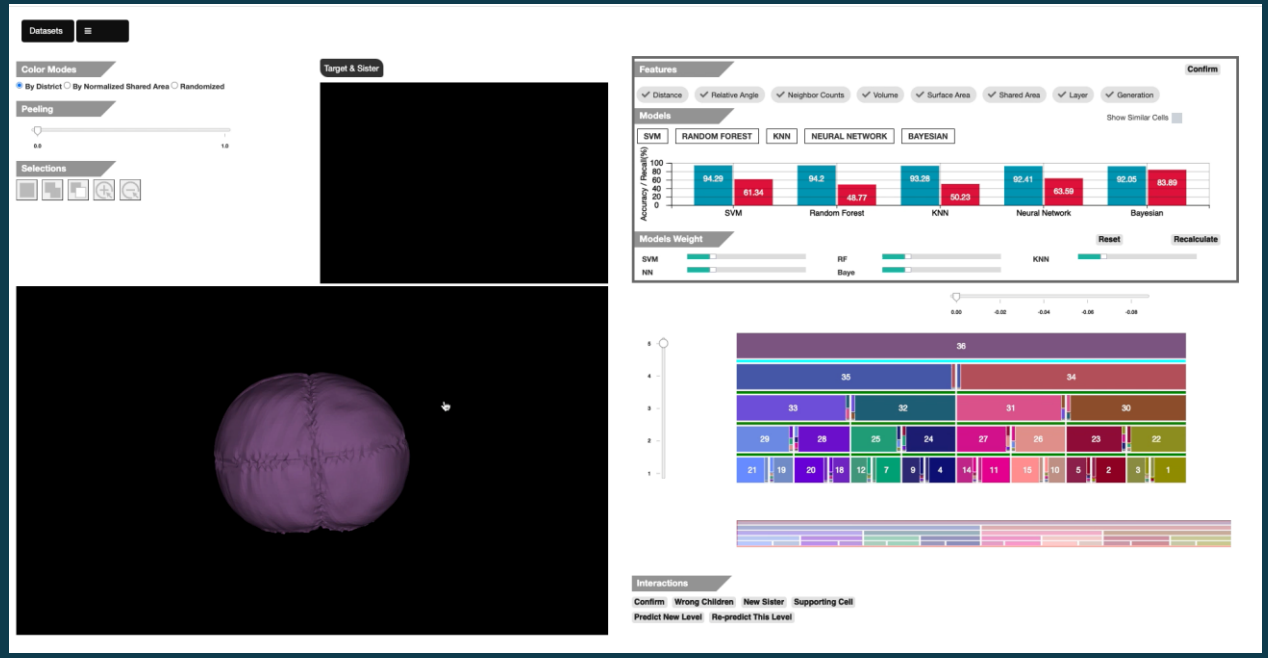




Visually and **Interactively Connected**



Visually and **Interactively Connected**



Advanced Workflow

Two-direction of Hierarchy Building

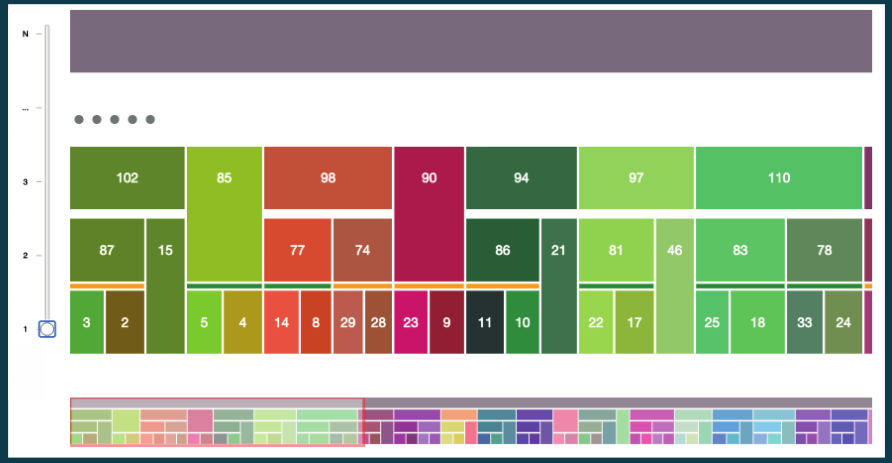
Enable biologists to build the hierarchy tree from both top-down and bottom-up approaches.

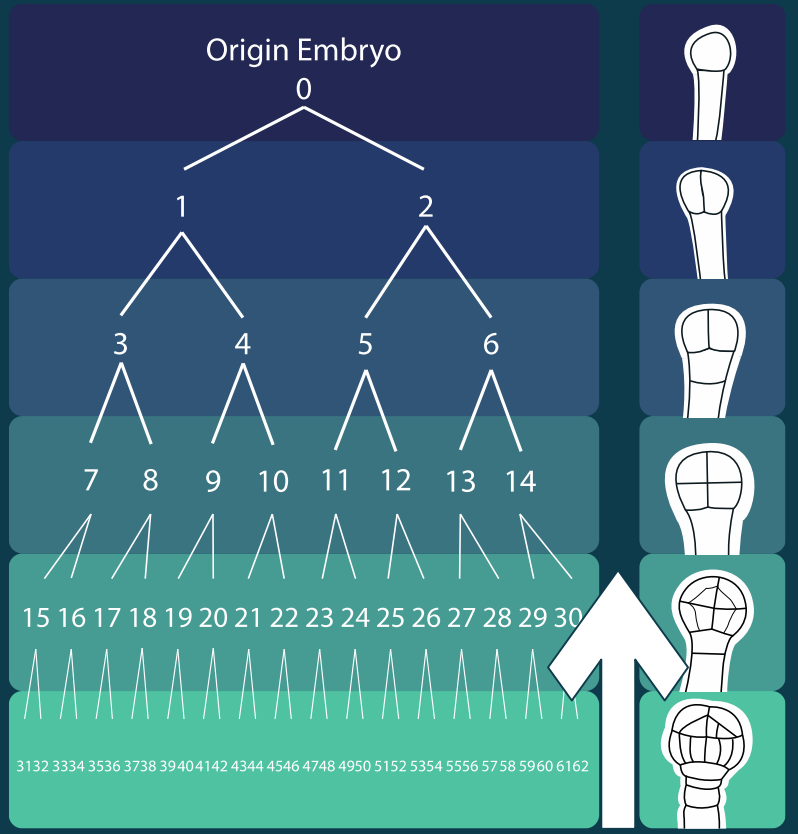
Machine Learning Predictions

Use ML to predict a single level as a basis for biologists to check and correct.

Bottom-up Approach

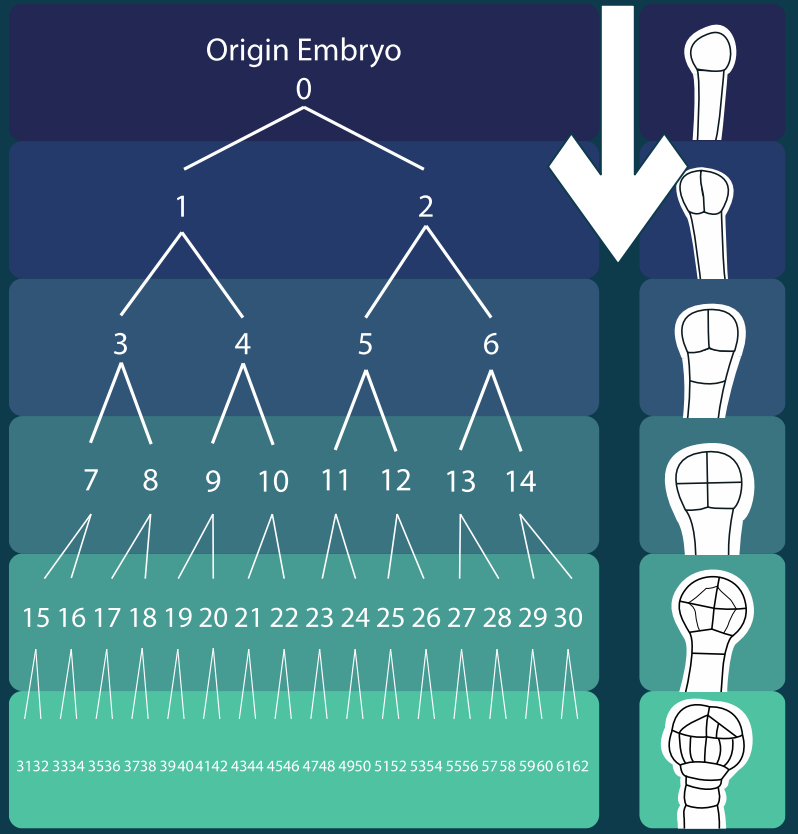
The traditional way to build the hierarchy.





Top-down Approach

Biologists have ideas about how the embryo could be divided in the beginning.



Datasets ≡

Color Modes

By District | By Normalized Shared Area | By Model Confidence | Randomized

Explosion Extent

0.0 0.5 1.0 1.5 2.0

Peeling

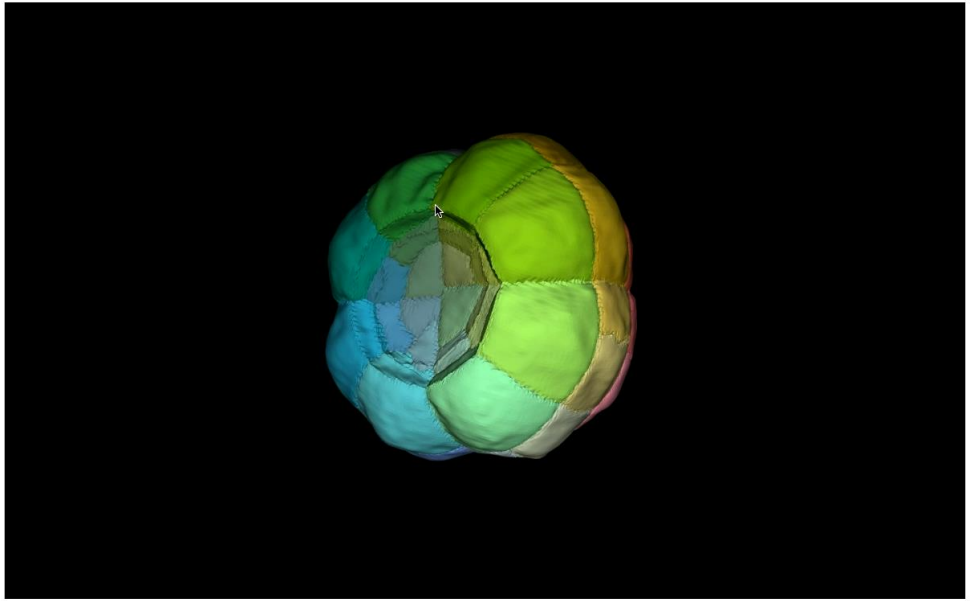
0.0 1.0 2.0

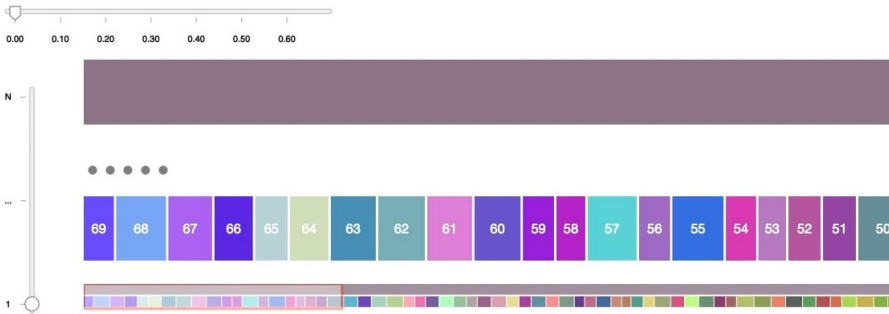
Interactions

Confirm | Wrong Children | New Sister | Supporting Cell | Predict One Level

Selections:

Target & Sister





Interactions

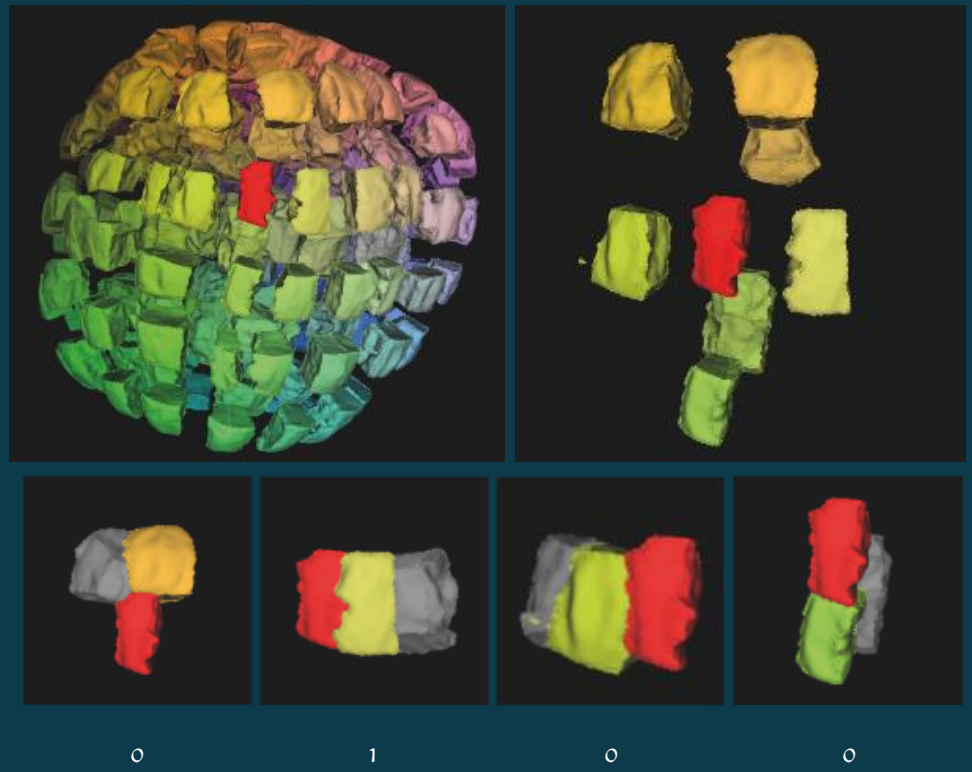
Confirm | Wrong Children | New Sister | Supporting Cell | Predict One Level

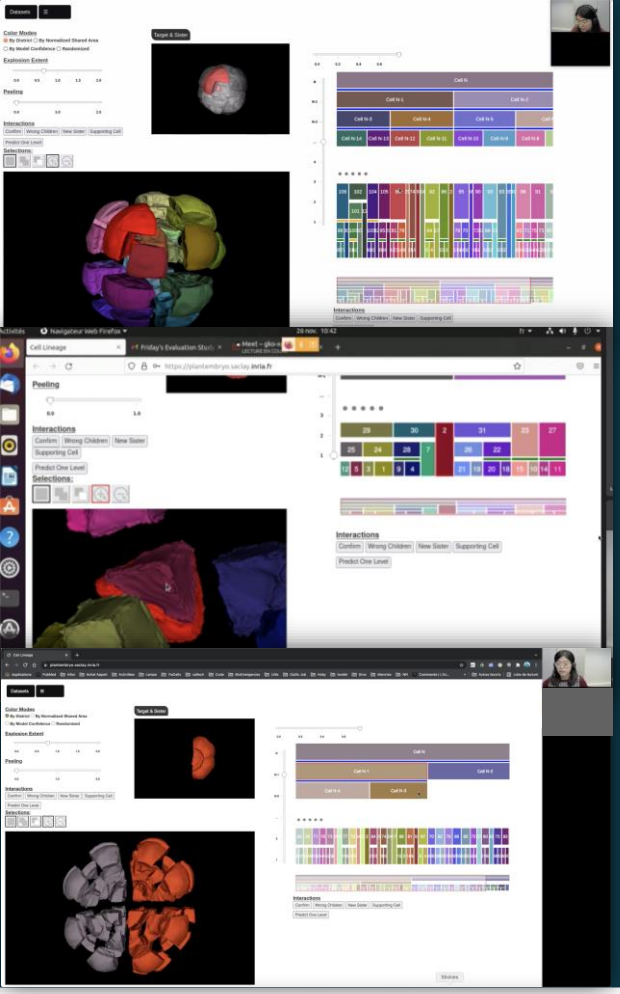
Machine Learning Model

Binary Classification Problem:

Sisters (1) / Non-sisters (0)

- Sequential Neural Network
- 93 Embryo datasets
- 47132 pairs
- 12 features





Evaluation Study

- Biologists took time to learn the functionalities of LineageD.
- They used both the 2D hierarchy and 3D views.
- They valued the visual representations in 2D and 3D.
- They thought LineageD helps to understand the embryos' development better.

ML Results not satisfying

How could we help biologists to have a better collaboration with ML in the case where **the training datasets are limited**?

Our Solution – Human-AI Teaming

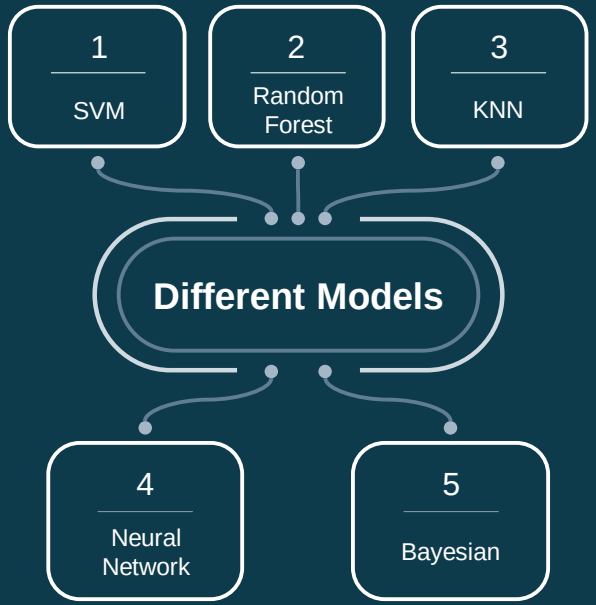
Make use of machine learning

We need to try our best to improve ML performances.

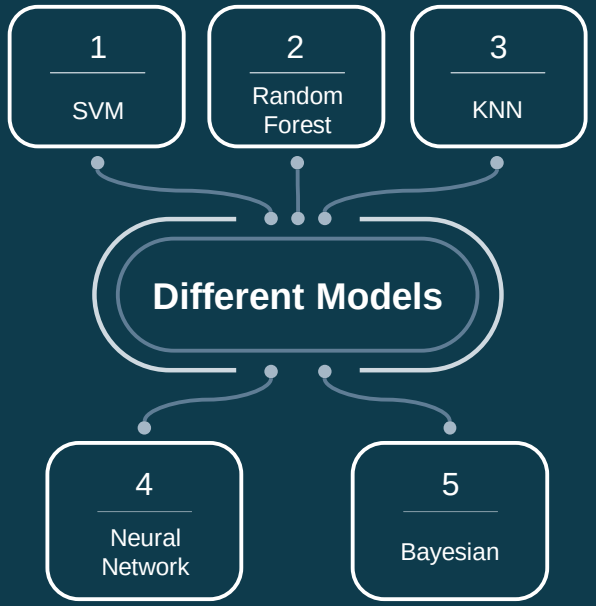
Allow people to control the results

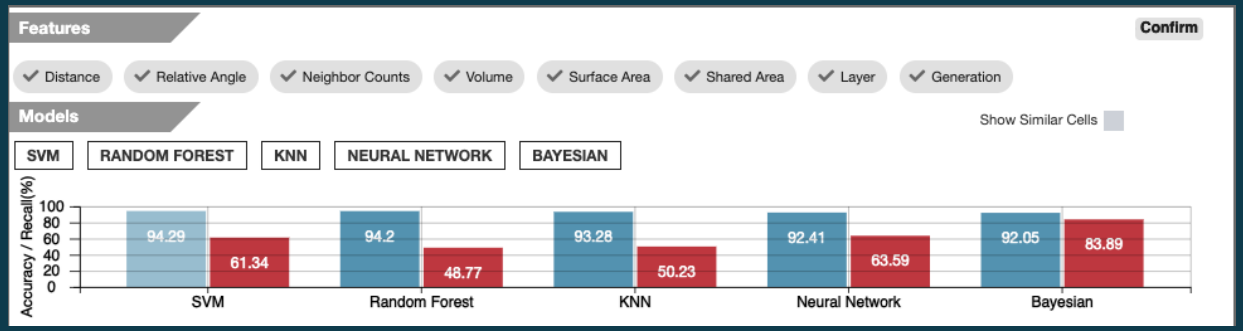
Human beings should have full control over the final decisions.

Model Training



Model Training





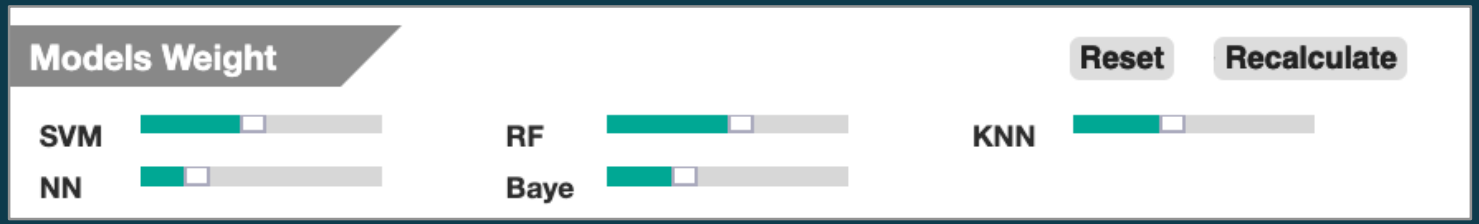
Accuracy

Recall

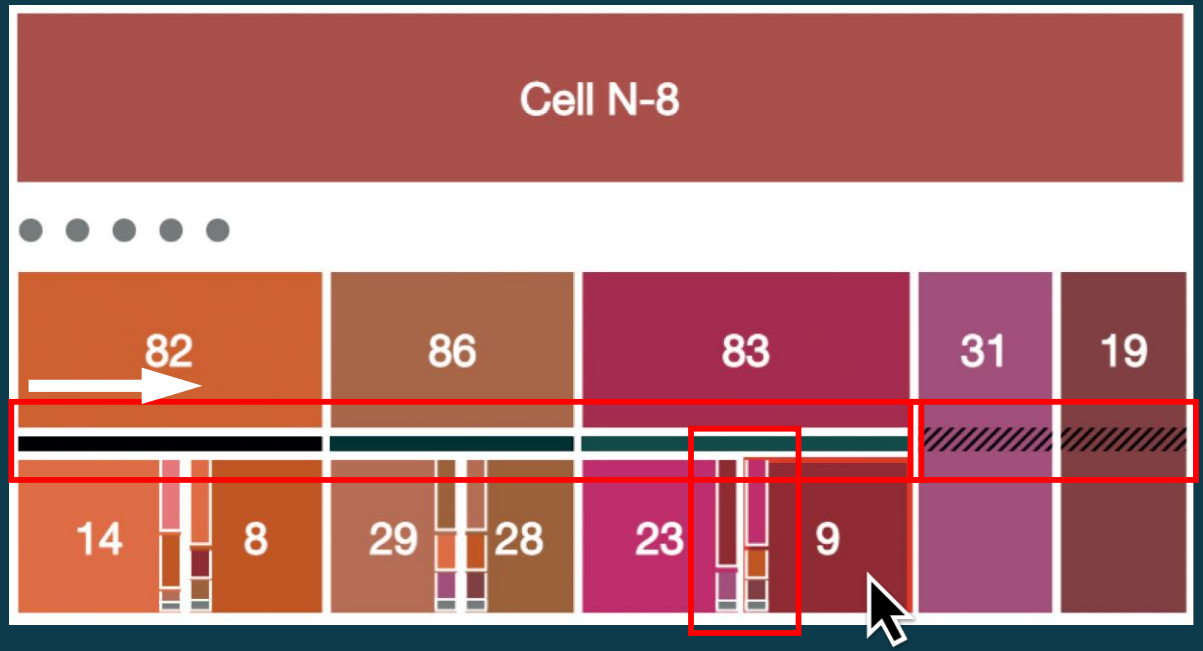
Models Weight

The overall predictions for each cell

The visualization of model prediction confidence

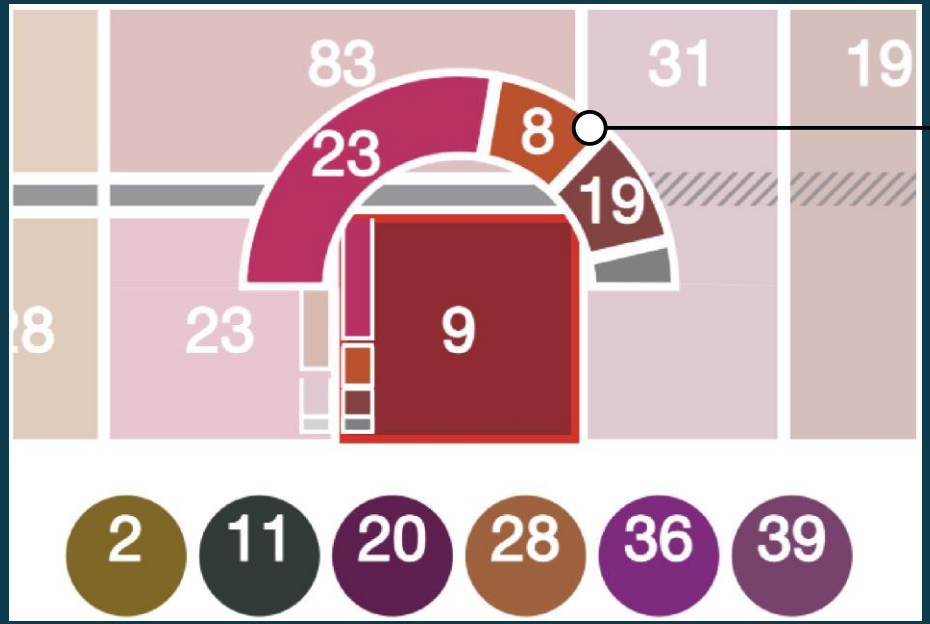


Prediction Visualization



Combining the prediction results from five models, we visualized them with stacked bar charts on each node.

Prediction Visualization

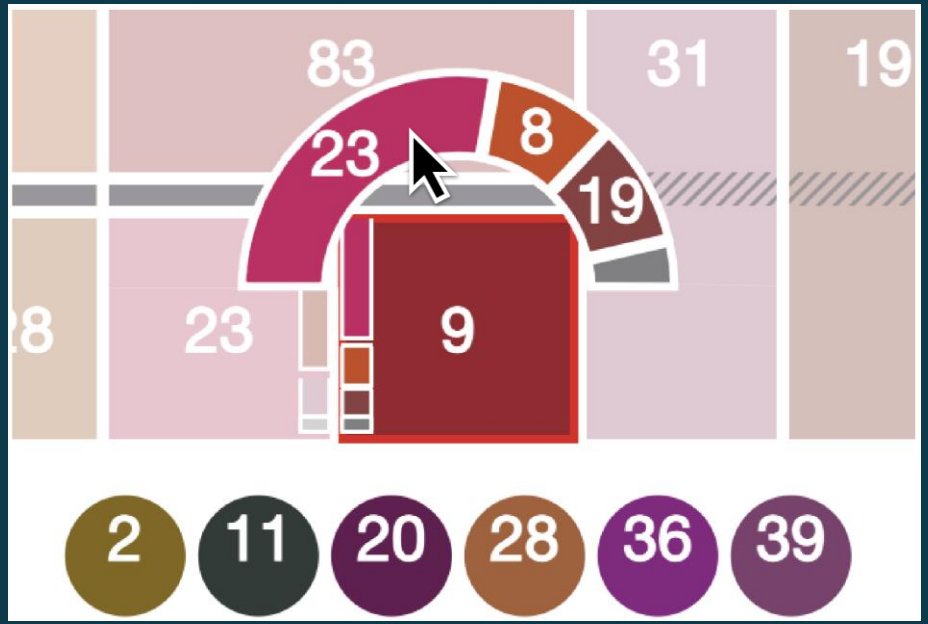


All predicted sisters

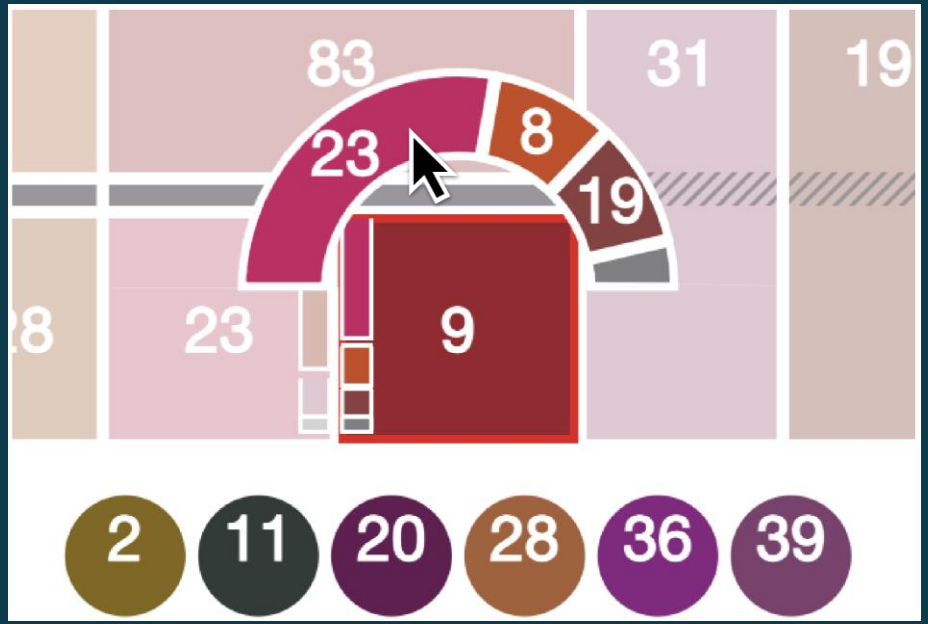
Vertical thumbnail of all predictions

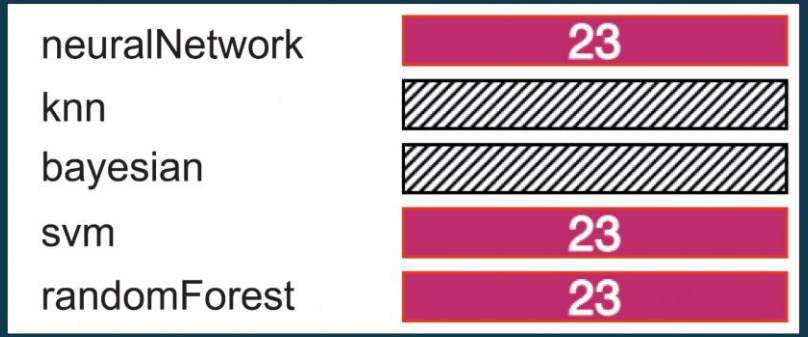
All the other neighboring cells

Prediction Visualization

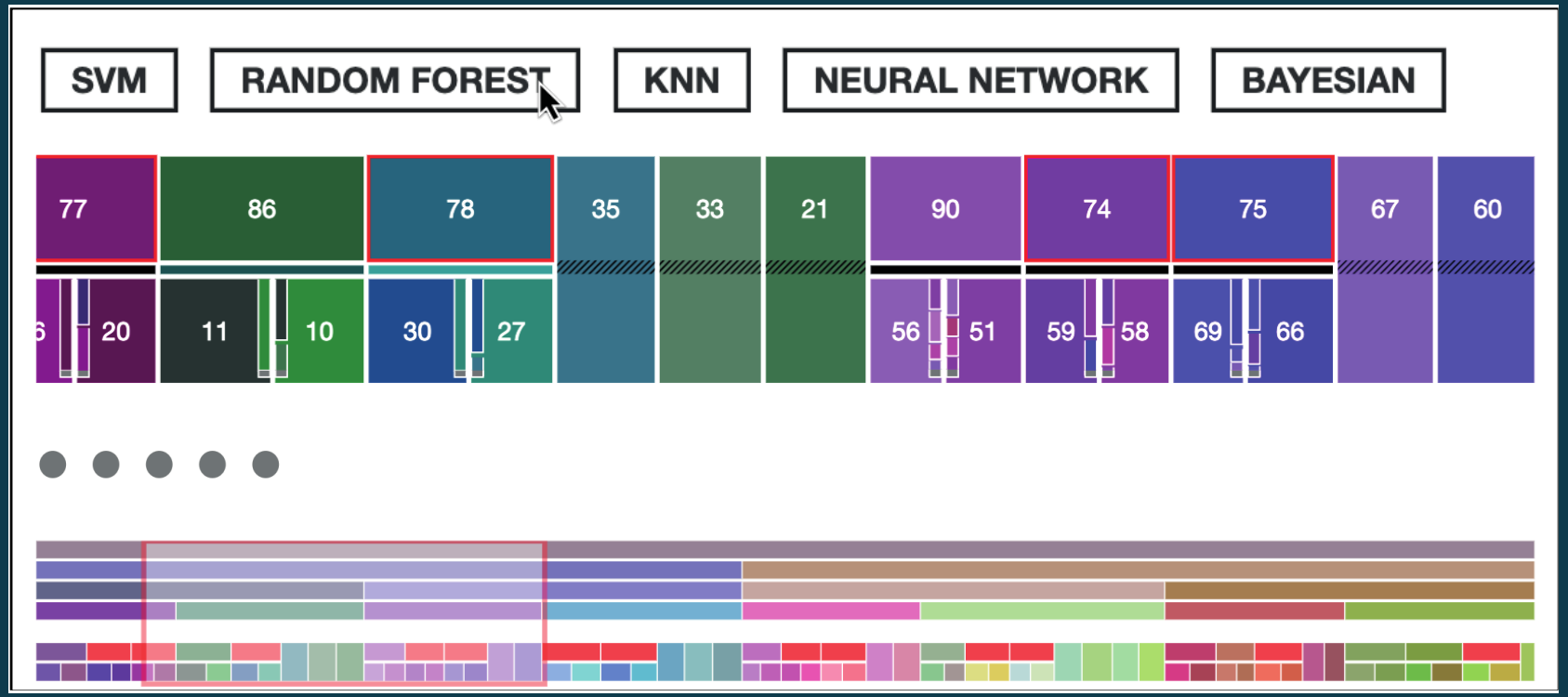


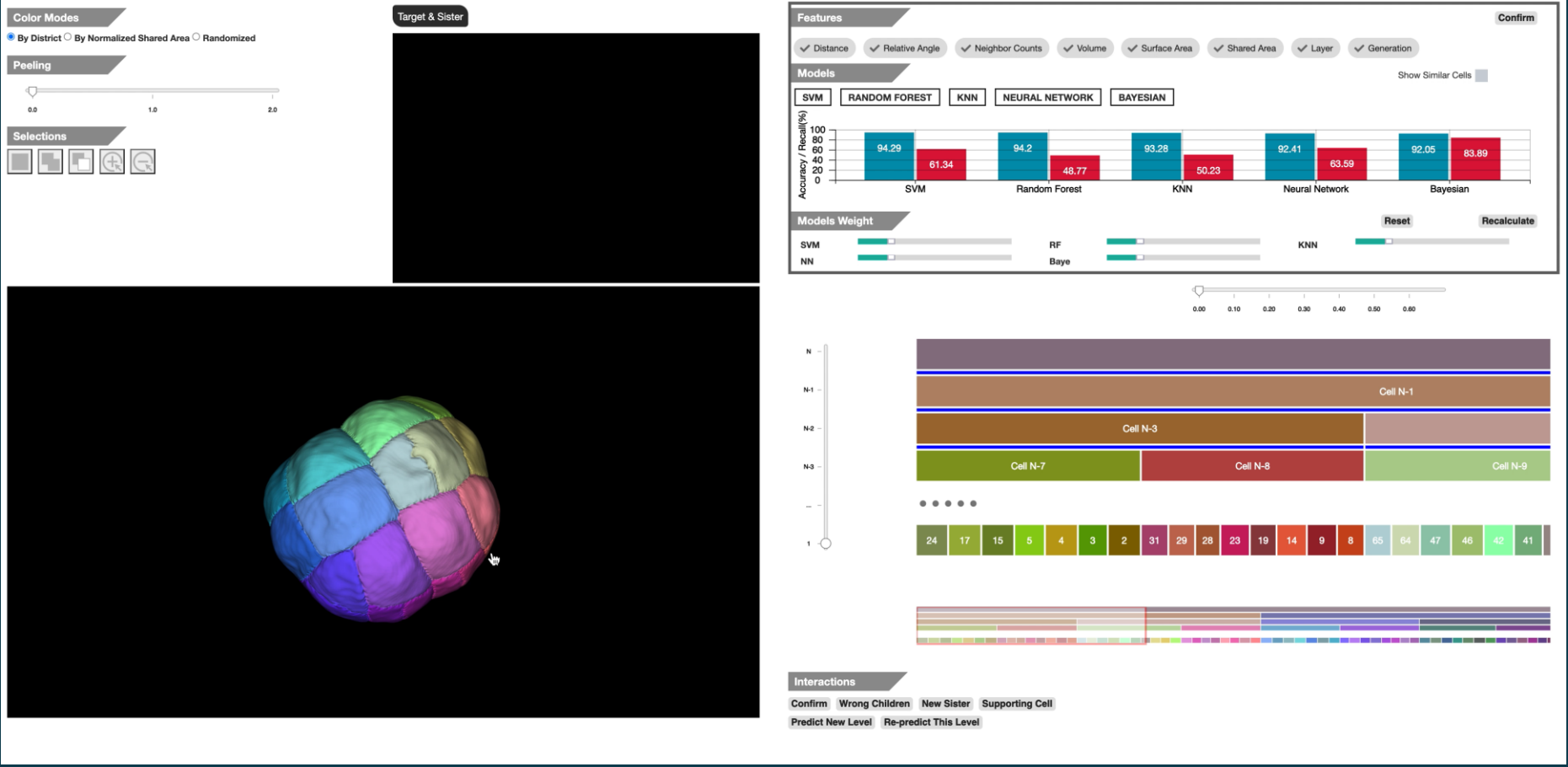
Prediction Visualization





Individual Model Prediction




Color Modes
By District
By Normalized Shared Area
Randomized
Peeling
0.0
1.0
2.0
Selections
Target & Sister
Features
Confirm
Distance
Relative Angle
Neighbor Counts
Volume
Surface Area
Shared Area
Layer
Generation
Models
Show Similar Cells
SVM
RANDOM FOREST
KNN
NEURAL NETWORK
BAYESIAN
Accuracy / Recall(%)
100
80
60
40
20
0
94.29
61.34
94.2
48.77
93.28
50.23
92.41
63.59
92.05
83.89
SVM
Random Forest
KNN
Neural Network
Bayesian
Models Weight
Reset
Recalculate
SVM
RF
KNN
NN
Baye
0.00
0.10
0.20
0.30
0.40
0.50
0.60
N
N-1
N-2
N-3
1
Cell N-1
Cell N-3
Cell N-7
Cell N-8
Cell N-9
24
17
15
5
4
3
2
31
29
28
23
19
14
9
8
65
64
47
46
42
41
Interactions
Confirm
Wrong Children
New Sister
Supporting Cell
Predict New Level
Re-predict This Level

Evaluation Study

- Biologists appreciated the prediction results and their visualization.
- They thought LineageD+ could help save time and change the traditional approach they used in the assignment process.
- One biologist expressed that interacting with ML made her feel like she was discussing with the computer in making decisions.

Plant Cell Lineage Specification based on the Interactive Visualization of Hybrid 3D and 2D Data and with the Support of Machine Learning

Thèse soutenue à Paris-Saclay, le 14 février 2023, par

Jiayi Hong