MulteeSum

A Tool for Comparative Spatial and Temporal Gene Expression Data

Miriah Meyer,[1] Tamara Munzner,[2] Angela DePace,[3] Hanspeter Pfister[1]

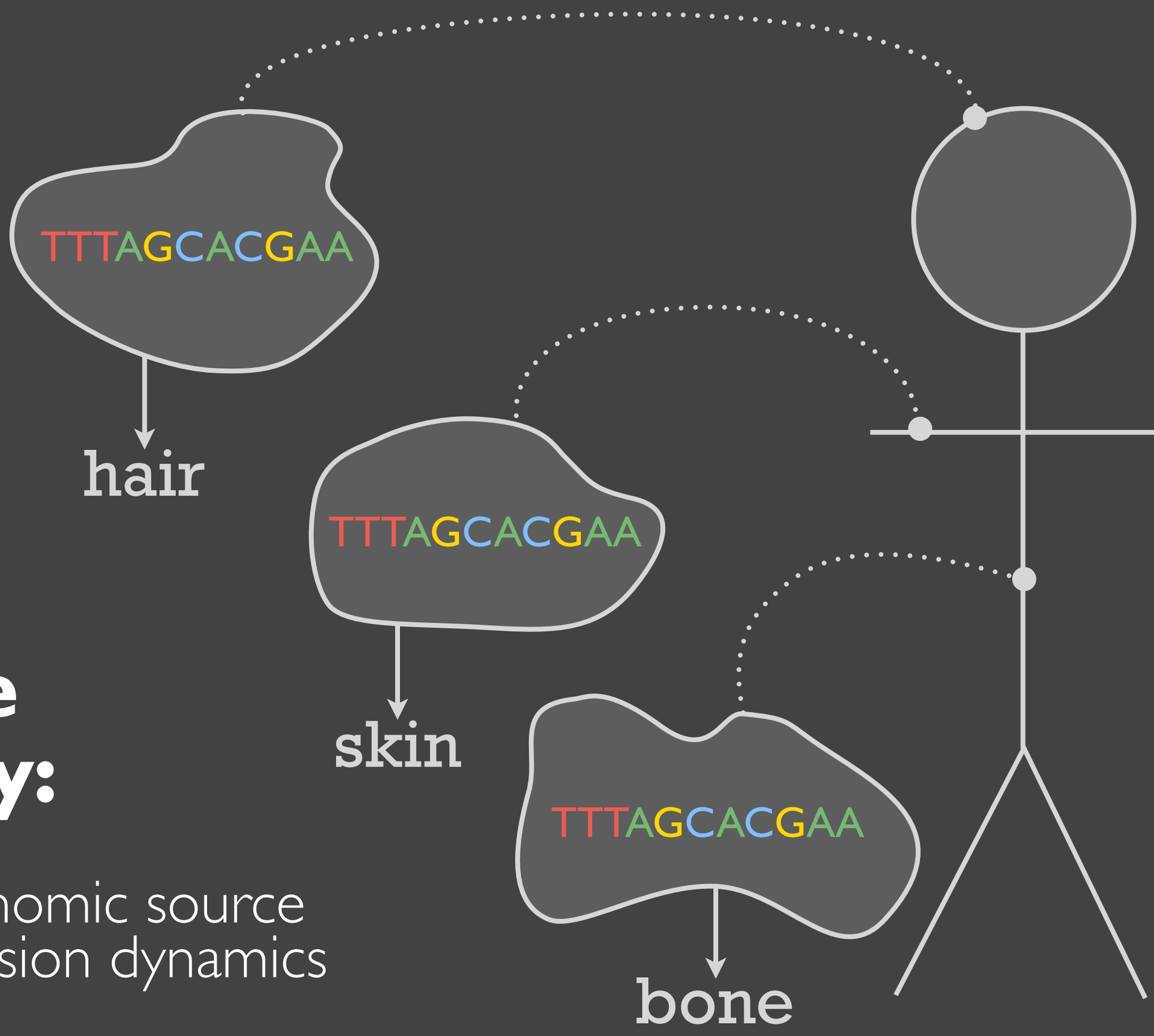
[1] Harvard University

[2] University of British Columbia

[3] Harvard Medical School

grand challenge in biology:

understand genomic source
of gene expression dynamics



collaborators: DePace lab at Harvard Medical School

model species: fruit fly

scientific goal: link changes in the regulatory part of the genome to species variation

requires: characterize differences in gene expression patterns between species

MulteeSum

process

two year collaboration

two early prototype systems

feedback from six biologists

informal interviews, emails

one day a week in biology lab

tool deployed

currently used several times a week

data & tool & tasks

summaries & groups

encodings & interaction

conclusions

gene expression

... the measured level of how much a gene is on or off.

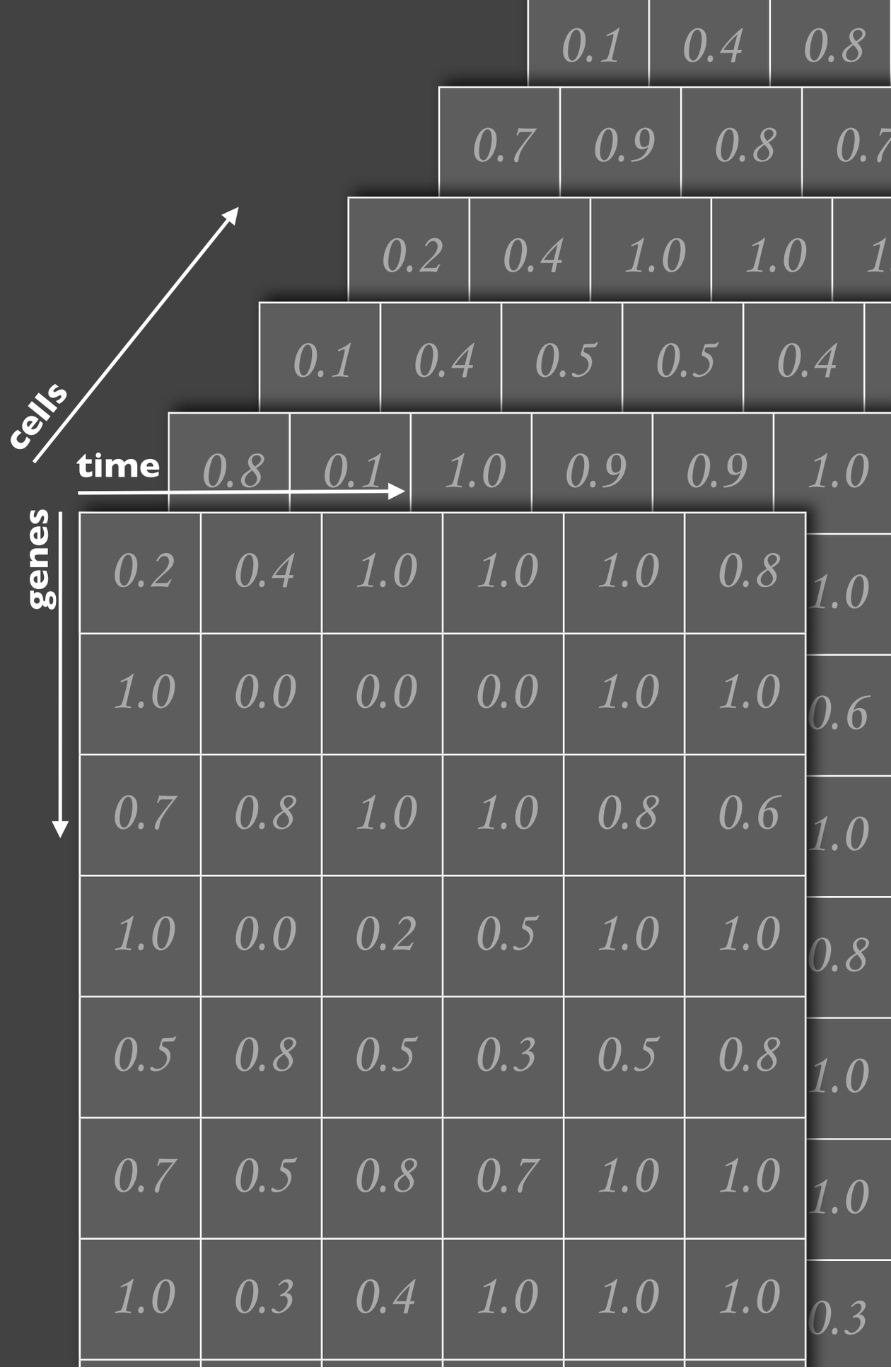
... a single quantitative value.

collaborators measure it ...

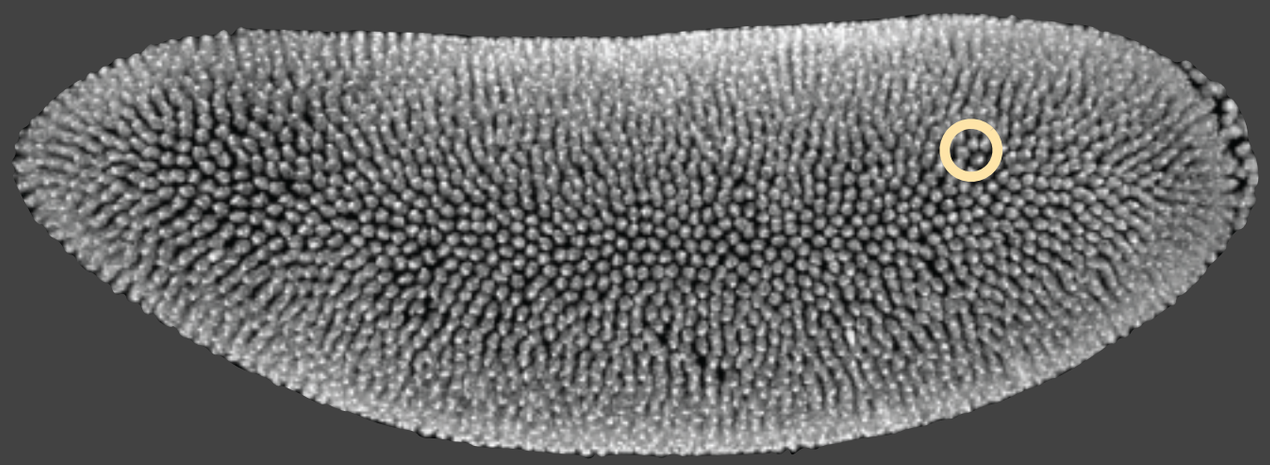
... for multiple time points.

... for multiple genes.

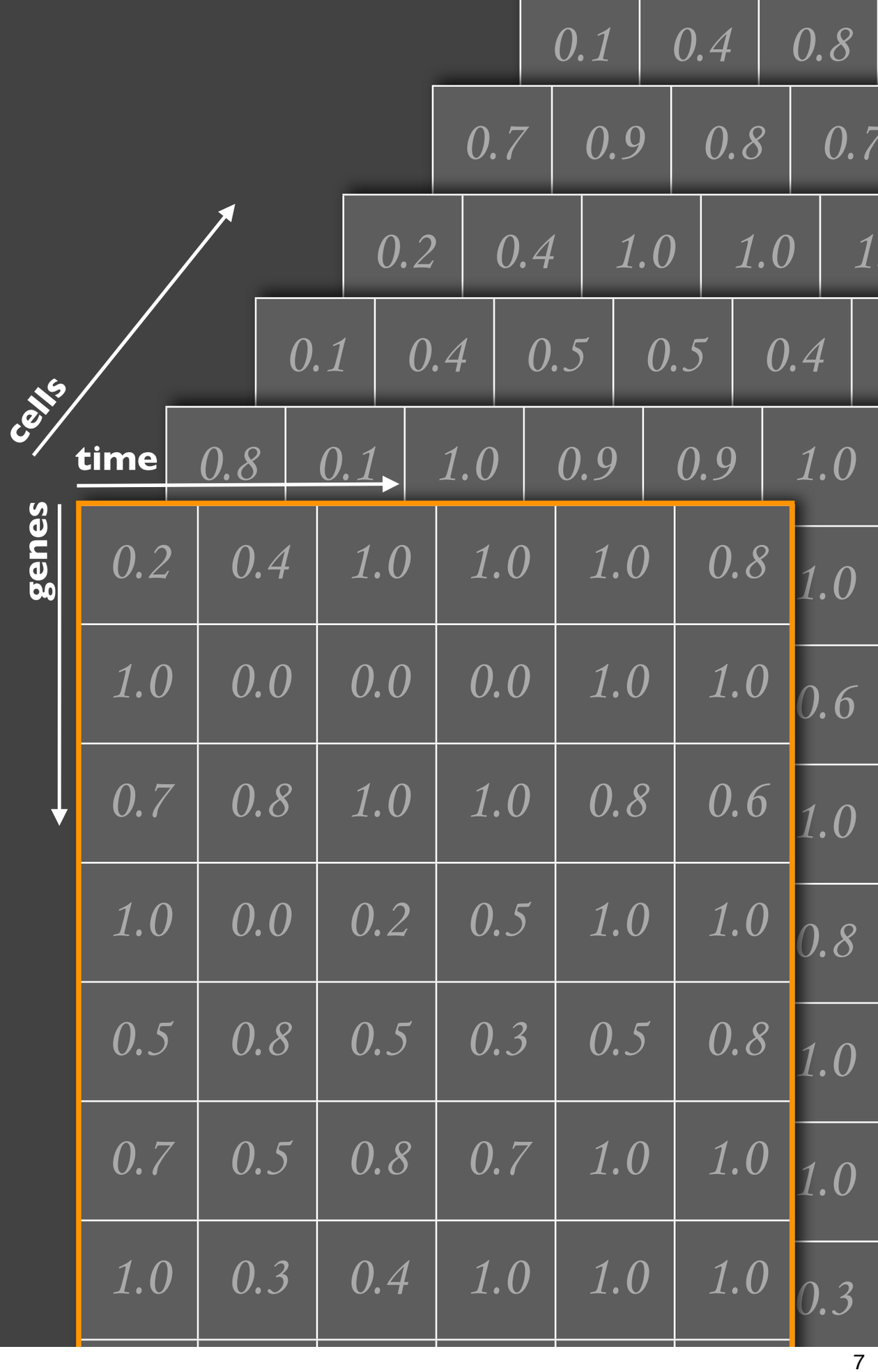
... in many cells.



measurements taken for every cell in a single fruit fly embryo.



correlate gene expression with spatial location.



virtual embryo

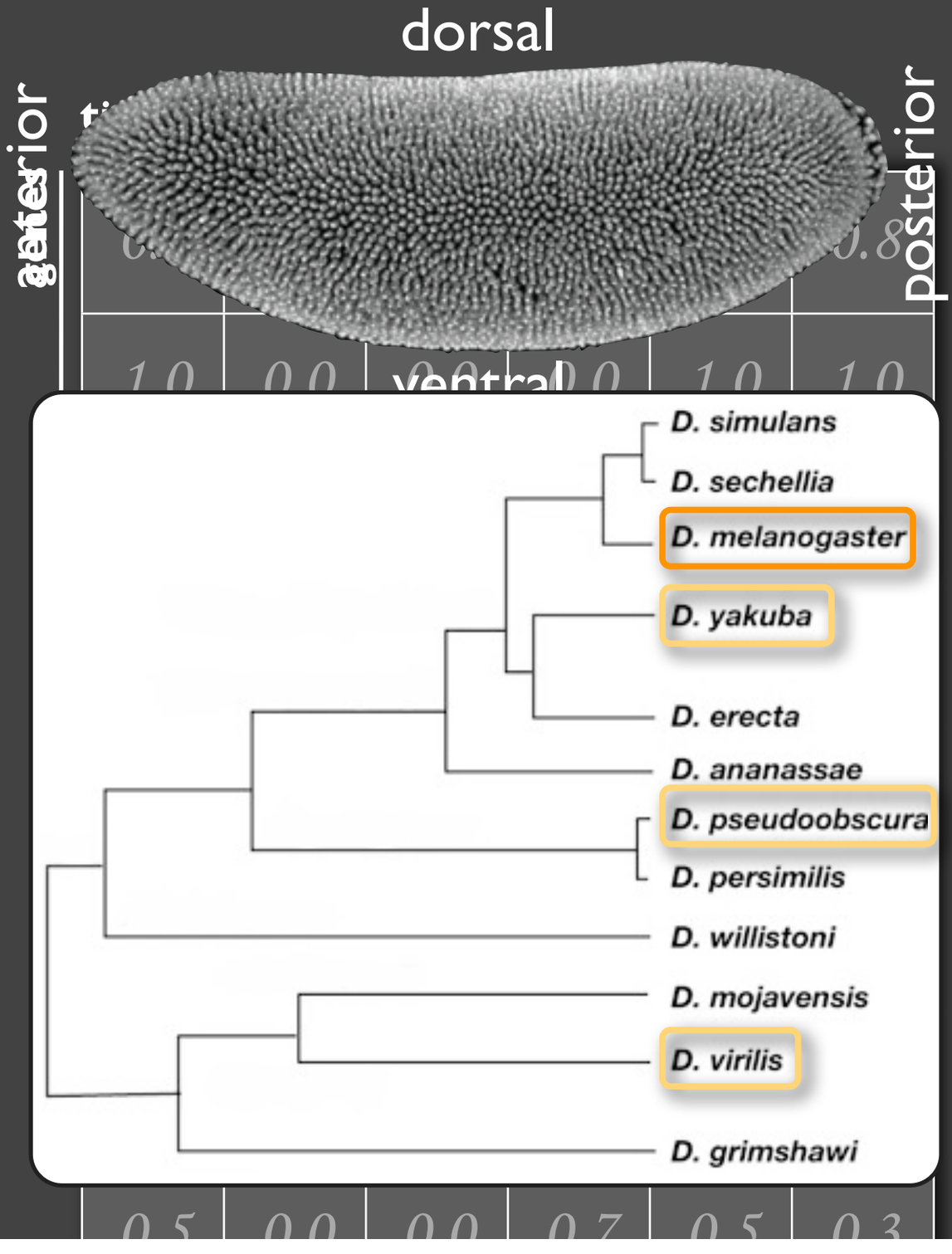
several thousand cells
[5,000 ± 1,000]

each cell has:

expression profile
[6 time points x 50 genes]

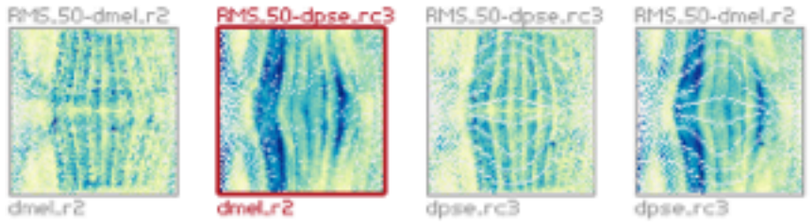
spatial position
[3D and 2D coordinates]

12 related species
one complete
three preliminary

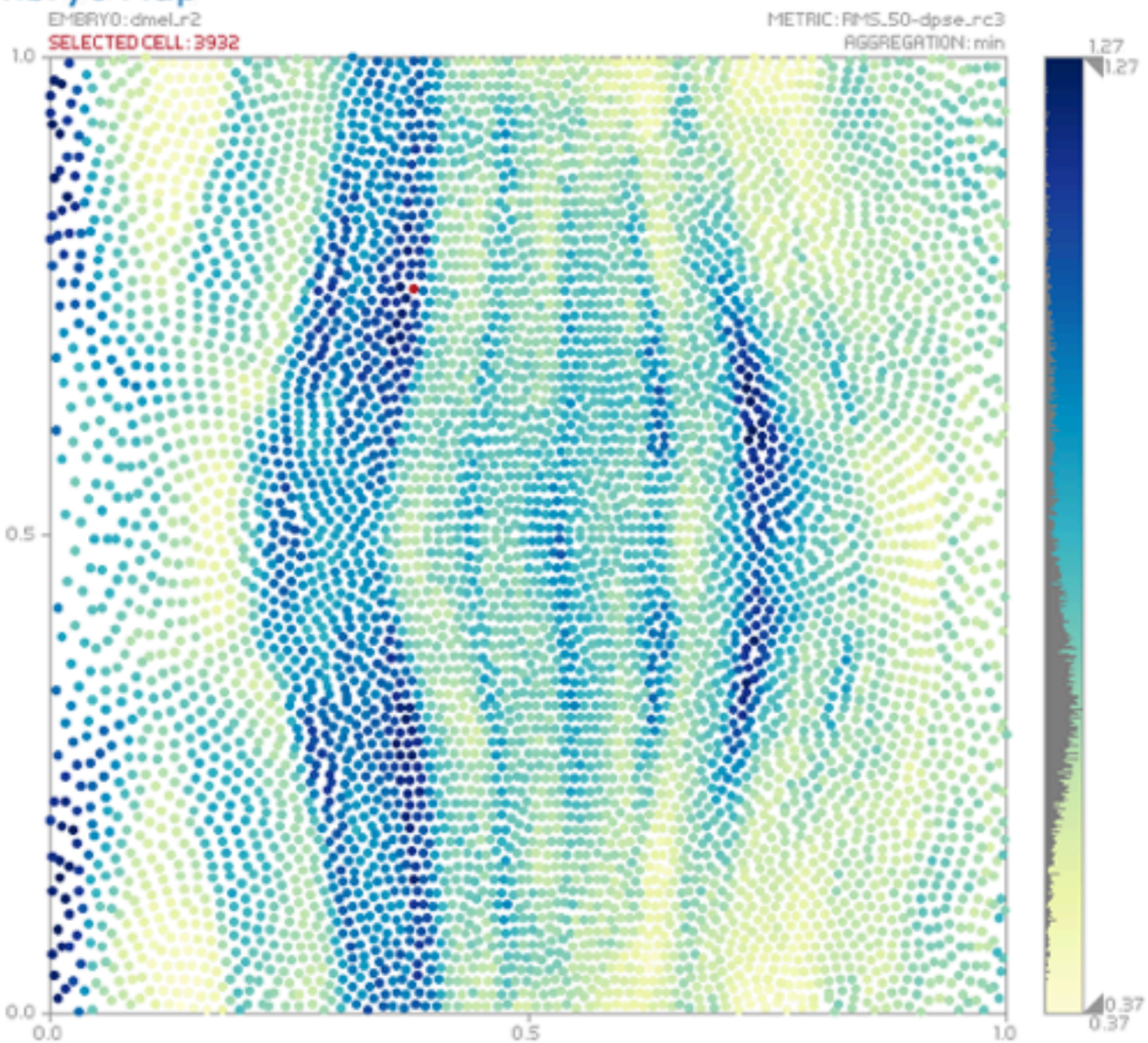


MulteeSum

Summaries



Embryo Map



Curvemap

MulteeSum

Summaries

RMS_50-dmel_r2 dmel_r2 | RMS_50-dpse_rc3 dmel_r2 | RMS_50-dpse_rc3 dpse_rc3 | RMS_50-dmel_r2 dpse_rc3

Embryo Map

spatial

EMBRYO: dmel_r2
SELECTED CELL: 3932
METRIC: RMS_50-dpse_rc3
AGGREGATION: min

embryo map

AGGREGATION GROUP | CREATED GROUP

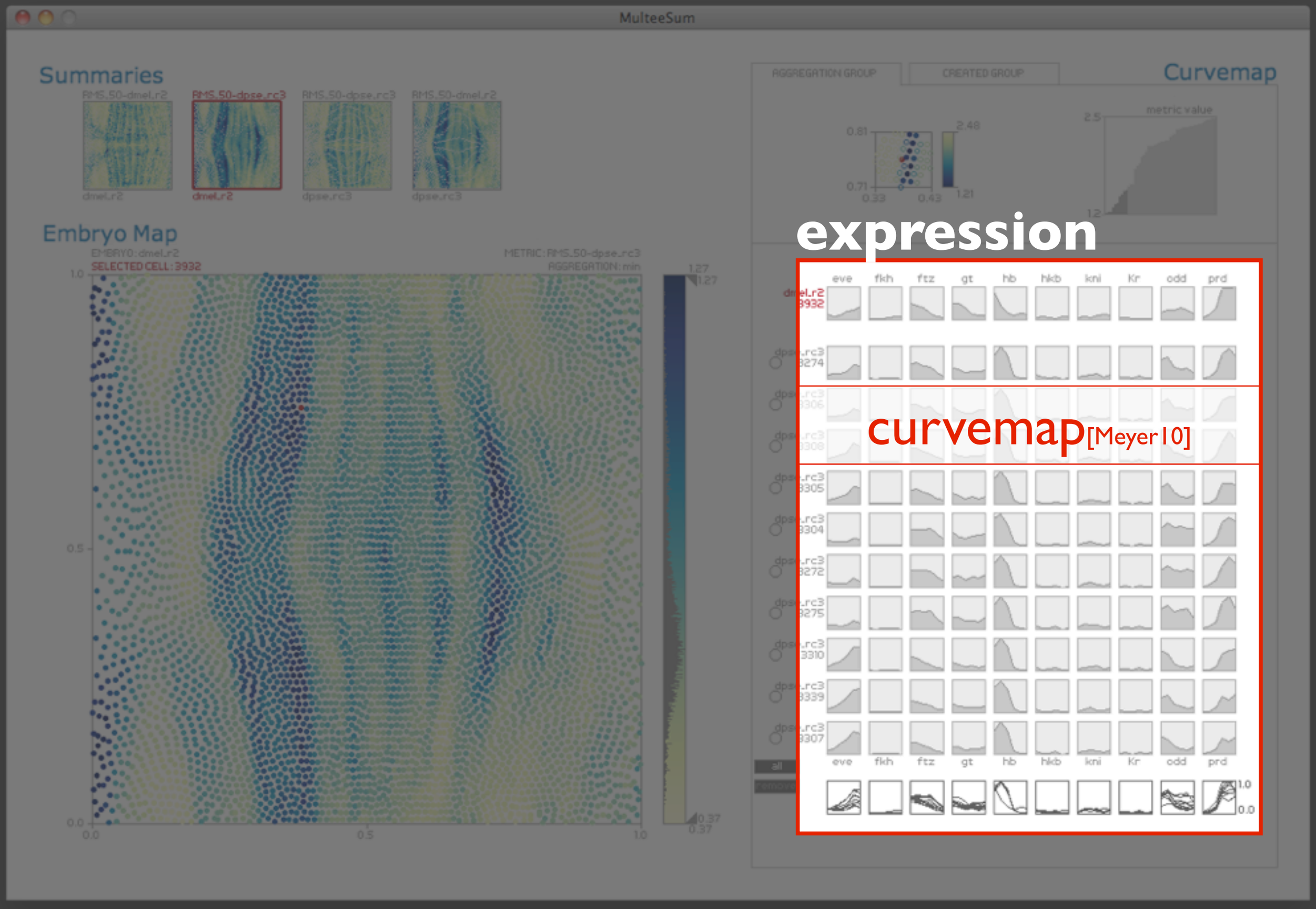
Curvemap

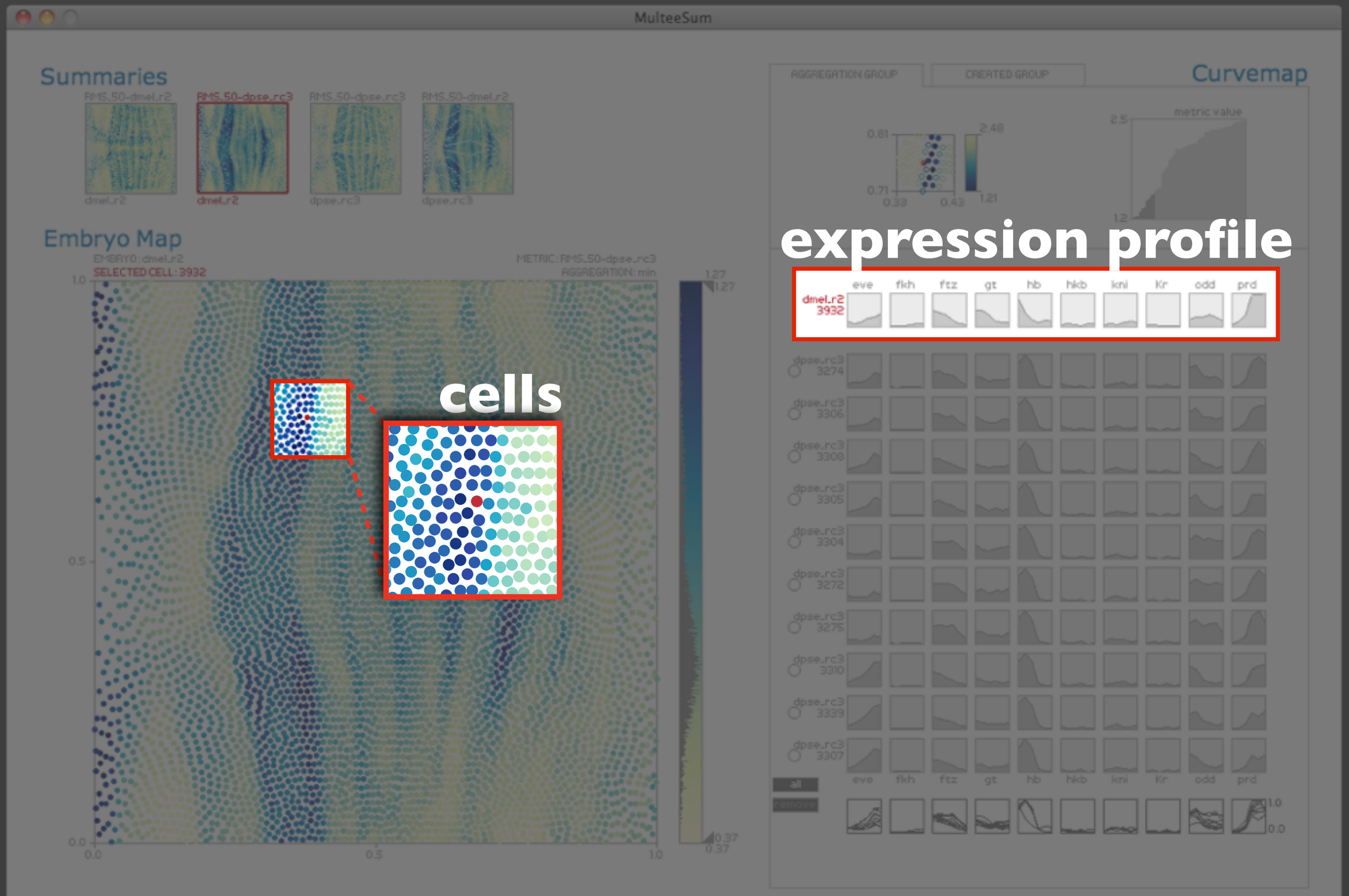
metric value

eve fkh ftz gt hb hkb kni Kr odd prd

dmel_r2 3932
dpse_rc3 3274
dpse_rc3 3306
dpse_rc3 3308
dpse_rc3 3305
dpse_rc3 3304
dpse_rc3 3272
dpse_rc3 3275
dpse_rc3 3310
dpse_rc3 3339
dpse_rc3 3307

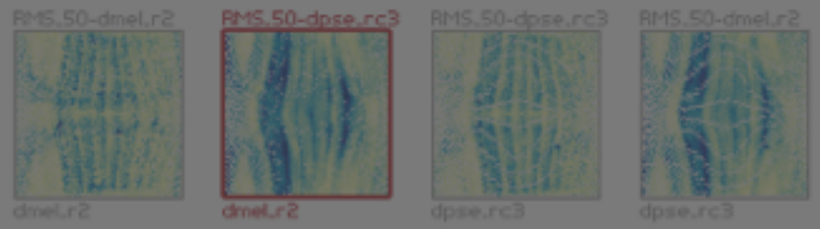
all
remove


MulteeSum
Summaries
RMS.50-dmel.r2
RMS.50-dpse.rc3
RMS.50-dpse.rc3
RMS.50-dmel.r2
dmel.r2
dmel.r2
dpse.rc3
dpse.rc3
Embryo Map
EMBRYO: dmel.r2
SELECTED CELL: 3932
METRIC: RMS.50-dpse.rc3
AGGREGATION: min
1.0
0.5
0.0
0.0
0.5
1.0
1.27
1.27
0.37
0.37
AGGREGATION GROUP
CREATED GROUP
Curvemap
0.81
0.71
0.33
0.43
2.48
1.21
metric value
2.5
1.2
expression
eve
fkh
ftz
gt
hb
hkb
kni
Kr
odd
prd
curvemap [Meyer10]
all
remove
1.0
0.0

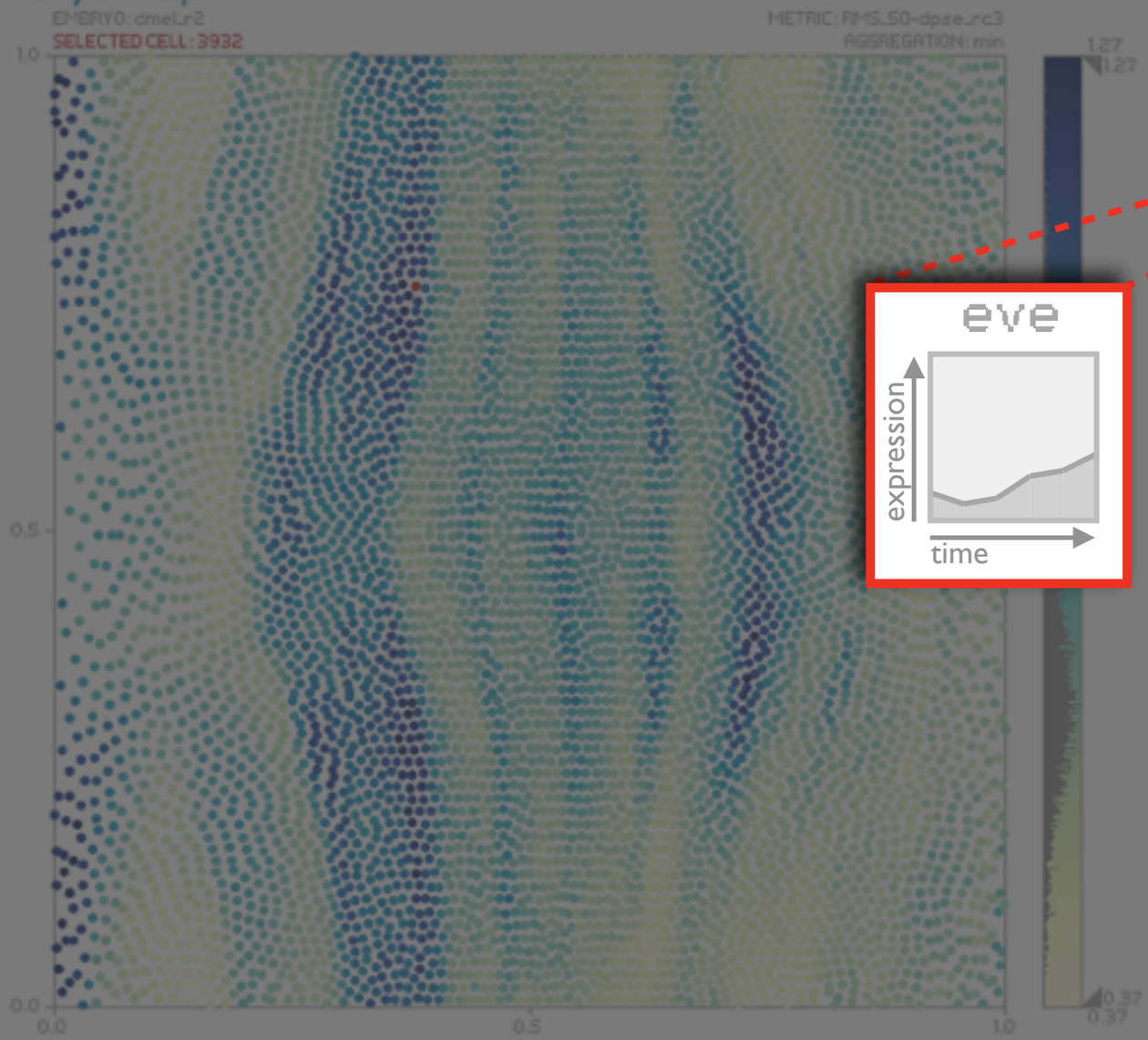

MulteeSum
Summaries
RMS.50-dmel.r2
RMS.50-dpse.rc3
RMS.50-dpse.rc3
RMS.50-dmel.r2
dmel.r2
dmel.r2
dpse.rc3
dpse.rc3
Embryo Map
EMBRYO: dmel.r2
SELECTED CELL: 3932
METRIC: RMS.50-dpse.rc3
AGGREGATION: min
cells
AGGREGATION GROUP
CREATED GROUP
Curvemap
metric value
expression profile
eve fkh ftz gt hb hkb kni Kr odd prd
dmel.r2 3932
dpse.rc3 3274
dpse.rc3 3306
dpse.rc3 3308
dpse.rc3 3305
dpse.rc3 3304
dpse.rc3 3272
dpse.rc3 3275
dpse.rc3 3310
dpse.rc3 3339
dpse.rc3 3307
all
none

MulteeSum

Summaries

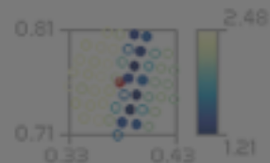


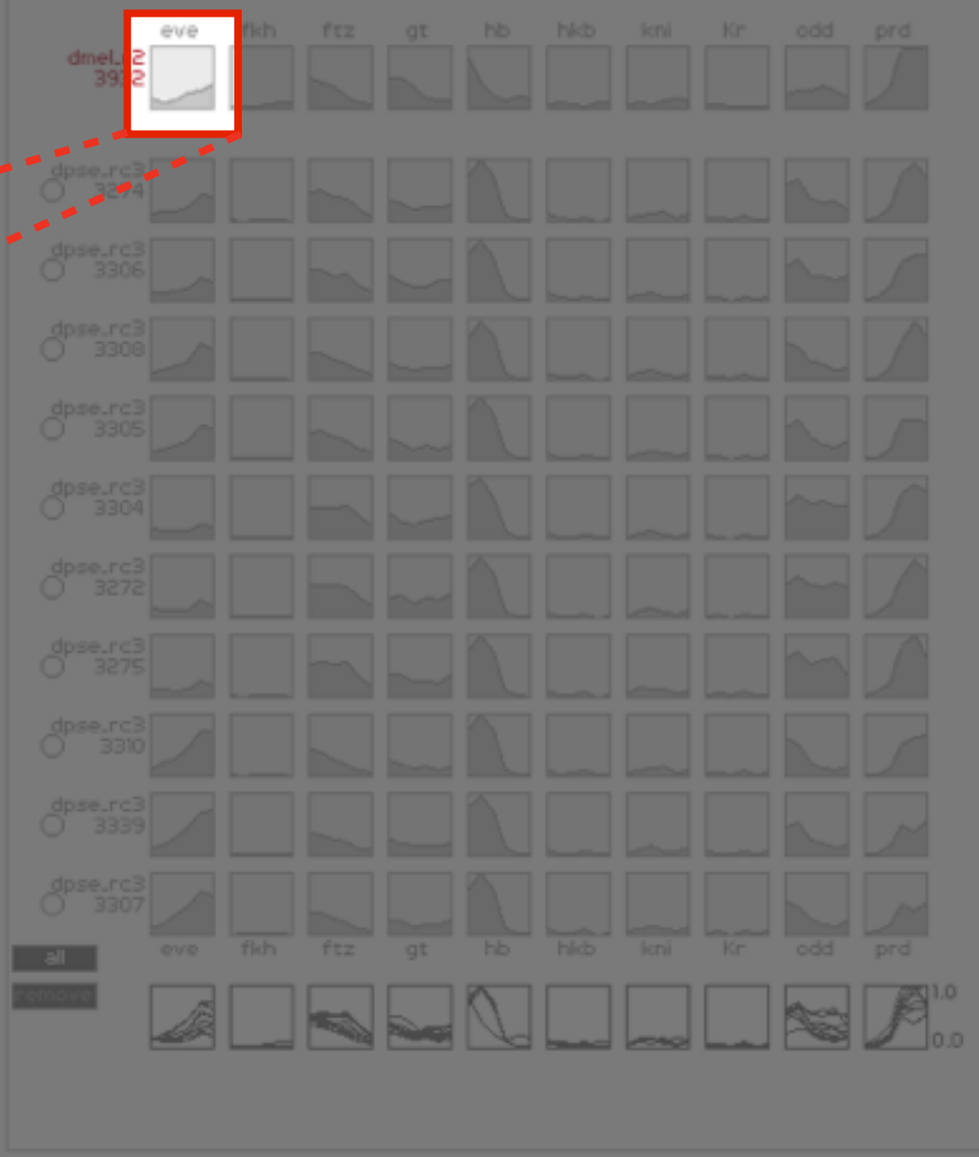
Embryo Map



AGGREGATION GROUP | CREATED GROUP

Curvemap





MulteeSum

Summaries

RMS_50-dmel_r2 dmel_r2 | RMS_50-dpse_rc3 dmel_r2 | RMS_50-dpse_rc3 dpse_rc3 | RMS_50-dmel_r2 dpse_rc3

Embryo Map

EMBRYO: dmel_r2
SELECTED CELL: 3932
METRIC: RMS_50-dpse_rc3
AGGREGATION: min

Curvemap

AGGREGATION GROUP | CREATED GROUP

metric value

eve fkh ftz gt hb hkb kni Kr odd prd

dmel_r2 3932
dpse_rc3 3274
dpse_rc3 3275
dpse_rc3 3310
dpse_rc3 3339
dpse_rc3 3307

all
none

MulteeSum

Summaries

RMS_50-dmel_r2 dmel_r2 | RMS_50-dpse_rc3 dmel_r2 | RMS_50-dpse_rc3 dpse_rc3 | RMS_50-dmel_r2 dpse_rc3

Embryo Map

EMBRYO: dmel_r2
SELECTED CELL: 3932
METRIC: RMS_50-dpse_rc3
AGGREGATION: min

Curvemap

AGGREGATION GROUP | CREATED GROUP

metric value

MulteeSum

Summaries

RMS.50-dmel.r2 dmel.r2 | RMS.50-dpse.rc3 dmel.r2 | RMS.50-dpse.rc3 dpse.rc3 | RMS.50-dmel.r2 dpse.rc3

Embryo Map

EMBRYO: dmel.r2
SELECTED CELL: 3932
METRIC: RMS.50-dpse.rc3
AGGREGATION: min

Curvemap

AGGREGATION GROUP | CREATED GROUP

metric value

genes

cells

eve fkh ftz gt hb hkb kni Kr odd prd

all
remove

MulteeSum

Summaries

RMS.50-dmel.r2 dmel.r2 | RMS.50-dpse.rc3 dmel.r2 | RMS.50-dpse.rc3 dpse.rc3 | RMS.50-dmel.r2 dpse.rc3

Embryo Map

EMBRYO: dmel.r2
SELECTED CELL: 3932
METRIC: RMS.50-dpse.rc3
AGGREGATION: min

summary value

Curvemap

AGGREGATION GROUP | CREATED GROUP

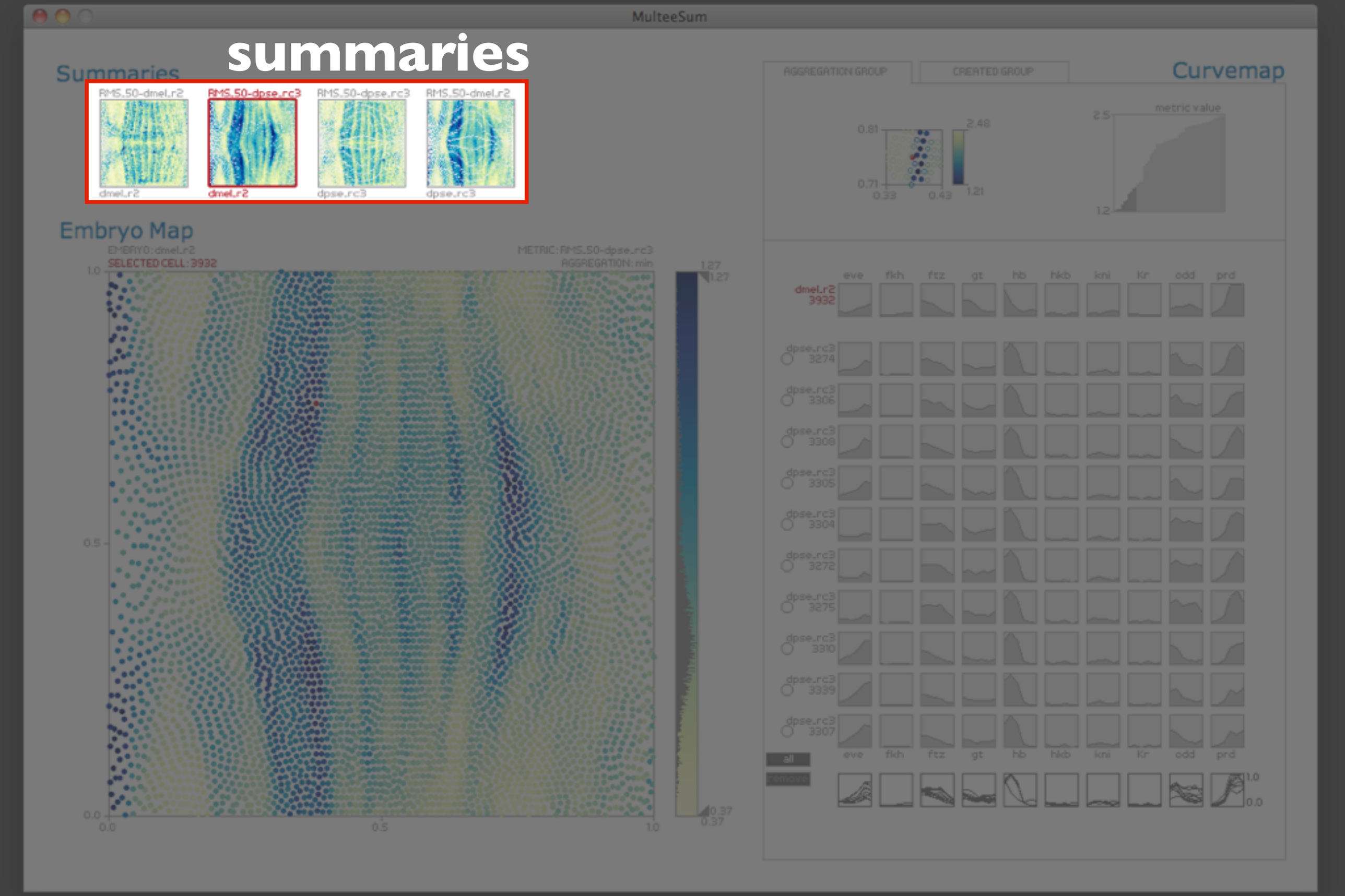
metric value

eve fkh ftz gt hb hkb kni Kr odd prd

dmel.r2 3932
dpse.rc3 3274
dpse.rc3 3306
dpse.rc3 3308
dpse.rc3 3305
dpse.rc3 3304
dpse.rc3 3272
dpse.rc3 3275
dpse.rc3 3310
dpse.rc3 3339
dpse.rc3 3307

all
none

summaries


MulteeSum
Summaries
RMS_50-dmel_r2
RMS_50-dpse_rc3
RMS_50-dpse_rc3
RMS_50-dmel_r2
dmel_r2
dmel_r2
dpse_rc3
dpse_rc3
Embryo Map
EMBRYO: dmel_r2
SELECTED CELL: 3932
METRIC: RMS_50-dpse_rc3
AGGREGATION: min
1.0
0.5
0.0
0.0
0.5
1.0
1.27
1.27
0.37
0.37
AGGREGATION GROUP
CREATED GROUP
Curvemap
0.81
0.71
0.33
0.43
2.48
1.21
metric value
2.5
1.2
eve
fkh
ftz
gt
hb
hkb
kni
Kr
odd
prd
dmel_r2
3932
dpse_rc3
3274
dpse_rc3
3306
dpse_rc3
3308
dpse_rc3
3305
dpse_rc3
3304
dpse_rc3
3272
dpse_rc3
3275
dpse_rc3
3310
dpse_rc3
3339
dpse_rc3
3307
all
remove
1.0
0.0

Characterize differences in gene expression patterns between species.

differences related to:

spatial position

gene expression profiles

complex combination

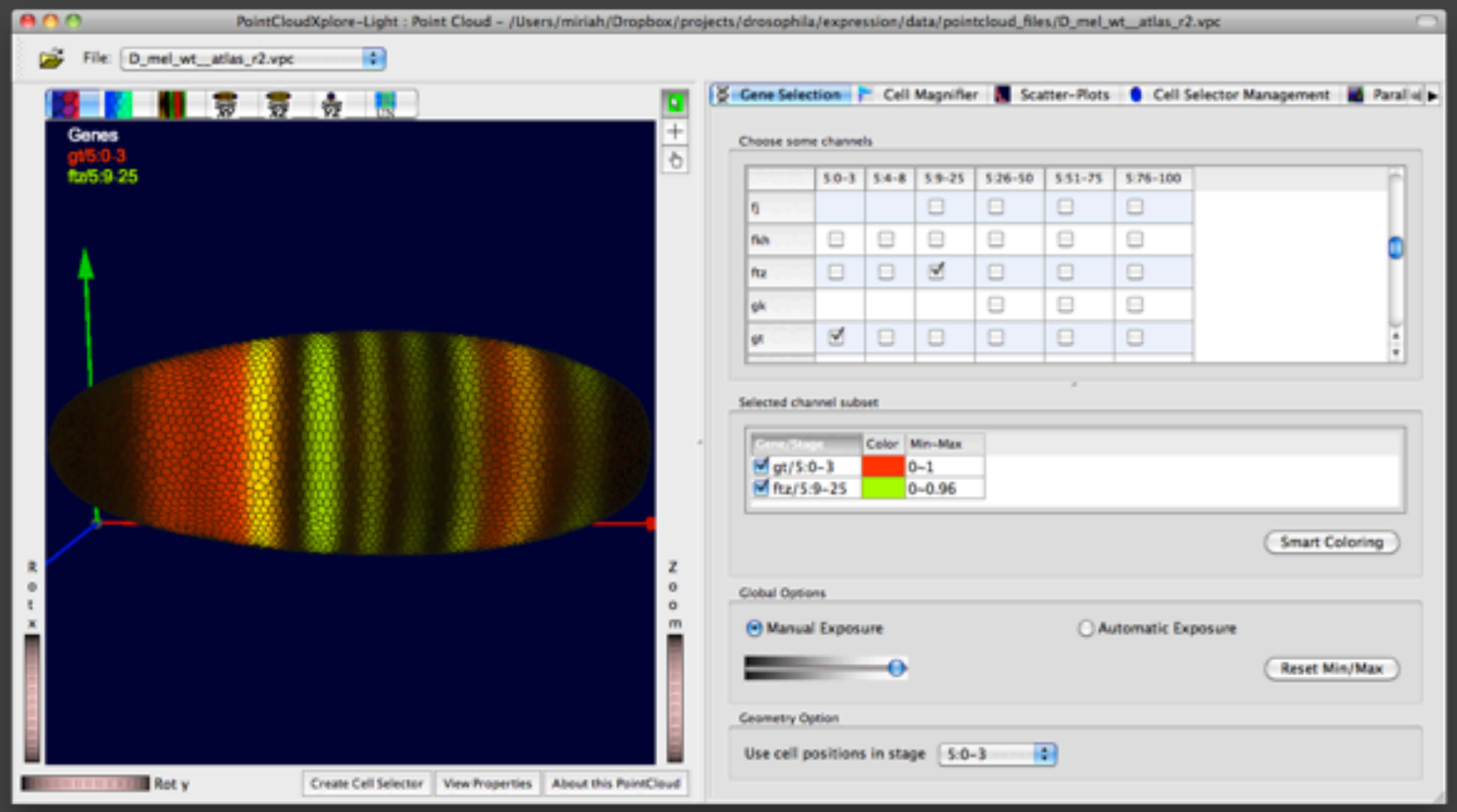
challenging to characterize manually

support mechanisms:

summaries, groups

previous work

PointCloudXplore
[Rübel06][Weber09][Rübel10]



selection:	multiple genes single time point	≠	single gene multiple time points
encodes:	a few genes for all of the cells		all the genes for a few cells
supports:	a single virtual embryo		multiple virtual embryos

data & tool & tasks

summaries & groups

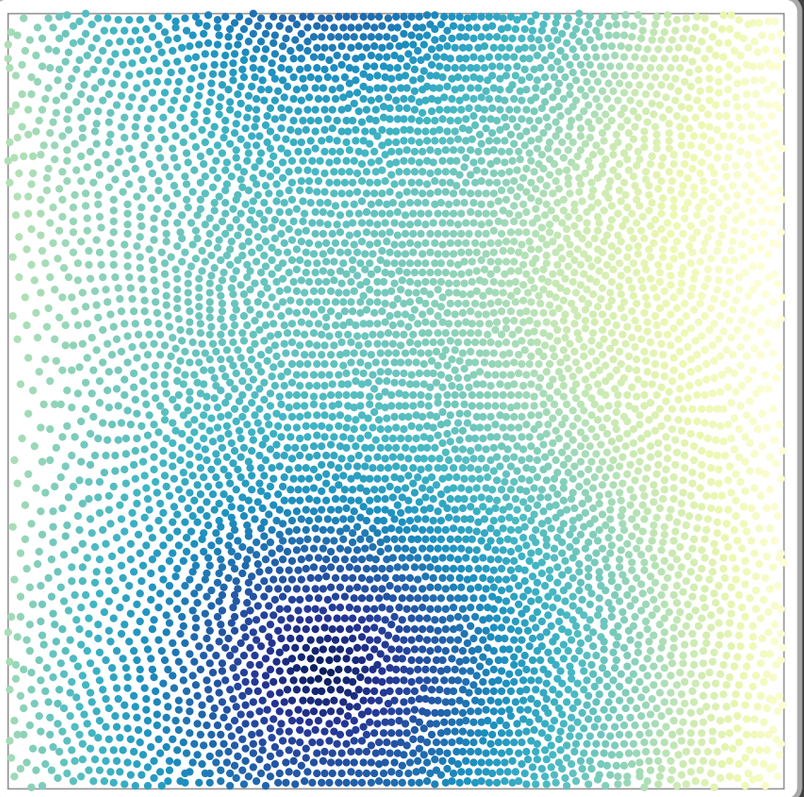
encodings & interaction

conclusions

summary

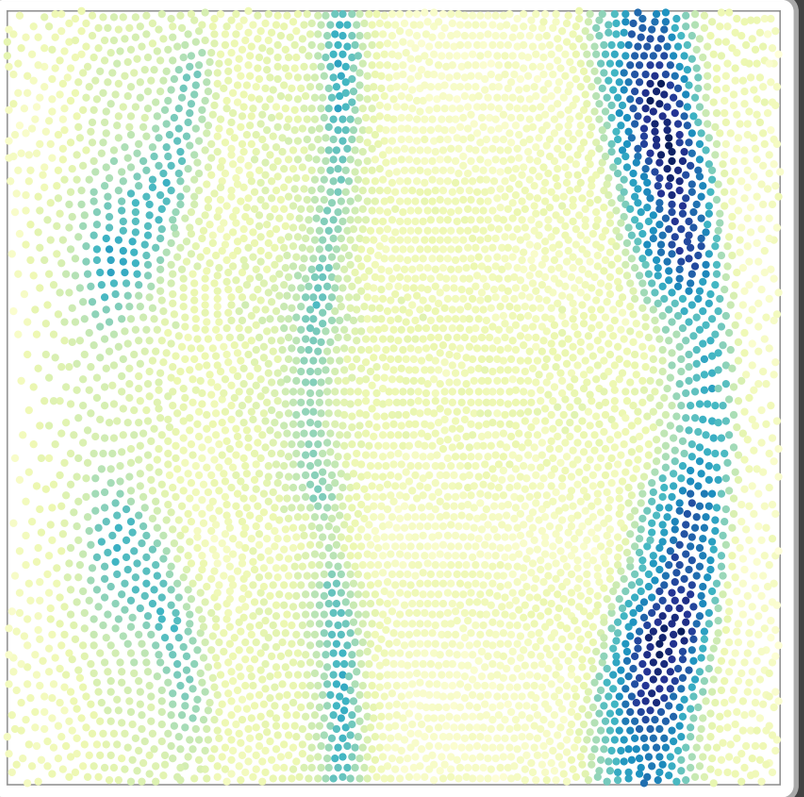
a value for each cell that summarizes the underlying data

spatial



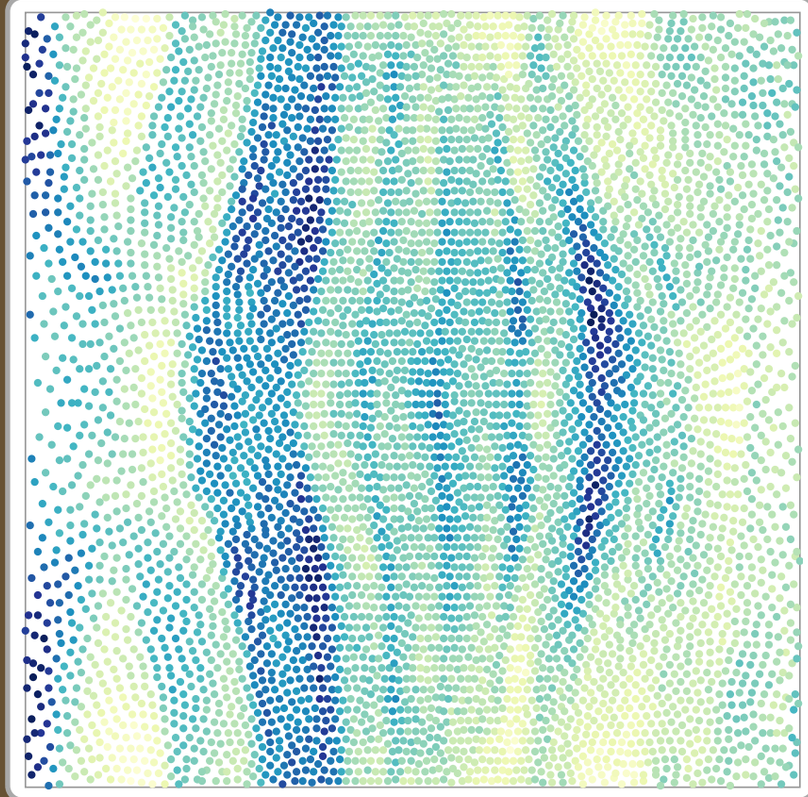
distance to cell *i*

expression



gene *j* at timepoint *k*

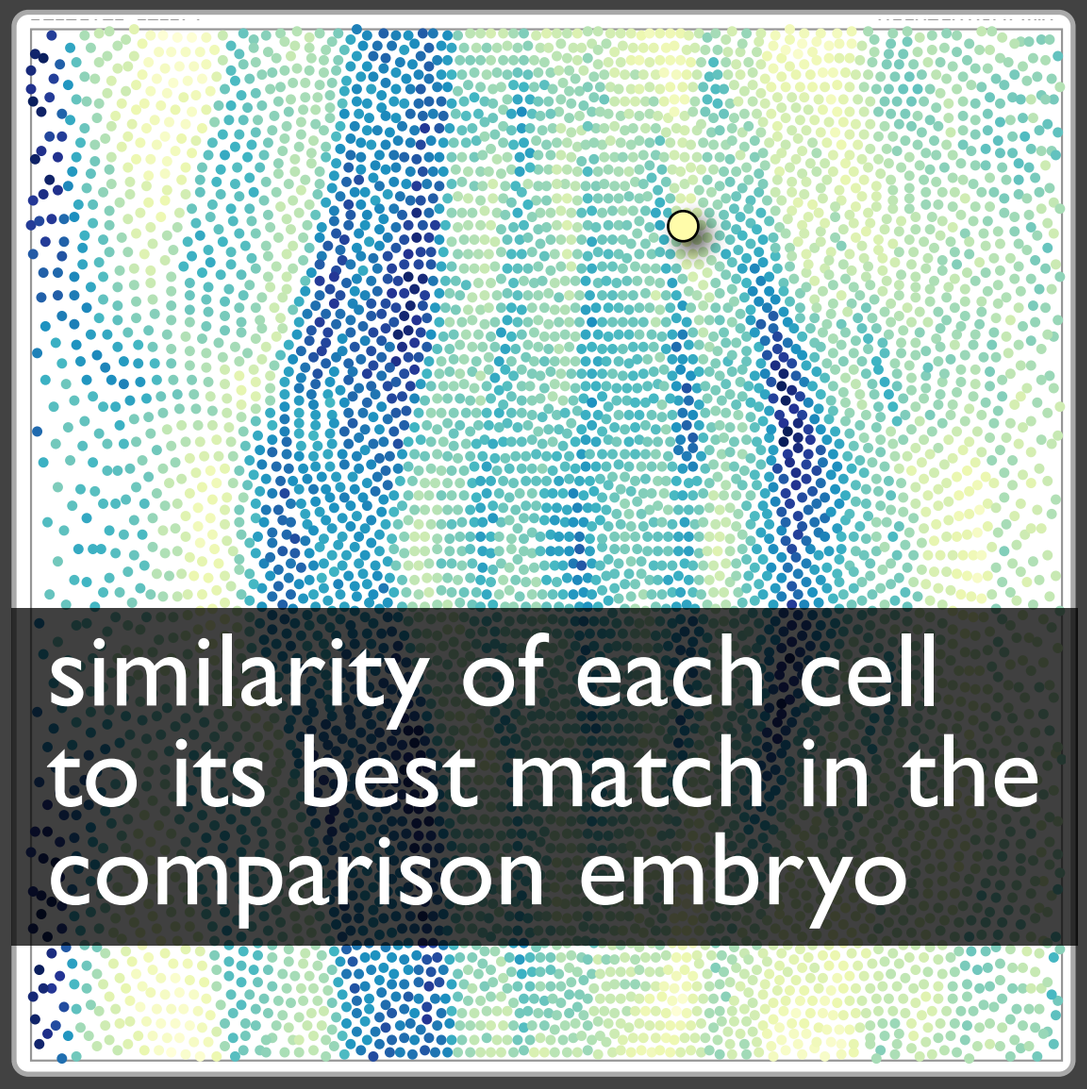
combination



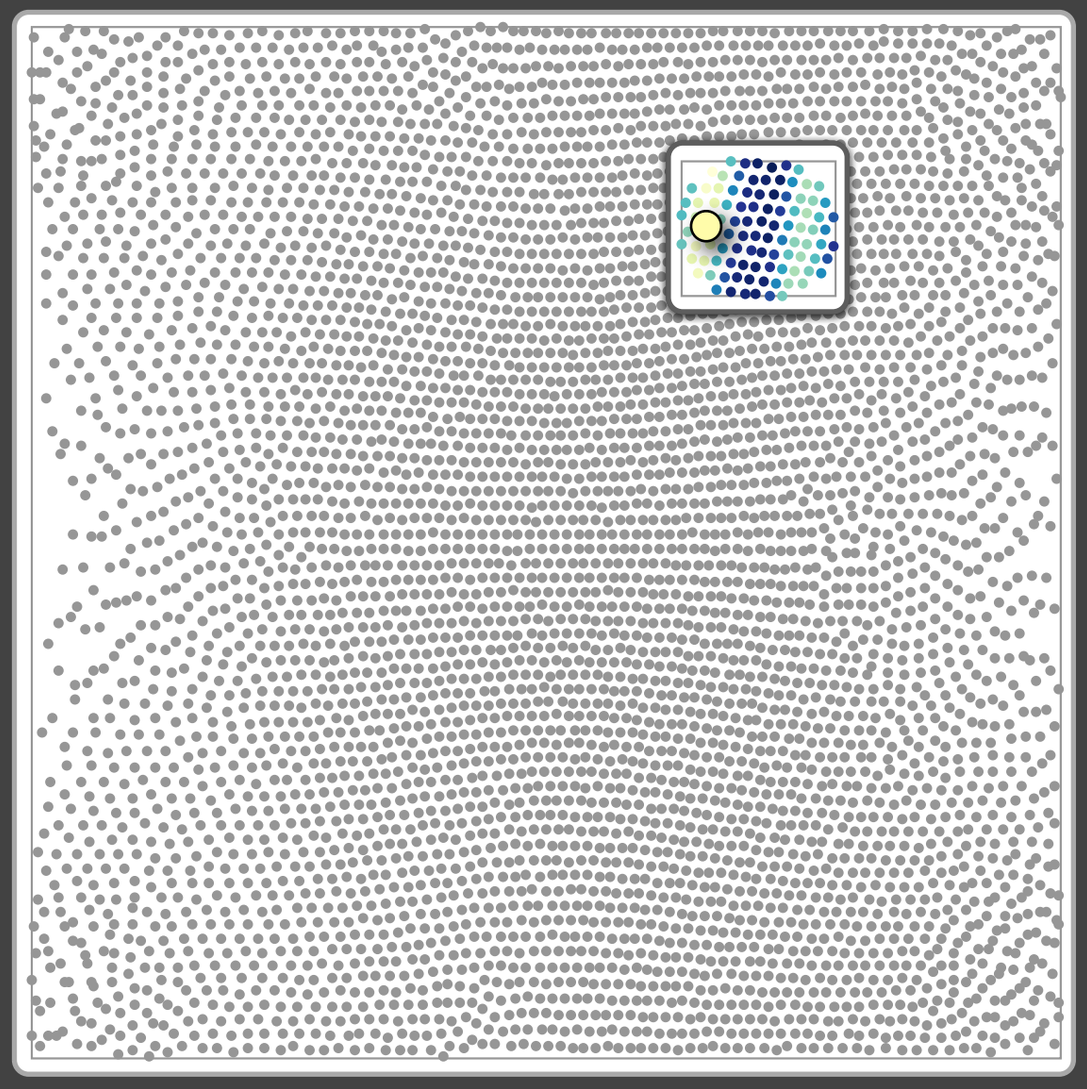
compare embryos *A* and *B*

creating a comparison summary

root-mean-square distance



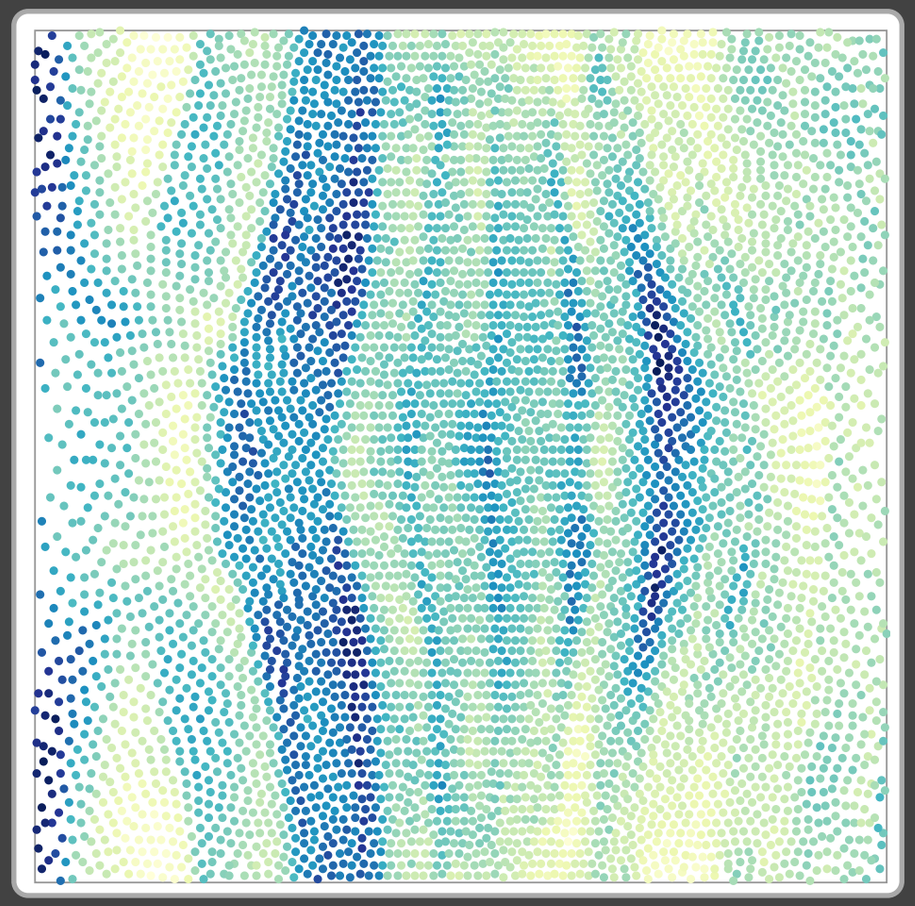
similarity of each cell to its best match in the comparison embryo



embryo A

embryo B

comparative summary components

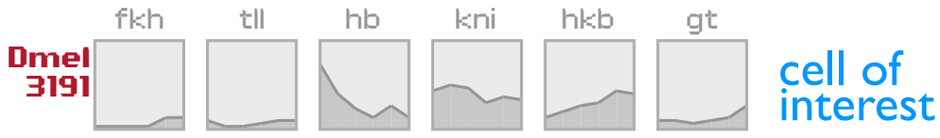


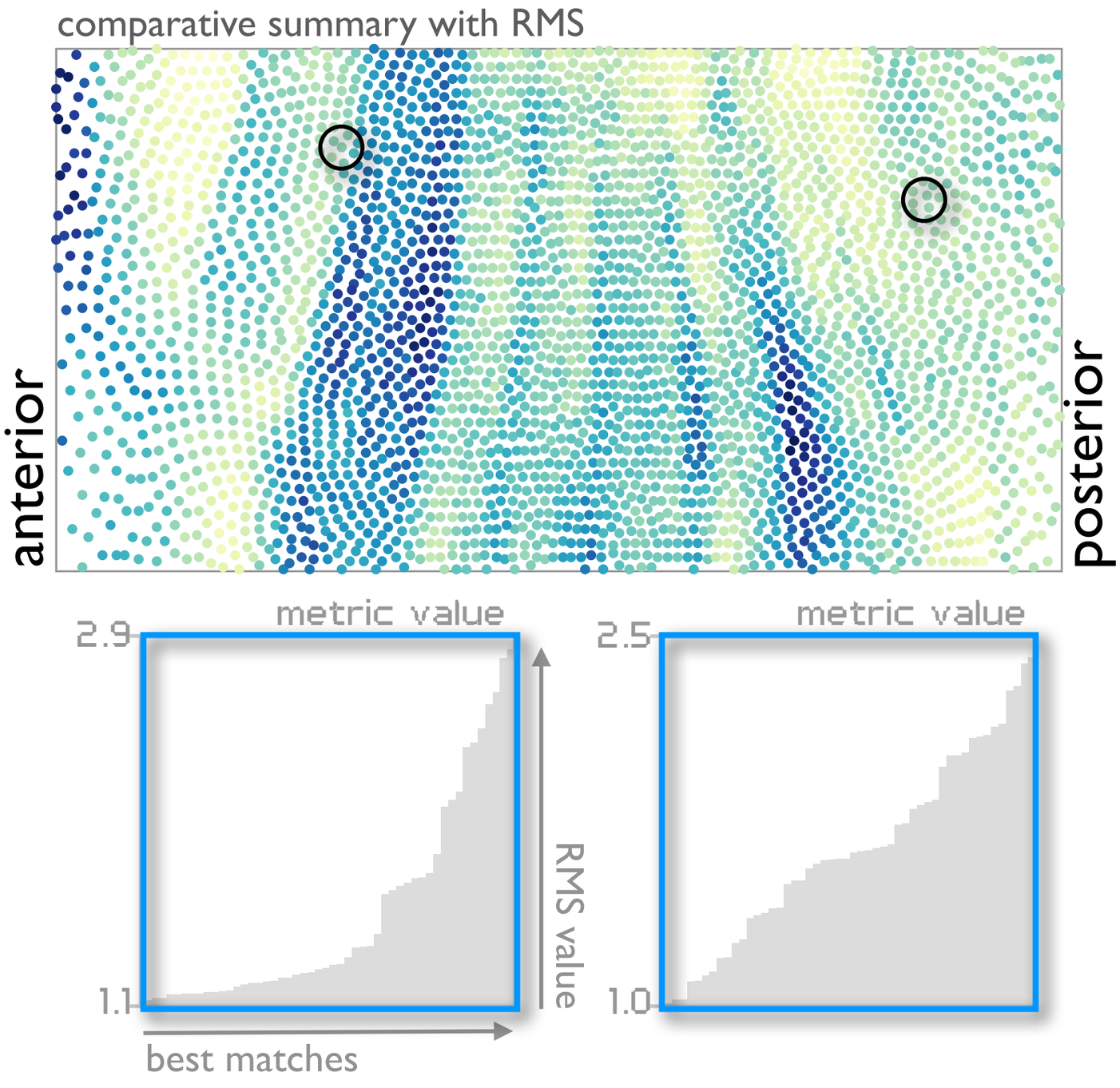
aggregation group

metric

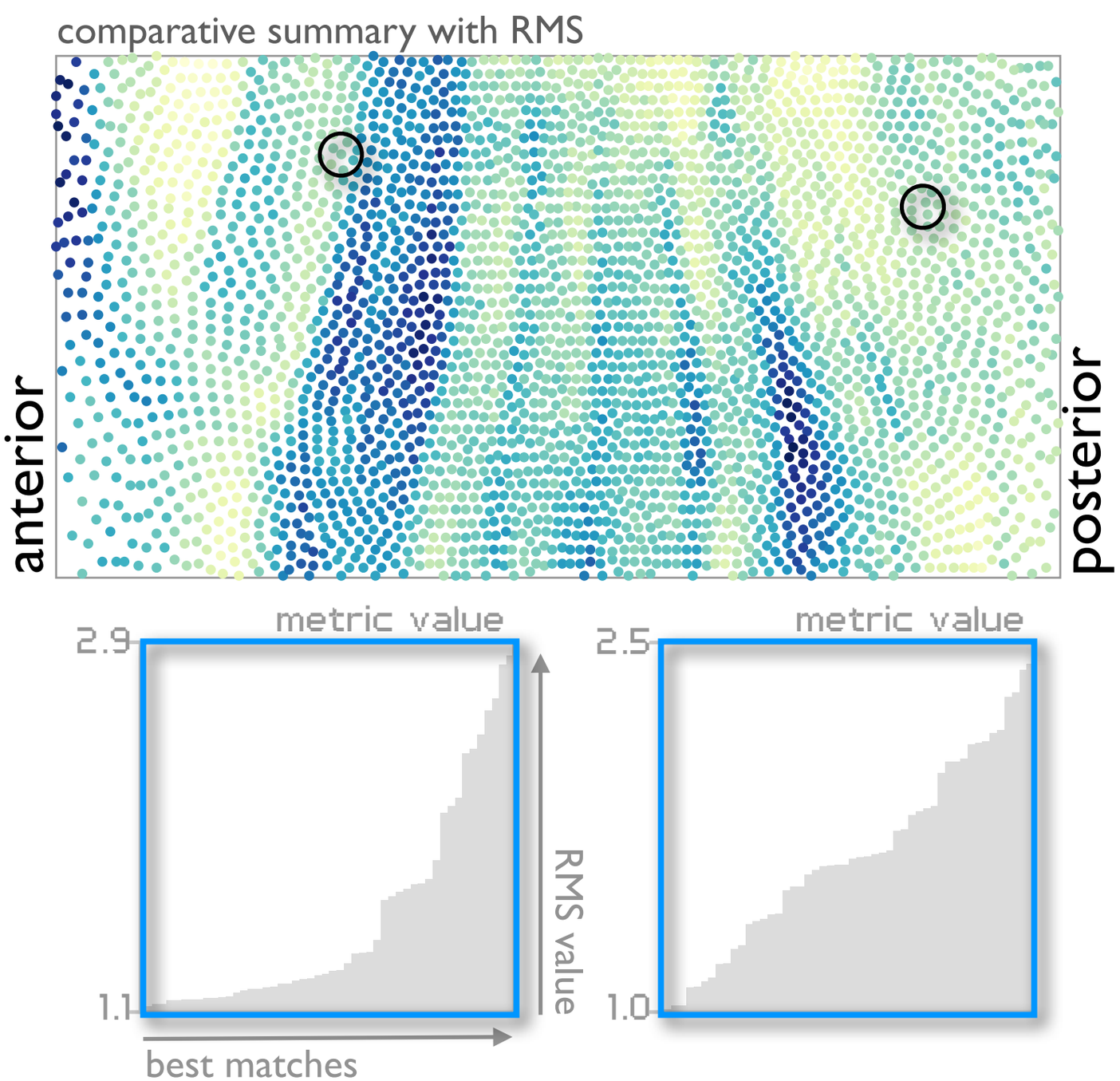
aggregation

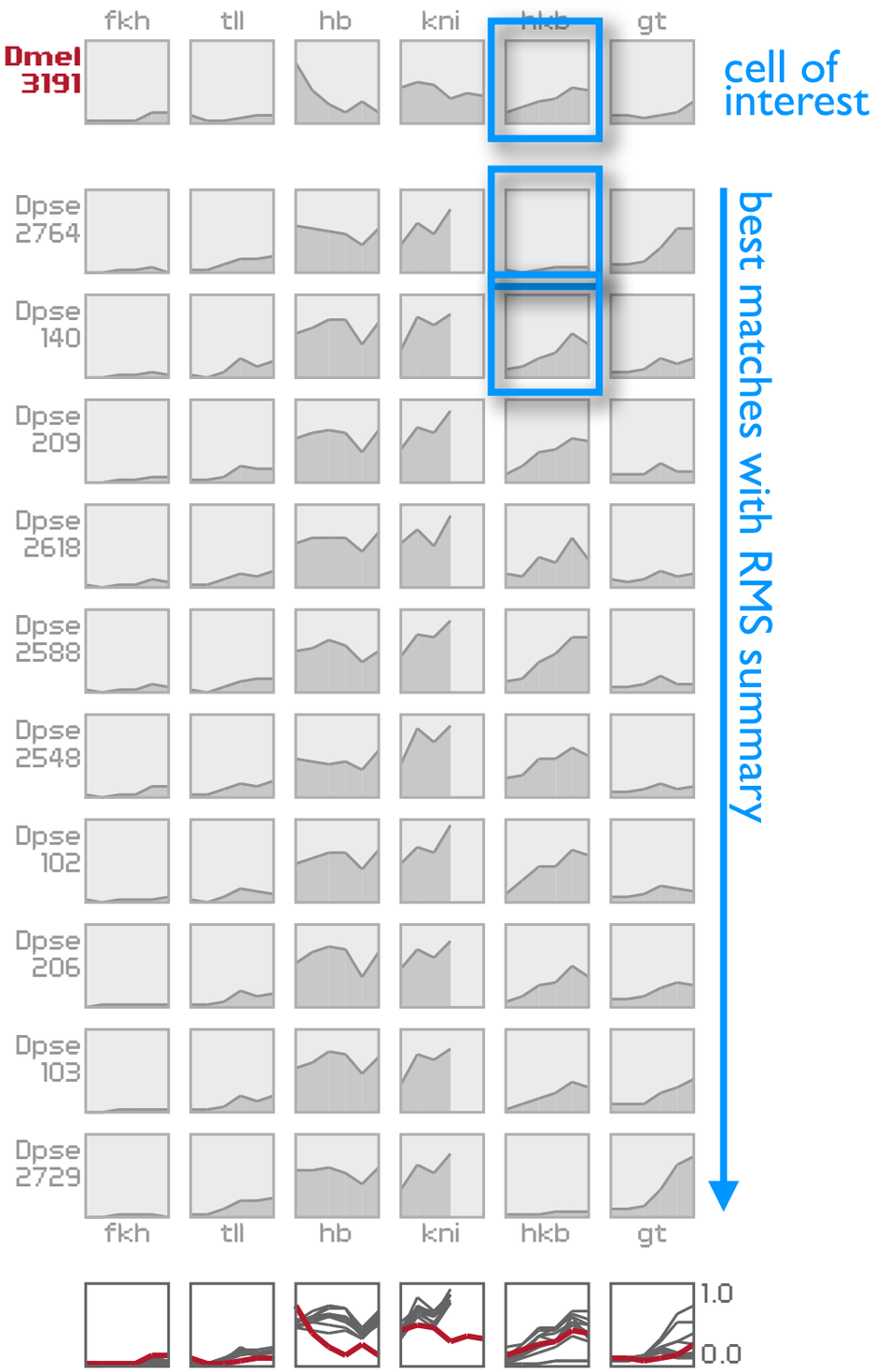
case study one:
characterizing a summary





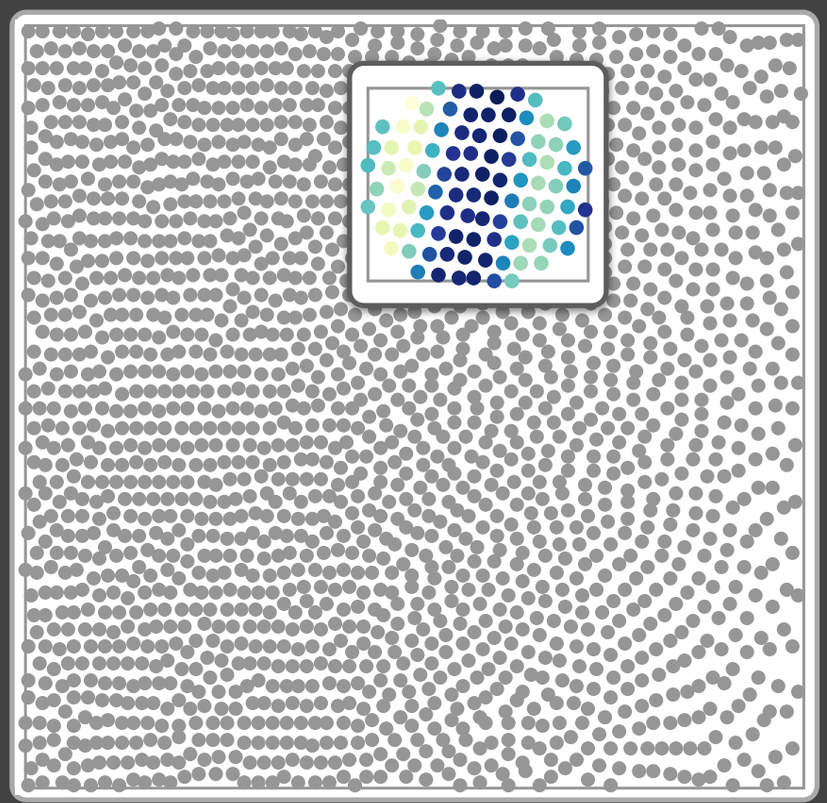
case study one:
characterizing a summary



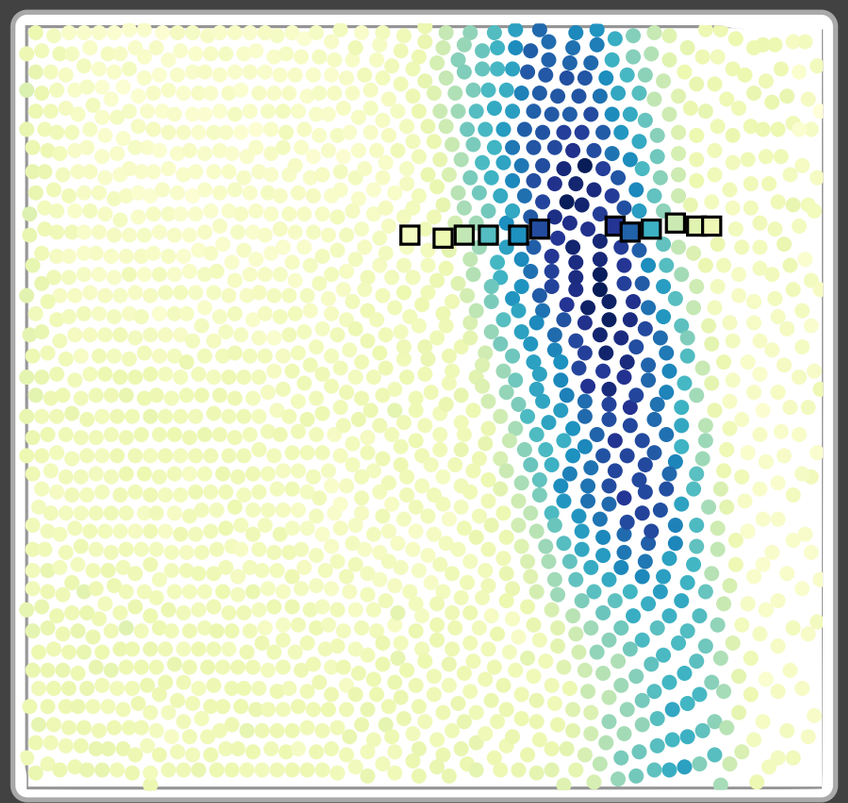


group
a set of cells

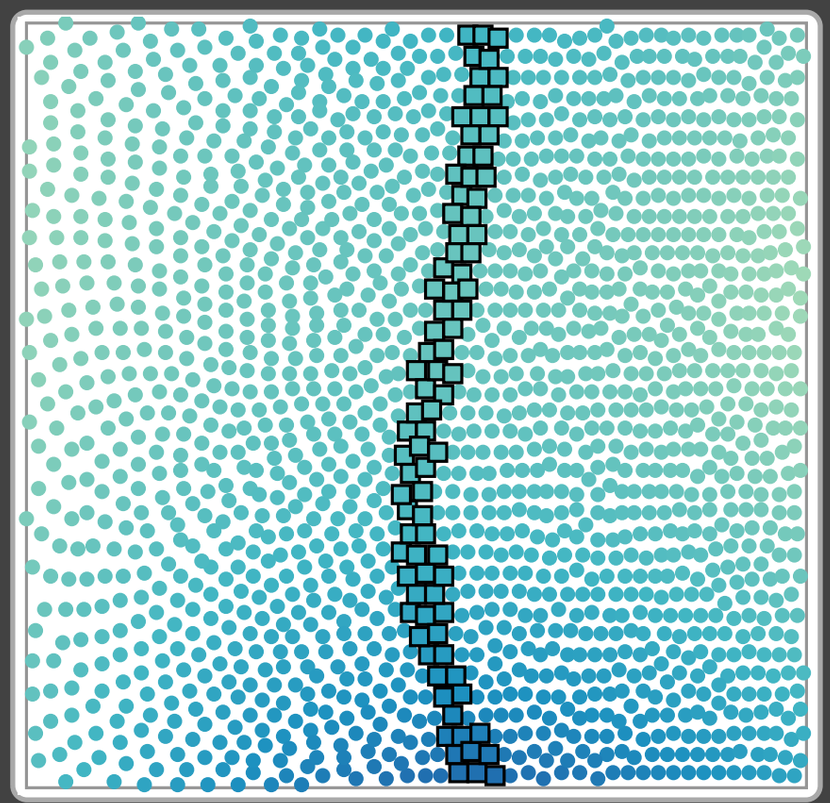
aggregation group



created group

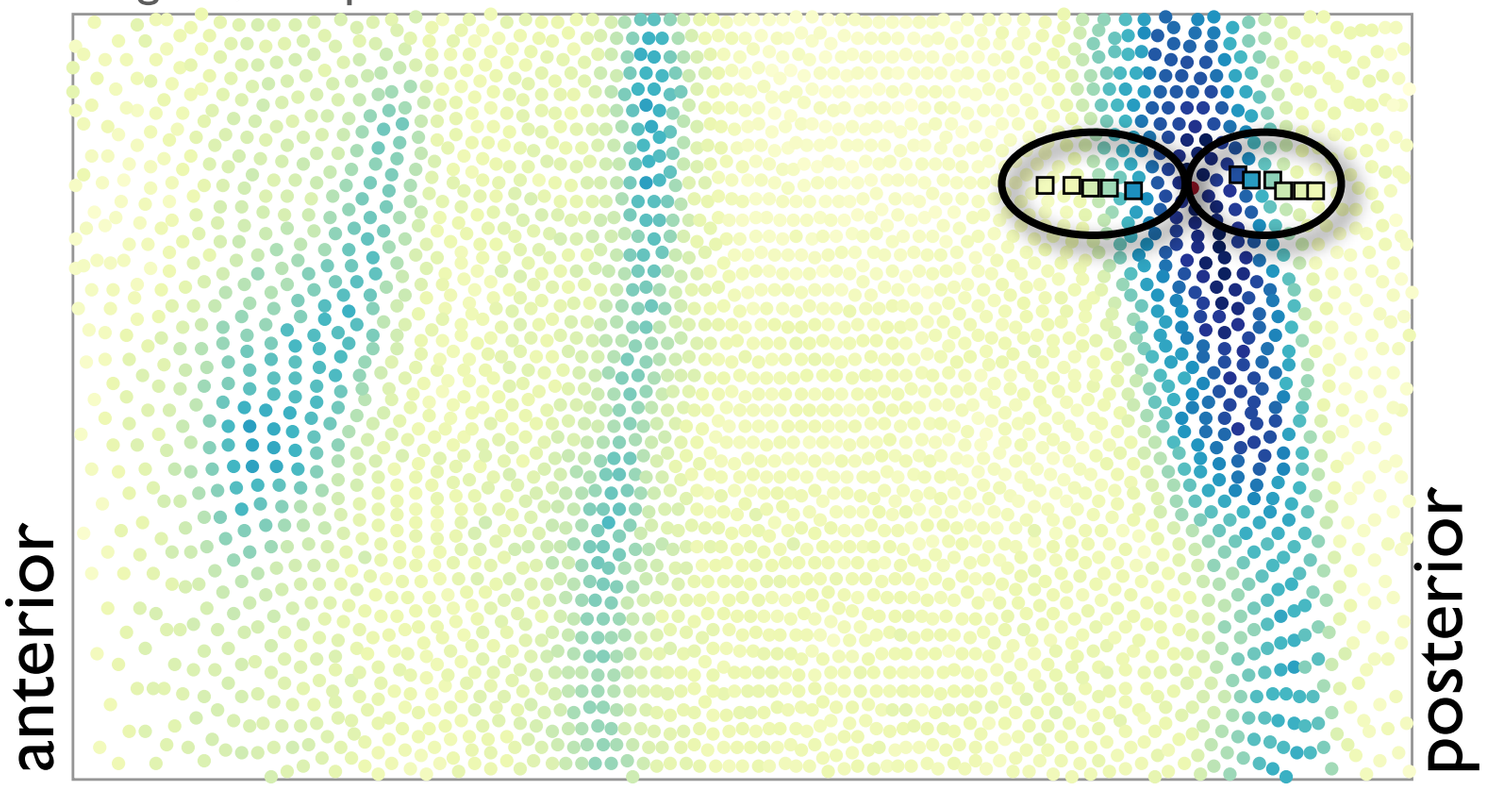


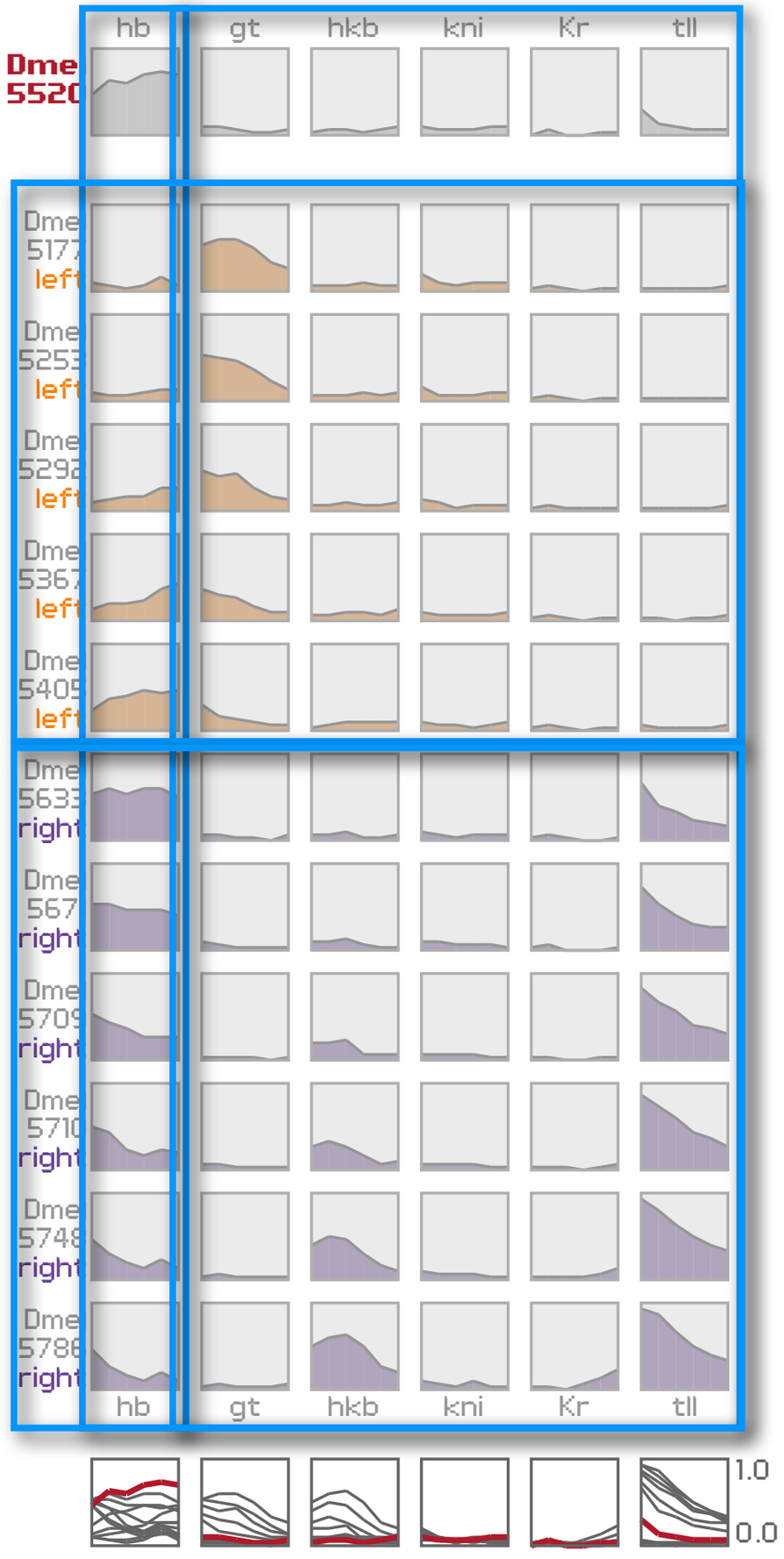
existing group



case study two:
exploring groups

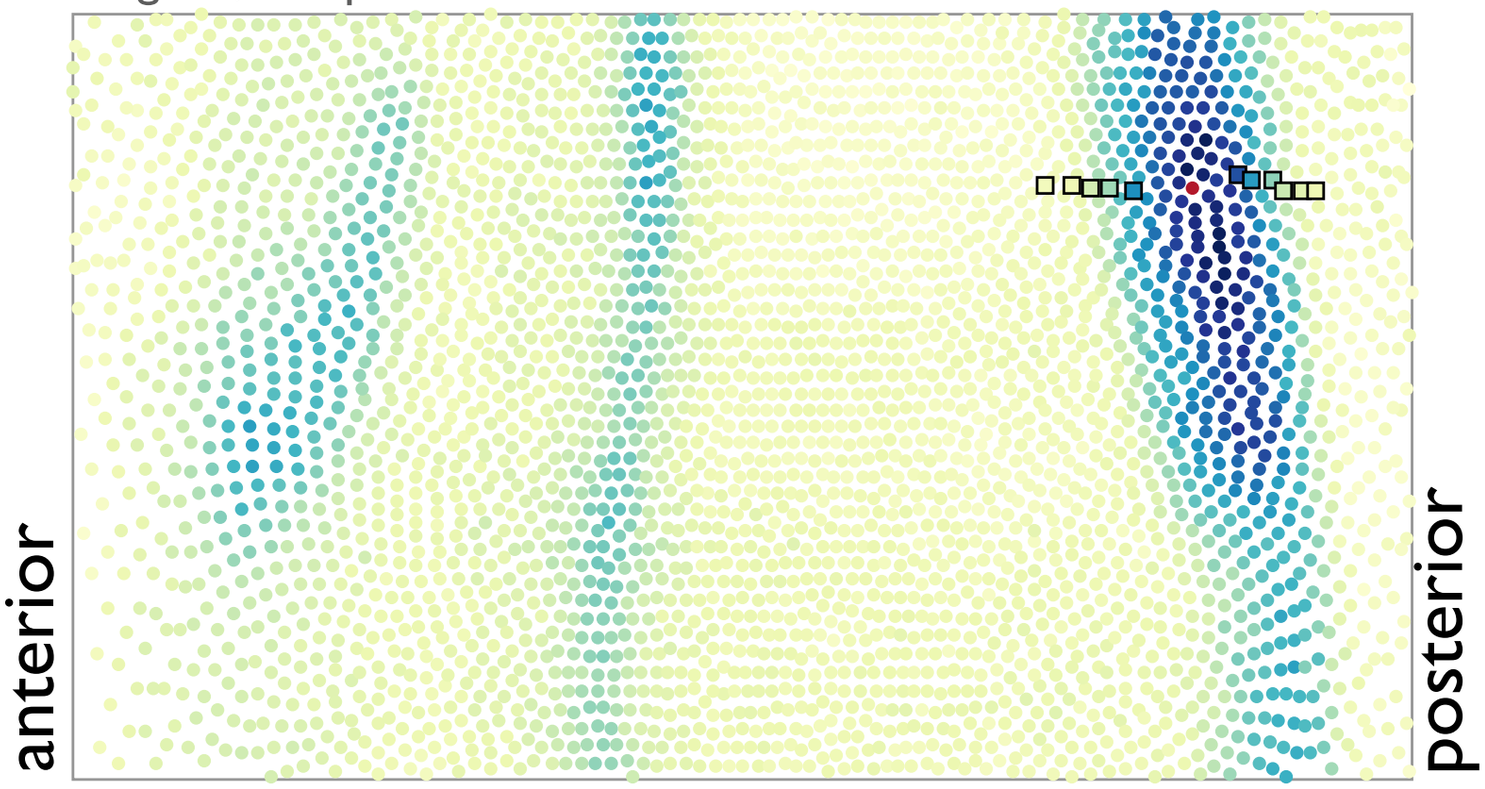
hb gene, timepoint 4

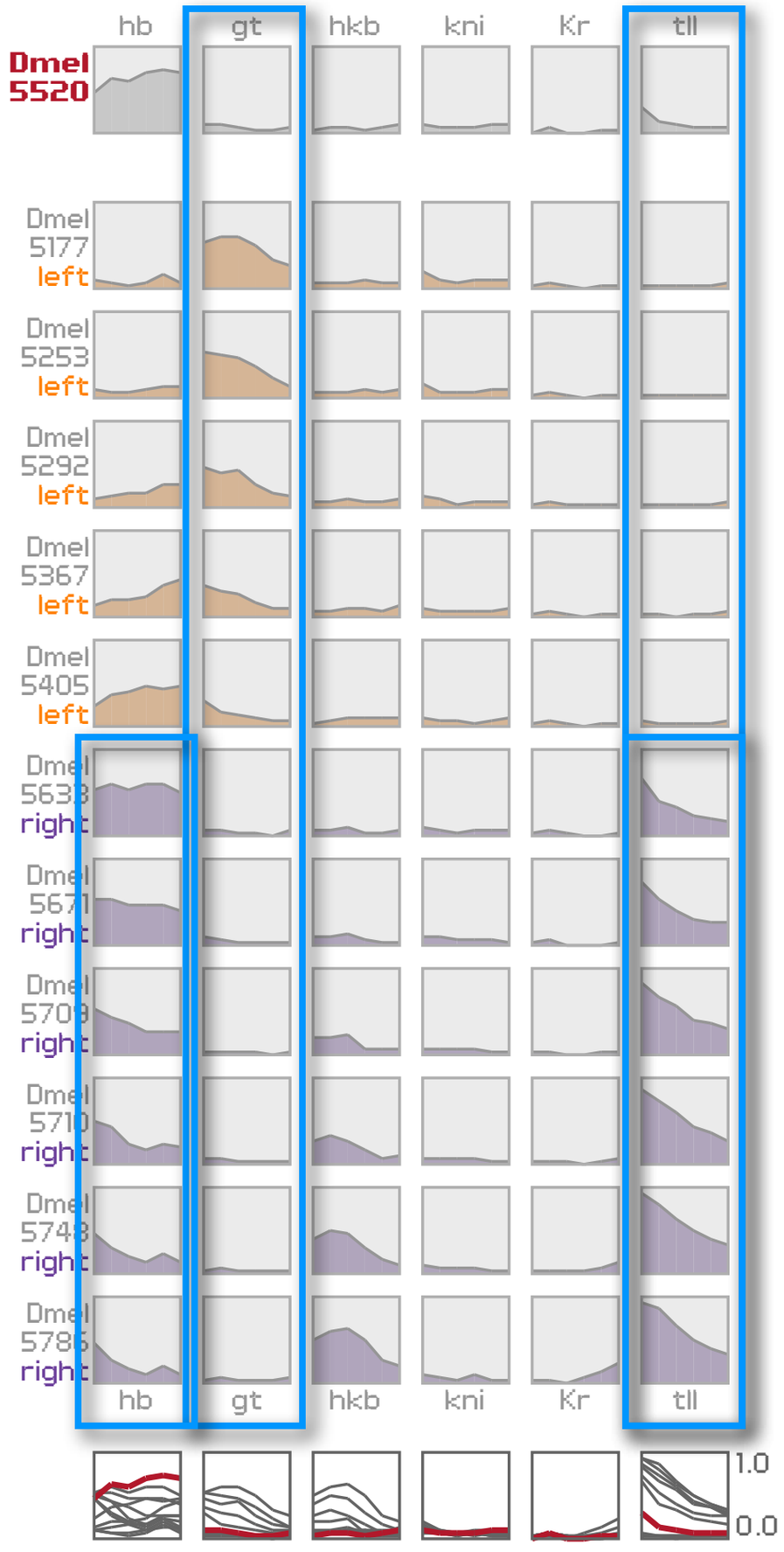




case study two:
exploring groups

hb gene, timepoint 4



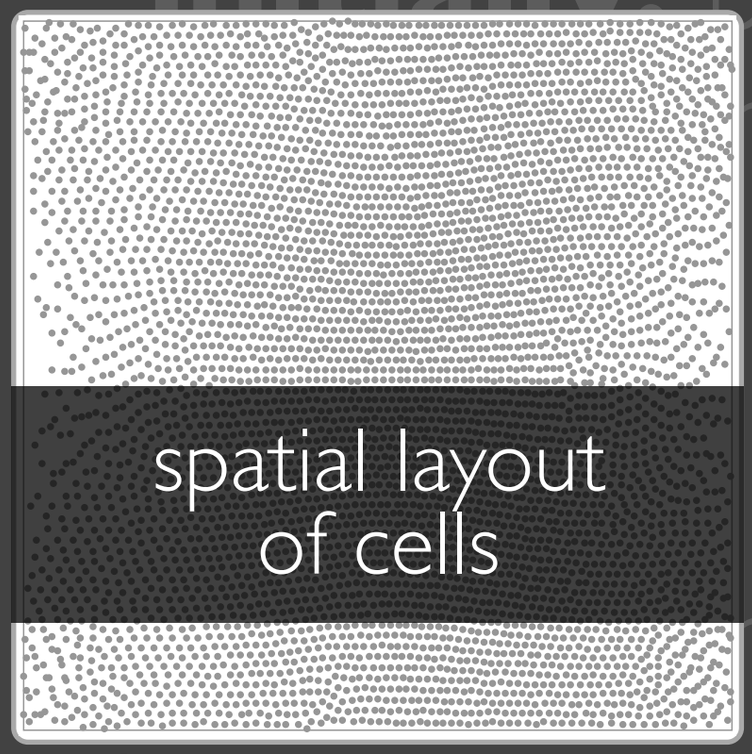


data & tool & tasks

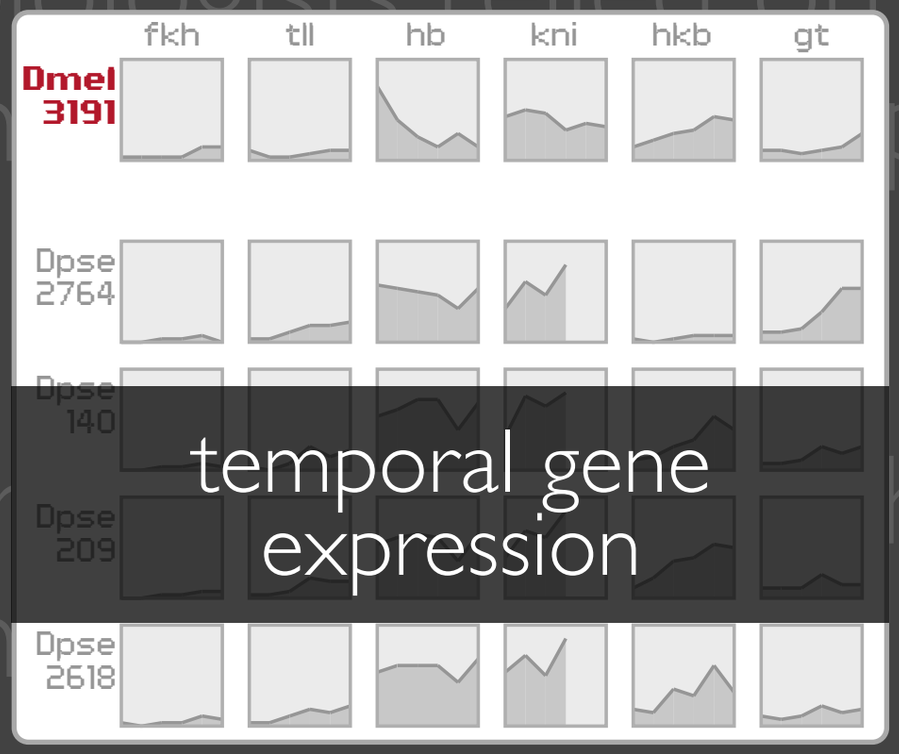
summaries & groups

encodings & interaction

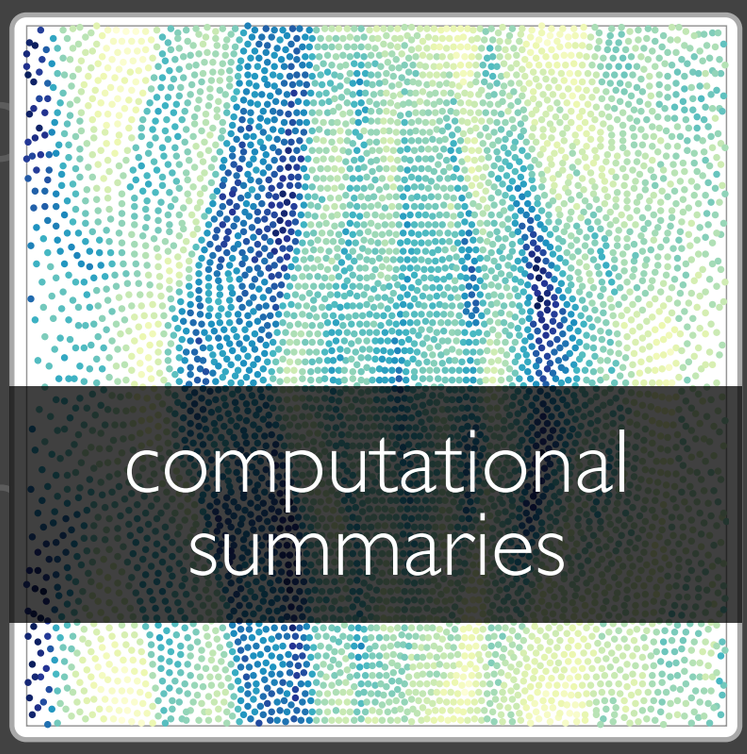
conclusions



spatial layout of cells



temporal gene expression



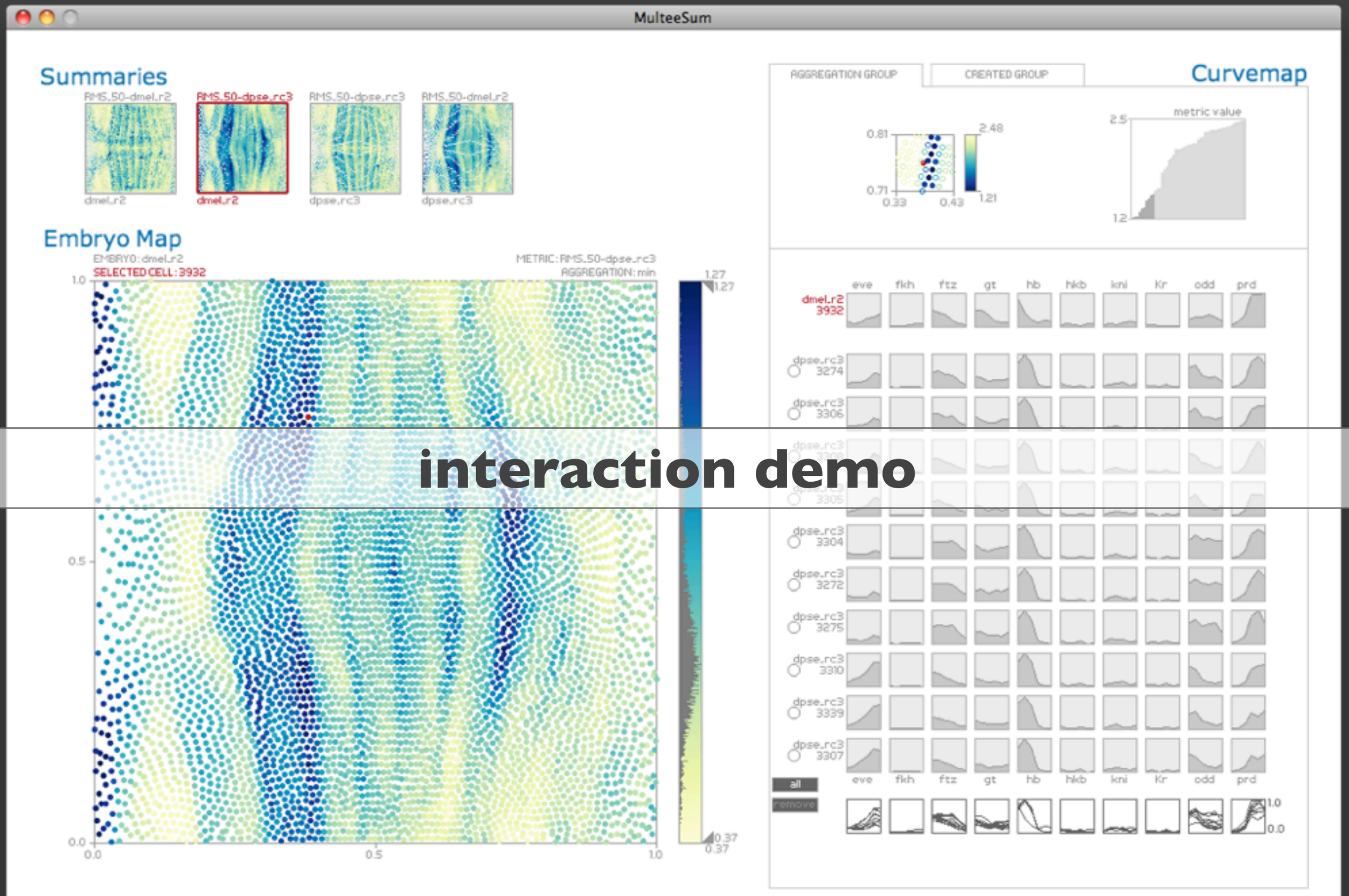
computational summaries

embryo map

curvemap [Meyer10]

colormap

visualize: triad of data



interaction demo

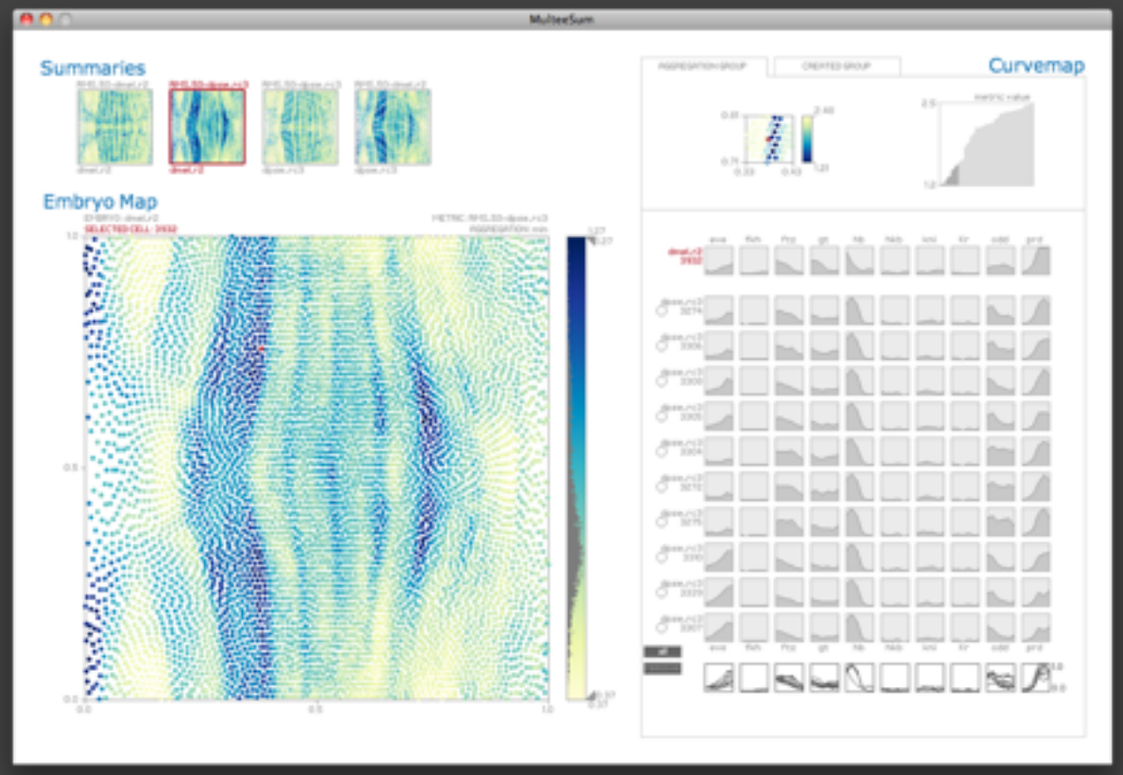
data & tool & tasks

summaries & groups

encodings & interaction

conclusions

contributions



MulteeSum

spatial and temporal gene expression data from multiple species

workflow

visualization supports upstream computation via summaries

validation

case studies, deployment

Cellular resolution comparison of gene expression in Drosophila reveals coordinated shifts in the segmentation network.

DePace et. al, in preparation.

future work

design is broadly applicable

- consider: scalability of curvemap
- consider: mapping of other spatial data to 2D

continuing to support collaborators

- additional features in MulteeSum
- new summary designs

multeesum.org

questions?

acknowledgements

Members of the DePace lab: Angela DePace, Kelly Eckenrode, Tara Martin, Sarah Saminadin-Peter, Max Stallar, Ian Sudbery, Zeba Wunderlich

Charless Fowlkes

The National Science Foundation under Grant 0937060 to the Computing Research Association for the CIFellows Project, the Helen Hay Whitney postdoctoral fellowship, the Armenise Foundation of Harvard Medical School, and the Natural Sciences and Engineering Research Council of Canada.