**MulteeSum**

A Tool for Comparative Spatial and Temporal Gene Expression Data

Miriah Meyer, Tamara Munzner, Angela DePace, Hanspeter Pfister

1. **grand challenge in biology:**
   - understand genomic source of gene expression dynamics

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

3. **correlate gene expression with spatial location.**

4. **data & tool & tasks**
   - summaries & groups
   - encodings & interaction
   - conclusions

5. **virtual embryo**
   - several thousand cells (5,000 ± 1,000)
   - each cell has:
     - **expression profile**
       - 6 time points × 50 genes
   - spatial position
     - 3D and 2D coordinates

6. **12 related species**
   - one complete three preliminary

7. **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

8. **tool deployed**
   - currently used several times a week

9. **embryo map**

10. **expression**

11. **genes**
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

- selection:
  - multiple genes
  - single time point
  - complex combination
- encodes:
  - single gene
  - multiple time points
  - all the genes for a few cells
  - multiple virtual embryos

comparative summary components

- aggregation group
- metric
- aggregation

case study one:
characterizing a summary

- summaries & groups
- encodings & interaction
- conclusions

data & tool & tasks

- summaries & groups
- encodings & interaction
- conclusions

case study one:
characterizing a summary

- grouping
- exploring groups

data & tool & tasks

- summaries & groups
- encodings & interaction
- conclusions

visualize: triad of data
**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

**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

**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 Armenian Foundation of Harvard Medical School, and the Natural Sciences and Engineering Research Council of Canada.

**questions?**

**multeesum.org**