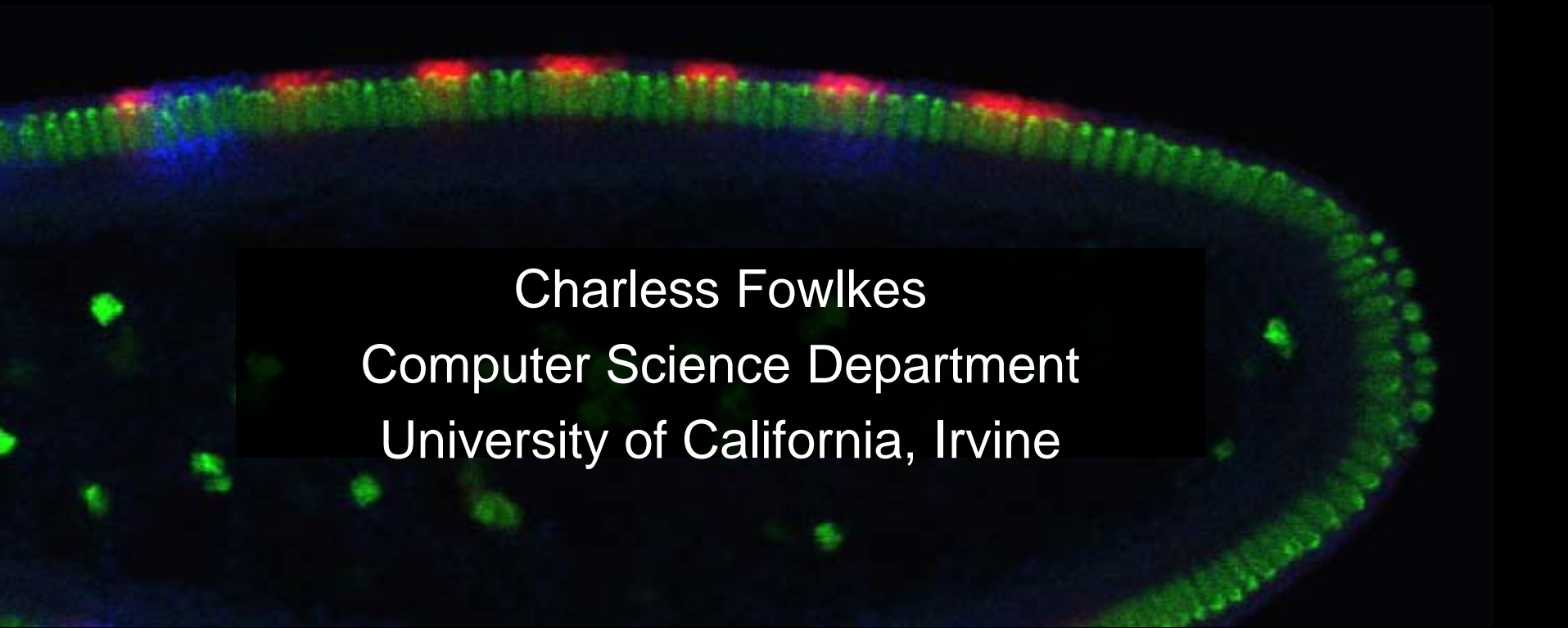
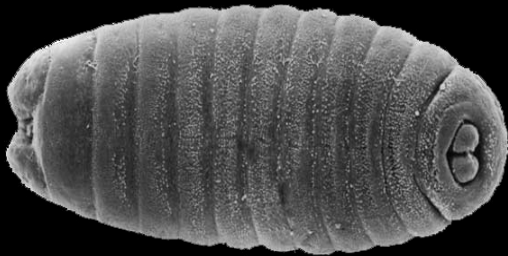
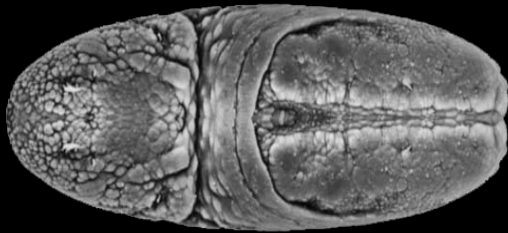
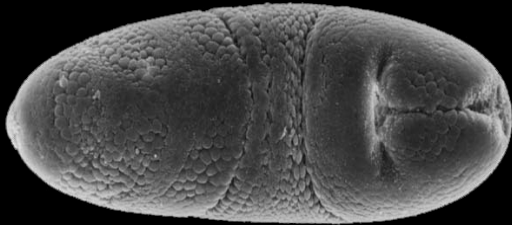
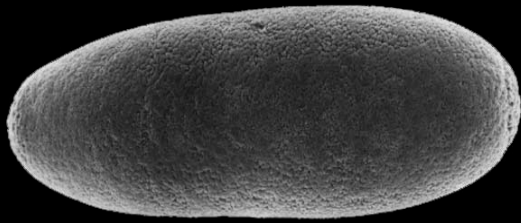


Characterizing variation in embryonic patterning among *Drosophila* species using a spatio-temporal atlas of gene expression



Charless Fowlkes
Computer Science Department
University of California, Irvine

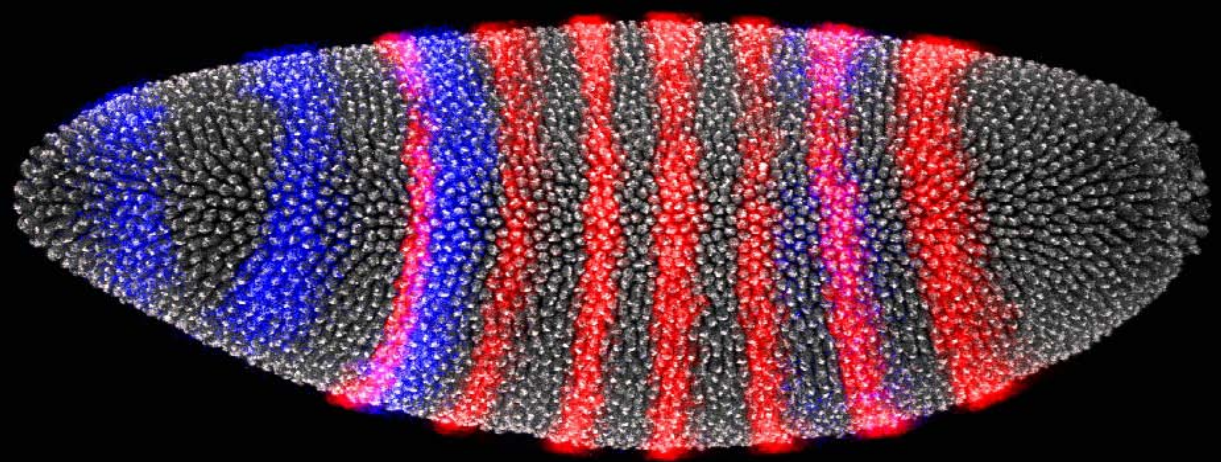
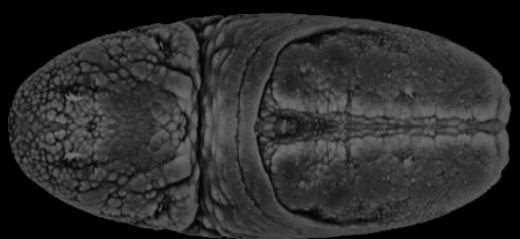
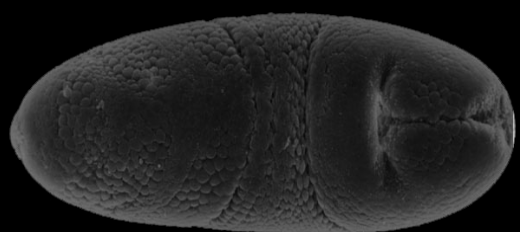
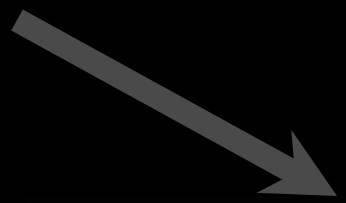
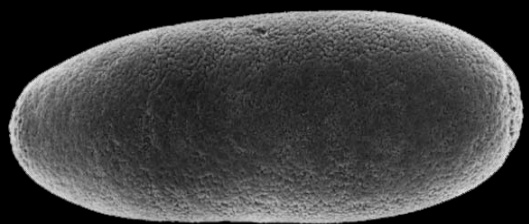
How can we relate variations
in biological form to
underlying molecular events?

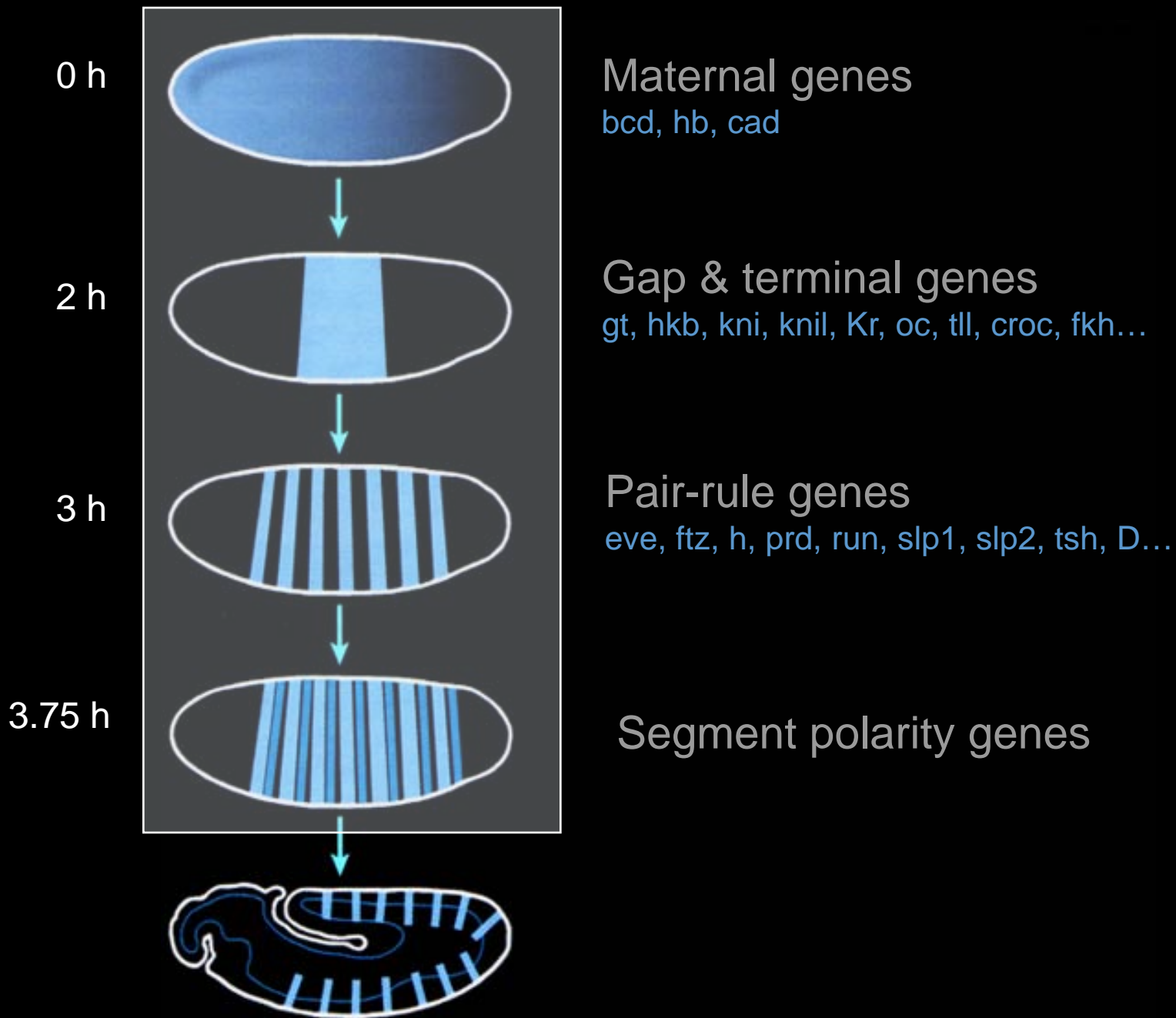


D. melanogaster



D. pseudoobscura





Maternal genes
bcd, *hb*, *cad*

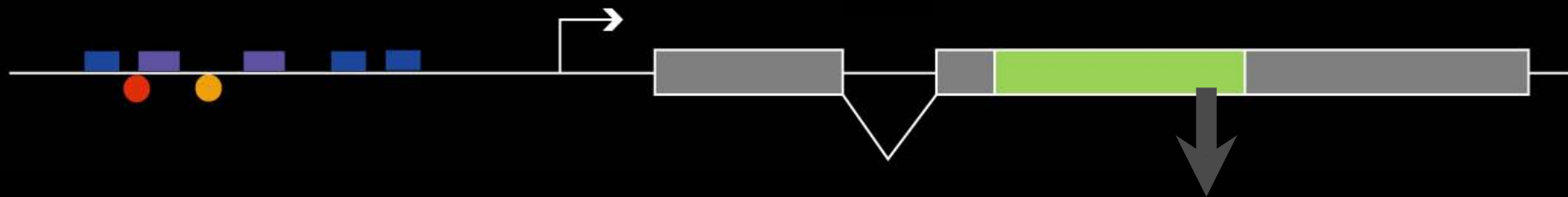
Gap & terminal genes
gt, *hkb*, *kni*, *knil*, *Kr*, *oc*, *tll*, *croc*, *fkf*...

Pair-rule genes
eve, *ftz*, *h*, *prd*, *run*, *slp1*, *slp2*, *tsh*, *D*...

Segment polarity genes

regulatory sequence

coding sequence



Gene product



Reaction Diffusion Model

$$\begin{aligned}dX/dt &= F(Y) - AX + B\Delta X \\dY/dt &= CX - DY + E\Delta Y\end{aligned}$$

$X_i(s,t)$ = concentration of mRNA for gene i at location s and time t .

$Y_i(s,t)$ = concentration of protein for gene i

F(Y) – transcription rate as a function of protein concentration (and DNA sequence).

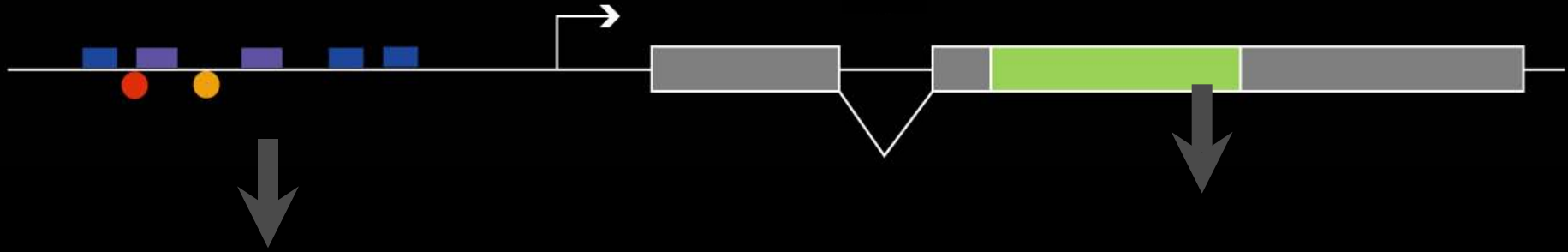
C - translation rates

A,D - decay rates

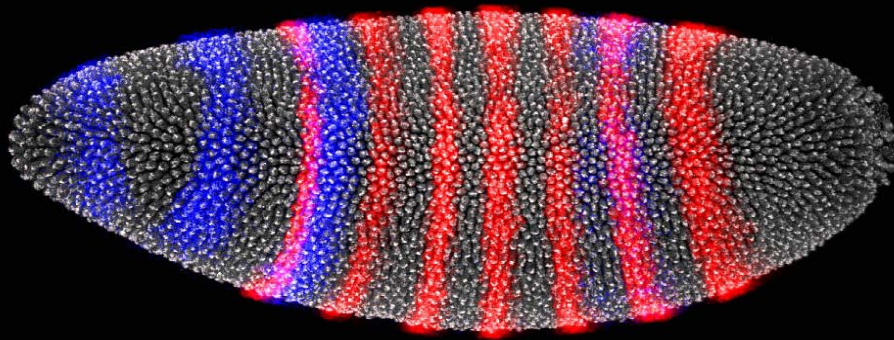
B,E - diffusion rates

regulatory sequence

coding sequence



Expression pattern

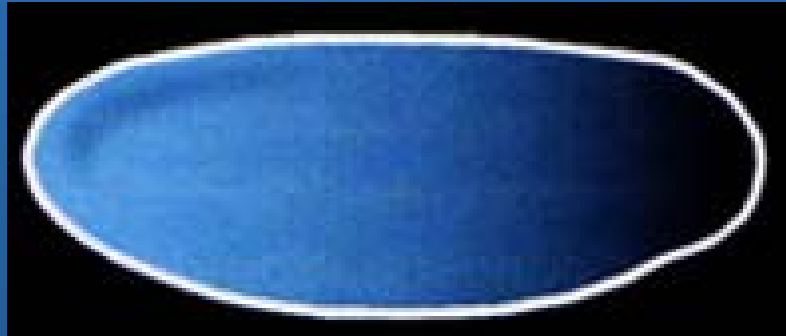


Gene product



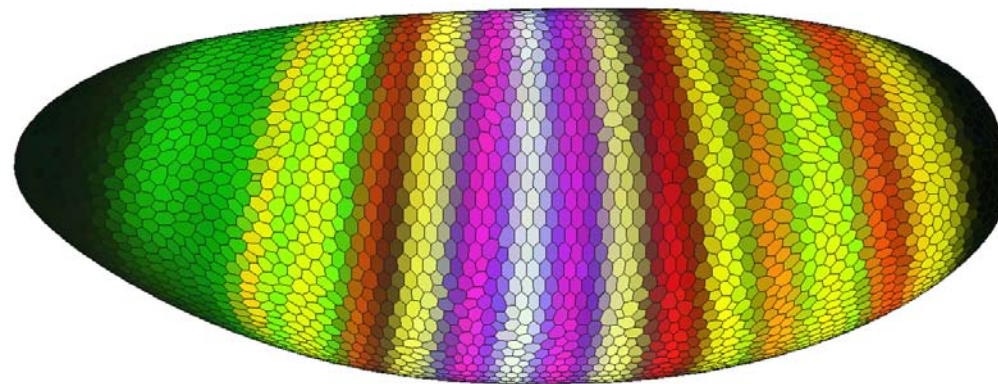
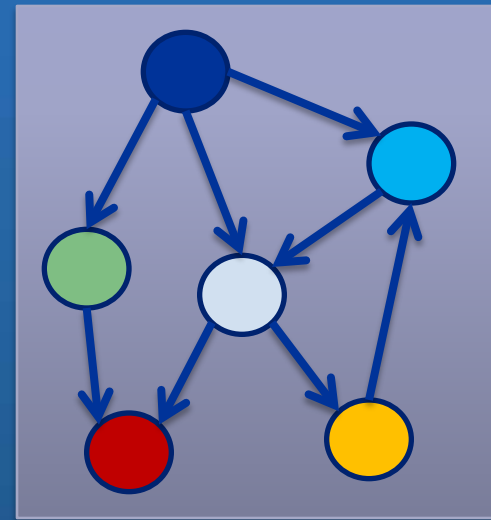
Goal: Understand general rules by which patterning is encoded in the cis-regulatory sequence in the genome.

Initial conditions (egg geometry, maternally established gradients)

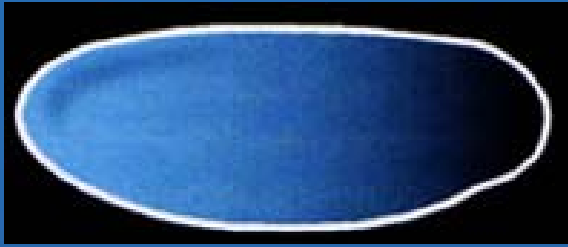


+

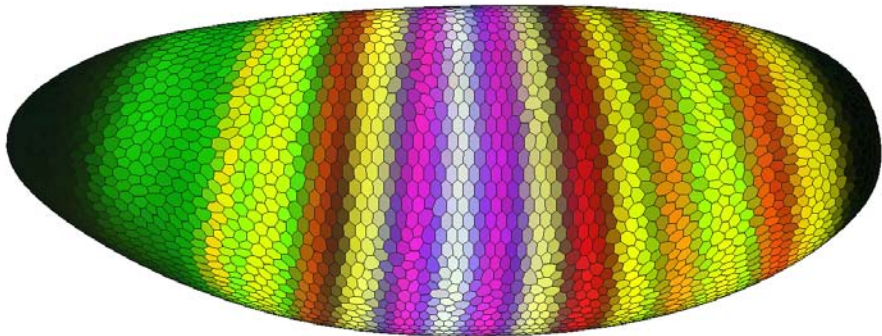
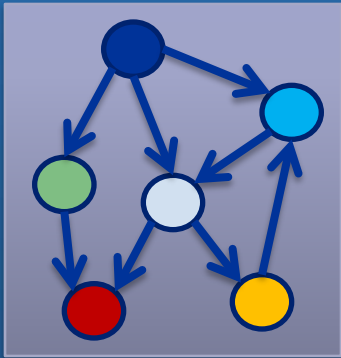
Zygotic regulatory network



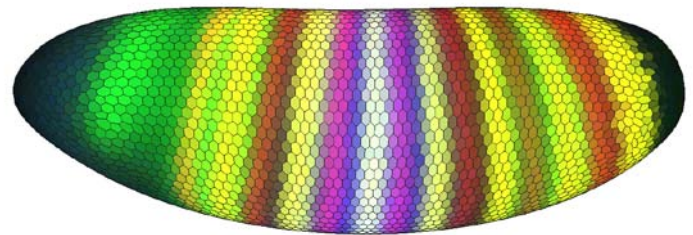
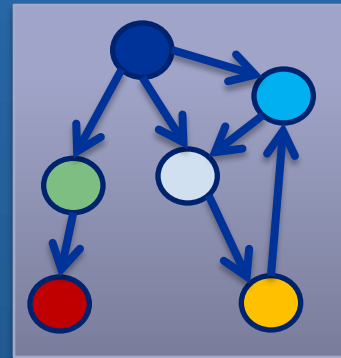
Patterned blastoderm (animal body plan)

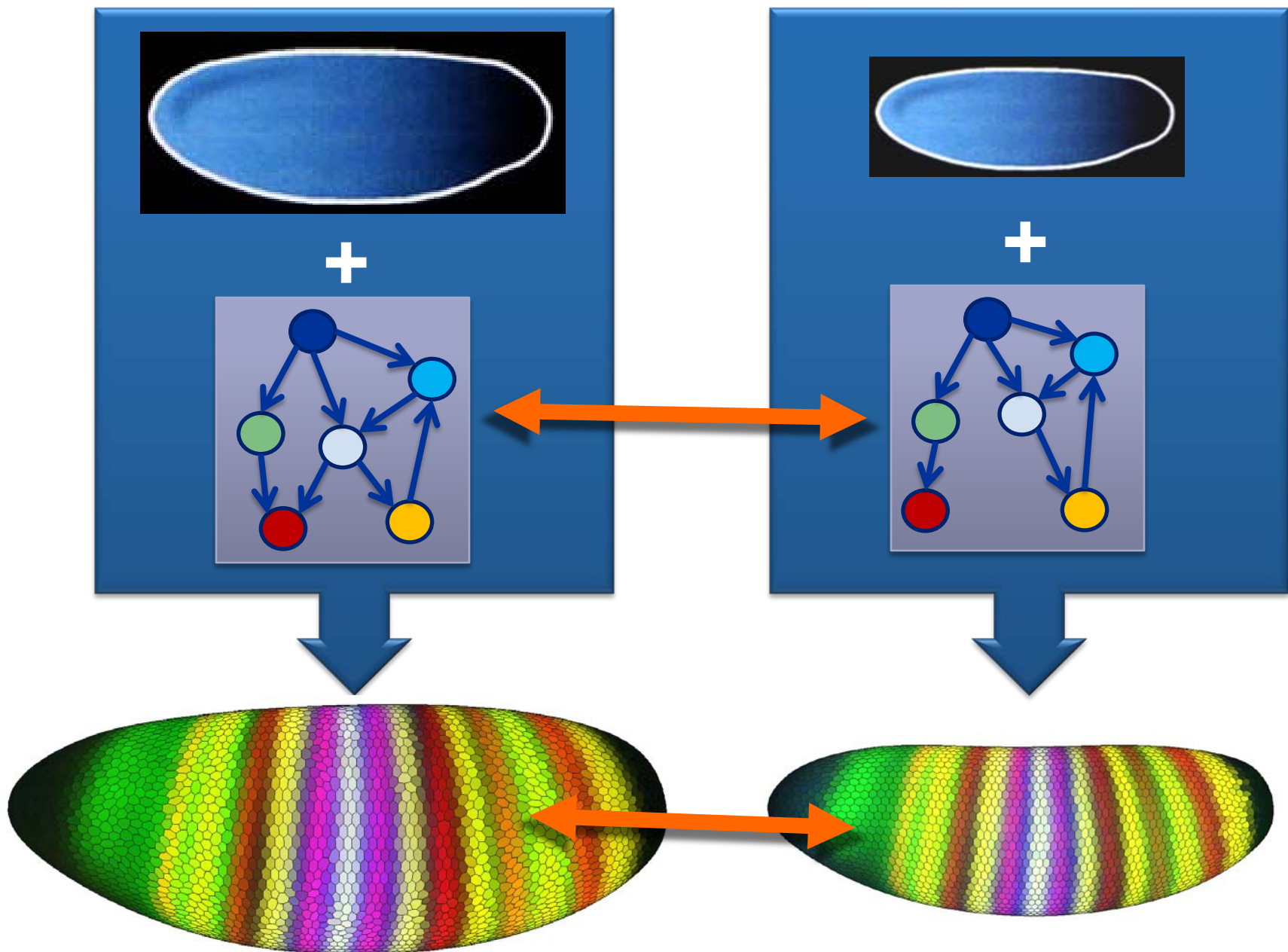


+



+





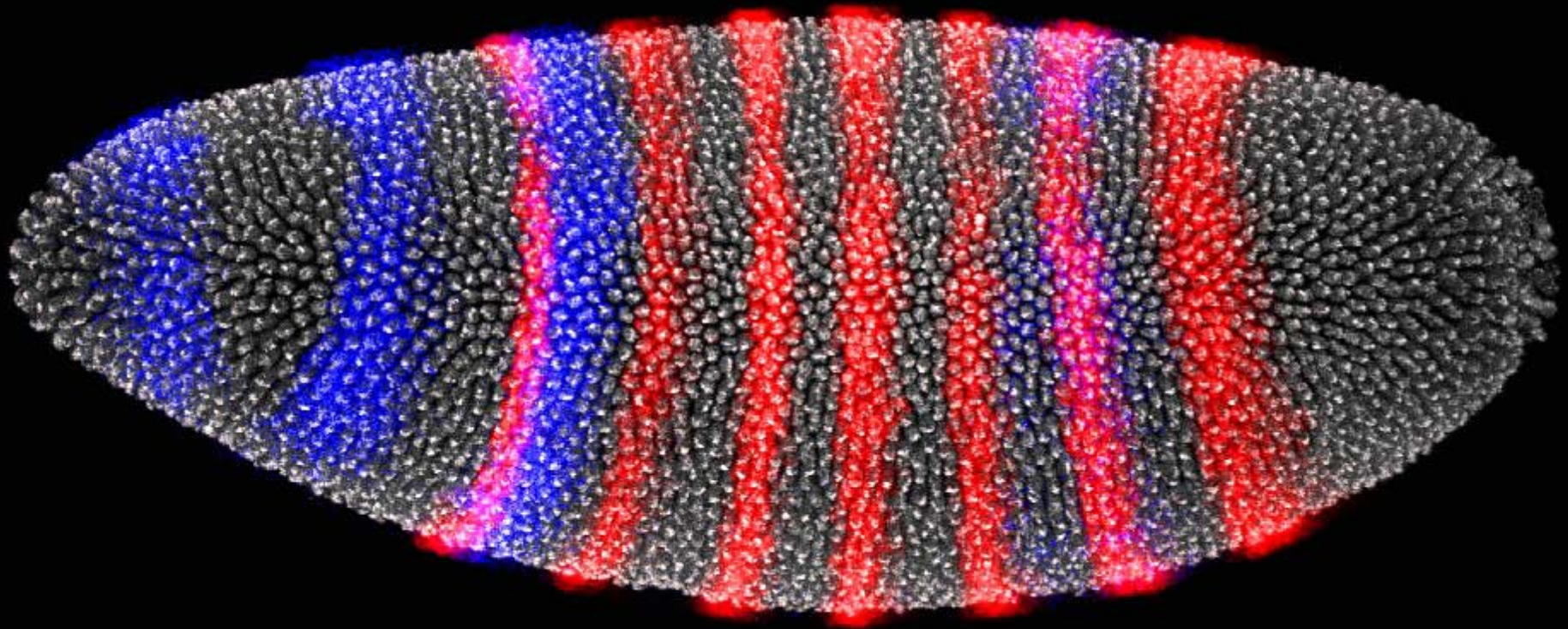
Outline

- Building a high resolution spatio-temporal atlas of gene expression in the *Drosophila* blastoderm
- Characterizing variation in expression patterns:
 - among individuals in a population
 - between species

Challenges for acquiring spatio-temporal expression data

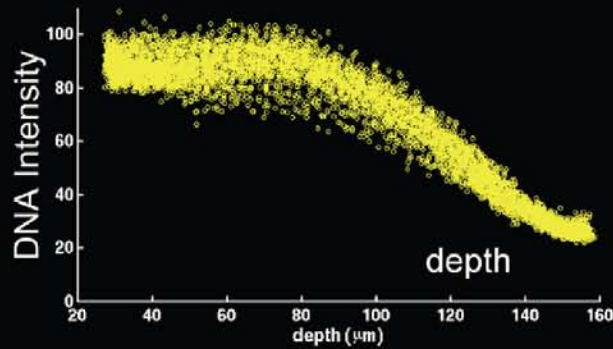
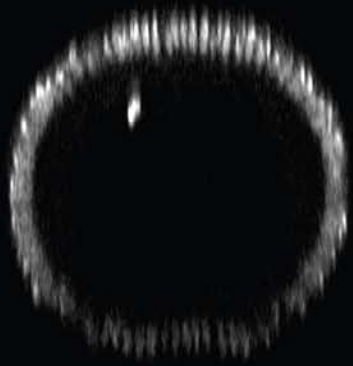
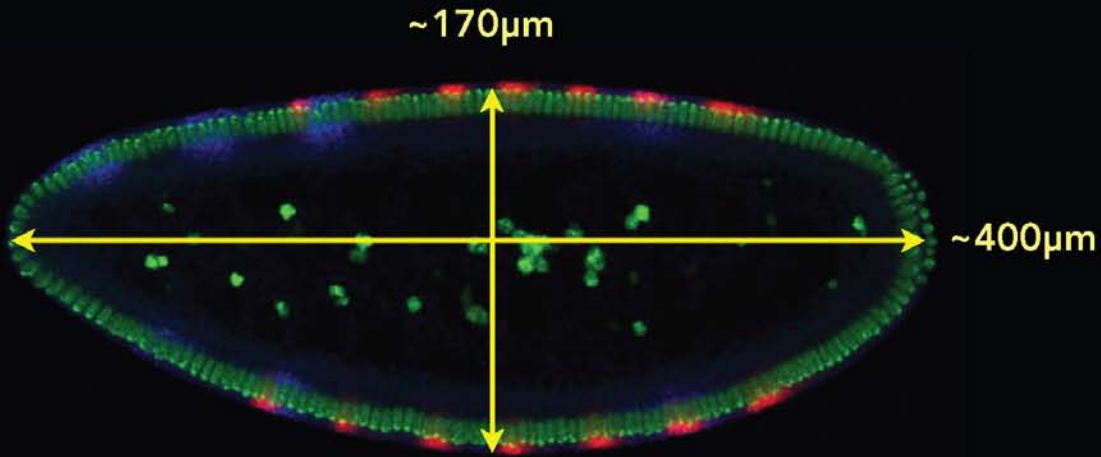
- Imaging large numbers of whole-mount embryos with cellular resolution
- Accurate quantitative measurements of gene expression and protein concentrations
- Building a high fidelity model of embryo morphology
- Combining expression data from different embryos into a common model atlas

3D gene expression imaging

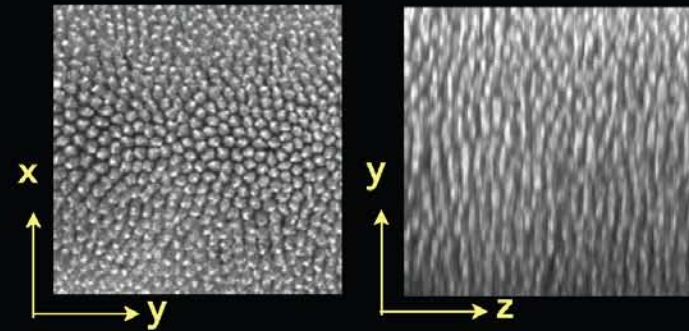


- ***In situ*–hybridization**
 - Anti-sense RNA-probes, Tyramide Signal Amplification
 - Sytox labeled DNA
- **2-photon fluorescence microscopy**

Imaging through the embryo is challenging



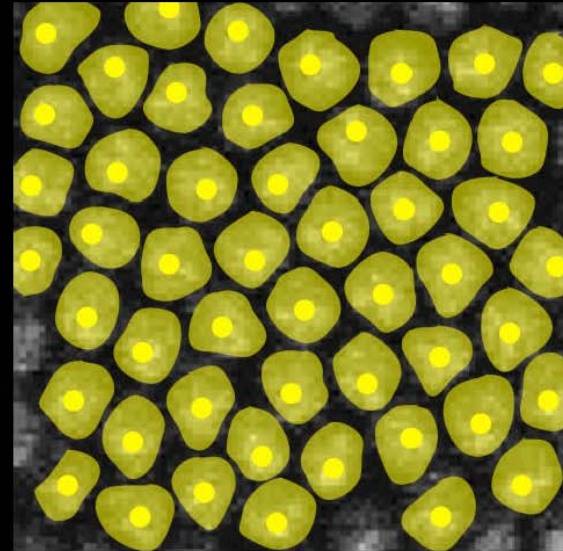
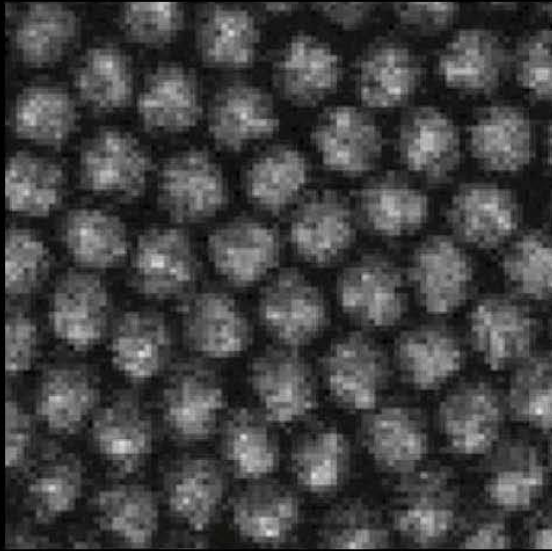
Penetration loss



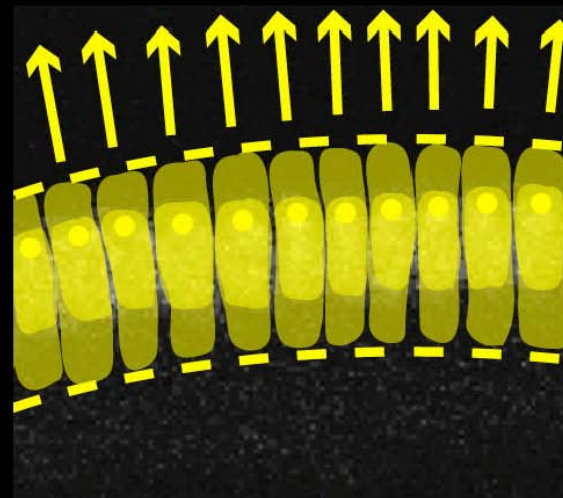
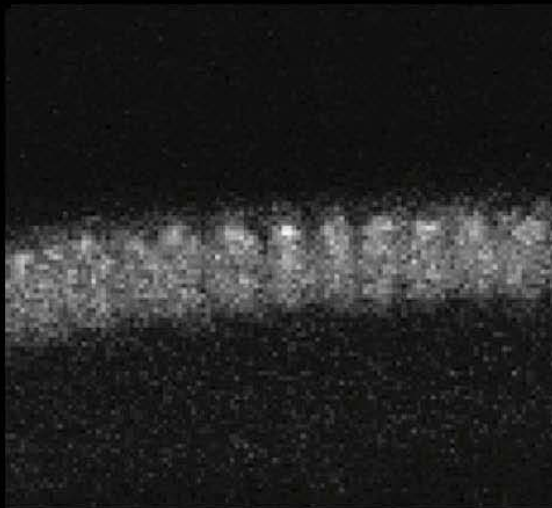
Decreased resolution in Z

Nuclei are segmented from the 3D stack

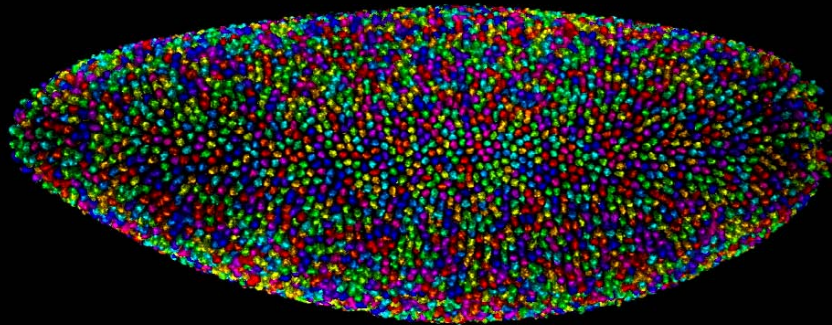
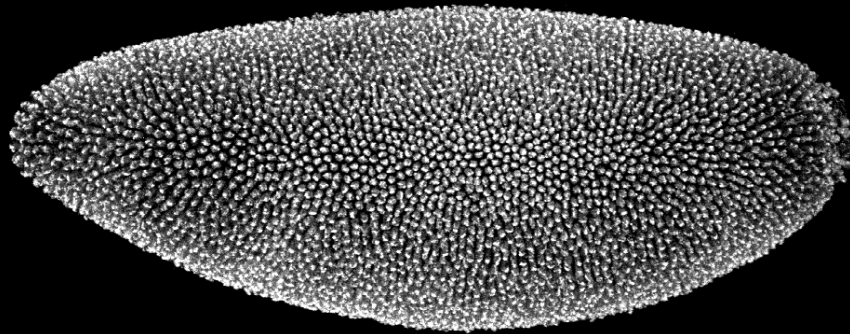
Top



Mid section



Extracting cellular resolution expression data

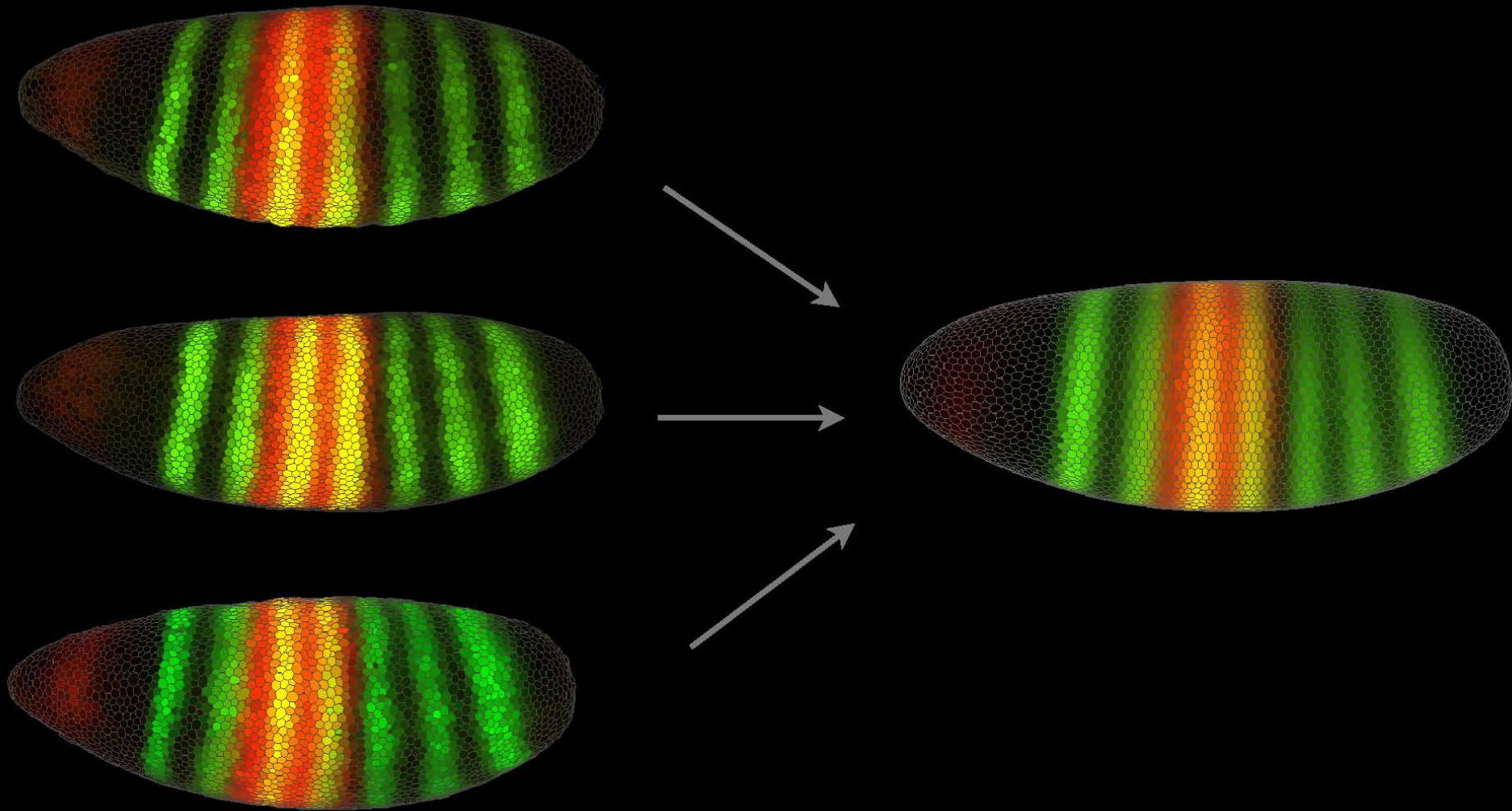


1. Segment out each individual nuclei
2. Record 3D location and expression levels associated with each nucleus.

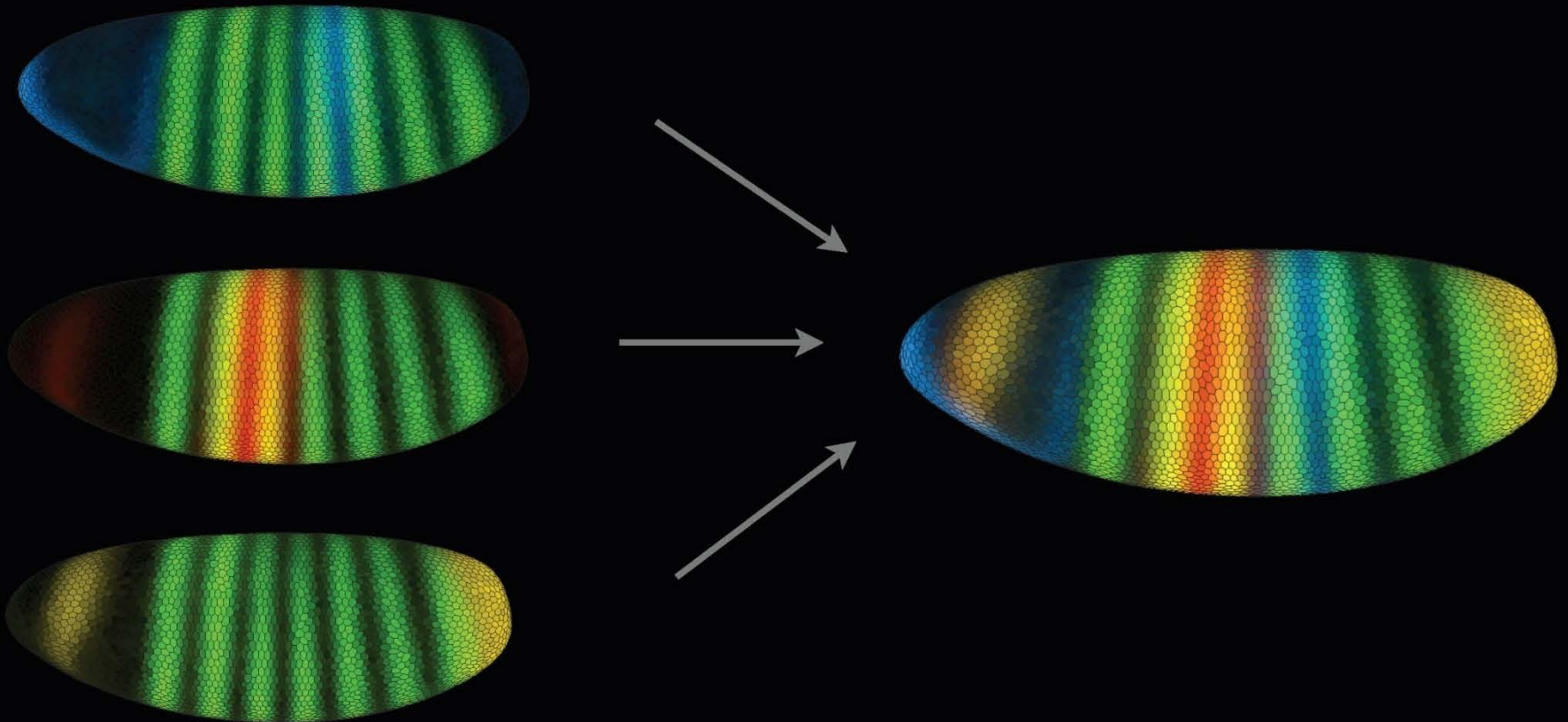
```
id, x,      y,      z,      Vn,      Vc,      Sytox, Cy3_n, ...  
1, 102.36, 142.14, 112.00, 207.96, 605.36, 52.18, 23.55, ...  
2, 264.63, 172.01, 79.36, 281.73, 599.90, 82.12, 31.67, ...  
3, 225.91, 174.99, 88.65, 185.79, 418.35, 85.32, 35.63, ...  
...  
...
```

pointcloud file

Average and compare expression across individuals



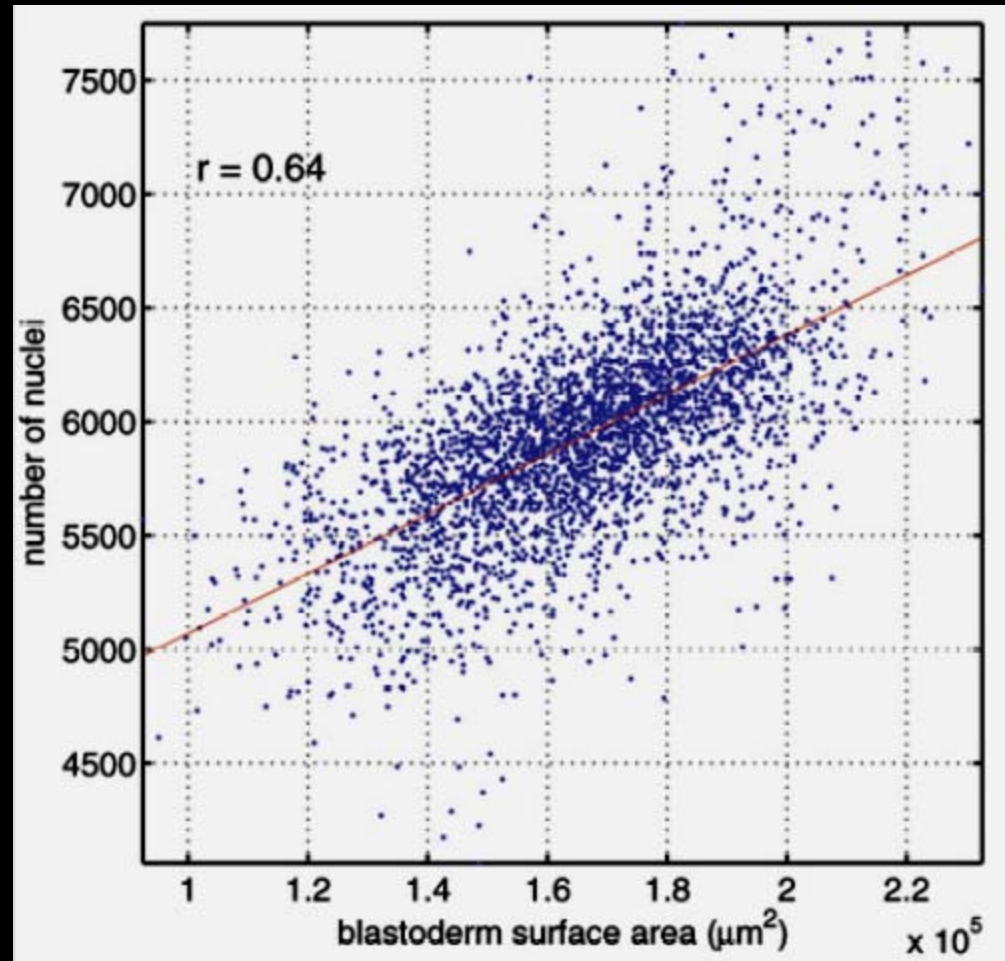
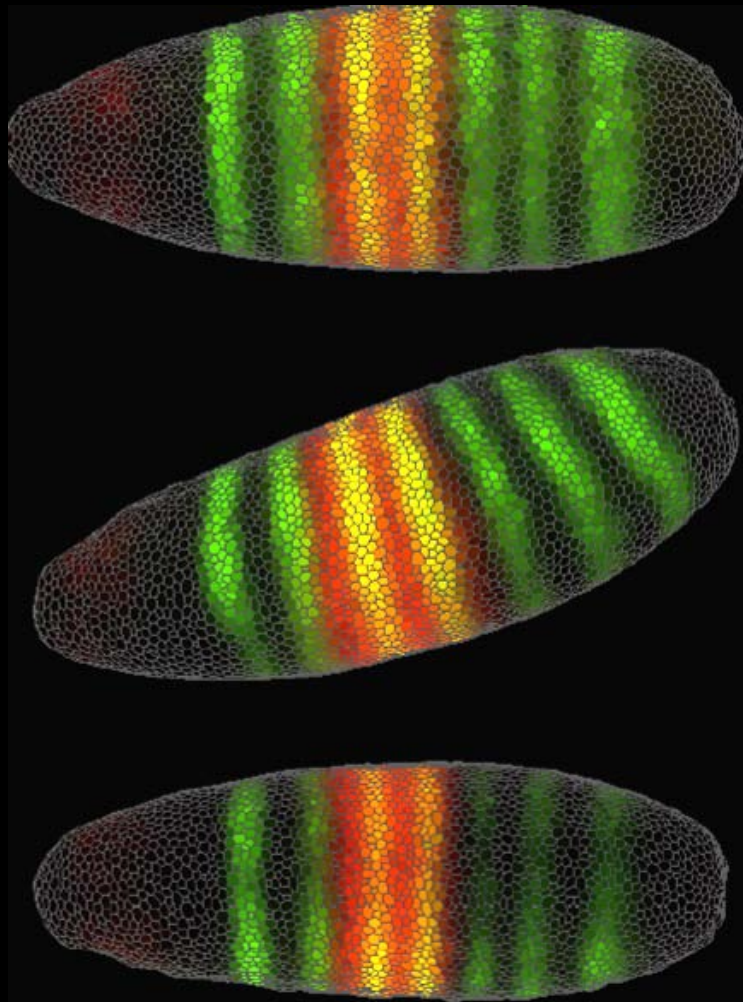
Assemble expression measurements for multiple genes



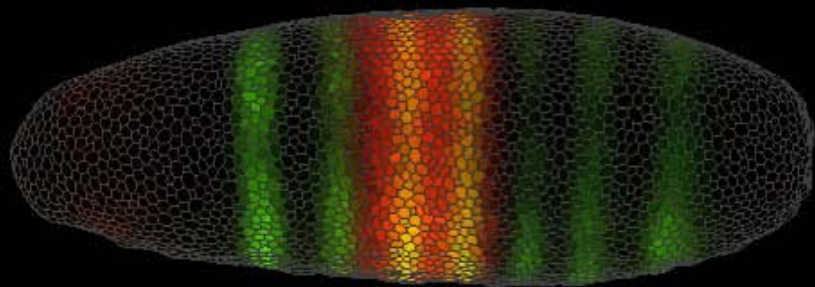
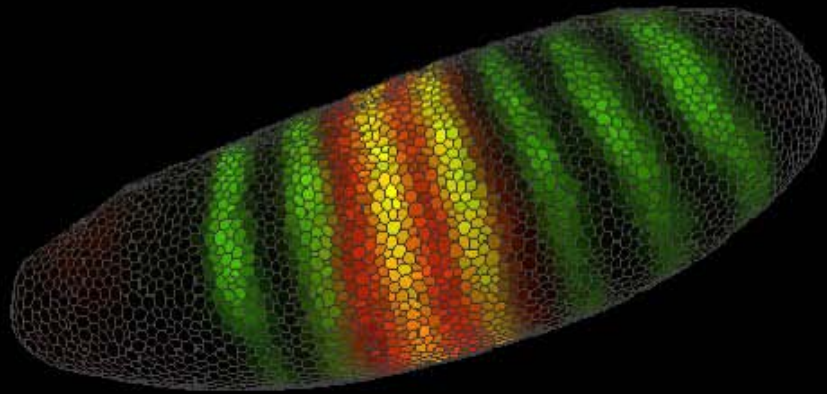
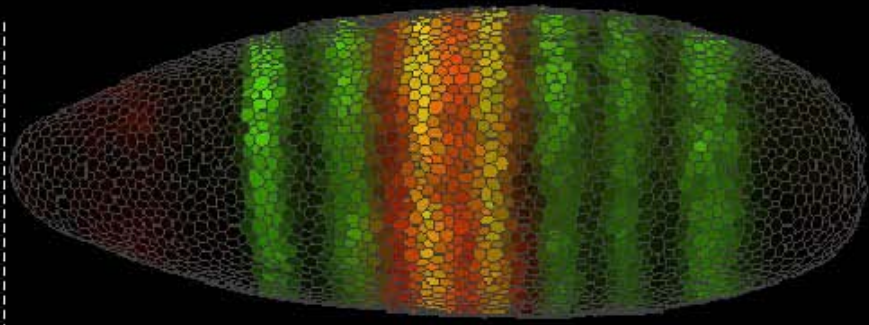
Model can include an arbitrary number of genes

Allows analysis of spatial relationship between many genes

Difficulty: significant variations among individuals within a population

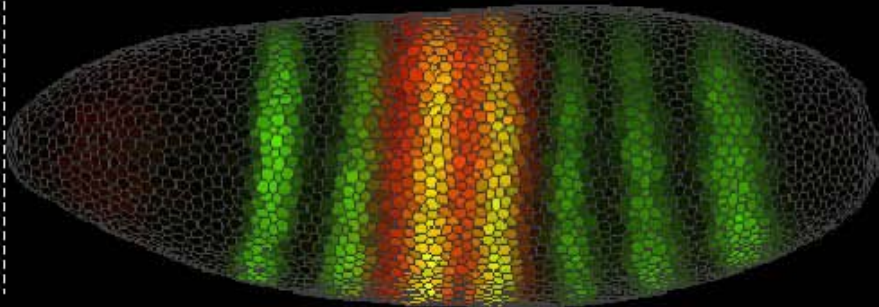
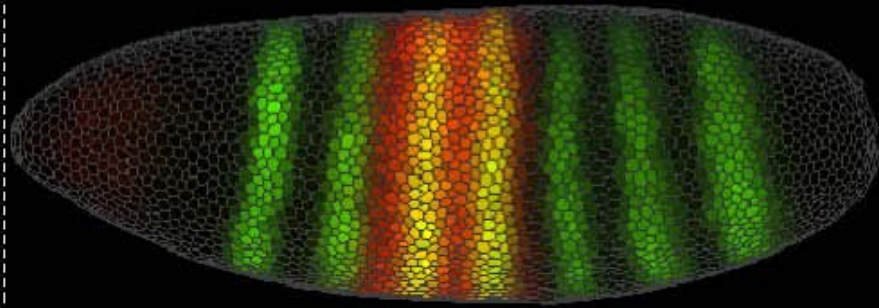
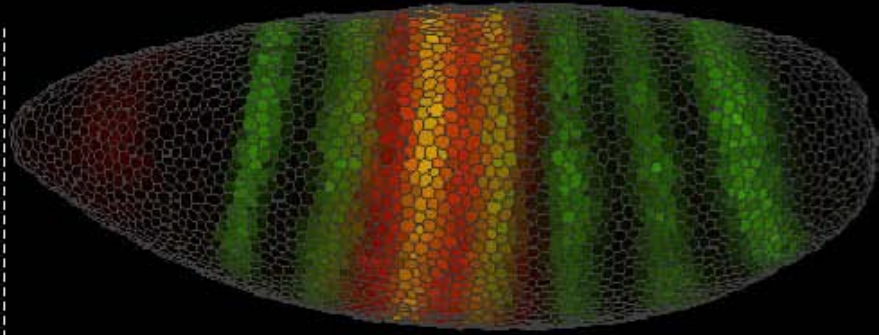


Coarse registration roughly aligns embryos



Align
Orient
Scale

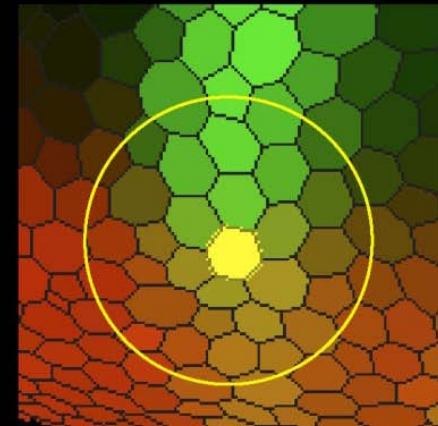
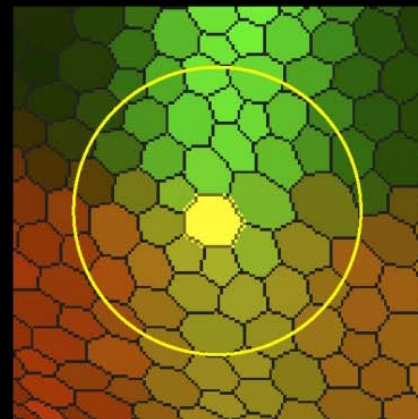
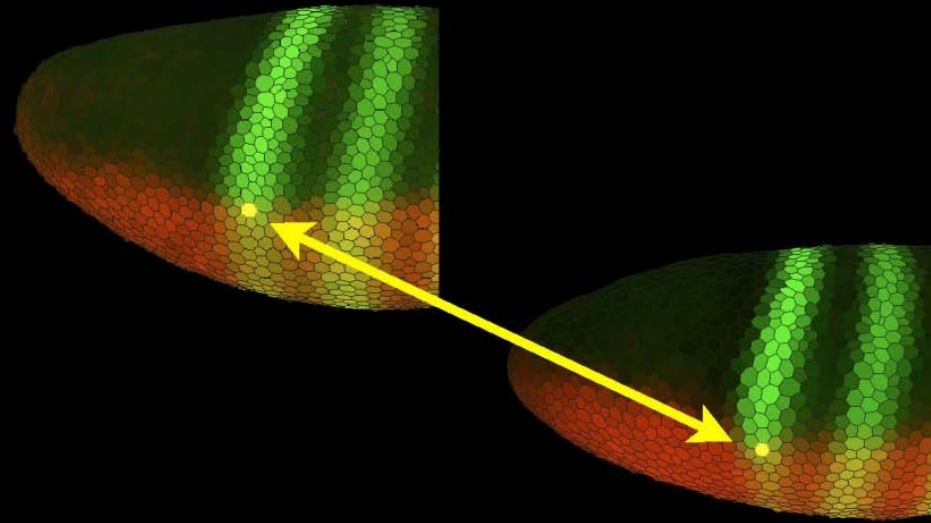
Coarse registration roughly aligns embryos

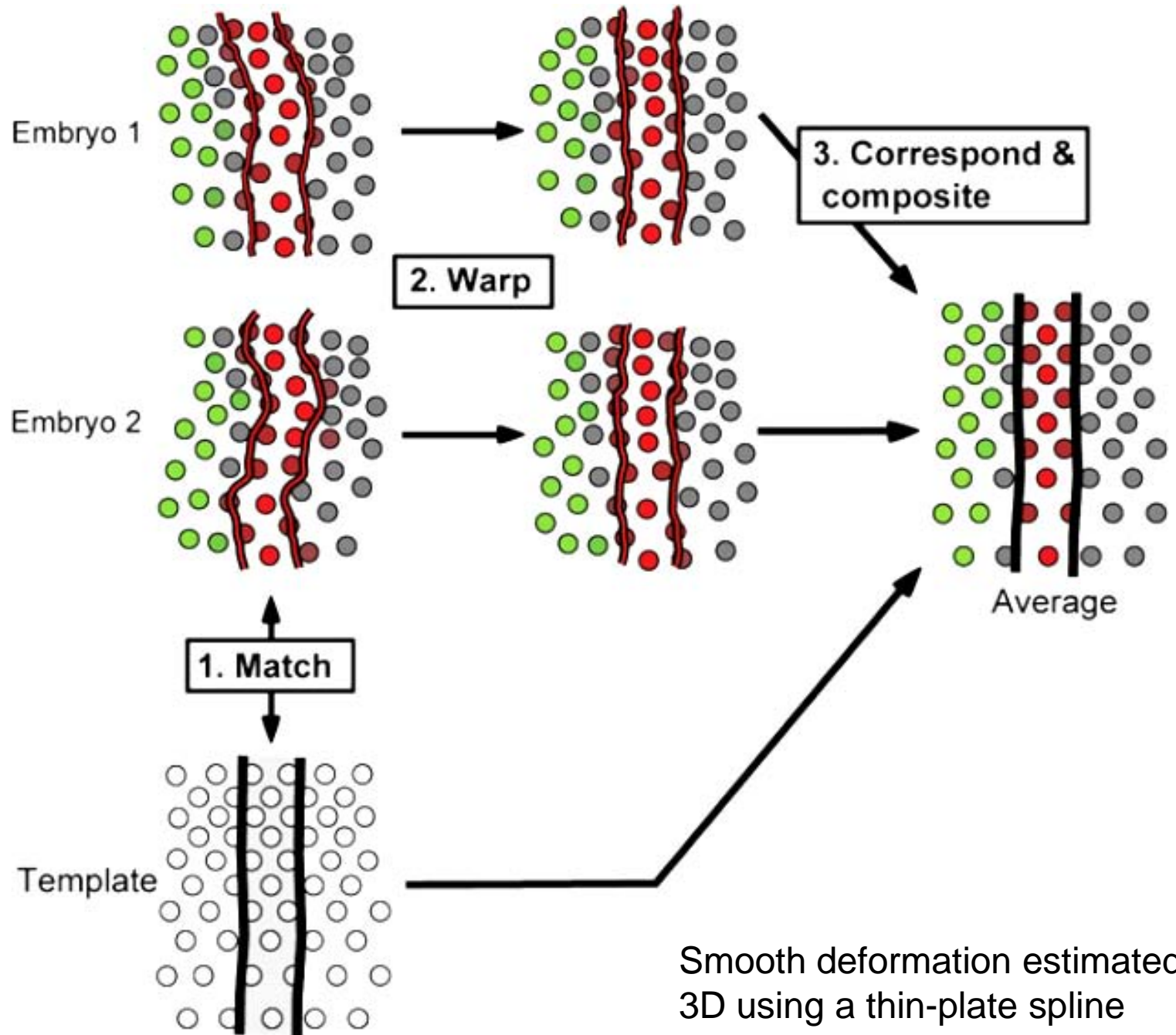


Align
Orient
Scale

Fine Registration

Idea: Find corresponding cells based on spatial location and local pattern of expression





Smooth deformation estimated in 3D using a thin-plate spline

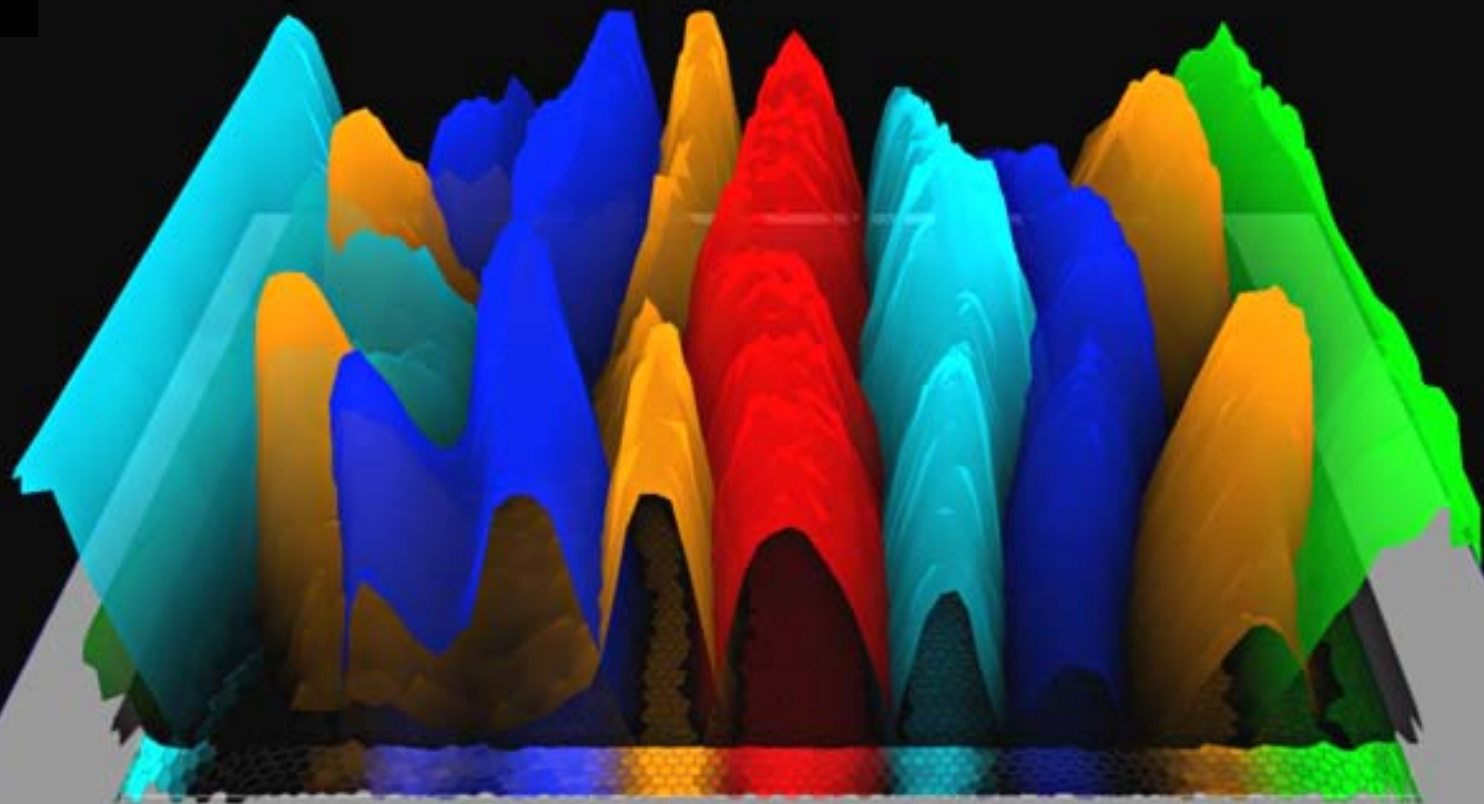
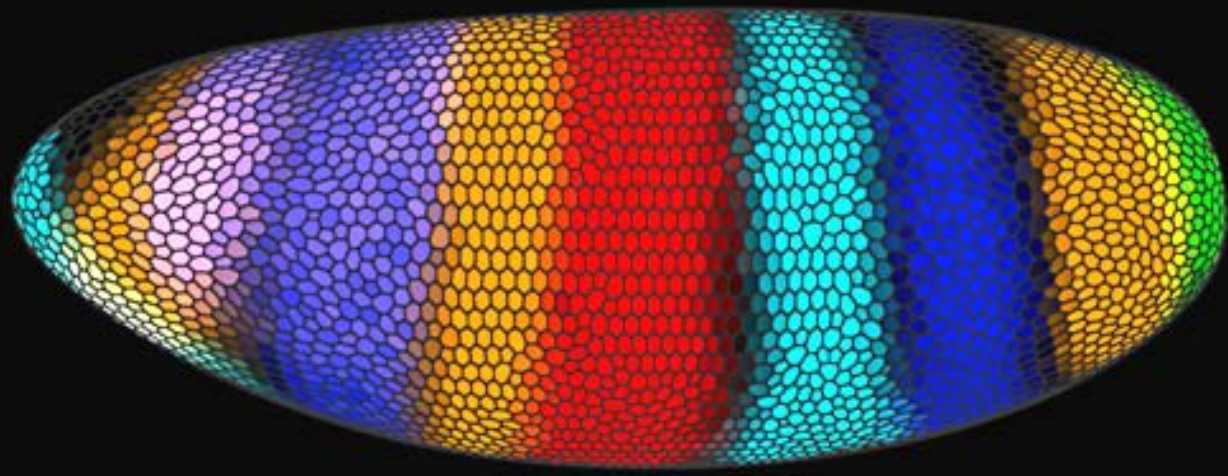
fkh

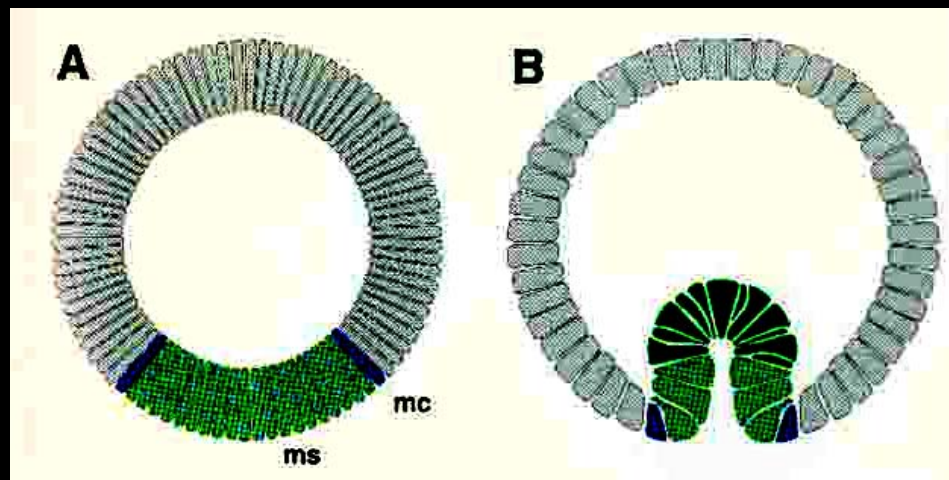
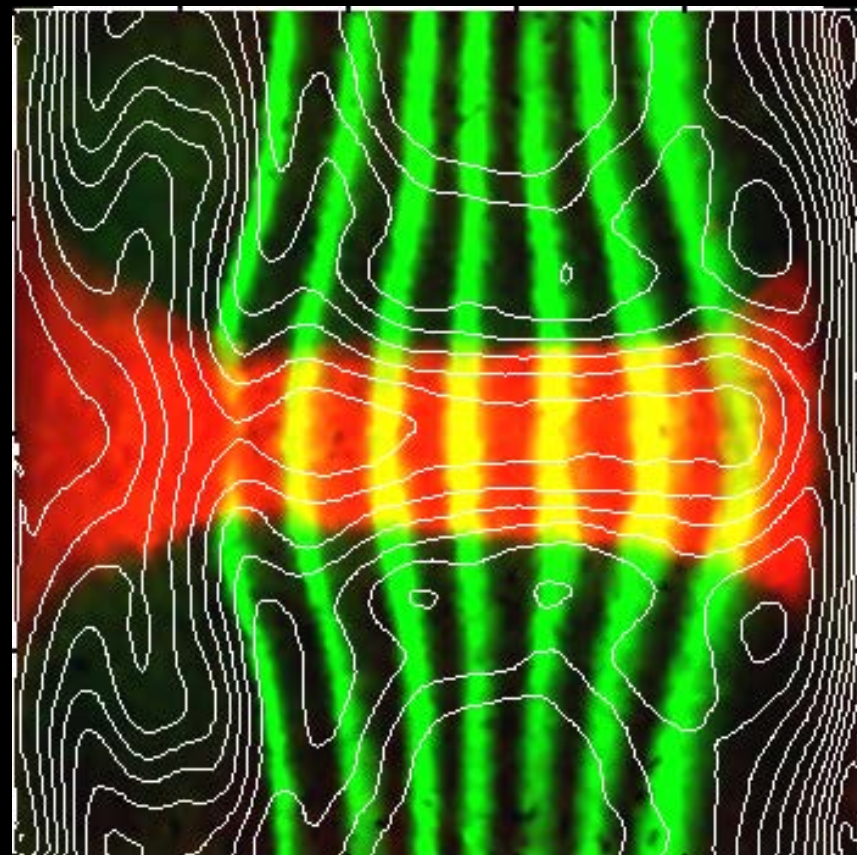
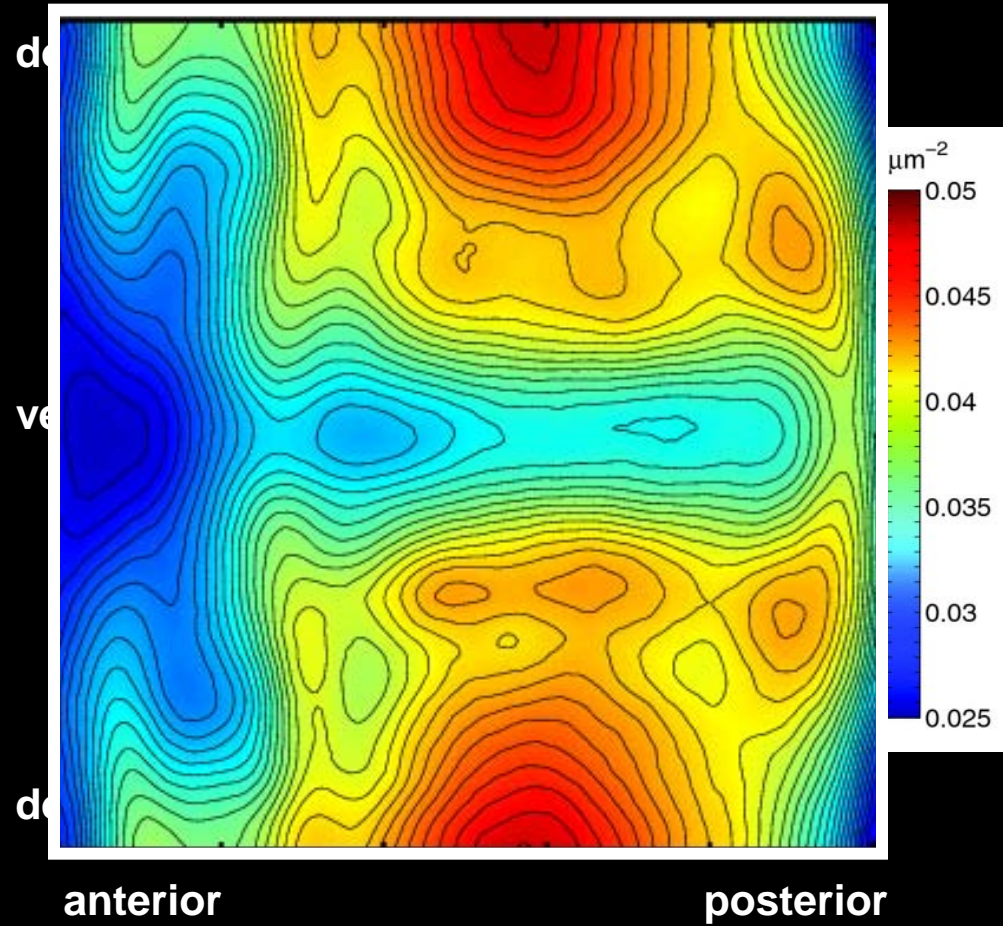
gt

hb

kni

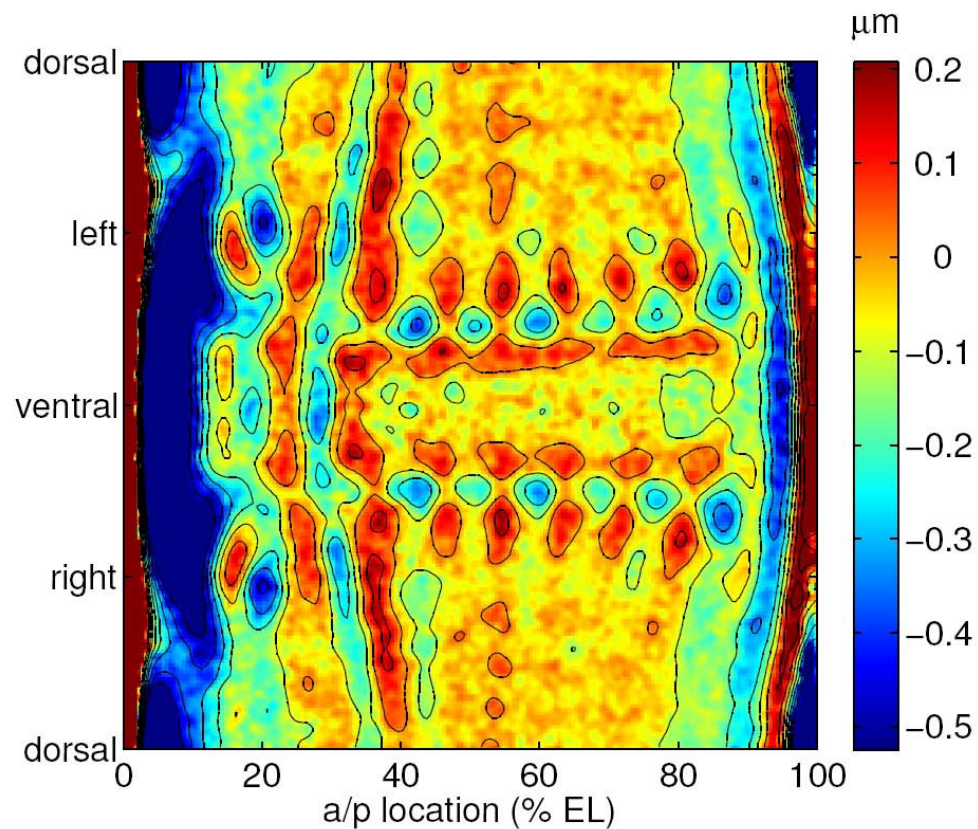
Kr



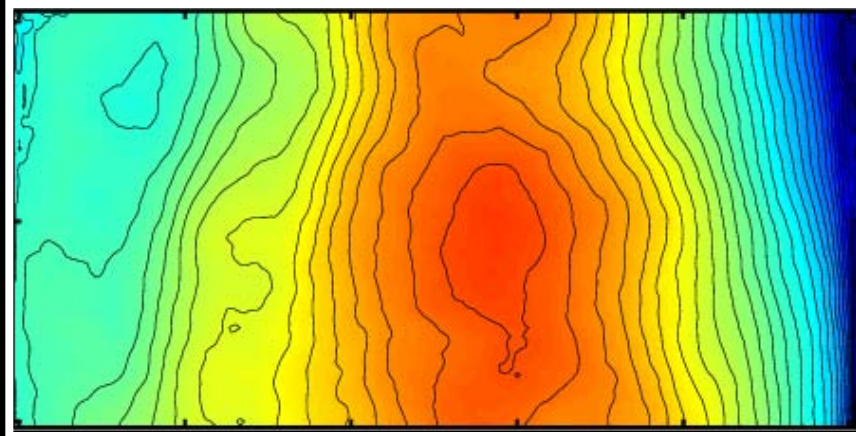


Average displacements of nuclei normal to the blastoderm surface

average displacement normal to
smoothed surface



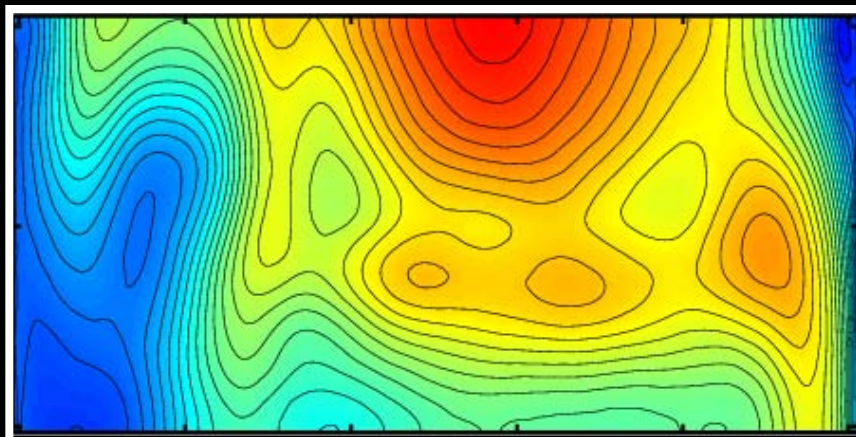
gd7 mutant
(dorsalized)



dorsal

ventral

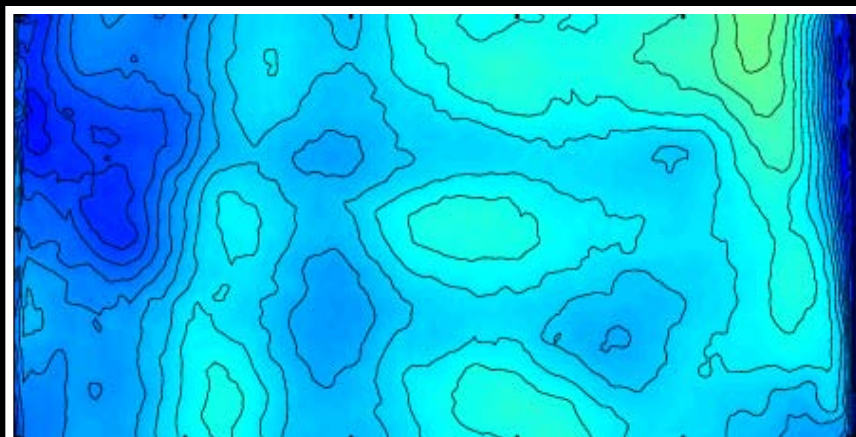
wildtype



dorsal

ventral

toll^{10B} mutant
(ventralized)



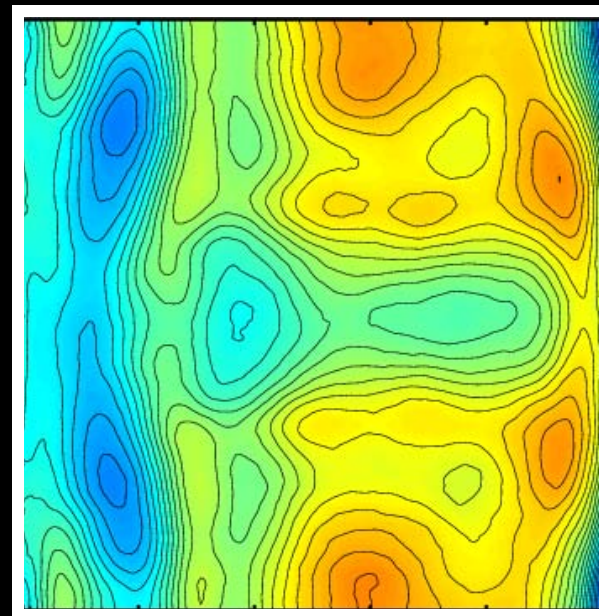
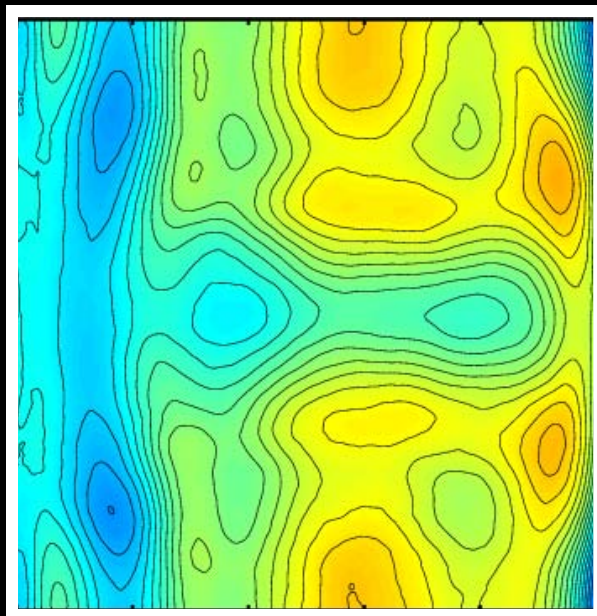
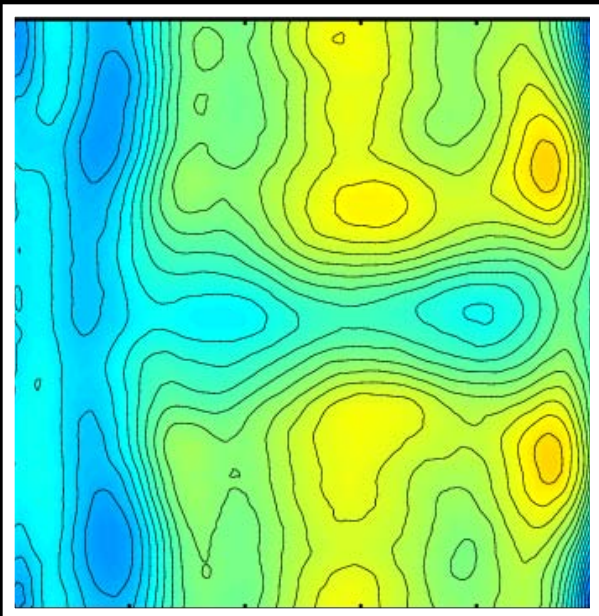
dorsal

ventral

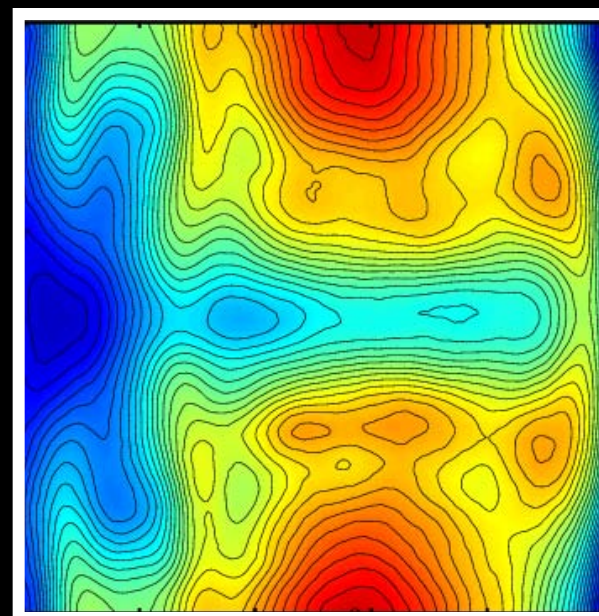
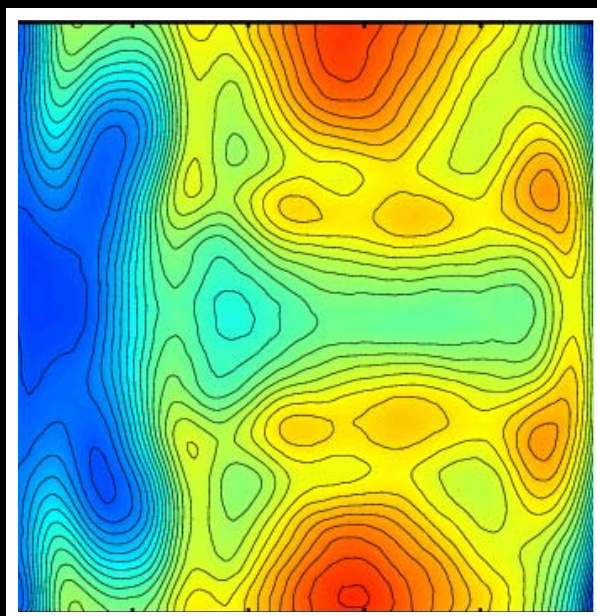
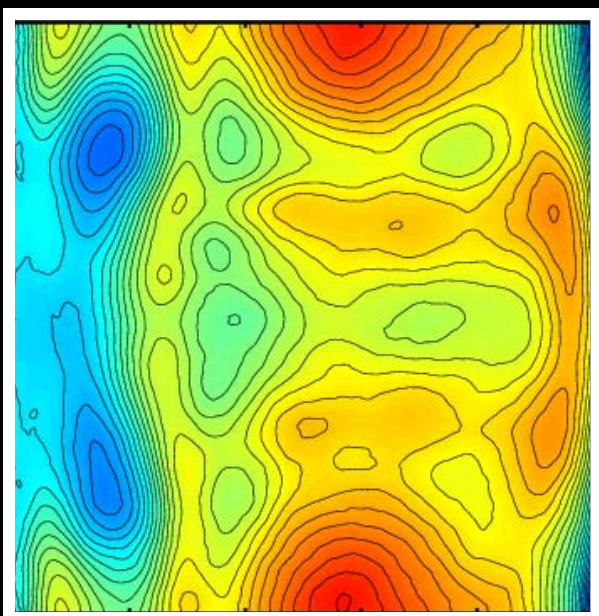
anterior

posterior

0% membrane invagination...

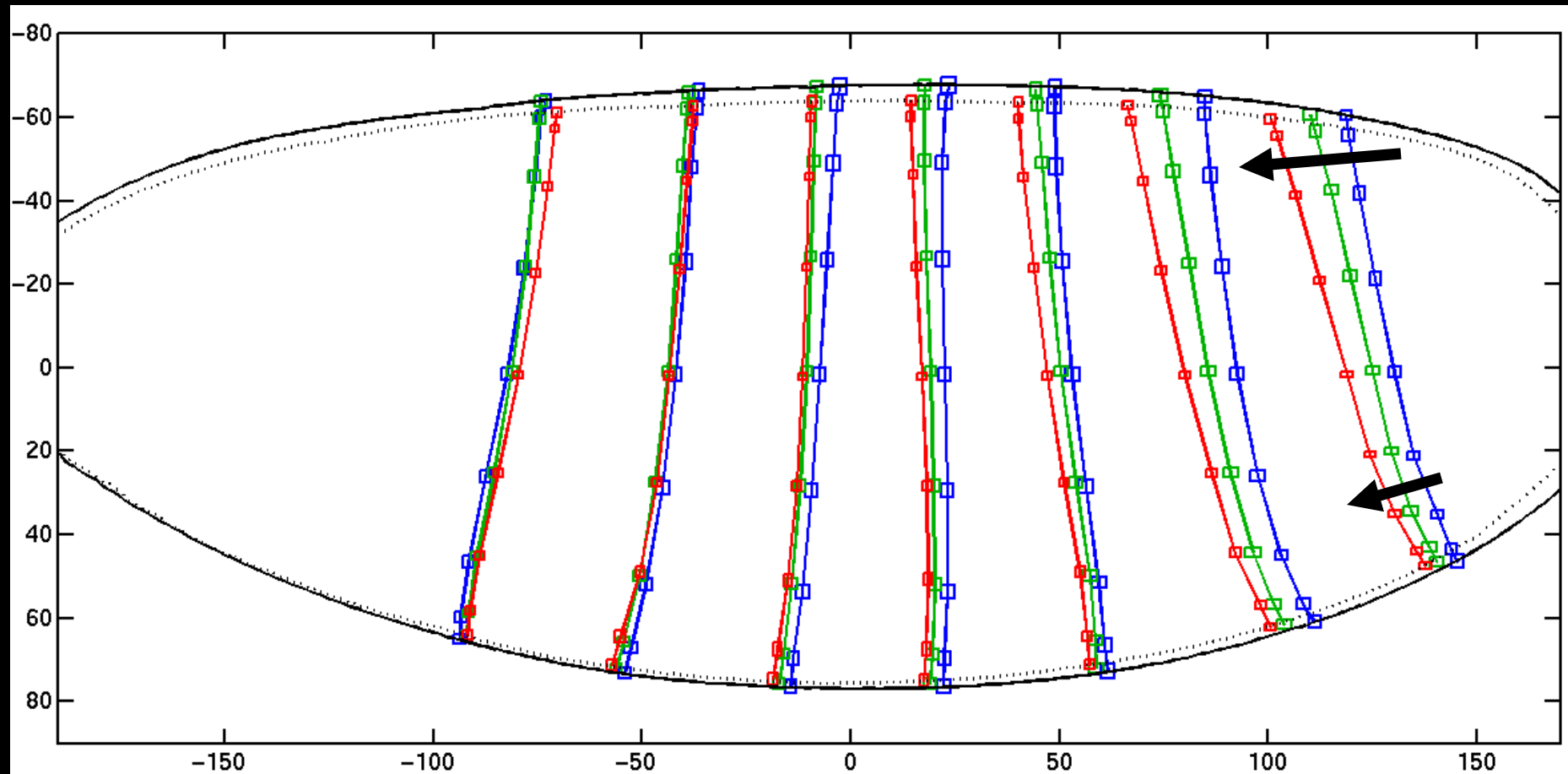
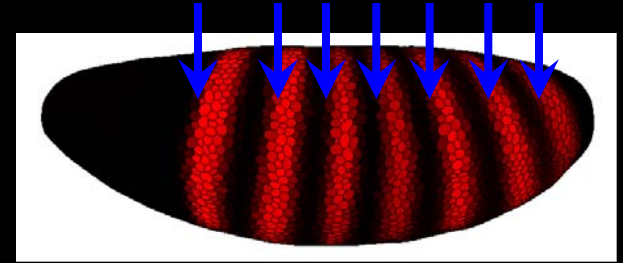


...100% membrane invagination

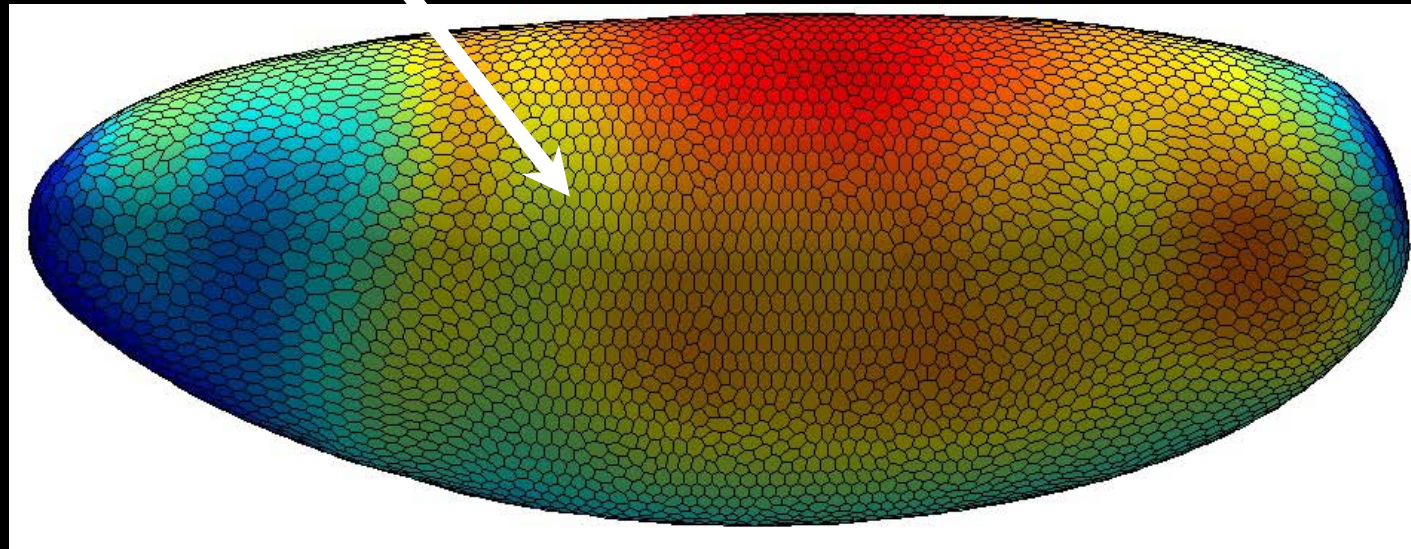
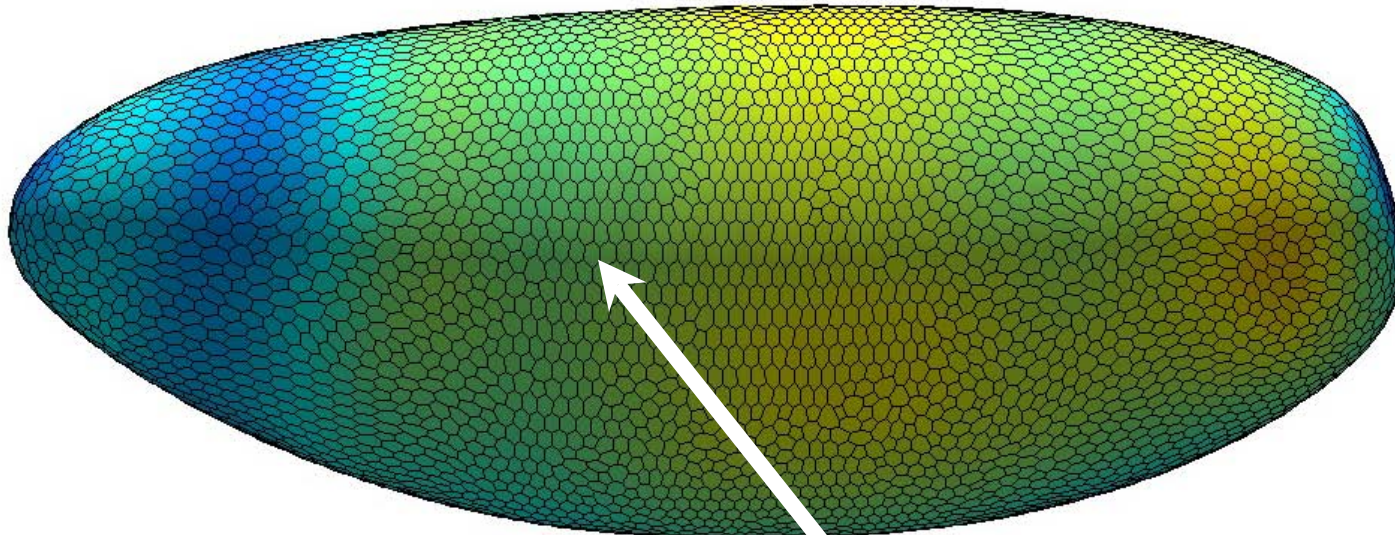


Movement of anterior *ftz* boundaries

early: 4-8% (44 embryos)
mid: 26-50% (42 embryos)
late: 76-100% (57 embryos)

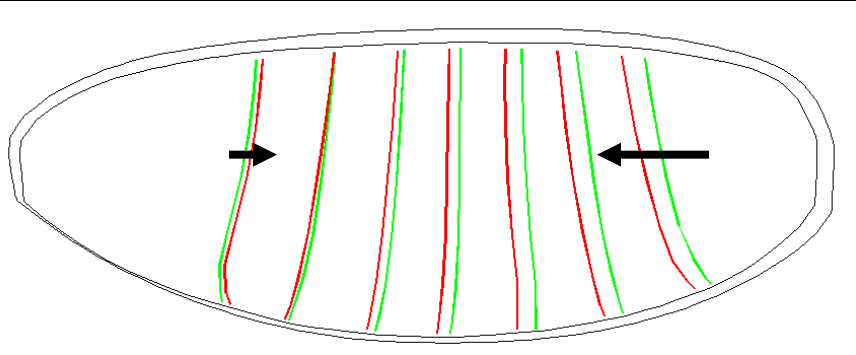


Average Morphological Template

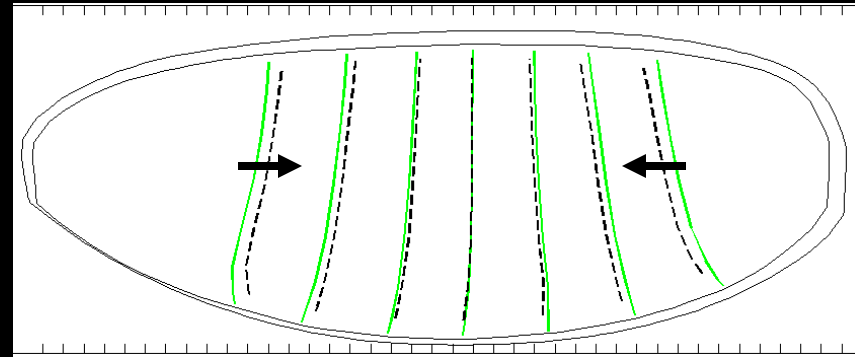


6000 nuclei with average shape and density for each of 6 time intervals

Pattern Dynamics

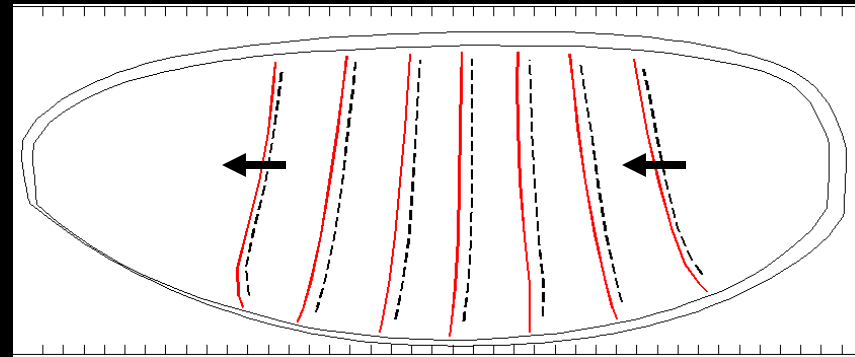


Nuclear movement

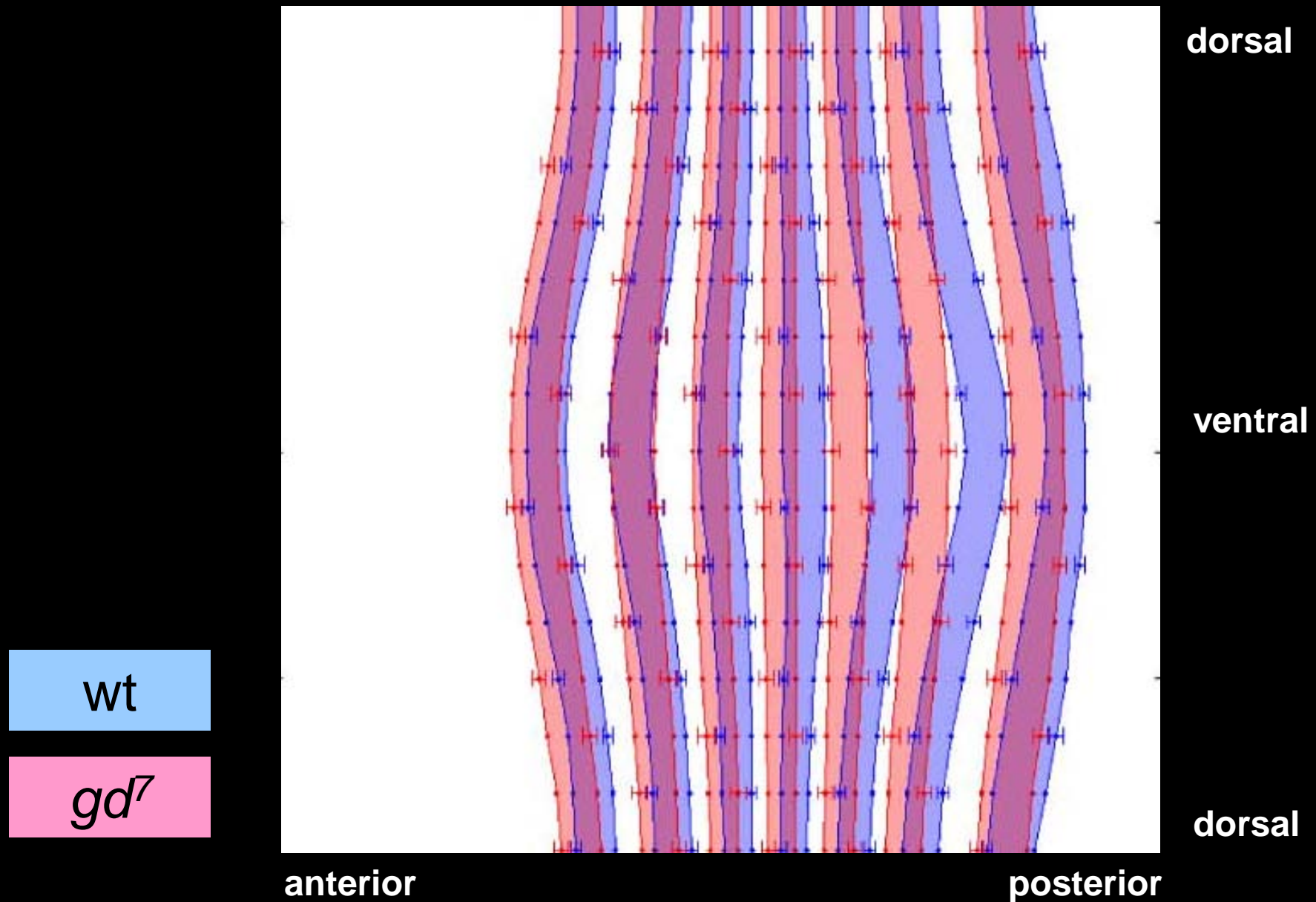


+

Changing expression

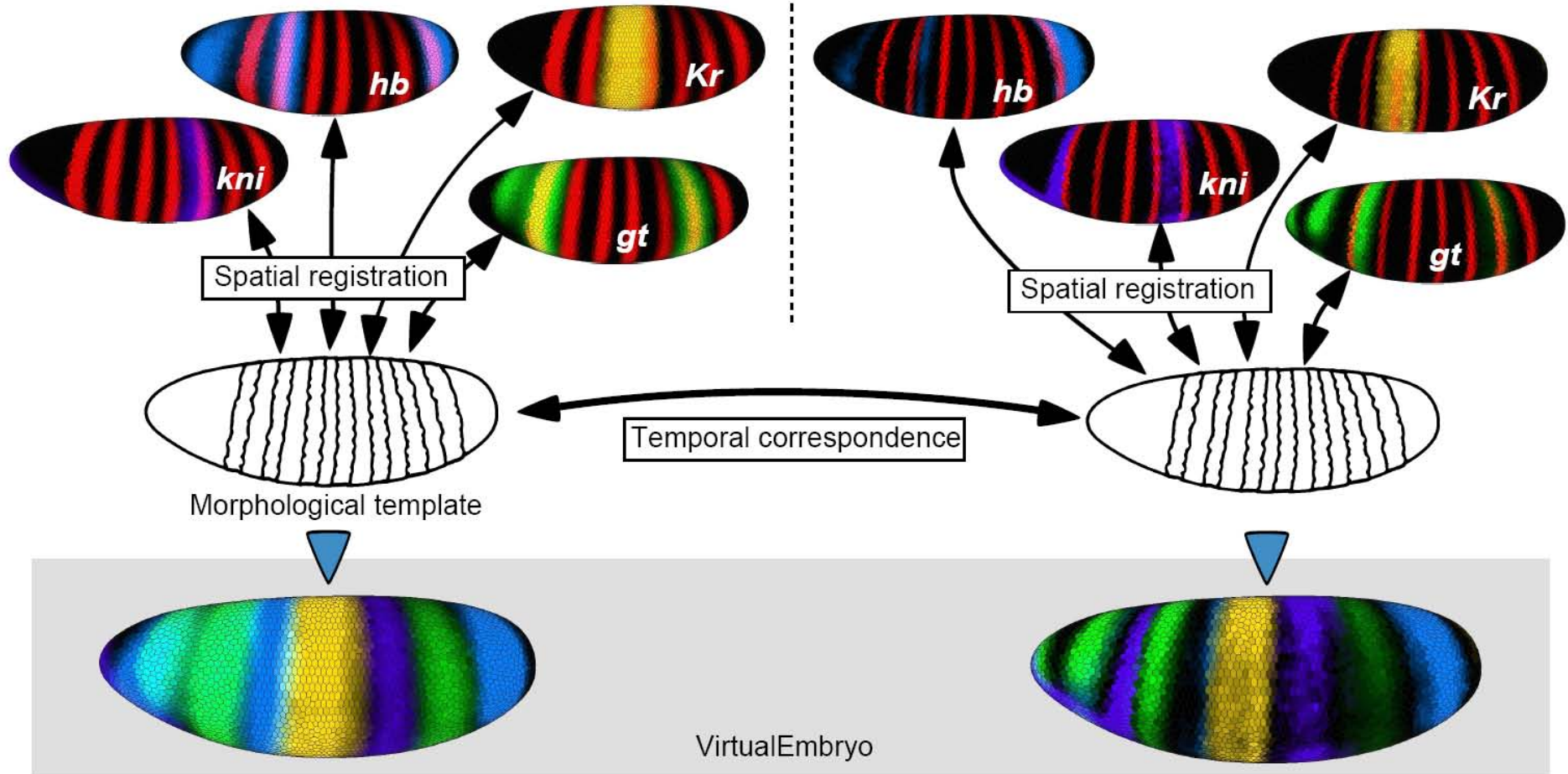


ftz stripe locations in *gd⁷* (dorsalized)

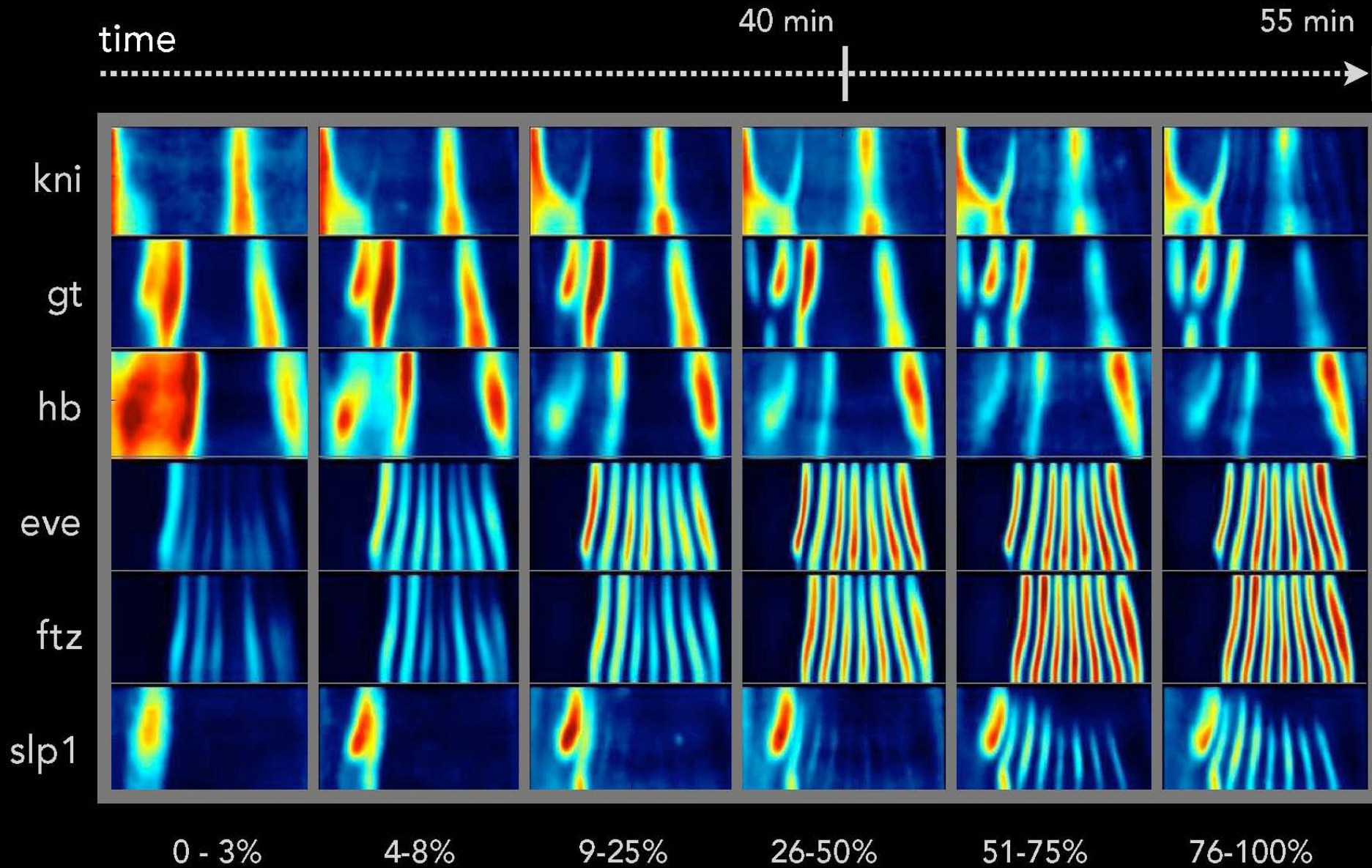


Early cohort

Late cohort



Expression varies over time

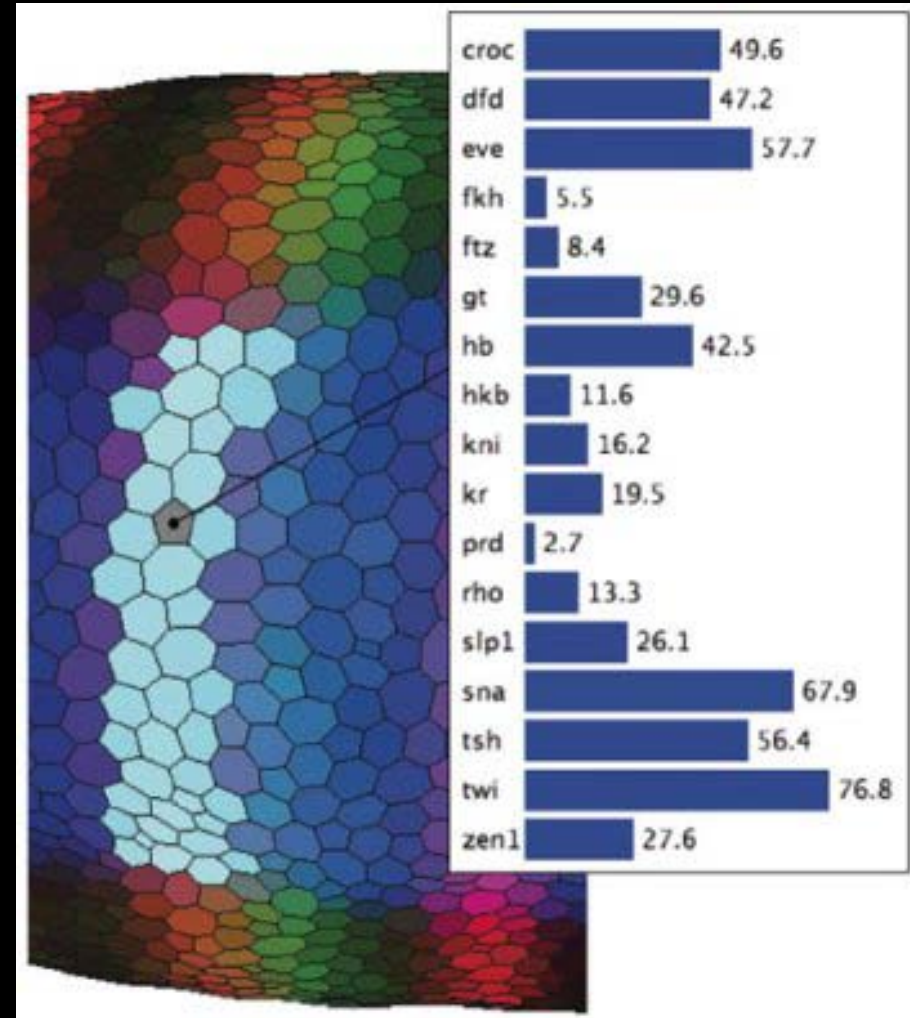


Current state of D. mel atlas

~4500 embryos imaged
>1 TB of raw image data
~100 genes stained for
mRNA

Ongoing imaging of protein
and transgenic lines

1st release online:
<http://bdtntp.lbl.gov>



Outline

- Building a high resolution spatio-temporal atlas of gene expression in the *Drosophila* blastoderm
- Characterizing variation in expression patterns:
 - among individuals in a population
 - between species

D'arcy Thompson

On Growth and Form (1916)

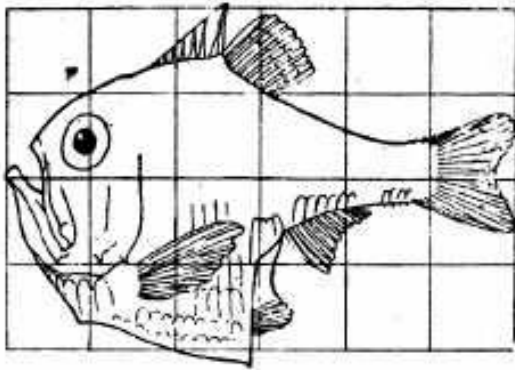


Fig. 517. *Argyropelecus Olfersi*.

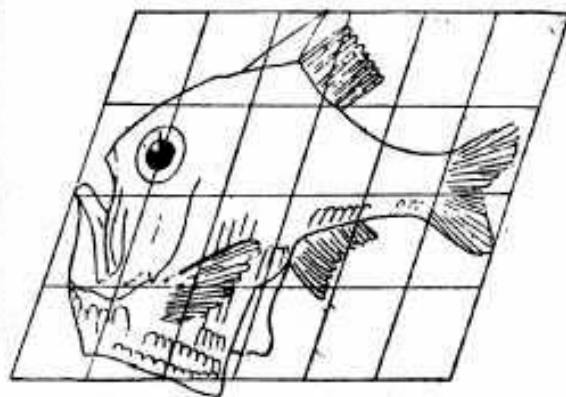
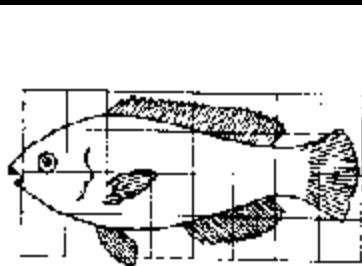
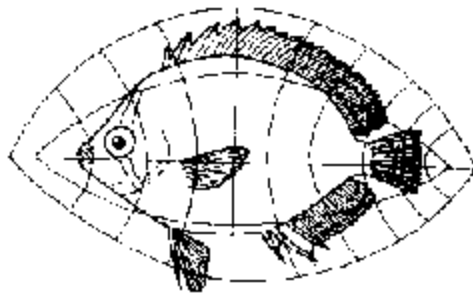


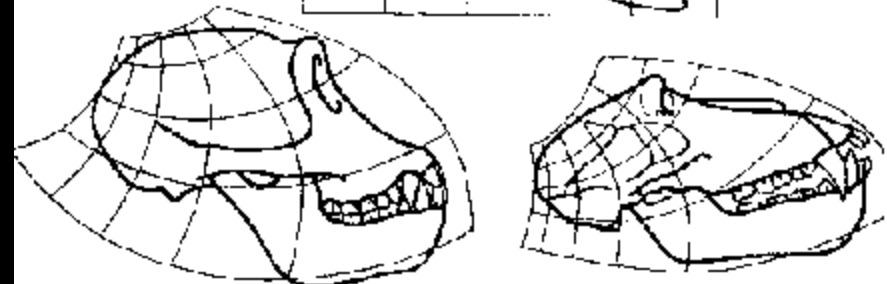
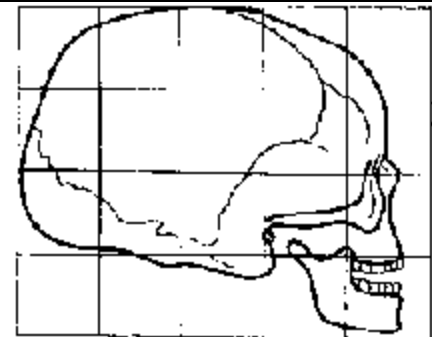
Fig. 518. *Sternoptyx diaphana*.



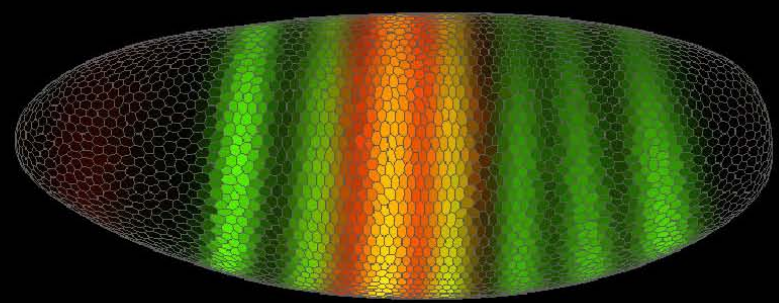
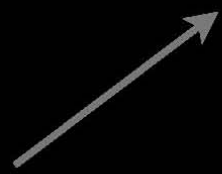
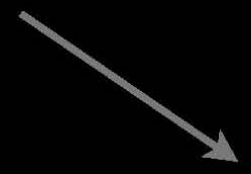
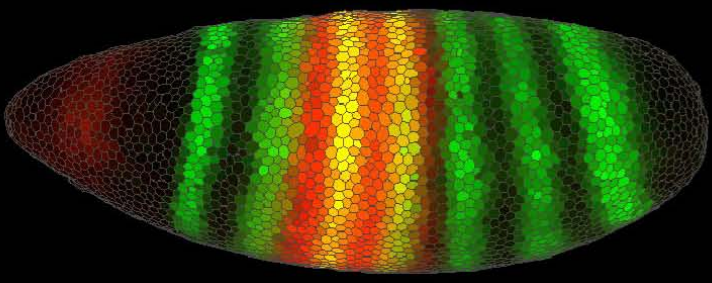
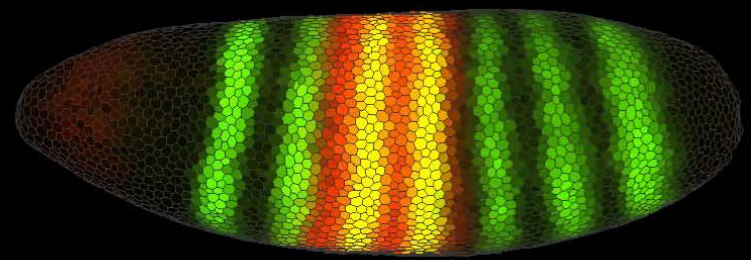
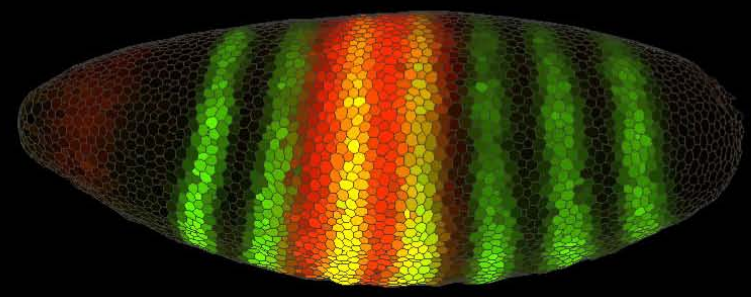
Scarus sp.



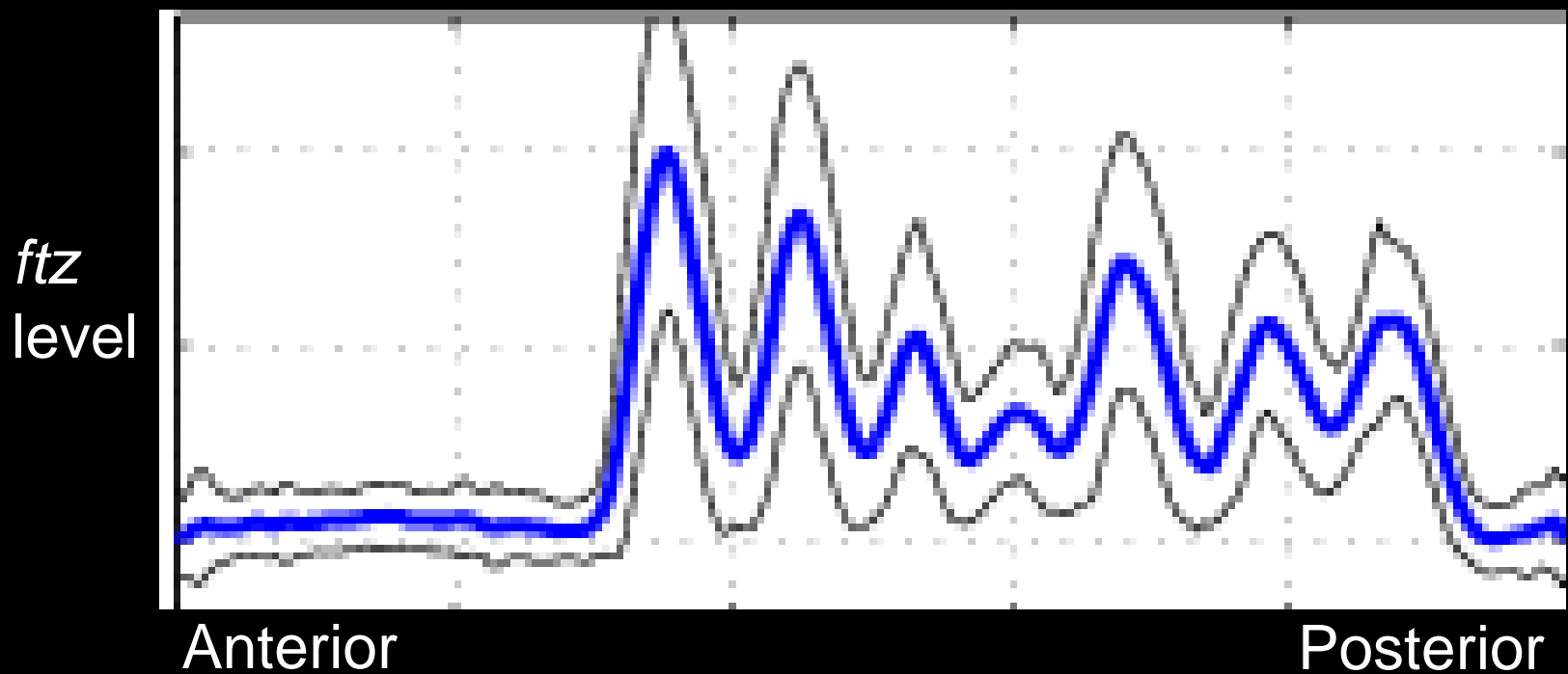
Pomacanthus.



Skulls of a human, a chimpanzee and a baboon
and transformations between them



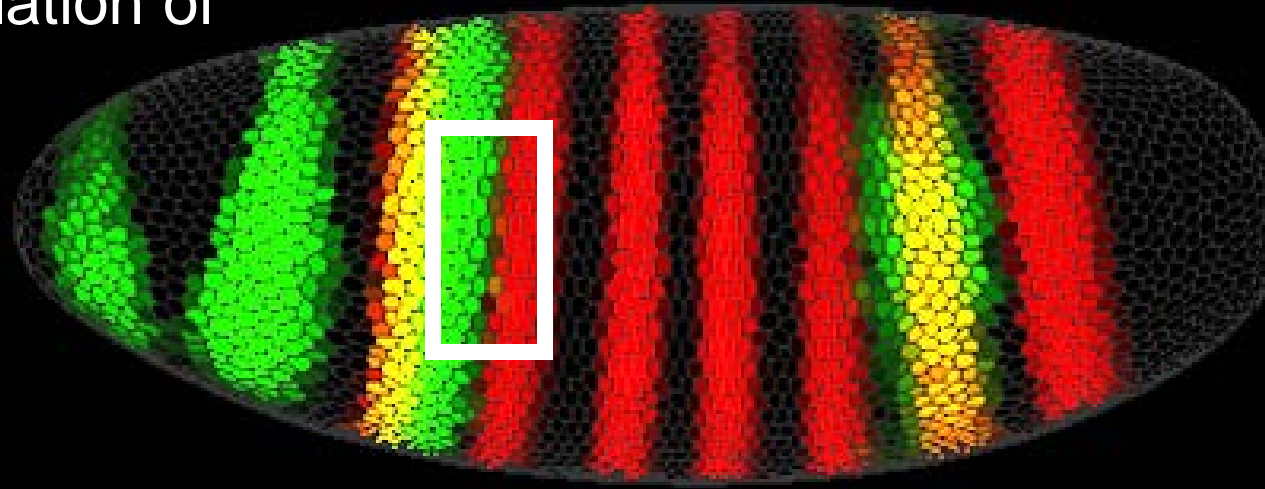
Variation in expression levels



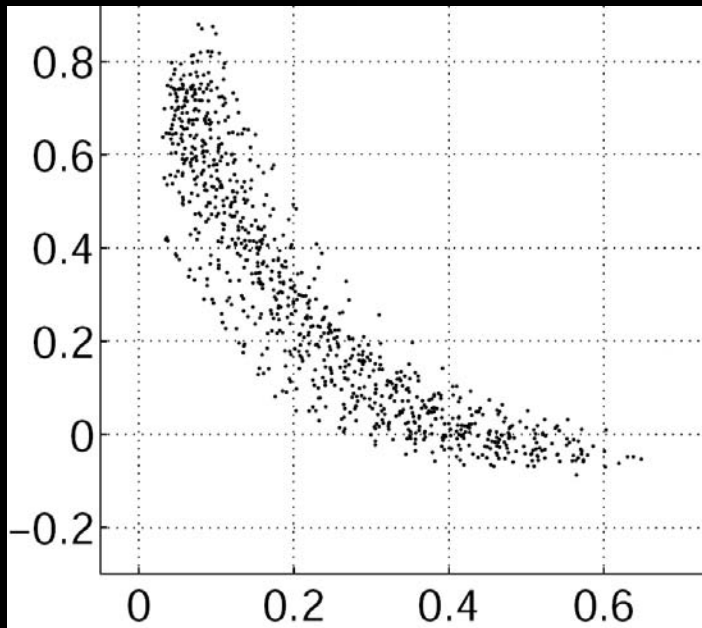
Sources of variation

1. *Regulatory variability* (stochasticity in expression levels)
2. *Geometric variability* (failure to accurately register embryos)
3. Measurement error

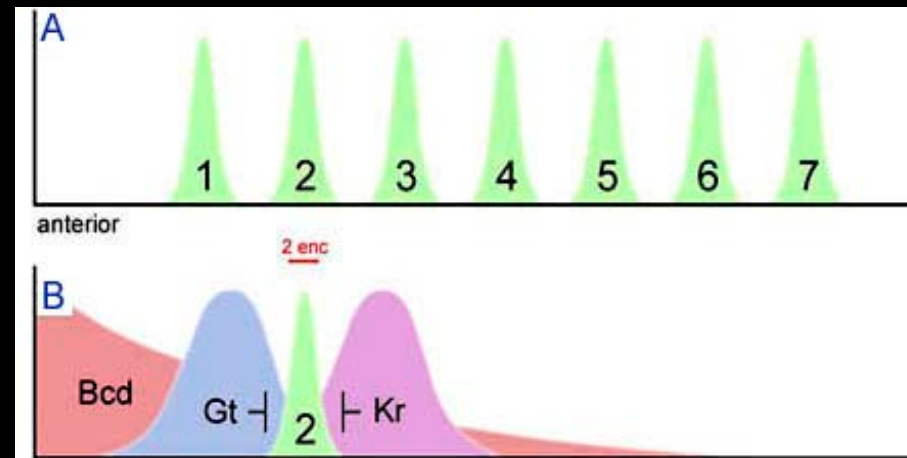
Transcriptional regulation of *eve* stripe 2.

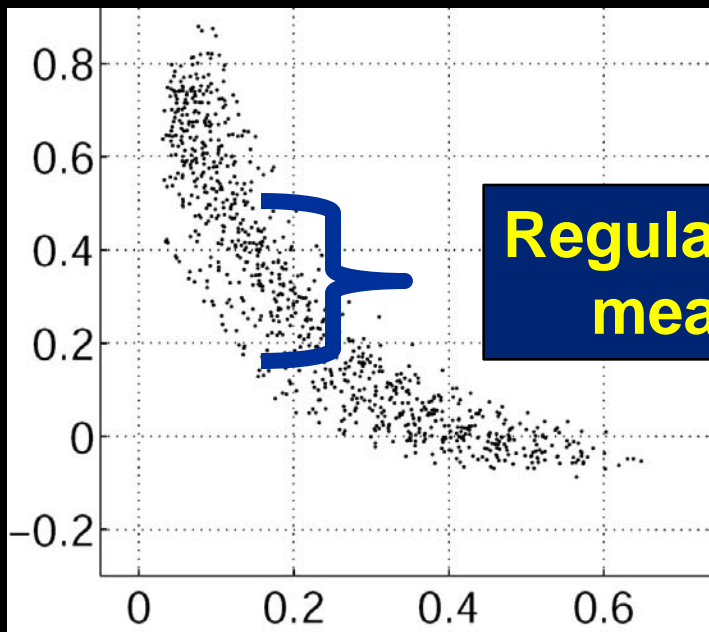
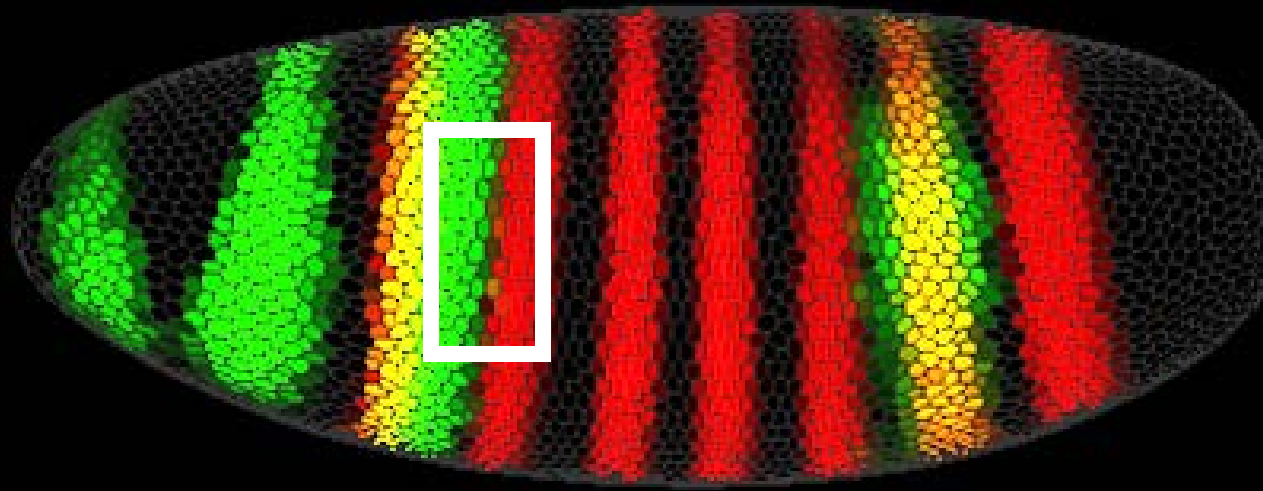


eve (2)



gt



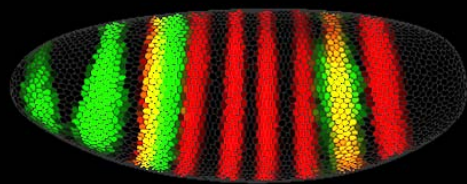


eve (2)

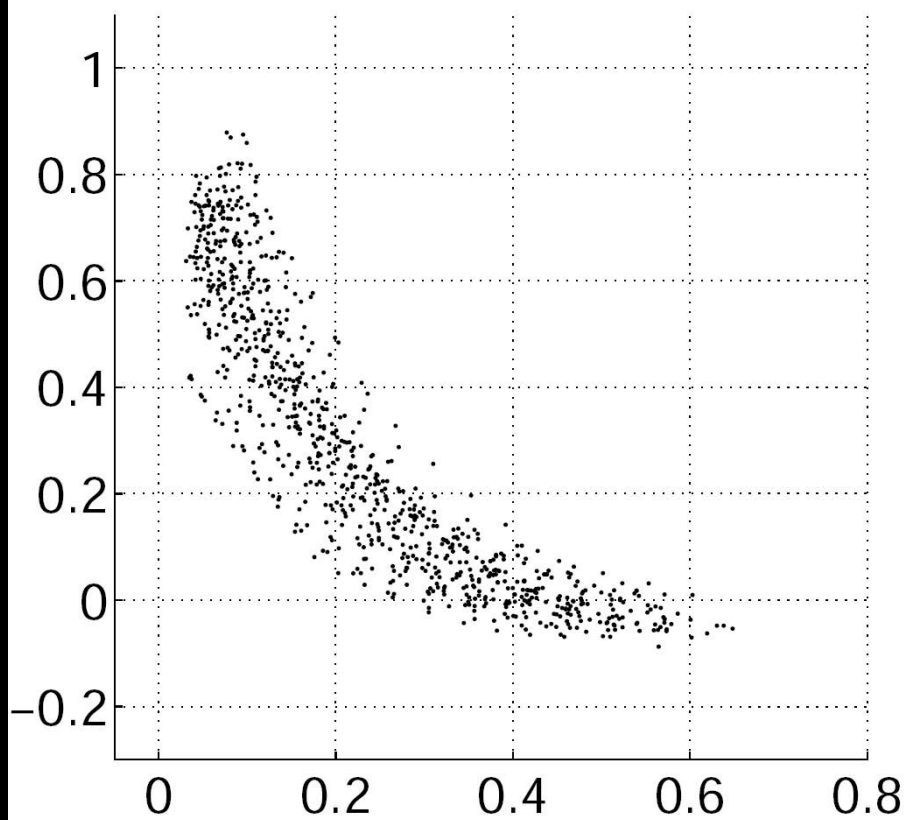
gt

Regulatory variability +
measurement error

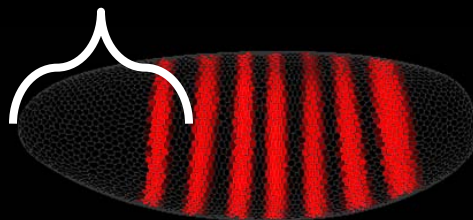
eve / gt



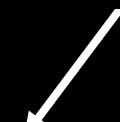
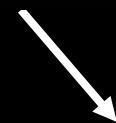
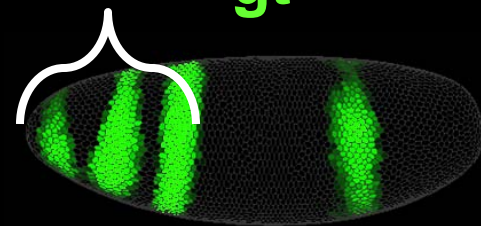
co-stained, n=60



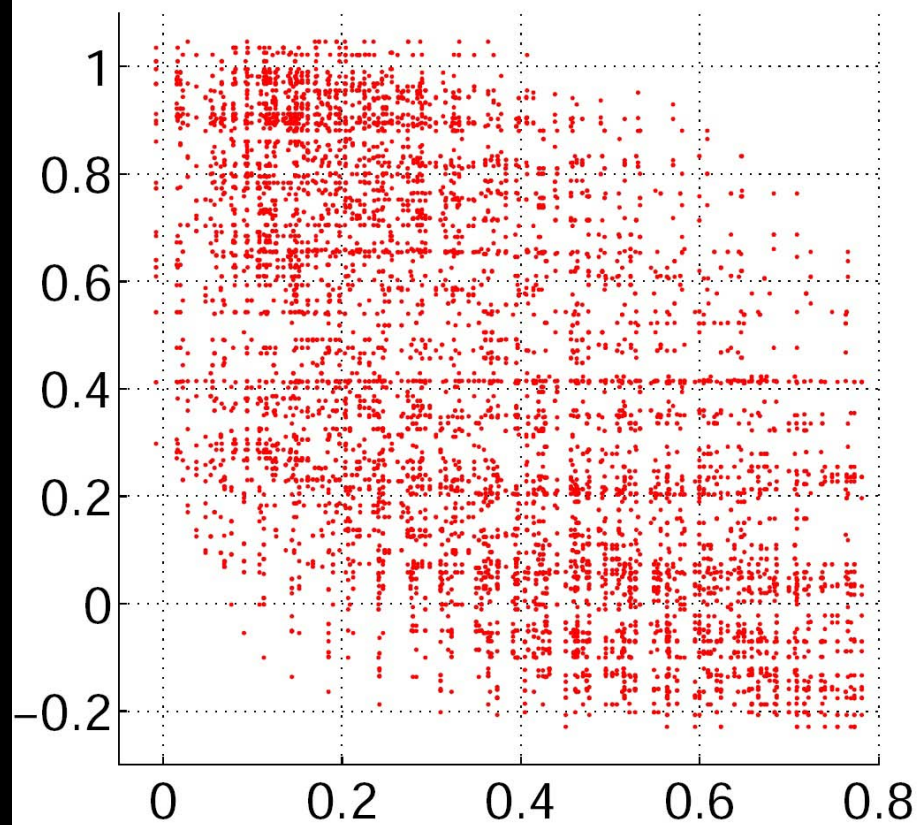
eve



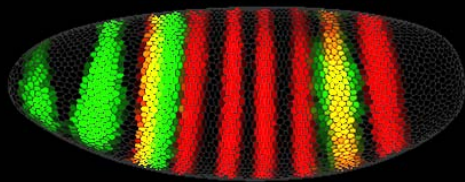
gt



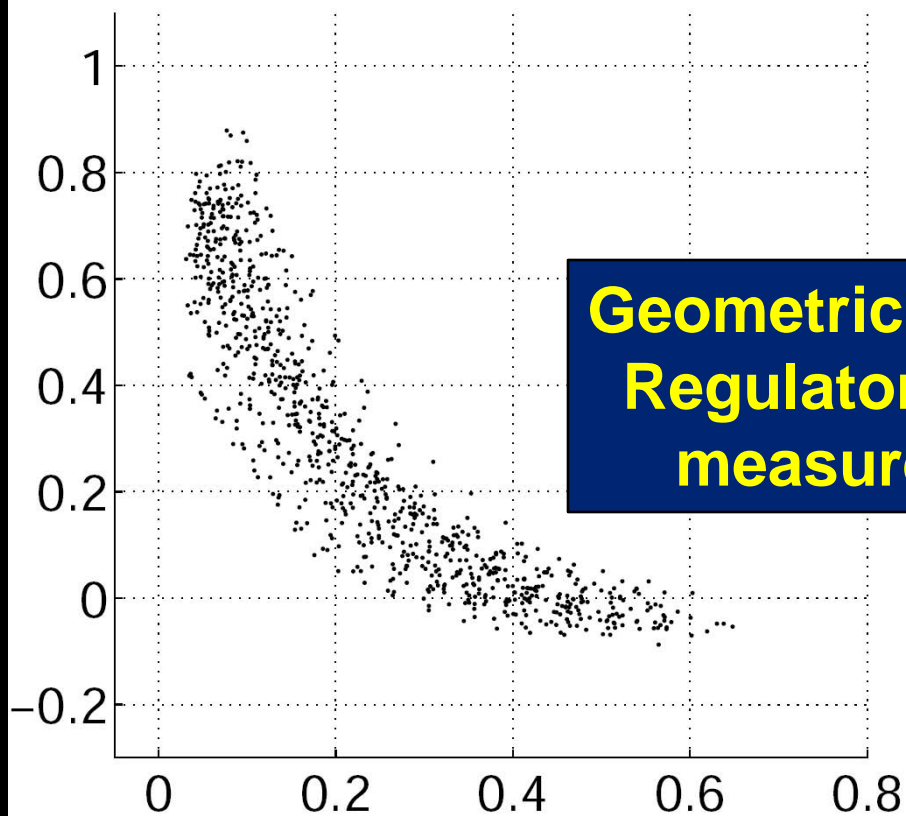
synthetic, coarse registration



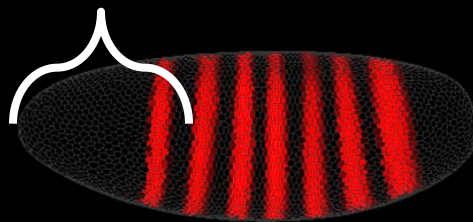
eve / gt



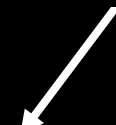
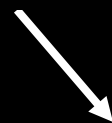
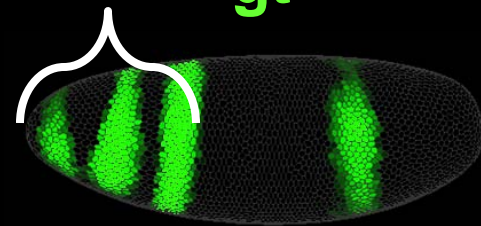
co-stained, n=60



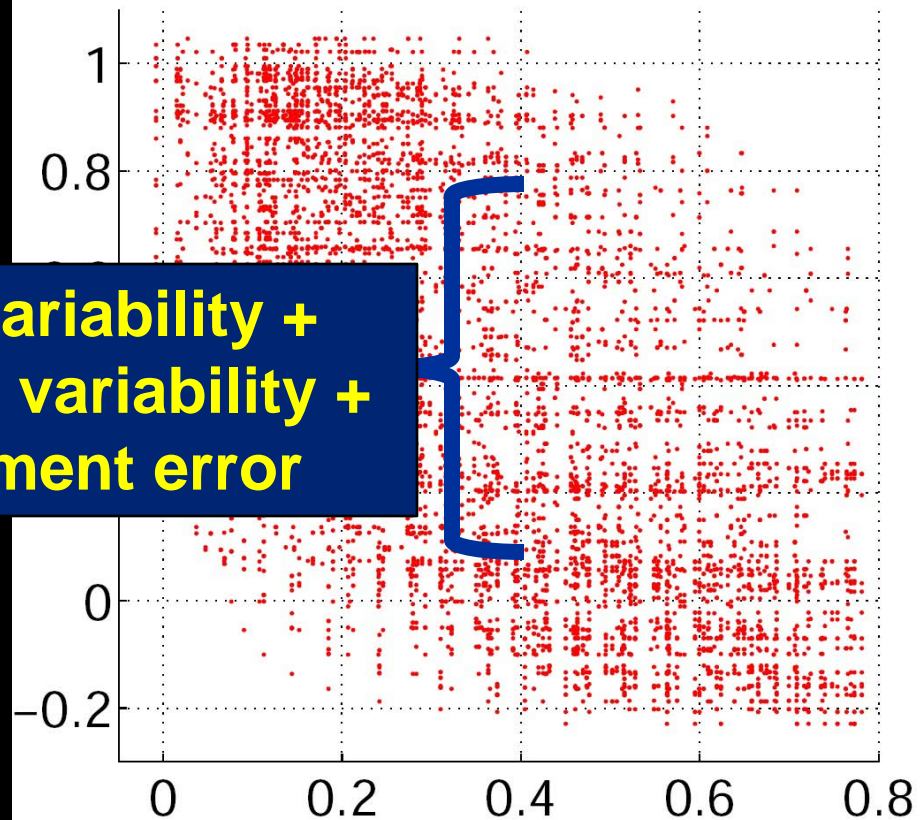
eve



gt

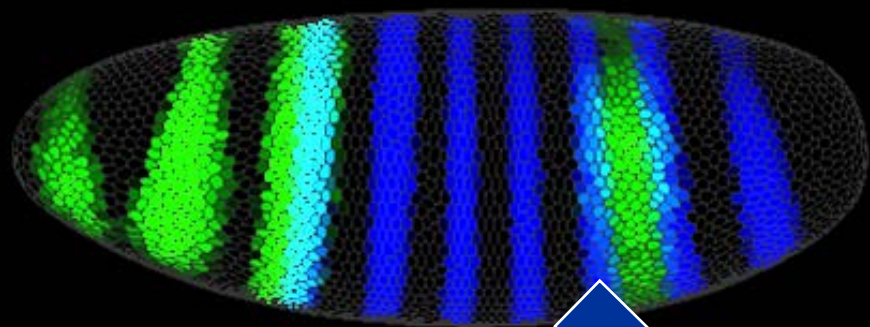


synthetic, coarse registration

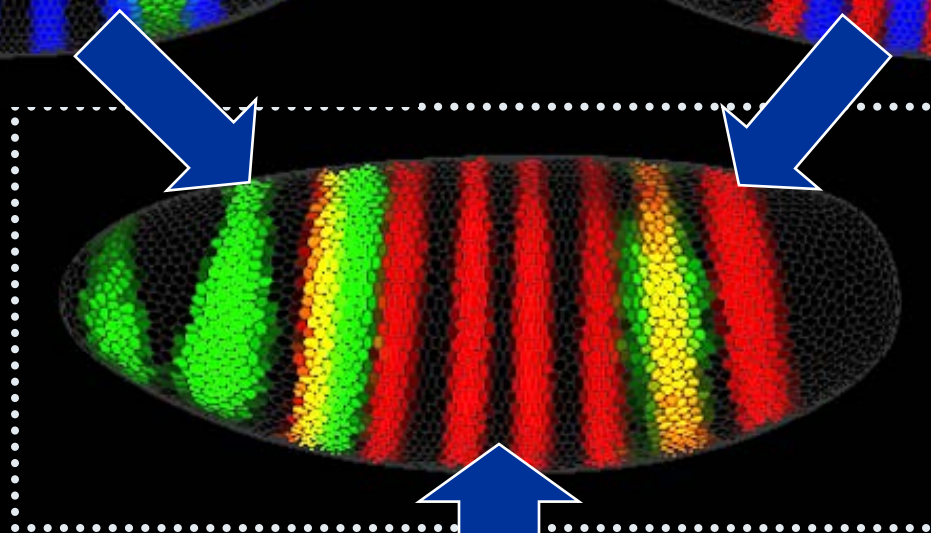
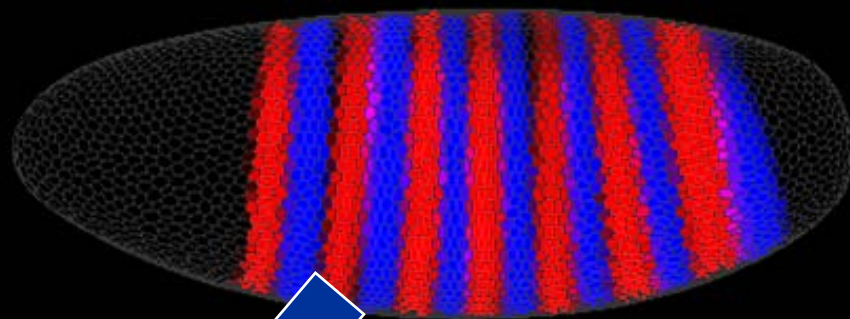


**Geometric variability +
Regulatory variability +
measurement error**

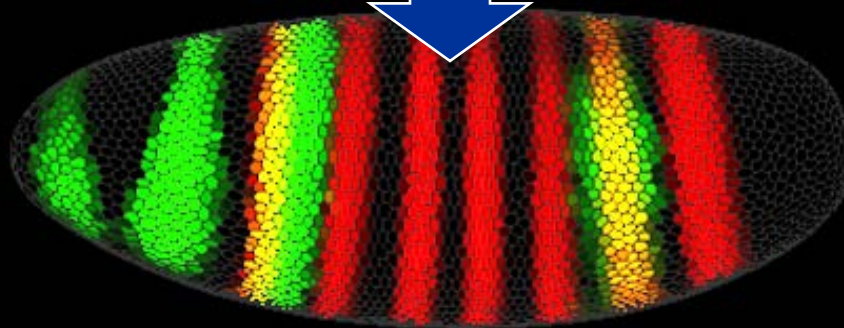
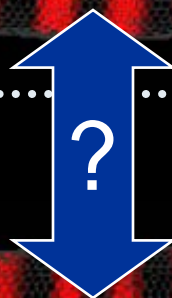
gt / ftz



ftz / eve

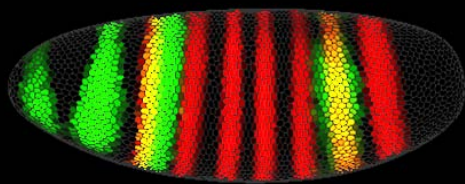


Virtual
Embryo

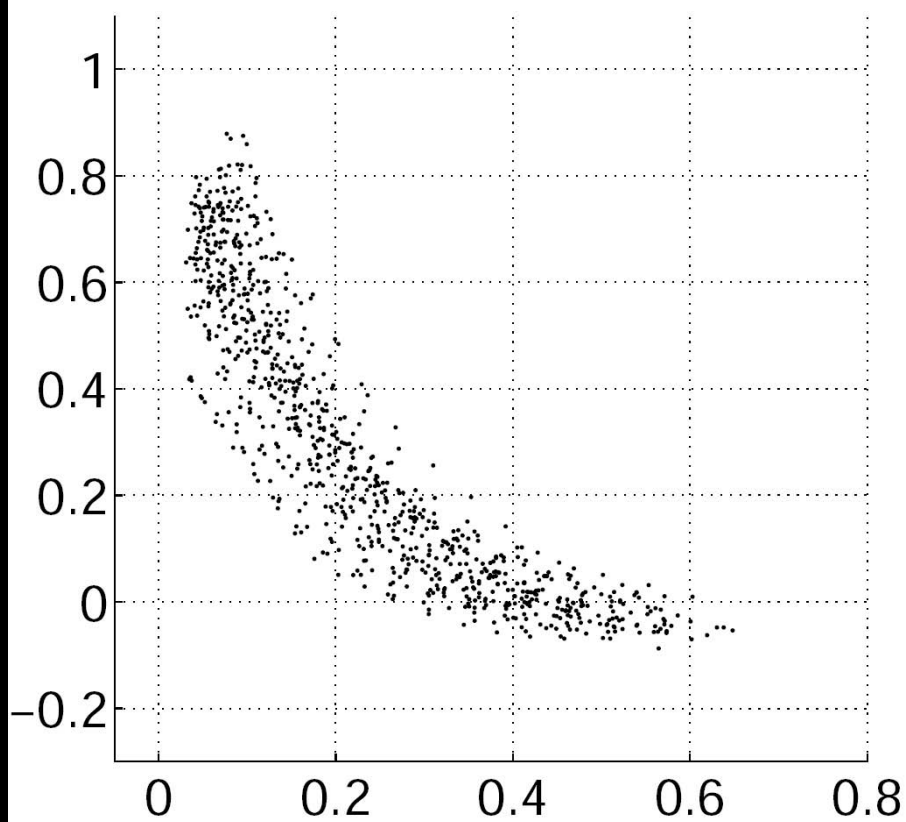


eve / gt
costain

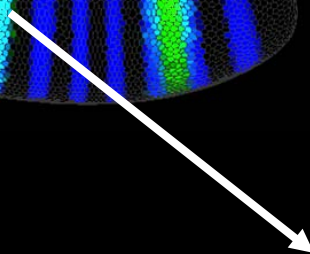
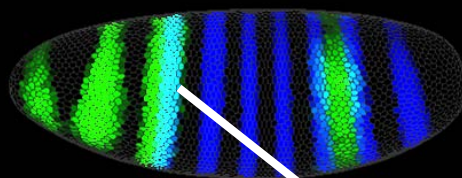
eve / gt



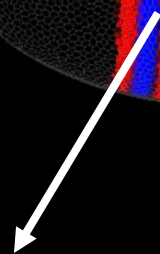
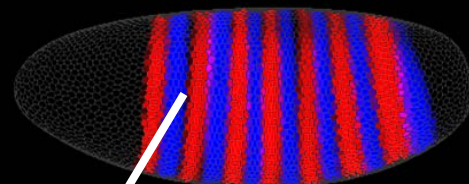
co-stained, n=60



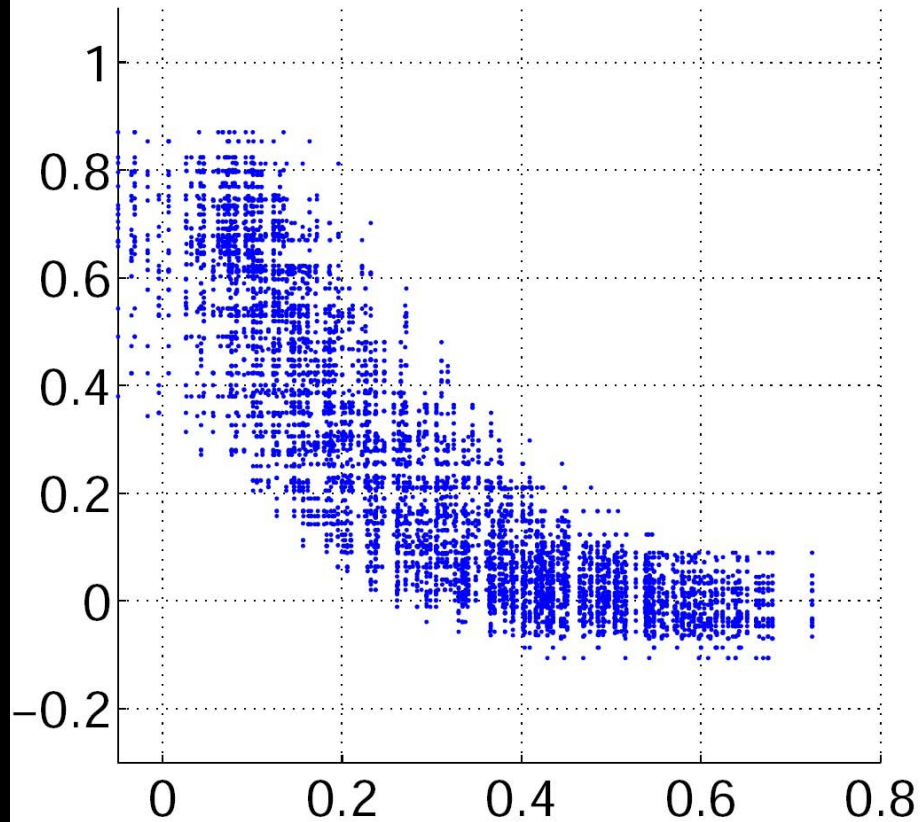
gt / ftz

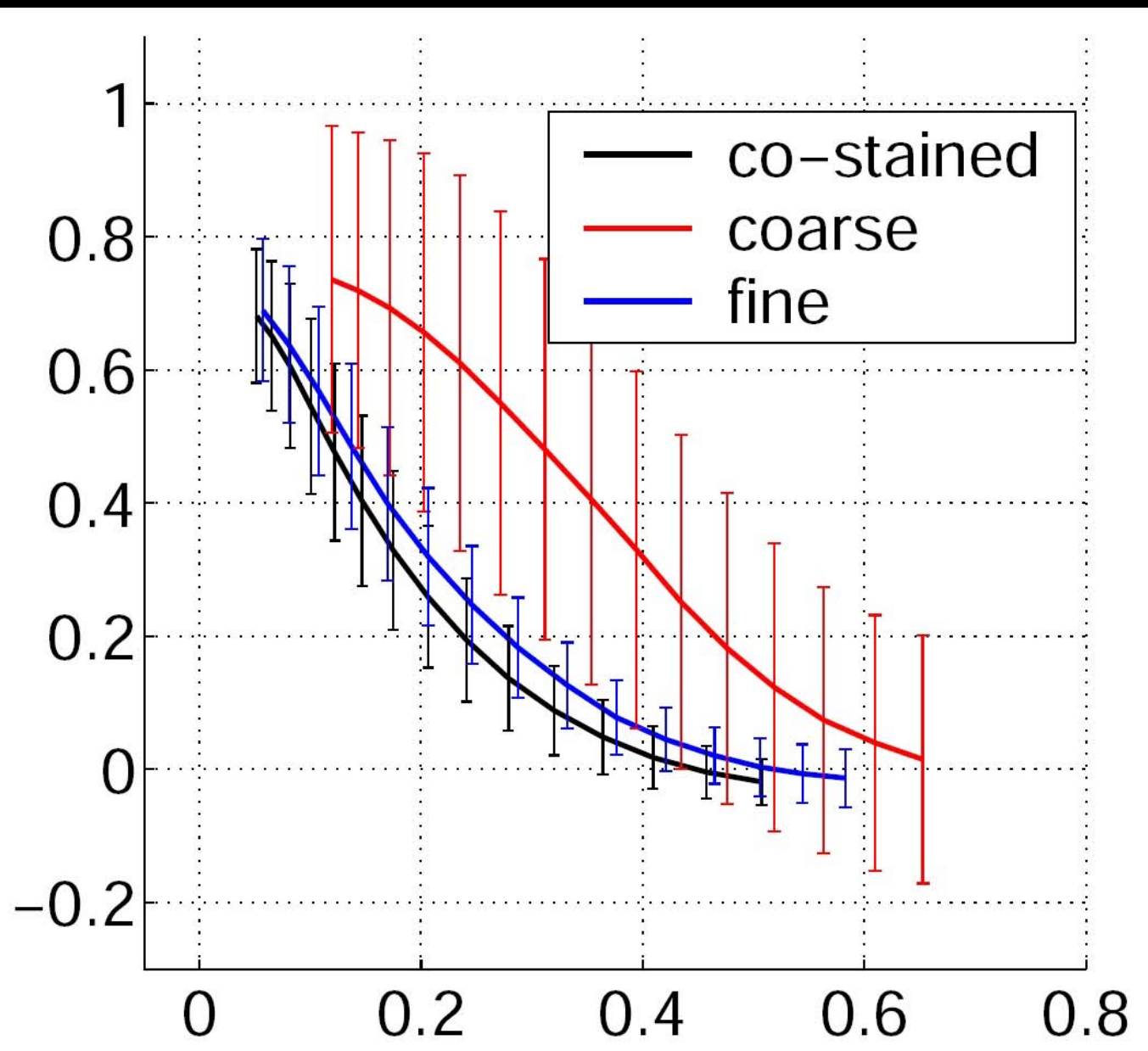


ftz / eve



synthetic, fine registration

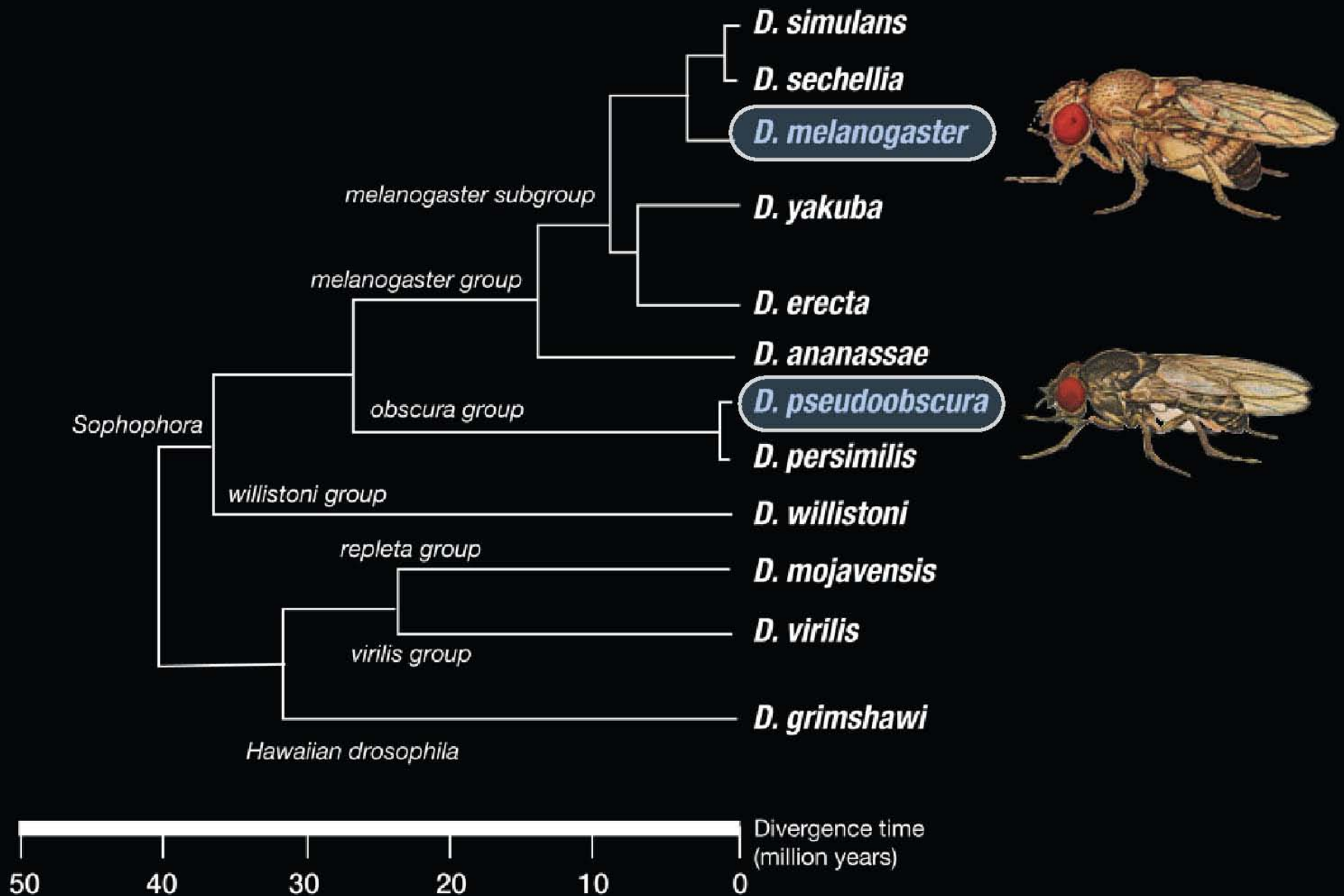




Outline

- Building a high resolution spatio-temporal atlas of gene expression in the *Drosophila* blastoderm
- Characterizing variation in expression patterns:
 - among individuals in a population
 - **between species**

There are exceptional genomic resources for *Drosophila*



Genetic variation between *D.pse* and *D.mel*



D. melanogaster

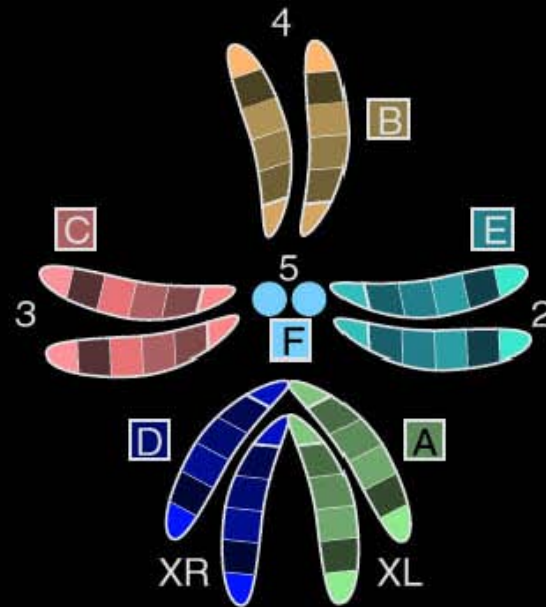
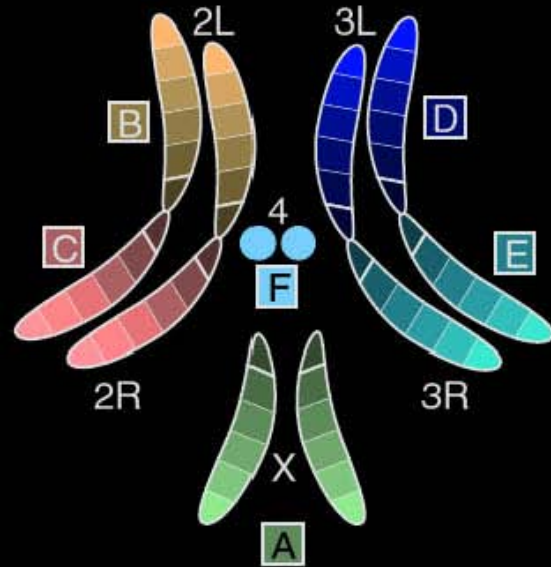


D. pseudoobscura

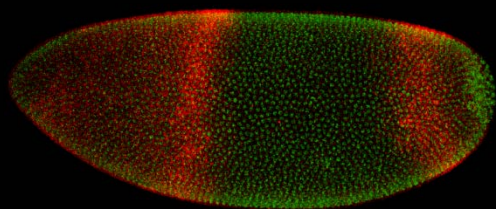
90% orthologous genes

70% mean identity in coding sequences

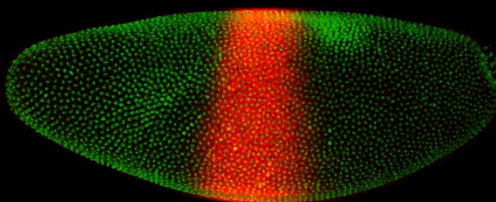
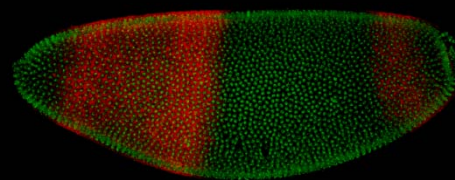
95-100% identity in TF DNA binding domains



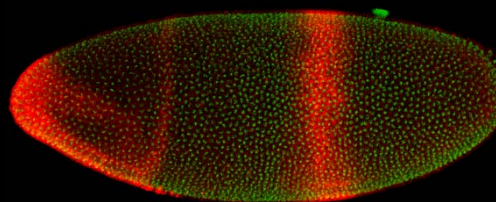
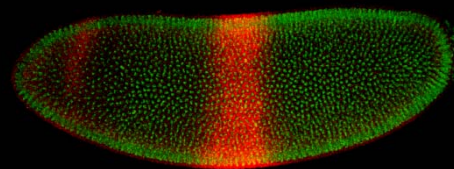
Variability in phenotype across different species



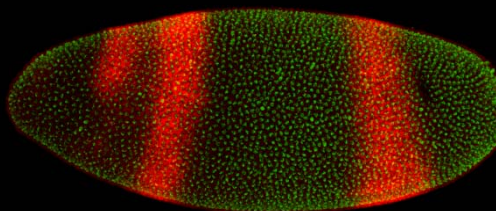
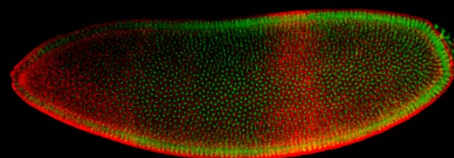
Hb



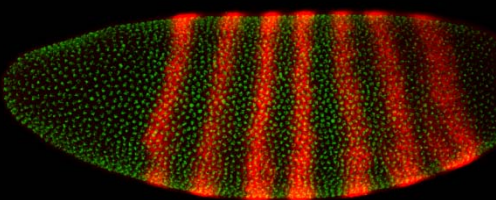
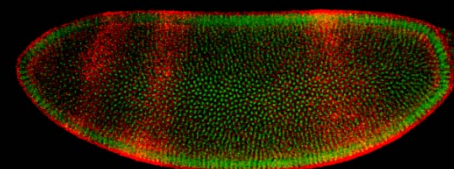
Kr



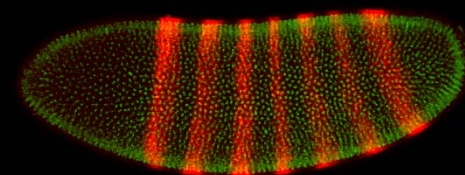
kni



Gt



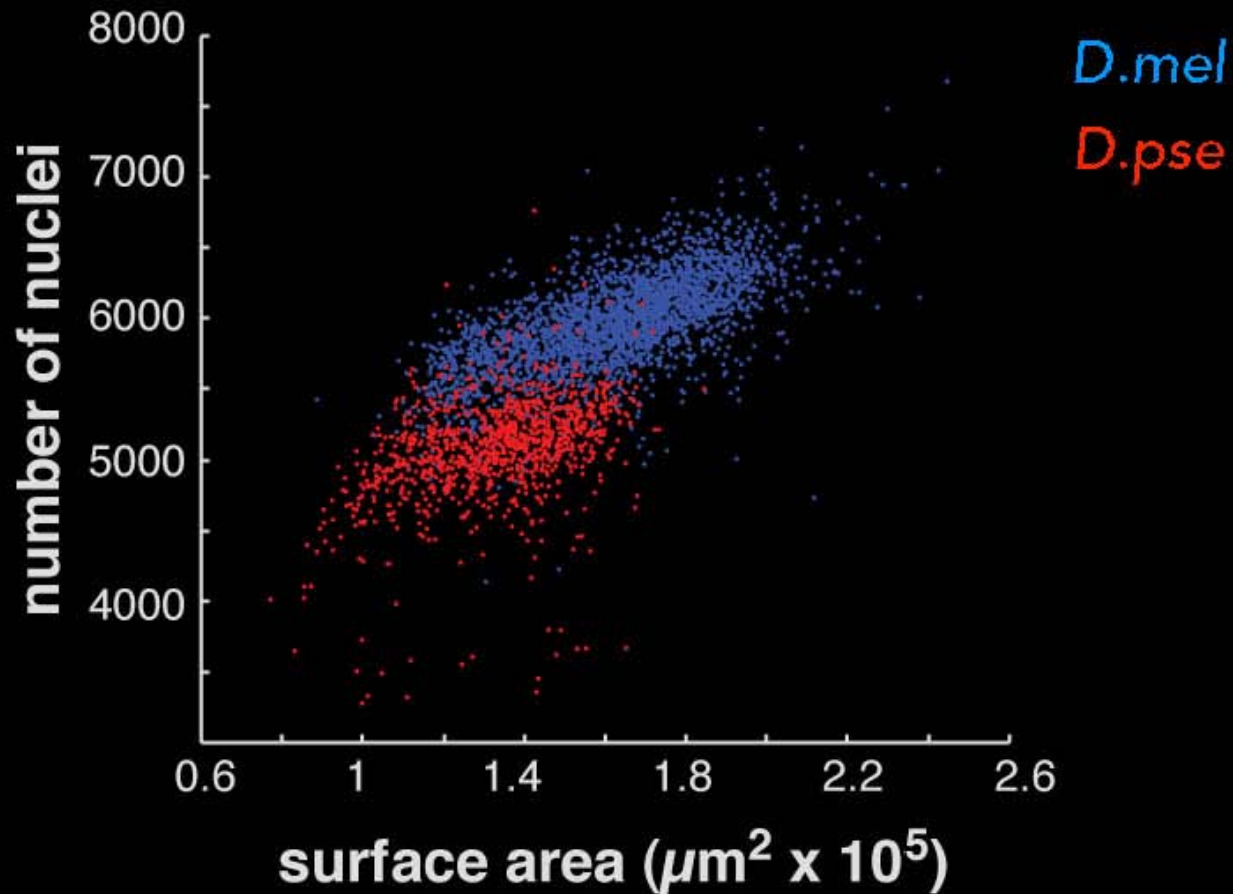
eve



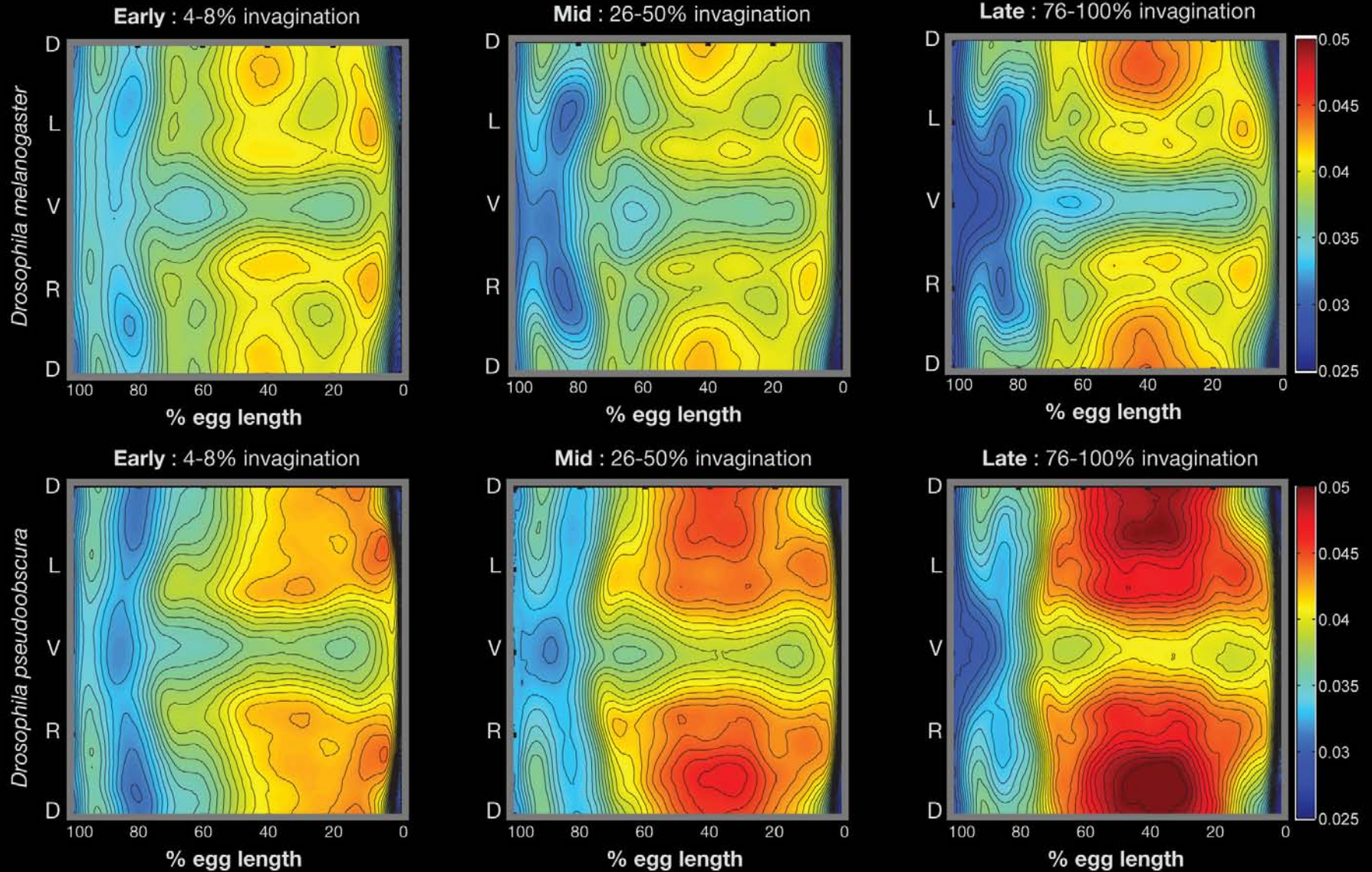
D. melanogaster

D. pseudoobscura

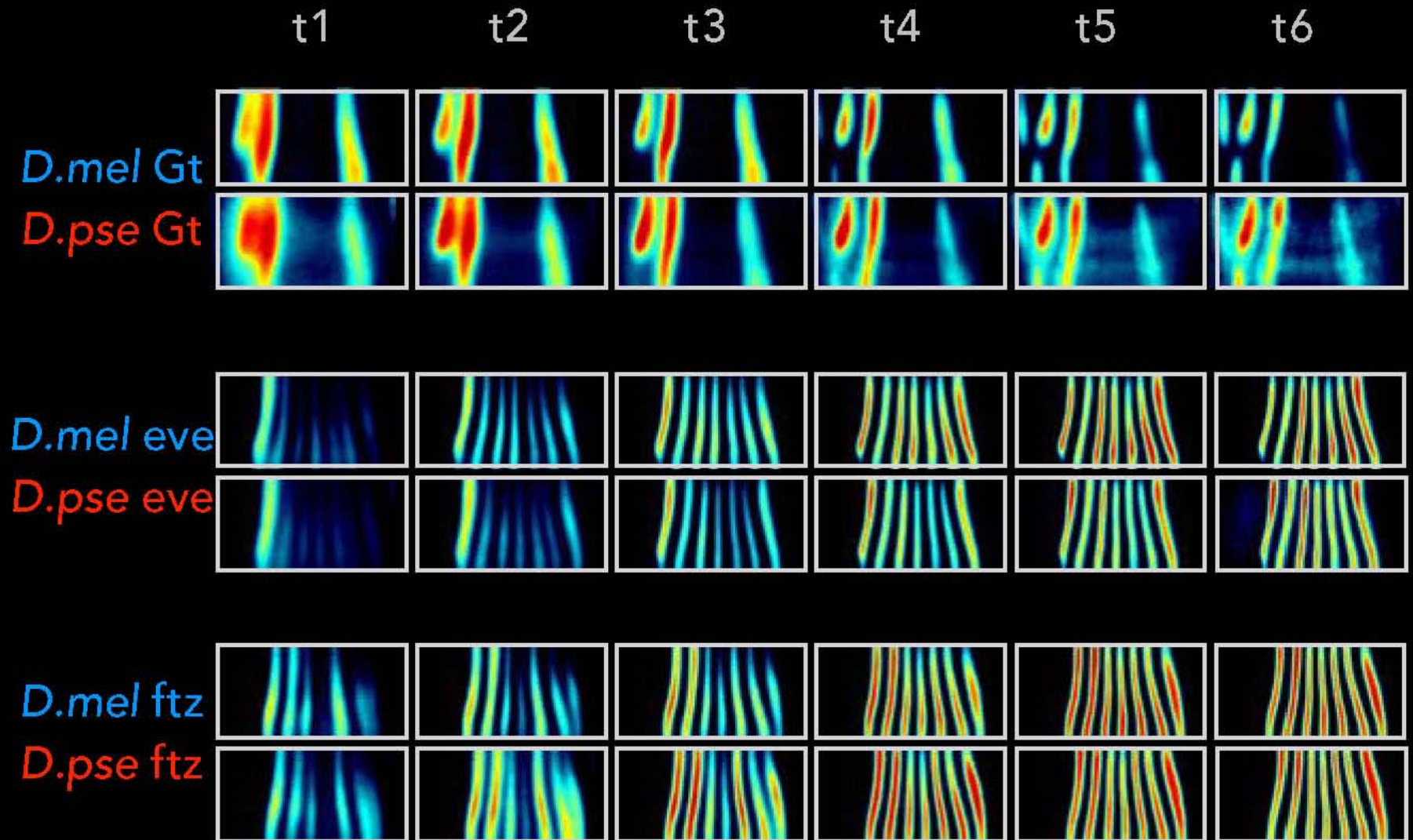
Blastoderm embryos from *D.mel* and *D.pse* have different numbers of nuclei.



Blastoderm embryos from *D.mel* and *D.pse* have distinct density patterns.



Some expression patterns are very similar in *D.mel* and *D.pse*.



Some expression patterns are distinct in *D.mel* and *D.pse*.

t1

t2

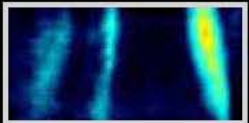
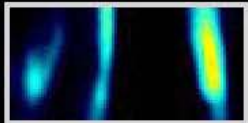
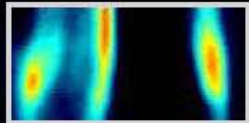
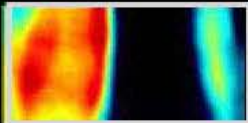
t3

t4

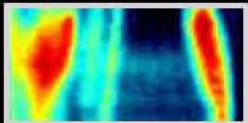
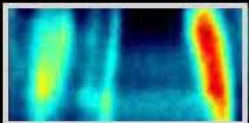
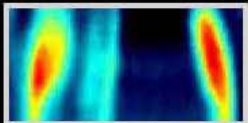
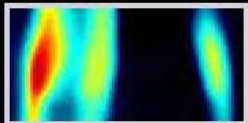
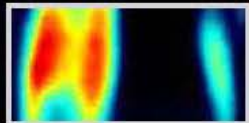
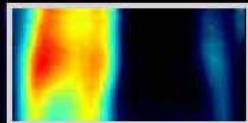
t5

t6

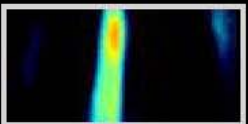
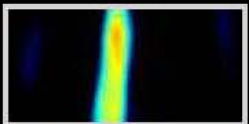
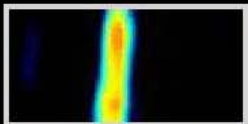
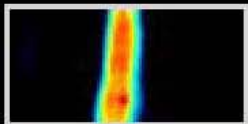
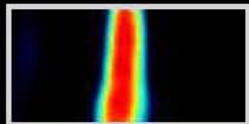
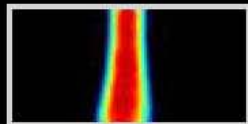
D.mel Hb



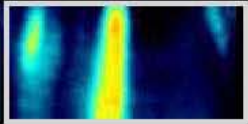
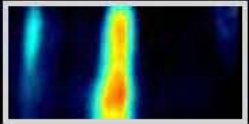
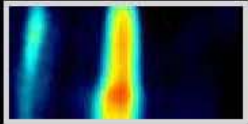
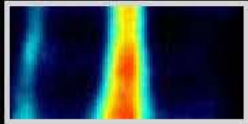
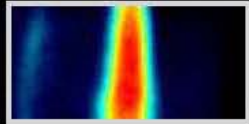
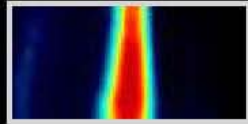
D.pse Hb



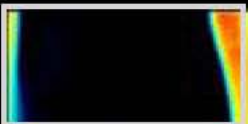
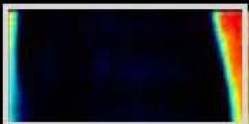
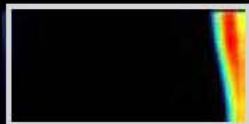
D.mel Kr



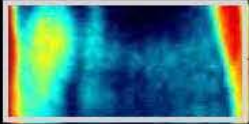
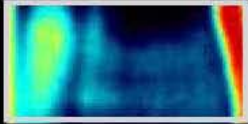
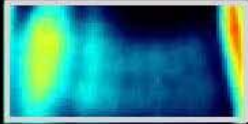
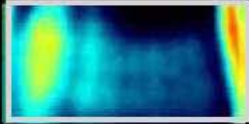
D.pse Kr



D.mel fkh

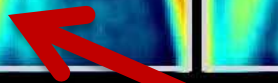
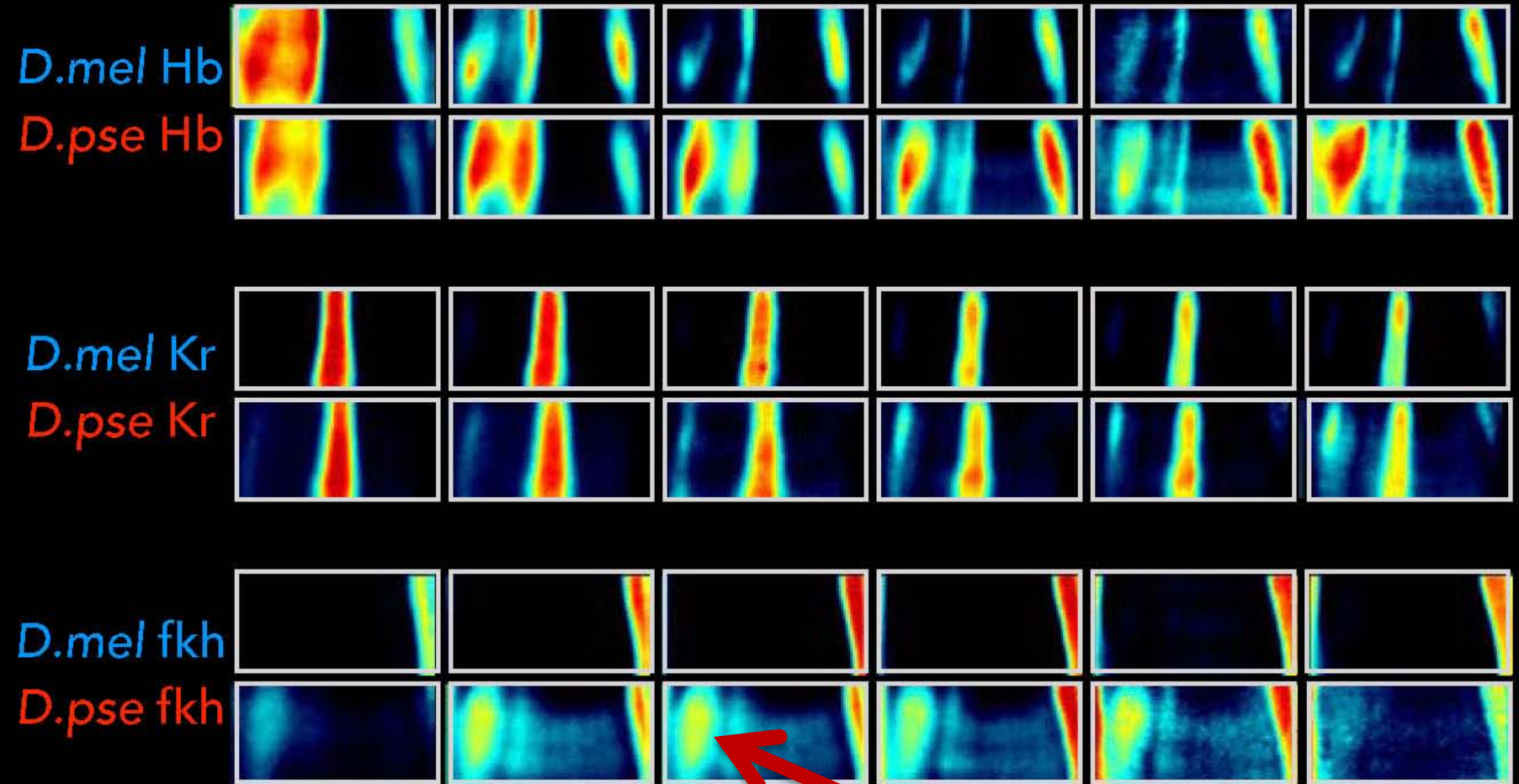


D.pse fkh



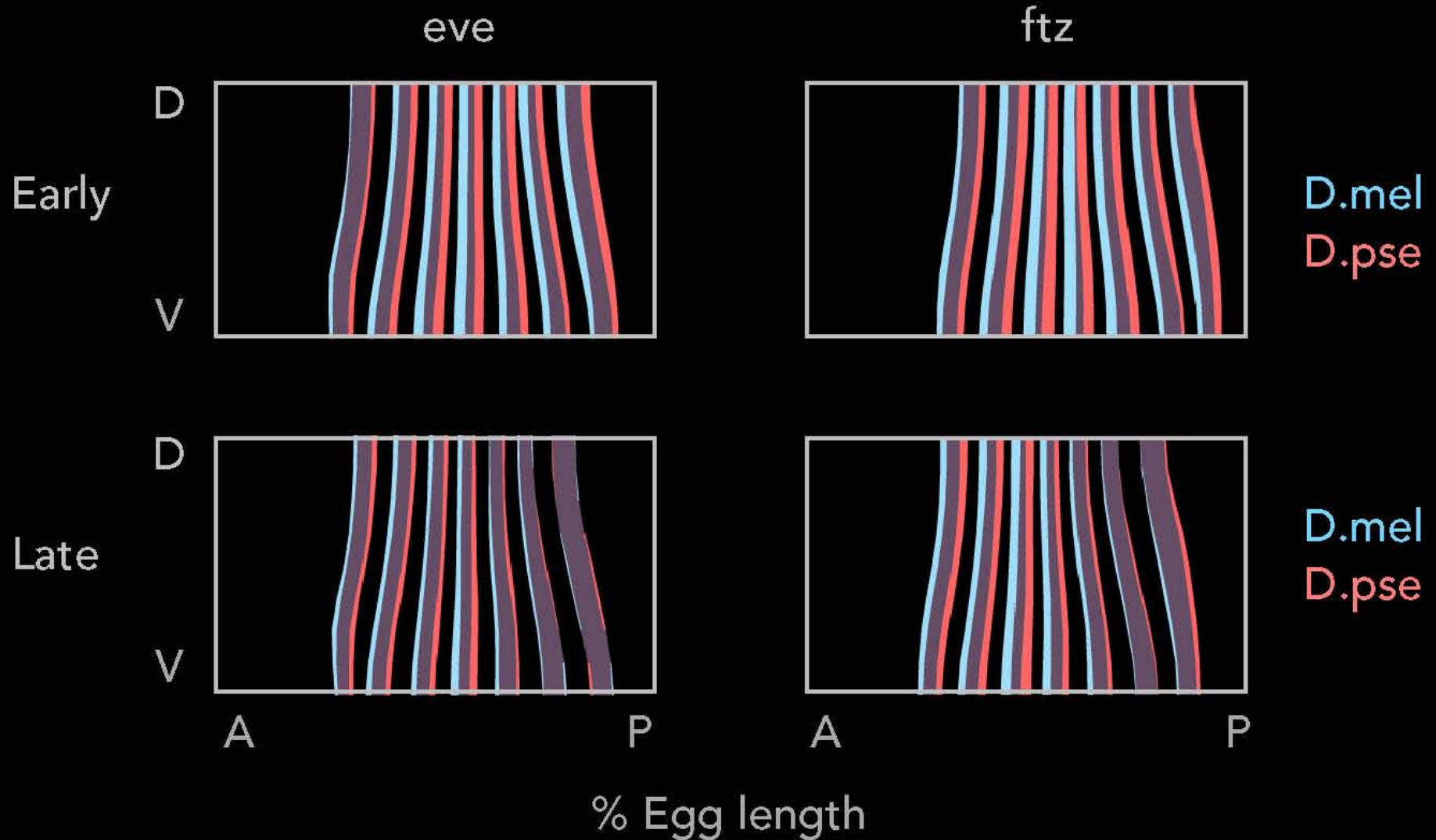
Some expression patterns are distinct in *D.mel* and *D.pse*.

t1 t2 t3 t4 t5 t6



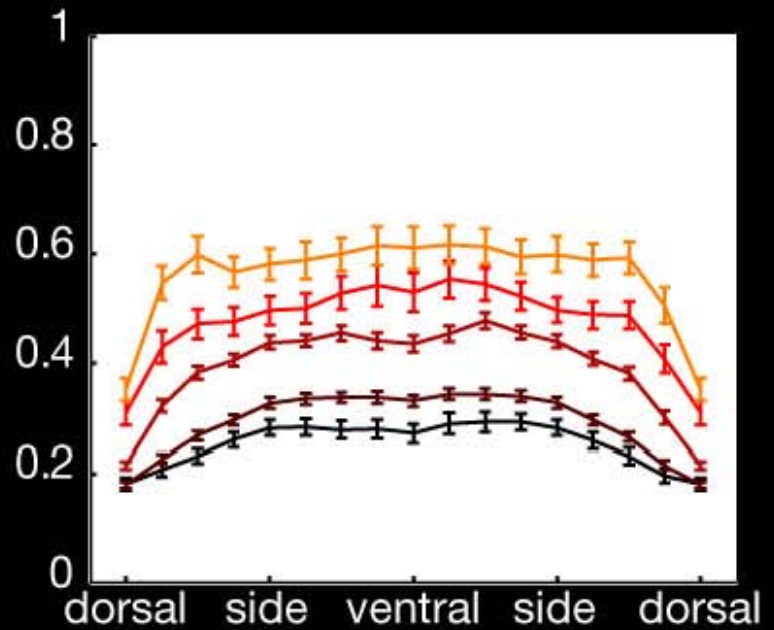
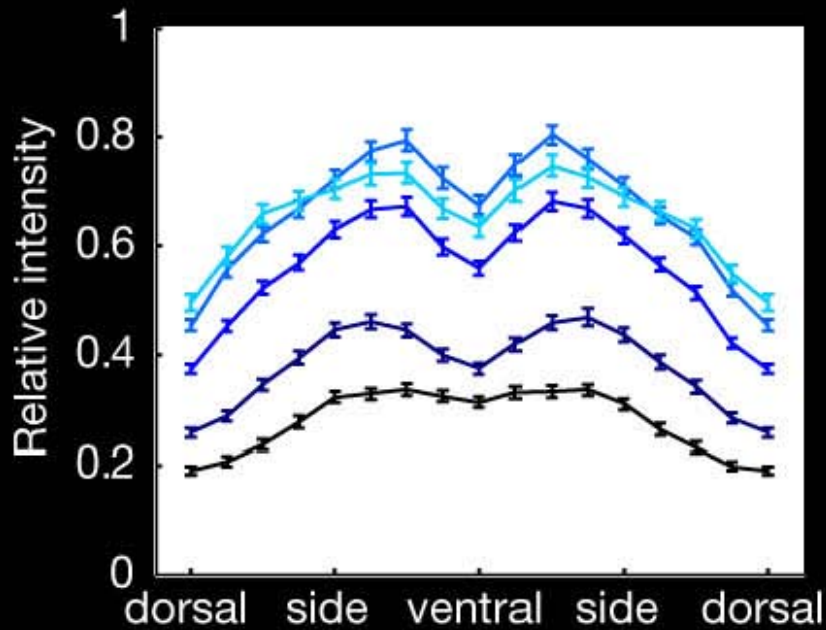
Sadly this is probably an artifact

Even for similar patterns, there are often quantitative differences in spatial location.

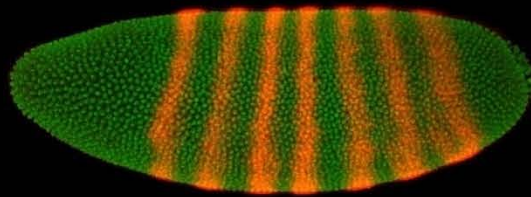


Even for similar patterns, there are often quantitative differences, in intensity profiles.

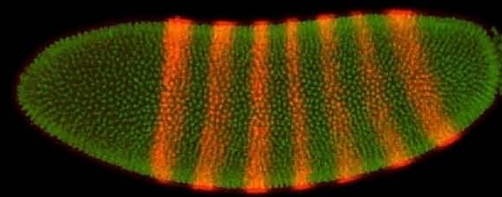
stripe 6



Even for similar patterns, there are often quantitative differences in cell number.

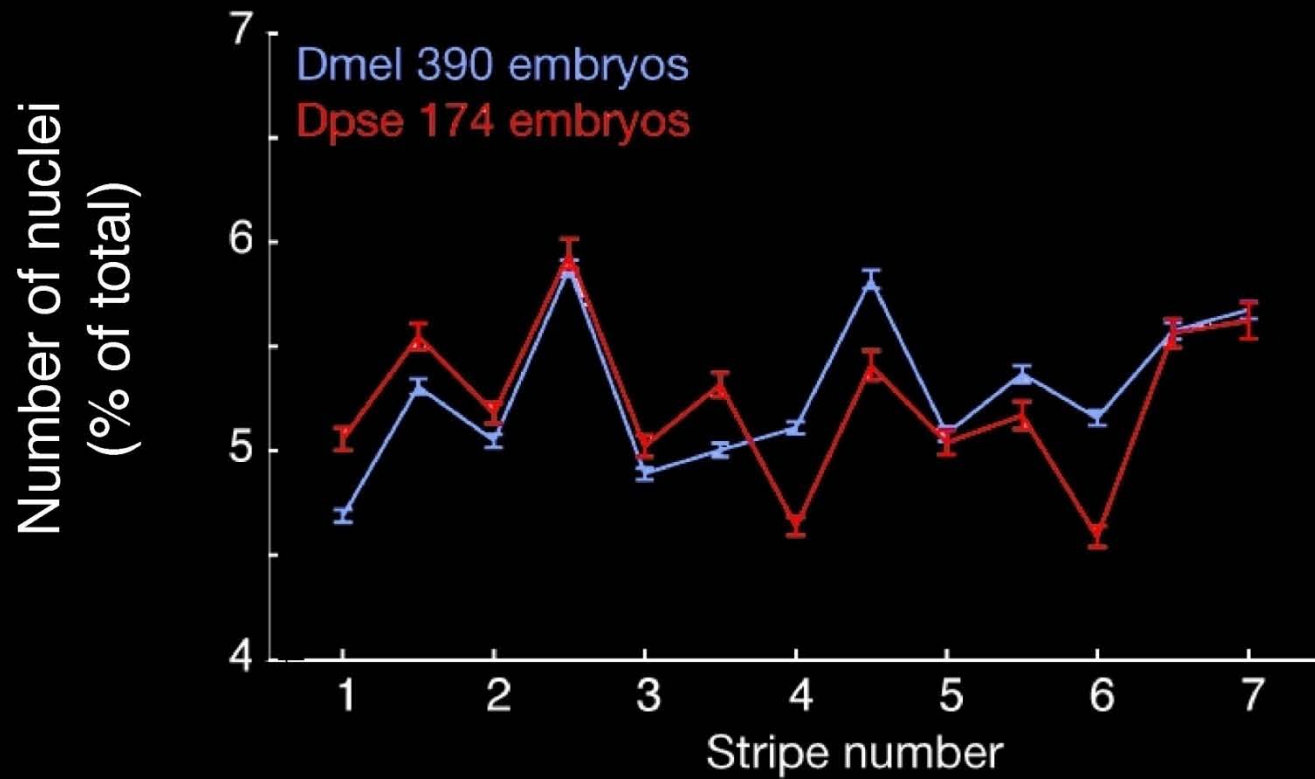


D. melanogaster

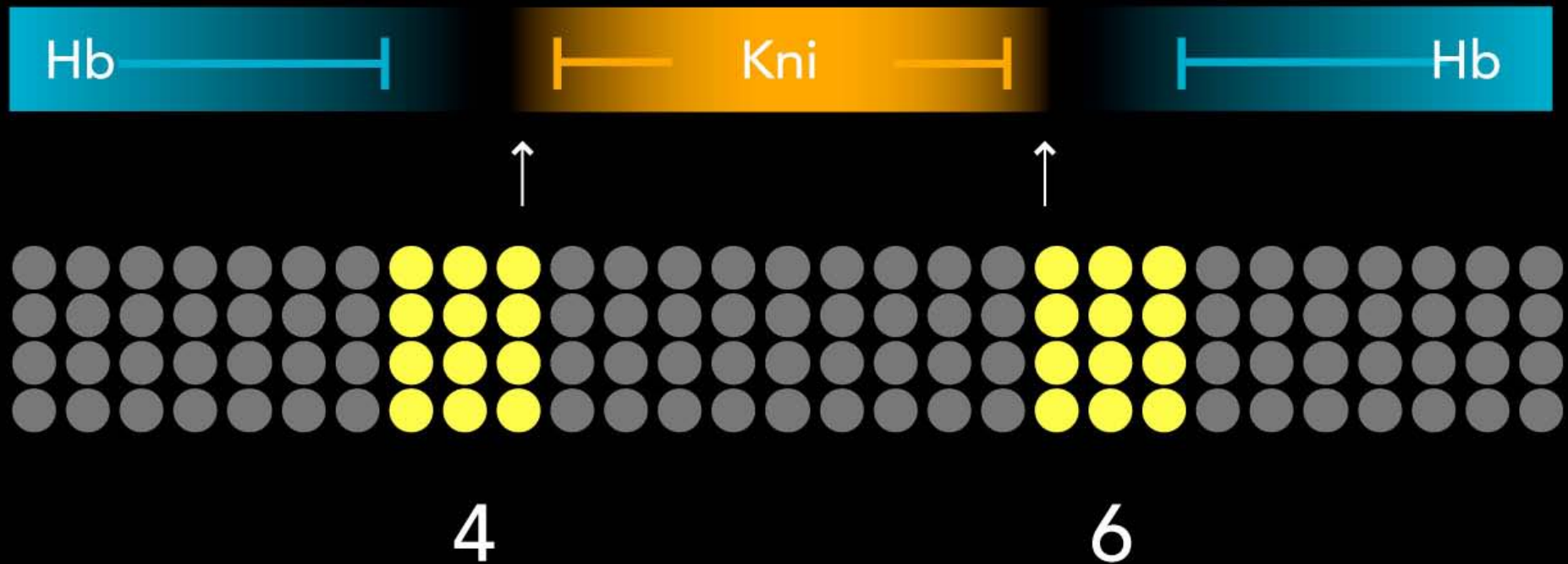


D. pseudoobscura

26-50% invagination

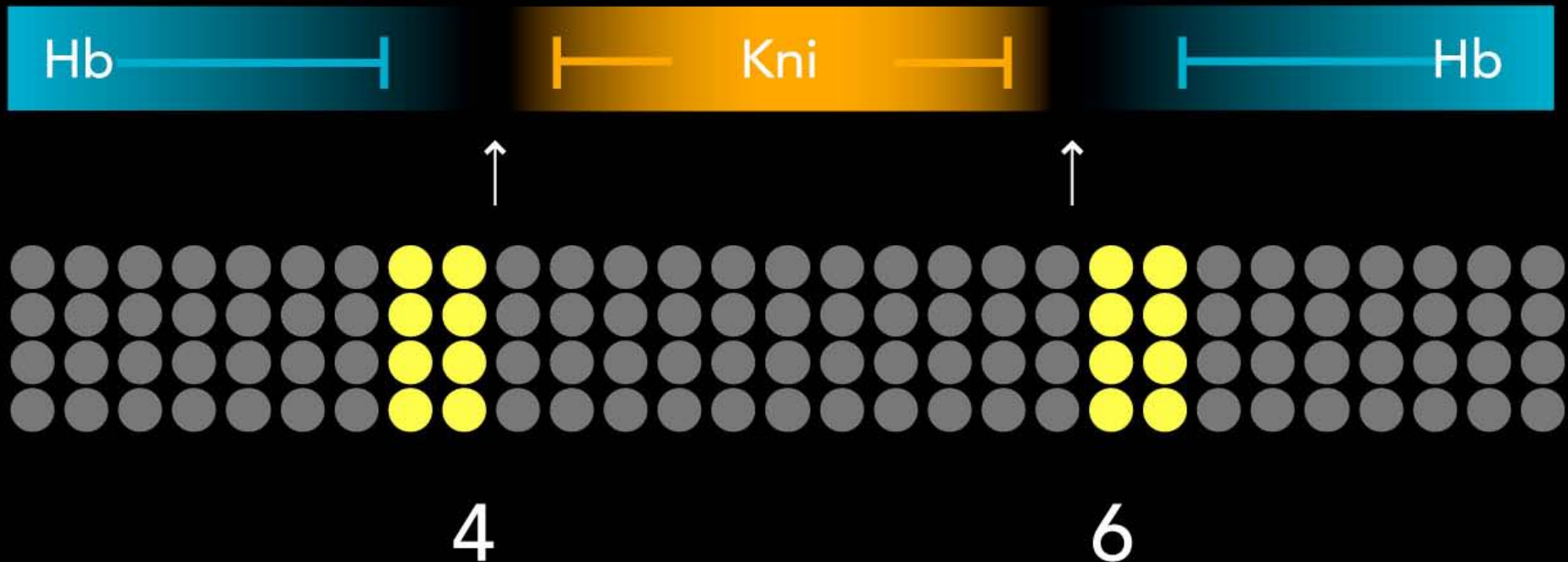


Kni and Hb form the boundaries of eve stripe 4+6



Kni and Hb form the boundaries of eve stripe 4+6

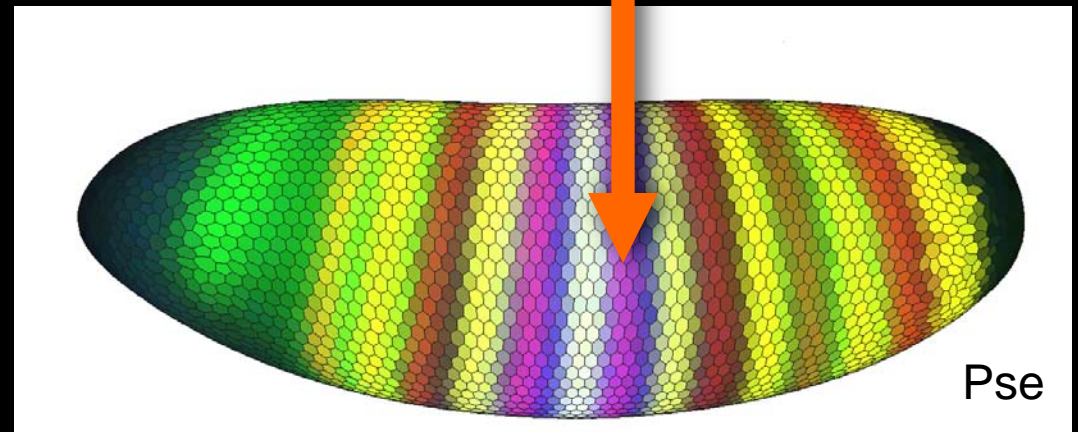
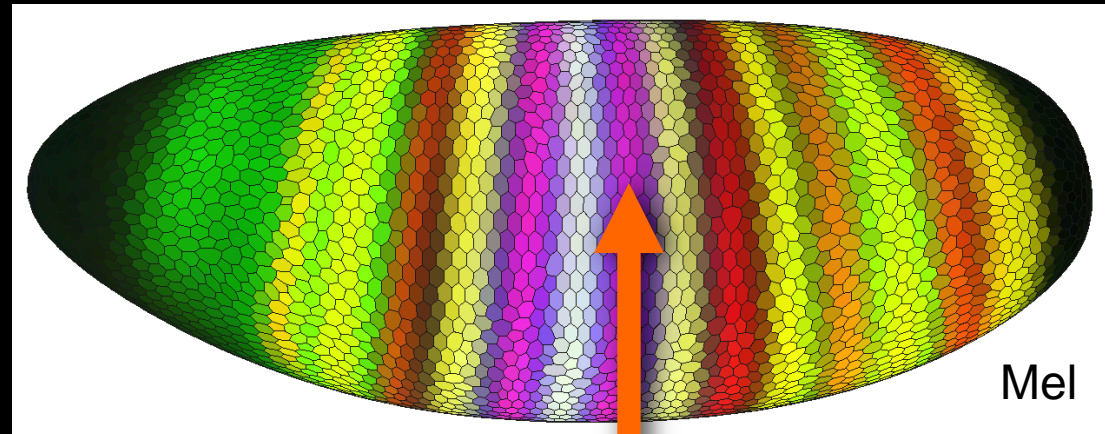
Decreased number of cells in stripe 4+6 may be explained by addition of predicted Kni binding in *D. pse* regulatory sequence



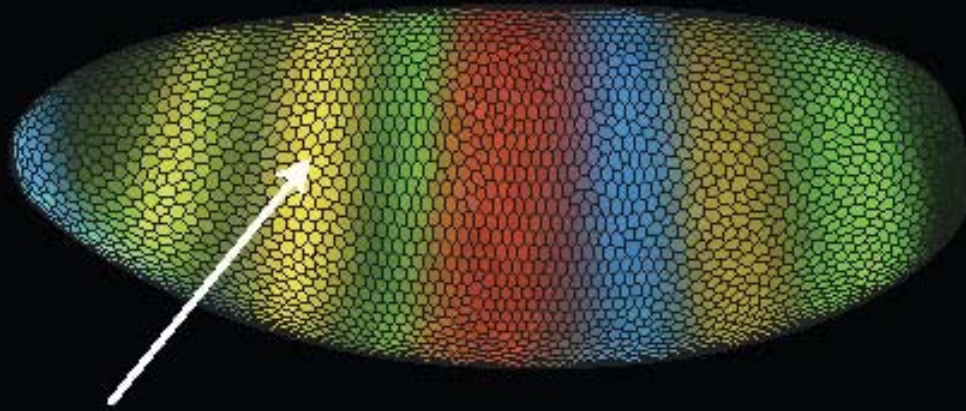
Global “Phenotype Alignment”

mel TGTG----CGTCAGC----

pse TGTGACTGCG-CTGCCTGC

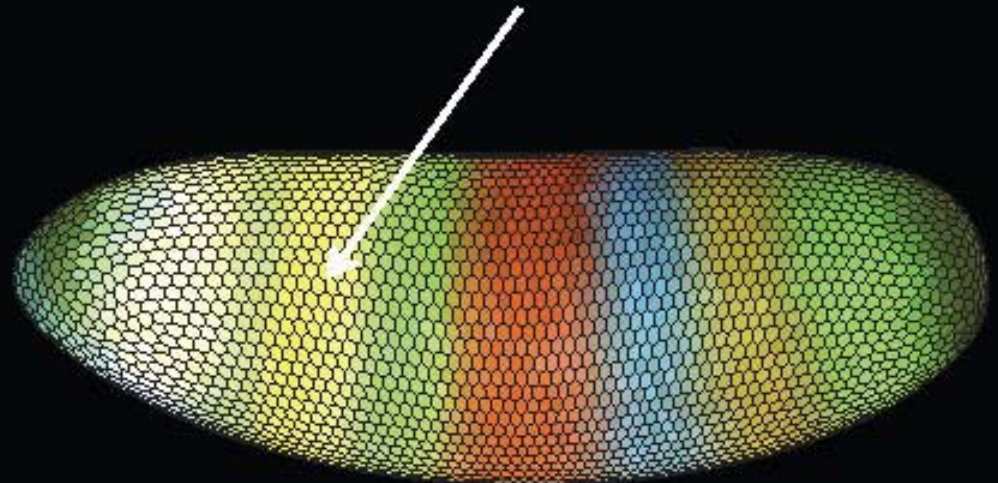


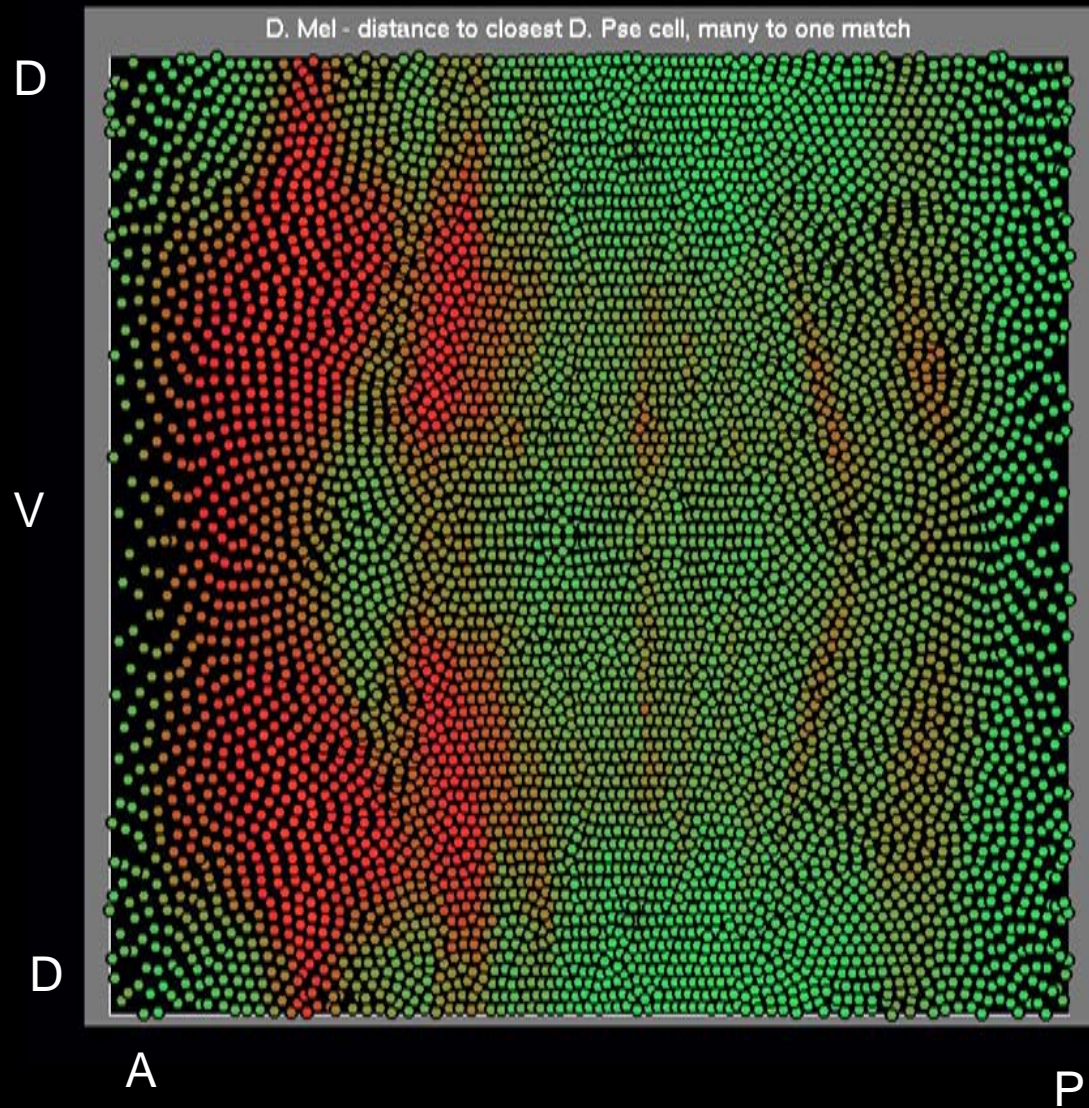
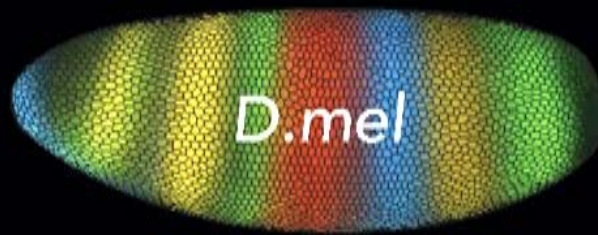
We want to find corresponding cells in the two embryo models.



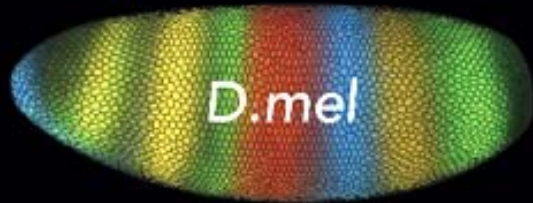
ID2
Gene 1 @ t1, t2, t3, t4, t5, t6
Gene 2 @ t1, t2, t3, t4, t5, t6
.
.
.
Gene 8 @ t1, t2, t3, t4, t5, t6

ID1
Gene 1 @ t1, t2, t3, t4, t5, t6
Gene 2 @ t1, t2, t3, t4, t5, t6
.
.
.
Gene 8 @ t1, t2, t3, t4, t5, t6

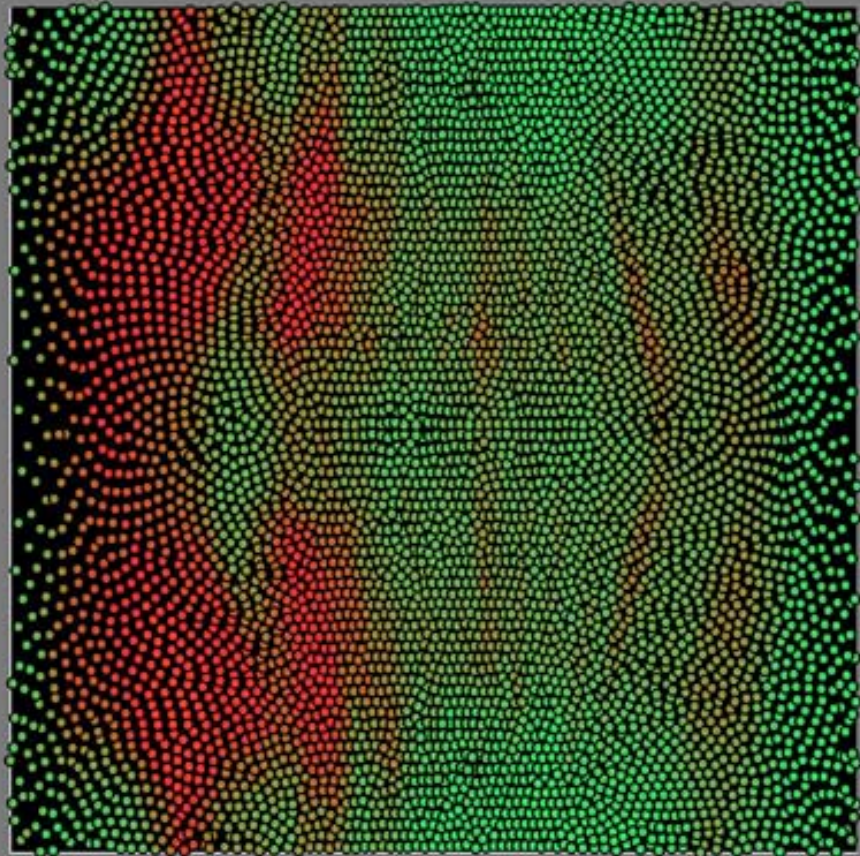




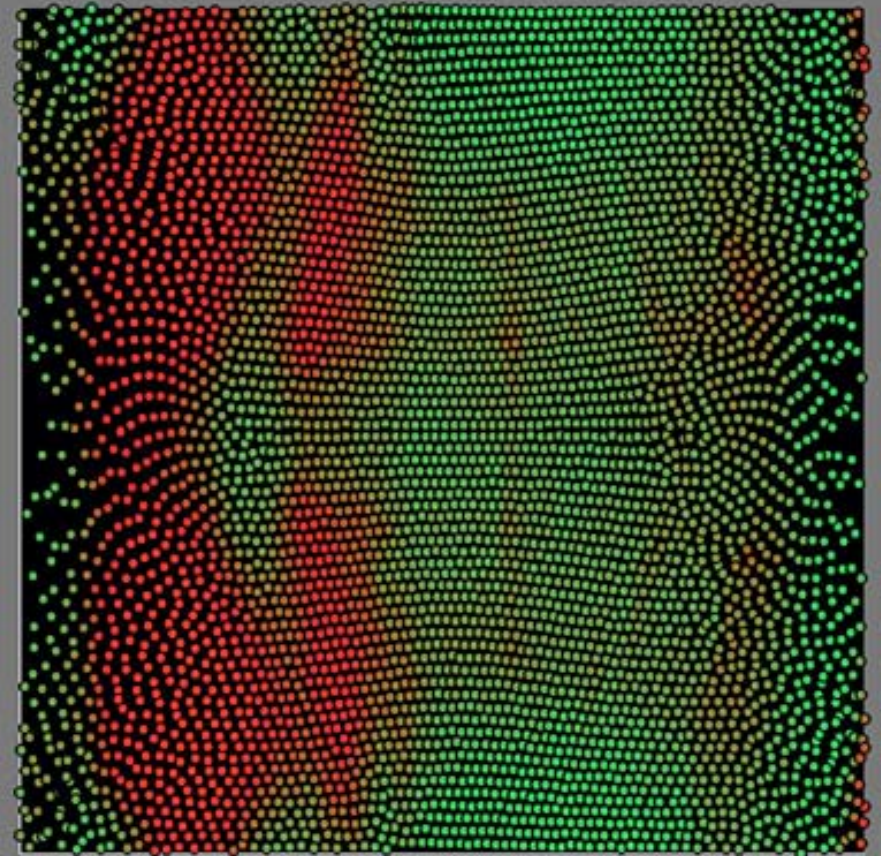
Expression differences create new transcriptional niches.



D. Mel - distance to closest D. Pse cell, many to one match



D. Pse - distance to closest D. Mel cell, many to one match



Conclusions

- Spatial registration and correspondence provides a basis for analyzing variations in biological form both within and between species
- Phenotype alignment provides a global analysis of variations in patterning
- Working to experimentally identify molecular bases for differences and develop systematic models of regulation



Angela DePace
Soile Keränen
Cris Luengo
Lisa Simirenko
Gunther Weber
Oliver Rübhel
Min-Yu Huang
Bernd Hamann
Damir Sudar
David Knowles
Jitendra Malik
Mark Biggin
Mike Eisen



<http://bdtnp.lbl.gov>