

An introduction to systems thinking

William Bialek

Joseph Henry Laboratories of Physics, and
Lewis-Sigler Institute for Integrative Genomics
Princeton University

Initiative for the Theoretical Sciences
The Graduate Center, City University of New York

Enumerating the molecular building blocks of life is one of the great triumphs of the 20th century

BUT:

the reductionist hypothesis does not by any means imply a “constructionist” one



More Is Different

Broken symmetry and the nature of the hierarchical structure of science.

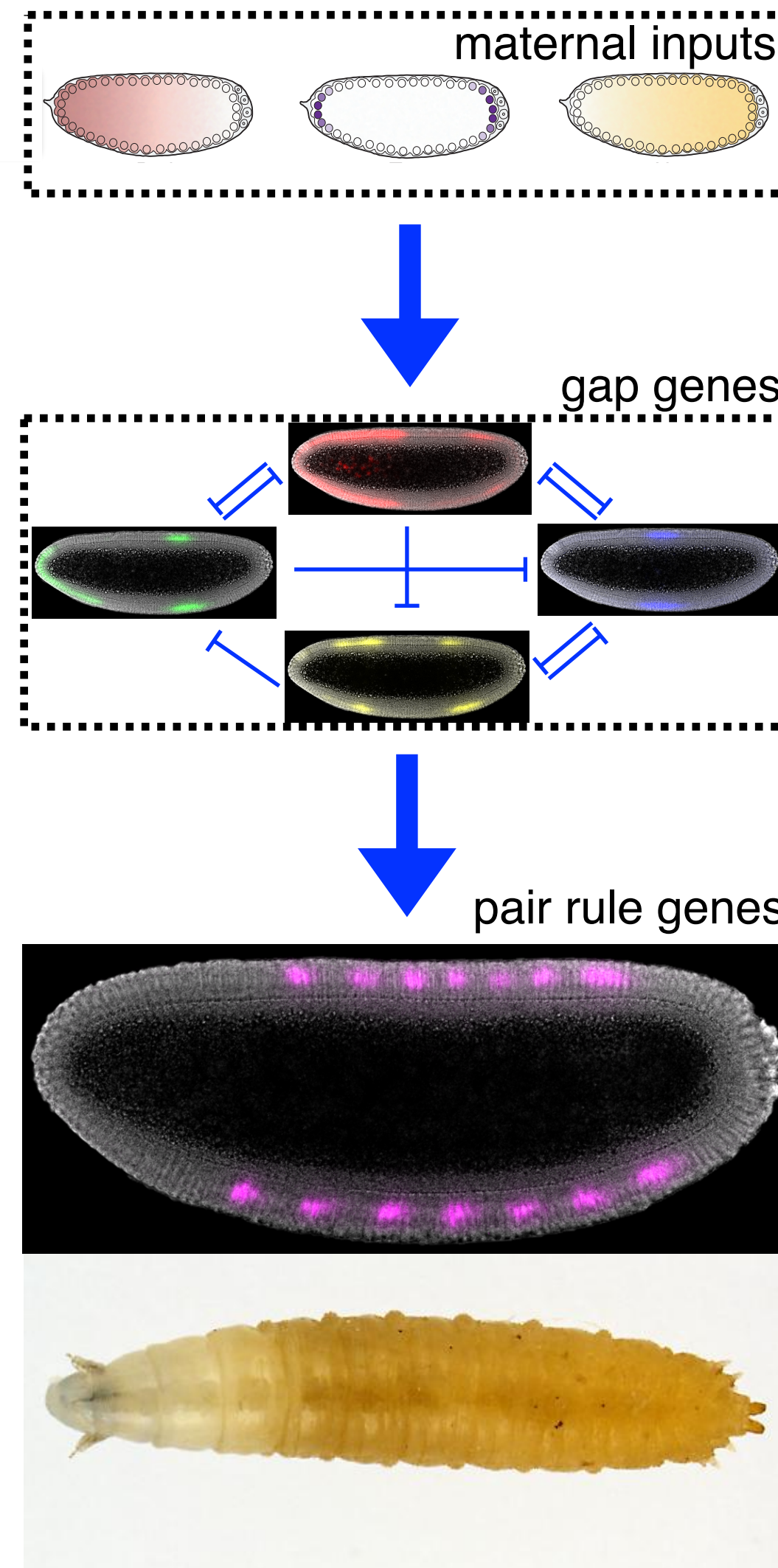
P. W. Anderson

(1923-2020)

4 August 1972, Volume 177, Number 4047

SCIENCE

Anderson's concerns were conceptual, but there are practical issues as well ...



e.g., If we identify all the relevant genes and their interactions in a network, what can we predict?

(concrete example from fly development)

What's the problem?

Every arrow has (at least) two parameters attached to it

Convergence of arrows hides an arbitrary function

4 genes, 3 inputs = 28 possible arrows

> 50 parameters

What should we do about all these parameters?
We need to add systems-level ideas to the reductionist description.

Give up, it really is that complicated.
More subtly, “quantitative biology” really is about learning complex models from limited data.

The only hope for a real theory is evolution.
Not about how things are, but about how they are related, historically.

Not all the parameters matter.
“Robustness” (?), or “sloppy models.”

Lots of parameters at the micro scale, but emergent functional phenomena are simpler.
Maybe this works as it does in condensed matter, via the renormalization group.

There is some principle that selects the parameters.
Random parameter choices don’t “work.”

**In many ways, this provides a classification scheme for
theoretical physicists’ approaches to the phenomena of life.**

Candidate principle: transmit as much relevant information as possible with limited resources.

Simplest case: $x \rightarrow y$
input $\longrightarrow$ output How much “information” does the output provide about the input?

Shannon (1948): Entropy is the unique measure of available information consistent with simple postulates.

Entropy reduction = Information gain

$P(y|x)$ describes noisy input/output transformation
 $Q(x)$ distribution of inputs
 $P(x|y) = \frac{P(y|x)Q(x)}{P(y)}$ Bayes' rule

average information gain $I = \Delta S = \left[- \int dx Q(x) \log Q(x) \right] - \int dy P(y) \left[- \int dx P(x|y) \log P(x|y) \right]$

entropy of inputs
entropy of inputs, conditional on observing a particular output
average over outputs $P(y) = \int dx P(y|x) Q(x)$

This “mutual information” is symmetric.

With fixed input/output relation and noise, information can be maximized by adjusting the distribution of inputs (“achieving capacity”).

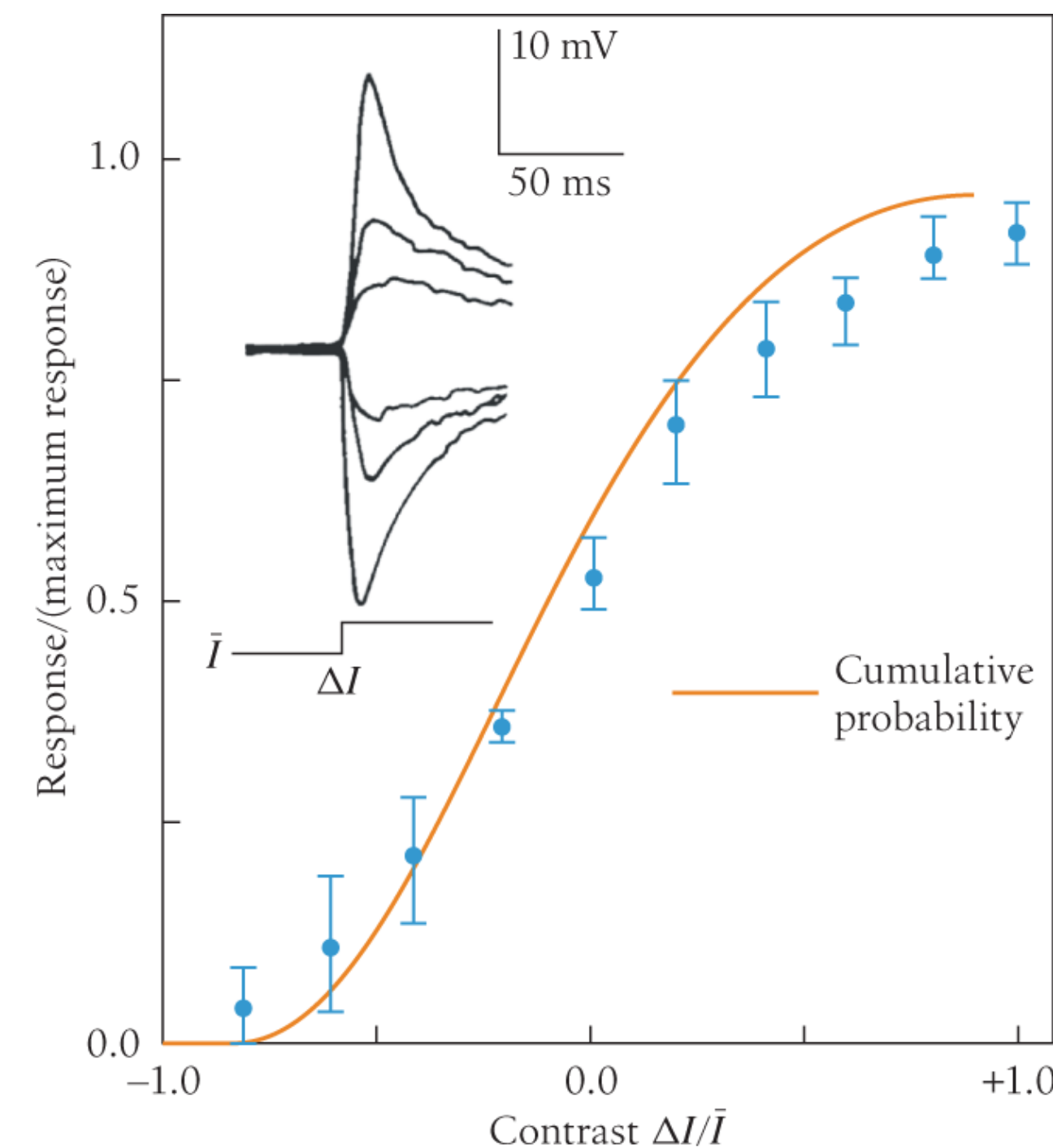
Or, with fixed noise and limited dynamic range at the output, information can be maximized by adjusting the (mean) input/output relation.

This “matching” is a parameter-free prediction.

input/output relation
↓
 $y = f(x) + \eta$ ← noise

e.g.,
 $0 < f(x) < 1$
 $\langle \eta^2 \rangle = \text{const} \ll 1$

$\Rightarrow f(x) = \int_{-\infty}^x dx' Q(x')$
↑
distribution of inputs



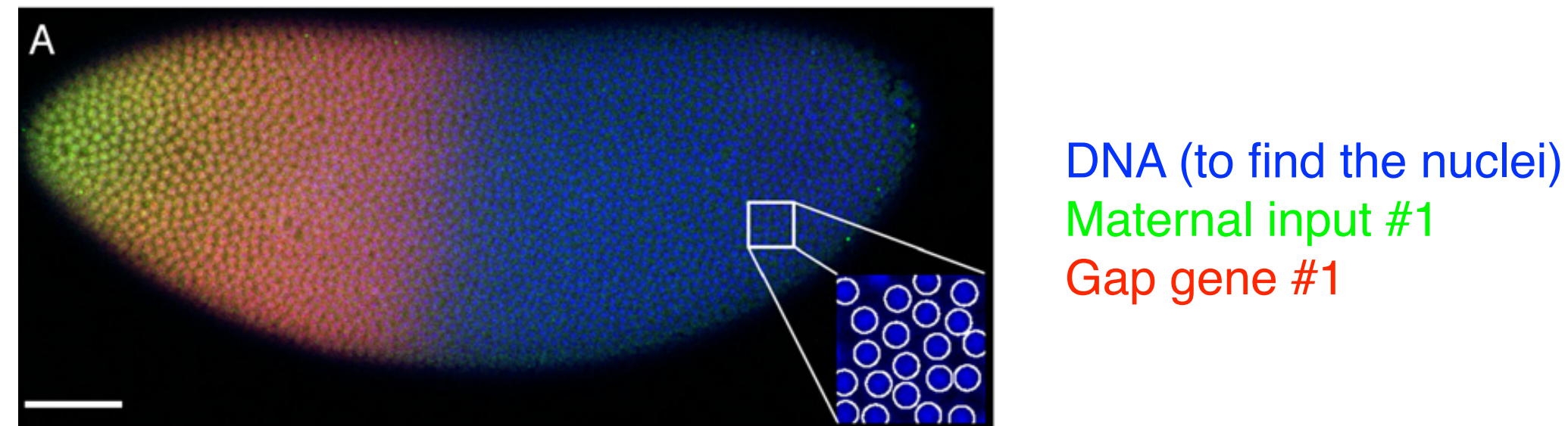
Input/output relation of neurons in the fly retina match the distribution of input light intensities.

SB Laughlin, *Z Naturforsch* **36c**:910 (1981).

Also leads to predictions about adaptation of neural input/output relations to changes in the distribution of inputs ...

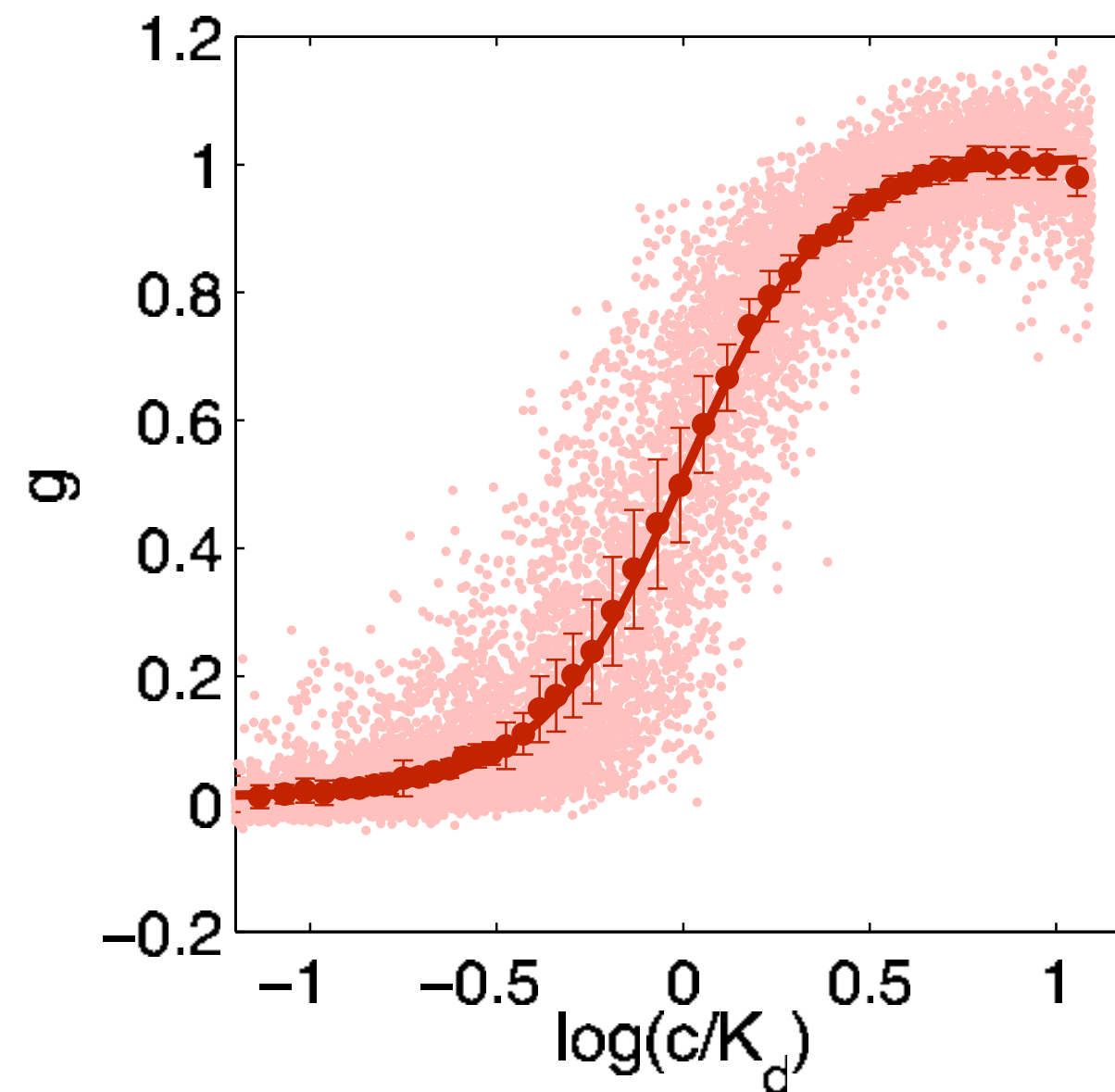
Can we do this for a genetic regulatory element?

Let's go back to the fly embryo ...



Expression of
Gap gene #1

y



Concentration of
Maternal input #1

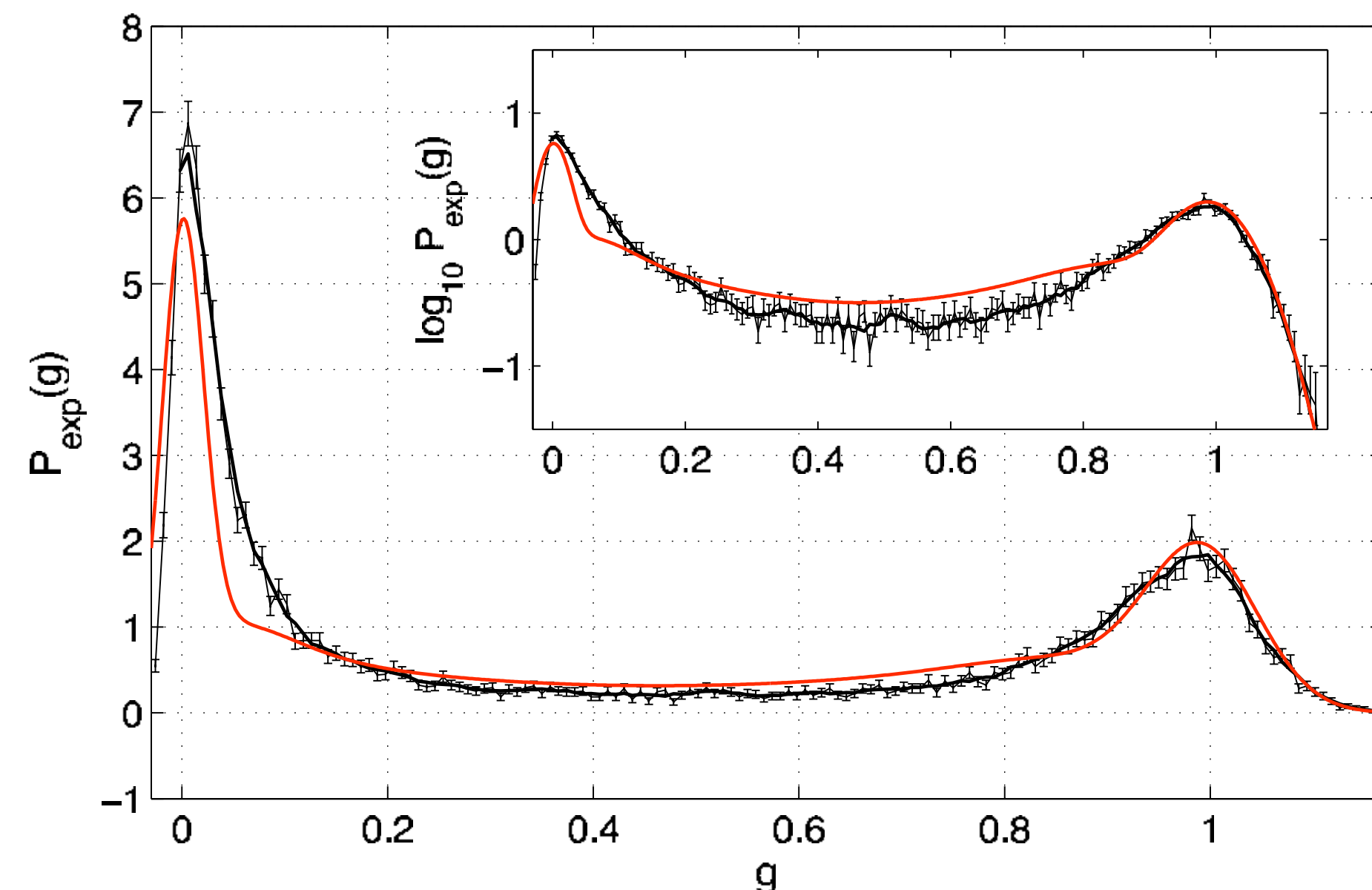
x

T Gregor et al, *Cell* **130**:153 (2007).

This gives us thousands of samples drawn from the relevant probability distributions.

So we can do the general matching problem, numerically. This predicts distribution of inputs, but more sensitive to distribution of outputs.

No free parameters.



G Tkačik et al, *Proc Natl Acad Sci (USA)* **105**:12265 (2008).

This is great, but the real system has several inputs and several outputs ...

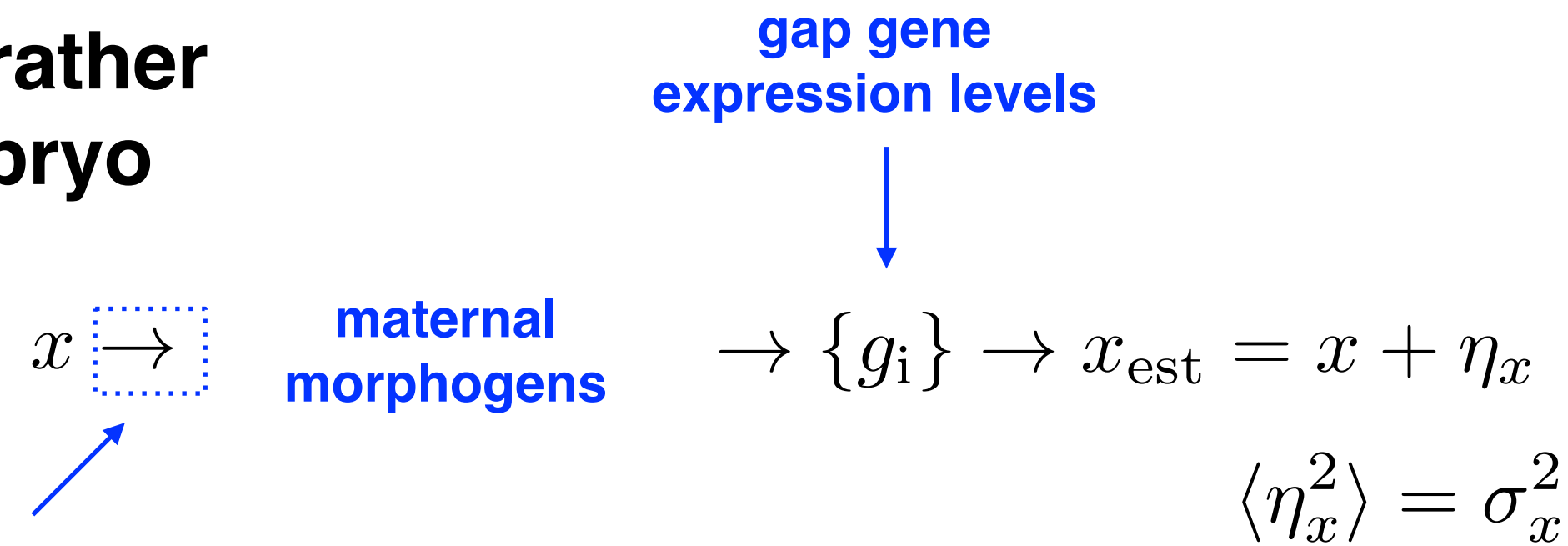
relevant information is not the morphogen concentration, but rather the position of cells in the embryo

if noise is small, the matching condition becomes

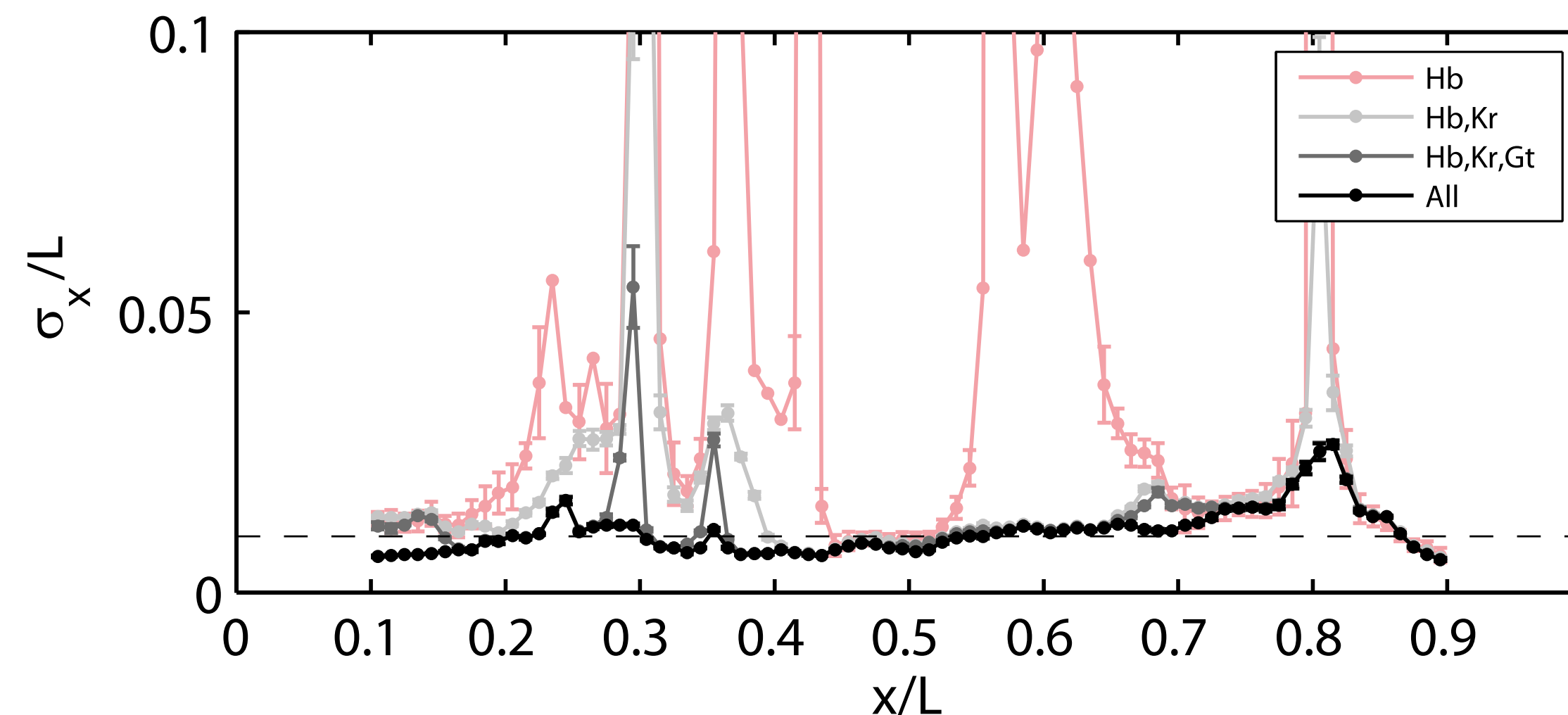
$$Q(x) \propto \frac{1}{\sigma_x}$$

but the distribution of positions is uniform!

maximize positional information by adjusting the mapping to maternal morphogen concentrations



Complicated patterns of gap gene expression lead to nearly uniform positional errors, as predicted from the matching condition.



Noise level is the same as in subsequent events.

$$I(x; \{g_i\})/I_{\text{max}} = 0.984 \pm 0.003$$

Can we make the step $\{g_i\} \rightarrow x_{\text{est}}$ more explicit?

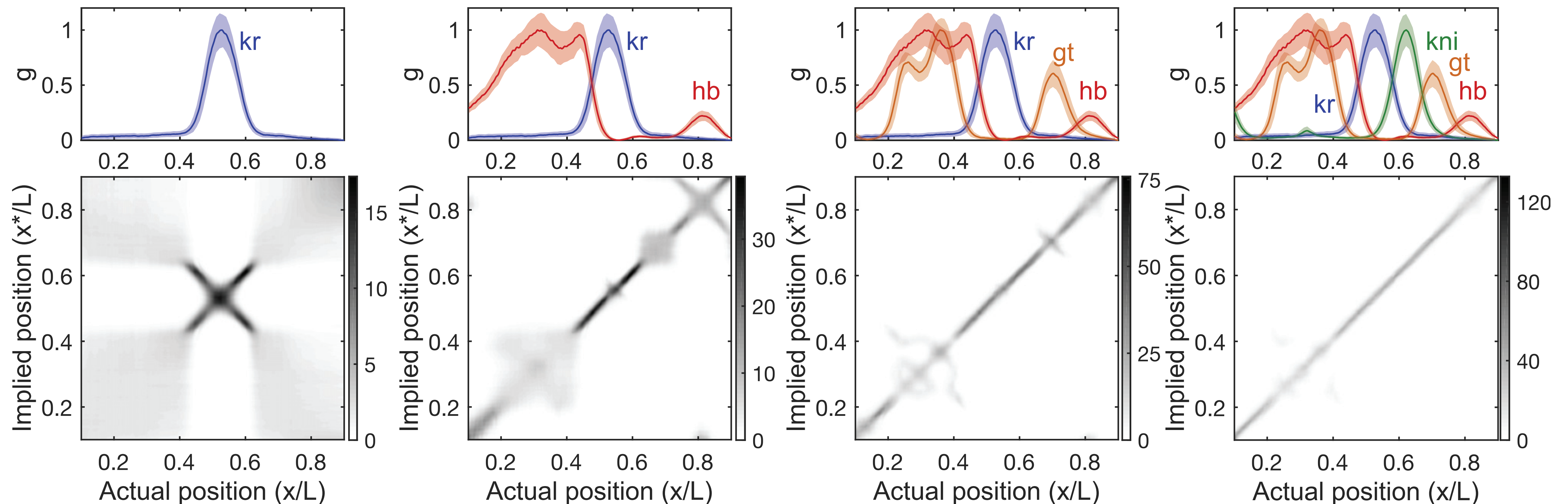
All information
about position from
expression levels

Measured!

$$P(x|\{g_i\}) = \frac{P(\{g_i\}|x)P(x)}{P(\{g_i\})}$$

1/L

Normalization



Once we include all the gap genes, ambiguities are removed and we recover the 1% precision from above.

Notice that this uses all the locally available information.

MD Petkova et al, *Cell* 176:844 (2019).

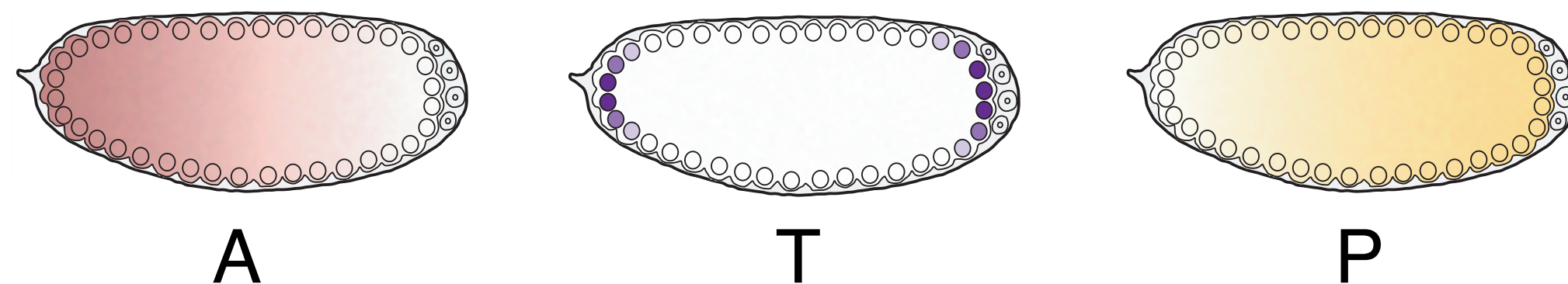
Does the embryo really do this computation?

Use the power of *Drosophila* genetics.

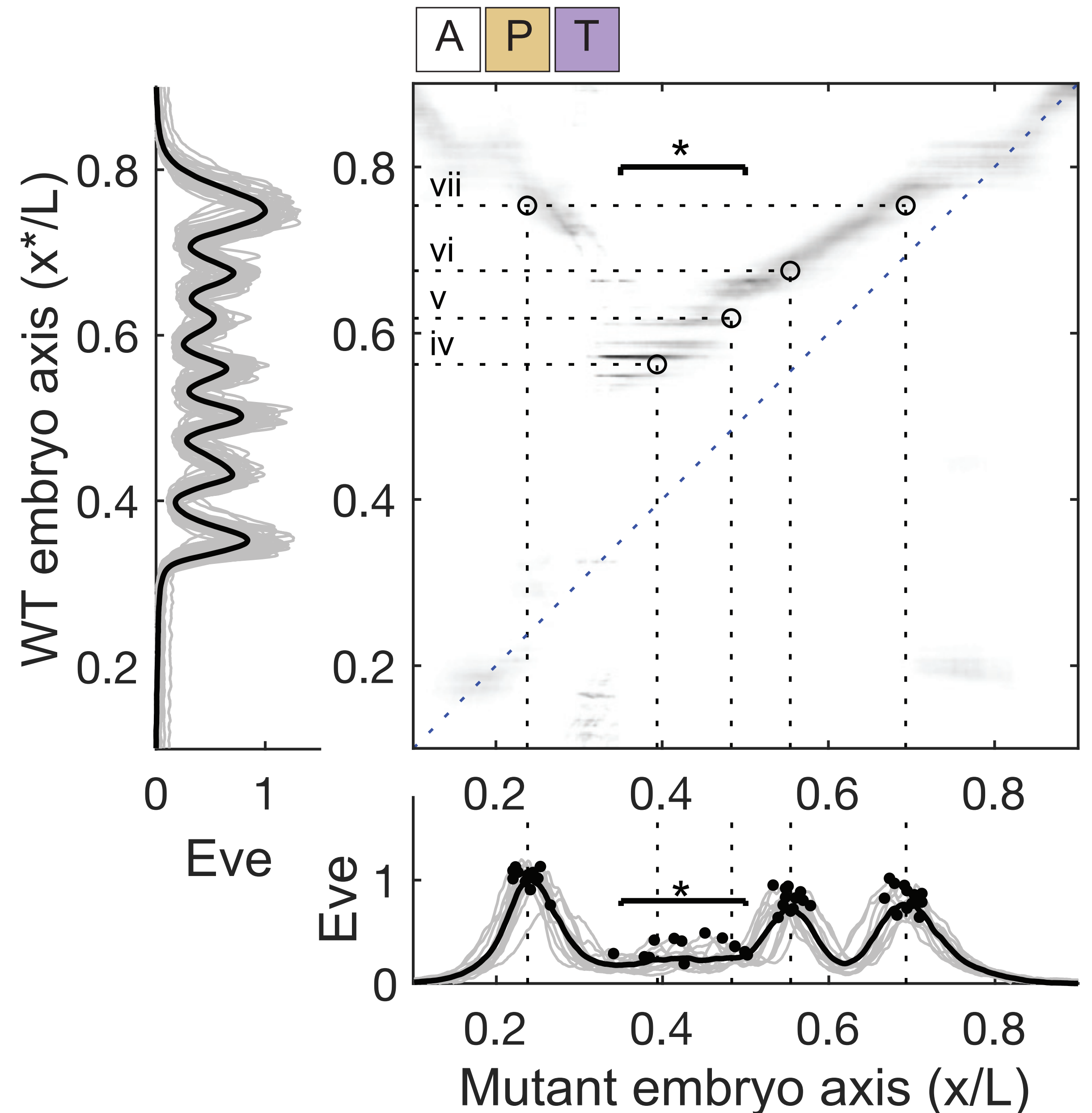
Mutants defective in maternal inputs produce distorted patterns of gap gene expression. We “decode” these with the optimal estimator.

We can see what cells “think” about their position by looking at pair-rule stripes.

Still no free parameters.



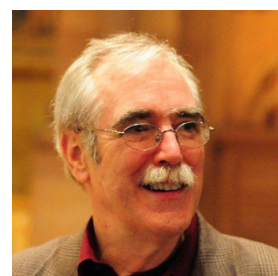
Many stripes, six different mutants ...



“Top-down” system-level principles can generate detailed, quantitative predictions, circumventing highly parameterized “bottom-up” models.

BUT, some of our most powerful experimental tools probe and manipulate biological systems at the molecular, not the system level.

Can top-down principles reach all the way to the bottom?



EF Wieschaus
Princeton



DW Tank
Princeton



RR de Ruyter
van Steveninck
Indiana



T Gregor
Princeton



S Setayeshgar
Indiana



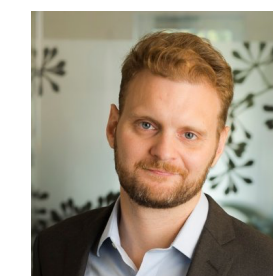
CG Callan
Princeton



G Tkačik
IST Austria



AM Walczak
ENS Paris



JO Dubuis
Clarify Health



T Sokolowski
Frankfurt



MD Petkova
Harvard