Filtering the noise of embryonic development

2005 J. Phys.: Condens. Matter 17 S1245

(http://iopscience.iop.org/0953-8984/17/14/014)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **17** (2005) S1245–S1258 doi:10.1088/0953-8984/17/14/014

Filtering the noise of embryonic development

Bahram Houchmandzadeh

Laboratoire de Spectrometrie Physique, 140, rue de la Physique, BP87, 38402 St-Martin d'Hères Cedex, France

Received 27 August 2004
Published 24 March 2005
Online at stacks.iop.org/JPhysCM/17/S1245

Abstract

Embryonic development is the generation of a multicellular organism from a single cell. During this process, tissues and organs are differentiated and positioned at different parts of the organism. In this paper, we deal with the *precision* (embryo to embryo variations) of this process by looking at the early development of the fruit fly embryo. We show that precision of the antero-posterior gene expression pattern is hard coded at the very first stages of development. The noise reduction mechanism is extremely robust and resistant to mutations, in contrast to the actual position of the tissues themselves.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Multicellular organisms such as ourselves, the fruit fly and the elm tree are generated from one fertilized egg. During the development process, as the number of cells in the organism increases, the subsequent new cells become differentiated and specialized, giving rise to organs and tissues (heart, eyes, neurones, . . .). These organs and tissues are placed at given positions of the body: the heart in humans inside the torso on the left side, neurones inside the head, itself on top of the body and so on. All these processes are of course under genetic control. All cells of an organism have however exactly the same genome, so the development process can be seen as multiple controlled symmetry breakings where cells, according to their position and their history, are instructed to express only a subset of all available genes. Understanding how these patterns are generated is *the* central problem of developmental biology. Tremendous advances have been made during the last 20 years and the general scheme is now more or less understood.

In this paper, we deal with a slightly different problem: the precision of the developmental process. Gene transcription, RNA localization, diffusion are all involved in embryonic pattern generation and are all subject to fluctuations and noise. Any given embryo development programme should be resistant to these noises, i.e. generate always the same patterns in *all* the embryos. All studies until now have focused on the *average* patterns, ignoring the importance

of fluctuations. But if the heart was positioned with a standard deviation of 10% body length, many of the embryos would be lost!

What is the importance of fluctuations? If fluctuations are important, how are they reduced during the unravelling of the development programme? Is that compatible with the current model of embryonic development? To answer these questions, we have looked at the early stages of fruit fly (*Drosophila melanogaster*) embryos [1]. This fly is the main animal model of embryonic development. Its early stage of development involves few genes and after two decades of work, their mutual interactions are reasonably well understood. The understanding of antero-posterior differentiation is what has set the current frame of mind about development and this is why this work has been focused on this process.

This paper is organized as follows. As biological processes are not well known among physicists, the second section is devoted to a brief recap of what genes, transcription and genetic networks are. In the third section, we will review current models of embryonic development and we reanalyse the main one from the 'resistance to fluctuations' point of view in the fourth one. We will describe our measurements and their incompatibility with models in the fifth section. The sixth section presents a detailed investigation of an AP network. A final section is devoted to concluding remarks and plausible models.

2. Genes, transcription and networks

Both computers and living cells are information processing units, with many similarities and profound differences. DNA, a linear array of four nucleic acids (A, T, C, G: adenine, thymine, cytosine, guanine), is similar to a hard disk: a repository for programmes. For computers, programs on a hard disk are more or less continuous sets of bits having a meaning for the processor when read in a continuous way. For living organisms, programmes are called *genes*, a more or less continuous interval of nucleic acid bases on the DNA. Executing a programme for the cell means reading and copying a given gene (*transcription*) and making a protein (*translation*) according to a universal dictionary which associates specific amino acids (building blocks of proteins) with specific triplets of nucleic acid.

On the hard disk of the computer, there is a very special table containing the position and length of each program to which the operating system refers each time it is asked to execute a program. There is no such table for the DNA. Instead, in front of each gene, there is a special interval of bases called a *promoter* which attracts the reading machines (called RNA polymerase) and positions them at the beginning of the gene. Different promoters have different attractivities for RNA polymerase and thus control the amount of protein made from the gene they are attached to. What makes the cell an information processing unit is that the attractivity of a promoter for RNA polymerase is not fixed, but can be modified (enhanced or reduced) by other proteins attaching themselves to the DNA close to the promoter (figure 1; see classic textbooks such as [2, 3]).

Proteins with the sole purpose of regulating the activity of other genes are called transcription factors and can group into complicated networks of auto-regulation. Part of such a network in the *Drosophila*, responsible for antero-posterior patterning and of special interest to us, is shown in figure 2(a). If a gene A enhances the production of a gene B it is called an activator and this is denoted by $A \rightarrow B$. It is generally believed that the rate of production of B should follow an expression proportional to $A^n/(K^n + A^n)$. The coefficient n is called the Hill coefficient. It is a rough estimate of how many A proteins have to bind *cooperatively* to the DNA to activate the transcription of B. For high n, the production of B is switch-like, with negligible production if $A < K$ and a high saturating level if $A > K$. The constant K thus plays the role of a concentration threshold. Measuring such laws in

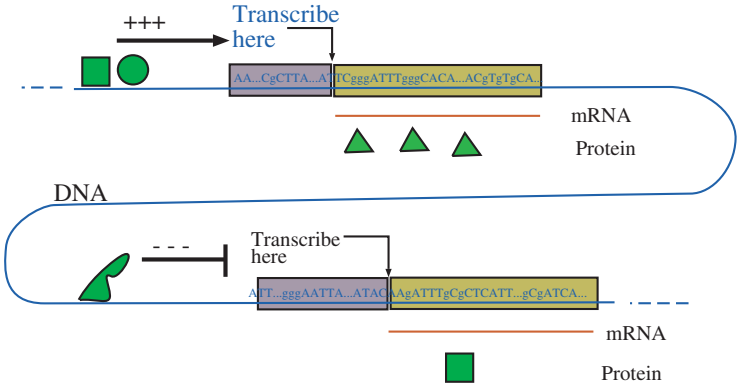

Figure 1. The transcription rate of a gene can be enhanced or reduced by the attachment of other gene products (proteins) to their controlling region. In this sketch, the production of the protein 'triangle' is activated by the protein 'square'. The transcription of this last gene is inhibited by the protein 'shapeless'.

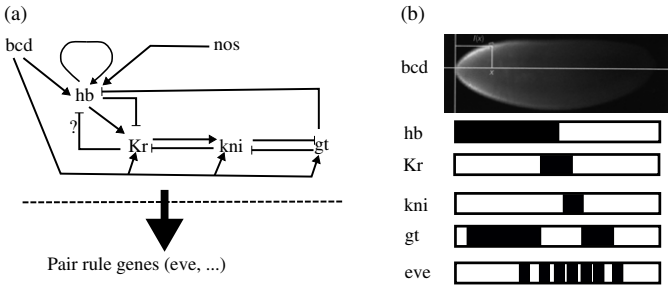

Figure 2. Left side: the regulatory network for the early antero-posterior pattern formation of *Drosophila* embryos. We do not show all the genes or all the interactions. Right side: the top photograph shows a *Drosophila* embryo (∼0.5 mm) stained by antibodies to show the expression domain of the gene *bicoid*. Other boxes sketch the expression domains of other genes discussed in this paper.

gene expression is extremely difficult. An attempt to measure the cooperative binding of the transcription factor bicoid has been made by Burz *et al* [4]. A transcription factor A can also inhibit the expression of a gene B. This interaction is denoted as $A \dashv B$ and the production rate would be proportional to $1/(K^n + A^n)$.

3. Pattern formation in embryonic development

As the embryo develops and its number of cells increases, cells differentiate and express different subsets of genes, even before any morphological change can be observed. The *Drosophila* embryo at cycle 14 (just before the gastrulation, when all kinds of folds are formed) is an ellipsoid egg formed of roughly 6000 cells equally dispersed on its surface. At this stage however, cells are well differentiated and express different genes as a function of *their position* inside the embryo (figure 2(b)) [5].

The fundamental problem of developmental biology is the following: how do cells measure their position inside the embryo, and how do they decide, on the basis of this information, which genes to express. There are two main classes of models dealing with this problem:

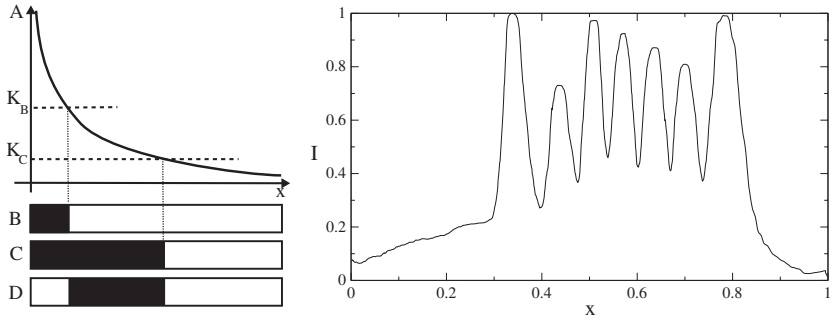

Figure 3. The morphogen gradient model. A gene is expressed in the region where the concentration of its activator is above (or of its inhibitor is below) a given threshold. The combinatorial game of secondary morphogens can produce complicated patterns, such as those of the *eve* gene in the *Drosophila* embryo shown on the left side.

the biological ones, called the *morphogen gradient* models; the physical ones, called *Turing* models (although different versions carry other names such as reaction–diffusion or lateral inhibition).

The morphogen gradient model has been formulated from the very beginning of embryogenesis history and it is based on many 'cut and paste' embryo experiments. It received a mathematical formulation in 1969 by Wolpert [6]: let us suppose that we have a transcription factor A (a morphogen) which activates genes B and C in a switch-like way with threshold values K_B and K_C; let us moreover suppose that A has a variable concentration through the embryo. Then only cells in the region where $A > K_B$ (or K_C) will express the gene B (or C). If B and C proteins are themselves transcription factors for the gene D, with the first one acting as an activator and the second one as an inhibitor, only cells in a stripe will express D (figure 3). This combinatorial game can be continued and any gene expression pattern, such as the one displayed in figure 3 (right), can be obtained. By simple calculations, Lewis *et al* [8] have estimated that a single gradient can reliably set 30 different thresholds.

The Turing models were first proposed by Alan Turing in a seminal 1952 paper [7]. Here, symmetry is spontaneously broken. Turing supposed the existence of two transcription factors X and Y (he was using chemical reactions to be precise; the nature of genes was not known at that time). If X activates itself and Y, and Y activates itself and inhibits X, then oscillatory patterns such as that shown in figure 3 (right) can be generated. It is necessary for the gene products (i.e. proteins) to have different diffusion coefficients. Many other patterns can be generated upon slight modifications of the model and have been extensively investigated by researchers such as Lacalli, Meinhardt and Murray (see [9, 10] for an example). The beauty of the Turing model is the ease of generating oscillatory patterns which are ubiquitous among insects: for example, genes belonging to the pair rule groups (such as *even-skipped* shown in figure 3) are all expressed in seven equally spaced stripes.

The debate about the relevance of each of these models raged until 1987, when Driever and Nusslein-Volhard [13, 14], and later Struhl [15], in a series of beautiful experiments, established the gradient model by discovering the first morphogen called bicoid. This protein is responsible for the antero-posterior pattern formation in *Drosophila* embryos and possesses all the properties that a morphogen should have:

(i) The bicoid (bcd) RNA is provided by the mother and *localized* to the anterior pole, acting as a continuous source for bcd protein production; the protein diffuses from this source

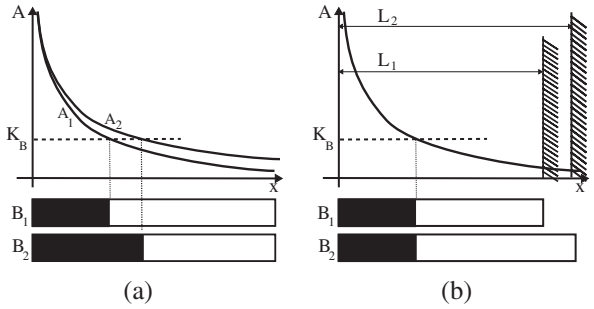

Figure 4. If the morphogen gradient varies in different embryos (a) or if not the gradient but the size of the embryos is variable (b), the relative expression domain of downstream genes fluctuates from embryo to embryo.

and is degraded throughout the embryo, establishing an exponential gradient (figures 2 and 5).

(ii) Bicoid is a transcription factor for the first A–P zygotic genes expressed in the embryo (the gap genes shown in figure 2).

(iii) Modifying the number of copies of the bcd gene on the mother shifts the expression domain of all downstream genes!

This last experiment was the most conclusive one. Suppose that the bcd protein concentration is $B(x) = B_0 \exp(-x/\lambda)$ and a downstream gene hunchback (hb) is activated in the region where $B(x) > K_{\mathrm{H}}$, i.e. for $x \in [0, x_{\mathrm{H}}]$ where $x_{\mathrm{H}} = \lambda \ln(B_0/K)$. Now if the number of copies of the gene is doubled, so is the amplitude of the bcd gradient and the expression domain of hb is shifted toward the posterior to $x'_{\mathrm{H}} = x_{\mathrm{H}} + \lambda \ln 2$. The above authors observed downstream gene expression domain shifts, even though they did not check the quantitative agreement.

Later it was demonstrated that each stripe of pair rule genes can be switched on and off by modifying the controlling region of the gene (see for example [11, 12]) exactly as predicted by the combinatorial code of successive morphogen gradients. Akam then wrote a revengeful article [16] definitively dashing the physical model.

4. Problems with the bicoid model

After many experiments, bicoid is now established as the paradigm of the morphogen gradient. This model however, when inspected carefully, has many weak points related to error propagation:

- The exponential bicoid gradient is established by diffusion from a source and degradation:

$$\partial_t B = D\partial_x^2 B - \tau^{-1} B + S(x)$$

  where D is the diffusion coefficient, τ the decay rate of the protein and the source $S(x)$ denotes the mRNA provided by the mother and localized at the anterior pole. To make a reliable gradient in all the embryos, the mother should precisely control the quantity of degrading proteins, the amount and position of localized RNA and so on. Fluctuations in these quantities will give rise to different gradients in different embryos and different patterns of gene expression in them (figure 4(a)).
- This model cannot give the proportionality between gene expression patterns and the size of the egg: *a priori*, we would think that leg and arm length should be proportional to

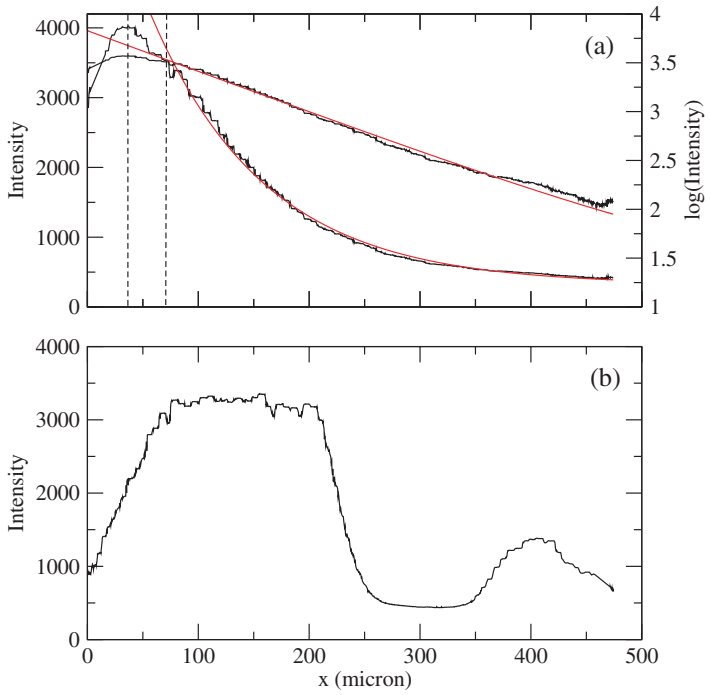

Figure 5. (a) The expression profile of *bcd* and its fit to an exponential on linear (left) and logarithmic (background extracted) (right) scales. The fit is made in the region $x >$ two times peak position. (b) The expression profile of *hb*.

body size. In the bicoid model, even if we suppose that the gradient is reliably established throughout the embryo, its exponential slope $\lambda = \sqrt{D\tau}$ is independent of the egg length (figure 4(b)).

- Environmental variables such as temperature fluctuate. Cooperative gene transcriptions and morphogen gradients themselves are however extremely sensitive to these variations. How the same adult fly can thus emerge from eggs maintained at 15 or 30 °C?

To answer these questions, one needs a tool to measure the embryo to embryo fluctuation of gene expression in a quantitative way. We have used antibody staining and image analysis to achieve this. Antibody staining is a technique used to visualize gene expression domains: a tissue is washed with a fluorescent antibody raised against a specific protein and protein concentration is thus converted to fluorescent intensity. Of course, there is experimental noise associated with this technique which has to be carefully analysed [1] before one can quantify the intrinsic fluctuations. The image analysis technique is straightforward and consists mainly of taking a moving window the size of a nucleus along the embryo's contour, computing the intensity inside and projecting onto the main axis of the embryo to get the signal $I(x)$ [1].

Figure 5 shows the spatial gene expression profile (all spatial units are relative to the embryo length EL) of *bcd* and *hb*, which is the first zygotic gene activated by bcd and itself a secondary morphogen [17]. As can be observed, the bcd gradient is indeed exponential (after a first peak located at $x \approx 0.1$) with an average slope (see below) of $\lambda = 0.27$. The hb expression switches from a high value to a low one at mid-embryo ($x_{\mathrm{H}} = 0.49$) over a $\Delta x = 0.1$ interval, which corresponds to a 30% change in bcd concentration. Supposing that *hb* is activated cooperatively by bcd, i.e. $H \sim B^n/(K_{\mathrm{H}}^n + B^n)$, and using the exponential

variation of bcd, one can evaluate the threshold value and the Hill coefficient:

$$K_{\mathrm{H}} = \exp(-x'_{\mathrm{H}}/\lambda) = 0.23$$
$$n = 4\lambda \left.\frac{\mathrm{d}H}{\mathrm{d}x}\right|_{x=x_0} \approx 11$$

where $x'_{\mathrm{H}} = x_{\mathrm{H}} - 0.1$ is the distance between the bcd peak (at $x = 0.1$) and the hb switching position. K_{H} is relative to the maximum bcd value (taken as unity).

The Hill coefficient seems extremely high, especially if compared to the estimate of Burz *et al* [4] of $n \approx 2$. Years ego, Lewis *et al* [8] proposed that sharp boundaries can be made with smooth gradients if a positive auto-feedback loop is added. Here, hb does indeed act as a transcription factor for itself [18] (figure 2(a)) and we think that this should be related to the sharp gradient of hb. This hypothesis can be checked by removing the corresponding regulatory element of the *hb* gene.

5. Embryo to embryo fluctuations

Figure 6 shows profiles of *bcd* and *hb* gene expression in more than 100 embryos at cycle 14. As can be seen, the bcd gradient is indeed extremely variable. The position x_{B} at which it crosses the $K_{\mathrm{H}} = 0.23$ threshold has a standard deviation of $\sigma_{\mathrm{B}} = 0.07$. Cell nuclei have a size of $\approx$0.01. The positional error of bcd is therefore greater than five nuclei in 50% of embryos! If embryos were made only on the basis of this signal, a majority of them would be lost. The more robust way of quantifying the positional information of bcd is to measure the exponential slope, which shows the same variability: $\sigma_\lambda = 0.05$, in agreement with the above value ($\sigma_{\mathrm{B}} = \sigma_\lambda \log(1/K_{\mathrm{H}})$). The amplitude B_0 of the gradient can only be evaluated through statistics, and we have therefore neglected fluctuations in this parameter (always normalized to unity). No correlation has been found between the absolute (fluorescent intensity) value of B_0 and λ. Fluctuations in B_0 will thus further degrade the positional information on the bcd gradient.

The hb profile in contrast is extremely precise. Its switching position x_{H} shows a variability of $\sigma_{\mathrm{H}} = 0.01$: in two thirds of embryos, the boundary of its expression domain is defined more precisely than one nucleus.

The noise in bcd is indeed huge, but it is filtered out at the *hb* expression level. Moreover, this precision arises very early: at cycle 13, when the zygotic gene expression begins, the hb boundary is already set precisely. The same precision is observed on the *hb* mRNA domain, indicating that the noise filtering is not due to post-transcriptional processes.

Our second objection was about proportions. x_{B} and λ (measured in absolute value), as expected, show no correlation with the embryo length. On the other hand, the *hunchback* switching position x_{H} shows perfect correlation and proportionality with the embryo length L: the linear (r_{l}) and Spearman (r_{s}) correlation coefficients between x_{H} and L are 0.84 and 0.82 ($P < 10^{-20}$) [1]. Thus, speaking of the hb switching position as $0.49L$ is legitimate.

These results are at odds with the model of bcd acting as a morphogen for *hb*. They can be further confirmed by maternal mutations such as *torso* and *exuberantia* which modify the amount and distribution of *bcd* mRNA deposited in the embryo. Figure 7 shows the effect of a harsh double mutation *torso–exu* on the bcd and hb profiles and its comparison to the wild type. As can be observed, the bcd profile is flattened and even its peak barely reaches the 0.23 threshold. The hb on the other hand is unaffected (the posterior expression of *hb* is not discussed in this paper).

Temperature variations between 9 and 29 °C have a similar effect: bcd profiles are extremely sensitive to them, but hb ones are again unaffected [1]. This happens despite

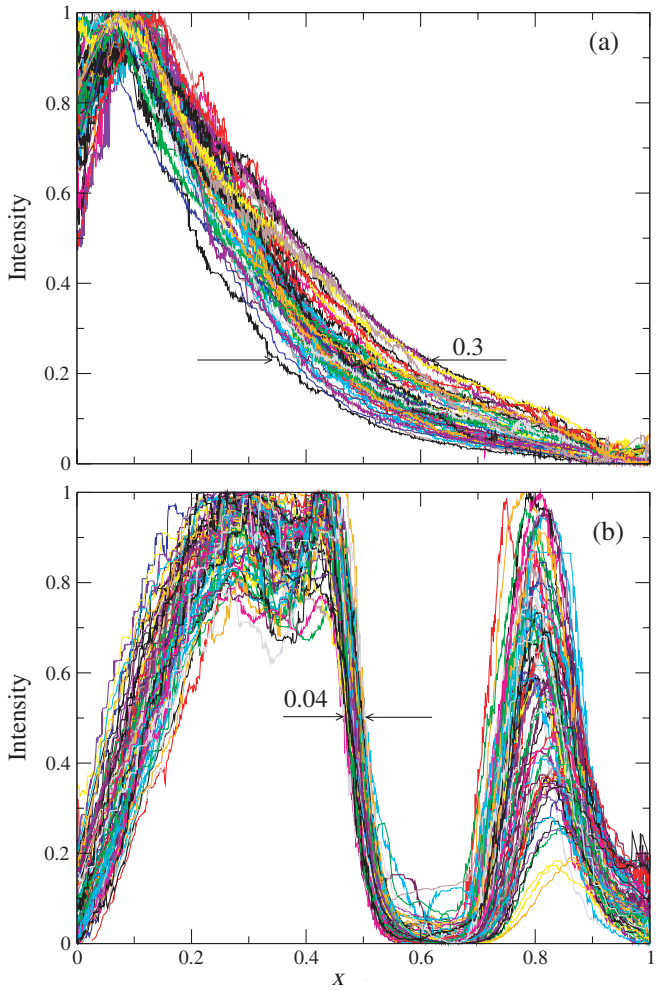

Figure 6. The expression profile of *bcd* (a) and *hb* (b) in more than 100 embryos.

huge variations in development time for reaching the mid-cycle 14, from 2 h at 25 °C to 20 h at 9 °C.

As mentioned earlier, bcd as a morphogen received wide acceptance because of dosage modification experiments: embryos derived from mothers with variable numbers of *bcd* gene copies compared to the wild type ($n = 0.5, 1$ (WT), 2, 3) displayed shifts in the expression domain of downstream genes. The earlier authors however did not check for numerical agreement with the morphogen model. As we mentioned, if *hb* was activated solely by bcd, the shift in its domain boundary x_{H} would be $\Delta x = \lambda \ln n$. We have used the same transgenes as Driever and Nusslein-Volhard. Table 1 shows that there is a large discrepancy between the expected and measured values for x_{H}. One can argue that transgenes can have a lower efficiency. By quantifying the fluorescence, we have measured the efficiency of the transgenes and they do indeed have a lesser efficiency of 50%. Even taking this effect into account, the discrepancy with respect to expectation remains high. The shift in the hb boundary instead follows $\Delta x = (\lambda/2) \ln n$.

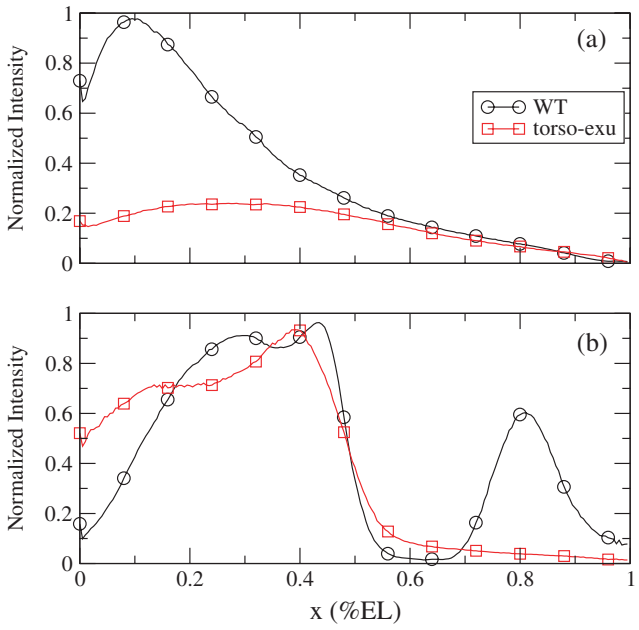

Figure 7. The bcd (a) and hb (b) in wild type (circles) and *torso–exu* (squares) mutant embryos. Profiles are averages over approximately 20 embryos. WT and *torso–exu* mutant embryos were collected at the same time, stained in the same bath and average fluorescent intensities were used for the expression level quantification.

Table 1. The hb boundary x_H as a function of the *bcd* copy number n of the mother.

Background	Measured	Expected (50% efficiency)	Expected (100% efficiency)
$n = 0.5$	0.41	NA	0.30
$n = 1$ (WT)	0.49	NA	NA
$n = 2$	0.56	0.60	0.68
$n = 3$	0.59	0.68	0.79

6. Effects of mutations

The results of the preceding section show clearly that the established picture of bcd is missing important elements and there should be other transcription factors filtering its noise, and establishing a sense of proportion in the embryo. To test this hypothesis, we have removed by mutation various genes of the AP network, and checked for decrease in hb *precision*: many genes are known to participate in setting hb boundaries, and their mutations shift the hb position. But here, we are especially looking for the gene (or genes) involved in the filtering of bcd noise. This means that mutating this gene will increase the variability of the hb domain to a level comparable to that of bicoid.

6.1. Maternal genes

The most natural candidate would be the maternal gene *nanos* (see figure 2(a)). mRNA of this gene is localized at the *posterior* pole of the embryo, the protein makes a posterior gradient and the product *inhibits* the translation of *hb* [19]. Table 2 shows indeed that *nanos* mothers

Table 2. The hb boundary position and precision as a function of the mutation. The superscripts denote the alleles used.

Mutation	x_H	σ_H (10^{-2})	n
WT	0.49	1	110
oskar^{6}	0.52	1.6	40
nos^{BN}	0.54	1.6	16
hb^{mat}, nos^{BN}	0.47	1.4	31
exu^{PJ42}	0.51	2	21
swa^{1}	0.52	2.5	17
torso^{PM51}	0.50	1.4	30
spg	0.51	1.7	9
Kr^{1}	0.49	1.1	29
kni(del)	0.48	1.0	11
gt^{YA82}	0.46	1.0	16
hb^{6N}	0.45	1.1	17
X^{-}	0.43	1.4	40
2R^{-}	0.49	1.5	19
2L^{-}	0.47	1.4	16
3L^{-}	0.50	2.0	39

make embryos with posterior shifted hb domains and the same holds for *oskar* mutant (*oskar* is upstream of *nanos*). The precision of hb however remains unaffected in these backgrounds.

One of the complexities of the AP regulatory network is that part of *hb* itself is provided maternally, i.e. the mother deposits *hb* mRNA homogeneously inside the egg, which acts as a secondary source of hb protein production (the primary being the zygotically produced one). No equivalent of *bcd* has been found in higher dipterans, where the first steps of AP boundary setting are supposed to be achieved solely by *nanos* and *hb* (either maternal or zygotic). It has been shown that even in flies, part of this old circuit can be artificially used to rescue thoracic segments in *bcd* mutants [20]. We have checked however that removing *nanos* and maternal *hb* does not affect hb precision (table 2).

Other maternal genes such as *swa*, *torso*, *exu* do not contribute directly to the activation of *hb*, but affect more or less the distribution and activity of *bcd* [21, 22]. The maternal gene *sponge* regulates the spatial distribution of cell nuclei [23]. Neither of these genes is involved in the precision of the hb boundary (table 2).

6.2. Zygotic genes

Regarding the above results, one could wonder whether Turing mechanisms have not been eliminated too fast. Can zygotic cross-regulation between gap genes be the main source of precision in the embryo? The bcd would then set the hb boundary coarsely, and leave its refinement to cross-regulations.

A look at figure 2(a) shows that there are indeed many feedbacks from gap genes to *hb*. The best candidate would be *Kruppel*, which has its expression domain adjacent to hb toward the posterior position (figure 2(b)). Struhl has shown that hb controls Kr [17], where, on the other hand, Jackle *et al* [24] have apparently shown that Kr represses *hb*. Moreover, a Turing-like numerical simulation of the AP network as shown in figure 2 displays all the noise filtering properties described above. There is indeed a strong linear correlation between hb and Kr boundaries which is maintained even when they are shifted in various backgrounds ($r_s = 0.90$) (figure 8(a)). Figure 8(b) shows the average (over approximately 20 embryos)

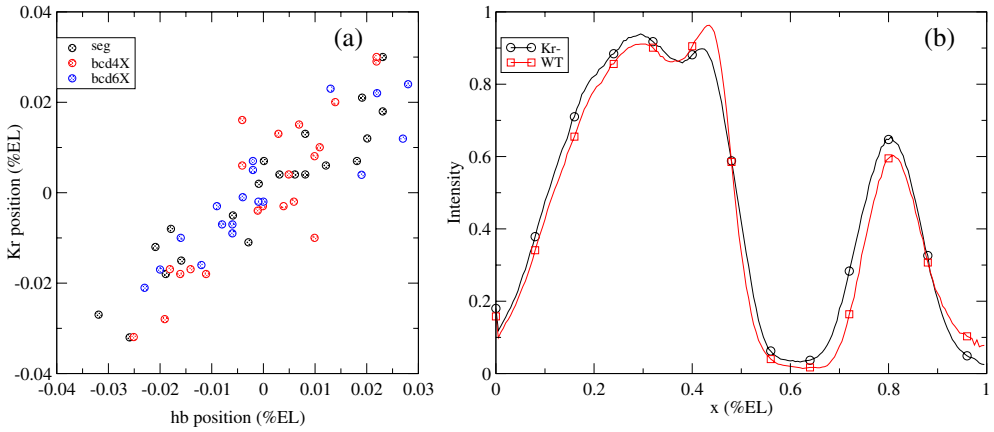

Figure 8. (a) The correlation between hb and Kr positions in three different backgrounds where their position is shifted. The average position of each background has been subtracted to centre the data around zero. (b) The average profile of *hb* in WT and Kr mutant embryos.

profile of hb in the wild type and in Kr^- mutants: no difference is observable (either in position or in precision) and there is therefore no repression (or activation) of *hb* by Kr! The Jackle *et al* result was obtained when antibody staining was in its infancy, using radioactive probes on slices of fixed embryos: angular variation in the cut section could be a possible source of error, and because of technical difficulties at the time, apparently no statistical analysis on several embryos has been performed.

We have also tested other candidate gap genes such as *Giant* and *kni* known to more or less cross-regulate with *hb* [25]. Also, as we mentioned above, hb is an activator for itself and therefore we have also investigated flies heterozygous for *hb*. None of them has an effect on the precision of the hb boundary (table 2)

In principle, the large scale mutation screening performed by Wieschaus *et al* [26] has uncovered all the zygotic genes of *Drosophila* development. A closer look at the data shows however that the allele distribution is not Poissonian and this is probably due to the size of the genes: the bigger a gene, the more susceptible it is to single-point mutation induced by mutagen agents. Taking the gene size into account, a rough estimate of the unknown number of zygotic genes involved in the development would be fifteen [27]. It is therefore possible that the gene responsible for the precision of hb is still unknown. To check this hypothesis, we have screened around 80% of the *Drosophila* genome by selectively removing each chromosome, except the right arm of chromosome 3 where *hb* itself is located. No difference in the precision of hb has been observed (table 2).

6.3. The gene staufen

We have finally found that mutation in *staufen*, a *maternal* gene, disrupts the fine positioning of the hb boundary (figure 9(a)). The hb boundary is widely fluctuating in embryos derived from homozygous mothers and shows a variability similar to that of bcd: $\sigma_H = 0.07$. The result is *a priori* strange and unexpected. The stau is mainly active during oogenesis and plays a role in the localization of *bcd* and *oskar* mRNA to anterior and posterior poles [28]. The effect of *stau* mutation on the bcd gradient is similar to that of *exu*. Moreover, by quantifying the fluorescence level, we have shown that embryos with posterior or anterior shifted hb boundaries have similar bcd gradients (in shape and amplitude) [1]. Even though we do not know the

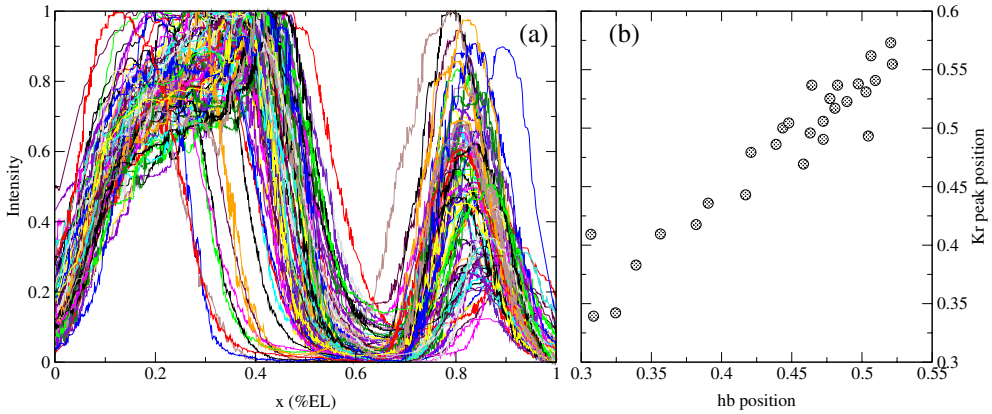

Figure 9. (a) Profiles of hb in about 100 embryos in a staufen background. (b) The correlation between hb and Kr positions in the same background.

effect on the *nanos* gradient, we have checked (see above) that neither *oskar* nor *nanos* have any effect on the hb precision. The stau itself does not form any kind of gradient in the embryo. At this stage, we do not understand the molecular mechanisms by which *stau* mediates the precision of the hb boundary.

One should note that hb acts as a real morphogen: other gap genes such as Kruppel follow precisely the shifts in its boundary and remain strongly correlated to it, even when the hb boundary fluctuates widely from embryo to embryo in the *stau* background (figure 9(b)).

7. Concluding remarks

We have shown above that bcd is not as good a morphogen as it should be: it fluctuates widely from embryo to embryo. The boundaries of downstream genes such as *hb* do not follow these fluctuations and are set precisely in all embryos. Turing mechanisms, i.e. cross-regulation and diffusion among gap genes, despite their elegance and their inherent error correction capabilities, are not involved in this noise filtering process. Instead, as the *stau* mutation shows, this process is under maternal control, which *a priori* does not suppose any transcription feedback loop.

A possible model which can explain all the above-mentioned properties of *hb* transcription is the following. There exist *two* gradients: one the well established anterior bcd acting as an activator for *hb*, and one a hypothetical posterior P acting as an inhibitor of *hb*. As we mentioned, bcd follows an exponential $B(x) = \exp(-x/\lambda)$, and the major measured source of noise in embryo to embryo fluctuations of the positional information is due to fluctuations in λ. We have to make the additional hypothesis that the posterior gradient $P(x) = \exp[(x - L)/\lambda]$ has the same slope λ as bcd in every embryo. This is not an unreasonable hypothesis. The slope depends on the diffusion coefficient D and the degradation rate τ^{-1}. This last parameter depends itself on the quantity of degrading protein deposited by the mother in each embryo and we suspect this last quantity to be the major source of noise. The required condition is that bcd and P be degraded through the same pathway, which will correlate with the slope of their exponential.

If the boundary of hb is set by the condition $B(x_H) = P(x_H)$, then $x_H = L/2$! This explains two major properties of the hb boundary: (i) it is set proportionally to the size of the embryo;

(ii) is does not depend on the fluctuations of λ and the bcd gradient. An additional hypothesis on the exact law of transcription could make this model also temperature compensated. Moreover, this model explains the dosage modification experiments. If the amplitude of the bcd gradient is modified n-fold, $B(x) = n \exp(-x/\lambda)$, then the corresponding shift in the hb boundary will be $\Delta x = (\lambda/2) \ln n$, in quantitative agreement with actual measurements.

Despite the good agreement of this model with measurements, and the fact that *staufen* is involved in the set-up for both anterior *and* posterior gradients, we cannot avoid the fact that the model is not backed by actual hard data. The hypothetical P is not the nanos posterior gradient, and after 20 years of screening for AP maternal mutations no other gene involved in the transcription of *hb* has been found. At our present level of understanding, the above model is just a plausible one.

Finally, we should mention that other morphogens involved in later stages of development can be made robust. A beautiful example has been shown for the BMP case, where internalization by cells can play a regulatory role and make the activating signal precise [29].

For the problem of bicoid discussed in this paper, further understanding of the exact role of *staufen* is needed before any theoretical progress can be made.

Acknowledgments

This work was the result of a close cooperation with Professor E Wieschaus (Princeton University, Princeton, NJ) and Professor S Leibler (Rockefeller University, New York, NY).

References

[1] Houchmandzadeh B, Wieschauss E and Leibler S 2002 Establishment of precision and proportionality in early Drosophila embryo *Nature* **415** 798–802
[2] Alberts B *et al* 2002 *Molecular Biology of the Cell* (New York: Garland Science)
[3] Ptashne M and Gann A 2002 *Genes and Signals* (New York: Cold Spring Laboratory Press)
[4] Burz D S *et al* 1998 Cooperative DNA binding by bicoid provides a mechanism for threshold-dependent gene activation in the Drosophila embryo *EMBO J.* **17** 5998–6009
[5] The FlyEx web site http://urchin.spbcas.ru/flyex maintained by J Reinitz and M Samsonova contains confocal photographs and quantitative data for nearly all the early gene expression patterns of the AP system
[6] Wolpert L 1969 Positional information and spatial pattern of cellular differentiation *J. Theor. Biol.* **25** 1–47
[7] Turing A M 1952 The chemical basis of morphogenesis *Phil. Trans. R. Soc.* **237** 37–72
[8] Lewis J, Slack M W and Wolpert L 1977 Thresholds in development *J. Theor. Biol.* **65** 579–90
[9] Lacalli T C and Harrison L G 1991 From gradient to segments *Sem. Dev. Biol.* **2** 107–17 This is a late coming paper, already incorporating gradients. But many significant concepts are discussed therein
[10] Meinhardt H 1989 Models for positional signaling *Development* (Suppl.) 169–80
[11] Stanjovec D, Hoey T and Levine M 1989 Sequence-specific DNA-binding of the gap proteins *Nature* **341** 331–5
[12] Howard K R and Struhl G 1990 Decoding positional information: regulation of the pair-rule gene hairy *Development* **110** 1223–31
[13] Driever W and Nusslein-Volhard C 1988 The bicoid protein determines position in the Drosophila embryo *Cell* **54** 95–104
[14] Driever W, Thoma G and Nusslein-Volhard C 1989 Determination of spatial domains of zygotic genes *Nature* **340** 363–7
[15] Struhl G, Struhl K and Macdonald P M 1989 The gradient morphogen bicoid is a concentration dependent transcriptional activator *Cell* **57** 1259–73
[16] Akam M 1989 Making stripes inelegantly *Nature* **341** 282–3
[17] Struhl G, Johnston P and Lawrence P A 1992 Control of Drosophila body pattern by the hunchback morphogen gradient *Cell* **69** 237–49
[18] Treisman J and Desplan C 1989 The product of the Drosophila gap genes hunchback and Kruppel bind to the hunchback promoters *Nature* **341** 335–7
[19] Gavis E R and Lehmann R 1994 Translational regulation of nanos by RNA localisation *Nature* **369** 315–8

[20] Wimmer E A *et al* 2000 Bicoid-independent formation of thoracic segments in Drosophila *Science* **287** 2476–9
[21] Schnorrer F, Bohman K and Nusslein-Volhard C 2000 The molecular motor dynein is involved in targeting Swallow and bicoid RNA to the anterior pole of Drosophila oocyte *Nat. Cell Biol.* **2** 185–90
[22] Ronchi E *et al* 1993 Down regulation of the Drosophila morphogen bicoid by the torso receptor-mediated signal transduction signal *Cell* **74** 347–55
[23] Postner M A, Miller K G and Wieschaus E F 1992 Maternal effect mutations of the sponge locus affect actin cytoskeletal rearrangements in Drosophila melanogaster embryos *J. Cell Biol.* **119** 1205–18
[24] Jackle H *et al* 1986 Cross-regulatory interactions among the gap genes of Drosophila *Nature* **324** 668–70
[25] Wu X, Vakani R V and Small S 1998 Two distinct mechanisms for differential positioning of gene expression borders involving the Drosophila gap protein giant *Development* **125** 3765–74
[26] Wieschaus E, Nusslein-Volhard C and Jurgens G 1984 *Roux Arch. Dev. Biol.* **193** 267–82
Wieschaus E, Nusslein-Volhard C and Jurgens G 1984 *Roux Arch. Dev. Biol.* **193** 283–95
Wieschaus E, Nusslein-Volhard C and Jurgens G 1984 *Roux Arch. Dev. Biol.* **193** 296–307
[27] Houchmandzadeh B and Wieschaus E, unpublished data
[28] St Johnson D, Beuchle D and Nusslein-Volhard C 1991 Staufen, a gene required to localize maternal RNAs in the Drosophila egg *Cell* **66** 51–63
[29] Eldar A *et al* 2002 Robustness of the BMP morphogen gradient in Drosophila embryonic pattering *Nature* **419** 304–8