DEVELOPMENTAL GENETICS 23:11–27 (1998)

Stripe Forming Architecture of the Gap Gene System

JOHN REINITZ,[1]* DAVID KOSMAN,[1] CARLOS E. VANARIO-ALONSO,[1] AND DAVID H. SHARP[2]

[1]*Brookdale Center for Molecular Biology, Mt. Sinai School of Medicine, New York, New York*

[2]*Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico*

ABSTRACT In this report, we show that gap genes encode exactly one set of pair-rule stripes, which occur in the native *even-skipped* position. The core of this work is a detailed analysis that shows how this conclusion follows from the arrangement of gap domains in the embryo. This analysis shows that: (1) pattern forming information is transmitted from gap to pair-rule genes by means of a nonredundant set of morphogenetic gradients, and (2) the stripe forming capability of the gap genes is constrained by the arrangement of these gradients and by the fact that each gap domain consists of a pair of correlated gradients. We also show that in the blastoderm, the regulatory sign of a transcriptional regulator is unlikely to change in a concentration dependent manner. The principal analytic tool used to establish these results is the gene circuit method. Here, this method is applied to examine hybrid data sets consisting of real gene expression data for four gap genes and hypothetical pair-rule expression data generated by translating native *even-skipped* data along the anterior-posterior axis. In this way, we are able to investigate the stripe forming capabilities of the gap gene system in the complete absence of pair-rule cross regulation. We close with an inference about evolutionary development. It is argued that the constraints on gap gene architecture identified here are a consequence of selective pressures that minimize the number of gap genes required to determine segments in long-germ band insects. Dev. Genet. 23:11–27, 1998. © 1998 Wiley-Liss, Inc.

Key words: pair-rule genes; even-skipped; positional information; morphogenetic gradients; gene circuit; simulated annealing; computational biology; *Drosophila*

INTRODUCTION

A critical step in the determination of the periodic segments in the fruit fly *Drosophila* is the transformation of aperiodic positional information encoded by gap domains and maternal gradients into the periodic pattern of pair-rule gene expression [Howard and Ingham, 1986; Goto *et al.*, 1989; Harding *et al.*, 1989; Howard and Struhl, 1990; Warrior and Levine, 1990; Gutjahr *et al.*, 1993; Klingler and Gergen, 1993; Yu and Pick, 1995]. This transformational step is modulated by pair-rule cross regulation [Carroll and Scott, 1986; Harding et al., 1986; Carroll et al., 1988; Carroll and Vavra, 1989; Hooper et al., 1989; Baumgartner and Noll, 1990]. The central question addressed here is the identification of those features of pair-rule expression patterns that are controlled by the gap gene system as opposed to those that are determined by pair-rule cross regulation. An important part of the analysis is the determination of those elements of the gap domain architecture that are utilized in an essential way in this determinative process.

The difficulty of this problem arises because pair-rule genes have many inputs. Their combined effect leads to a complex but precisely positioned set of overlapping patterns [Carroll *et al.*, 1988]. The correct set of overlaps is absolutely required for embryonic viability. The pair-rule genes *fushi-tarazu* (*ftz*) and *even-skipped* (*eve*) are normally expressed in separate, complementary stripes at gastrulation [Frasch and Levine, 1987], but any overlap in expression at this time results in lethality [Frasch *et al.*, 1988].

An analysis of the relative contributions of gap and pair-rule genes to the pair-rule patterns would begin by identifying which periodic patterns can be generated by the gap gene system alone. This question could be answered by a simple experiment. Monitor the expression of each pair-rule gene in embryos mutant for all of the seven other pair-rule genes. Such an experiment is probably not feasible by genetic methods, because the animal would die as a heterozygote.

The difficulty of constructing a multiply mutant genetic stock can be circumvented by a hybrid approach that utilizes modeling and computation in addition to experimental data. This approach, which we call the gene circuit method, provides a way to represent the

Contract grant sponsor: NIH; Contract grant numbers: RR07801 (JR, DK, DHS) and HD27937 (CEV-A).

Carlos E. Vanario-Alonso is now at the Universidade Federal do Rio de Janeiro, Instituto de Biofisica Carlos Chagas filho, Laboratorio de Biologia Molecular Maury Miranda, Rio de Janeiro, Brazil.

*Correspondence to: John Reinitz, Brookdale Center for Molecular Biology, Box 1126, Mt. Sinai School of Medicine, One Gustave L. Levy Place, New York NY 10029. E-mail: reinitz@kruppel.molbio.mssm.edu

Received 1 April 1998; Accepted 15 May 1998

interactions in a network of genes. It takes protein concentrations as observable state variables and describes their variation in space and time by equations for the rate of protein synthesis. Unknown parameters in these equations are determined by fits to experimental data. The gene circuit method has been described in previous work [Mjolsness *et al.*, 1991; Reinitz *et al.*, 1995; Reinitz and Sharp, 1995]. The utility and validity of this method have been demonstrated by its successful application to several important problems, including the correct prediction that *hunchback* (*hb*) and *bicoid* (*bcd*) act together to specify positional information in the *Drosophila* blastoderm [Reinitz *et al.*, 1995]. Further, the gene circuit methodology enabled us to identify the specific gap gene expression domains that controlled each of the eight borders of *eve* stripes 2–5 to predict the early *eve* transient pattern and to understand why *eve* stripes form at the time they do. The analysis also led to an understanding of why *eve* and other pair-rule mRNA's must be confined to an apical compartment [Reinitz and Sharp, 1995].

In this previous work, gene circuits were used to uncover the details of how a directly observed expression pattern was established. The analysis used in that work depended on a particular biological fact: *eve* is the only pair-rule gene whose stripes are neither eliminated nor shifted in mutants for other pair-rule genes prior to gastrulation [Frasch and Levine, 1987; Carroll and Vavra, 1989]. Had this not been true, the analysis would have required the inclusion of other pair-rule genes in the model. This implies that *eve* ought to be the only pair-rule gene all of whose stripes can be made by gap gene input alone, although to our knowledge this question has never been directly investigated because of the technical difficulties mentioned above.

In this report, we use the gene circuit method to investigate whether a hypothetical arrangement of expression domains is biologically realizable. Specifically, we investigate whether or not the gap gene system can encode periodic patterns of stripes that are shifted relative to one another by one or more nuclei. Also, by numerical modeling of the expression patterns of four gap genes, but only a single pair-rule gene, we can ask about the stripe forming capabilities of the gap gene system in the complete absence of pair-rule cross regulation. We find that gap genes encode exactly one set of stripes, in the native *eve* position.

The core of this work is the detailed analysis that shows how our conclusions follow from the arrangement of gap domains in the embryo. This analysis shows that the essential positional information in the gap domain system is contained in a nonredundant set of morphogenetic gradients. The stripe forming capability of the gap genes is constrained both by the arrangement of the gradients and by the fact that a single gap domain consists of pairs of correlated gradients.

MATERIALS AND METHODS

Theoretical and Numerical Methods

The gene circuit method has been carefully described elsewhere [Reinitz *et al.*, 1995; Reinitz and Sharp, 1995, 1996]; the main ideas are reviewed here. The computational model takes gene expression data as input and gives a regulatory circuit as output. We represent a circuit by the elements of a matrix **T**. Each element T^{ab} of this matrix characterizes the regulatory effect of one gene on another by a single real number for each possible pair a and b. Thus if T^{ab} is positive, gene b activates gene a; if T^{ab} is negative, gene b represses gene a, and if T^{ab} is zero, gene b has no effect on gene a. In a very basic sense, this is a minimal model of gene regulation: We do not know at the outset which gene regulates which other gene. Consequently, we must allow for the possibility that each gene regulates every other gene. For N genes, this leads to an $N \times N$ matrix.

Most important developmental processes are too complex to permit useful computational modeling at present. The *Drosophila* blastoderm is an exception for three reasons. First, the blastoderm is a syncytium, so that cell-cell signaling can be neglected. Spatial interactions can be treated in terms of a diffusive coupling between blastoderm nuclei. Second, segment determination happens because of differential gene expression, which is directly observable. The segmentation genes have been cloned, and hence their level of expression can be monitored by antibody methods. Third, the system of segmentation genes is not coupled to other developmental processes until after gastrulation, because morphological alterations due to mutations in segmentation genes are not apparent before gastrulation. The second and third reasons together mean that the state variables of this system are directly observable. This is a very unusual occurrence in biology. It affords an opportunity to understand certain features of morphogenesis at an unprecedented level of detail.

We can use the concentrations of protein products of gap and pair-rule genes as state variables while neglecting RNA, because we are aware of no evidence for a direct role of RNA in the regulation of zygotic segmentation genes. Similarly, we do not take into account the phosphorylation state of segmentation gene proteins. In the part of the blastoderm that gives rise to segments, the expression of segmentation genes is a function of position along the anterior-posterior (A-P) axis only. This means that even though the blastoderm is an ovoid, the essential features of the segment determination process are well described by considering a one-dimensional strip of nuclei running along the A-P axis.

In the blastoderm, changes in concentrations of proteins are governed by three basic processes: (1) direct regulation of protein synthesis from a given gene by the protein products of other genes (including auto-regulation as a special case), (2) transport of molecules

between cell nuclei, and (3) decay of protein concentrations.

We combine these considerations into a coarse-grained chemical kinetic equation as follows. Let the position of a cell nucleus along the A-P axis be indexed by i, such that nucleus $i + 1$ is immediately posterior to nucleus i. Each cell nucleus contains a copy of a regulatory circuit composed of N genes and which is determined by an $N \times N$ matrix $\mathbf{T}$. The concentration of the ath gene product in nucleus i is a function of time, denoted by $v_i^a(t)$. Then,

$$\frac{dv_i^a}{dt} = R_a\, g_a\left(\sum_{b=1}^{N} T^{ab} v_i^b + m^a v_i^{bcd} + h^a\right) + D^a(n)[(v_{i-1}^a - v_i^a) + (v_{i+1}^a - v_i^a)] - \lambda_a v_i^a, \quad (1)$$

where N is the number of zygotic genes included in the circuit. The first term on the righthand side of the equation describes gene regulation and protein synthesis, the second describes exchange of gene products between neighboring cell nuclei, and the third represents the decay of gene products.

In (1), T^{ab} is the previously discussed matrix of genetic regulatory coefficients. The *bcd* input is given by $m^a v_i^{bcd}$, where v_i^{bcd} is the concentration of *bcd* protein in nucleus i and m^a is the regulatory coefficient of *bcd* acting on zygotic gene a. g_a is a "regulation-expression function," which we assume takes the form $g_a(u^a) = (1/2)[u/\sqrt{u^2 + 1}) + 1]$ for all a, where the total input to gene a is $u^a = \Sigma_{b=1}^{N} T^{ab} v_i^{bcd} + h^a$. R_a is the maximum rate of synthesis from gene a, and h^a summarizes the effect of general transcription factors on gene a. The diffusion parameter $D^a(n)$ depends on the number n of cell divisions that have taken place and varies inversely with the square of the distance between nuclei. We assume that the distance between adjacent nuclei is halved after a nuclear division. λ_a is the decay rate of the product of gene a. Nuclear divisions are incorporated by shutting down synthesis for a time equivalent to one mitosis and doubling the number of nuclei.

At the outset, we do not know the values of the T^{ab} and the other parameters in (1). We do know the experimentally observed *solutions* of (1): they are simply the gene expression patterns observed. For fixed initial conditions, the solutions of (1) depend on what parameters are chosen: We seek the set of parameters that minimize the summed squared deviations between the observed data and the solutions of (1), which are obtained by numerical integration of this system of equations. This leads to a least-squares optimization problem, which we solve by the method of simulated annealing. The advantage of this method is that under appropriate conditions, it will reliably give the global minimum. This advantage is purchased at the cost of intensive computation. In part, this is because a single annealing run requires several million integrations of the equations. Also, because simulated annealing is a stochastic method, it is necessary to do repeated fits to ensure that the answer is consistent and reliable. All the fits performed for this work were done using the annealing schedule described [Reinitz and Sharp, 1995, Appendix], under conditions appropriate for convergence to the global minimum.

When the gene circuit method produces a fit to data in which pair-rule stripes form under gap control, it is straightforward to assign control of stripe borders to particular regulators. This is a significant strength of the approach and is done as follows. A stripe border represents a spatial location across which the state of activation of a pair-rule gene switchs from off to on, or vice versa. Consequently, the input to pair-rule gene a, u^a above, changes from negative to positive across such a border. By graphically examining combinations of regulators, as well as u^a itself, in the vicinity of a border one can unambiguously determine which regulator or regulators lead to the change of sign of u^a. Usually a single regulator is unambiguously responsible (Fig. 3).

Experimental

Rabbit anti-Eve serum was obtained from M. Frasch. Guinea pig anti-Hb was raised as described, and staining and microscopy were performed as described [Kosman and Reinitz, 1998]. Photometric studies were performed by analyzing each channel independently as follows. The image of the entire embryo was cropped to a narrow A-P strip a few nuclei wide extending along the center of the image. The cropped image was convolved with a Gaussian of radius equal to about a half nucleus to eliminate the effect of punctate staining, and the average pixel values around the center of each nucleus were read off by hand.

RESULTS

Gap Genes Encode Only One Set of Stripes

The first step in our strategy for investigating the unique role of *eve* in the pair-rule system is to check whether gene circuit models share with the biological system the property that only pair-rule stripes in the *eve* position can be generated by the gap gene system alone. We can use gene circuits as a tool for exploring this question only if this property is present in the gene circuit model without having been explicitly introduced a priori—i.e., if it is implicitly present in the expression patterns. Once this fact has been established, we can analyze the gene circuit to discover why the gap gene system has this property.

Our strategy for accomplishing the first step is based on approximating the stripe pattern of a variety of pair-rule genes by the expression of a single pair-rule gene, shifted along the A-P axis relative to the gap domains. We are lead to this approach by the following argument. The most straightforward way to examine stripe forming capabilities of the gap gene system

would be to use experimental data about gap expression domains and the expression pattern of each pair-rule gene, one per data set. As yet, we have mapped only the expression of four gap genes and the pair-rule gene *eve* with respect to one another. However, the stripe forming capabilities of the gap gene system can still be analyzed using this data if the following approximation is made. The initial striped patterns of all pair-rule genes, with the exception of *prd*, are quite similar, containing seven stripes each 3–4 nuclei wide. They differ primarily with respect to the location of the stripes relative to one another. We model this property by considering a data set derived from our experimental data on *eve* and gap genes by translating the *eve* stripes relative to the gap domains along the A-P axis. Because this data has a periodicity of seven nuclei, this procedure defines seven possible locations for the pair-rule stripes, relative to the gap domains. The question we address here is whether the gap genes can form stripes at each of these seven locations.

We begin with a data set used in an earlier study [Reinitz and Sharp, 1995]. This data set was generated from fluorescence double-antibody-stained embryos. The data described a row of 32 nuclei running along the lateral equator in an A-P direction, extending from the middle of *eve* stripe 1 to the interstripes between *eve* stripes 5 and 6. Note that this will limit our analysis, approximately, to the five most anterior pair-rule stripes. Data were obtained for *eve*, *Kruppel*, (*Kr*), *giant* (*gt*), *knirps* (*kni*), *hb*, and *bcd*. We used gap gene data at the early, middle, and late parts of the period between the thirteenth synchronous nuclear division and gastrulation, and *eve* expression data were included at the early and late parts of this period, as well as during cleavage cycle 13. The *eve* expression data in cleavage cycle 13 and early cleavage cycle 14A [Foe and Alberts, 1983] were spatially uniform. In the late cleavage cycle 14A data, *eve* is approximately periodic, with the spatial period of seven nuclei mentioned above.

The hypothetical pair-rule expression data was generated by translating the native *eve* data anteriorly or posteriorly by one or more nuclear positions. We needed to do this only for the data in late cleavage cycle 14A, because only at this time (within the time interval being studied) was the pattern nonuniform. The late cleavage cycle 14A data sets generated in this way are shown in the lefthand column of Figure 1, whereas the righthand column shows the displacements of stripes relative to the position of native *eve* stripes. We denote a set of stripes displaced one nucleus anterior from the position of the native *eve* stripes as "1+," stripes displaced one nucleus posterior as "1−," and so on. We refer to this collection of seven stripe positions as a family of stripe datasets.

Thus we have a hybrid data set consisting of actual experimental data for four gap genes and the native *eve* pattern and synthetic data derived from translated *eve* patterns. Fits to this family of data sets were performed in a number of ways. In one set of fits, all parameters were determined by a least-squares procedure. The data sets used identical wild-type and mutant pair-rule and gap gene expression data in order to take into account the fact that gap domains are unaltered in single mutants for pair-rule genes. This results in circuits in which connections from the pair-rule gene back to the gap gene are close to zero, so that the pair-rule gene does not contribute to formation of the gap genes' expression pattern. Thus a family of stripe data sets differs only with respect to pair-rule stripe location, not gap gene domains. For this reason, it is also possible to fix the gap gene circuit and fit only to parameters specifically associated with the pair-rule gene. Fits done in this way gave the same results as fits to the entire set of parameters, including those associated with gap genes, with the advantage of significantly greater computational efficiency.

The results of a set of fits performed with gap gene parameters described elsewhere [Reinitz and Sharp, 1995] are shown in the middle column of Figure 1. Note that a full set of five stripes forms only in the 0 data set (native *eve* position) and the 1− data set. Closer inspection of the 0 and 1− fits, shown in Figure 2, indicates that although the 1− data is displaced by one nucleus, the model forms stripes at the native *eve* position rather than the position of the data. The chief effect of the displacement in the data is to make the stripes less distinct, in the sense that there is more expression in the interstripes.

The fits in which stripes do not form in the locations specified by the data indicate the impossibility of stripe formation at those locations and hence that the configuration specified by the hypothesized data is forbidden by regulatory constraints. In the next section we present an analysis that makes this result an understandable consequence of the architecture of the gap domain system.

Why the Gap Gene System Can Make Only One Set of Stripes

We have shown that a gene circuit model captures a key property of the gap gene system: Gap genes by themselves encode only one set of pair-rule stripes. Now we discuss why this is so.

To show why the gap genes encode only one set of pair-rule stripes using a two-step analysis, we first ask whether each stripe in each member of the family of stripe data sets can be made individually. Since different stripes form by different mechanisms, if it is impossible to form any individual stripe contained in a given periodic pattern, then it is impossible to form the entire pattern. Moreover, a single stripe is itself composed of two borders. A single stripe cannot form unless both of its borders form, so if a stripe fails to form in a particular location, we check to see if one or both of its borders fails to form correctly. This amounts to assaying the ability of the gap gene system to make

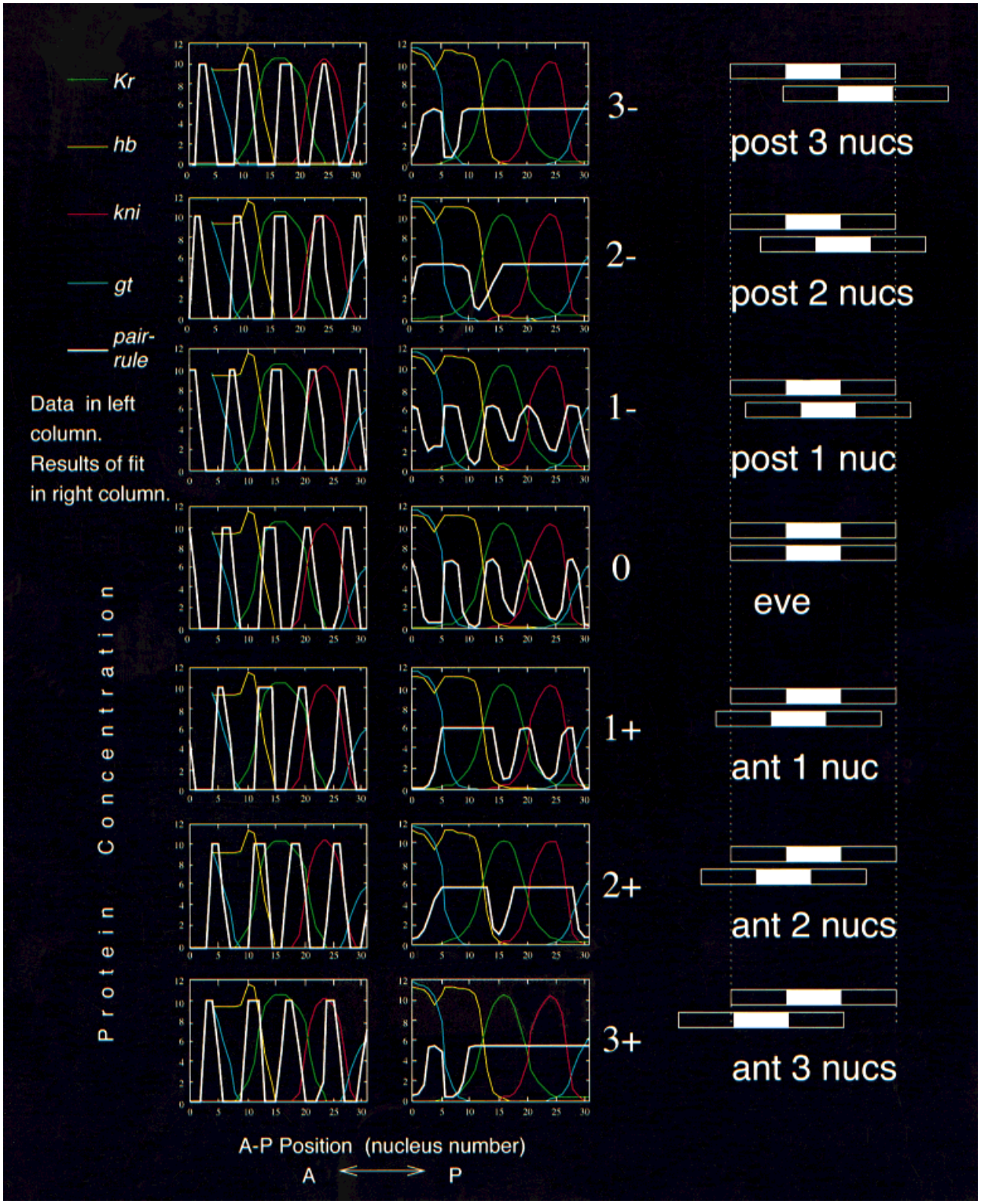

Fig. 1. Results of fitting the gene circuit equations to the seven possible positions of pair-rule stripes translated from the *eve* pattern. Each row shows the data and circuit results for a single position of pair-rule stripes and is labeled accordingly. "0" denotes native eve stripe placement, "1+" denotes stripe placement one nucleus anterior to the native *eve* position, "1−" denotes stripe placement one nucleus posterior to the native *eve* position, and so on. The lefthand column shows the data sets used, the central column the results of fits, and the righthand column the displacement of pair-rule stripes from the native *eve* position. In the lefthand column there is no gap expression data for the four most anterior nuclei, but our gap gene data set is complete for earlier times [Reinitz *et al.*, 1995]. Each graph shows expression levels at the onset of gastrulation. The vertical axis of each graph shows relative protein concentration. The horizontal axis of each graph shows position along the anterior-posterior axis, with anterior to the left, in terms of the number of nuclei posterior to the middle of *eve* stripe 1. One nucleus is ~1% egg length. Each graph shows the expression of *Kr*, *gt*, *kni*, *hb*, and the hypothetical "pair-rule" gene, as shown in the key on the left.

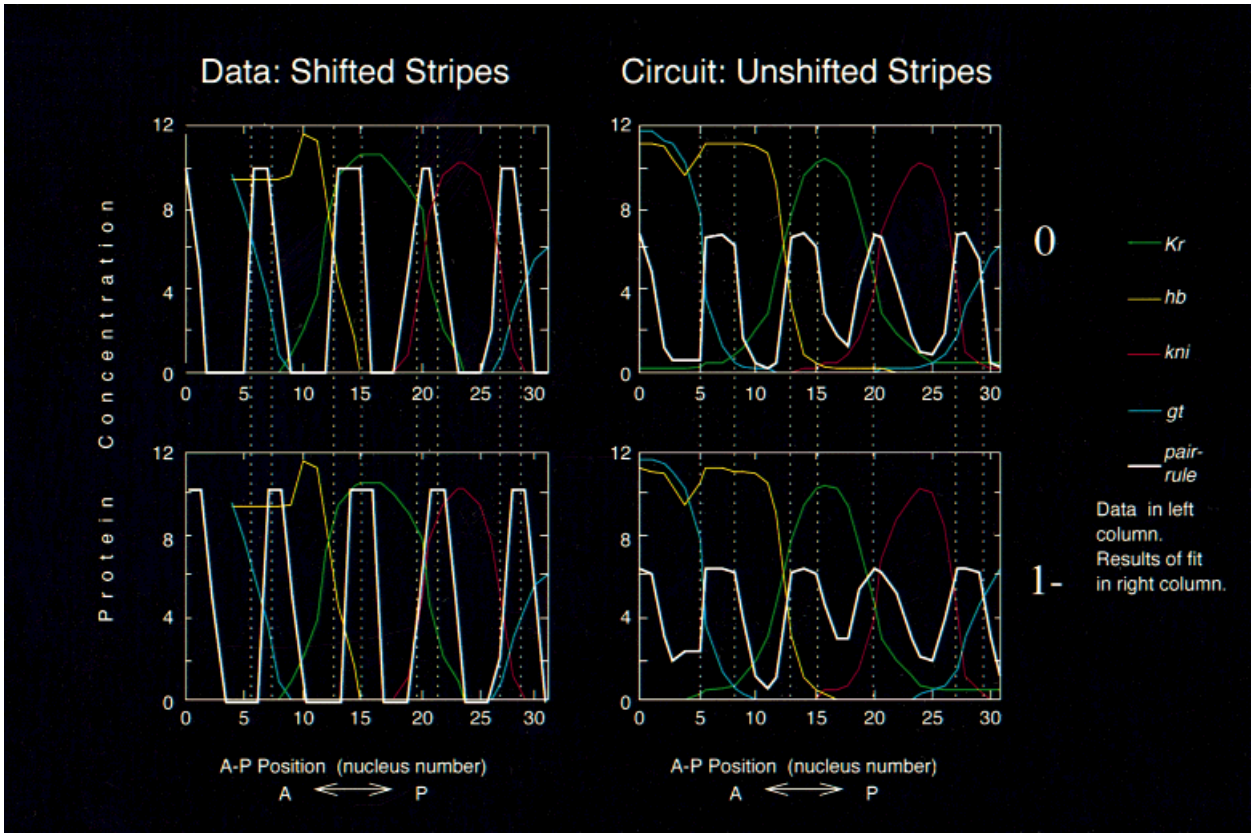

Fig. 2. Magnified view of the data and circuit results for the "0" (native *eve*) and "1−" stripe placements shown in Figure 1. The dotted lines facilitate comparison of the relative locations of stripe features between the two datasets and the two circuits. The stripes in the two input data sets are displaced by one nucleus, but the stripes in the two output circuits are perfectly aligned.

anterior and posterior borders at every location in the region under consideration. A principal result of our simulations is that in the absence of additional pair-rule input, the anterior borders of pair-rule stripes are generated by the posterior slopes of gap domains, whereas the posterior borders of pair-rule stripes are generated by the anterior slopes of gap domains. In general, there is at most one such gradient at each location. In regions of the blastoderm where these positional cues are not present, a stripe cannot form, and hence the periodic pattern containing that stripe cannot form. We show that this local constraint eliminates five out of seven members of the family of stripe data sets. We then show how a global constraint eliminates one of the two remaining sets of stripes.

Local constraints. From each of the seven members of the family of stripe data sets, we generated a subfamily of 4–5 new *single stripe* data sets. Each of these contained the same gap data as the original periodic data set and the same pair-rule data in cycle 13 and early cycle 14A. In late cycle 14A, each dataset contained pair-rule data consisting of a single stripe and the surrounding interstripes, a region 8 nuclei wide. This amounts to asking for a circuit with a given pair-rule pattern in an 8-nucleus-wide stripe and leaving expression elsewhere unconstrained. To avoid weighting the fit toward uniform expression in cycle 13 and early 14A, we actually used four copies of the single stripe data at the four 1.33 minute timesteps immediately preceding gastrulation, so as to maintain a balance of 32 striped datapoints (late cycle 14A) to 64 uniform datapoints (cycle 13 and early 14A). For similar reasons we ensured that each single stripe dataset contained exactly 8 nuclei, although it should be noted that for stripes near the edge of the region considered (e.g., stripe 5 in 3−), this results in the inclusion of a small amount of expressing region from an adjacent stripe. An example of such a dataset and the results of a fit to it are shown in Figure 3.

As demonstrated elsewhere [Reinitz and Sharp, 1995], it is a relatively straightforward task to determine which regulators set the borders of a stripe if there is a gene circuit that makes that stripe. The first step is to note that a stripe border forms when $u^{\text{pair-rule}}$ changes sign. Next, we check which genes dominate the contribution to $u^{\text{pair-rule}}$ in that region. Large contributors to

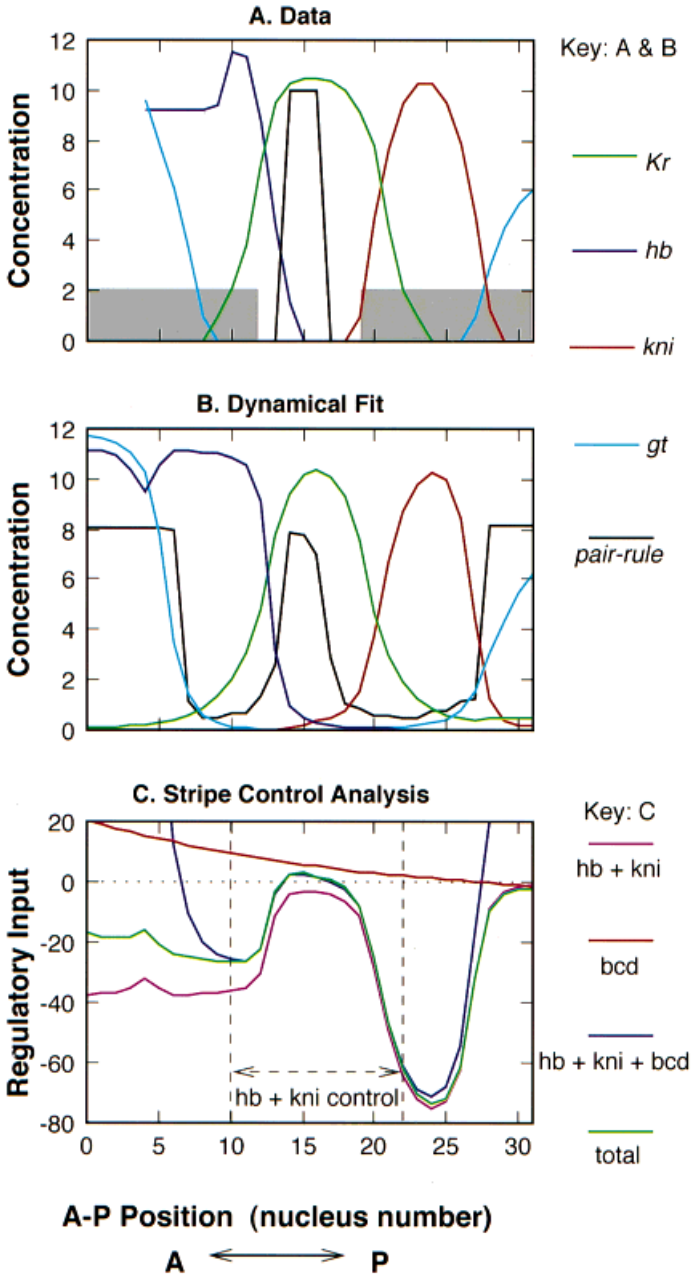

Fig. 3. Three steps in the analysis of control of a single stripe: stripe 3 of the 1− data set. **A.** Input data. Axes are as in Figure 1, color coding as shown. The gray bars on the bottom show where there was no pair-rule data. **B.** Gene circuit fit to the data; axes as in Figure 1. **C.** Stripe control analysis. The x-axis shows anterior-posterior position as described for Figure 1. The y-axis shows regulatory input to a hypothetical pair-rule gene (dimensionless units). The magenta line represents $T^{pr\leftarrow kni}v^{kni} + T^{pr\leftarrow hb}v^{hb}$, where *pr* represents the hypothetical pair-rule gene. This is the main repressive input to the stripe. The red line represents $m^{pr}v^{bcd}$, which is the main activating input to the stripe. The net effect of these three inputs is shown by the the green line $T^{pr\leftarrow kni}v^{kni} + T^{pr\leftarrow hb}v^{hb} + m^{pr}v^{bcd}$, and the total input to the pair-rule gene by the blue line u^{pr}. Where the blue line overlaps the green line, *kni*, *hb*, and *bcd* provide all input to the pair-rule gene leaving it under the effective control of *kni* and *hb*. The region of this overlap is marked "kni + hb control."

$u^{\text{pair–rule}}$ may vary with position or be nearly constant; only spatially varying regulator concentrations carry the information necessary to establish a border. Note that a regulator that is spatially unvarying around a stripe (e.g., Hb around *eve* stripe 2) may be indirectly essential to the formation of that stripe, even if it does not set the border, by establishing the proper sensitivity for the regulator that does set the border to act. These points are illustrated in Figure 3.

The studies are summarized in Figure 4. In Figure 4A, the genes that control anterior stripe borders at each position are shown graphically with the gap gene pattern superimposed. The figure shows that anterior borders controlled by a given gap gene are located on the posterior slope of that gap domain. Similarly, Figure 4B shows that the posterior borders controlled by a gap gene are located on the anterior slope of that gap domain. We also see that there is at most one concentration gradient of a given polarity at each location. Note there are no borders in Figure 4B in the vicinity of nucleus 10, meaning that the model did not allow a border to be formed at that location. It can be seen that this region is approximately co-extensive with a domain whose lefthand boundary is marked by

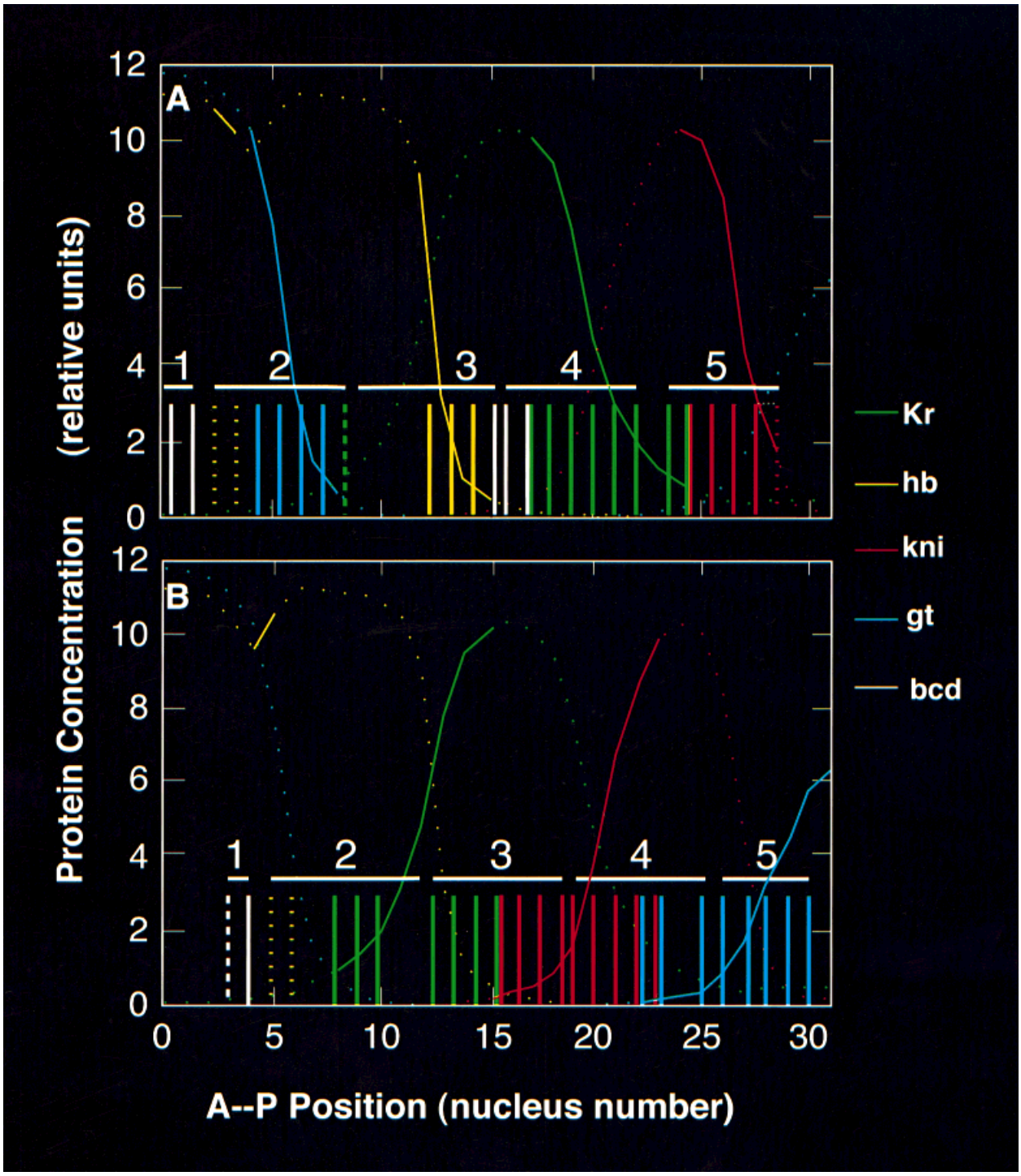

Fig. 4. Control of anterior and posterior pair-rule stripe borders by gap genes. **A.** Anterior borders. **B.** Posterior borders. Each graph shows gap gene expression patterns (solid and dotted colored curves) and border control information (thick colored tickmarks). The vertical and horizontal axes are as in Figure 1. Solid tickmarks are placed at the locations of stripe borders for the appropriate stripe. Solid tickmarks indicate a repression-generated border, dashed tickmarks denote an activation-generated border, and dotted tickmarks indicate repression-generated borders we believe to be artifacts of the model (see text). Occasionally, control of a border involved the cooperation of two gap genes; these cases are shown by two colored tickmarks in contact. The absence of a tickmark shows that no border could form in that position. The gap gene expression levels are drawn as solid lines in regions where they control borders and dotted elsewhere. The tickmarks and expression domains are color-coded as shown.

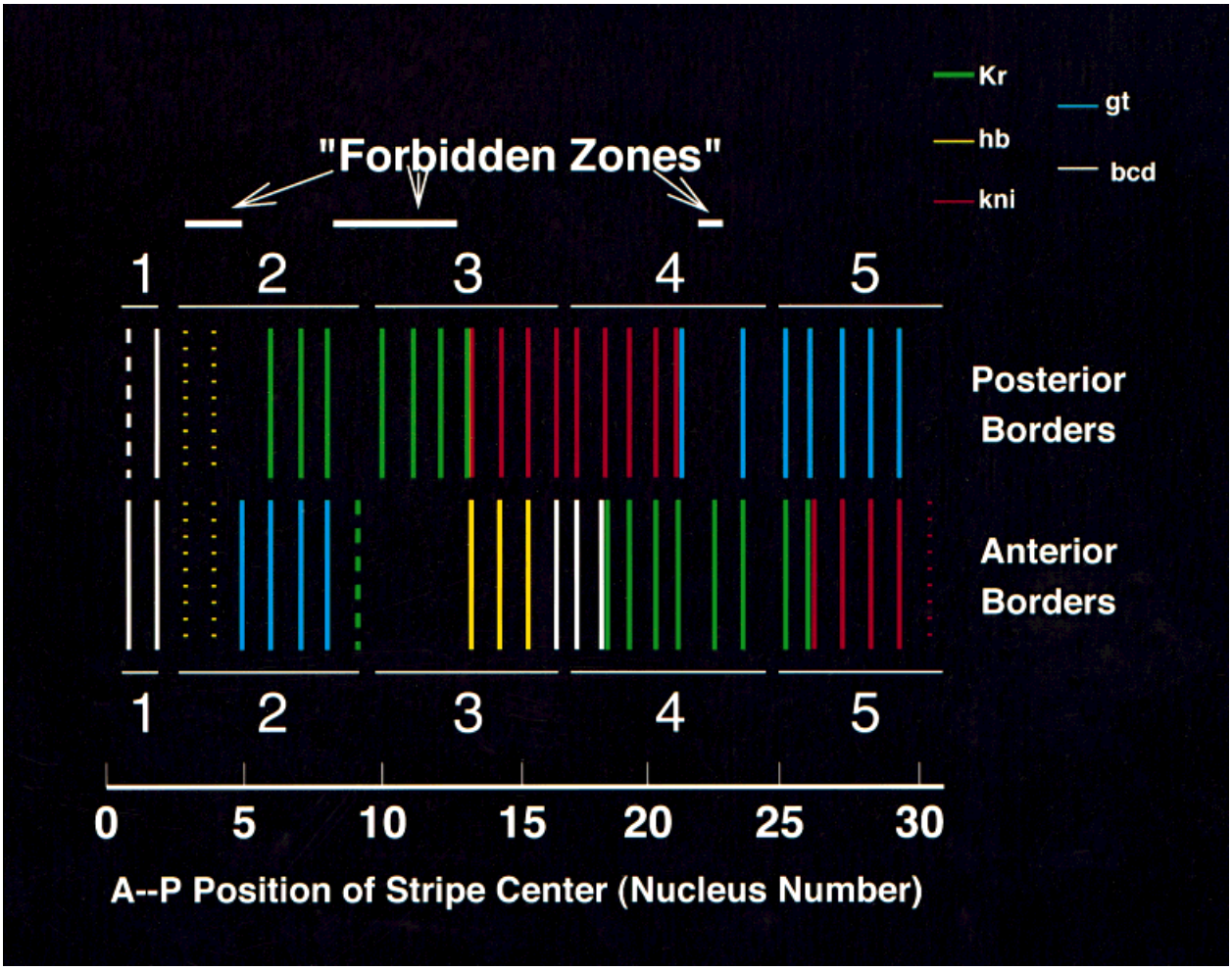

Fig. 5. Stripe border control and the forbidden zones. Two tick marks denoting the control of anterior (top) and posterior (bottom) borders, if formed, are placed at the locations of stripe centers for the appropriate stripe. Solid tick marks indicate a repression-generated border, dashed tick marks denote an activation generated border, and dotted tick marks indicate repression-generated borders. We believe that the borders represented by dashed and dotted lines to be spurious (see text). Occasionally, control of a border involved the cooperation of two gap genes; these cases are shown by two colored tick marks in contact. Three forbidden zones, shown at the top, are regions where a border could not form.

the near vanishing of the *gt* gradient and whose right-hand boundary approximates the location of the steep *hb* gradient.

Figure 5 shows that there are regions of the embryo where stripe borders, and hence stripes, cannot form. We refer to these as "forbidden zones." Figure 5 also shows the regulators that set borders when borders can form. We note that each member of the family of seven stripe data sets, with the exception of 0 and 1−, contains at least one stripe that lies in a forbidden zone. These are exactly the stripe data sets that failed to give a full set of stripes in Figure 1.

There are two borders in Figure 4A and two borders in Figure 4B that we believe to be spurious, in the sense that they are generated by erroneous features of the gap gene circuit. They are both borders of stripe 2 in positions 2+ and 3+ and are denoted by dotted tick marks. Specifically, we call attention to the "dip" in the *hb* pattern in the vicinity of nucleus 5. This dip reflects an error in the gene circuit: Its occurrence appears to be connected with the timing of the formation of pair-rule stripes in the sense that circuits not containing this feature form stripes much earlier than is observed experimentally. However, the dip reduces *hb* repression and allows the formation of a stripe by an incorrect mechanism.

We confirmed this by repeating the entire single stripe analysis with a different fixed gap gene circuit. Note that we cannot perform this analysis simply by using experimental gap gene expression data. The behavior of the model reflects the entire course of expression dynamics from cleavage cycle 11 to the onset

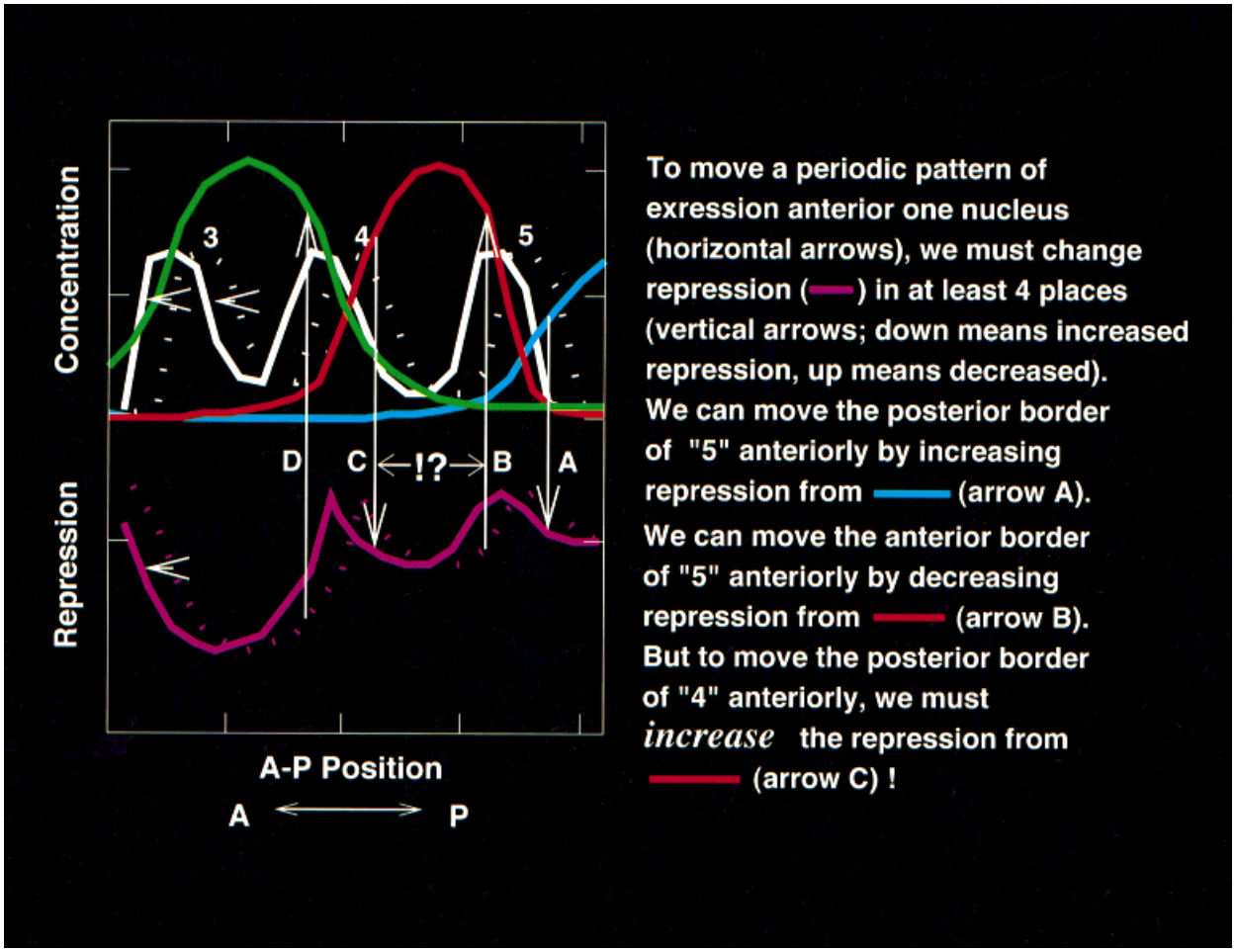

Fig. 6. How symmetric gradients lead to global constraints.

of gastrulation, whereas our expression data consists of "snapshots" at fixed time. The circuit (data not shown) contained an anterior *hb* domain with no noticeable errors, although the *Kr* domain was significantly wider (2–3 nuclei on each side at ~20% maximum expression). This circuit gave a poorer fit overall, but a better fit to the anterior *hb* domain. With this gene circuit, stripe 2 in the 2+ position has its anterior border set by *bcd* repression and its posterior border by the spuriously expanded *Kr* domain. Stripe 2 in the 3+ position has its anterior border set by *bcd* repression and its posterior border is set by *gt* activation. In both of the gap gene circuits used in this analysis, the control of the posterior border of stripe 2 in the 2+ position is due to an erroneous feature of the circuit for the gap gene expression pattern. For this reason, we believe that it is not possible for a border to form at this location, in view of the fact that our gap gene data (see Fig. 1) shows no gradients at this point.

Global constraints. The two periodic patterns of pair-rule stripes that do not lie in any forbidden zones are 0 and 1−. Here, we show how global constraints over the entire gap gene system prevent stripes from being expressed in the 1− position. Our major point is that in a system in which a periodic pattern of stripes is generated from a series of bell-shaped expression domains by repression, it is impossible to shift the pattern of stripes in a coordinated way by changing the strength of repression from these domains. This strong argument for the uniqueness of a gap gene generated pair-rule stripe pattern is supplemented by a more heuristic argument that indicates why stripes in the 0 position are preferred to those in the 1− position.

Figure 6 shows a pair of stripes under the control of three gap genes. Imagine that one wished to translate each of these stripes by one nucleus in an anterior direction. To do that, it is necessary to move each of four borders anterior by one nucleus. The posterior border of stripe 5 can be moved anteriorly by increasing repression from the blue gap gene (arrow A), and its anterior border can be moved anteriorly by decreasing repression from the red gap gene (arrow B). To shift stripe 4

anteriorly, however, we must decrease repression from the green gap gene (arrow D) and *increase* repression from the red gene (arrow C)—in direct contradiction to what had to be done to shift stripe 5. The difficulty is that each gap domain forms a *pair* of gradients that, because they are in fact made of the same morphogen, have the same regulatory action. Since the anterior slope controls posterior pair-rule borders and the posterior slope controls anterior borders, changing the repressive coupling will make a pair of borders on either slope move together or apart, but it cannot make them both move in the same direction. Decreasing the concentration of a gap gene product has the same effect as reducing its repressive strength an appropriate amount at fixed concentration. It is interesting to note that this "coming together" of stripes is seen is *Kr* heterozygotes, where Kruppel concentration is presumably lowered to half its normal level [Frasch and Levine, 1987; Warrior and Levine, 1990].

Given that only one set of stripes should be expected, why is the 0 position to be preferred to 1−? Consider stripe 3. The anterior border of stripe 3 in the 1− position lies at a lower concentration of Hb than the anterior border of stripe 3 in the 0 position, and thus requires greater repression from Hb. At the same time, stripe 2 in the 0 or 1− position is in a region where Hb is expressed at a high level. In the native *eve* case [Reinitz and Sharp, 1995], this Hb repression is balanced by increased activation by Bcd. The balance, however, is a delicate one that would be disrupted by increased Hb repression. This problem is actually more acute than the circuit results would indicate: In our data (Fig. 2), there is a "lip" of Hb expression in the region of presumptive parasegment 4 at the edge of the large anterior domain. This lip has never been reproduced by circuits, but stripe 2 in the 1−position would overlap with it.

Validity of the Gene Circuit Model

The above analysis, although conceptually satisfying, depends on the fundamental assumptions of the gene circuit model. In this section we discuss the credibility of the assumptions in the light of current information. We also present additional quantitative data on the relative concentrations of Eve and Hb in the neighborhood of *eve* stripes 2 and 3, which supports them. Careful consideration of these quantitative results combined with an assessment of evidence in the literature lends support to our modeling assumptions, as applied to the blastoderm of *Drosophila*.

Does the gene circuit model describe the real world? A frequently misunderstood point about modern molecular genetics concerns the amount of ambiguity in the form of the regulation-expression function of an intact eucaryotic gene. It would appear correct to assume that the the synthesis rate S of the product v of gene a must depend on the concentration of regulators, so that

$$S = f([reg_1], \ldots [reg_i], \ldots [reg_n]),$$

and a discussion of the validity of the gene circuit model must certainly include a discussion of the correct form for f. In fact, there is *no* direct experimental information available on the form of f for any intact segmentation gene and almost none for any other eucaryotic gene transcribed by RNA polymerase II. In the absence of such information, indirect methods must be used, as discussed below.

One key question about f is whether an activator of a given gene also can be its repressor. Here we use the word "repression" to denote any instance where increasing the concentration of a regulator lowers the transcription rate of its target; we do not distinguish here between the mechanisms of classical repression, quenching, or squelching.

Two major lines of evidence concerning the form of f come from studies of embryos containing transformed *lac*Z promoter constructs, on one hand, and tissue culture transfections, on the other. To consider what is learned from transformant lines, it is useful to consider the *eve* promoter itself as a particularly well studied example (summarized in Fig. 7) and to focus on the regulatory action of *hb* on *eve*. Available evidence indicates that Hb has an activating effect on MSE2 [Goto *et al.*, 1989; Stanojevic *et al.*, 1991; Small *et al.*, 1991, 1992] and a repressive effect on MSE3 [Small et al., 1993, 1996]. What is the effect of Hb on the intact gene? In other words, suppose the concentration of Hb protein rose 10% in the vicinity of an *eve* gene in its native chromosomal milieu. Would *eve* transcription rise (Hb an activator) or fall (Hb a repressor)?

This question is central to eucaryotic molecular genetics, because it amounts to asking what are the rules by which the regulatory activities of single enhancers are combined to give the regulative properties of an intact gene. The MSE3 element of the *eve* promoter illustrates the difficulty well: it behaves differently from the intact promoter in *kni* mutants [Frasch and Levine, 1987; Small *et al.*, 1993, 1996; Reinitz and Sharp, 1995], and its behavior is altered when it is placed adjacent to MSE2 [Small *et al.*, 1993].

No mechanistic picture that does not account for these phenomena can possibly be complete. In particular, small-scale mechanistic models invoking mechanisms such as dimerization of a ligand [Sauer and Jackle, 1993, e.g.] do not address interactions between enhancers. In general, it is widely supposed that in a reasonable model of transcription, the effect of a given protein on a promoter depends on the presence of other proteins and that these other proteins may exert synergistic effects and possibly change the regulatory sign of a given protein. Within the context of the gene circuit model, the sigmoidal function $g(u)$ in equation (1) provides for modulation of the regulatory effect of one regulator according to the presence or absence of oth-

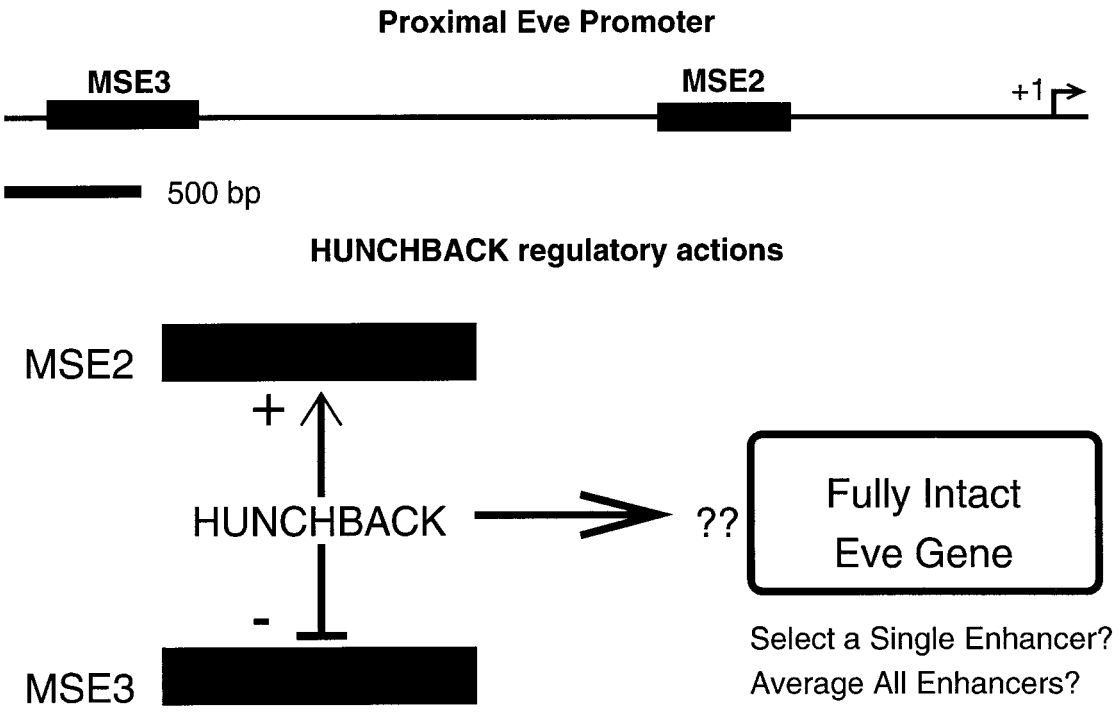

Fig. 7. Schematic diagram of Hb regulator actions on the proximal eve promoter. This information has been summarized from the literature [Goto *et al.*, 1989; Harding *et al.*, 1989; Stanojevic *et al.*, 1991; Small *et al.*, 1991, 1992, 1993, 1996].

ers: Such effects can range from synergistic potentiation of regulatory effect (if other proteins move *u* close to threshold) all the way to the abolition of regulatory effect (if other proteins move *u* far from threshold). The cooperative effects of dimerization or cooperative binding are representable by a large value of the appropriate T^{ab}, leading to a steep sigmoid similar to a higher order Hill function [Hill, 1985]. The gene circuit model does not, however, provide for a change in regulatory sign. It does not explicitly consider the presence of individual enhancers, but one can imagine the gene circuit equations as approximations to a more precise (and much more complicated) model in which all binding interactions were included. This would argue for a picture in which regulatory actions at MSE2 and MSE3 are somehow averaged to give the net regulatory action for the intact gene, which the gene circuit method identifies as a net repressor of *eve*.

It is possible to argue that MSE2 is preferentially associated with the TATA complex in stripe 2 and that MSE3 is associated with the TATA complex in stripe 3, but that begs the question because it does not explain how the respective enhancers are selected. Nevertheless, the differing regulatory actions of Hb on MSE2 and MSE3 suggest that the regulatory sign of this protein on *eve* can change. This change in sign, should it take place, must be dependent on the difference in concentrations of some set of factors in the vicinity of stripe 2 as compared to those in the vicinity of stripe 3. In fact, the blastoderm contains a very limited set of spatially distributed factors at the time the *eve* stripes form: gap genes and dorsal-ventral gene products. The only difference between stripe 2 and stripe 3 is the relative concentration of Hb itself and of Kr. This fact implies that any change in sign of the regulatory action of Hb on *eve* must be dependent on concentrations of Hb itself or of Kr. This inference is reinforced by results from tissue culture transfection experiments [Sauer and Jackle, 1991; Zuo *et al.*, 1991] where the regulatory sign of both Hb and Kr have each been shown to change in a concentration-dependent manner.

We believe that this argument for a change of regulatory sign is incorrect. To explain, we first acknowledge that we know of no biophysical reason why the regulatory sign cannot change. Moreover, the regulatory sign of an activator *must* change over a sufficiently high concentration range. This point is easy to understand—any activator that is present at sufficiently high concentrations (e.g., half of total cellular protein) will be toxic to the cell, and hence there must be a certain concentration of each activator that gives a maximum rate of transcription. The real question concerning the credibility of the assumption used in the gene circuit model is whether this change of sign occurs at *physiological* concentrations. We argue here that it does not.

First, we note that the only direct evidence for a change of regulatory sign comes from transfection studies in tissue culture cells. Until a satisfactory

method of quantitating intra-cellular fluorescence is developed, there is no way directly to relate the dosage of transfecting DNA to the intracellular concentration of effector protein elicited. Moreover, in one tissue culture study where Schneider cells transfected with Zen protein were afterward monitored using fluorescent antibody techniques, it was found that 90% of the cells contained no effector protein [Han *et al.*, 1989]. This indicates that the average levels of effector protein generated in the culture were substantially less than the actual intracellular levels that effected the measured CAT expression. In this regard it is significant that the change in regulatory sign observed for both Kr [Sauer and Jackle, 1991] and Hb [Zuo *et al.*, 1991] took the form of activation at low concentrations followed by repression at high concentration, which we expect to be generic for unphysiologically high concentrations of activators.

The argument *in favor* of the form of *f* used in the gene circuit model comes, not from biophysics but from the selective pressure of evolution. The segment determination system appears to have evolved so as to be insensitive to a variety of perturbations, including variations in gene dosage. In particular, all segmentation genes are recessive lethals. In heterozygotes, a full set of pair-rule, *en*, and *wg* stripes are produced, with the sole alteration being a slight compression of pair-rule stripes in the vicinity of the heterozygotic gap domain [Frasch and Levine, 1987]. This puts serious constraints on possible reversals of regulatory sign. Below, we give an argument relating to Hb and based on experimental data that indicates why we believe a change of regulatory sign in the blastoderm is unlikely to occur.

Photometric studies of levels of *hb* and *eve* expression were performed on six late cycle 14A embryos. A typical example of such an embryo is shown in Figure 8, together with photometric measurements of the fluorescence signals for Hb and Eve. Note that the level of Hb expression in the vicinity of stripe 2 is nearly maximal, but that the expression level at the anterior margin of stripe 3 is about half maximal. If the regulatory sign of Hb on *eve* is concentration-dependent, one would expect that at half maximal concentration, Hb would act as a repressor and at maximal concentration as an activator. However, in a *hb* heterozygote, the maximal Hb concentration is about half that seen in wild type but all seven *eve* stripes form normally. But then *stripe 2 is "activated" by the same concentration of Hb we had just determined to be "repressive."* This is a contradiction, which shows that the regulatory sign of Hb is not concentration-dependent.

We chose Hb rather than Kr for this analysis because its regulatory action on both MSE2 and MSE3 is unambiguously determined. We note, however, that one objection that could be raised to the above argument is that there is also a maternal component to the anterior Hb expression, and so the net Hb concentration in a late

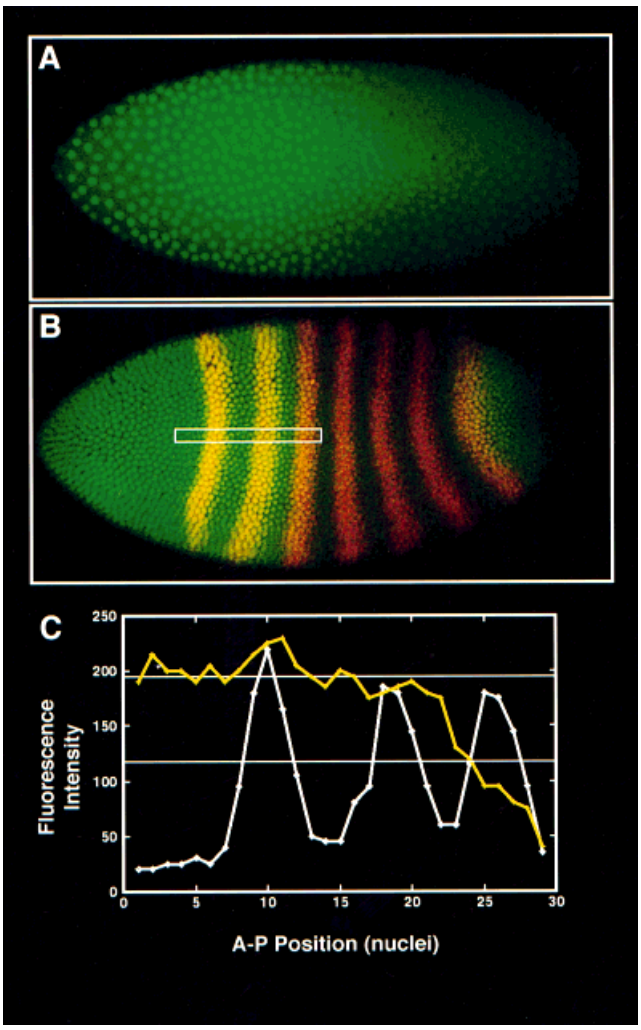

Fig. 8. *eve* and *hb* expression. **A.** Cleavage cycle 12 embryo fluorescently stained for Hb protein. At this embryonic stage, Hb protein originates from maternally expressed transcripts. Note that expression extends across two thirds of the embryo. **B.** Late cleavage cycle 14A embryo fluorescently stained for Eve (red) and Hb (green) proteins. Note the restriction of the anterior Hb domain compared to A. In both A and B, anterior is to the left. Panel C was generated from the boxed area. **C.** Average pixel values of nuclei along a 28 nucleus strip of the embryo shown in B (see boxed area). Hb levels are denoted by a yellow line, Eve levels by a white line. Two white lines cross the graph. The upper one shows maximal Hb levels. The lower one shows half maximal Hb levels. The strip extends from a point in the presumptive head region to the interstripe between eve stripes 3 and 4. The vertical axis shows average fluorescence intensity values in dimensionless 8 bit units and the horizontal axis shows A-P position as measured by nucleus number.

cleavage stage 14A embryo would be more than half its concentration in a heterozygote. In fact, the maternal Hb contribution has decayed to an undetectable level by late cleavage cycle 14A, as shown by a comparison of the posterior limits of Hb expression in the cleavage cycle 12 embryo, also shown in Figure 8 with the posterior limits of late cycle 14A expression. Maternal

Hb expression in cycle 12 extends farther posteriorly than total (maternal plus zygotic) expression in late stage 14A, indicating that the maternal contribution has decayed away by this time.

DISCUSSION

We have established four main results in this report. First, the transition from aperiodically expressed gap genes to the periodic pattern of pair-rule stripes happens because stripe borders are established by nonredundant gradients of gap gene products. Second, repression by posteriorly sloping gradients forms the anterior borders of pair-rule stripes, and repression by anteriorly sloping gradients forms posterior borders. These gradients are said to be nonredundant because at most, one such gradient can form a pair-rule border at each point along the A-P axis. In certain regions along the A-P axis, there is no suitable gradient, and stripes cannot form in such a region. Third, we have shown there is a global constraint in the placement of stripes that originates from the fact that anteriorly and posteriorly sloping morphogenetic gradients of a gap domain are coupled, because they are part of the same expression domain. Fourth, we have given arguments that show that it is highly unlikely that the regulatory sign of a transcriptional regulator changes in a concentration-dependent manner.

In this section we assess the validity of our analysis on the basis of comparison to results on pair-rule genes in the literature. We find that our analysis is well supported overall, but that limitations appear when it is applied to fine scale features of expression patterns. We present independent evidence for the existence of forbidden zones in the predicted positions, but also point to data that are in conflict with the analysis at the resolution of a single nucleus. We then point to possible implications of our findings for the evolution of long-germ band insects.

Forbidden Zones in the Primary Pair-Rule Gene System

Here, we discuss the correctness of our results in light of the known biology of *eve, runt* (*run*), and *hairy* (*h*). We focus on these so-called "primary pair-rule genes" because members of this group are not regulated prior to gastrulation by any pair-rule genes outside the group [Carroll and Scott, 1986]. The three "primary" pair-rule genes are unique in that they are the only pair-rule genes whose response to gap genes in the absence of pair-rule cross regulation can be determined. This is done by monitoring the expression of one "primary" pair-rule gene in embryos mutant for the other two. Hence their behavior is most easily interpretable in light of the results presented here.

Whether the behavior of other pair-rule genes supports or argues against the results presented here is difficult to assess because of the complexity of the inputs such genes receive. *ftz* is of particular importance. Early suggestions that gap input to the rest of the pair-rule system was mediated solely through "primary" genes [Ingham and Martinez-Arias, 1986; Howard and Ingham, 1986], based largely on studies of *ftz* and *h*, have now been superseded [Yu and Pick, 1995]. The work of Yu and Pick [1995] shows that *ftz* initially establishes seven stripes in the correct order in mutants for any "primary" pair-rule gene, strongly indicating direct input to *ftz* from the gap gene system. Nevertheless, the positions of *ftz* stripe borders are altered in *run* mutants. *ftz* stripes 3 and 4 in *run*– appear narrower than normal [Yu and Pick, 1995, compare Figs. 3H and 1G). There is also the issue of overall change of phase in pair-rule mutants. At 31°C, the allele eve^{ID19} has been reported to be functionally null [Nusslein-Volhard *et al.*, 1985], while still expressing detectable Eve protein [Frasch *et al.*, 1988]. In embryos of this genotype, *ftz* stripes are expressed one nucleus posterior to *eve* stripes instead of in the *eve* interstripes, as is the case in wild-type embryos. Taken together, the above results indicate that *ftz* receives direct patterning input from gap genes and "modulatory" input from pair-rule genes, but it is difficult to say more. Similar ambiguities exist for other "secondary" and "tertiary" pair-rule genes. For example, *paired* (*prd*) exhibits nearly uniform expression in *eve* mutants [Baumgartner and Noll, 1990; Gutjahr *et al.*, 1993], but studies of its promoter clearly indicate the presence of direct input from the gap gene system [Gutjahr *et al.*, 1993]. Understanding the interplay of regulation by the gap and pair-rule systems in genes like *ftz* and *prd* is vital for gaining an understanding of the pair-rule system, but such complexity is beyond the scope of the present work. Here, we focus on *eve*, *run*, and *h*. The phenomenology of the regulation of these genes is simple enough to give an indication of the correctness of our results.

An important question is whether the behavior of *eve*, *h*, and *run* is consistent with the existence of forbidden zones. There is good reason to believe that this is so in the case of *h*, which is expressed in stripes that are located approximately one nucleus anterior to those of *eve* [Hooper *et al.*, 1989; Howard and Struhl, 1990; Warrior and Levine, 1990], in approximately the 1+ position. There is some variation in the displacement, so that *h* stripes 1, 2, and 7 are displaced by two nuclei, whereas stripes 4 and 5 are displaced by slightly less than one nucleus [Hooper *et al.*, 1989]. *h* cannot form all of its stripes from gap gene cues alone. Specifically, in the absence of pair-rule cross regulation (*eve; run* double mutants), *h* stripe 2 does not form, nor do the interstripes between *h* stripes 3 and 4 or 6 and 7 [Ingham and Gergen, 1988; Carroll and Vavra, 1989; Hooper *et al.*, 1989]. These six *h* stripe borders are clearly under pair-rule control, and in fact *h* stripe 2 is *eve*-dependent and the 3–4 and 6–7 interstripes are *run*-dependent. *h* stripe 2 lies in a forbidden zone, and the

fact that it requires *eve* input to form supports the existence of this forbidden zone. We did not analyze data extending to stripes 6 and 7, so that the *run* dependence of that interstripe cannot be compared to our results. In the case of the *h* 3–4 interstripe, the model does not preclude gap gene control of the interstripe nor does it require it.

We next consider *run*. Since *run* is expressed posterior to *eve* in the 1− or 2− position, we are primarily concerned with *run* stripe 2, which is expressed in the forbidden zone posterior to *eve* stripe 2. We are not aware of any studies of *run* in an *eve*−; *h*− double mutant background, but changes in the relative spacing of stripes 1–4 have been detected in *eve*− [Klingler and Gergen, 1993]. These data are ambiguous with regard to the forbidden zones, because seven stripes still exist, but their borders have shifted. Thus a pair-rule controlled border lying in a forbidden zone might shift by one or two nuclei in a pair-rule mutant to a place where it can respond to gap cues. Some additional complexities in the *run* pattern relevant to later times or the ventral region of stripe 1 [Ingham and Gergen, 1988] are not discussed here, because they do not pertain to forbidden zones.

Analysis of the *run* promoter is consistent with results presented here. The only *run* stripes controlled by discrete enhancers are stripes 1, 3, and 5 [Klinger *et al.*, 1996]. A proximal 5 kb region of promoter expresses 7 stripes when fused to *lac*Z, and these stripes are strongly dependent on *eve* and to a lesser extent on *run*. The relative contributions of the seven stripe element and the discrete elements are obscure in the intact gene, but there is no evidence that *run* stripe 2 is controlled by gap genes and partial evidence that it is controlled by *eve*. This is consistent with the presence of a forbidden zone posterior to *eve* stripe 2, but the evidence is not as strong as that supporting a forbidden zone anterior to *eve* stripe 2.

Gap Gene Control in the Primary Pair-Rule Gene System

The result that gap genes generate pair-rule stripes via repression by nonredundant gradients is consistent with earlier interpretations of experimental data [Carroll and Vavra, 1989; Hooper *et al.*, 1989; Carroll, 1990; Langeland *et al.*, 1994]. It provides a unifying picture of the pattern-forming capacity of the gap gene system. In principle it does considerably more. Figure 4 amounts to a prediction of which gap gene controls a pair-rule stripe (where such control is possible) at every location between the middle of *eve* stripe 1 and the interstripe between *eve* stripes 5 and 6. In fact, the current level of resolution of the model is not sufficient to provide reliable predictions at this level of resolution. Here, we discuss examples where the predictions are correct and others where they are incorrect and consider possible sources of error.

As a first example of a prediction with mixed results, we discuss *run* stripes 3 and 5. *run* is expressed in a pattern roughly complementary to *h*, with the *run* stripes occurring in a position 1–2 nuclei posterior to *eve* [Kania *et al.*, 1990]. This corresponds to data in the 1− or 2− positions. Figure 5 indicates that *run* stripe 3 should be regulated at its anterior margin by *hb* and at its posterior margin by *kni*, and that stripe 5 is regulated on the anterior margin by *kni* and on the posterior margin by *gt*. For stripe 3 border control, the model's predictions are consistent with experiment, but for stripe 5 it was found experimentally that although the posterior border was set by *gt*, the anterior border was set by *kni* [Klingler *et al.*, 1996]. In this case the model was in error, but it is instructive to note that the anterior border of *run* stripe 5 lies only two or three nuclei posterior to the point where the model predicts control by *Kr* should turn into control by *kni*.

Similar situations occur with *h*. For example, the model predicts that *h* stripe 4 (1+) should have its anterior margin controlled by *Kr* and its posterior margin by *kni*. In fact, the posterior margin is indeed under control by *kni* [Hartmann *et al.*, 1994], but the anterior border is under pair-rule control by *run* [Ingham and Gergen, 1988; Carroll and Vavra, 1989; Hooper *et al.*, 1989; Hartmann *et al.*, 1994]. Here, it is interesting to note that the location of the anterior margin of stripe 4 of the 1+ register (Fig. 5) is only 2 nuclei anterior of the zone of repressive *bcd* control lying on the broad apex of the *Kr* domain. We speculated that this region of *bcd* control was spurious and represented a forbidden zone; thus the results for *h* stripe 4 also suggest a broadly correct picture, but one subject to error at a fine scale of analysis.

One last example is of interest. We reported in earlier work that the posterior border of *eve* stripe 3 was under the control of *Kr* based on a fit to a data set with multiple stripes [Reinitz and Sharp, 1995], consistent with results of expression studies of the intact *eve* gene in gap mutants [Frasch and Levine, 1987]. In this work, we found that the posterior margin of a *single stripe* in the *eve* stripe 3 position (0) was set cooperatively by *kni* and *Kr*. By this we mean that the inputs of *hb*, *Kr*, and *kni* were all contributing to $u^{\text{pair-rule}}$ over the region containing stripe 3 and its immediate interstripe neighborhood; the observed cooperativity is phenomenological and implies nothing about the presence or absence of specific cooperative mechanisms. For this reason, the stripe 3 (0) posterior border tickmarks in Figures 4B and 5 each have two colors, indicating that *Kr* and *kni* both control the posterior border. This result is in near agreement with an experimental study of the enhancer that expresses stripes 3 and 7 [Small *et al.*, 1996], which showed that the posterior margin of stripe 3 as expressed by the minimal enhancer is controlled by *kni*. We believe it significant that a fit to periodic stripe data gave results consistent with experimental work on the intact gene and its periodic pattern, whereas fits to

single stripe data gave results much closer to experimental work on single stripes. It also indicates, with respect to both theory and experiment, that under some circumstances the behavior of an individual stripe alone will differ from its behavior as part of a periodic pattern.

Let us summarize potential sources of error in our analysis. (1) Errors in input data. We mention especially that data for pair-rule stripes, other than *eve*, were obtained by translation of the native *eve* stripes. Other inaccuracies in input data, including uncertainties in scale factors, are discussed elsewhere [Reinitz and Sharp, 1995]. (2) The analysis is based on *fits* to data, which may contain errors. (3) An analysis of border control done on a stripe-by-stripe basis may give different results from one performed on an intact embryo. Taking into account these sources of error, we believe the level of agreement between the model predictions and biological facts is striking.

The key importance of the gene circuit method in this problem is that it makes possible an entirely new type of analysis of expression data. In order to explore the capacity of the gap gene system to form a stripe border at an arbitrary location *experimentally*, one would need a selective assay for the ability of enhancers to form such a border. Such experiments are not presently feasible.

Evolutionary Origin of Constraints on Gap Gene Architecture

Our findings have a bearing on important evolutionary questions. One of these questions concerns the extent to which specific characteristics exhibited by living species reflect historical contingencies, as opposed to fundamental constraints imposed by the organism's developmental organization [Buss, 1987]. The results discussed here are illustrative of a constraint that results from the inherent structure of a well-characterized regulatory network.

As background for our observations, we note that it is believed [Sander, 1975, 1988] that insects with long germ bands like *Drosophila* evolved from short germ band insects that determined segments sequentially through the growth of the segmented germ band. Most segmentation genes play a role in the nervous system of both long and short germ band insects, and this was probably their original role. The existence of intermediate germ band species such as *Tribolium*, in which more anterior segments are determined in a syncytium and the posterior segments are generated by growth, suggests that segmentation genes were gradually recruited into syncytial segment determination. Syncytial segment determination is usually interpreted as selection for speed of development [Sander, 1975, 1988]. The specification of segments by means of morphogenetic gradients in a syncytium in turn imposes requirements for patterning accuracy. One of the conclusions of this work is that it is very unlikely that morphogenetic gradients change regulatory sign. We propose that this property is necessary to ensure accuracy and reliability of developmental patterning in a naturally varying population.

Selection for speed of development, although maintaining accuracy, also provides a framework for interpreting other constraints on the architecture of the gap gene system identified in this report. We note that the process by which multiple steep gap protein gradients are formed from broader maternal gradients tends to produce bell-shaped domains. Because it is impossible to have a concentration discontinuity in a syncytium, the only way to make a gradient that is not bell shaped is to have a one-sided, flat-topped domain like the anterior Hb domain. But gradients of both polarities are needed, so implementing flat-topped domains would require recruiting twice as many gap genes into segment determination. This is presumably why most gap domains are bell-shaped. Moreover, given a bell-shaped domain, the presence of the same genome, and hence the same set of promoters, in nuclei on either side of the domain requires both of its slopes to be morphologically active. This fact, combined with the accuracy constraint that prevents a regulatory change of sign, imposes a restriction on stripe placement. This results from global constraints of the type described here, which prevent the occurrence of stripes in the 1− position (see Fig. 6).

These global constraints thus appear to implement a developmental strategy that minimizes the number of gap genes required. We next point out that this global constraint is the key determinant of the architectural constraints in the gap gene system reported here. Of the six members of the family of stripe data sets that could not form seven stripes, five could not do so because of the local constraint of forbidden zones, and the sixth (1−) could not do so because of a global constraint. The global constraint, however, is remarkably strong. It says that bell-shaped domains acting by repression can make exactly one set of pair-rule stripes. If all of the gap domains were bell-shaped, they could make exactly one set of pair-rule stripes. If such a system were to be rendered capable of specifying two sets of pair-rule stripes, it would need another full set of bell-shaped domains. Forbidden zones would thus be a side effect of having recruited close to the minimum possible number of gap genes needed to make a periodic pattern. The existence of the global constraint itself is a consequence of the same thing, because determining segments with flat-topped domains also would increase the number of genes that would need to be recruited into segment determination.

ACKNOWLEDGMENTS

We thank M. Frasch for the gift of antibodies; M. Frasch, J.P. Gergen, G. Marnellos, and S. Roth for helpful comments on the manuscript; and V. Friedrich for help and advice on confocal microscopy. This is

publication # 200 from the Brookdale Center for Molecular Biology.

REFERENCES

Baumgartner S, Noll M (1990): Network of interactions among pair-rule genes regulating *paired* expression during primordial segmentation of *Drosophila*. Mech Dev 33:1–18.

Buss LW (1987): "The Evolution of Individuality." Princeton, NJ: Princeton University Press.

Carroll SB (1990): Zebra patterns in fly embryos: Activation of stripes or repression of interstripes? Cell 60:9–16.

Carroll SB, Laughon A, Thalley BS (1988): Expression, function, and regulation of the *hairy* segmentation protein in the *Drosophila* embryo. Genes Dev 2:883–890.

Carroll SB, Scott MP (1986): Zygotically active genes that affect the spatial expression of the *fushi tarazu* segmentation gene during early *Drosophila* embryogenesis. Cell 45:113–126.

Carroll SB, Vavra SH (1989): The zygotic control of *Drosophila* pair-rule gene expression: II. spatial repression by gap and pair-rule gene products. Development 107:673–683.

Foe VA, Alberts BM (1983): Studies of nuclear and cytoplasmic behaviour during the five mitotic cycles that precede gastrulation in *Drosophila* embryogenesis. J Cell Sci 61:31–70.

Frasch M, Levine M (1987): Complementary patterns of *even-skipped* and *fushi tarazu* expression involve their differential regulation by a common set of segmentation genes in *Drosophila*. Genes Dev 1:981–995.

Frasch M, Warrior R, Tugwood J, Levine M (1988): Molecular analysis of *even-skipped* mutants in *Drosophila* development. Genes Dev 2:1824–1838.

Goto T, MacDonald P, Maniatis T (1989): Early and late periodic patterns of *even skipped* expression are controlled by distinct regulatory elements that respond to different spatial cues. Cell 57:413–422.

Gutjahr T, Frei E, Noll M (1993): Complex regulation of early *paired* expression: Initial activation by gap genes and pattern modulation by pair-rule genes. Development 117:609–623.

Han K, Levine MS, Manley JL (1989): Synergistic activation and repression of transcription by *Drosophila* homeobox proteins. Cell 56:573–583.

Harding K, Hoey T, Warrior R, Levine M (1989): Autoregulatory and gap gene response elements of the *even-skipped* promoter of *Drosophila*. EMBO J 8:1205–1212.

Harding K, Rushlow C, Doyle HJ, Hoey T, Levine M (1986): Cross-regulatory interactions among pair-rule genes in *Drosophila*. Science 233:953–959.

Hartmann C, Taubert H, Jackle H, Pankratz MJ (1994): A two-step mode of stripe formation in the *Drosophila* blastoderm requires interactions among primary pair-rule genes. Mech Dev 45:3–13.

Hill TL (1985): "Cooperativity Theory in Biochemistry: Steady-State and Equilibrium Systems." New York: Springer-Verlag.

Hooper KL, Parkhurst SM, Ish-Horowicz D (1989): Spatial control of *hairy* protein expression during embryogenesis. Development 107: 489–504.

Howard K, Ingham P (1986): Regulatory interactions between the segmentation genes *fushi tarazu, hairy,* and *engrailed* in the *Drosophila* blastoderm. Cell 44:949–957.

Howard K, Struhl G (1990): Decoding positional information: Regulation of the pair-rule gene *hairy*. Development 110:1223–1231.

Ingham P, Gergen P (1988): Interactions between the pair-rule genes *runt, hairy, even-skipped,* and *fushi-tarazu* and the establishment of periodic pattern in the *Drosophila* embryo. Development 104 (Suppl): 51–60.

Ingham PW, Martinez-Arias A (1986): The correct activation of *antennapedia* and bithorax complex genes requires the *fushi tarazu* gene. Nature 324:592–597.

Kania MA, Bonner AS, Duffy JP, Gergen JP (1990): The *Drosophila* segmentation gene *runt* encodes a novel nuclear regulatory protein that is also expressed in the developing nervous system. Genes Dev 4:1701–1713.

Klingler M, Gergen JP (1993): Regulation of runt transcription by *Drosophila* segmentation genes. Mech Dev 43:3–19.

Klingler M, Soong J, Butler B, Gergen JP (1996): Disperse versus compact elements for the regulation of *runt* stripes in *Drosophila*. Dev Biol 177:73–84.

Kosman D, Reinitz J (1998): Rapid preparation of a panel of polyclonal antibodies to *Drosophila* segmentation proteins. Dev, Genes, Evol (in press).

Langeland JA, Attai SF, Vorwerk K, Carroll SB (1994): Positioning adjacent pair-rule stripes in the posterior *Drosophila* embryo. Development 120:2945–2955.

Mjolsness E, Sharp DH, Reinitz J (1991): A connectionist model of development. J Theor Biol 152:429–453.

Nusslein-Volhard C, Kluding H, Jurgens G (1985): Genes affecting the segmental subdivision of the *Drosophila* embryo. Cold Spring Harbor Symposiums on Quantitative Biology 50:145–154.

Reinitz J, Mjolsness E, Sharp DH (1995): Cooperative control of positional information in *Drosophila* by *bicoid* and maternal *hunchback*. J Exp Zool 271:47–56.

Reinitz J, Sharp DH (1995): Mechanism of formation of eve stripes. Mech Dev 49:133–158.

Reinitz J, Sharp DH (1996): Gene circuits and their uses. In Collado J, Magasanik B, Smith T (eds): "Integrative Approaches to Molecular Biology," chap. 13, Cambridge: MIT Press, pp 253–272.

Sander K (1975): Pattern specification in the insect embryo. In Porter R, Rivers J (eds): "Cell Patterning," Vol. 29. Ciba Foundation Symposia, pp 241–263.

Sander K (1988): Studies in insect segmentation: From teratology to phenogenetics. Development 104 (suppl):111–121.

Sauer F, Jackle H (1991): Concentration-dependent transcriptional activation or repression by *Kruppel* from a single binding site. Nature 353:563–566.

Sauer F, Jackle H (1993): Dimerization and the control of transcription by *Kruppel*. Nature 364:454–457.

Small S, Arnosti DN, Levine M (1993): Spacing ensures autonomous expression of different stripe enhancers in the *even-skipped* promoter. Development 119:767–772.

Small S, Blair A, Levine M (1992): Regulation of *even-skipped* stripe 2 in the *Drosophila* embryo. EMBO J 11:4047–4057.

Small S, Blair A, Levine M (1996): Regulation of two pair-rule stripes by a single enhancer in the *Drosophila* embryo. Dev Biol 175:314–324.

Small S, Kraut R, Hoey T, Warrior R, Levine M (1991): Transcriptional regulation of a pair-rule stripe in *Drosophila*. Genes Dev 5:827–839.

Stanojevic D, Small S, Levine M (1991): Regulation of a segmentation stripe by overlapping activators and repressors in the *Drosophila* embryo. Science 254:1385–1387.

Warrior R, Levine M (1990): Dose-dependent regulation of pair-rule stripes by gap proteins and the initiation of segment polarity. *Development* 110:759–767.

Yu Y, Pick L (1995): Non-periodic cues generate seven ftz stripes in the *Drosophila* embryo. Mech Dev 50:163–175.

Zuo P, Stanojevic D, Colgan J, Han K, Levine M, Manley JL (1991): Activation and repression of transcription by the gap proteins hunchback and Kruppel in cultured *Drosophila* cells. Genes Dev 5:254–264.