Dynamic modeling of genes for spatial patterning in embryo development on the example of the Drosophila segmentation gene hunchback

Alexander V. Spirov

State University of New York, Stony Brook, Stony Brook, New York, USA, Alexander.Spirov@sunysb.edu

Ekaterina M. Myasnikova

St.Petersburg Polytechnical University, St.Petersburg, Russia, myasnikova@spbcas.ru

David M. Holloway

British Columbia Institute of Technology, Burnaby, B.C., Canada, David_Holloway@bcit.ca

Background. The hunchback (hb) gene, which forms an anterior-posterior (AP) pattern in early *Drosophila melanogaster* development and controls the positioning of downstream segmentation genes, has been the subject of many recent studies quantifying the temporal and spatial details of its expression (see references in Holloway & Spirov, 2015). It has also been the subject of a number of theoretical projects concerning how zygotic expression is controlled by maternally-derived spatial signaling gradients, in particular how hb is activated by the Bicoid (Bcd) gradient. The degree to which hb is understood, from sequence-level information of its cis-regulatory modules (CRMs) to the temporal and spatial dynamics of its expression (and of its transcriptional regulators), allows for the development and testing of detailed quantitative models for the different aspects of control (e.g. maternal-, self-, and cross-regulatory) in this classic example of spatial gene expression in development. Since the Hb protein itself forms a spatial gradient controlling several expression boundaries in downstream segmentation genes (Hülskamp et al., 1990; Yu and Small, 2008), precise control of its expression is critical in the robust development of flies and other insects (Bonneton et al., 1997; McGregor et al., 2001).

One of the fundamental developmental problems being studied with hb is how expression maintains robustness in the face of numerous sources of variability and noise. Earlier theoretical (Lacalli and Harrison, 1991) and experimental work (Houchmandzadeh et al., 2002) indicated that Hb expression is far more precise, in spatial positioning, than the

upstream Bcd maternal signal. This has now been greatly refined, towards an understanding of the many factors involved in Hb precision and synchronized expression (Bergmann et al., 2007; Manu et al. 2009; Perry et al., 2011; Porcher et al., 2010; He et al., 2011; He and Ma, 2013; Liu and Ma, 2013). Our work has focused on the dynamics of hb expression in the 50 minute nuclear cleavage cycle 14A (NC14), during which hb and the other gap-class genes reach their highest expression, downstream targets (*pair-rule* and *segment polarity* genes) are expressed and refined, and cellularization occurs. Developing a theoretical framework from these models involves both extensive simulations and mathematical analysis (sensitivity analysis (Gutenkunst et al., 2007; Bieler et al., 2011), dynamical systems analysis (Gursky et al., 2011).

Results. Here we worked from the known structure and organization of the hb gene to investigate several aspects of hb expression dynamics which reduce developmental errors. Here, we extended the gene circuit approach (e.g. Jaeger et al., 2004; Manu et al., 2009ab) to a 'mid-grained' level by including separate regulation of hb promoter 1 (P1) and promoter 2 (P2) transcription and additionally allowing for the effects of TFs working together, as co-factors (see eqs 1). The model also has distinct transcription and translation steps. The relative simplicity of gene circuits makes them suitable for large automated parameter searches. By optimizing the model to Hb expression data (constrained by experimental data for regulatory TFs), we found parameter sets which produced robust hb expression despite high variability in maternal input factors (Bcd and maternal Hb (Hb_{mat})). In addition, the model has a Kr-dependence which correctly produces the loss of parasegment 4 (PS4) expression recently reported by Kozlov et al. (2012) in Kr- mutants.

Extending a standard gene circuit approach, we can write the following general differential equations for Hb protein and the P1 and P2 transcripts:

$$d[Hb]/dt = R_{ap}([P2mRNA]+[P1mRNA]) - \lambda_p[Hb] + Dif\,\Delta[Hb] \quad (1a)$$

$$d[P1mRNA]/dt = R_{aR}\,REF(C_c\,prodP1) - \lambda_R[P1mRNA] \quad (1b)$$

$$d[P2mRNA]/dt = R_{aR}\,REF(C_c\,prodP2) - \lambda_R[P2mRNA] \quad (1c).$$

The R_{aR} terms are for regulation, λ is for decay, and *Dif* is for protein diffusion. The sigmoidal function $REF(u) = u/\sqrt{u^2+1}$ normalizes regulation to a value between 0 and 1.

Parameters in the equations are optimized through EC to fit an experimental Hb pattern. Bcd, Hb, Gt, Kr and Kni were entered as potential regulators in the equations (through the production P1, *prodP1* term, and *prodP2* term); the EC process finds optimal TFs and parameters with respect to Hb expression. Experimentally-derived spatial patterns of the potential TFs were used as model input. In this 'mid-grained' extension of the gene circuit approach, co-action of multiple TFs on a target gene is allowed (including hb self-activation): the EC process can try different 3rd order (e.g. [Hb]*[Hb]*[Kr], etc.), 2nd order (e.g. [Hb]*[Hb], etc.) and 1st order (e.g. [Hb], etc.) TF combinations in the *prodP1* and *prodP2* terms.

An initial screen of runs with different combinations of TFs in *prodP2* & *prodP1* found some simple models showing key characteristics of hb expression. With EC, we found hundreds solutions of these equations (with different parameter sets) producing anterior-high Bcd- and Hb-dependent P2 expression and striped Hb- and Kr-dependent P1 expression. Combined translation of P2 and P1 creates a Hb protein profile with a distinct PS4 peak anterior of the mid-embryo boundary. In addition to the shape of the profile, the Hb output is relatively robust to variability in the maternal Bcd input gradient. Introducing experimental levels of between-embryo Bcd variability (by increments/decrements of the Bcd exponential parameters; same technique used in Manu et al. 2009a) causes little change in Hb output from about 0 to 35%EL.

Maternal Hb (Hb_{mat}) has previously been shown to influence zygotic hb expression (e.g. Simpson-Brose et al., 1994; Porcher et al., 2010; Porcher et al. 2010). Analysis of a gap gene circuit model has also indicated Hb_{mat} to be a major factor in pattern formation (Manu et al., 2009). We find that Hb output is more strongly influenced by variability in Hb_{mat} than by Bcd variability. Published data shows high variability of Hb PS4 amplitude (Perry et al., 2012, Fig. 3A); it would be of interest to investigate the correlation of these fluctuations with Hb_{mat} variability.

Conclusions. The results indicate that upstream variability can be reduced by downstream regulation acting differently through different transcripts of the target gene. In particular, regulation of the different transcripts can depend on different TFs with different levels of TF

co-action. The hb dynamics may serve as a model for how variability reduction occurs for other genes with multiple transcript types. In particular, we find that P1, P2 expression dynamics optimized to fit experimental data shows robustness to maternal variability; and that dual action of Kr as a TF for hb can generate PS4, increase precision of the hb mid-embryo boundary position, and reduce within-nucleus transcriptional variability.

The given CRMs which have evolved for hb provide a framework of binding sites (BSs) to regulate expression. In the development of individual embryos with these CRMs, though, there are numerous levels of dynamic regulation which can be involved in making expression patterns robust. In looking at developmental trajectories, we can use temporally resolved data and finer resolution spatial data (for mRNA and for protein) to extract some of these dynamics.

Simulation of the Kr- mutation by removing Kr causes the *prodP1* term to go to zero. The combined translation of the P1 and P2 transcripts produces a Hb profile lacking the PS4 peak. The model is consistent with data concerning the expression of P1 in the WT PS4 (Margolis et al., 1995); the Kr BSs in the 'stripe' enhancer (Perry et al., 2012); and the loss of PS4 in Kr- embryos (Kozlov et al., 2012).

Acknowledgements. The research is supported by RFBR grants 15-04-07800, 13-04-02137, and 15-04-06480.

1. V.V.Gursky, L.Panok, E.M.Myasnikova, et al., (2011) *BMC Systems Biology*, 5, 118.
2. D.M.Holloway, A.V.Spirov, (2015) *PLoS ONE*, .
3. Manu, S.Surkova, A.V.Spirov, et al. (2009) *PLoS Biol*, **7**: e1000049.
4. Manu, S.Surkova, A.V.Spirov, et al. (2009) *PLoS Comp Biol*, **5**: e1000303.
5. E.Myasnikova, K.Kozlov (2014) *Journal of Bioinformatics and Computational Biology*, **12**(2) DOI: 10.1142/S0219720014410029
6. K.Kozlov, S.Surkova, E.Myasnikova, et al., (2012) *PLoS Comp Biol*, **8**: e1002635.
7. E.A.Zagrijchuk, M.A.Sabirov, D.M.Holloway, A.V.Spirov (2014) *Journal of Bioinformatics and Computational Biology*, **12**(2): 1441009.