Chapter 28. Promoter architecture, TF expression patterns and extended regulatory domains.

Following the mid-blastoderm transition, a set of Dl-dependent morphogenetic functions are expressed in lateral stripes along the A/P embryonic axis. A medial stripe of *zen* along the dorsal midline is flanked by successive, lateral bands of *dpp*, *shortened gastrulation* (*sog*), *rhomboid* (*rho*, aka *veinlet*^{rho}), *fgf8* (fibroblast growth factor8), *heartless* (*htl*), *sna*, *twi* around the D/V (L/R) embryonic axis. To either side of the ventral mid-line, inwardly migrating cells express broad stripes of *twi*, *sna* and *heartless* (*htl*); while the mid-lateral and dorsal ectodermal domains are delineated by stripes of *rho*, *sog* and *dpp* [1 2 3 4 5]. Most of these genes encode TFs, except the growth factor ligands (*dpp* and *fgf8*), the *fgf8* receptor (*htl*), and the *rho* protease. Critically, these lateral stripes are colinear with the promoter affinities for Dl: the *twi*, *sna* and *fgh8* promotors having weak Dl affinity, while the *sog*, *dpp* and *zen* promotors show progressively stronger binding, reviewed in [6]. Thus, Dl activity drives downstream responses leading to invagination along the ventral midline, giving rise to the mesodermal and neurogenic lineages. Meanwhile, the lateral stripes of the *msh*, *ind* and *vnd* Hox TFs allocate tissue-specific fates along the long (A/P) embryonic axis [7 8]. The nuclear localisation of Dl spreads outwards from the ventral midline, generating the D/V (L/R) AMS and the sequential progression of internal tissue fate: mesoderm > glia > neurogenic ectoderm > lateral ectoderm > dorsal ectoderm > aminoserosa.

The medio-lateral stripe of the Rho protease regulates EGFr ligand activity [9 10]. Strikingly, the catalytic site of Rho is buried within a hydrophilic, membrane-spanning pocket [11 9], such that its protease activity remains anchored to the signal-receiving cell. Meanwhile, *WntD* and *single minded* (*sim*) are transcribed along a single cell wide medio-lateral stripe, that may be coincident with the medio-lateral boundaries of *twi* and *sna* expression [12]. In general, WntD activity decreases nuclear Dl localisation, while Sim acts as transcriptional activator, in combination with the E-box TFs: Mad, Twi and MyoD [13 14]. In contrast to the compact *WntD* transcript, the *sim* TU (20.37 kb) has three promoters, multiple exons and a single 3' UTR. This chromosomal structure should allow transcriptional regulation of the different Sim protein isoforms, with separate promoters, common protein-coding exons and extended intronic segments. In this context, promoter interactions with putative regulatory domains have been analysed using a synthetic Glass Multiple Repeat (*GMR*) promoter, with *TATA, Inr, MTE* and *DPE* sequence motifs fused to a bacterial *Gal-4* TF. The native *glass* gene is expressed in the larval Bolwig organ and adult eye, however, short (2-3 kb) genomic fragments inserted into a *GMR* vector drive *Gal4* expression in multiple tissues and developmental stages [15 16 17 18] (https://flweb.janelia.org/cgi-bin/flew.cgi). In the case of *sim*, two of the six tested *GMR-sim* fragments drive embryonic *UAS-GFP* patterns that monitor the ventral midline (Fig. 41), while three of the six strains are inviable (or lethal), when homozygous with the *UAS10X-GFP* reporter. In principle, the long perdurance of the Gal4 protein may drive *UAS-GFP* transcription throughout several cell divisions. High levels of GFP expression should not be cytotoxic, but overloading the protein synthesis machinery may disrupt the balance of endogenous functions. Alternatively, the genomic inserts may modify transcription of the *Mocs1* gene at the *GMR-sim* insertion site, or off-target transcriptional responses to *Gal4* at other loci. Whatever the mechanism, the intronic *sim* fragments alter the distribution of the endogenous Eve protein, without causing visible morphogenetic alterations, or embryonic lethality. Thus, while the expression of *GMR-sim* constructs may reflect some features of the inserted genomic fragment, novel "non-sense" expression patterns may also be generated (https://flweb.janelia.org/cgi-bin/flew.cgi).

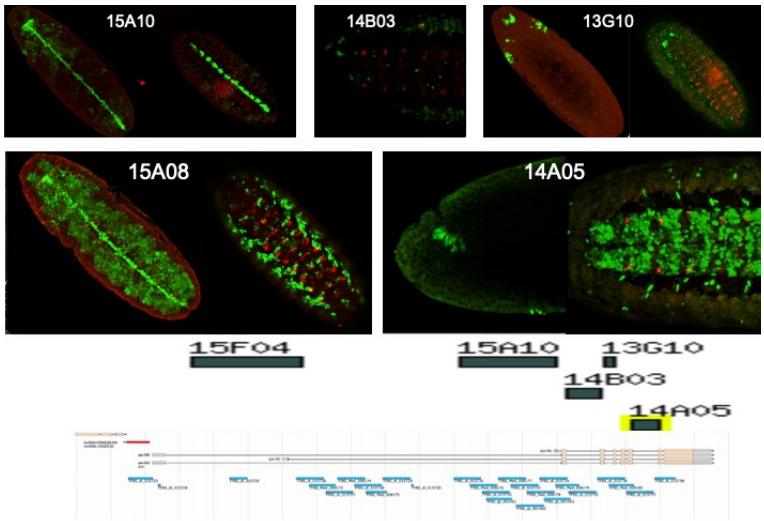

Fig. 41. Embryonic *GMR-sim* expression. Two of the six strains carrying intronic segments of *sim* monitor the ventral midline, like the endogenous TU. Each panel shows germ band extension to L and stage 16 to R (after germ band retraction, except 14B03, which shows stage 16 only. Eve antibody (red), GFP^{nls} (green). Black bars indicate the donor *GMR* fragments. From flweb.janelia.org. TFBS Dl, Med and Gt footprints (blue), from FlyBase, JBrowse view.

Similar novel synthetic patterns appear to be a general feature of the *GMR* strains, including *eve* (Fig. 42) and *dpp* (Fig. 43). In particular, the endogenous Eve protein distribution in *GMR-eve* strains rarely coincides with fluorescence of the *GFP* reporter. The *GMR*-driven alterations in endogenous Eve localisation appear not be associated with visible defects, or embryonic lethality.

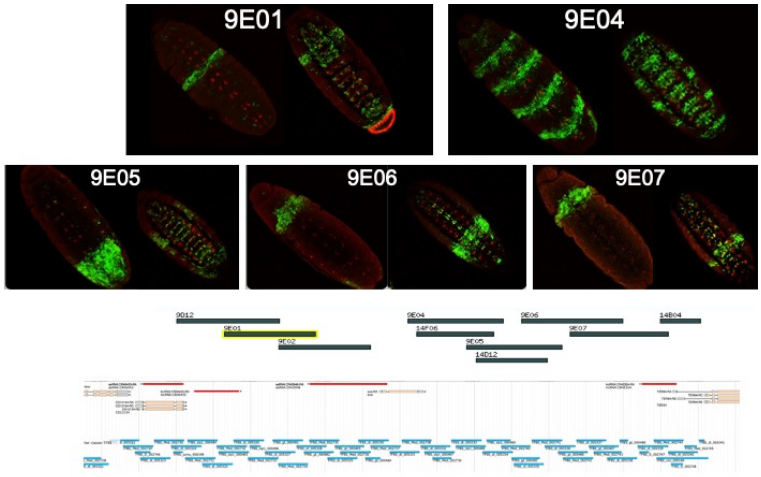

Fig. 42. Embryonic *GMR-eve* expression. *GMR-eve* strains, with intragenic regulatory segments 5' and 3' to *eve*, modify the distribution of the endogenous Eve protein (red), which rarely colocalises with the GFP^{nls} reporter driven by the *GMR* construct (green, nuclear localisation signal). Each panel shows germ band extension L and stage 16 R. Black bars indicate donor *GMR* fragments. From flweb.janelia.org. TFBS Dl, Med, Slp1 and Gt footprints (blue), from FlyBase, JBrowse view.

By contrast to *eve*, the *dpp* TU spans 35.5 kb, with four promoters, a single 3' UTR and extensive 3' regulatory domains, although all the splice variants encode the same protein isoform [19], https://flybase.org. In addition to novel embryonic patterns (Fig. 43), intronic *GMR-dpp* fragments generate aberrant imaginal disc expression, with associated morphogenetic defects [20]. In particular, the Pr/Dist ring and D/V loop of Wg expression are modified in *GMR-dpp* wing discs, with altered progression of the morphogenetic furrow in the eye-antennal disc (Fig. 44).

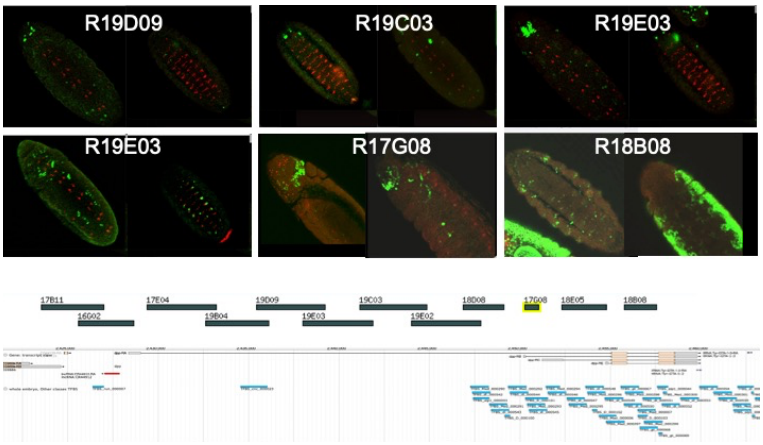

Fig. 43. Embryonic *GMR-dpp* expression with intronic fragments of *dpp.* **A.** Eve antibody (red), GFP^{nls} fluor. Each panel shows germ band extension L and stage 16 R. Black bars indicate donor *GMR* fragments. From flweb.janelia.org. TFBS Dl, Med, Gt, Slp1 and D footprints (blue), from FlyBase, JBrowse view.

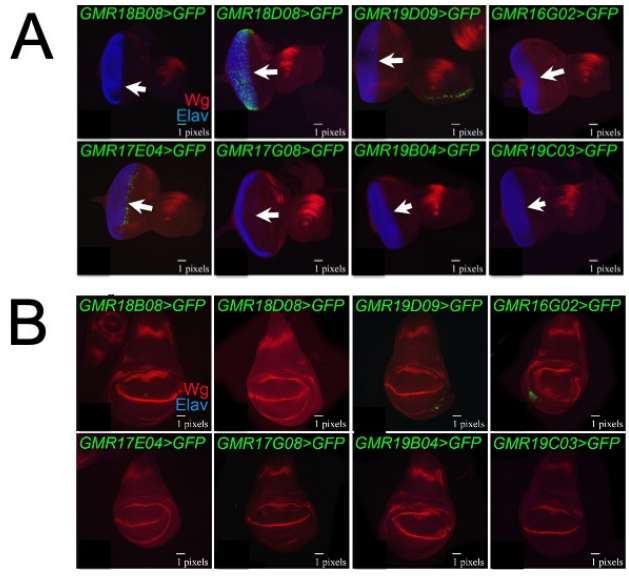

Fig 44. Imaginal disc *GMR-dpp* expression *GMR-dpp*; *UAS-GFP^{nls}* **A.** Eye-antennal disc. The eye twin-field boundaries may be distorted, with altered recruitment of ommatidia. White arrow indicates equator. ***GMR18B08*** shows a ventral shift of the equator. GFP^{nls} (green) is visible in scattered ommatidia particularly near the polar margins, but only in the V twin-field of the pupal retina (Sarkar et al 2018). ***GMR16G02*** shows progressively delayed P > A, and equator > pole, ommatidial recruitment, like *B* mutants. ***GMR19D09*** shows strong punctate GFP^{nls} fluorescence, which may correspond to nuclei in the peripodial membrane. From,

Sarkar et al. supp Fig. 1 **B.** Wing disc. Presumptive wing blade may be distorted, without the V ring of Wg in *GMR18B08*, *GMR19D09*, *GMR17G08*, *GMR19B04* or *GMR19C03*. The D/V marginal loop of Wg is absent in *GMR16G02*. The GFP^{nls} fluor is not detected, except for a few nuclei in *GMR19D09* and *GMR16G02*, which may be peripodial. From, Sarkar et al. supp. Fig. 2.

Taken together, these results confirm that regulatory functions may be smeared across extensive chromatin domains, including protein-coding, intronic and intragenic segments. Promotor architecture is critical for gene expression but is insufficient to provide stage- and tissue-specific transcription of morphogenetic functions.

Summary:

The morphogenetic functions expressed to either side of the D/V (L/R) midline determine tissue-specific fate and impose the ventral AMS. The promotors of these morphogenetic functions show graded affinities for Dl. However, promoter architecture is insufficient to confer stage- or tissue-specific expression. Instead, the genetic regulatory domains may extend across promoters, introns, exons and intergenic spacers. In general, early zygotic genes have short primary transcripts encoding single protein isoforms, that are either intronless, or contain a single, short intron. By contrast, later morphogenetic functions tend to be transcribed from extended TUs, with multiple promoters. Such extended TUs may encode different protein isoforms, or the same isoform with differential stage and tissue-specific regulation. The fusion of a synthetic promoter with short genomic fragments may drive transcription patterns with some features of the donor gene. However, such novel expression patterns are rarely stage- or tissue-specific.

References:

1. Irish, V & Gelbart, W. The decapentaplegic gene is required for dorsal-ventral patterning of the Drosophila embryo. *Genes Dev* **1**, 868–897 (1987).
2. Arora, K. & Nusslein-Volhard, C. Altered mitotic domains reveal fate map changes in Drosophila embryos mutant for zygotic dorsoventral patterning genes. *Development* **114**, 1003–1024 (1992).
3. Jiang, J. & Levine, M. Binding affinities and cooperative interactions with bHLH activators delimit threshold responses to the dorsal gradient morphogen. *Cell* **72**, 741–52 (1993).
4. Francois, V, Solloway, M, O´Neill, J, Emery, J, & Bier, E. Dorsal-ventral patterning of the Drosophila embryo depends on a putative negative growth factor encoded by the short gastrulation gene. *Genes Dev.* **8**, (1994).
5. Stathopoulos, A. & Levine, M. Dorsal Gradient Networks in the Drosophila Embryo. *Dev. Biol.* **246**, 57–67 (2002).
6. Gilbert, S. F. *Developmental Biology*. (Sinauer Associates, 2006).
7. Kosman, D. *et al.* Multiplex detection of RNA expression in Drosophila embryos. *Science* **305**, 846 (2004).
8. Mizutani, C. M. *et al.* Formation of the BMP activity gradient in the Drosophila embryo. *Dev. Cell* **8**, 915–924 (2005).
9. Urban, S, Lee, J R, & Freeman, M. A family of Rhomboid intramembrane proteases activates all Drosophila membrane-tethered EGF ligands. *Embo J* **21**, (2002).

10. Ligoxygakis, P., Roth, S. & Reichhart, J. M. A serpin regulates dorsal-ventral axis formation in the Drosophila embryo. *Curr. Biol.* **13**, 2097–2102 (2003).
11. Bier, E., Jan, L. Y. & Jan, Y. N. rhomboid, a gene required for dorsoventral axis establishment and peripheral nervous system development in Drosophila melanogaster. *Genes Dev.* **4**, 190–203 (1990).
12. Ganguly, A., Jiang, J. & Ip, Y. T. Drosophila WntD is a target and an inhibitor of the Dorsal/Twist/Snail network in the gastrulating embryo. *Development* **132**, 3419–3429 (2005).
13. Zelzer, E., Wappner, P. & Shilo, B. Z. The PAS domain confers target gene specificity of Drosophila bHLH/PAS proteins. *Genes Dev.* **11**, 2079–2089 (1997).
14. Jones, S. An overview of the basic helix-loop-helix proteins. *Genome Biol.* **5**, 226–226 (2004).
15. Pfeiffer, B. D. *et al.* Tools for neuroanatomy and neurogenetics in Drosophila. *Proc. Natl. Acad. Sci.* **105**, 9715 (2008).
16. Jenette, A *et al.* A GAL4-Driver Line Resource for Drosophila Neurobiology. *Cell Rep.* **2**, 991–1001 (2012).
17. Jory, A. *et al.* A Survey of 6,300 Genomic fragments for cis-regulatory activity in the imaginal discs of Drosophila melanogaster. *Cell Rep.* **2**, 1014–1024 (2012).
18. Manning, L. *et al.* A resource for manipulating gene expression and analyzing cis-regulatory modules in the Drosophila CNS. *Cell Rep.* **2**, 1002–1013 (2012).
19. Blackman, R. K., Sanicola, M., Raftery, L. A., Gillevet, T. & Gelbart, W. M. An extensive 3’ cis-regulatory region directs the imaginal disk expression of decapentaplegic, a member of the TGF- family in Drosophila. *Development* **111**, 657–666 (1991).
20. Sarkar A, Gogia N, Farley K, Payton L, & Singh A. Characterization of a morphogenetic furrow specific Gal4 driver in the developing Drosophila eye. **13**, (2018).