Patterning, Segmentation, and Hox genes in Drosophila and vertebrates

Notes and reading are for both Class 17 and 18
Reading:
8e: Chp 9: 278-287; Chp 14: 443-450
9e: Chp 6: 226-231;234-236; Chp 8: 312-316

Learning Goals Be able to:

- Describe how repeating spatial expression patterns of the *pair rule* genes are regulated and maintained.
- Cite evidence to demonstrate the existence of stripe specific promoter modules in the *pair-rule* genes.
- Interpret or predict phenotypes of mutations in particular patterning or segmentation genes.
- Describe what homeotic genes (Hox genes) are and explain their role in the process of A-P patterning in Drosophila.
- Design/interpret experiments that demonstrate the principles of Hox gene regulation and expression, including mutant phenotypes.
- Compare vertebrates and Drosophila with regard to how Hox genes are activated, and their roles in patterning.
- Explain the general mechanism of segmentation (somitogenesis) in vertebrates.
- Compare the mechanism and identity determination of segmentation in invertebrates and vertebrates

All animals have repeated structural units with corresponding parts (metameric organization), from the Pn cell lineages in *C. elegans* (the vulval precursor cells) to the vertebrae of vertebrates. In insects, such segmentation is apparent from the regular bumps and grooves in the external covering of the animal (the cuticle). In all animals, differences in these metameric units or segments along the A-P body axis are controlled by a class of genes called the Hox genes, which will be discussed in the next class. Prior to making the units different, insects must be able to form a series of complex repeating structures that are identical at equivalent points in the pattern. This organization has evolutionary implications: segments can duplicate and then later evolve special functions for different segments. When comparing Drosophila segmentation with vertebrate segmentation, the mechanisms at first appear quite different, but turn out to have features and genes in common that suggest homology.

In Drosophila, as discussed last week, the A-P and D-V axes are established by maternal transcripts, followed by specification of broad regional identities along the D-V and A-P axes and then division of the embryo into a repeating pattern of segments along the A-P axis. The segmentation process along the A-P axis is well described in the text; these notes provide a summary. In particular, you should think about how the patterning genes sequentially activate each other, and the mechanisms by which their expression patterns are controlled (remember combinatorial control), and finally, how the Hox genes give each of these segments their identity.

Segmentation along the A-P axis: segments and parasegments

Segments are determined at the blastoderm stage and appear as morphological units during germ band elongation a few hours later. At the time of segment determination, the embryo has about 6,000 cells. At blastoderm, each segment primordium is a circumferential strip about four cells wide. The larva has three thoracic segments (T1-T3), eight abdominal segments (A1-A8), a tail structure made of two additional segments (the telson) and a set of three partially fused head segments (the acron) that end up inside the anterior of the larva. Each external larval segment has a distinctive pattern of ventral denticle belts and dorsal hairs.

Studies of gene expression described below showed that physical morphological repeating units, the segments, are actually composed of the posterior of one parasegment and the anterior of the next parasegment. It is the parasegments that are patterned during early embryogenesis, and then converted into segments later. There are 14 parasegments established in the embryo.

As discussed in the last class period, the egg polarity gene products, acting as morphogens, repress translation of maternal *gap* gene mRNAs in specific regions and activate transcription of the several embryonic *gap* genes in different positions along the A-P axis, thereby specifying different identities for broad regions of the embryo. This all works by combinatorial control. In class we'll think some more about how changes in expression patterns of the egg polarity or of other gap genes will impact the expression of other gap genes, and thus impact the patterning of the entire embryo.

The *pair-rule* genes (all embryonically transcribed) encode transcription factors that control position within each segment primordium (parasegment). They are transcribed in the embryo in overlapping repeating patterns; i.e. each is expressed at the equivalent position in each segment primordium, which corresponds to 7 "stripes" within the embryo. Thus, the mutant phenotype of a loss of function mutation in one of the pair rule genes is the loss of alternating stripe. Their transcription is ALSO under elaborate combinatorial control, this time by the *egg-polarity* and *gap* gene transcription factors, in such a way that each *pair-rule* gene can be activated by different combinations of transcription factors at the appropriate points along the A-P axis. The promoters of these genes consist of several modules of enhancer and silencer response elements, each module specific for a parasegment. Activation of any one module is sufficient to cause expression. The mutant phenotype of the pair rule genes is that they are missing an equivalent portion of every other segment, different for each gene. Make sure you look at the figures in your book, and be prepared to think about how a mutation in any one of the egg polarity and gap genes will affect the expression of the pair rule genes.

The segment polarity genes

The *segment polarity* genes, which primarily encode cell signaling ligands, receptors, and pathway components, are activated after the blastoderm has cellularized, in specific positions of each parasegment, defined by the pair rule genes, which act by combinatorial control to activate the segment polarity genes. The mutant phenotype of a loss of function in a segment polarity gene is that part of each segment is deleted, replaced by a mirror-image duplication of another part (A,21-60). The segment polarity genes encode both cell-signaling molecules (*wg, hh, ptc)* and transcription factors (*en*, homeodomain class; *ciD*, Zn-finger class).

Summary of Drosophila segmentation:

During the segmentation process, two things are happening along the A-P axis.

1) Specification of broad regional identities. Overlapping patterns of *gap* gene expression and short range diffusion of the resulting proteins through the syncytial blastoderm define several regions of the body plan, which will later be refined by *Hox* gene expression.

2) Segmentation. *Gap* transcription factors act combinatorially on promoters of the *pair-rule* genes, activating transcription of these genes in periodic patterns that define seven pairs of parasegments. Until this point, the patterning process depends entirely on localized expression and diffusion of transcription factors, which then interact to turn on subsequent transcription factors in nuclei of the syncytial blastoderm.

Then, as cellularization occurs, local, periodic combinations of *pair-rule* transcription factors activate *segment-polarity* genes at specific locations within each segment, and interactions between these genes (e.g. *wg* and *en*) sharpen the segment boundaries, by mechanisms that involve cell signaling.

Note that the *pair-rule* and *segment-polarity* genes are expressed periodically, with the same pattern in each segment. They have nothing to do with determining segment identity, just with the process of segmentation per se. The *gap* genes, and later the *Hox* genes, must supply the information for different segment identities.

Hox genes

The *homeotic selector genes* or *Hox* genes of the *homoeotic complex (Hom-C)* are responsible for giving each segment a unique identity, expressed as different structures in the larva and the adult animal. They are perhaps the most interesting and universally important developmental genes so far discovered. Their expression and action are complex, and not yet completely understood.

Homeotic mutants

The Hox genes were first genetically defined in Drosophila by so-called *homeotic* mutations. These cause not deletions of pattern elements, but rather transformation of certain regions of the body into others. That is,

segments lose their normal identity and take on the identity of other segments. In general, recessive loss-of-function Hox gene mutations cause posterior-to-anterior transformations, and dominant gain-of-function mutations cause the opposite.

For example, *Antennapedia (Antp)* null mutants die as embryos or larvae in which T2 develops as T1, suggesting that a normal function of the *Antp* gene product is to determine T2. *Ant* is also expressed at low levels in T3, but as you will see below, so is the next Hox gene, *Ubx*. Dominant gain of function mutants of *Antp*, often resulting from chromosomal rearrangements, cause a variety of defects such as transformation of antenna to leg or eye to wing.

For example, the dominant mutation $Antp^{73b}$ is a small inversion that connects the *Antp* coding sequences to a promoter that causes expression in the eye/antennal disc, resulting in the classic homeotic transformation of antenna to leg.

Likewise, loss-of function mutations in the *Ultrabithorax Ubx* gene can cause transformation of T3 into a second T2, resulting in a four-winged fly. For almost every one of the 14 segments, homeotic mutations have been described that cause transformation into another segment. Therefore, in general, the Hox genes function to make each segment different from the adjacent anterior segment.

Genetic definition. The eight *Hox* genes defined by these mutations comprise the *homoeotic complex (HOM-C)*, which in Drosophila is split into two clusters or (sub)complexes on chromosome 3. The *Antennapedia* complex (ANTP-C) includes 5 genes: *labial (lab), proboscipedia (pb), Deformed (Dfd), Sex combs reduced (Scr),* and *Antp*. The bithorax complex (BX-C) includes only 3: *Ubx, abdominal A (abdA),* and *Abdominal B (AbdB)*. Curiously, the map order of these genes on the chromosome is the same as the order of segments in the body that they affect. However, note that there are not as many genes as there are segments, although mutations that specifically alter each segment have been found. The regulation of these genes (below) can explain this discrepancy.

Function and expression. Segment identity is specified by the combination of Hox genes expressed in each region of the body. To achieve normal development, the correct combination must be expressed at the appropriate level in each segment of the embryo. The ANTP-C genes control primarily the head and thoracic segments, while the BX-C genes control primarily the abdominal segments; however, both complexes are involved in specifying the posterior thoracic segment T3. ANTP-C genes are initially expressed at the cellular blastoderm stage, in essentially non-overlapping successive regions of a segment or two, from anterior to posterior, beginning with *lab* in the head. They are expressed along the A/P axis in the same order as their genetic map positions. For example, *Antp* is expressed in parasegments 4 and 5 (posterior T1, T2, anterior T3). Later *Scr* and *Antp* are expressed in more posterior segments as well during central nervous system (CNS) development. BX-C genes are expressed similarly, in broader overlapping regions, beginning in successively more posterior segments. Thus each segment expresses a unique combination of Hox proteins at different levels.

Regulation of expression. Gap genes provide the initial positional information for regional Hox gene expression. This is as expected, since gap gene expression is the only way to distinguish pairs of segments from each other. E.g. *Kr* protein (in the middle of the embryo) activates transcription of *Antp* and inhibits that of *AbdB*. Pair-rule proteins fine tune the expression pattern. Unique combinations of gap and pair-rule proteins establish the anterior boundary of expression of each Hox gene.

In addition, the Hox genes regulate each other's expression in a manner that exhibits *posterior dominance*: transcription of each Hox gene is negatively regulated by the proteins encoded by the more posteriorly expressed genes. This mechanism establishes the posterior boundary of expression of each Hox gene.

The patterns of gap and pair-rule gene expression, which establish the initial anterior boundaries of Hox gene expression during cellular blastoderm stage, are transient. Additional genes that act to maintain these boundaries throughout larval development were identified by mutations that cause opposite transformations, from anterior to posterior identities. For example, mutations defining a group of genes called the Polycomb group cause all segments to develop like A8. In these mutants, the early anterior boundary regulation breaks

down, so that *AbdB* is eventually expressed in all body segments. Polycomb proteins maintain the boundaries by controlling local changes in chromatin structure (chromatin remodeling) that are inherited through the cell divisions of larval development.

Targets of Hox gene regulation. Understanding how the Hox genes actually play their roles as master transcription factors that dictate segment identity will require identifying the target genes they regulate, sometimes called the *realisator* genes, and only some of these have so far been identified. Besides other transcription factors, they include genes involved in downstream patterning events (e.g. the TGFβ-family growth factor *decapentaplegic*), as well as genes that control morphogenesis (e.g. CAM genes).
Note that the anterior-posterior pattern of Hox gene expression coupled with the pattern of dorsal-ventral patterning factors (e.g. the Dpp and Dorsal transcription factors) provides a coordinate system of two-dimensional positional information that can specify any particular region of the ectoderm. This system appears to be used in dictating the position of salivary gland formation near the lateral midline in the second parasegment, and it is probably used to specify formation of other organ primordia in other segments as well.

Vertebrate Segmentation and Hox genes

Vertebrates are not segmented in quite the same way as Drosophila, where different units of the body have both a clearly segmented look, and also clear differenced (legs vs. wings or antennae). However, in some regions of the developing vertebrate embryo, there are clearly defined regions of cells with different identities that fit into the general idea of segments. These two regions are the developing somites and the rhombomeres of the hindbrain. The somites form as discrete packages of mesodermal tissue on either side of the neural tube along the A-P axis. The somites differentiate into several tissue types, but it is their formation of the vertebrae which allows us to compare them most easily to the segments of Drosophila. They also contribute cells to the back, body wall and limb muscles, as well as the dermis (underlying mesodermal skin layer) of the back. Somite formation (unlike Drosophila segmentation, which works by coordinated subdivision) is temporally sequential. Neurulation begins at the head (rostral) region of the neural plate and proceeds sequentially toward the tail (caudal) region, following the posterior regression of the node at the anterior end of the primitive streak. As the sides of the neural tube begin to fold up at each point, the adjacent paraxial mesoderm (presomitic mesoderm, PSM) becomes segmented into somites, discrete bundles of mesodermal cells organized into an epithelium. The number of somites differs among different vertebrate species (65 pairs in mice, 50 in chick, up to 500 in snake!).

Somites are formed in a sequential fashion from the presomitic mesoderm, PSM. The most P region of the PSM is composed of dividing stem cells, which continue to add mesenchymal cells as the more anterior cells are budding off into epithelialized somites. The most posterior region of the PSM is the Node, which continues to regress towards the posterior of the animal until gastrulation is complete.Cells exit the mitotic area after they are produced simply because cells are continuously added posteriorly as the primitive streak extends to the most P region of the embryo.

Production of somites by a molecular clock mechanism

Somite formation is precisely timed: every 90 minutes, a new pair of somites buds off from the PSM. This timing led scientists to suggest there might be a molecular clock, or rhythmic cycling of mRNA expression that could translate into the basic spatial pattern of the somites. Recent evidence in chicks supports this idea. Several genes have been identified that seem to regulate this timed formation of somites, but it is complicated, and not yet clear whether the clock creates the gene expression, or the gene expression creates the clock!

Because the somites form due to boundaries being established between each chunk of mesoderm, scientists have also studied this process. It appears that boundary formation may be dependent on the Notch signaling pathway: mutations in Notch or its ligands lead to defects in somite size and shape, but do not disrupt the actual formation of the somites. Notch (receptor) and Delta (ligand) are expressed in the P half of each new somite (see figure in ppt slides). These expression patterns are static. On the other hand, the expression of another molecule called Lunatic fringe, which is involved in Notch activity, is expressed for a time, then not expressed, then expressed again, etc. Notch activation occurs only along boundary between Lfng expression and no Lfng expression. This may limits the size of each somite (where a somite begins and ends).

Regional specification: Somite identity along the anterior-posterior axis

Somite structures (e.g., vertebrae--cervical, thoracic, lumbar, sacral) show variation along the A-P axis.

Specification is achieved early, already present in the PSM. If transplant PSM cells from the thoracic region into the cervical region, still get thoracic-looking ribs.

Within a somite, there are clearly separate anterior and posterior halves. In fact, vertebrae are derived from the posterior half of 1 somite and the anterior half of the next most caudal somite, just as in Drosophihla a segment is formed from the posterior half of one parasegment and the anterior half of the next parasegment.

Hox genes in vertebrates. In vertebrates, the *Hox* genes are arranged similarly to their respective homologs in *Drosophila*, but they have expanded from 8 genes to 13. Moreover, the *Hox* cluster of flies has duplicated twice, so that in mammals there are 4 separate clusters, none of them complete, on separate chromosomes. The clusters are called *Hoxa, b, c,* and *d,* and the genes in each are called *Hoxa-1, a-2, a-3,* etc (Box 4a). Thus, there are multiple homologs of each *Drosophila Hox* gene: these are called **paralogs**. In mice it has been possible to show by reverse genetics that mammalian *Hox* genes are also functionally homologous to their counterparts in flies. The clearest results come from knocking out all the members of a particular paralog group, e.g. *Hoxb-8, c-8*, and *d-8*. The result is a clear posterior-to anterior transformation of the first lumbar vertebra into a thoracic vertebra (4.13).

A-P patterning in mouse and human. These are the most difficult organisms to study, because they're very small, and normally develop within the uterus. And for humans, of course there are moral and legal barriers to experimental work on embryos. In the mouse, however, there is the possibility of knocking out homologs of *Drosophila* patterning genes, which has given additional information. Presumably, some early events are homologous to those in frog and chick embryos, but some of the signals come from extra-embryonic tissues and so may be different. Eventually, however, a gradient of molecules present along the A-P axis helps to specify which Hox genes will get activated in which place. One of the main molecules that leads to the activation of different Hox genes is Retinoic Acid. This molecule is in high concentration at the posterior end of vertebrates, and at lower concentrations towards the head. Anterior Hox genes are activated in culture by low levels of RA; posterior Hox genes require high levels of RA. Thus, although the mechanism for activating the Hox genes is different, the vertebrate embryo, just like Drosophila, ends up with different domains of Hox gene expression along the A-P axis that then specify cell fate. The identity of the somites is determined by their patterns of Hox gene expression, so just as in invertebrates, the identity of each segment is determined by the Hox genes

Review questions

1) What is a parasegment in Drosophila development? What is the relationship between parasegments and segments?
2) How are the pair-rule genes regulated to give "stripes" of expression? What is the logical structure of pair rule gene upstream regulatory sequences that makes this pattern of expression possible?
3) Describe the key experiment (will be discussed in class) that demonstrated the modular nature of the *even-skipped* promoter.
4) Interpret Hox loss of function and gain of function mutant phenotypes and be able to describe the regulation of the Hox genes.
5) What is the evidence for evolutionary conservation of the patterning mechanisms that have been discovered in *Drosophila*? What is the relationship between the Hox genes in *Drosophila* and those in vertebrates?
6) How is retinoic acid involved in A-P patterning in vertebrates?
7) What are somites? What region of the embryo do they arise in, and what embryonic tissues do they give rise to?
8) How do somites arise during somitogenesis, and how is this process controlled?