

Using genetics to study development

Reading: these notes (many topics in this lecture and the next two are not well covered in Gilbert) 8th ed: pp. 90-93, 96-98, 106-108, 114-115, 133-135, 243-250 (*C. elegans*).

9th ed: p 22-23; 41-42 (transgenic animals), p 32 (sidelights and speculations), p 61-65, 192-201 (*C. elegans*)

Learning goals: Be able to:

Explain the rationale behind using exploratory (“forward”) or manipulative (“reverse”) genetics to understand developmental processes, and identify situations in which one or the other can be used.

Design experiments using the exploratory genetics approach to identify and characterize the genes that control a particular developmental process of interest

Explain why and how you would use a particular genetic technique to answer a specific question

Developmental genetics: rationale and strategy of two genetic approaches

Genetic analysis provides powerful approaches to understanding how development works. Some of the central questions that have been or are being answered by developmental genetics are: How do cells in a cleavage stage embryo control their division orientations? What kinds of molecules are the cell-autonomous determinants of cell fate? How do these determinants become localized in cell divisions where they are partitioned asymmetrically to only one of the two daughter cells? What are the specific signaling pathways utilized when one cell or tissue induces a fate change in another? How is the development of regional structural differences dictated in patterning of the body plan during embryogenesis? How are particular cells and tissues recruited to form an organ? How is organ size controlled? Body size?

Exploratory genetics – from mutant phenotype to gene:

Exploratory genetics (sometimes also called “forward” genetic analysis) can be used to “dissect” functionally and define the important components of any biological process *without knowing anything about it in advance*. Why should you understand it? 1) Because many of the important transcription factors and signaling components we will discuss were discovered this way (which accounts for the funny names – they were originally gene names based on mutant phenotypes), and 2) because this same basic approach is used to identify the genes that are altered in human heritable diseases.

The analysis proceeds by isolating mutations that affect the process, using these to define the genes that control it, and then characterizing the products of these genes and their functions, thereby revealing the mechanism of the process. Below is a list of the general steps in this approach as applied to model organisms (you do not need to memorize them, or their order, but rather should be familiar with each one and the reason it is done):

- 1) Choose a defective phenotype that is specific and selectable or easily recognizable.
- 2) Treat animals with a mutagen to increase the mutation frequency.
- 3) Carry out a saturation screen: in a mutagenized population, screen for mutants with the defective phenotype, identifying enough mutants so that a mutation is likely to be found in each gene involved in the process.
- 4) Map the location of the genes
- 5) Identify the types of mutations found in this gene (loss-of-function, gain-of-function, hypomorphic, antimorphic, neomorphic), and determine the null phenotype (phenotype in complete absence of gene product).
- 6) If possible, use epistasis tests to establish whether the genes interact with each other in a developmental pathway.
- 7) Genetically map genes more precisely and isolate them by positional cloning.
- 8) Use each genomic sequence as a probe to isolate corresponding cDNAs from a suitable library, and sequence them to determine the predicted amino acid sequence of the encoded protein(s).

9) Carry out a similarity search, comparing this sequence with those in available computerized databases, which may provide information on a functional class to which the protein belongs.

10) Determine where and when the gene or protein is expressed during development using various techniques. Some of these might be

a) In situ hybridization: localizing where the message is within the embryo

b) Reporter constructs (described in more detail below). Reporters carry the promoter region of the cloned gene fused to a gene encoding a marker that can be seen in the microscope, usually either an enzyme whose activity can be detected by a histochemical stain on fixed preparations, such as the *E. coli* β -galactosidase gene LacZ, or a jellyfish fluorescent protein (GFP, YFP, RFP) which can be seen in live preparations.

c) Immunofluorescence—using antibodies made against the protein product of the cloned gene.

After all this, you still often want to test how genes function if it is not clear from their identification (i.e., similarity to other known genes). Do the protein products act as signaling molecules or transcription factors? One way to do this is using genetic mosaic analysis, which will be described in more detail later. Ultimately, to be certain of a gene's function, you need to isolate the protein products of cloned genes and analyze their functions biochemically.

This exploratory genetics approach is not fast, but is certain to yield valuable information. The positional cloning step, formerly the bottleneck, is becoming steadily easier with the development of molecular markers, physical genome maps, and complete genome sequences. *However, the complete approach, that is, the 10 steps above, can be used conveniently only in organisms that reproduce rapidly and are amenable to genetic analysis.* Two such organisms are the nematode *C. elegans* and the fruit fly *Drosophila*. This approach was initiated for several developmental processes in worms and flies in the late 1970's, and has continued since then to pay off richly with results that we will be considering in the next several lectures. It has also provided valuable information about vertebrate development through the discovery of vertebrate genes that are similar to those identified in worms and flies.

Manipulative genetics – from gene to mutant phenotype (also sometimes called “reverse” genetics):

Exploratory genetic analysis is impractical or at least relatively slow and resource-limited for vertebrates, which have much longer reproductive life cycles than worms or flies and are more expensive to maintain.

However, most of the genes discovered in worms and flies by these techniques have vertebrate homologs, which can be easily cloned by sequence similarity. The expression patterns of these vertebrate genes are often consistent with functions similar to those of their invertebrate homologs.

Manipulative genetics allows a direct test of how these genes function in vertebrates, by inactivating them and analyzing the resulting mutant phenotype. Often a vertebrate may have several homologs of an invertebrate gene. These techniques can establish which is the functional ortholog of the invertebrate gene and which are paralogs. The most commonly used methods are gene inactivation, like knockouts of mouse genes by targeted recombination (discussed in class 19), silencing of gene expression by RNA interference (RNAi), and overexpression or ectopic expression by generation of transgenic organisms.

Important methods in forward and reverse genetics

Positional cloning (in organisms with sequenced genomes) of mutationally identified genes:

Mapping (using genetic markers and DNA markers like SNPs) of a mutation identifies a small chromosomal region where gene must be located. Predicted genes in that region are candidates. Tests of whether a candidate gene is the mutant gene: 1) Can mutant phenotype be “rescued” by transformation of mutant animals with cloned genomic DNA including only the wild-type gene? 2) Does inactivation of the gene in a wild-type animal by RNAi or gene knockout give the same phenotype as the mutant? 3) Does sequencing of the gene from mutant animals identify a change from wild type that could be the mutation?

Mosaic analysis:

A mutant phenotype is generally due to lack of (or abnormal) gene function in one or a few types of cells, tissues or organs; in the rest of the animal the mutation may be irrelevant because the gene is not expressed. Thus, in mosaic analysis, the result of removing function of the gene of interest is studied only in the relevant group of cells. A mosaic animal (generated by different techniques in different animals—more information later) consist of partly mutant and partly wild-type cells, marked in some way that allows them to be distinguished by the investigator. One then looks at the phenotypes of cells that are physically next to each other but genotypically different to answer the question of how the gene product of interest functions. More on this later.

Targeted expression or mis-expression: generating transgenic organisms:

Once a gene has been cloned (its sequence is known), it can be introduced into other organisms, or reintroduced into the same organism, but made to be expressed in a different set of cells. In all organisms, the DNA construct (gene of interest plus a promoter that will determine where it is expressed) is injected into the oocyte (or just fertilized egg) at the one-cell stage. At a low frequency, DNA injected in this way can integrate prior to first cleavage at an apparently random chromosomal location, so that all cells of the resulting animal carry the integrated sequence. Sometimes multiple copies of the DNA line up and are integrated in a head-to-tail array (still just into one chromosomal location). Once a gene is integrated, it is generally expressed, although the levels can vary due to position effects of the integration site. If the construct includes a tissue specific promoter, the gene will be expressed only in cells in which that promoter can be activated. If the construct did not include a promoter, the integrated gene will be expressed dependent on whatever promoter is near its integration site.

Transgenic animals have many uses. One use is to visualize expression of a gene. In this case, the enhancer or promoter region of the gene is isolated and hooked to a lacZ or GFP-encoding sequence. When the construct is integrated, the regulatory region for the gene of interest now “drives” expression of the marker. Thus, one can see where the gene is normally activated. Another use is to take a normal copy of a gene of interest and inject it into an oocyte mutant for this gene. The resulting animals are then studied to determine whether the introduced DNA is sufficient to rescue the mutant phenotype. Lastly, a gene can be purposely expressed under control of a different promoter: this is called artificial regulation of expression. This can be used to induce “ectopic” expression and observe the phenotype (an excellent example of this is the example given in class 11 of the *Drosophila* gene *pax-6* homologue (*eyeless*) being expressed in leg cells, and converting the leg cells into eye cells).

Knock-downs—RNAi:

When double-stranded (ds) RNA corresponding to the sequence of a particular gene is introduced into most cells, cellular proteins process it into fragments that associate with other proteins, complex with the corresponding mRNA, and cause its degradation, thereby preventing function of the gene. The dsRNA fragments may be transported from one cell to another in the animal, so that the effect is systemic. The normal function of such small dsRNAs, and why cells should have this machinery, are still incompletely answered questions, currently a hot research topic. But whatever the answers, RNAi provides an extremely useful simple technique for specifically inactivating gene functions in worms, flies, and cultured mammalian cells. In zebrafish and *Xenopus*, a similar technique (morpholinos, very stable antisense oligomers) is used to prevent translation of specific messages.

Introduction to *C. elegans*:

The common free-living soil nematode *Caenorhabditis elegans* has become a valuable model animal for approaching basic questions in development. Its small size, small cell number, short life cycle, and transparency make it ideal for analysis by microscopy and genetics. Despite its simplicity, its mechanisms of development have much in common with those of higher animals. The contributions of *C. elegans* research were recognized

by award of the 2002 Nobel Prize in Physiology and Medicine to Brenner, Sulston, and Horvitz.

<http://www.nobel.se/medicine/laureates/2002/press.html>.

*Description of the cell lineage of *C. elegans**

Fate map. The entire cell lineage, which is the ultimate fate map, is known from embryo to adult, for both sexes. It provides an excellent starting point for studying development. The fates of cells in nematodes were classically thought to be "lineally programmed," i.e. dictated by cell ancestry ("European plan"), without input from neighboring cells. Although many cell fates indeed are lineally determined, we now know that there are many essential cell-signaling interactions as well, as described below.

Founder cells are six cells in the embryo which give rise to the major embryonic lineages, named AB, MS, E, C, D, and P4. Founder cell lineages *do not* correlate consistently with germ layers. Three of the five somatic founder cells (AB, MS, C) contribute to more than one of the three germ layers. Founder cell lineages also do not correlate consistently with tissues; three founder cells (E, D and P4) contribute to only one tissue type, while the other three contribute to multiple tissue types (G,8.43).

Fates of early embryonic cells (blastomeres). Some of the early cleavages are unequal (giving daughters of different size), and at the 4-cell stage, each of the blastomeres has a different fate – that is, will give rise to different kinds of cells and tissues.

How are the fates of early embryonic cells (blastomeres) determined?

This question has been addressed using techniques of experimental embryology as well as genetics--many of these ideas will already be familiar to you.

Components of the egg cytoplasm segregate to certain blastomeres (review)

We have already seen the most striking example--P granules, which segregate specifically to germ line cells (the P lineage) in early cleavages, as seen by antibody staining. These granules, probably analogs of similar germ-line-specific inclusions in many other embryos, including those of *Drosophila*, *Xenopus*, and mice, appear to be partitioned by an actin-based process in *C. elegans*. They contain proteins and RNAs, which include germ-line determinants. A few of their components are known to be regulatory proteins; the rest are not yet characterized. In the early *C. elegans* embryo they provide a useful assay for normal cytoplasmic segregation.

Embryo manipulations demonstrate both cell-autonomous and inductive determination mechanisms

Cell isolation and laser ablation experiments showed that the gut, germ-line, and body-wall muscle lineages are programmed autonomously by the mid-4-cell stage. Other such experiments indicated that inductive events are also required. If one of the cells of the 4-cell stage embryo (EMS) is ablated, the ABa descendants fail to form anterior pharynx. If P₂ is ablated, the ABp descendants behave like ABa descendants. And finally, experiments with isolated blastomeres showed that gut development from the E lineage requires brief contact between P₂ and EMS (the parent of E) during the first few minutes of the 4-cell stage. These details (cell names and what the signals are) are not important; however, the way in which genetics can be used to help dissect these signaling events, is!

So, in *C. elegans* as in most embryos, some cell fates are determined by ancestry (cytoplasmic determinants) and others by inductive interactions (cell signaling). The molecular nature of the determinants and signaling components involved have been discovered through study of mutants, using the genetic approach (more details in upcoming lectures!).

Review Questions

- 1) Outline the steps by which the molecular components in a developmental process can be identified and functionally characterized using a forward genetic approach.
- 2) What is meant by the null phenotype of a gene? What is the experimental (genetic) test for a null mutation?

- 3) What is meant by saying that an animal is genetically “mosaic” for a mutant gene? What can the phenotype of such an animal tell you about the function of the corresponding gene product?
- 4) How are transgenic animals made, and why would you want to generate a transgenic animal? What kinds of questions can be answered by studying transgenic animals?
- 5) If you had managed to clone and sequence a mutationally defined gene encoding an unknown product involved in cell determination, what is the simplest thing you could do next to try to find out something about its molecular function?
- 6) What is the difference between forward and reverse genetics?
- 7) If you have a cloned gene of potential interest and you wish to know where in the organism it is expressed, how would you find out?
- 8) If you have a cloned gene of potential interest and you wish to know whether its normal function is important for development, what experiment would you do?