Chapter 5. Genetic Interactions and Pathways

Genes do not act alone. We can begin to infer the functional connections between genes by examining how alleles of different genes interact within an organism. For example, consider the process of programmed cell death. If you had a set of genes whose phenotypes suggest involvement in the regulation of programmed cell death, how might you use genetics to determine how these genes work together?

In this chapter we will consider how genetic interactions can be used to build models of genetic pathways. So far we have discussed how the behavior of a single gene can be analyzed, either through its inheritance patterns (Chapter 3) or through assessing its particular genetic characteristics (Chapter 4). In this chapter, we will discuss how to infer the functional relationship between two genes by analyzing a cell or organism that contains mutations in more than one gene. When trying to determine whether two genes display a meaningful genetic interaction, the questions that should be asked when two mutations are combined within an organism include: Is this phenotype additive or does the double mutant display something not simply explained by the phenotypes of the single mutants? Does one phenotype mask the other? Does the multiply mutant animal more closely resemble a wild-type animal than either single mutant?

Genetic Interactions

Using genetic interactions to define functional relationships among genes

When do two genes display a genetic interaction? A genetic interaction occurs when two alleles affecting different genes combine within an organism to yield a phenotype not simply explained by adding together the phenotypes associated with each of the two alleles. A genetic interaction indicates a functional connection between two genes.

What do we mean by an animal having a phenotype that is not additive? Consider an example of a genetic interaction between two genes, A and B, found in a diploid worm. In this case, there are two mutant alleles, a1 and b1, respectively. Either allele alone does not cause a phenotype. However, when both alleles are present in the same animal in a homozygous fashion, all the worms die during embryonic development (Figure 2 [SGF-1246]). The genes A and B are considered to display a genetic interaction, because a simple additive effect between the alleles does not explain the phenotype of the double mutant animal. We define double mutant as an organism containing mutant alleles in the two genes of interest. This is not to be confused with an organism that is a heterozygote for two mutant alleles of the same gene.

It is important to note that genetic interactions are distinct from physical interactions, which involve a direct interaction of gene products (Figure 1 [SGF-1240]). Two proteins that bind to one another (even transiently) within a cell are considered to have a physical interaction. This type of interaction is often defined biochemically. This physical proximity of two gene products can suggest a functional connection, although further experiments need to be done to demonstrate the functional purpose of the physical interaction. On the other hand, a genetic interaction reflects a functional interaction between gene products, but these gene products may or may not physically interact.

Although genetic analysis by itself will not allow you to elucidate the detailed molecular mechanism by which gene products interact, the interpretations of genetic interactions can be a powerful tool for making inferences about the relationships between genes.

The example above illustrates only one of the possible phenotypic outcomes from a double mutant experiment. In this chapter we will discuss how to analyze the results of experiments examining genetic interactions between pairs of genes. How genes interact to give...
continuous phenotypes will be discussed in Chapter 7, while the analysis of multiple mutations in naturally varying populations will be covered in Chapter 16.

Once a genetic interaction is discovered, this information can be used to infer gene relationships. Our focus this chapter is on the logic used to dissect genetic interactions between pairwise combinations of genes. Specifically, we will cover how mutant animals can be used to understand the architecture of the pathway in which they act. To order genes within a pathway, we will describe methods for genes that act as switches in regulatory pathways and genes that interact in a metabolic pathway, as the logic used to order genes in these types of pathways differ. Included in this chapter are examples of how the analysis differs for recessive and dominant genes, and how using non-null alleles can affect the analysis.

**Double mutant analysis can elucidate genetic interactions**

When a genetic interaction occurs, the double mutant will show a phenotype that is not simply explained by the independent actions of the two alleles. In the example considered above, alleles $a_1$ and $b_1$ not having a phenotype individually but do have a phenotype that leads to dead animals when placed together in an organism. This phenotype is not simply explained because the phenotypic analysis of the single mutants in $a_1$ and $b_1$ does not suggest that the combination would have such a severe phenotype. This interaction of $a_1$ and $b_1$ is a **synthetic genetic interaction**, because the resulting double mutant phenotype displays a mutant phenotype not observed in either individual mutants. Other types of common genetic interactions are summarized in Table 1, and will either be discussed in more detail when relevant in this chapter, or in future chapters. In particular, **enhancement** (phenotype gets more severe) and **suppression** (phenotype gets less severe) will be discussed in Chapter 6, in the context of genetic screens.

For genetic interaction studies to be most informative, it is useful to have at least an inkling about the type of biological process in which the genes of interest are involved. Genetic interaction analysis has been particularly informative when ordering genetic pathways (see below). Such pathways may involve a series of genes involved in the production of a gene product or a metabolic product, or control a series of steps that create a particular state in the cell. In these types of examples, if the analysis of the mutant allele suggests that the gene is involved in a particular process, you can then gather a series of mutations in genes that may be involved in a similar or related process. This set of genes can then be used for double mutant analysis, to help determine the order of action among these genes. A broad, non-targeted set of mutations can also be analyzed for a genetic interaction without any assumptions; this type of approach will be discussed in Chapter 6 when the type of analyses described in this chapter is carried out at a genome-wide level in the form of a genetic screen.

Any type of measureable or observable phenotypes can be used to analyze the effects of genetic interactions between two genes. Observable does not necessarily refer to a phenotype that can be seen by the naked eye. Sometimes these phenotypes might be biochemical and involve the measurement of a metabolic product or enzyme activity. Sometimes these phenotypes might involve sophisticated microscopy, to examine the localization of a specific gene product within a cell and how its localization is perturbed in various mutants. Sometimes these phenotypes might be measured molecularly, as by direct examination of the expression levels of a particular gene. Sometimes these phenotypes manifest as a behavior that is elicited from an animal, such as sleep or feeding. As long as the phenotype can be robustly measured, it can be used for studying genetic interactions between two genes.

**The importance of understanding the alleles and strains used for genetic analysis**

For the proper analysis of double mutants, it is important to understand the type of mutant alleles that you have in hand. Before attempting genetic interaction studies between two
alleles, it is most useful to understand the nature of your allele (for example, is it a loss-of-function allele that retains some gene activity or is it a null?), as you need to have an idea of what kind of gene product is there in the mutant animals from your gene of interest. In fact, the more information you have about your mutant allele, the stronger the interpretation of your results from genetic interactions. Genetic interaction studies build from studies that assess individual gene function (discussed in Chapter 4).

To construct double mutant strains, it would be best if each of the mutant alleles were in the same strain background, sometimes known as an ecotype (see Box 5-1). This is because the purpose of these analyses is to see what happens when the two different alleles are in the same organism. If strain backgrounds differ, mutations in additional genes may contribute to differences in the phenotypes observed in these analyses. Dissecting the contribution of genetic heterogeneity would require more sophisticated methods, which will be discussed in later chapters (7 and 16).

As with any type of experiment, careful design of a genetic interaction experiment will likely yield a better result. In other words, the thoughtless combining of any two genes and looking for an interaction can be less than informative. For example, consider a situation where two genes have no functional relationship. One gene (A) is involved in fly wing development while the other (B) is involved in fly leg development. Animals carrying allele a1 do not have wings and animals carrying allele b1 do not have legs. When you construct the a1b1 double mutant, you find a fly that lacks both wings and legs (Figure 3 [SGF-1287]). This may not be a very informative answer, as from the single mutant analysis you had already deduced that each of these genes were important in the development of their respective tissues. This double mutant phenotype is consistent with the interpretation that these two genes likely act independently in their respective tissues to regulate their proper development. In this case, the a1b1 double mutant fly lacking both wings and legs would not be considered to be displaying a genetic interaction, as the wingless- legless- phenotype seen can be simply explained by the simple action of both alleles a1 and b1 without any interactions with each other.

**Synthetic genetic interactions suggest genes act together in a way that buffers the organism from their loss**

In Figure 2 [SGF-1246] above, we diagram a situation where the combination of the a1b1 alleles within the organism leads to dead worms. The appearance of dead worms can be referred to as a lethal phenotype, and the situation where the combination of the a1b1 alleles together is referred to as a synthetic lethal phenotype. Genes A and B show a synthetic interaction, because neither the a1 or b1 allele alone leads to any lethality; the lethal phenotype is only seen when both alleles are combined.

What does this type of synthetic genetic interaction tell you? In this case, worms only die when both alleles a1, b1 are present. This means that normally in the organism, genes A and B work together to keep worms alive. Somehow, the activity of gene A buffers the organism against any effect due to the loss of gene B, and vice versa. In other words, without both gene products, something goes awry.

A synthetic genetic interaction typically results because of **apparent functional redundancy**. One common reason for redundancy involves genetic pathways that act in parallel to elicit a similar outcome, where either pathway can functionally compensate for the other (see Figure 4 [501]). When null alleles are used for genetic analysis, synthetic phenotypes are typically due to redundant genetic pathways. In this case, two pathways converge to affect the function of a protein, M, which is required for life. Pathways 1 and 2 buffer each other, as losing the function of a protein in either pathway (Figure 4B-C) does not cause enough loss of protein M function so the organism will still live. It is only when the organism loses function in both pathways that M cannot function. Figure 4D illustrates a situation where a mutation in gene A and gene H lead to a lethal phenotype.
Duplicated genes can have synthetic interactions

A simple case of synthetic interactions can result from the duplication of one ancestral gene into two genes. These two duplicated genes will sometimes show a synthetic interaction because they retain redundant function that was originally performed by the ancestral gene in the ancient organism.

For example, a cell could contain two genes encoding similar protein kinases that can phosphorylate the same substrate. Phosphorylation by either kinase will lead to activation of the substrate so it can perform its essential function. The deletion of the gene that encodes either of these kinases will not cause a phenotype, because either genes D or E alone are sufficient for this function. However, when cells now lack both gene D and E, the substrate is no longer phosphorylated and a phenotype ensues. Sometimes gene D and gene E could act in parallel pathways (Figure 5A [503]). Alternatively, a synthetic interaction can also mean that gene D and E represent bifurcations within a pathway (Figure 5B [503B]). Distinguishing between these two possibilities would depend on the architecture of the pathways upstream of gene D and E, as assayed though epistasis analysis (described below). A real example of how two duplicated genes act in a synthetic fashion to affect the development of the nervous system is described in Box 5-2.

Synthetic genetic interactions within a pathway can occur among hypomorphic alleles

The synthetic phenotypes discussed above involve pathways that converge on the same target. However, when non-null alleles are used, synthetic phenotypes can be observed even when two genes act in a linear pathway. In the case of hypomorphic alleles, a mutation may partially cripple the function of a gene within a pathway so that the overall function of the pathway is only slightly compromised because the missing activity can be compensated for by the activity of other genes within the pathway. For this example, the combination of two hypomorphic alleles within a pathway can lead to a non-functional pathway because the activity of the pathway is now sufficiently compromised in that it no longer yields a wild-type phenotype. This is illustrated in Figure 6 [502], where a pathway essential for life is diagrammed. When a pathway is overspecified, the organism is buffered when either gene B or gene C suffers a mutation that disrupts (but does not eliminate) the function of their respective gene products. However, when both hypomorphic b and c alleles are combined, there is not enough output from the pathway and now the organism dies.

Ordering Genes in Pathways

Epistasis analysis is used to order genes within a pathway

Genetic interaction studies have been particularly useful when analyzing genes involved within a similar pathway. These types of studies are sometimes called epistasis analysis, where epistasis involves a mutation in one gene masking the effect of a mutation in a second gene (see Box 5-3 for a discussion on the use of the term epistasis). An epistatic interaction differs from the types of interactions discussed above, where an unexpected phenotype was observed when two mutations with the same phenotype are combined within an organism. In epistasis analysis, two mutations with different phenotypes are combined in the same organism. The resulting double mutant exhibits the phenotype associated with only one of the mutants. Note that if an intermediate phenotype is obtained in the double mutant, epistasis analysis cannot be simply carried out.

The logic behind the interpretation of epistasis analysis differs, depending on how the genes are related to each other. The two types of pathways we will consider are the substrate
dependent pathway (Figure 7A [SGF-1321]) and the switch regulation pathway (Figure 7B [SGF-1321]).

Substrate dependent pathways are sometimes referred to as assembly or metabolic pathways, and involve the formation of a product from a substrate through a series of obligate sequential steps. The glycolysis pathway and the TCA (tricarboxylic acid or Krebs) cycle are well known metabolic pathways that involve a series of enzymes responsible for producing intermediate metabolites along a pathway. Although the terms substrate and product may imply a chemical reaction, this type of pathway can refer to a set of genes involved in the biosynthesis of a chemical within a cell (such as a metabolite), a set of genes involved in the biosynthesis of a larger structure (such as a bacteriophage), or a set of genes involved in the generation of a particular type of differentiated cell from a precursor cell. In order for the genetic pathway to be considered a substrate dependent pathway, the genes that work in this pathway are involved in the sequential steps necessary for the substrate to be changed into the product.

A switch regulation pathway is sometimes referred to as a regulatory pathway or as a negative pathway. For a switch regulatory pathway, the genes involved are important for determining the particular state, or condition, of an event. Mutations in genes involved in these pathways will typically represent one of two states, or conditions, of this event. An ON or OFF switch is a simple example of two states; mutations in a switch regulation pathway would either represent the consequence of an ON state or an OFF state of a regulatory pathway. Because there are two opposite states, these genes in the switch regulation pathway negatively regulate each other, where the state of the upstream gene affects the state of the downstream gene. These types of pathways can be seen in developmentally important signal transduction pathways regulating cell fate, like the pathway used to regulate sex determination in C. elegans.

We choose to refer to these types of pathways as switch regulation pathways, as substrate dependent pathways also involve regulation. The term “negative pathway” has been used because negative regulation is often seen in switch regulation pathways; however, we avoid this term because switch regulation pathways sometimes are regulated through positive regulatory elements (and sometimes are entirely composed of positive regulation).

**Mutations in substrate dependent pathways lead to phenotypes where intermediates are accumulated**

How can you tell in which kind of pathway your genes are involved? This is critical to determine for the set of genes being analyzed, as the logic of how to interpret your double mutants will vary depending on the type of pathway involved.

Genes involved in substrate dependent pathways will have null alleles whose phenotypes suggest a progression of events. Sometimes this is easier to deduce, as for example, when you have a morphogenetic event; each mutant in this pathway will have completed a portion of the morphogenetic process (Figure 8 [SGF-1288]). For example, various genes control the production of different parts of the bacteriophage, which use previously assembled components to self-assemble from its parts to a mature phage. Gene 1 is used to produce the (capsid subunit), which is then added to the product of gene 2 to make the (capsid). A null mutation in gene 2 would allow for the formation of the (capsid subunit) but no (capsid) would be produced. More generally, a null mutation in a gene involved in producing the parts required for phage morphogenesis would lead to a block at the step where the missing part is needed. Mutations in the various genes in this process would lead to a series of incompletely assembled phages, arrested at intermediate stages in phage formation (Figure 9 [SGF-1289]). Thus, it is important to appreciate that an entire series of phenotypes will be observed, not just two that represent and "ON" or "OFF" state of the pathway as would be seen in a switch regulation pathway.

Sometimes the progression along a pathway is harder to immediately determine, such as when we are examining metabolic intermediates within a substrate-dependent pathway. In
this case, a mutation in this type of pathway will lead to the accumulation of an intermediate in the synthesis of the metabolite can only be detected if assays exist for the distinction of one intermediate from the other.

For example, the production of adrenaline from tyrosine involves a substrate dependent pathway (Figure 10 [SGF-1249]). Adrenaline is an important metabolite used as both a hormone and as a neurotransmitter. During the production of adrenaline, tyrosine is converted into intermediate metabolites by different enzymes that catalyze the chemical reactions that ultimate convert it into adrenaline. A null mutation in any of the genes that produce the enzymes would lead to a block in the pathway and a buildup of the intermediate metabolite, as it cannot then be converted to the next metabolite in the pathway. Specifically, a null mutation in the gene encoding dopa decarboxylase would lead to a buildup of the Dopa metabolite, which could not get converted into dopamine (Figure 11 [SGF-1295]). On the other hand, a null mutation in the gene encoding the hydroxylase would lead to a buildup of dopamine. Assays able to detect Dopa and dopamine would be necessary to know which intermediate metabolite was accumulating in the mutant organism. Animals lacking both dopa decarboxylase and the hydroxylase would have the same phenotype as those lacking only dopa decarboxylase.

**Mutations in switch regulation pathways lead to phenotypes with distinct and opposite phenotypes**

Unlike genes involved in a substrate-dependent pathway, genes involved in a switch regulation pathway have alleles that have distinct and opposite phenotypes. In this case, the opposite phenotypes seen in these alleles represent one of the two states the pathway can take (for example, either an ON state or an OFF state).

In these types of pathways, dominant mutations that flip the switch the opposite direction from the loss-of-function allele can be used for epistasis (see below). An example of a switch regulation pathway is sex determination in *C. elegans*. Sex determination is typically a genetically determined phenotype, as discussed in Chapter 3. Normally, sex determination in *C. elegans* depends on the ratio of X chromosomes to autosomes (non-sex chromosomes). Animals with one X chromosome (XO) and two sets of autosomes are male (1X:2A), while animals with two X chromosomes (XX) and two sets of autosomes make female bodies (2X:2A) (Figure 12A [511A]). (Although the 2X:2A animals are technically hermaphrodites, they are considered to have female bodies with a germ line that does make some sperm.) In this particular example, the two states being examined are either the production of a male or female body. The normal wild-type state is an intermediate state, where both males and females are produced according to the X:A ratio.

Two important genes for sex determination are *her-1* and *tra-1*. Animals carrying a null mutation in *her-1* will lead to a female body, even when they are 1X:2A (Figure 12B [511]). Conversely, animals carrying a null mutation in *tra-1* will lead to a male body, even when they are 2X:2A (Figure 12C [511]). This data suggests that *her-1* is thus necessary for male development and *tra-1* is necessary for female development. Thus, these null alleles lead to either the male state or the female state being produced in the incorrect context. As these two genes have alleles of distinct and opposite phenotypes, they are good candidates for genes that act together in a switch regulation pathway. How epistasis was used to dissect the relationship of *her-1* and *tra-1* is discussed in Box 5-4.

How is a switch regulatory pathway different from a substrate-dependent pathway? A switch regulatory pathway involves two states: for this example, the states are male and female. Other genes that are involved in the sex determination pathway also show either a male or female body. Two other genes in the sex determination pathway are the *tra-2* and *fem-2* genes. Animals with null mutations in *tra-2* also produce a male body like the *tra-1* mutant animals. Similarly, animals with null mutations in *fem-2* produce a female body like *her-1* mutants. Thus, there appears to be only two states for genes in the pathway, unlike the adrenaline biosynthesis
pathway discussed above (Figure 10 [SGF-1249]). With adrenaline biosynthesis, a null mutation in any one of the enzymes involved in adrenaline synthesis leads to an accumulation of a different metabolic intermediate. Specifically, for the four steps illustrated above, there are four distinct phenotypes. A mutation in the gene encoding tyrosine hydroxylase leads to an accumulation of tyrosine, a mutation in the gene encoding dopa decarboxylase an accumulation of Dopa, a mutation in the gene encoding hydroxylase leads to an accumulation of dopamine, and a mutation in the gene encoding transmethyllase leads to an accumulation of norepinephrine.

**Substrate-dependent pathways can be ordered through the examination of intermediate phenotypes**

The ordering of genes in a substrate-dependent pathway depends on the ability to observe the various intermediates within the pathways. Sometimes this analysis can involve directly assaying metabolic intermediates. In some cases, the substrate can be a cell or structure in an organism.

Consider a simple case, with two genes, C, and D, that both act in a cell called AR. Gene C is important for the survival of the AR cell such that when an animal lacks gene C function, these animals do not have cell AR (Table 2; Figure 13 [SGF-1294]). On the other hand, in a normal animal, the AR cell will differentiate into a neuron called neuroAR. The ability for AR to become a neuron is determined by gene D such that without gene D, an organism will have a non-differentiated AR cell that will never become neuroAR. If allele c₁ and allele d₁ are null alleles for genes C and D respectively, these can be used for epistasis analysis by combining the two alleles to create a doubly mutant animal carrying alleles c₁ and d₁.

For this example, the doubly mutant animals look like animals carrying allele c₁; that is, these animals are missing AR. In this case, allele c₁ is masking the ability of allele d₁ to produce the differentiated neuroAR cell; C is considered epistatic to D. This analysis suggests that gene C acts on cell AR to promote its survival before gene D acts to affect differentiation. In other words, the surviving cell is necessary so that gene D can act and cause the cell to differentiate into a neuroAR. This result is consistent with the phenotypic analysis of the single mutant suggesting a substrate dependent pathway. The substrate is the AR cell, which first must survive and then will differentiate into a neuron. From this example, you can see that for a substrate dependent pathway, the epistatic gene is the one acting first. This is because the intermediate product must be produced before the next gene can act.

If we extend this example to include gene E, a gene involved in the production of a protein, E, that allows neuroAR to function. Cells carrying the null allele e₁ do not make E, and although neuroAR will differentiate and take on the morphology of a neuron, it cannot function and fire an action potential. Double mutant analysis involving c₁, d₁, and e₁ yield the data in Table 2. From these data, C is epistatic to both D and E, and D is epistatic to E. These data are consistent with the pathway shown in Figure 4.

**The ordering of switch regulation pathways depends on the altered states**

How are switch regulation pathways analyzed? For these types of pathways, there are usually two states being analyzed. Consider a case of a set of genes important for the color of a particular cell, OM. Normally, the OM cells are grey. However, mutants exists that either have unpigmented (or white) cells, or cells with too much pigment (and thus are black). It is thought that there is a signal from a neighboring cell (NM) that leads to the normal grey pigmentation (Figure 14 [SGF-1293]). Here, the two states are white ("OFF") and black ("ON"), with the wild-type intermediate state being grey.

You want to undertake a genetic analysis of genes likely involved in the pigmentation of the OM cell. You have a collection of genes that you suspect act in this pathway, because null alleles in these genes have OM cells that are either white or black. Table 3 shows the
phenotypes of this set of genes. Construction of double mutants between alleles that have opposite phenotypes yields the results in Table 4.

For simplicity’s sake, let’s first consider only two genes from this set of data, $H$ and $J$ (Figure 15 [SGF-1291]). Gene $H$ is necessary for OM to be white, while gene $J$ is necessary for OM to be black; these are opposite phenotypes. The opposite phenotypes suggest that these genes antagonize each other, and thus the relationship between $H$ and $J$ involves negative regulation. How might you order these two genes? By examining the double mutant, you see that $h_j$ animals have white OM cells. This double mutant behaves like the single $j$ mutant, and thus in the double mutant animal, the $j$ allele is considered to mask the effect of the $h$ allele. Another way of saying this is that gene $J$ is epistatic to gene $H$. For a switch regulatory pathway, the downstream gene is the epistatic gene. Thus, gene $H$ acts upstream of gene $J$ and negatively regulates gene $J$.

If we now consider the all data with the double mutant combinations with the $h$ allele, you can see that for all combinations examined, the double mutant phenotype masks the phenotype of $H$. Thus, all these genes are epistatic to $H$. For a switch regulatory pathway, the epistatic gene is the downstream gene. These data suggests that $J$, $K$, $N$, and $P$ act downstream of $H$ (Figure 16A [500-7A]).

Now consider the situation with the $m$ double mutants, which is more complicated. Although $K$, $N$, and $P$ are epistatic to $M$, $M$ is epistatic to $J$. This suggests that $M$ acts upstream of $J$ and downstream of $K$, $N$, and $P$, to yield the following pathway (Figure 16 [500-7B]). If we combine the results from double mutant analyses using the $h$ and $m$ alleles together, you get the following pathway (Figure 16C [500-7C]).

These studies suggest that $H$ is the most upstream component of this pathway. Although $j$, $k$, $n$, and $p$ have the same phenotypes, this double mutant analysis shows that gene $J$ acts differently from genes $K$, $N$, and $P$.

## Finer aspects of pathway analysis

### Epistasis analysis should be carried out using null alleles

The use of null alleles is important for carrying out epistasis analysis. This is because the logic underlying these studies is based on the assumption that what is being measured involves what happens when the gene activity for the particular gene is absent.

First, imagine what might happen with a substrate dependent pathway. Recall the example from Figure 13 [SGF-1294], about how genes $C$ and $D$ regulate the development of the neuroAR cell. If instead of using null alleles $c$ and $d$, we conduct the analyses using weak loss-of-function hypomorphic alleles, $c_2$ and $d_2$ (Figure 17 [SGF-1296]). Animals with the $c_2$ allele have an AR cell surviving in 80% of animals. Animals with the $d_2$ allele will form neuroAR in 50% of animals. Thus, when a $c_2d_2$ double mutant is made, a normal functioning neuroAR is seen in 40% of animals, a non-differentiated AR cell is seen in 40% of animals, and 20% of animals have no AR or neuroAR cells. This complicated phenotype does not show clear epistasis: some of the $c_2d_2$ double mutant animals display a phenotype expected from loss of $C$ (no AR or neuroAR cells), some display a phenotype expected from loss of $D$ (a non-differentiated AR cell), and some display a normal wild-type phenotype! In the absence of clear epistasis, interpretation of gene order should not be attempted, as these results are uninformative and can even be misleading.

How might a hypomorphic allele behave in an epistasis experiment with a switch regulatory pathway? Using a hypomorphic allele can lead to incorrect conclusions. Consider the example of the OM cell, where gene $H$ negatively regulates gene $J$. When null alleles are used, we see that $h_j$ animals produce white OM cells, and $H$ acts upstream of $J$ (Figure 15 [SGF-1291]).
Let's consider what will happen if you have a hypomorphic allele of \( J (j_3) \), whose \( J \) gene product retains 80% of its function and produces a white OM cell. Suppose that the negative regulation on \( J \) is removed in the \( hj_3 \) double mutant. Now, the \( J \) gene product functions weakly, conferring a light grey non-wild type intermediate phenotype. In this situation, no clean epistasis is seen because the phenotype produced is not similar to what is seen with either single mutant allele, and no interpretation should be made. Alternatively, the removal of a negative regulatory function could increase the activity of the hypomorphic \( J \) product to such an extent that the \( hj_3 \) double mutant mimics the loss of \( H \), leading to a black OM cell. In this case, an incorrect interpretation would be made (Figure 18 [SGF-1297]), leading to the incorrect idea that \( H \) is epistatic to \( J \) and thus acts downstream.

**Reduction of function mutations can sometimes be informative**

Although the examples in the section above illustrate how the use of hypomorphic alleles can be misleading or non-informative, null alleles are not always available for the genes of interest. In some cases this can be because the gene has an essential function and null alleles are dead, making double mutant analysis difficult. In other cases this is because of the limitations of particular organisms and the difficulty in acquiring or engineering double mutants. In some organisms, the use of RNAi technology to knockdown gene activity can allow for the ready creation of hypomorphic (non-null) alleles. Thus, although not ideal, in this section we will discuss how hypomorphic alleles can be used in interaction studies.

The use of hypomorphic alleles in double mutant analysis is a one-sided experiment: one result is informative while the other is not. Let's first consider the informative case. You wish to order genes \( F \) and \( G \), both which act together in a switch regulatory pathway. You have a null allele in \( F \) (allele \( f_{null} \)) that causes black pigment to be formed, but only a hypomorphic allele in gene \( G \) (allele \( G_{hypo} \)) that causes white pigment to be formed. If the double mutant analysis shows that the \( g_{hypo} f_{null} \) double mutant has the phenotype of the \( G_{hypo} \), white pigment in this example, this is an informative result and you can conclude that \( G \) is epistatic to \( F \). In this case, even though the gene product produced by the \( G_{hypo} \) allele retains some function, it doesn't matter because you still see the phenotype conferred by the \( G_{hypo} \) allele (Figure 19 [SGF-1298]).

On the other hand, if the double mutant analysis shows that the \( g_{hypo} f_{null} \) double mutant has the phenotype of the \( f_{null} \) allele, black pigment, this would be uninformative because it is possible that the residual activity left in the \( G_{hypo} \) allele has enough function left that there enough is activity through the pathway to get to the next step. In other words, only if the epistatic gene is the hypomorphic allele, the result is informative.

**Using hyperorphic dominant alleles to order genes in a pathway**

Hypermorphic alleles can be useful for epistasis analysis, particularly when ordering genes whose loss of function phenotypes are the same and so cannot be ordered using their null alleles. Additionally, for switch regulatory pathways, hypermorphic alleles can help confirm the two states of the pathway, since hypermorphic mutations flip the switch the opposite direction from the hypomorphic loss-of-function allele (see Chapter 4 for such genes).

Epistasis can be carried out by combining hypermorphic and hypomorphic alleles. In the OM cell example above, we see that genes \( K, N \), and \( P \) all act downstream of gene \( H \) and upstream of gene \( M \) (Figure 16 [500-7c]). The relationships between \( K, N \), and \( P \) cannot be determined through analysis of null alleles. Animals with null alleles in these genes all have the same phenotype, so double mutant analysis between these genes will be uninformative for epistasis analysis. These three genes also show the same relationships with other genes with the pathway.

How might you order the functions of \( K, N \), and \( P \)? By using gain of function mutations in \( K, N \), and \( P \), you might be able to order these genes with respect to each other. Let's say that you have gain-of-function alleles \( k_{gf}, n_{gf} \), and \( p_{gf} \) that create black OM cells. Using these gain-of-
function alleles, you carry out more double mutant analyses and get the results shown in Table 5.

From these data, you find that \( P \) is epistatic to both \( N \) and \( K \), while \( K \) is epistatic to \( N \). These results show that \( N \) and \( K \) do not behave similarly. These newly uncovered relationships allows you to order these three genes, with \( N \) acting before \( K \) which acts before \( P \), as seen in Figure 20 [SGF-9A]. The arrows designating the relationship between these three genes are positive arrows because all three genes have the same null phenotype, and thus act on the pathway in a similar direction. By combining the information in Figure 16 [SGF-7C] with that in Figure 20 [SGF-9A], you can come up with the order of genes in the pathway seen in Figure 20 [SGF-9B].

It is difficult to be as sure that a mutations locks on a gene in its ON state as it is to be sure that a mutation eliminates gene activity. Thus any double mutants analysis using gain-of-function alleles should be considered carefully much as you might consider hypomorphs. If a gain-of-function allele is epistatic to a null allele, then the inference is solid. If a null is epistatic to a hypermorph, then the interpretation is weak. For example some gain-of-function alleles encode proteins that nonetheless depend on upstream activation, and thus the upstream null would be epistatic to the downstream hypermorph.

**Mosaic analysis can be combined with epistasis data to refine pathway inference**

The analysis of genetic interactions can also be combined with the results of other types of genetic experiments. For example, results from mosaic analyses (discussed in Chapter 4), which allow for the deduction of the site of action of the gene products, can be combined with genetic epistasis data, allowing for more interpretation and refinement the results from the genetic analysis.

Let's reconsider the pathway controlling the pigmentation of the OM cell, discussed above in Figure 15 [SGF-1296], Figure 16A-C [SGF-7], Figure 20A-B [SGF-9] and Tables 2, 3, and 4. From our double mutant analysis above, we have placed genes \( H \) as the most upstream acting genes. From the gain-of-function mutant analysis, we are also able to see that of genes \( N \), \( K \), and \( P \), gene \( N \) acts most upstream of this particular set of genes.

If you were able to further analyze these genes using mosaic analysis (discussed in Chapter 4), you could then determine the site-of-action of these genes. In other words, you could determine which genes acting in the OM cell, and which genes act in the NM signaling cell. When you carry out mosaic analysis, you get the results such that genes \( H \) and \( N \) act in the NM cell, while genes \( K \), \( P \), \( M \), and \( J \) act in the OM cell (also see Table 6). Since it is known that the NM cell sends a signal to OM, leading to its pigmentation, locating the site-of-action of genes \( H \) and \( N \) in the NM cell suggests that genes \( H \) and \( N \) are important in the signaling cell, while genes \( K \), \( P \), \( M \), and \( J \) act in the receiving cell OM. These studies also allow for the separation of \( N \)'s function from \( K \) and \( P \), as \( N \) acts in a different cell (Figure 21 [SGF-1251]).

**Genes that act together can be inferred from null mutations with incomplete penetrance**

In cases where many mutations confer the same phenotype but with incomplete penetrance (the proportion of organisms displaying the phenotype), we can associate genes into pathways by deviations from additive expectation.

Consider again a genetic interaction between two genes, in this case, \( R \) and \( S \), found in a diploid worm. Either of two complete loss of function alleles, \( r \) and \( s \), result in 30% survival of the embryos. When both alleles are present in the same animal in a homozygous fashion, only 10% of the worms survive embryonic development (Figure 22A [SGF-1219]). This is simply an additive phenotype – 30% of 30% is 9% survival (approximately 10%) – and thus we infer that \( R \) and \( S \) act independently. By contrast, we analyze loss-of-function alleles of genes \( T \) and \( U \), with alleles, \( t \) and \( u \), that also confer a 30% survival phenotype. If we find that the \( t \), \( U \) double mutants also display 30% survival (Figure 22B [SGF-1219]), we infer that \( T \) and \( U \) act together in
the same pathway. Loss of one gene eliminates the function of a pathway, and eliminating both components makes the phenotype no more severe.

**Larger scale pathway structure from analysis of incompletely penetrant null mutants**

We can extend this type of analysis with many mutations that confer the same phenotype but with different penetrance, and thereby associate genes into distinct pathways. This type of analysis is especially informative if we know something about the molecular function of the genes involved. Let’s say we have mutations in two ligands and two receptors. However, we do not know whether ligand-1 acts upstream of receptor-1 or receptor-2. We also do not know whether ligand-2 acts upstream of receptor-1 or receptor-2. Using double mutant analysis, we can determine this.

As shown in Figure 23 [SGF-1328], there are two phenotypes for viability: 30% penetrance and 70% penetrance. Mutation of ligand-1 or receptor-1 results in a 70% penetrant defect (30% viability). A ligand-1 receptor-1 double mutant has the same phenotype of 30% viability (Figure 23A [SGF-1328]). The expectation if they were independent would be that the double mutant would have about 10% viability (30% of 30% is 9%). This observation is consistent with ligand-1 working through receptor-1.

Similarly, mutation of ligand-2 or receptor-2 results in a 30% penetrant defect (70% viability). A ligand-2 receptor-2 double mutant has the same phenotype of 70% viability (Figure 23B [SGF-1328]). The expectation if they were independent would be that the double mutant would have about 50% viability (70% of 70% is 49%). This observation is consistent with ligand-2 working through receptor-2.

What happens if we combine mutations in the two pathways? If the two pathways are completely independent, we might expect that about ~20% viability (30% of 70% is 21%). However, in this example, we find that lethality is fully penetrant (0% viability) (Figure 23C [SGF-1328]), suggesting that the two pathways are synergistic. We would infer that there are two “parallel” pathways that converge to a common target (Figure 23D [SGF-1328]). The other double mutant combinations (ligand-1 receptor-1 and ligand-2 receptor-2) are consistent with the two pathways acting independently, with a synergistic effect if both pathways are eliminated.

If we had no knowledge of the likely order of action based on the molecular nature of the genes involved, we would just separate the genes into two functional groups (ligand-1, receptor-1) and (ligand-2, receptor-2) without ordering their action. This type of reasoning will be used in Chapter 12. Here, because we know that one molecule is a ligand and another is a receptor, we can infer that the ligand works to activate the receptor based on our knowledge of how these molecules work. How this type of analysis was used to associate the WNT ligands with their appropriate receptor is described in Box 5-5.

**Most pathways have both Substrate-Dependent and Switch Regulatory parts**

Although we have separated the discussion of analyzing substrate-dependent and switch regulatory pathways, most processes have both of these types of parts within a single pathway. How is this type of analysis used to inform our understanding of biological processes? For the study of apoptosis, genetic analysis was key to both identifying important molecules involved in the process as well as teasing apart the relationships of the genes involved. The pathway controlling the apoptotic process involves both switch regulatory and substrate dependent processes. Specific examples of how epistasis analyses were used to understand apoptosis are described in Box 5-6.
HIGHLIGHTS

- Genetic interactions are distinct from physical interactions of gene products. Epistasis implies that genes act in a common pathway.

- Pathways are of two general types.

- Substrate-dependent pathways have a series of intermediates.

- Mutations in Substrate-dependent pathways not only block the output of the pathway but accumulate intermediates.

- In a substrate-dependent pathway the epistatic gene acts earlier.

- A switch regulatory pathway has a series of binary states.

- In a switch regulatory pathway the epistatic gene acts later.

- More than additive interactions suggest parallel function.

- Simply additive interaction implies that genes act in independent pathways.

- Epistasis with non-null alleles can be informative in a limited sense.

- Gain of function alleles can help order gene function.

- Many pathways have both substrate-dependent and switch regulatory parts.

- Additional information can inform the model based on genetic interactions.

- Less than additive phenotypes imply genes act in a common pathway.
Reading

References:


Additional reading:


Systems Genetics Chapter 5: Figures and Tables

Physical Interaction

![Diagram of A and B physically interacting](image)

Genetic Interaction

![Diagram of C and D with an outcome](image)

Figure 1 [SGF-1240]. Physical interaction (A•B or A binds B). Genetic interaction (Genes work together somehow to do something.) whether or not they physically interact.

![Diagram of four different genotypes: a⁺b⁺, a⁻b⁺, a⁺b⁻, a⁻b⁻](image)

100% alive 100% alive 100% alive 0% alive

Figure 2 [SGF-1246]. Synthetic lethality in worms
Table 1. Major types of genetic interactions

<table>
<thead>
<tr>
<th>genetic interaction</th>
<th>definition</th>
<th>gene relationships ((a_1) and (b_1) are mutant alleles; wt = wild type)</th>
<th>Expected result (E) if no interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>synthetic</td>
<td>Double mutant results in a phenotype not observed in either single mutant</td>
<td>(wt = a_1 = b_1 \neq a_1 \ b_1)</td>
<td>E = wt</td>
</tr>
<tr>
<td>enhancement</td>
<td>Double mutant results in a phenotype enhanced to a severity beyond expected</td>
<td>([a_1 \ b_1 &gt; E \geq wt \ AND \ a_1 \geq wt \ AND \ b_1 \geq wt, \text{ but } a_1 \ \text{and } b_1 \ \text{cannot both } = \text{wt}] OR ([a_1 \ b_1 &lt; E \leq wt \ AND \ a_1 \leq wt \ AND \ b_1 \leq wt \ \text{but } a_1 \ \text{and } b_1 \ \text{cannot both } = \text{wt}]</td>
<td>E = (a_1 + b_1)</td>
</tr>
<tr>
<td>suppression</td>
<td>Double mutant results in a phenotype less severe that would be expected</td>
<td>([E &lt; a_1 \ b_1 &lt; wt \ AND \ a_1 \leq wt \ AND \ b_1 &lt; wt]] OR ([wt &lt; a_1 \ b_1 &lt; E \ AND \ a_1 \geq wt \ AND \ b_1 &gt; wt]]</td>
<td>E = (a_1 + b_1)</td>
</tr>
<tr>
<td>epistasis</td>
<td>Double mutant results in a phenotype that is equal to only one of the single mutant phenotypes</td>
<td>(a_1 \neq b_1 \ AND \ a_1 b_1 = a_1 \ OR \ a_1 b_1 = b_1)</td>
<td>E = (a_1 + b_1)</td>
</tr>
</tbody>
</table>

Figure 3 [SGF-1287]. Combining two genes without a functional relationship results in simply additive phenotypes.
Figure 4 [SGF-501]. Pathway 1 and Pathway 2 converge on protein M, which is essential. Panel A depicts the wild type, or normal situation.

Figure 5 [SGF-503]. Synthetic genetic interactions can come from bifurcations in pathways. Panel A shows two parallel pathways converging on the substrate encoded by F. Panel B shows a bifurcated pathway where genes D and E are redundant for their action on the substrate.
Figure 6 [SGF-1324]. Hypomorphic alleles can behave as synthetic alleles within a pathway.
A. Single or double hypomorphic mutations in a pathway affect pathway output. If pathway output is low enough (below the threshold shown in B) the organism is inviable.

Figure 7 [SGF-1321]. Substrate dependent (A) and switch regulation pathways (B). In panel A, arrows denote the chemical transformation from one substrate to the next, as catalyzed by the product of the gene above the arrow. In panel B, arrows denote positive regulatory interactions while lines ending in bars represent negative regulation.
Figure 8 [SGF-1288]. Phage Morphogenesis; a substrate-dependent pathway. Simplified version of real genes and real pathway.

Figure 9 [SGF-1289]. Epistasis in a phage morphogenesis pathway.
Figure 10 [SGF-1249]. Adrenaline biosynthesis: a substrate-dependent pathway

Figure 11 [SGF-1295]. Epistasis in adrenaline biosynthesis.
Animals with mutations in the sex determination pathway have phenotypes with opposite states. Mutants in the sex determination pathway have phenotypes such that all animals have female bodies, or all animals have male bodies.

**Figure 12 [511].** Animals with mutations in the sex determination pathway have phenotypes with opposite states. Mutants in the sex determination pathway have phenotypes such that all animals have female bodies, or all animals have male bodies.
### Table 2. Data for a substrate dependent pathway

<table>
<thead>
<tr>
<th>genotype</th>
<th>does AR survive?</th>
<th>does neuroAR form?</th>
<th>can neuroAR function?</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_+d_+e_+$</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>$c_1d_+e_+$</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>$c_+d_1e_+$</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>$c_+d_1e_1$</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>$c_1d_1e_+$</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>$c_1d_1e_1$</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

Figure 13 [SGF-1294]. The substrate dependent pathway controlling the formation of neuroAR.
Figure 14 [SGF-1293].

Table 3: Null allele phenotypes for a switch regulation pathway

<table>
<thead>
<tr>
<th>genotype</th>
<th>color of OM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h, j, k, m, n, p$ (wild type)</td>
<td>grey</td>
</tr>
<tr>
<td>$h_1$</td>
<td>black</td>
</tr>
<tr>
<td>$j_1$</td>
<td>white</td>
</tr>
<tr>
<td>$k_1$</td>
<td>white</td>
</tr>
<tr>
<td>$m_1$</td>
<td>black</td>
</tr>
<tr>
<td>$n_1$</td>
<td>white</td>
</tr>
<tr>
<td>$p_1$</td>
<td>white</td>
</tr>
</tbody>
</table>

Table 4: Double mutant analysis data for switch regulation pathway

<table>
<thead>
<tr>
<th>genotype</th>
<th>color of OM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h, j, k, m, n, p$ (wild type)</td>
<td>grey</td>
</tr>
<tr>
<td>$h_1 j_1$</td>
<td>white</td>
</tr>
<tr>
<td>$h_1 k_1$</td>
<td>white</td>
</tr>
<tr>
<td>$h_1 n_1$</td>
<td>white</td>
</tr>
<tr>
<td>$h_1 p_1$</td>
<td>white</td>
</tr>
<tr>
<td>$j_1 m_1$</td>
<td>white</td>
</tr>
<tr>
<td>$k_1 m_1$</td>
<td>black</td>
</tr>
<tr>
<td>$m_1 n_1$</td>
<td>black</td>
</tr>
<tr>
<td>$m_1 p_1$</td>
<td>black</td>
</tr>
</tbody>
</table>

Figure 15 [SGF-1291[[old504]. Gene J is epistatic to gene H. 
Figure 16 [500-7]. The relationships for the switch regulation pathway.

Figure 17 [SGF-1296]. Confusing epistasis with non-null alleles in substrate-dependent pathway.
**Figure 18** [SGF-1297]. Confusing epistasis results with non-null alleles in a switch-regulatory pathway.

<table>
<thead>
<tr>
<th>genotype</th>
<th>phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_+ g_+$</td>
<td>white pigment</td>
</tr>
<tr>
<td>$f_{null} g_+$</td>
<td>black pigment</td>
</tr>
<tr>
<td>$f_+ g_{hypo}$</td>
<td>white pigment</td>
</tr>
<tr>
<td>$f_{null} g_{hypo}$</td>
<td>white pigment</td>
</tr>
</tbody>
</table>

**Figure 19.** [SGF-1298]. If a hypomorph is epistatic to a null, an inference can be made.
Table 5. Analysis of switch regulation pathway using gain-of-function mutants

<table>
<thead>
<tr>
<th>genotype</th>
<th>color of OM cell</th>
</tr>
</thead>
<tbody>
<tr>
<td>k&lt;sub&gt;gf&lt;/sub&gt;n&lt;sub&gt;1&lt;/sub&gt;</td>
<td>black</td>
</tr>
<tr>
<td>k&lt;sub&gt;gf&lt;/sub&gt;p&lt;sub&gt;1&lt;/sub&gt;</td>
<td>white</td>
</tr>
<tr>
<td>n&lt;sub&gt;gf&lt;/sub&gt;k&lt;sub&gt;1&lt;/sub&gt;</td>
<td>white</td>
</tr>
<tr>
<td>n&lt;sub&gt;gf&lt;/sub&gt;p&lt;sub&gt;1&lt;/sub&gt;</td>
<td>white</td>
</tr>
<tr>
<td>p&lt;sub&gt;gf&lt;/sub&gt;k&lt;sub&gt;1&lt;/sub&gt;</td>
<td>black</td>
</tr>
<tr>
<td>p&lt;sub&gt;gf&lt;/sub&gt;n&lt;sub&gt;1&lt;/sub&gt;</td>
<td>black</td>
</tr>
</tbody>
</table>

A.  
\[\text{gene } N \rightarrow \text{gene } K \rightarrow \text{gene } P\]

B.  
\[\text{gene } H \rightarrow \text{gene } N \rightarrow \text{gene } K \rightarrow \text{gene } P \rightarrow \text{gene } M \rightarrow \text{gene } J\]

Figure 20 [500-9].

Table 6. Results from mosaic analysis of genes important for OM cell pigmentation

<table>
<thead>
<tr>
<th>gene</th>
<th>site of action</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>NM cell</td>
</tr>
<tr>
<td>J</td>
<td>OM cell</td>
</tr>
<tr>
<td>K</td>
<td>OM cell</td>
</tr>
<tr>
<td>M</td>
<td>OM cell</td>
</tr>
<tr>
<td>N</td>
<td>NM cell</td>
</tr>
<tr>
<td>P</td>
<td>OM cell</td>
</tr>
</tbody>
</table>

Figure 21 [SGF-1251]. Site of action plus epistasis refines pathway
Figure 22 [SGF-1219]. Additivity and less than additivity with incompletely penetrant loss-of-function alleles. A. The $r_f$ and $s_f$ alleles display additive phenotypes. B. The $t_f$ and $u_f$ alleles display less than additive phenotype.
Figure 23 [SGF-1328]. Pathway inference by scoring quantitative phenotypes. A. One pathway contributes 30% and the lack of additivity of mutations define the pathway. B. A second pathway contributes 70% and the lack of additivity of mutations define the pathway. C. The additivity from combining one mutation in each pathway establishes the independence. D. The inferred model for the overall pathway architecture.
Chapter 5: Boxes

Box 5-1 What is genetic background and why is it important?

Much like a photograph in which the subjects are present in the foreground and are viewed in the context of a background, a genetic experiment also has a foreground and a background. In a genetic experiment, the genes explicitly under study represent the genetic foreground. The rest of the genome, which is not being explicitly interrogated, provides the genetic background.

When carrying out an experiment, you can best understand the contribution of a particular variable by keeping all else constant and examining what happens when you only vary a single or a few parameters. This same principle holds true when undertaking genetic analysis. Analyzing the effect of two mutant alleles is easiest when all the other alleles contributing to the phenotype are the same. Since one cannot know a priori the entire set of genes whose variation can effect the gene of interest, scientists attempt to make all regions of the genome whose allelic variation is not under study as uniform as possible. Various model systems have tried to address this problem by using methods that allow for as constant a genetic background to be utilized as possible when conducting genetic analyses.

Isogenic strains are ideal for genetic analyses

An isogenic strain is a strain that is essentially genetically identical, and it the ideal background for genetic analysis. Most model organism communities have tried to generate an isogenic strain that can be considered “wild type.” An isogenic strain is one which has been extensively inbred such that most of the heterozygosity at differing loci is minimized. As most naturally occurring diploid organisms are heterozygous at many genetic loci, isogenic strains do not usually exist naturally and are usually created within a lab for research use.

How do you get an isogenic strain? Isogenic strains are created by inbreeding. During inbreeding, a brother-sister pair is crossed to each other and this procedure is repeated for multiple generations. Since by definition a brother-sister pair represents limited genetic diversity, the choice of a single brother-sister pair each generation over many generations results in the elimination of much of the genetic diversity initially apparent in the original pair.

Some model organism communities will decide to designate a single strain as the “wild type” strain for genetic analysis. An example a model organism community using this method is the C. elegans research community, where scientists decided to try to minimize the problem of different research groups working on different strain backgrounds by designating a wild-type worm strain called N2. Because C. elegans can be stored as a frozen stock and thus frozen and dispersed to researchers, laboratories can easily adopt this wild-type strain as the background for genetic analyses. Nonetheless, strains will diverge in time due to spontaneous mutation (see Chapter 13).

Different strain backgrounds may be used within a model organism community

For some model organism communities, there is more than one “wild type” strain. For example, researchers in S. cerevisiae commonly use one of several distinct strain backgrounds for their experiments. A strain refers to a particular inbred laboratory stock; these strains are all considered the same species of yeast. Each yeast strain has distinct properties that has led to its use for particular types of research. The strain S1278B can form pseudohyphae while the originally sequenced S288c strain cannot. The strain SK1 sporulates efficiently and synchronously unlike many other yeast strain backgrounds. These designated strain backgrounds are relatively stable, as a wild type version of the yeast strain can be stored in a frozen form and accessed periodically to minimize genetic changes that can occur during strain propagation.
When there is more than one strain background, it is best to confine the analysis to a single strain background to minimize the effect of potential genetic modifiers. In this case, it is important to know the strain background for the particular genetic experiment. For example, the particular yeast strain used is typically reported in research publications. Since different labs may utilize differing yeast strains, there are occasions where a genetic interaction seen in a particular strain background that cannot be recapitulated in a different strain background. These differences are likely due to silent genetic modifiers, as these various yeast strains do differ in their genomes (Liti et al., 2009).

Other model organisms that use the concept of strain background include Arabidopsis, Drosophila, and mice, which use various ecotypes. The ecotypes can be naturally occurring variants, as well as ecotypes that have been especially bred for the lab to minimize genetic variation. In plants, the backgrounds can be quite stable as strains can be stored as seeds which can last for decades.

For organisms where freezing a stock of a background or ecotype is not possible (i.e., Drosophila), genetic drift of the originally isolated genotype does occur over time as the organism is propagated for thousands of generations within the lab. Thus, while uniformity of genetic background is the ideal, in practice, scientists simply attempts to make the background as uniform as possible and are sensitive to the issues that may arise from a slightly non-uniform background.

A striking example of genetic background effects is seen in mice lacking the epidermal growth factor receptor (Egfr). Scientists created targeted null alleles of the EGFR in three distinct strain backgrounds, and found that EGFR is essential for life, although the stage of life that Egfr-null mice would die at depended on the strain background. In a CF-1 background, EGFR deficiency results in death around the time of implantation due to degeneration of the inner cell mass. In a 129/Sv background, EGFR deficiency results in death at mid-gestation due to placental defects. In a CD-1 background, EGFR mutants live as long as three weeks but have abnormalities in skin, kidney, brain, liver, and gastrointestinal tract, among other tissues.

References

**Box 5-2 Two receptor tyrosine phosphatases show synthetic interactions during Drosophila nervous system development**

During the development of the nervous system, axons must correctly reach their targets. Axon guidance depends on the proper activity of signaling molecules, including kinases, which phosphorylate their substrates, and phosphatases, whose actions oppose kinases by dephosphorylating their targets. A particular class of phosphatases called receptor tyrosine phosphatases (RPTPs), found in humans, worms, and flies, are important for proper development of the nervous system.

How do we know this? Originally, it wasn't so clear that these RPTPs played an important role in axon guidance. The *Drosophila* genome contains several genes encoding RPTPs, including *Ptp10D* and *Ptp4E*. Intriguingly, both *Ptp10D* and *Ptp4E* appear to be expressed in the developing *Drosophila* nervous system. However, animals with null mutations in either *Ptp10D* or *Ptp4E* were viable, fertile, and did not display nervous system defects. Did this lack of a phenotype indicate that these genes were not important for the developing nervous system? Alternatively, might these genes act in a synthetic fashion?

To test the hypothesis that these genes act in a synthetic fashion, animals carrying both *Ptp4E¹* and *Ptp10D¹* alleles were constructed. Analysis of the *Ptp4E¹ Ptp10D¹* double mutant demonstrated a phenotype in the developing central nervous system (CNS). By staining the axons of the CNS, it became clear that the longitudinal axons of the CNS were wavy and discontinuous compared to wild type and the single mutant animals (Fig 510A). This phenotype suggested that *Ptp4E* and *Ptp10D* were redundant and function together to ensure proper nervous system development.

The idea that *Ptp4E* and *Ptp10D* are redundant makes sense molecularly. Both of these genes encode RPTP proteins that are orthologous to the *C. elegans dep-1* and human PTPβ. By analysis of genomic sequences from various *Drosophila* species, mosquito species, and other insects, *Ptp4E* is thought to have arisen from a recent duplication of an ancestral gene closely resembling the modern *Ptp10D*. This duplication appears to have occurred recently, as it is only found in the drosophilid lineage of insects (Fig 510B).
Figure 510. Ptp4E and Ptp10D show a synthetic phenotype for central nervous system development in Drosophila. A. The developing embryonic nervous system of Drosophila. Axons stained in brown. Arrowheads point to areas where the longitudinal tracks of axons are discontinuous. from the Jeon paper B. [SGF-1290] Evolutionary analysis demonstrates that Ptp4E is likely due to a recent gene duplication.

Related readings:


**Box 5-3 Why is the definition of epistasis so confusing?**

Epistasis seems like such a specific term, so why is there confusion about its meaning? Classical geneticists and population biologists use this term to refer to different types of genetic phenomena. To understand this disparity, it is worth understanding how this confusion arose.

The term “epistatic” was first used by the British geneticist William Bateson, a British geneticist. Bateson coined the term during his discussion of genetic interactions between alleles at different loci. In this discussion, he was trying to explain a variation from the Mendelian ratios that were expected when two genes are independently assorting. Bateson used epistatic to describe the observed phenomenon of how an allele of one gene can prevent the allele of another gene from having an effect. This original definition of epistasis, where an allele masks the effect of another allele, is the one typically used by molecular geneticists, and is sometimes referred to as **physiological epistasis** or classical epistasis (and is the definition we use in Chapter 5).

The confusion in the meaning of epistatic arose when in 1918, R.A. Fisher, a British evolutionary biologist, uses “epistacy” to refer to genetic interactions in a broader sense, including those that demonstrated non-additive interactions among different genetic loci (this is often called **generalized epistasis**). To this day, quantitative geneticists and epidemiologists, particularly those who look at statistical models of inheritance, commonly use Fisher’s definition. This term covers many different types of genetic interactions, and thus is broader in its potential meanings. In much of this book, we will use the narrower classical Bateson definition. The epistasis analysis we describe in this chapter involves two-locus physiological epistasis: that is, the analysis of genetic interactions by examining the phenotype in double mutants.
**Box 5-4 Using epistasis to order genes during sex determination in C. elegans**

The ratio of sex chromosome to autosome ratio determines the sex of many organisms. How is this ratio interpreted? Epistasis analysis was key to ordering the gene important in the pathway used to read this ratio so that animals would develop as either males or females.

**Epistasis analysis allowed for the ordering of tra-1 and her-1**

As mentioned in this Chapter, the tra-1 and her-1 genes act together in a switch regulation pathway important for somatic sex determination in *C. elegans* (see Figure 1 [SGF](https://doi.org/10.1094/SGF-1250A)). Recall that her-1 mutants produce female bodies even when they are 1X:2A, while tra-1 mutants produce male bodies even if they are 2X:2A. Analysis of gain of function alleles in her-1 and tra-1 were consistent with these genes acting in a switch regulation pathway. Gain of function alleles resulted in the opposite phenotype of what was seen in null alleles for these genes. Gain of function alleles in her-1 result in animals with female bodies even when they are 2X:2A, and gain of function alleles in tra-1 result in animals with male bodies even when 1X:2A.

In what order do tra-1 and her-1 act in the pathway for somatic sex determination? This information can be gained by constructing an organism carrying null alleles in both of these genes. The double mutant animal defective in both of these genes (her-1; tra-1) produces a male body regardless of the number of X chromosomes. Thus, tra-1 is epistatic to her-1, because in the double mutant the presence of the tra-1 mutation prevents the her-1 mutation from having an effect. From this epistasis, we infer that tra-1 acts downstream of her-1 (Fig 1 [SGF](https://doi.org/10.1094/SGF-1250A)).

A.  

B.  

**Figure 1. [SGF](https://doi.org/10.1094/SGF-1250A). tra-1 is epistatic to her-1**

**Epistasis allows inference of states of gene activity in the sex determination pathway**

What does this epistasis mean? In other words, does the ordering of this pathway make sense with the data? Let's first consider what happens in the normal animal. When an animal is 1X:2A, the HER-1 protein is activated. ([1X:2A]; Figure 2A [SGF](https://doi.org/10.1094/SGF-1250A)). HER-1 then represses the TRA-1 protein, which normally acts to shut down the pathway leading to the male body and turns on the pathway leading to the female body. Thus, when TRA-1 is repressed and not functional, the default male program of development will be carried out. When the ratio is 2X:2A, HER-1 is not active, leading to a functional TRA-1, which activates the female program of development and represses the male program of development (Figure 2 [SGF](https://doi.org/10.1094/SGF-1250A)).

Does this interpretation make sense with the data? A her-1 mutant does not produce HER-1 protein. Thus, HER-1 cannot get activated to repress TRA-1. In other words, having active TRA-1 always leads to the production of a female body and the repression of the male program of development, regardless of the X:A ratio (Figure 2B [SGF](https://doi.org/10.1094/SGF-1250B)). A tra-1 mutant does not produce TRA-1 protein, and thus cannot repress the male program of development even in 2X:2A animals. Thus, tra-1 mutant animals always produce a male body (Figure 2C [SGF](https://doi.org/10.1094/SGF-1250C)). In the tra-1 her-1 double mutant animals, there is no TRA-1 to repress the male program of development. Thus, the tra-1 her-1 double mutant animal will activate the male program of development, even in 2X:2A animals (Figure 2D [SGF](https://doi.org/10.1094/SGF-1250D)).
Figure 2 [SGF-1250]. Epistasis in the *C. elegans* sex determination pathway. Grey arrows and bars are inactive.

**Molecular identification of her-1 and tra-1 confirm epistasis ordering**

The identification of the *her-1* gene demonstrated that *her-1* encodes a secreted protein that promotes male development in a cell non-autonomous pathway. The *tra-1* gene encodes a transcription factor that effects the program of gene expression. The HER-1 protein represses the activity of a signaling pathway in the responding cell that will ultimately turn off the transcription factor encoded by *tra-1*. When the TRA-1 transcription factor is active, it promotes the expression of genes that lead to the development of the female body and prevents male development. Of course, there are other gene products that act between HER-1 and TRA-1.

**Related readings:**


Box 5-5 Example of using incompletely penetrant nulls to associate ligands and receptors

Analysis of double mutants with complete loss-of-function mutations that are nonetheless incompletely penetrant was particularly useful in studies of WNT receptors in *C. elegans*. WNTs are the ligands for proteins that are transmembrane WNT-receptor, which bind WNTs and transduce a signal across the plasma membrane. WNTs are secreted glycoproteins. Wnt signaling is important, as WNTs and their receptors are used for intracellular signaling during development to control cell fate and cell polarity. WNTs are also used for stem cell development, and cancer can occur when cells misregulate the output of WNT signaling.

In *C. elegans*, two WNT ligands were identified that each acted independently on two WNT receptors. However, mutations in each ligand and receptor only gave a partially penetrant phenotype. The key to associating the WNTs with their appropriate receptor came from analysis just like described in Chapter 5, Figure 23.

The phenotype in question is the polarity of a vulval precursor cell called P7.p (posterior daughter to cell P7), which divides asymmetrically, with each daughter generating distinct set of progeny cells (Figure 1). However, P7.p can be oriented either to the anterior or the posterior of the animal. In wild-type animals it is oriented to the anterior. Another precursor cell, P5.p, is oriented to the posterior.

In the absence of the WNT-receptor *lin-17*, only 24% of the animals have P7.p with correct orientation (Table 1). In the absence of another receptor, *lin-18*, only 59% of the animals have P7.p with correct orientation. A *lin-17 lin-18* double mutant has a completely penetrant phenotype, with no P7.p with correct orientation. Thus, these two receptors are partially redundant. What about the ligands? Mutants of the ligands *lin-44* and *mom-2* have essentially wild-type phenotypes. However, a *lin-44 mom-2* double mutant has a striking phenotype, even though still partially penetrant. The fact that the ligand double mutant is less severe than the receptor double mutant could be due to *mom-2* mutant not being null, or another ligand also acts on these receptors. Can we associate ligands and receptors? If there is a one-to-one match of ligand and receptor, we expect that eliminating activity of a ligand and its receptor would not be more penetrant than the most severe of the single mutants. By contrast, if a ligand acts independently of a particular receptor, then we expect a more penetrant phenotype. Consider *mom-2*. The *lin-17 mom-2* double mutant is 3% normal, indicating that this receptor and ligand act independently. However, the *lin-18 mom-2* double mutant is 56% normal, not different than the *lin-18* single mutant (59%). Therefore MOM-2 is more likely the ligand for LIN-18. For *lin-44*, the data are a bit sloppier but are similarly informative. *lin-44* weakly suppresses *lin-17* (and thus certainly does not enhance *lin-17*!), but enhances *lin-18*. Therefore LIN-44 is likely the ligand for LIN-17.

**Figure 1** [SGF-1334]. Vulval lineage polarity.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>%WT</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>100</td>
<td>wild-type control</td>
</tr>
<tr>
<td><em>lin-17</em></td>
<td>24</td>
<td>strongly penetrant mutant</td>
</tr>
<tr>
<td>Genotype</td>
<td>Frequency</td>
<td>Phenotype</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>lin-18</td>
<td>59</td>
<td>weakly penetrant mutant</td>
</tr>
<tr>
<td>lin-44</td>
<td>100</td>
<td>essentially wild type</td>
</tr>
<tr>
<td>mom-2</td>
<td>99</td>
<td>essentially wild type</td>
</tr>
<tr>
<td>lin-17; lin-18</td>
<td>0</td>
<td>fully penetrant</td>
</tr>
<tr>
<td>lin-44 mom-2</td>
<td>41</td>
<td>weakly penetrant mutant</td>
</tr>
<tr>
<td>lin-17 lin-44</td>
<td>42</td>
<td>weakly penetrant mutant</td>
</tr>
<tr>
<td>lin-17 mom-2</td>
<td>3</td>
<td>almost fully penetrant</td>
</tr>
<tr>
<td>lin-18 mom-2</td>
<td>56</td>
<td>weakly penetrant mutant</td>
</tr>
<tr>
<td>lin-18 lin-44</td>
<td>12</td>
<td>almost fully penetrant</td>
</tr>
</tbody>
</table>

**Diagram:**
- lin-44 ➔ lin-17 ➔ P7.p anterior orientation
- mom-2 ➔ lin-18

*Image not included*
Box 5-6 Epistasis analysis was key for identifying and ordering key players in the cell death pathway

Programmed cell death, or apoptosis, is an important process that organisms use to remove cells that are no longer needed or to remove cells that might be harmful to the organism (such as those with extensive DNA damage). Apoptosis is also critical for the normal shaping and sculpting of organs and tissues during development. For example, apoptosis occurs during the development of the human hand, as the cells between the fingers are eliminated. We now know that apoptosis is extensively regulated in cells and involves a cascade of serine proteases called caspases. The control of caspase activity is critical for apoptosis. When caspases are activated, they destroy the contents of the cell in a regulated fashion and also cause dying cells to send out the appropriate signals so that they are neatly removed by phagocytic cells. Apoptosis is distinct from necrotic cell death, which occurs upon acute cell injury. During necrosis, cells spill their contents onto neighboring cells, often triggering potentially harmful inflammatory responses.

Pioneering studies in *C. elegans* were important for identifying and ordering the key players in apoptosis. This important work led to the Nobel Prize in Medicine awarded in 2002 to three *C. elegans* researchers, Sydney Brenner, H. Robert Horvitz, and John Sulston, for their discoveries important in organ development and programmed cell death. Apoptosis involves a switch regulation pathway important for regulating the decision to undergo programmed cell death, followed by a substrate-dependent pathway important for the dying cell getting engulfed by a phagocytic cell and for the DNA degradation that occurs in apoptotic cells.

*C. elegans* was a particularly good system for studying apoptosis because scientists could watch the death of the 131 cells that died during the development of the nematode. This allowed for the isolation of genes important in the decision to undergo apoptosis, where mutations in these genes led to cells living that were supposed to die, or the opposite case, where cells that normally live would undergo apoptosis.

A switch regulation pathway regulates the decision to undergo apoptosis

Genetic analysis suggested that *ced-3, ced-4*, and *ced-9* were important in the decision to undergo apoptosis. Loss of function mutations in the *ced-3* and *ced-4* genes led to cells surviving that normally died, sometimes referred to as "undead" cells. Gain of function mutations in the *ced-9* gene led to the prevention of apoptosis, while null mutations in *ced-9* resulted in animals dying as embryos from ectopic cell death (Figure 1 [SGF-1300]).

![Wild type and photo of dying cell(s)](image)

![No cell death and all cell death](image)

Fig. 1 [SGF-1300]. Phenotypes of the *ced-3, ced-4*, and *ced-9* genes
These genes were good candidates for being in a switch regulation pathway regulating cell death. Double mutants analysis of loss of function ced-9 mutations with either ced-3 or ced-4 loss of function mutations demonstrated that ced-3 and ced-4 were epistatic to ced-9. In the double mutants, the lethality associated with the ced-9 null phenotype was masked by the mutation in either ced-3 or in ced-4. These data suggested ced-3 and ced-4 were required for ced-9 to act, and ced-9 negatively regulates these genes.

As ced-3 and ced-4 have similar loss-of-function phenotypes, the ordering of ced-3 and ced-4 relative to each other was carried out using gain of function mutations. The overexpression of ced-3 or ced-4 results in a cell-autonomous apoptosis phenotype. Double mutant analysis showed that the killing caused by ced-4 overexpression was masked in the ced-4(gf); ced-3(lf) double mutant. However, the converse experiment showed in the ced-3(gf); ced-4(lf) double mutant, the killing caused by ced-3 was not blocked. These experiments are consistent with ced-4 acting upstream of ced-3, leading to the pathway in Figure [SGF-1301].

**Fig. 2 [SGF-1301]. The ced-3 and ced-4 genes were ordered using gain of function alleles.**

We now know that ced-9 encodes a protein similar to the mammalian BCL2 protein known to be involved in cell death in mammals. ced-4 encodes a protein similar to the mammalian Apaf-1 protein that was biochemically purified using an in vitro assay to identify proteins involved in programmed cell death. The identification of the ced-3 protein product was particularly important for the cell death field, because ced-3 encodes a caspase, and caspases were not previously known to function in apoptosis. Thus, these studies were important in both the ordering of the pathway and the identification of important components in apoptosis. Subsequent biochemical and structural studies supported the genetic conclusions, demonstrating that CED-9 and CED-4 proteins form a complex that inhibits CED-4 from activating CED-3. CED-3 is produced in an inactive form that becomes active by interacting with CED-4.

**A substrate-dependent pathway is involved when the dying cell is to be engulfed by a phagocytic cell**
Once a cell dies by apoptosis, it is neatly removed from the organism by a phagocytic cell that engulfs it to prevent the dying cells’ content from spilling into the organism. The ced-1 gene is involved in the engulfment of dying cells, as animals lacking ced-1 contain dying cells that are not engulfed. When the ced-1(lf) ced-3(lf) double mutant is analyzed, the ced-3 "undead" cell phenotype is seen. Because the loss of function phenotypes for ced-1 and ced-3 are not the opposite of each other, the pathway containing ced-1 and ced-3 is considered a substrate dependent pathway. The phenotype seen in the ced-1 ced-3 double mutant suggests that the cell needs to first make the decision to die (in a process which requires ced-3) before its corpse can be engulfed (through a process that uses ced-1). The ced-1 gene encodes a transmembrane protein and functions in the engulfing cell, likely acting as a receptor for the signal present on the dying cell (Figure 3 [542]).

A.

picture of unengulfed dead cell in ced-1?

B.

switch-regulatory pathway

ced-9  \[\rightarrow\]  ced-4  \[\rightarrow\]  ced-3  \[\rightarrow\]  ced-1  \[\rightarrow\]  cell engulfment

substrate-dependent pathway

Figure 3 [542]. The apoptotic pathway is a complex pathway involving both switch-regulatory and substrate-dependent portions.

The DNA degradation pathway is also a substrate-dependent pathway

The engulfment of cells is important not only for eliminating the dead cell, but can also promote the changes seen in dying cells. For example, the degradation of DNA in the dying cell is another hallmark of apoptosis, and occurs in a substrate-dependent fashion. DNA degradation occurs in several steps during apoptosis, as determined by various staining methods. Intact DNA first undergoes an initial degradation process, then is further degraded by the nuc-1 nuclease, and finally converted into free nucleotides. Interestingly, unengulfed dying cells in the ced-1 mutants did not initiate DNA degradation, suggesting that the engulfment process mediated by ced-1 was necessary for promoting the initial step of DNA degradation. The ced-1 nuc-1 phenotype for DNA degradation is like ced-1, consistent with nuc-1 acting after the step requiring ced-1. The initial degradation process is likely mediated by the CRN nucleases, which were initially identified through RNAi screening (Figure 4 [SGF-1302//543]).
Figure 4 [SGF-1302]. The DNA degradation that occurs when cells undergo apoptosis involves a substrate-dependent pathway.

Related Readings:


