

Chapter 4: Individual gene function

As a geneticist, interesting genes to study are often identified by their phenotype, because of some compelling reasons to examine the difference between the organism containing the mutant gene and the wild type organism. When studying human beings, a person with a disease may be interested in understanding the genetic component of their disease. When using model organisms, certain aspects of that organism can be used to understand the basic biology underlying normal functions. In *C. elegans*, the development of an egg laying structure, the vulva, has been used to learn about signal transduction pathways important for human development and disease. How can mutations that lead to the inability to lay eggs in *C. elegans* tell us about the function of the gene the mutations are in (Figure 1)? How can information about how the alleles of this gene behave be used to understand the function of this gene? Can these alleles be used to understand more about what tissues are important for controlling the development or behavior of the organism? As geneticists, we are often trying to infer the normal function of a gene from mutations. How a gene functions can often be inferred by analyzing the relationship of a mutant allele and its relationship with other alleles of the gene. Sometimes, specific types of alleles can be created for these genetic analyses. Mutant alleles can also be used to infer when a gene acts and in which tissue the gene acts, through time-of-action and site-of-action experiments. These types of genetics experiments are important complements of molecular and biochemical experiments, and this chapter will focus on how the genetic analysis of a single gene is done.

4A. Inferring gene function from mutations

The function of an individual gene can be assessed by perturbing its activity and observing the consequences for the phenotype of the organism. In order to understand what the phenotype of the mutant allele tells us, it is important to know what type of allele you are working with. In Chapter 3, we briefly touched on how alleles represent different variants of a gene, and how some alleles are recessive while others are dominant. In this section, we discuss in more depth the types and classifications of mutant alleles and how these various alleles can help inform us about gene function.

Alleles alter gene activity

In Chapter 3, we discussed mutant alleles and how they are inherited. Molecularly, alleles are alterations in genes at the DNA level, and can range from a complete deletion of the entire gene to nucleotide-level alterations that alter gene function. Usually, the most common allele is considered the “wild-type” allele, and this allele is most often fully functional and dominant.

Even without knowing the molecular function of the gene product, alleles can be placed into different classes based on their inferred **gene activity**. Gene activity is a term used to describe the abstract concept of gene function. The concept of activity includes any function of a gene, including the ability of its product to catalyze a reaction, to bind to another molecule, or to act as a nutrient reservoir. Geneticists typically measure gene activity by observed phenotype. For example, if you had a mutant organism carrying a recessive mutation in a single gene, and saw that this mutant organism lacked pigment in their fur when compared to wild-type organisms, you might infer that the normal activity of this gene was involved in producing pigment. Different alleles of this gene that led to organisms with different amounts of pigment in their fur would lead to the inference that these alleles all have different amounts of gene activity

(Figure 2 [SGF-1141]). Further genetic studies could be carried out to determine how much gene activity this allele has.

Mutations can be classified by their effect on gene activity

In order to properly infer the wild-type activity of a gene from the properties of the mutation, we thus need to know what the mutation does to the activity of the gene. In the example above, we inferred that the lack of gene activity leads to a lack of pigment in the fur. This inference would only be valid if the alleles we are examining lack gene activity; if those alleles led to increased gene activity, a different conclusion would be correct.

How do we know whether an allele leads to a lack of gene activity? Careful studies can be carried out to determine and properly classify how the mutation affects gene activity. Hermann Muller, at the International Congress of Genetics in 1932, defined six different classes of alleles and how they alter gene activity compared to a wild-type reference allele (Table 4-1). An **amorph**, which is sometimes called a **null** completely eliminates activity of the gene. All null alleles are also **loss-of-function alleles**, which refers to alleles where gene activity has been eliminated. The distinction between these two terms is further discussed below. A **hypomorph** has less than the wild-type level of activity; these alleles are also known as reduction-of-function alleles. Gain-of-function mutations were broken down into three classes: **hypermorph**, **neomorph** and **antimorph**. A hypermorphic allele led to an increase in gene activity, and is usually dominant. A neomorphic mutation causes the gene to acquire a novel function, and is also usually dominant. An antimorph encodes a gene product that now interferes or antagonizes the wild-type gene function, and is sometimes referred to as a **dominant negative** allele. How genetic tests can be used to determine the classification of mutant alleles is discussed in this section.

Null alleles are the gold standard for inferring gene function

The strongest inference of gene function can be made by completely eliminating the function of the gene. Other types of alterations in gene activity can be highly informative but occasionally misleading (discussed below). By examining the phenotype of an organism that completely lacks the function of the gene, you can then infer how that gene is used in the organism.

For example, a null allele that causes a fly to lack wings may lead you to infer that this gene is necessary for flies to form wings (Figure 3). However, imagine a case where a gene in flies is used to make both eyes and legs. If you examined a hypomorphic allele that only disrupted the gene's function in the eye, you would have a fly without eyes. If you mistakenly thought this allele you were examining was a null allele, you might then infer that this gene is necessary for eye formation but not necessary for leg formation. A null allele of this gene should result in flies without both eyes and legs.

Because a null allele is the gold standard for genetic inference, great care is taken to assess whether a particular allele is really a null, or loss-of-function allele. A molecular definition of a loss-of-function allele is that the DNA corresponding to the gene is absent from the organism. In this case, without the ability to produce the gene product, the allele must lack all gene activity. Although creating alleles that precisely delete the gene from the genome may sound simple, achieving this is not always trivial. First, only a few organisms have the experimental techniques available to allow for the routine removal of desired sequences; the techniques involved in these manipulations are described below. Second, gene structure can be complex, particularly in multicellular eukaryotes, making it difficult to simply remove a gene.

Since precise deletions of genes are not typically available in most organisms, other criteria are used to assess whether an allele behaves as a loss-of-function allele. A **DNA null** lacks the DNA required for the gene product (Figure 4 [SGF-1142]). An **RNA null** lacks detectable mRNA produced from the gene. A **protein null** lacks a detectable protein product. One

important caveat to using gene products (RNAs or proteins) as evidence that an allele is a null is that a gene could have activity at a low level of product that cannot be easily detected. It is possible that even when we cannot detect a gene product, there may be enough gene product present in the right place and right time to provide sufficient gene activity.

Some loss-of-function alleles may still produce DNA, RNA, or protein, but still lack gene activity. This might happen if a gene lacks the ability to make the crucial functional domain of its product, it may not be able to function even if detectable protein product is present. For example, a gene encoding an ion channel, which lets ions pass through a cell membrane usually in a regulated fashion, might have an allele that would not be able to make the trans-membrane pore, and thus would not function despite the mutant gene being transcribed and translated. Similarly, a gene encoding a protein tyrosine kinase might have an allele that lacks the ATP-binding region of the kinase domain, effectively disabling the important enzymatic activity of the kinase. We refer to these types of alleles as loss-of-function alleles; these types of alleles are not true “null alleles”, as gene product is produced from these alleles. Thus, loss-of-function alleles describe a wider scope of allele compared to null alleles (which are also loss-of-function). The genetic tests described below can allow you to determine whether the allele behaves as a loss-of-function allele.

Genetic tests can indicate whether a mutation is a loss of function allele

Genetic tests can be used to assay whether an allele can have its gene activity further reduced. If it can, that allele cannot be a loss-of-function allele, as loss-of-function alleles do not have any gene activity. This type of genetic determination can be particularly useful when the molecular nature of the gene is not known, or, if molecular tools available cannot clearly assess whether an allele is a loss-of-function allele. For example, molecular tools would not be available when a gene is defined by mutation, which is typical of mutations acquired through a genetic screen (as described in Chapter 7). Alternatively, the corresponding DNA for mutations that are naturally occurring mutations may not be known. In these types of cases, it is important to properly infer gene function from mutant gene activity.

In diploid organisms, chromosomal **deficiencies** (*Df*; see Box 317) can be used as a simple test of whether a mutation is a loss-of-function allele. These *Dfs* are useful because they typically lack the chromosomal region containing the gene of interest, and thus must be a null for the locus. An organism that has the mutant allele on one chromosome and a genetic *Df* of the locus on the other chromosome is created and analyzed. If the heterozygote b_1/Df does not have the same phenotype as the homozygous b_1/b_1 , we can conclude that the b_1 allele is not complete loss-of-function (Figure 5 [SGF-1107]). If b_1/Df behaves like the homozygous b_1/b_1 , this result would be consistent with the b_1 allele being a loss-of-function.

Ideally, the *Df* used for this analysis would be a small deletion that completely removes the gene of interest but nothing else. Practically, this situation can be difficult to achieve in most organisms, as typically this analysis is done using pre-existing *Dfs* that remove small regions of the chromosome. Creating of a *Df* that only removes the gene of interest can be achieved through a non-complementation screen using a mutagen that leads to chromosomal deletions (see Chapter 7), or through homologous recombination (see below). Collections of strains that have small, defined deletions of chromosomes throughout much of the genome (sometimes called **deficiency collections**) can be used for this purpose. The *Dfs* in these collections usually remove more than one gene. Deficiency collections exist for some model organisms (i.e., *C. elegans* and *D. melanogaster*). Strains containing the *Dfs* are usually heterozygous because then the wild-type copy of the gene on the non-*Df* chromosome can usually compensate for the missing genes on the *Df* chromosome.

An important caveat to remember when using a *Df* that removes more than the locus of interest is that some of the other genes removed by the *Df* may affect your results. For example, if the *Df* used also removes one copy of a negative regulator (or other suppressor; see Chapter

5), your analysis of the heterozygous b_1/Df strain can be affected. Consider the case with a mutation in the gene C , which you place in *trans* to a Df , to create $c_1 +/ Df$. If the chromosomal deficiency removed both gene C and gene D , and if the lack of gene D acts as a dominant suppressor (d_{sup}) of the loss-of-function phenotype of gene c , then a deficiency that removes both c and d would alter the observed phenotype ($c_1 +/ Df[c_{if} d_{sup}]$). This issue can be addressed by using multiple Dfs , each which take out different regions surrounding the gene of interest.

An alternative method for assaying whether an allele is a loss-of-function allele involves using molecular tools to potentially further reduce gene activity. For organisms for which RNAi (or equivalent knock-down method, such as **morpholinos** [Box ---]) is available, RNAi knock-down (Box 309) can be used to test whether any gene activity remains in a particular mutant allele. For example, the phenotype of an untreated homozygous strain [b_1/ b_1] is compared to the same homozygous strain b_1/ b_1 treated with RNAi against gene B . If the untreated strain does not have the same phenotype as the strain treated with RNAi, then b_1 is not complete loss-of-function, as gene activity can be further reduced in this strain. This type of analysis is simplest to interpret if the phenotype can be measured by some quantitative means, such as by examination of penetrance or expressivity (Table 4-2). For a discrete phenotype, this type of analysis will not distinguish a further reduction in gene activity below the threshold necessary for a non- wild-type phenotype.

Loss-of-function mutations can be constructed

With a genome sequence, sometimes we may see potential gene products where we may wish to know the importance of this gene product on an organism. For example, perhaps there is a new protein that has been associated with a particular disease, but the mechanism of how this protein works in a cell is unknown. In this case, one useful line of investigation may be to remove the function of the gene encoding the homologous protein in a model organism and study its mechanism of action there. Alternatively, genetic analysis may suggest that existing alleles are not loss-of-function alleles, thus creating the desire to engineer a null mutation.

Given a genome sequence, a definitive way of obtaining a loss-of-function mutation for genetic analysis is to construct one. Constructing loss-of-function mutations, which are also known as **gene knockouts**, is possible in some organisms, including *S. cerevisiae*, mouse, and flies. The general concept of creating these gene knockouts is to use recombination to remove an endogenous gene and replace it with other DNA sequences. The small genome size of the budding yeast, along with its propensity for homologous recombination makes the construction of complete DNA nulls a straightforward process. This process (Figure 6 [SGF-1108]) involves using the known sequence to define the portion of the genome to be removed. For most genes, the removal of the open reading frame from the start codon to the stop codon is a useful strategy for creating a DNA null. *S. cerevisiae* will undergo homologous recombination when stimulated by linear DNA, as usually (not in the lab!) the presence of linear DNA suggests a broken chromosome that needs to be fixed (Figure 7 [SGF-1091]). The yeast cell will scan for homologous sequences, and undergo recombination. Researchers exploit this ability by creating a double stranded linear piece of DNA that has homology to the ends of the gene that is to be excised from the genome. The insertion of this DNA can be selected for using a marker such as a drug resistance marker. The proper insertion of the exogenous DNA into the defined DNA region needs to be verified, to ensure that the gene that was targeted for removal was indeed removed.

Strategies in other organisms use similar experimental designs to create loss-of-function mutations. However, the larger gene size in flies and mice means that sometimes these techniques will lead to a situation where only part of the gene is removed. In these cases, it is important to try to remove as much of the functional domains of the gene products as possible, and to assay the created allele for gene activity, as described above. In organisms where targeted recombination is not facile, using designer nucleases (BOX 316), screening for random

alterations in the gene of interest can yield the desired mutation, or examining natural variants to find mutations within the gene of interest. Specifics about screens will be discussed in Chapter 7. Chapter 15? describes finding alleles by re-sequencing the genomes of mutant strains.

Multiple alleles of genes can help define gene function.

So far we have been mainly discussing the use, creation and characterization of null, or loss-of-function alleles. However, the other types of alleles (Table 4-1 above) are also important tools for genetic analyses. For example, hypermorphic and hypomorphic alleles can be useful in screens (Chapter 7) and for the ordering genes within pathways (Chapter 5). Having more than one class of allele in hand can more accurately indicate gene function (Table 404). Loss-of-function mutations define the **necessity** of a gene for a particular process. A gain-of-function mutation implies the **sufficiency** of a gene for a particular process. (Since genes do not act alone, by sufficiency we mean all other things being equal.) A neomorphic allele might cause a phenotype unrelated to the normal function of the gene as defined by loss-of-function; in this case we can say the gene is **sufficient but not necessary**. For example, if the neomorphic allele results in a protein being expressed in an inappropriate place. Another case of sufficiency without necessity is redundant genes: loss of function has no effect, but hyperactivity does. To use these other types of alleles (see Table 4-1 above), it is important to understand the nature of these alleles, and we will further discuss the characterization of these other types of alleles.

Frequency of obtaining mutations can suggest the nature of the mutation.

In general, loss-of-function or reduction-of-function alleles are more common than gain-of-function alleles. This is because there are usually more ways to reduce or disrupt the activity of a gene by mutation compared to the ability of a mutation to alter or increase gene activity.

Let's first consider the types of mutations that can occur within the open reading frame of protein coding genes. Alterations of nucleotides within the open reading frames can lead to amino acid changes, premature stop codons, or **frameshift mutations**.

Amino acid changes occur when a nucleotide within a codon is changed such that the codon now codes for a different amino acid (see Table 101 for genetic code). The type of allele created by an amino acid change depends on the characteristic of the introduced amino acid compared to the original. A reduction-of-function allele may occur if the new amino acid is similar to the original amino acid. In this case, only a slight change in gene activity may occur. For example, a very similar change may lead to a protein that is mostly functional, but occasionally may misfold and be non-functional. On the other hand, if this new amino acid has a different characteristic compared to the normal original amino acid, this new amino acid can change, and possibly inactivate, the function of the protein. Inactivation of the protein would lead to a loss-of-function allele. Sometimes, a single amino acid change can lead to a gain-of-function phenotype, if the amino acid change leads to a change that mimics an activated form of the protein. One example is a protein that is activated by phosphorylation, where replacing the endogenous amino acid with a negatively charged amino acid can sometimes cause the protein to take on its activated role, leading to a gain-of-function phenotype (Figure 8 [SGF-1119]). Negatively charged amino acids (glutamic acid and aspartic acid) can act as phosphomimetics because the covalent attachment of a phosphate group to an amino acid during phosphorylation introduces three negative charges to the region of the protein (Figure 9 [SGF-1113]).

The creation of a premature stop codon can lead to a truncated protein product. Premature stop codons are created when mutation causes a codon that normal codes for the insertion of an amino acid into a stop codon that terminates translation. For example, mutation of tyrosine codon TAA in the DNA to stop codon TAA would terminate translation (Figure 10 [SGF-1109]). Truncated proteins that lack most of the protein typically behave as loss-of-function alleles. However, if the premature stop codon removes a regulatory domain that normally keeps

the protein in an off state until activated, the premature stop codon may lead to a hypermorphic allele. Similarly, a truncated protein that removes proper cellular localization information may create a neomorphic allele that now acts in the wrong compartment within the cell.

Frame-shift mutations involve insertion or deletion of 1-2 base pairs. This type of deletion will alter the reading frame, effectively deleting all downstream codons. Often times, a frame-shift mutation also leads to the insertion of some number of abnormal codons after the mutation, until a stop codon is reached. The effect of a frameshift mutation will thus depend on the protein and the location of the frame-shift. For example, in [Figure 11B \[SGF-1052\]](#), a frame-shift in the first half of the protein would likely disrupt activity, but one later in the gene would disrupt the negative regulatory domain, potentially leading to a hyperactive protein. In many cases, a long untranslated region after a stop codon or the stop codon in the altered reading frame will result in the mRNA being degraded by the Nonsense Mediated Decay (NMD) pathway, a cellular surveillance pathway that scans for errant mRNAs. This degradation by the NMD pathway could thus result in a mutation acting as a loss-of-function allele even though the conceptual translation product suggests a gain-of-function phenotype due to the removal of the negative regulatory domain!

Mutations within the regulatory domains of genes can be more difficult to understand by purely examining the location of the mutation. Some examples of these types of mutations may be changes that affect mRNA splicing, or changes within the promoter regions governing gene expression. For example, if a binding site for a critical transcriptional activator was abolished by mutation, there may be no transcription and the mutation would create a loss-of-function allele. Alternatively, if a binding site for a crucial transcriptional repressor was eliminated by mutation, the gene may be produced at inappropriate times, leading to either a gain-of-function allele or a neomorphic allele.

During genetic analysis, if there are many alleles of the same gene that are caused by different molecular changes, the most frequent phenotype can sometimes be thought of as the reduction- or loss- of function phenotype, since the frequency of obtaining these types of alleles are more common. However, this is only the case if there was no selection for the phenotype. For example, if you are selecting for alleles of a gene that cause a particular phenotype, you may obtain many alleles that have that phenotype, which would of course then tell you nothing about the frequency of a mutation that has that phenotype. Such screens are discussed in Chapter 7.

Dosage studies can distinguish hypomorphs and hypermorphs

For the interpretation of genetic experiments, it is critical to understand the nature of the allele. This understanding is acquired through examining the relationship of the activities of the wild-type to mutant alleles. Although loss-of-function alleles are considered the gold-standard for genetic analysis, the analysis of hypomorphs and hypermorphs can also be useful. An obvious example of this utility would be cases where loss-of-function alleles lead to the death of an organism, since it is difficult to analyze dead organisms. A more subtle use may be a gene that is used at different times in the development of an organism. In this case, a loss-of-function allele will remove the gene activity at the first needed instance, and if that activity is required for further development, later uses of the gene cannot then be assayed using a loss-of-function allele. When using hypomorphs and hypermorphs for genetic analysis, it is important to distinguish these alleles from neomorphs and antimorphs.

Dosage studies refer to experiments where the number of copies of an allele in an organism is experimentally manipulated. In these types of studies, the allele of interest is placed in trans to another allele of the same gene. To start, this analysis is typically done with a wild-type allele and a mutant allele. The severity of the phenotype (i.e., whether the phenotype is less penetrant or less highly expressed) is assayed relative to the phenotype of the homozygous mutant organism.

A hypomorphic allele is one that causes a partial reduction of gene activity (see Table 401). In a dosage study, a hypomorphic allele differs from a null allele because an organism that is h_1 /deficiency will have a more severe phenotype than an organism that is homozygous for the hypomorphic allele (h_1/h_1). This is the case because if h_1 is a hypomorphic allele, an organism that is h_1/h_1 will have more gene activity than an organism that is h_1 /deficiency. Using similar logic, if h_1 has partial activity, multiple copies of h_1 (for example: $h_1/h_1/h_1$) would cause to a less severe phenotype than two copies of h_1 . More than two copies of a gene can be created in some organisms for dosage studies by the introduction of extrachromosomal copies of a gene.

A hypermorphic allele creates an increase in normal gene activity. If we consider a hypermorphic allele H , an organism that has H /deficiency will have a less severe phenotype compared to a strain that is homozygous H/H . For this test, a hypermorph will behave differently from a hypomorph, which has a more severe phenotype in trans to a deficiency. Similarly, additional copies of H or the wild-type allele should exacerbate the phenotype, as either of those alleles will continue to increase the amount of gene activity.

An example of how a hypermorphic allele might behave is illustrated in Table 403. In this example, we assume that a wild-type allele provides 50% activity, and the diploid with two wild-type alleles has 100% activity. A hypermorphic mutant allele H provides 150% activity (i.e., three times the wild-type level). H is dominant, as $H/+$ has 200% gene activity activity, and displays a phenotype. Homozygous H has 300% activity, and the phenotype is more severe than organisms that are heterozygous $H/+$. A strain that is hemizygous for H (H/Df where Df is deficiency) has 150% activity, and thus mutant in trans to a deficiency is less severe than the homozygote, by contrast to the example using a hypomorphic allele discussed above. Additional copies of either H or wild-type allele exacerbate the phenotype, as adding either of these alleles will increase gene activity. Three copies of H would have 450% activity, while $H/H/+$ would have 350% activity compared to 300% for H/H . In this example, we have not assumed any specific relationship between gene activity and phenotype, except that more activity has a more severe phenotype.

Dosage studies can distinguish antimorphs and neomorphs

Distinguishing hypermorphs from antimorphs and neomorphs can be difficult, and dosage studies can help with determining the novel nature of these alleles. A **neomorph** has an altered activity that is qualitatively distinct from the wild-type allele. Both neomorphs and hypermorphs behave in a dominant fashion in trans to a wild-type allele. Neomorphs are usually dominant, since the altered function is present to cause a phenotype. Neomorphs can be distinguished from hypermorphs because a hypermorph is suppressed in trans to a null, while neomorphs are not. This is because removing the wild-type gene product when a neomorphic allele is present should not make the organisms any better, since it is now left with only the altered gene product.

There are two classes of neomorphs: those whose phenotypes are not affected by the presence wild-type copy of the gene (dosage-independent neomorphs) and those whose phenotypes are sensitive to the presence of wild-type copy of the gene (dosage-dependent neomorphs). If the neomorph is acting in a dosage-independent fashion, the neomorph in trans to the wild-type allele would behave similarly to the neomorph in trans to a Df . Furthermore, addition of extra wild-type copies to an organism carrying a neomorphic allele does not change the severity of the phenotype caused by the neomorphic allele (Table 451). This behavior is distinct from a hypermorphic allele. Sometimes neomorphic products behave in a dosage-dependent fashion, where the neomorphic allele is rescued by the presence of wild-type products (Table 452). In these cases, increasing the copy of the wild-type gene and hence gene product can diminish the effect of the neomorph.

An **antimorph** encodes a poison product. These are sometimes called **dominant negatives**, an evocative name, but strictly speaking, antimorphs do not have to be dominant.

Antimorphs work because the gene product antagonizes the activity of the normal gene. If an allele is an antimorph, it will get better as more copies of the wild-type allele are added back, because then the amount of the poison product is being titrated down relative to wild-type gene product (Table 453). If an allele is an antimorph, it will still display a loss-of-function phenotype in trans to a deficiency. This effect because there is no wild-type gene activity in those organisms. A hypermorphic allele typically has a less severe phenotype in trans to a deficiency compared to in trans to a wild-type allele, while an antimorphic allele typically will have an equal or more severe phenotype.

An allelic series can imply gene activity

Once the nature of different alleles is understood, these alleles can be used for further dosage studies. These types of studies involve building an **allelic series**. An **allelic series** is a set of alleles that display graded effects on penetrance or expressivity; this set of allele is ordered with respect to levels of increasing gene activity. Having an **allelic series** can be useful for inferring normal gene activity. The hypothesis behind examining an allelic series is that these alleles fall along a single always increasing (monotonic) dose-response curve (Table 462; Figure 12 [SGF-1117]). This hypothesis can be tested by examining the phenotype of the various alleles in trans to each other, with the prediction that the activity of a two alleles in trans would lie between the level of activity seen in each of the two homozygotes. For example, in Figure 12 [SGF-1117], 95% of j_1/j_1 homozygotes are wild-type and 10% of j_5/j_5 homozygotes are wild-type, but 50% of the j_1/j_5 heterozygotes are wild-type. The alleles can be ordered in decreasing activity as $j_+ > j_1 > j_2 > j_3 > j_4 > j_5 > j_6$, and the inference is that the null phenotype of J is 0% wild-type even though the strongest allele is j_6 , with a 97% penetrant phenotype as a homozygote. We infer that a complete loss-of-function would have 100% penetrance. In BOX 315, a real example shows how this type of argument was used.

By ordering the alleles in order of inferred gene activity, these alleles can be used to study how the effects of decreasing gene activity affects phenotype. Although using non-null alleles is the best for inferring gene function, sometimes the loss-of-function mutation in a gene leads to a less than useful phenotype. For example, perhaps the gene is essential for an organism to live, and results in a dead organism. In this case, by examining allele combinations within an allelic series, you might be able to extrapolate to the null state and thus infer what that phenotype would be.

Phenotypic series reveals multiple functions of a gene corresponding to different activity levels

Genes often affect many aspects of phenotype. In some cases each phenotype arises at a different activity level. This phenomenon results in a **phenotypic series**. Figure 12.5 [SGF-1125] illustrates how distinct thresholds of the gene activity required for each phenotype results in a phenotypic series. Consider a gene Q , mutation of which results in three different phenotypes, **Q1**, **Q2** and **Q3**. Loss of Q function by the q_{null} allele results in the **Q1**, **Q2** and **Q3** phenotypes; severe reduction of Q by the q_2 allele results in **Q1** and **Q2**, and mild reduction of Q by the q_1 allele results in only the **Q1** phenotype. The simplest hypothesis is that **Q1**, **Q2** and **Q3** form a phenotypic series with non-**Q1** requiring the most Q activity, non-**Q2** a medium amount and non-**Q3** requiring the least activity. A q_1/q_{null} heterozygote would now display the **Q1** and **Q2** phenotypes as the activity of Q has dropped below the non-**Q2** threshold. Similarly, a q_2/q_{null} heterozygote would now display the **Q1**, **Q2** and **Q3** phenotypes as the activity of Q has dropped below the non-**Q3** threshold. In this example, the q_1/q_2 heterozygote would display a partially penetrant **Q2** phenotype as the activity of Q is intermediate between q_1/q_1 and q_2/q_2 . Table 433 shows this in tabular form.

Haploinsufficiency is less common because cells can usually function within a range of gene activity

Sometimes, a loss-of-function of a hypomorphic allele may be dominant, a situation referred to as **haploinsufficiency** (Figure 13 [SGF-1114]). In these cases, the organism is particularly sensitive to gene dosage, which is why a less than wild-type level of gene activity causes a defect. In diploid organisms, haploinsufficiency tends to be the exception. If the relationship of gene activity to phenotype were strictly linear, an organism with a null allele in trans to a wild-type allele would have half the gene dosage compared to an organism that had two wild-type alleles and might be expected show a phenotype. Instead, most hypomorphic and loss-of-function alleles behave in a recessive fashion. This is likely because the cell has enough of a buffer of gene activity such that a phenotype is not manifested even when half the wild-type gene activity is present. The extreme dose-sensitivity that underlies cases of haploinsufficiency is typically caused by either subunit imbalance, a rate-limiting step in metabolism, or by changing the dosage of developmental regulators such as transcription factors.

An example of haploinsufficiency involves the *Pitx1* transcription factor. A study of the *Pitx1* gene in mouse revealed that about 9% of *Pitx1*(+/-) heterozygous mice have clubfoot. Either a missense mutation that lowers *PITX1* activity or a deletion of chromosome 5q31 that affects *PITX1* can cause this phenotype. Thus, a reduction-of-function or loss-of-function allele can behave in a dominant fashion. Interestingly, loss of one copy of *Pitx1* also results in clubfoot in humans, and 1/1000 live human births is affected by clubfoot.

Haploinsufficiency at a genome-wide level has mainly been assayed for the ability to support life of the organism. In *Drosophila*, where haploinsufficiency has been analyzed systematically, there are about 50 loci (out of ~15,000 genes) that are sensitive to reduction of one gene copy. The number of haploinsufficient loci in *C. elegans* is thought to be similarly few, based on the ability to recover heterozygous *Df* for much of genome, and the observations that most dominant mutations are gain-of-function rather than loss-of-function. Likewise, only 3% of *S. cerevisiae* genes are haploinsufficient for growth on rich medium. However, as we dig deeper into particular phenotypes that may be more dosage sensitive, we will likely uncover additional cases of haploinsufficiency.

Gain-of-function mutations can be recessive

Just like some loss-of-function alleles can behave in a haploinsufficient manner, sometimes a gain-of-function mutation can be recessive (Table 461). One example how this can happen involves a protein that acts within a protein complex. If protein works in a complex in a manner where having a wild-type protein in the complex is enough to compensate for the gain-of-function mutant allele, this gain-of-function allele would behave as a recessive allele. This mutation would be considered a gain-of-function allele because without the wild-type allele, it may lose its activity to be regulated to its “off” state, and thus has an activity greater than that seen with a wild-type allele.

For example, Figure 14 [SGF-1115] illustrates a case where the wild-type protein is a homodimeric protein with a regulatory domain and a catalytic domain. Normally, the presence of the regulatory domain restricts the activity of the catalytic domain until the appropriate time and place. The gain-of-function mutant lacks this regulatory domain, and is constitutively active. In the heterozygote, assuming that both types of protein can form a dimeric complex equally well and the dimers form randomly, only 25% of complexes will completely lack the regulatory domain. In the heterozygote, 75% of the complexes have a wild-type protein and the presence of a single regulatory domain allows for proper regulation of the catalytic activity. In this case, what you see is less than the normal amount of gene activity, and thus the gain-of-function mutation behaves as a recessive allele.

Mutations that cause gain-of-function can be constructed

Gain-of-function alleles can be useful for genetic analysis and can sometimes be created by molecular engineering of the gene product (Figure 15 [SGF-1120]). To increase gene activity, sometimes altering the transcriptional control of a gene by the addition of a strong enhancer or strong promoter can lead to increased amounts of gene product and thus the creation of a hypermorphic allele. Similarly, providing an organism with multiple copies of the wild-type gene can often increase the amount of wild-type gene product made, leading to a hypermorphic phenotype. Also, proteins that have deletions of regulatory domains can also have gain-of-function effects (Figure 16 [SGF-1121]).

Neomorphic alleles can be created by adding an inappropriate cellular localization signal to a protein or RNA, to cause the gene to now function in an inappropriate time or place. For example, if a product is localized to one cellular compartment, such as a neuronal synaptic compartment, then one might want to mis-localize that product to the dendritic compartment to determine its effect there. Because antimorphs can exert their effect by causing non-productive interactions with wild-type alleles, they are sometimes created to block wild-type gene activity. For example, the dominant negative protein may bind to a target but does not have catalytic activity, and thus prevents the wild-type protein from accessing the target and being active (Figure 17 [SGF-1122]). Antimorphs have been created using site-specific DNA-binding proteins such as transcriptional regulators, which often have separable DNA-binding and regulation domains. A protein with intact DNA-binding but no activation domain can prevent the wild-type protein from functioning.

Genes with opposite loss- and gain- of function phenotypes likely control biological processes

Having both loss-of-function (lf) and gain-of-function (gf) alleles can help define gene function, and provide strong genetic evidence that that gene plays a specifying role in a phenotype. The simplest cases and most common case is that loss-of-function alleles cause one phenotype while hypermorphic gain-of-function alleles (if they exist) cause the opposite phenotype. For example, the *C. elegans lin-12* gene, involved in anchor cell production, has two major classes of alleles: loss-of-function alleles with two anchor cells and gain of function alleles with no anchor cells (Table 406). Other similar examples include the *Antp* gene in the fruitfly and the *Lmbr1* gene in mouse. Loss of *Antp* activity leads to flies that have an antenna where there should be a leg, while gain off *Antp* function leads to leg where there should be an antenna. Loss of *Lmbr1* activity leads to mice with fewer numbers of digits while increased *Lmbr1* activity leads to extra digits. The general interpretation of these results is that the state of gene activity specifies the outcome. In some contexts, the gene specifies one fate or inhibits the alternative fate (Figure 18 [SGF-1124]), and thus acts as a component of a binary (two-state) switch.

Genes can have similar loss- and gain- of function phenotypes

Sometimes, loss- and gain- of-mutations can have the same phenotype. This is typically seen when a protein must cycle through two states, and both states are utilized by the organism such that the inability to cycle causes a phenotype. For example, in *S. cerevisiae*, the gain-of-function and loss-of-function phenotype for the *SEC4* gene is the same: a secretion defect. Yeast *SEC4* encodes a GTPase, a type of protein that exists in three states (Figure 19 [SGF-1110]). They exist in three states: empty, bound to GDP, and bound to GTP. The GTP-bound state is the active state. Because GTP concentrations are much higher within the cell compared to GDP concentrations, the empty protein will quickly bind GTP, and will bind GTP more often than GDP. Once a protein binds to the GTP state, its intrinsic GTPase activity releases Pi, leaving it GDP bound. The GDP then falls off the G protein. Because this cycle of GTP binding, Pi release, and GDP exit are important to the functioning of the protein, an allele locked in a GTP bound state behaves similarly to an allele that cannot bind GTP. The simplest

interpretation of this apparently paradoxical result is that it is crucial for SEC4 to complete its GTPase cycle in order for secretion to proceed.

4B. Complementation

Complementation analysis examines the phenotypic effect of having two or more alleles present in the same cell. In more concrete terms, these types of studies compare situations in which there are multiple forms of one gene product within the same cell. Inferences from the types of genetic interactions seen among these gene products can inform about gene function. Above, we have discussed the use of an allelic, which includes allelic combinations, to examine levels of gene activity. Here, we discuss other uses for making different allelic combinations.

Complementation tests are typically used to define functional units of genes and to help connect DNA to mutant alleles. Typically, two recessive alleles in a gene will fail to **complement** and display a mutant phenotype while two recessive alleles in two different genes will complement and display a wild-type phenotype. Sometimes, exceptions to typical patterns of complementation are seen. For example, two alleles of the same gene can complement, a phenomenon known as **intragenic complementation**. On the other hand, two alleles of a different gene may fail to complement, which is described as **non-allelic non-complementation** or **extragenic non-complementation**. These exceptions to typical complementation can be highly informative, particularly when studying genes encoding multidomain proteins or for studying genes that have multiple functions. **Non-allelic non-complementation** can be useful for genetic screens (Chapter 7) and for the study of gene-gene interactions (Chapter 5). In this section, we will discuss analysis of complementation and its exceptional cases.

Complementation tests define functional units

Complementation tests were originally used to determine whether two different recessive alleles affect the same functional unit, which is typically defined now as a gene. In these tests, recessive alleles were placed in *cis* (on the same chromosome) and in *trans* (on a different chromosome). If the two alleles displayed a wild-type (non-mutant) phenotype in *trans*, these two alleles were considered to be in different functional units (Figure 20 [SGF-1092]). When a wild-type phenotype is seen, we infer that the presence of each allele provided some function that complemented each other such that the mutant phenotype was not expressed. Thus, the simple interpretation of this result was that these two alleles are in different genes. However, if the two alleles in *trans* still displayed the mutant phenotype, these two alleles fail to complement, and the simple interpretation of this case is that these two alleles are not in the same gene. Because complementation tests were used so often to define a functional unit, geneticists sometimes formally refer to alleles in the same gene as alleles in the same **complementation group**. Alternatively, the functional unit of a gene is sometimes referred to as a **cistron**, a word derived from the *cis-trans* test used for complementation analysis.

Complementation in transgenic organisms can connect DNA to alleles

Besides defining genes, complementation tests can be used to define DNA segments related to the gene of interest. This type of complementation experiment is sometimes referred to a functional test, as this experiment tests the function of the DNA segment. Specifically, complementation analyses can be used to determine whether the correct fragment of DNA corresponding to a mutant has been isolated. This type of experiment involves the introduction of DNA into an organism displaying a mutant phenotype, to test whether it can complement the mutant allele by restoring a wild-type phenotype. The introduced DNA is sometimes referred to

as a **transgene**, and animals carrying exogenous DNA are sometimes referred to as **transgenic animals**.

For example, imagine a diploid organism with a recessive mutation that causes phenotype **A** (Table 430). When DNA containing gene *A* is introduced into this organism, it restores the phenotype to **non-A** (usually wild type) (Figure 21 [SGF-1197]). Because typically these experiments involve restoring the mutant phenotype to a wild-type phenotype, this restoration of phenotype is sometimes referred to as **rescue** of the mutant phenotype. As a control, a non-functional version of gene *A* (for example, a gene with a premature stop codon so that most of the protein cannot be produced) can be introduced into an organism, which should not have an effect on phenotype **A**. If the DNA containing the functional *A* gene was introduced into an organism that is wild type for *A* but mutant at another locus, *B*, there should be no effect on phenotype **B**. That DNA fragment would be said not to **complement** the phenotype, or, that DNA would not rescue the phenotype **B**.

Exceptions to complementation are highly informative

Sometimes, two alleles of the same gene can complement and display a non-mutant phenotype when placed in *trans*, a phenomenon called **intragenic complementation** (Figure 22A [SGF-1090]). Since these alleles complement, they do not pass the typical test for being part of the same genetic locus. These alleles instead are determined to be part of the same genetic locus by fine-structure mapping (i.e., intragenic recombination mapping), by DNA sequence that shows the alleles to be mutations within the same gene, or by the failure of these alleles to complement other alleles of the same gene. Typically, intragenic complementation is seen when the two alleles produce gene products that interact, or, when the normal gene product has multiple independent functions. Molecular examples and genetic interpretations of intragenic complementation are discussed below.

Similarly, if you put two alleles in *trans* and see a mutant phenotype, sometimes this can be due to two alleles in different genes that fail to complement (Figure 22B [SGF-1090]). These cases are referred to as **extragenic non-complementation** or **non-allelic non-complementation**. Extragenic non-complementation can be distinguished from intragenic non-complementation by mapping, as alleles from a single gene should be tightly linked while alleles in different genes should not display tight linkage. Extragenic non-complementation is suggestive of a functional relationship between the genes involved and can imply a physical relationship between the two gene products. These types of genetic interactions and their interpretations are described further in Chapters 5 and 6 and are utilized during non-complementation screens described in Chapter 7.

Distinct functional domains of multidomain proteins can be identified by intragenic complementation

When intragenic complementation is seen, the each allele must be providing some gene activity that supplements that activity of the other allele. Intragenic complementation can occur when a gene encodes a product with independent functions, each necessary in the organism. Loss of a single function will result in a distinct phenotype from loss of a different function. Alternatively, a mutation that disrupts a protein can be compensated for by the right type of mutation in another allele. These cases are discussed below.

If a gene has two independent functions, each necessary for the wild-type phenotypes, a complicated pattern of intragenic complementation may be seen when using non-null alleles. A gene may have independent functions because the gene encodes a protein with multiple domains with distinct activities, or because distinct binding surfaces on the gene product interact with distinct binding partners. Because these different functions are encoded by distinct regions on the gene product, it may be possible to obtain mutant alleles that disrupt one function while keeping other functions of the protein intact.

Consider a gene *R* that has two functional domains (Table 434 and Figure 23 [SGF-1126]). Allele *r*₁ lacks a function that causes phenotype **R1** but not **R2**. Similarly, allele *r*₂ lacks a function that causes phenotype **R2** but not **R1**. Because the null allele lacks both domains, it displays both the **R1** and **R2** phenotypes. In this case, neither allele complements the null allele. However, the *r*₁/*r*₂ combination leads to intragenic complementation for phenotypes **R1** and **R2** because each allele brings a functional domain that is enough to confer wild-type activity.

This type of intragenic complementation was seen with the epidermal growth factor (EGF) receptor gene in *Drosophila*, which encodes a protein with two major domains separated by a single-span transmembrane domain. The extracellular domain binds a ligand (the signal), while an intracellular domain contains a kinase domain important for conveying the signal within the cell. When a EGF receptor gene containing a mutation within the ligand binding domain is placed in *trans* to an allele with a mutation within the kinase domain, intragenic complementation was seen. Either of these alleles in *trans* to a null allele does not complement.

Distinct functional regions of a single domain protein can be identified by intragenic complementation

Similarly, a single domain protein that contains distinct interaction surfaces on its surface can also display intragenic complementation (Figure 24 [SGF-1127]). The protein calmodulin is the best example of this type of intragenic complementation. When bound to particular interaction partners, calmodulin will perform a specific function. As long as the cell contains calmodulin proteins that can interact with its different partners, calmodulin can perform its duties. For example, consider two alleles of calmodulin, one that cannot bind to partner A (allele *s*₁) while the other cannot bind to partner B (allele *s*₂). Homozygous organisms that are *s*₁/*s*₁ will display phenotype **A** while homozygous organisms that are *s*₂/*s*₂ will display phenotype **B**. However, *s*₁/*s*₂ organisms display a non-**A** non-**B**, or wild-type, phenotype.

Interactions within a protein can be identified by intragenic complementation

Sometimes, compensatory mutations in different alleles can cause intragenic complementation. A simple example of a compensatory mutation involves a protein that normally works as a homodimer. In this case, salt-bridges in the protein (which are electrostatic interactions between oppositely charged side-chains) mediate the dimerization of the protein (Figure 25 1118 [SGF-1118]). Mutations of a positively charged amino acid to a negative charged amino acid will prevent the salt-bridge from forming, leading to a lack of dimerization (and often times, a lack of proper folding). Similarly, mutation of the negatively charged amino acid to a positively charged amino acid can also disrupt salt-bridge formation. However, if a positive amino acid is changed to a negative amino acid in one allele while the other allele contains the complementary negative amino acid changed into a positive amino acid, salt bridge formation can once again form, resulting in a functional dimer and a wild-type phenotype.

Different temporal and spatial requirements of a gene can lead to intragenic complementation

Sometimes genes have different functions because of a requirement for the protein at different times in an organism's life. For example, a gene product may be required at different times and places during the development of the organism. This differential expression is typically mediated by differential gene regulation. These regulatory elements can sometimes be independently mutated to create alleles that only affect a particular temporal or spatial expression of the gene. In this case, these alleles can lead to intragenic complementation, because when these types of alleles are placed in *trans* to each other, the gene activity should now be present at all the proper times and places needed (Table 436).

4C. Site and time of gene action

Site and time of gene action experiments will help distinguish between the place and time at which a gene **functions** in a particular process from where and when it is **expressed** (Figure 26 [SGF-1143]). This information is crucial, as gene can be expressed in places in which its function is irrelevant to the phenotype of interest. Alternatively, a gene can be expressed in one cell yet exert an effect on the phenotype of another cell. This could occur, for example, if a gene encodes a diffusible signaling protein whose role is to act on cells that do not express that gene. A gene might also be expressed over a long time period, but only be required at a particular point in time. Molecular experiments (such as visualizing the mRNA by *in situ* hybridization experiments) can tell you where and when a gene is expressed, but site of action experiments and time of action experiments are needed to find out where and when a gene's activity is required. Genes may also function in places or times where their expression cannot be detected. For genes expressed at very low levels, genetic assays of gene activity can be more sensitive than the ability to experimentally detect gene expression.

Site of action experiments utilize genetic manipulations to create situations where a gene is expressed only in a subset of cells or tissues within the organism, to determine where a gene acts (Figure 27 [SGF-1140]). Time of action experiments manipulate gene expression in a temporal fashion, to allow the dissection of when the gene activity is needed. Ultimately, these types of experiments provide complementary information to molecular experiments detailing mRNA and protein expression.

Site of action experiments indicate where a gene acts

Site of action experiment create genetic mosaics in which some cells are of one genotype and others are of a different genotype. Analyses of these genetically mosaic organisms allow for the inference of whether a gene acts **autonomously** within the cell in which it is expressed, or **non-autonomously**, on cells which it is not expressed (Figure 28 [SGF-1096] and Tables 428 and 429).

Consider the simple case of determining the site of action of genes *B* and *C*, with the wild-type alleles (b_+ and c_+) and the mutant alleles (b_1 and c_1). In this organism, there are two cells: type 1, that can either be round or square, and type 2, which is a triangle shaped cell that influences the shape of the type 1 cells. Genes *B* and *C* are important for the shape of type 1 cells, as animals that have the b_+ and c_+ allele have type 1 cells that are round while animals that are have the b_1 and c_1 allele have type 1 cells that are square.

If gene *B* acts cell autonomously, the genotype of the type 1 cells will reflect their phenotype. In this example, if the type 1 cells have the b_+ allele, they will be round. If they have the b_1 allele, they will be square. It does not matter which *B* allele the triangle shaped type 2 cells contains. Furthermore, if different type 1 cells have different *B* alleles, the individual cell's phenotype will reflect its specific genotype, resulting in a mixture of type 1 cells that are square and round.

On the other hand, if gene *C* acts cell non-autonomously in the type 2 cell to determine the shape of the type 1 cell, the phenotype of the type 1 cell will depend on the genotype of the type 2 cell. In this case, a genotypically c_+ type 1 cell next to a type 2 cell is c_1 will be square. Since gene *C* acts non-autonomously, the genotype of the type 1 cell is irrelevant. On the other hand, if the type 2 cell is c_+ , the type 1 cell will be round no matter the genotype of the type 1 cell, which could be either c_+ or c_1 . The effect of gene *C* is seen in the type 1 cells, and affects their cell shape. In addition, the type 2 cell is a triangle, regardless of the *C* allele it carries, as gene *C* has no effect on the shape of the type 2 cell.

Site of action experiments work even in complex tissues

Let's now consider the case where there are three type 1 cells (Figure 29 [SGF-1101]). If the organism is wild type ($b_+/b_+ c_+/c_+$), all the type 1 cells are circles. In either a b_1/b_1 or a c_1/c_1 mutant, the type 1 cells are squares. In mosaic organisms, if the triangular type 2 cell has the b_1/b_1 genotype, there is no effect on the type 1 cells since b_1 is autonomous. On the other hand, if the type 2 cell has the c_1/c_1 genotype, all type 1 cells are square since c_1 is non-autonomous. However, if only one type 1 cell is b_1/b_1 in an otherwise wild-type background, the b_1/b_1 type 1 cell will be a square (Figure 29A [SGF-1101]). If the genotypes of the type 1 cells are heterogeneous for gene C, they will all be circle or square depending on the genotype of type 2 cell (Figure 29B [SGF-1101]). This is the case even when one of the type one cells is c_1/c_1 .

How do genes behave in a non-autonomous fashion? Sometimes it is because the non-autonomous gene acts in a cell that will influence nearby or adjacent cells. These effects would result from cell contact, or molecular signaling that acts only at close proximity. On the other hand, a non-autonomous effect can affect the entire organism, acting in a global fashion. Such global effects can result from genes affecting endocrine systems, where the gene product may get sent throughout the entire body. Similarly, genes that act in the nervous system can affect the entire organism. For example, neuronal-specific genes can affect lifespan, a whole-organism phenotype.

It is important to note that although genetic mosaic experiments can be carried out with loss-function non-null alleles, a negative result can be tricky to interpret, because it is difficult to assess whether the residual activity present was enough in the mutant cells to confer the wild-type phenotype. Thus, if null alleles are available, it is preferable to do the experiment with those alleles.

Non-autonomous genes acting in multiple cells may be subject to majority rules

For genes that act non-autonomously from more than one cell, the numbers of cells can have an effect on site-of-action experiments. This situation typically occurs when there is a threshold level of gene product necessary from the non-autonomous cell to affect the phenotype of the other cells. In these cases, the number of cells that have a particular allele will determine the phenotype, with the majority genotype determining the phenotype (Figure 30 [SGF-1102]). This interpretation is suggested when the presence of mosaic organisms in which wild-type cells are effected by mutant cells (Figure 30C [SGF-1102]) along with other mosaic organisms in which mutant cells are affected by wild-type cells (Figure 30D [SGF-1102]). For example, this phenomenon has been proposed in the determination of sexual phenotype in *C. elegans*, where mosaics containing a mixture of male and female cells that have a consistent whole-organism sexual phenotype in *C. elegans*.

Relative, and not absolute, levels of gene activity can determine outcome.

Sometimes, the relative level of gene activity (not absolute levels) is important in determining the phenotype. For some genes, this level is measured relative to adjacent cells. Figure 31 [SGF-1103] illustrates this situation. Precursor cells (pentagons) can differentiate into either circles or squares. In h_1/h_1 loss-of-function mutants, precursor cells become circles. In mosaic animals, the pentagon with the least activity will become a circle. Consider a pentagon with two copies of the h_+ allele. If it is near an h_1/h_+ pentagon it will become a square. By contrast, if the h_+/h_+ pentagon is next to an $h_+/h_+/h_+$ cell, it will instead become a circle. The same genotype cell will have two different outcomes depending on the relative level of the *H* gene!

The *C. elegans lin-12* gene and the *D. melanogaster Notch* gene are two examples of genes that behave this way. Both these genes encode cell surface receptors that were discovered for their role in lateral inhibition during development. These genes encode similar proteins, and mosaic analysis showed that for two adjacent cells, the one expressing the higher amount of gene product relative to its neighbor would adopt a particular fate. Specifically, a wild-

type cell next to a null mutant would result in the wild-type cell adopting a particular cell fate, **A**. On the other hand, a wild-type cell next to a cell expressing multiple copies of the wild-type gene would result in the overexpressing cell adopting the **A** cell fate.

There are multiple methods for carrying out site of action experiments

How might a genotypically mosaic organism be created for site of action experiments? Depending on the tools available in the organism being studied and the specific problem being studied, different methods are appropriate. Sometimes, a physical mixing or transplantation of cells can be used to create a genetic mosaic. In organisms that are more genetically manipulable, the targeted inactivation or activation of a specific gene within a subset of cells can also generate mosaic organisms. Classically, chromosome manipulations such as mitotic recombination or the loss of an unstable chromosome (or chromosome part) were used.

Physical manipulations can be used to create genetic mosaics

Sometimes, site of action experiments can be carried out by simply mixing tissues or cells of distinct genotypes (Figures 32 [SGF-1129] and 33 [SGF-1159]). For example, in slugs of the social amoeba (slime mold) *Dictyostelium discoideum*, cells spontaneously aggregate to form a slug that will develop into a fruiting body. If cells of two different genotypes are mixed, a genetically mosaic slug would form, and the autonomy of genes involved in fruiting body formation could then be assayed.

Similarly, mosaic mammalian embryos can be created by mixing mammalian blastocyst cells of different genotypes. As the embryo of mixed genotype develops, the autonomy of genes can be assessed for various cells or tissues. Alternatively, tissue of a particular genotype can be transplanted into an organism of another genotype to create a mosaic organism. The ability to use cell or tissue mixing to examine the site of action depends on the ability to mark, or recognize, which cells or tissues have a particular genotype. One way this can be carried out is by having a relatively innocuous marker such as green fluorescent protein (BOX 320) selectively expressed in cells of one genotype, allowing for their distinction from the other genotype.

Knockdown of individual genes can be manipulated to create genetic mosaics

Sometimes engineered genes can be used to create genetic mosaics. When genes are engineered such that they are only activated or only inactivated in a particular cell (or tissue), an organism can be produced that is a mixture of wild-type and mutant cells (Figure 34 [SGF-1134]). Typically, the types of alleles used in these experiments are recessive to the wild-type allele.

A typical method for turning off genes only in a specific cell or tissue involves the use of RNAi expressed under a promoter that causes its expression in the place of interest (see BOX 309). Using this method, a mosaic animal can be created that is mostly wild type but has a set of cells or tissues that now have less than wild-type levels of the gene of interest (Figure 35 [SGF-1131]). Another approach to this type of experiment is if the promoter for the gene of interest is well understood, it is possible to remove a specific regulatory element controlling expression in a specific set of cells, to create an organism lacking the gene in that particular set of cells.

For example, a cell-type 1-specific promoter can be used to drive the expression of an RNAi construct targeting the gene of interest, *U*. Knock down of *U* only in cell-type 1 results in a phenotype, **U**. By contrast, knock down of *U* only in cell-type 2 has no phenotypic consequence; similarly, knock down of *U* only in cell type-3 also has no phenotypic consequence. Thus, gene *U* is required in cell type 1.

Note that a negative result in an RNAi experiment (where no effect is seen when knockdown is attempted) is not informative, as it is difficult to assess the extent to which gene activity has been reduced using this methodology, as RNAi allele are typically hypomorphic and not nulls. It is for this reason, that the positive control is crucial: knockdown of gene *U* in all cells should cause a phenotype, suggesting that the knockdown is indeed reducing gene activity.

Expression of individual genes can be manipulated to create genetic mosaics

The forced expression of a gene in an otherwise mutant background can also be informative. This experiment can be carried out by expressing a wild-type copy of a gene under the control of a tissue- or cell-type-specific promoter in an otherwise mutant animal. For example, a muscle-specific promoter can be used to drive the expression of the wild-type gene of interest in the muscle in an otherwise mutant background. Using these animals, you can then assess whether the expression of that gene in the muscle will affect the phenotype of the organism. Sometime instead of using transcriptional regulation to control gene expression, viruses can be used to deliver genes to a subset of cells within an organism. In these cases, the virally-encoded gene can be targeted to specific cells or tissues by including a tissue-specific transcriptional control element in the virus. In this way, a cell will only express the gene if it is of the right type of cell and is infected by the virus.

An example is given in **Figure 36 [SGF-1116]**, where gene *F* is expressed in either cell 1, 2 or 3, to determine its requirement. In this example, expression in cell 3 is sufficient to restore the phenotype of the cells to normal.

Sometimes, the site of action can be assayed by examining animals where a gain of function allele is ectopically expressed in specific cells or tissues in an otherwise wild-type background (**Figure 37 [SGF-1130]**). This type of analysis is important in organisms where loss-of-function genetics is difficult, or in cases where apparent functional redundancy prevents you from observing a loss-of-function phenotype. It is conceptually the same as expressing a wild-type gene in an otherwise recessive mutant background.

Experiments involving the forced expression of a gene in a particular tissue allows you to determine whether expression of that gene in those cells is sufficient, but, it does not tell you whether that gene is normally necessary in those cells. For example, if the gene was required in a subset of muscle cells and a general muscle-specific promoter was used to examine site of action, rescue of the mutant phenotype could be seen even though the gene was expressed in more cells than needed. For a secreted factor, the ectopic expression in a non-normal tissue may lead to rescue, even if that factor is not normally expressed there. Thus, these experiments provide complementary information to those experiments where gene activity is reduced within particular cells in an otherwise wild-type background.

More efficient transcriptional control can be accomplished by bipartite transcriptional activators that allow independent construction of the transcriptional control and target genes so that they can be recombined by crossing. The GAL4-UAS system is the most popular of these systems (**Figure 37.5 A [SGF-1137]**) but others are in use such as the *q* system. Rather than place gene *S* directly under the control of an enhancer of choice, the mRNA encoding the yeast transcriptional activator GAL4 is produced from that enhancer. The GAL4 protein then binds its target, the UAS_{GAL4} sequence, that has been engineered to control expression of gene *S*. If this is done once it is extra work, but the same GAL4 “driver” can control expression of many target genes, *S_i*, each of which is engineered to be under control of a UAS_{GAL4} sequence (**Figure 37.5 B [SGF-1137]**).

Regions of chromosomes can be manipulated to create genetic mosaics

Instead of manipulating the expression of a single gene using promoters or viruses, the manipulation of entire regions of chromosomes can create genetic mosaics. These types of techniques include using loss of an extrachromosomal genetic element within specific cells to create a mosaic or the induction mitotic recombination in specific cells to create mosaics.

In some organisms, genetic material can be carried on an extrachromosomal element. These extrachromosomal elements include small plasmids, extrachromosomal arrays of genes, and unstable chromosomal *Dps* (duplications). These extrachromosomal elements tend not to be maintained faithfully during cell division (unlike chromosomes) and thus are stochastically

lost. When these extrachromosomal elements are lost, genetic mosaics can be created. This specific method is most useful if there is some way to know which cells have lost the genetic element. Sometimes these extrachromosomal elements can be marked with a cell-autonomous ubiquitously expressed dominant marker, such as green fluorescent protein, so cells lacking the extrachromosomal element can be easily identified.

For example, if the site of action of a recessive allele, d_1 , is being investigated, the organism can have a wild-type copy of gene D (allele d_+) on the unstable genetic element (Figure 38 [SGF-1133]). As these elements are unstable, they can be stochastically lost. Because this element also contains a GFP marker, cells that have lost this element are not green, while cells that maintain the element are green. When this loss occurs, a genetic mosaic has been created. To find a loss in a particular cell or tissue requires the screening for loss of the marker in the cell or tissue of interest, and then assaying the phenotype of these cells.

In diploid organisms, mitotic recombination can also be used to generate cells of distinct genotypes (Figure 39 [SGF-1132]). Classically, mitotic recombination was induced in a heterozygous organism by the creation of low level DNA damage (typically by double stranded breaks). Mitotic recombination occurs as cells try to repair the damaged DNA using information from the other strand, leading to mitotic crossing over and the creation of sister cells that are now homozygous at the gene of interest (Figure 40 [SGF-1093]). Induction by DNA damage creates recombination at stochastic sites, and thus organisms must be screened for the appropriate recombination event in the cells of interest.

More efficient genetic mosaics can be carried out using site-specific recombination

A more efficient way of doing mosaic analysis by mitotic recombination is to use site-specific recombination (Figure 41 [SGF-1097]). In general, site-specific recombination requires addition (by engineering) of sites to the genome that can be recognized by an ectopically expressed recombinase, along with some kind of marker so that cells where recombination has occurred can be recognized. The controlled expression of the recombinase in specific cells or tissues results in recombination, and thus stimulation of mitotic recombination, in those cells.

Several systems have been developed to create genetic mosaics using site-specific recombination. In *Drosophila*, the FLP-FRT system is commonly used. This system utilizes the yeast FLP recombinase protein that stimulates recombination at specific FRT (Flip recombination target) sequences. By introduction of FRT sites at specific places within the genome, recombination at these sites can be stimulated by expression of the FLP recombinase (Figure 42 [SGF-1095]). For example, if a specific promoter driving expression in a subset of cells was used, recombination within those cells could be stimulated, resulting in the creation of cells now heat shock promoter can be used to drive expression of the recombinase; if flies are then subjected to a low level of heat shock, the stochastic expression of FLP can then cause the creation of mitotic recombinants that can then be screened through for animals that have created homozygous mutant cells in the appropriate places. Collections of chromosomes with FRTs integrated to allow mitotic recombination on specific arms of chromosomes of *Drosophila* exist, and are for mosaic analysis. In mice, the *cre-lox* system has been used for similar applications, where the *cre* recombinase is used to target the *lox* sites in the genome (Figure 43 [SGF-1098]). Or, genes can be activated using these same techniques (Figure 44 [SGF-1135]).

Intersectional methods can be valuable for determining the site of action of a gene

In practice, the molecular genetic reagents sometimes do not have sufficient specificity to uniquely determine the site of gene action. Trying to achieve tissue- or cell-type specific expression of an RNAi construct, a wild-type gene, or a site-specific recombinase is not always trivial. Transcriptional control elements do not often restrict expression to single cell types. In addition, the expression patterns of many genes or regulatory elements are not known exhaustively, and it is only when many people use the reagents that new expression patterns

are discovered. These caveats must be kept in mind when interpreting site of action experiments.

In practice, different methodologies can be used to build consistency arguments for a gene's site of action. Sometimes, promoters that have an expression pattern on in the cell or tissue of interest are not available. In these case, using several cell-type specific promoters with overlapping expression patterns can give you information whether the gene of interest is required in a particular expression domain. For example, if you were interested in examining the site of action of a specific gene in cell 6. However, there are no known promoters that allow for expression only in cell 6. Thus, you use a set of promoters with overlapping expression (Figure 45 [SGF-1195]). You could use a promoter that expressed gene *W* in cells 6, 7, 8 and 9, in cells 6 and 9, in cells 6, 7, and 8, and in cells 7 and 8. From these studies, you can try to infer what the outcome would be if you were to only express the gene in cell 6. In this example, you infer that expression in cell type 6 is sufficient to confer a non-*W* phenotype.

A more direct intersectional strategy is to have two constructs with different patterns that only overlap in one cell (Figure 45.3 [SGF-1138]). One way this can be accomplished is to use a split GAL4 transcriptional activator (Figure 45.7 [SGF-1138]) where the parts necessary to achieve expression of gene *W* overlap only in cell type 6 (Figure 46 [SGF-1196]). In particular, suppose one construct expresses in cells 6 and 7, while the other expresses in 7 and 9. Then, the only place both constructs are expressed is cell 6.

Time of action experiments are used to determine when a gene is needed.

Time of action experiments are carried out to determine when during an organism's lifespan a particular gene is needed. These studies utilize situations where the gene is on at one time and off at another time. By inactivating or activating a gene at a particular time, you can determine a critical time period at which the gene activity is required. These experiments are carried out by examining the phenotype of an organism where the gene is only active at particular times.

One method for carrying out time of action experiments involves the use of temperature-sensitive alleles. By shifting an organism at a particular time, the gene of interest can be inactivated and the phenotype assayed. Let us again consider the type 1 cells that can form either round or square shapes, and the type 2 triangle-shaped cell that can influence this decision. From the previous example (see Figure 28 [SGF-1096] above) we saw that gene *B* acts in a cell autonomous fashion in the type 1 cells while gene *C* acts in a cell non-autonomous fashion in type 2 cells. How might we determine when these genes act? If the type 1 and type 2 cells form at a particular time in the development of an organism, we can carry out time of action experiments using allele b_2 and c_2 , which are temperature sensitive alleles. In this example, type 1 cells are formed from pentagonal type 1 precursor cells. These pentagonal type 1 precursor cells will become either round or square 6 hours after they are formed (Figure 47 [SGF-1128]). The triangle shaped type 2 cell is formed at the same time that the pentagonal type 1 precursor cells are formed. Organisms carrying either allele b_2 or c_2 have round cells at the permissive 30°C temperature, while organisms carrying alleles allele b_2 or c_2 have square cells at the restrictive 15°C temperature. With these alleles, at 15°C, the gene activity of *B* and *C* is considered inactive in the cold.

To determine the time of action for the *B* gene, first consider an experiment using organisms that are genotypically b_2 . By shifting the organism from the restrictive to the permissive temperature and vice versa, we can pinpoint the critical time at which the activity of gene *B* is needed. Consider a case where these temperature shifts yield the data shown in (Figure 48 [SGF-1144]). If we examine experiment number 2 in this table, we see that shifting cells from the permissive temperature to the restrictive temperature at 5 hours leads to square cells. On the other hand, any experiment where cells are at the permissive temperature during the 4 and 5 hour time window (experiments 1, 7, 8, and 9) yield round cells. Consistently, experiments

that have cells at the restrictive temperature during the 4 and 5 hour time window (2, 3, 4, 5, and 6) seem to yield square cells. Thus, we can conclude that the activity of gene *B* is needed during this 4 and 5 hour time window.

Now, consider a similar time of action experiment using organisms that are genotypically *c₂*. Data from this experiment is presented in [Figure 49 \[SGF-1145\]](#). In this case, cells are square any time gene *C* is inactivated (at the restrictive temperature) during the third hour. Gene *C* does not appear to be required before the third hour (experiments 17 and 18), nor is it needed after the third hour (experiments 12 and 13).

Genes can have multiple times of action

Of course, with these types of experiments, a wild-type normal control organism must also be grown at the various temperatures, to ensure that the process being examined is not inherently temperature sensitive. If this were the case, these types of temperature experiments could not be carried out as simply as described.

The type of experiment described above is sometimes called a **reciprocal shift** experiment. The data for reciprocal shift experiments are often plotted as the penetrance of the phenotype on the y-axis against the time of a shift from permissive to restrictive or restrictive to permissive ([Figure 50 \[SGF-1106\]](#)). By visualizing the data in this manner, a critical period for when the gene is needed can be more easily seen. In some cases there are multiple critical periods for distinct phenotypes that are revealed by reciprocal shifts. An example of how these data might look like is shown in ([Figure 51 \[SGF-1111\]](#)).

Genes can be inactivated for time of action experiments using different types of perturbations

Just as with site of action experiments, there are multiple methods for carrying out type of action experiments. Classically, these experiments have been done with temperature sensitive alleles, as described in the example above. The creation of temperature sensitive alleles by mutagenesis is described in Chapter 7. Alternatively, molecular genetic tools can be used to inactivate or activate genes at particular times. If a promoter that acts in a temporal fashion is known to exist, that promoter can be used to either drive the expression of a gene (and thus provide gene activity) or inactivate the gene (for example using RNAi) at a specific point in time, similar to the experiments described for site-of-action determination.

Sometimes, promoter modifications can be made to allow for temporal control of gene expression. Specifically, the gene of interest can sometimes be put under the control of a drug-sensitive transcriptional activator. As addition of a drug can then either turn off or turn on a gene, the addition of this drug at a specific point in time can allow for temporal control. An example of this scheme with an inhibitory drug is shown in [Figure 52 \[SGF-1156\]](#). A specific example of this type of perturbation involves the use of the tetracycline transcriptional activator (tTA) and the doxycycline (Dox) drug. By placing tetracycline response elements in the promoter of the gene of interest, tTA can now regulate the expression of the gene in a doxycycline-responsive fashion. In the Tet-On system, the tTA can only bind to DNA when bound to doxycycline. In the Tet-OFF system, doxycycline-binding prevents binding of tTA to DNA.

Sometimes, engineering of the protein can be carried out to create a temperature sensitive allele. For example, temperature-sensitive protein stability can be engineered into a gene using a heat-inducible degradation signal (degron). A degron destabilizes a protein by targeting it for the ubiquitin-mediated proteasome degradation. A degron that is more active at high temperature has been identified and can be added to proteins to allow any protein to be rendered temperature-sensitive.

Time of action experiments have to be interpreted carefully, depending on the nature of the perturbation used to create temporal expression. By using a temperature sensitive allele, we are assuming that the gene activity is immediately eliminated at the restrictive temperature.

Sometimes this is not the case, as the gene product may perdure and remain active. Another issue is that the synthesis of a protein might be perturbed at the restrictive temperature, and thus using such a temperature-sensitive allele might reveal the time at which the protein is synthesized rather than when it functions.

Genes can be activated to define their time of action

In addition to turning genes off at specific times, genes can be activated to define their time of action. This allows reciprocal shifts to be carried out. In many organisms, heat-shock promoters can be used to turn on gene expression with reproducible kinetics (Figures 53 [SGF-1139] and 54 [SGF-1160]).

Example of genetic analysis: *C. elegans* vulva induction

C. elegans vulva induction was of interest to scientists as an example of organogenesis. Although geneticists were studying the *C. elegans* vulva for its interesting biological properties long before any direct connections to human health were known, we now know that many of the genes involved in regulating vulval development are critical components of signal transduction pathways important in human development and disease. The genetic analysis of the *C. elegans* vulva utilized many of the techniques discussed in this chapter.

During vulval development, three of six epidermal precursor cells, called vulval precursor cells (VPCs) are induced by the gonadal anchor cell (AC) to proliferate and generate specialized cells that form the mature vulva (Figure 55 [SGF-1149]). One reason this system was of particular interest was because it involved the development of a tissue in response to an extracellular signal. During the development of the vulva, the anchor cell produces a signal during the L3 larval period that directs three of the six vulval precursor cells to divide and produce vulval tissue after the L3 molt. The other cells that do not adopt vulval fates become epidermal cells. It was known that the anchor cell produced a signal important for the production of vulval cells, because laser ablation studies demonstrated that removal of the anchor cell led to a situation where none of the vulval-precursor cells differentiated into vulval tissue but instead, all vulval-precursor cells became epidermal cells (Figure 56 [SGF-1228]).

Genes involved in the development of the vulva were identified in genetic screens of the type that will be discussed in Chapter 7. The mutations either led to lack of vulva formation (vulvaless phenotype) or additional vulva formation (multivulva phenotype) (Figure 57 [SGF-1229]). Individual genes were analyzed to determine their normal role and whether they seemed to be particularly important. It was of particular interest to study genes not merely necessary for vulval development but also sufficient to direct aspects of vulval development. Thus, emphasis was placed on genes that fit the criterion such that a particular gene had hypermorphic alleles with phenotypes opposite to the loss-of-function phenotypes. Specifically, scientists were particularly interested in examining genes that behaved such that some mutant alleles produced a vulvaless mutant phenotype and other different alleles produced a multivulva mutant phenotype. By studying these alleles and determining the nature of these alleles, the genes *let-23*, *lin-3*, and *let-60* were found to have both reduction-of-function alleles that produced vulvaless animals and gain-of-function alleles that produced multivulval animals.

Cell ablation (by laser microbeam irradiation) experiments were performed to determine whether effect of the gene product came from the anchor cell or the vulva precursor cell. These experiments were particularly informative when performed on the gain-of-function alleles that conferred a multivulva phenotype. When the anchor cell was ablated in the *let-23* and *let-60* gain-of-function mutants, a multivulva phenotype was still seen, suggesting that the anchor cell signal was not required for this phenotype. This result was consistent with these two genes acting in the vulva precursor cells to receive the signal. On the other hand, when the anchor cell was ablated in the *lin-3* gain-of-function mutant, the animals were vulvaless, suggesting that

the *lin-3* gene requires the anchor cell to function, consistent with a role in the production of the anchor cell signal (Figure 58 [SGF-1230]).

Ultimately, the molecular identity of these genes was elucidated by the mapping and cloning, using techniques similar to those described in Chapter 9. The molecular identities of the proteins demonstrated that *lin-3* encodes a diffusible ligand similar to epidermal growth factor (EGF) and *let-23* encodes *C. elegans* EGF-receptor responsible for receiving the *lin-3* signal. *let-60* encodes a RAS protein, similar to that mutated in human cancers, and important for transducing the cellular response to the signal once it was received by the *let-23* receptor. These properties had previously been inferred by careful assessment of individual gene function by studying the mutant alleles in these genes.

Essential Concepts

- Gene function can be inferred from alleles that alter gene activity.
- Alleles of genes can have different activities.
- The phenotype of a complete loss of function allele indicates the normal function of a gene, that is, the processes for which the gene is necessary.
- Partial reduction of function alleles can reveal functions that are masked by the loss-of-function phenotype.
- A series of reduction of function alleles (an allelic series) can imply the loss-of-function phenotype.
- Gain of function alleles that increase gene activity can reveal the processes for which a gene is sufficient, all other things being equal.
- Dominant negative (antimorphic alleles) interfere with normal gene function and suggest gene function.
- Complementation tests among alleles define functional units that typically correspond to genes.
- Exceptions to expected results of complementation tests are informative, and reveal physical or genetic interactions among parts of genes or between genes.
- Site of action experiments indicate where a gene acts to affect a particular phenotype.
- Site of action experiments are carried out by construction or identification of mosaic organisms that have distinct genotypes in different cells.
- Time of action experiments indicate when a gene acts to affect a particular phenotype.
- There are many ways to construct genetic mosaics.
- Site-specific recombination systems allow efficient production of mosaic organisms with defined spatial and temporal alterations in gene activity.
- Intersectional gene expression strategies allow finer grained control.

Reading

Veitia RA. Dominant negative factors in health and disease. J Pathol. 2009 Aug;218(4):409-18. Review. PubMed PMID: 19544283.

Herskowitz I. Functional inactivation of genes by dominant negative mutations. Nature. 1987 Sep 17-23;329(6136):219-22. Review. PubMed PMID: 2442619.

Yook K. Complementation. WormBook. 2005 Oct 6:1-17. Review. PubMed PMID: 18023121.

Muller, H. J. in *Proceedings of the Sixth International Congress of Genetics* (ed. Jones, D.F.) 213–255 (Brooklyn Botanic Gardens, Menasha, Wisconsin, 1932).

Lindsley DL, Sandler L, Baker BS, Carpenter AT, Denell RE, Hall JC, Jacobs PA, Miklos GL, Davis BK, Gethmann RC, Hardy RW, Steven AH, Miller M, Nozawa H, Parry DM, Gould-Somero M, Gould-Somero M. Segmental aneuploidy and the genetic gross structure of the Drosophila genome. Genetics. 1972 May;71(1):157-84. PubMed PMID: 4624779; PubMed Central PMCID: PMC1212769. *Systematically analyzes haploinsufficiency in Drosophila.*

Deutschbauer AM, Jaramillo DF, Proctor M, Kumm J, Hillenmeyer ME, Davis RW, Nislow C, Giaever G. Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast. Genetics. 2005 Apr;169(4):1915-25. Epub 2005 Feb 16. PubMed PMID: 15716499; PubMed Central PMCID: PMC1449596.

Wilkie AO. The molecular basis of genetic dominance. J Med Genet. 1994 Feb;31(2):89-98. Review. PubMed PMID: 8182727; PubMed Central PMCID: PMC1049666.

Kacser, H., and Burns, J.A. (1981). The molecular basis of dominance. Genetics 97, 639–666.

References for examples

Nurse P, Thuriaux P. Regulatory genes controlling mitosis in the fission yeast Schizosaccharomyces pombe. Genetics. 1980 Nov;96(3):627-37. PubMed PMID: 7262540; PubMed Central PMCID: PMC1214365.

Struhl G. A homoeotic mutation transforming leg to antenna in Drosophila. Nature. 1981 Aug 13;292(5824):635-8. PubMed PMID: 7254358.

Greenwald IS, Sternberg PW, Horvitz HR. The lin-12 locus specifies cell fates in Caenorhabditis elegans. Cell. 1983 Sep;34(2):435-44. PubMed PMID: 6616618.

Hodgkin J. Two types of sex determination in a nematode. Nature. 1983 Jul 21-27;304(5923):267-8. PubMed PMID: 6866126.

Han M, Sternberg PW. Analysis of dominant-negative mutations of the Caenorhabditis elegans let-60 ras gene. Genes Dev. 1991 Dec;5(12A):2188-98. PubMed PMID: 1748278.

Bogdanove AJ, Voytas DF. TAL effectors: customizable proteins for DNA targeting. Science. 2011 Sep 30;333(6051):1843-6. Review. PubMed PMID: 21960622.

Ohya Y, Botstein D. Diverse essential functions revealed by complementing yeast calmodulin mutants. Science. 1994 Feb 18;263(5149):963-6. PubMed PMID: 8310294.

Wood AJ, Lo TW, Zeitler B, Pickle CS, Ralston EJ, Lee AH, Amora R, Miller JC, Leung E, Meng X, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Meyer BJ. Targeted genome editing across species using ZFNs and TALENs. *Science*. 2011 Jul 15;333(6040):307. Epub 2011 Jun 23.

Systems Genetics Chapter 4: Figures and Tables

Figure 1: [[add PHOTO OF Vulvaless on Background of WT]].

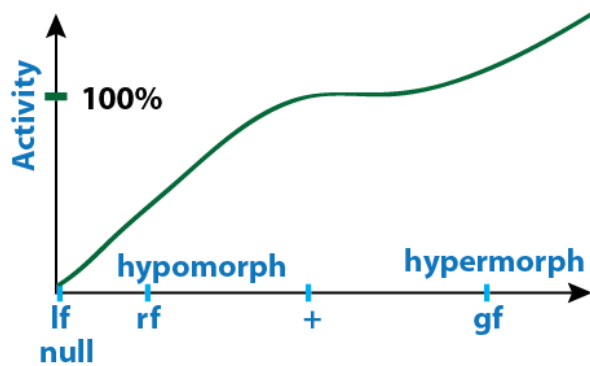


Figure 2 [SGF-1141]. General relationship between alleles and activity.

Table 4-1. Muller's classification of mutant alleles

classification	also known as	definition
wild type	reference allele	most common allele
amorph	null, loss-of-function	complete loss of gene activity
hypomorph	reduction-of-function	partial reduction of gene activity
hypermorph	gain-of-function	increase in normal gene activity
neomorph	gain-of-function	gene acquires novel function (may or many not retain normal function)
antimorph	dominant-negative, gain-of-function, poisonous product	antagonizes or interferes with wild-type gene function

Figure 3. Fly photos.

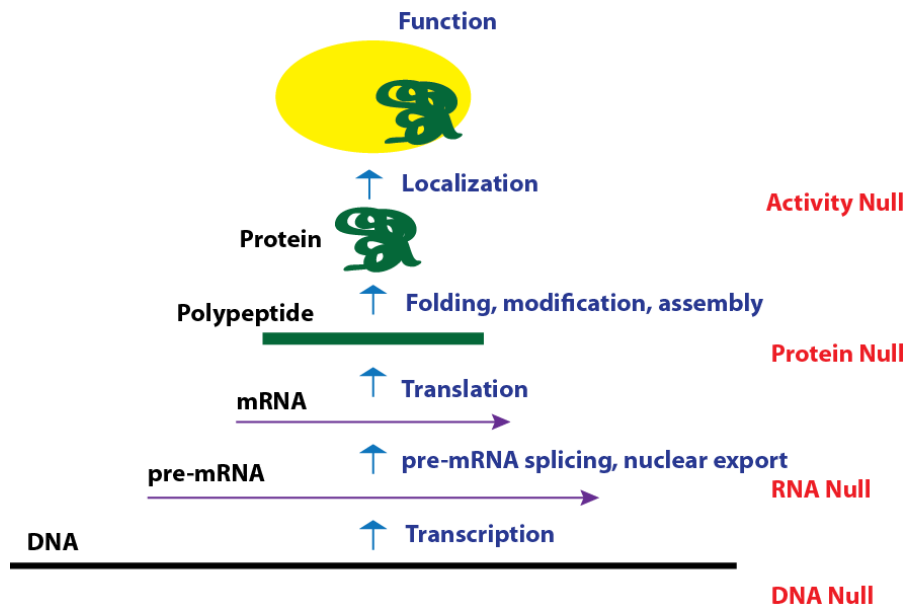


Figure 4 [Figure 1142]. Types of null alleles.

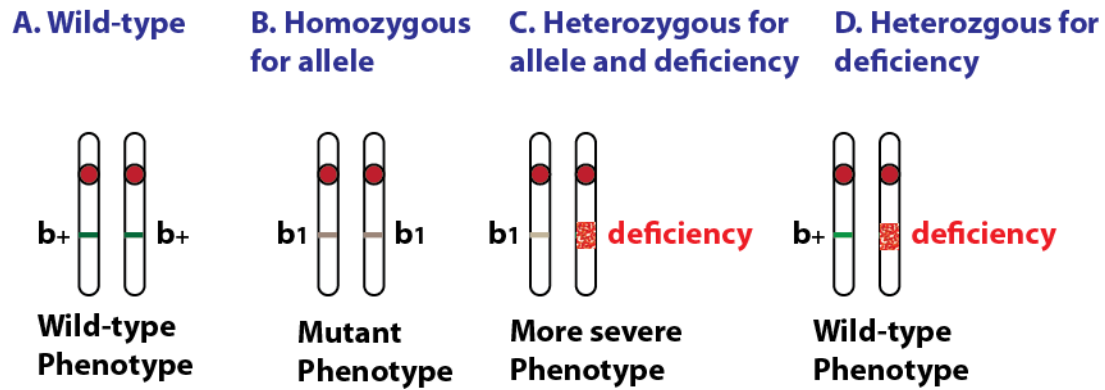


Figure 5 [SGF-1107]. Ruling out that an allele is null.

Table 4-2. RNAi test for complete loss of function

genotype	RNAi	phenotype	interpretation
c_+ / c_+	none	0% blue	wild-type control
c_1 / c_1	none	60% blue	mutant has a phenotype
c_+ / c_+	against C	30% blue	RNAi is partially effective
c_1 / c_1	against C	90% blue	c_1 is hypomorphic
c_2 / c_2	none	60% blue	control
c_2 / c_2	against C	60% blue	c_2 consistent with null

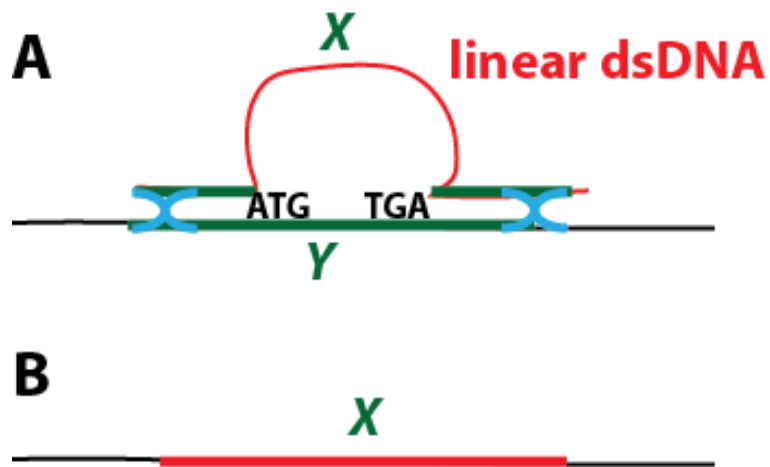


Figure 6 [SGF-1108]. Generalized knock-out strategy. [SGF-1108]. A. Chromosome, shown in black has targeted gene Y shown in green. Transformation with linear, double-stranded DNA that has homology to Y (green) and distinct sequences (red, X) results in (B) recombination that converts the chromosome to the X sequence (red).

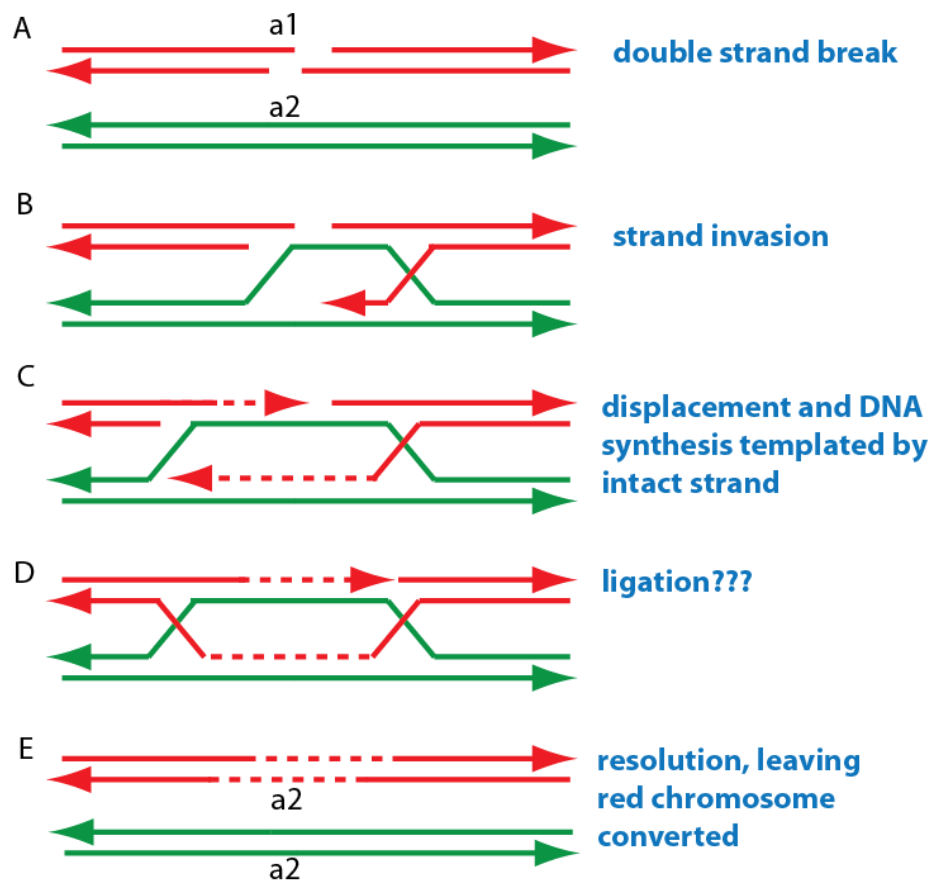


Figure 7 [SGF-1091]. Simplified model for double strand break repair. Two homologous chromosomes are shown in red and green. Dashed lines indicated new DNA replication. E. The repaired red chromosome now has allele *a2*.

Table 404. Opposite loss and gain of function phenotypes

S Genotype	Phenotype	Interpretation
+	+	wild type
lf	lack of X	S is necessary for X
gf	excessive X	S is sufficient for X

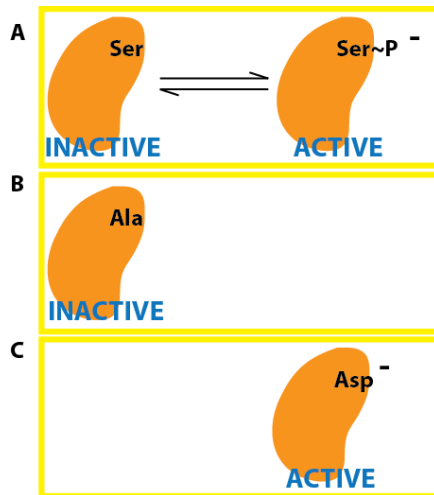


Figure 8 [SGF-1119] Regulation by phosphorylation. A. phosphorylation of a protein on a particular serine residue leads to its activation. Mutation of serine to alanine, a neutral amino acid, is inactive (lf). C. Mutation of serine to aspartic acid, a negatively charged amino acid, results in a constitutively active protein (gf).

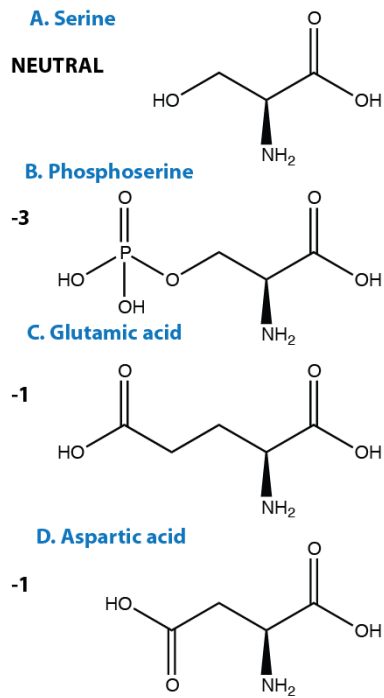


Figure 9 [SGF-1113] Glutamic and aspartic acid side chains mimic the negative charge of phosphoserine (or phosphothreonine).

A. wild-type human EGFR cDNA

cagttgggcacttttgaagatcattttctcagcctccagaggatgttcaataactgtgaggtggccttgggaatttgaaattaccta⁺gtgcagaggaattatgat
Q L G T F E D H F L S L Q R M F N N C E V V L G N L E I T Y V Q R N Y D

B. nucleotide substitution leading to a stop codon

cagttgggcacttttgaagatcattttctcagcctccagaggatgttcaataactgtgaggtggccttgggaatttgaaattaccta⁺agtgcagaggaattatgat
Q L G T F E D H F L S L Q R M F N N C E V V L G N L E I T - V Q R N Y D

C. insertion of a nucleotide leading to a frameshift and truncation

cagttgggcacttttgaagatcattttctcagcctccagaggatgttcaataactgtgaggtggccttgggaatttgaaattaccta⁺gtgcagaggaattatgat
Q L G H F - R S F S Q P P E D V Q - L - G G P W E F G N Y L C A E E L -

Figure 10 [SGF-1109]. Example of mutation to stop and to frameshift. Nucleotide sequence is shown in lower case; amino acid sequence in upper case below.

A. wild-type cDNA for human epidermal growth factor receptor. B. a nucleotide substitution leads to a stop codon (-). C. Insertion of a nucleotide leads to a reading frame shift and hence a premature stop codon.

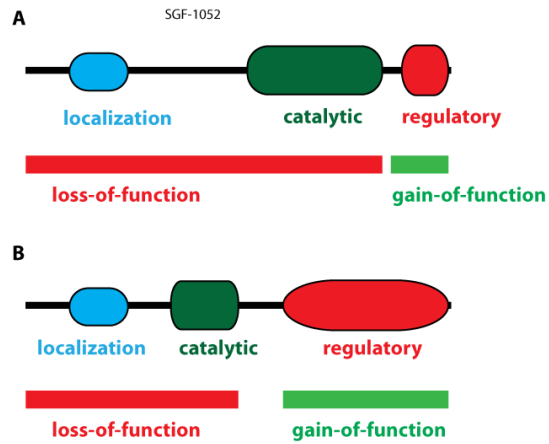


Figure 11 [SGF-1052]. The frequency of obtaining classes of mutations. A. The domains of one type of typical protein. B. Particularly pernicious protein with big negative regulatory domain!

Table 401. Dosage analysis of hypomorphic alleles

Genotype	% WT phenotype	Interpretation
$+/+$	100%	wild-type control
$h_1/+$	100%	allele is recessive
h_1/h_1	30%	mutant control
h_1/Df	15%	allele is enhanced in trans to a null
$h_1/h_1/h_1$	60%	increased copy number has a less severe phenotype

Table 403. Dosage analysis of hypermorphic alleles

Genotype	Activity	Interpretation
<i>H/H/H</i>	450%	increased copy number has a more severe phenotype
<i>H/H/+</i>	350%	increased copy of wild-type has a more severe phenotype
<i>H/H</i>	300%	mutant control
<i>H/+</i>	200%	allele is dominant
<i>+/+/+</i>	150%	increased copy of wild-type has a more severe phenotype
<i>H/Df</i>	150%	allele is suppressed in trans to a null
<i>+/+</i>	100%	wild-type control
<i>+/Df</i>	100%	<i>Df</i> is not dominant

Table 451. Dosage analysis of dosage-independent neomorphic allele

Genotype	Mutant?	Interpretation
+/+	no	wild-type control
N/+	yes	allele is dominant
N/+/+	yes	increased copy of wild-type allele does not suppress phenotype
+/+/+	no	increased copy of wild-type allele has no phenotype
N/Df	yes	allele not suppressed in trans to a null
+/ <i>Df</i>	no	<i>Df</i> is not dominant

Table 452. Dosage analysis of wild-type suppressible, dosage-dependent neomorphic allele

Genotype	Mutant?	Interpretation
+/+	no	wild-type control
N/+	yes	allele is dominant
N/+/+	no	increased copy of wild-type suppresses phenotype
+/+/+	no	increased copy of wild-type has no phenotype
N/Df	yes	allele not suppressed in trans to a null
+/ <i>Df</i>	no	<i>Df</i> is not dominant

Table 453. Dosage analysis of dominant antimorphic allele.

Genotype	Mutant?	Interpretation
+/+	no	wild-type control
N/+	yes	allele is dominant
N/+/+	maybe	increased copy of wild-type suppresses phenotype
+/+/+	no	increased copy of wild-type has no phenotype
N/Df	yes	no wild type function to poison, but acts as a loss-of-function since there is no wild-type gene activity
+/ <i>Df</i>	no	<i>Df</i> is not dominant

Table 462. Idealized allelic series can define loss of function phenotype

genotype	phenotype
$+/+$	100%
j_1/j_1	95%
j_2/j_2	70
j_3/j_3	40
j_4/j_4	20
j_5/j_5	10
j_6/j_6	3
j_1/j_5	50%
j_5/j_6	5%
j_2/j_6	36%
j_3/j_4	30%
j_2/j_5	40%

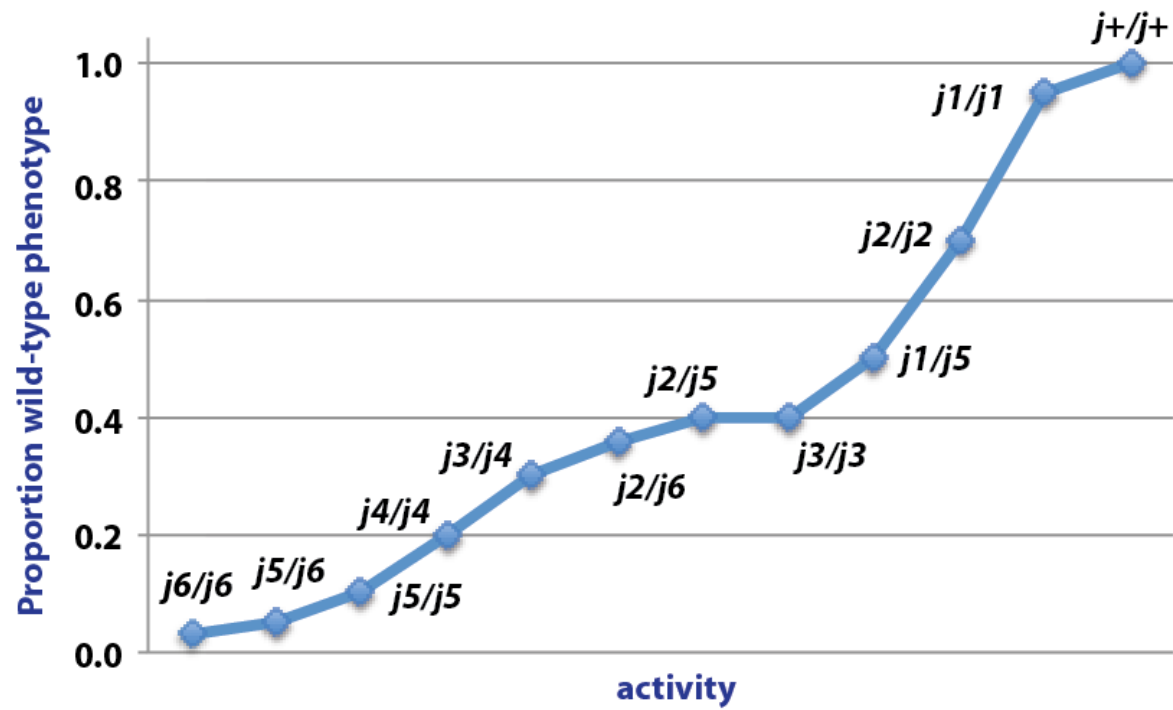


Figure 12 [SGF-1117]. Phenotype as a function of activity. Graph of phenotype and alleles.

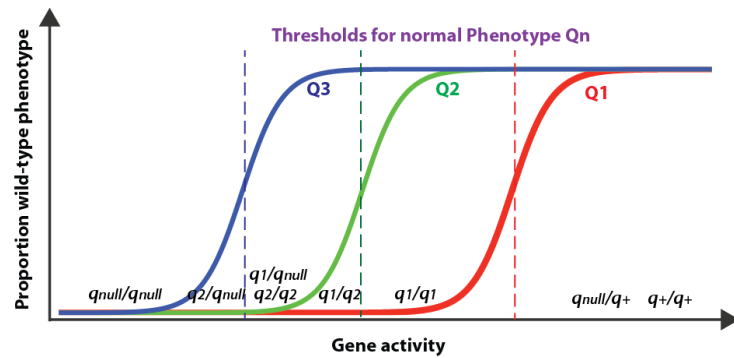
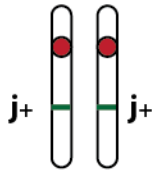


Figure 12.5 [SGF-1125] Phenotypic series.

Table 433. Level Hypothesis

Genotype	Phenotype		
	Q1	Q2	Q3
q_+/q_+	+	+	+
q_1/q_1	-	+	+
q_2/q_2	-	-	+
q_{null}/q_{null}	-	-	-
q_1/q_2	-	+/-	+
q_1/q_{null}	-	-	+
q_2/q_{null}	-	-	+/-

A. Diploid



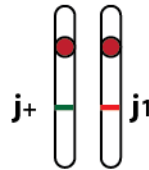
**Wild-type
Phenotype**

**B. Heterozygous
for deficiency**



Phenotype J

**C. Heterozygous for
dominant allele**



Phenotype J

Figure 13 [SGF-1114] Haploinsufficiency. A haploinsufficient locus has a phenotype (e.g., lethal) if one copy is removed, for example (B) by a genetic deficiency of the locus or (C) by a loss-of-function mutation.

Table 461. Dosage analysis of dose-dependent neomorphic allele

Genotype	Mutant?	Interpretation
+/+	no	wild-type control
N/+	yes	allele is recessive
N/N/+	no	two copy of mutant necessary to cause phenotype
N/N/+/+	no	two copy of mutant necessary to cause phenotype
N/Df	yes	allele not competed by wild-type

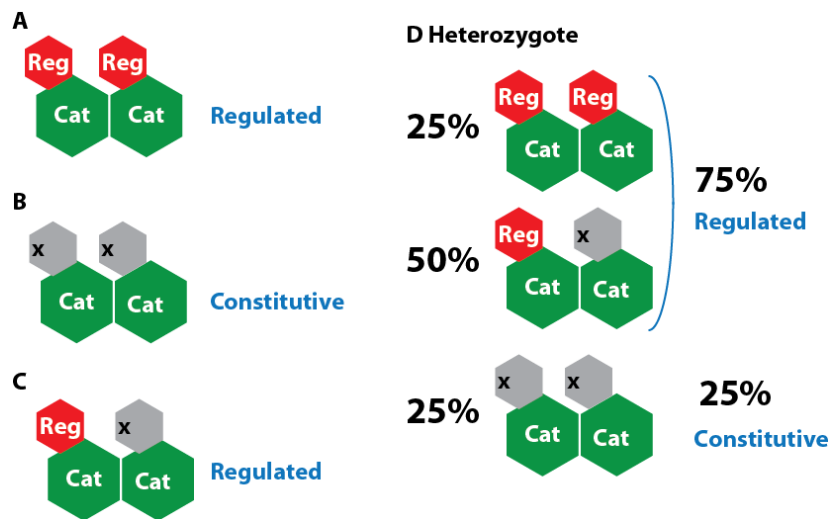


Figure 14 [SGF-1115]) How a gain of function could be recessive. D. A heterozygote would have a mixture of dimers with only 25% unregulated.

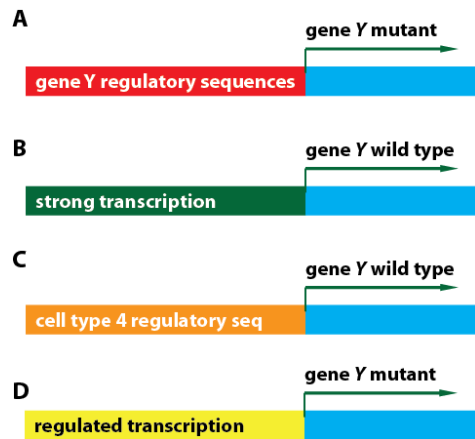


Figure 15 [SGF-1120]. Engineering of gene products to make gain of function genes.

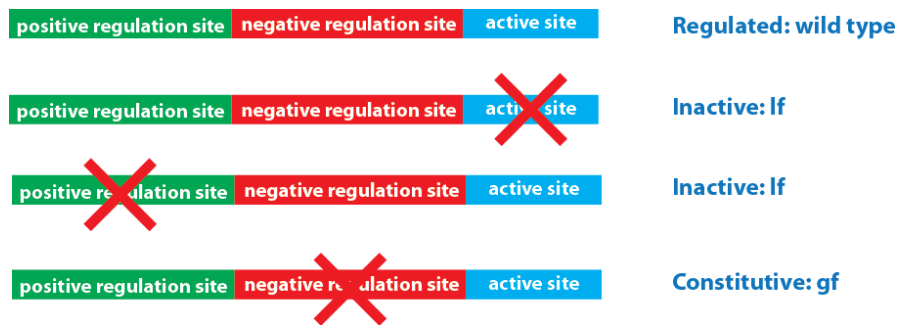


Figure 16 [SGF-1121]. Concept of making gain of function mutant proteins

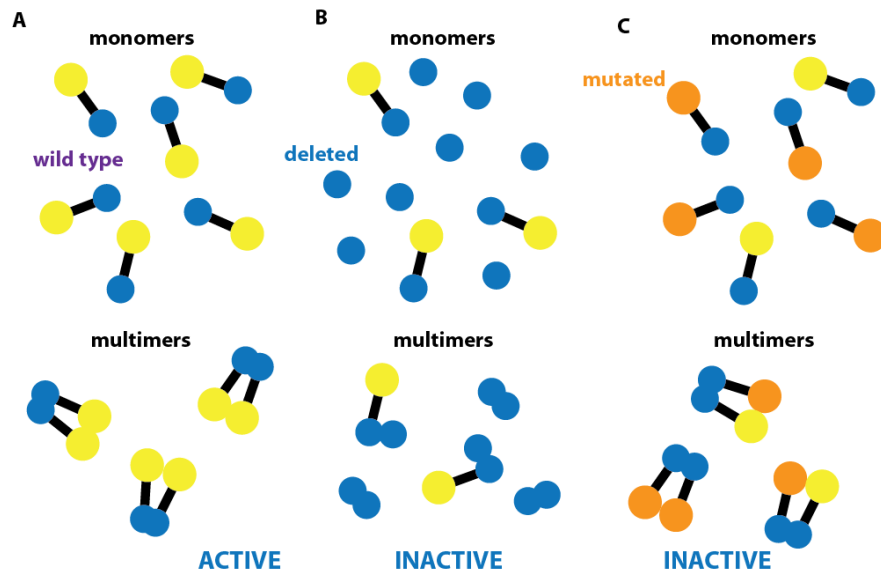


Figure 17 [SGF-1122]. Dimeric protein with dominant negative derivatives. A. monomer with DNA-binding domain (yellow) and dimerization domain (blue) form active dimers. B. Proteins with deleted DNA-binding domains can nonetheless dimerize and thus poison the wild-type monomers. C. Monomers with mutant DNA-binding domains (orange) can dimerize but don't function; they thus poison wild-type monomers.

Table 406. Opposite gain and loss of function

Gene	Organism	lf	wt	gf
<i>Antp</i>	<i>Drosophila</i>	leg to antennae	leg and antennae	Antennae to leg
<i>cdc2</i>	<i>S. pombe</i>	no cell cycle	regulated cell cycle	early cell cycle
<i>lin-12</i>	<i>C. elegans</i>	two AC	1 AC; 1 VU	two VU
<i>tra-1</i>	<i>C. elegans</i>	XX males	XX female; XO males	XO females
<i>ced-9</i>	<i>C. elegans</i>	extra apoptosis	regulated apoptosis	no apoptosis
<i>Lmbr1</i>	<i>M. musculus</i>	oligo-dactyly	normal digits	polydactyly

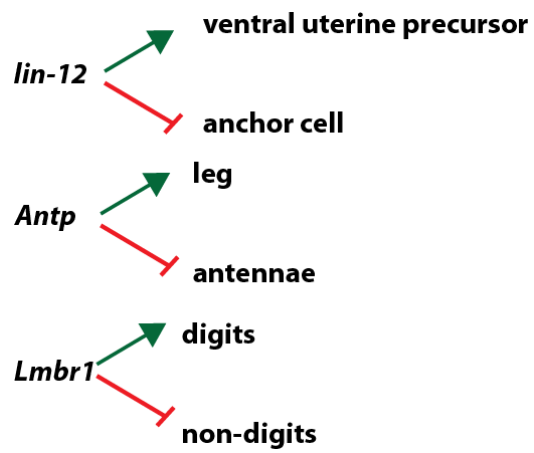


Figure 18 [SGF-1124]. Switch genes. Green arrow is positive regulation; red bar is negative regulation.

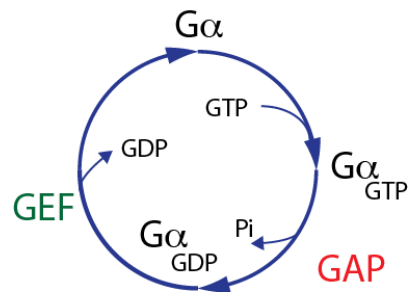


Figure 19 [SGF-1110]. G protein cycle.

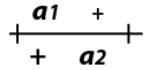
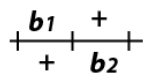
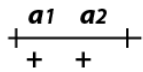
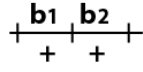
A. trans**Phenotype A****C. trans****Wild-type****Phenotype non-A****B. cis****Wild-type****Phenotype non-A****D. cis****Wild-type****Phenotype non-A**

Figure 20. [SGF-1092]. The complementation test. By constructing organisms that are heterozygous for each of two recessive mutations one can test the functional relationship of those mutations. If the two mutations are in the same functional unit, then the heterozygote ($a_1 +/+ a_2$) will have the mutant phenotype, but will be wild type in trans. If the two mutations are in different functional units, then the heterozygote ($a_1/+ b_1/+$) will display the wild-type phenotype whether in *cis* or in *trans*.

Table 430. Complementation rescue of mutant phenotype

Genotype		Phenotype
<i>gene A</i>	DNA	
+/+	none	non-A
+/-	none	non-A
-/-	none	A
-/-	A(+)	non-A
-/-	A(-)	A
+/+	A(+)	non-A

Figure 21 [SGF-1197]. Complementation rescue of a mutant. To make. **RESCUED WORMS.** Wt, mutant, wt and addition of active and inactive gene

**A. Intragenic
Complementation**

a1	+	+	Wild-type Phenotype non-A
+	a2	+	

a1	+	+	Phenotype A
+	+	a3	

+	a2	+	Phenotype A
+	+	a3	

**B. Extragenic
non-Complementation**

d1	+	Phenotype A
+	e1	

d1	+	Wild-type Phenotype non-A
+	+	

+	e1	Wild-type Phenotype non-A
+	+	

Figure 22 [SGF-1090] Exceptions to complementation. A. Intragenic complementation occurs when two alleles complement. B. Extragenic non-complementation occurs when alleles of distinct loci fail to complement.

Table 434. Multiple Function Hypothesis

Genotype	Phenotype		
	R1	R2	R3
r_+/r_+	+	+	+
r_1/r_1	-	+	-
r_2/r_2	+	-	-
r_{null}/r_{null}	-	-	-
r_1/r_2	+	+	-
r_1/r_{null}	-	+	-
r_2/a_{null}	+	-	-

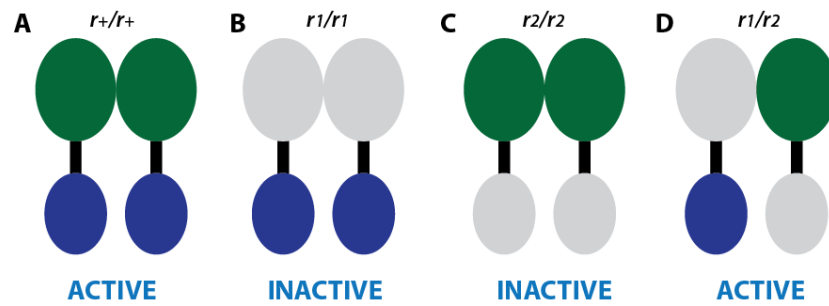


Figure 23 [SGF-1126]. Intragenic complementation in proteins with two domains. Receptor tyrosine kinases such as EGF-receptor are the key examples.

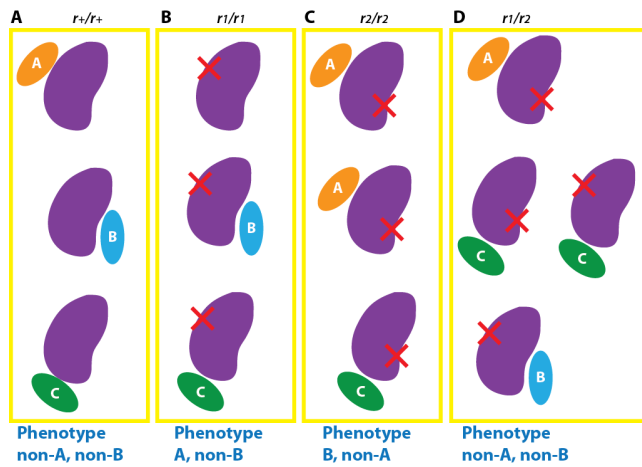


Figure 24 [SGF-1127]. intragenic complementation involves monomeric proteins. Wild-type protein (purple), a product of the *R* gene, independently binds three other proteins (A, B and C). The *r1* allele is defective in the A interaction; the *r2* allele in the B interaction. The types of monomers present in each genotype are shown. In the heterozygote, monomers that interact with A, B and C are present.

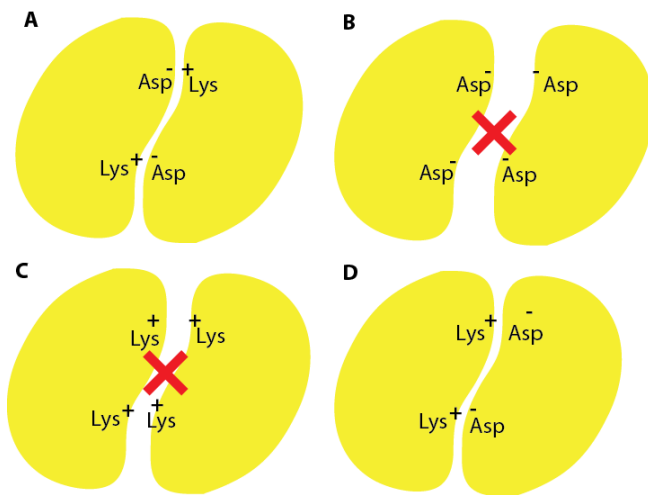


Figure 25 [SGF-1118]. Intragenic complementation in a homodimeric protein.
 Lys, positively charged lysine residue; Asp, negatively charged aspartic acid residue.

Table 436. Intragenic complementation via different expression domains

Genotype	Phenotype	
	Tail	Belly
$+/+$	+	+
h_1/h_1	-	+
$h_1/+$	+	+
h_2/h_2	+	-
$h_2/+$	+	+
h_1/h_2	+	+

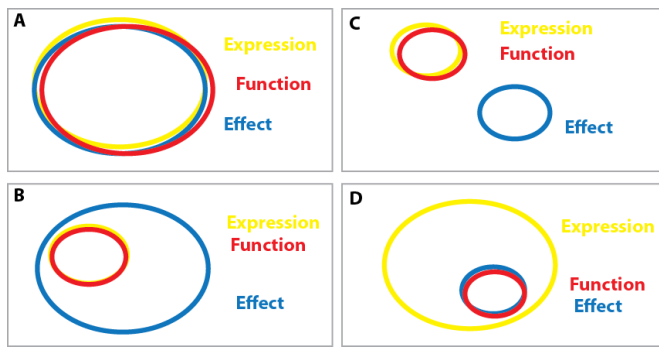


Figure 26 [SGF-1143]. Relationships among Gene Expression, Function and Phenotypic Effect. A. Expression, function and phenotypic effect are concordant. B. The effect of a gene is broader than its expression and function. C. The Effect of a gene is on cells other than where it is expressed and functions. D The function and effect of a gene are a subset of its expression.

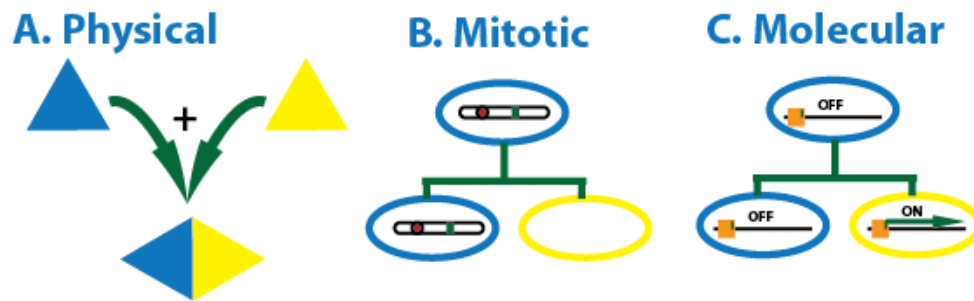
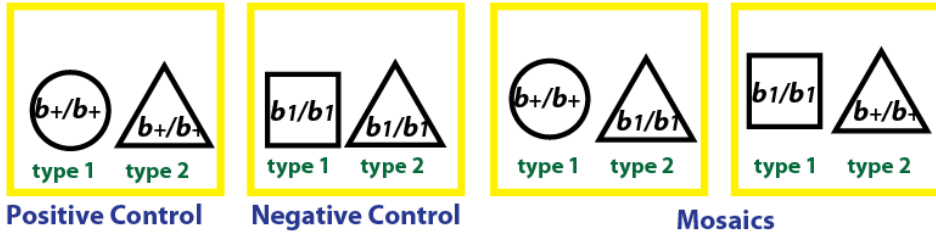


Figure 27 [SGF-1140]. General ways to make mosaic organisms. A. physical combination. B. change in genotype at mitosis. C. molecularly-engineered changes in gene expression. **SHOULD C not have mitosis?**

A. Autonomous



B. Non-Autonomous

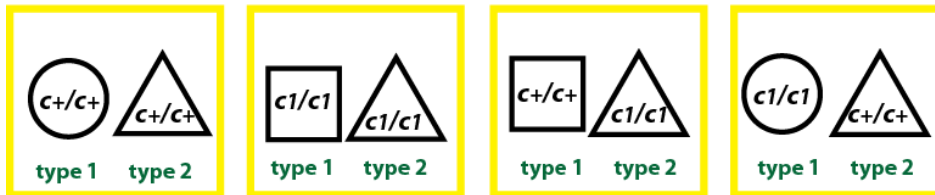


Figure 28 [SGF-1096]. Autonomy and non-autonomy. The wild type has a circle. Mutations b_1 and c_1 transform the circle to a square. A. The B gene acts in the circle cell. B. The C gene acts in a distinct cell (triangle) but affects the phenotype of the circle cell.

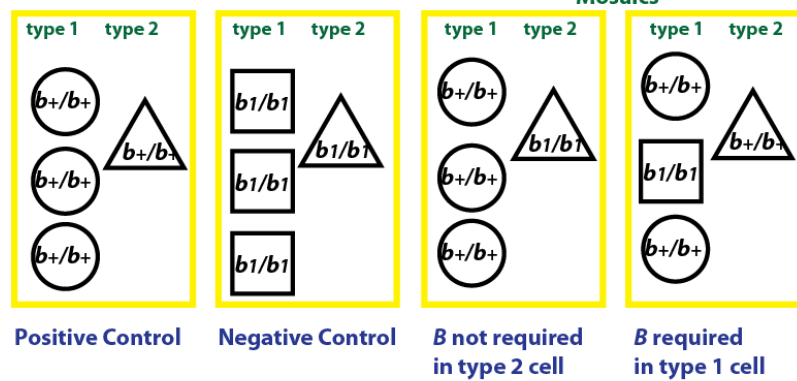
Table 428. Cell autonomous

gene B Genotype of cell		Phenotype of type1 cell
type 1	type 2	
+	+	round
+	-	round
-	+	square
-	-	square

Table 429. Non cell autonomous

gene C Genotype of cell		Phenotype of type 1 cells
type 1	type 2	
+	+	round
+	-	square
-	+	round
-	-	square

A. Autonomous



B. Non-Autonomous

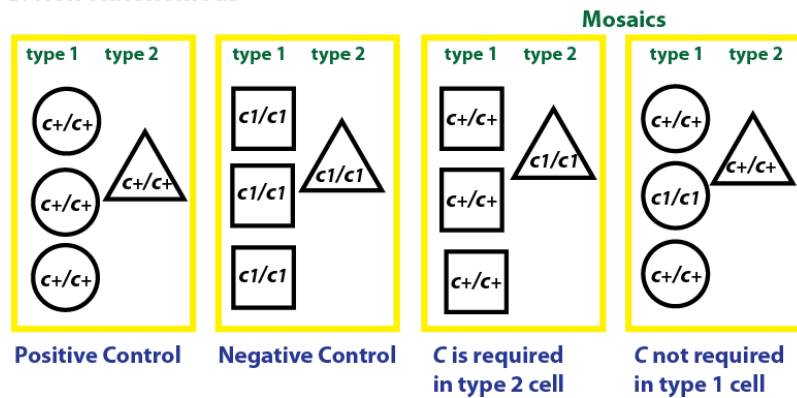


Figure 29 [SGF-1001]. Autonomy and non-autonomy A. Autonomous. B. non-autonomous.

Non-Autonomous: majority rules

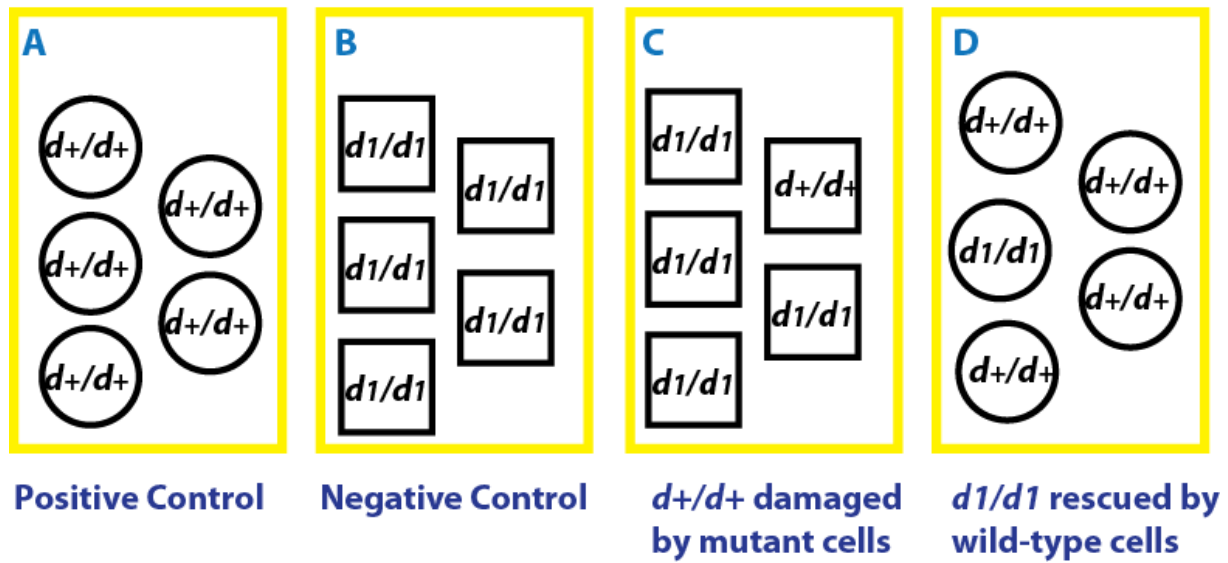


Figure 30 [SGF-1102] Majority rules.

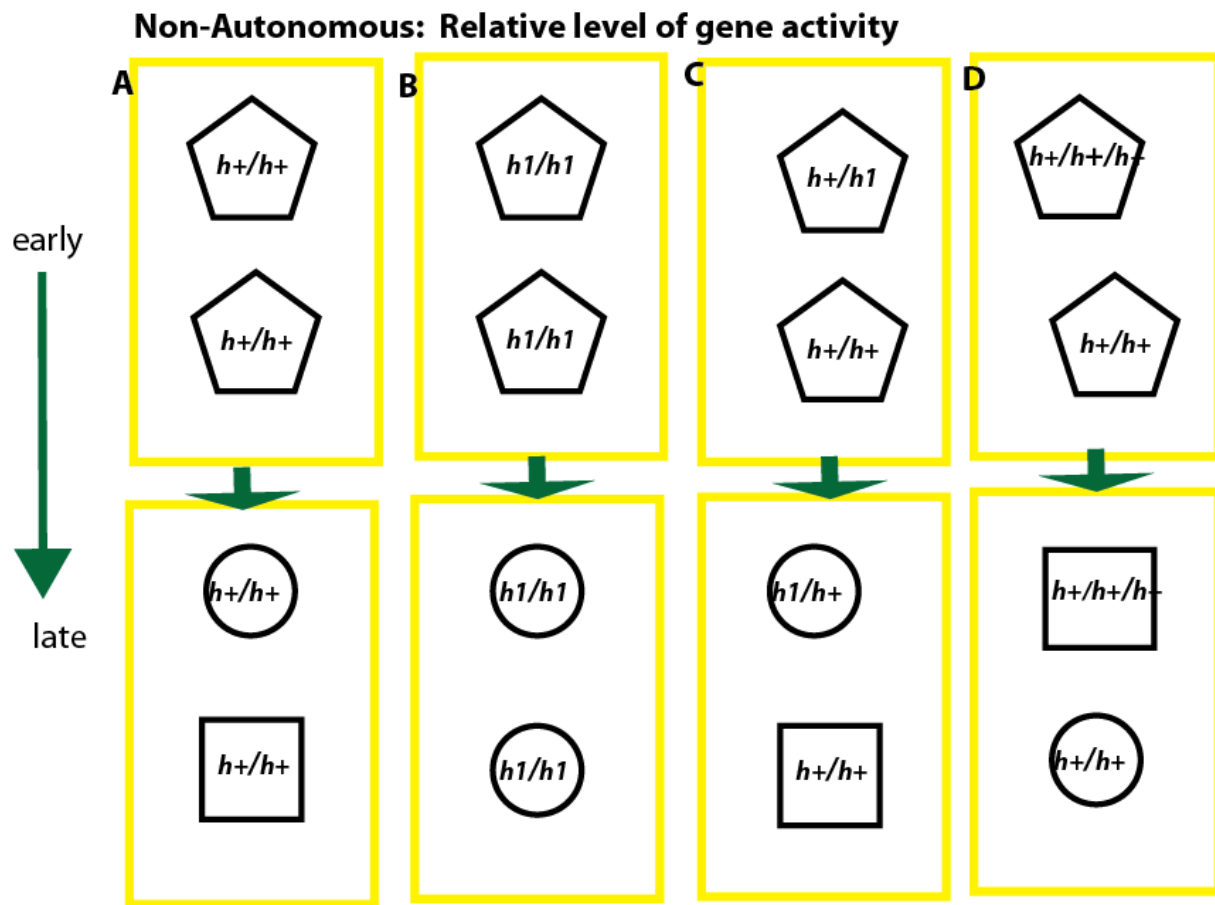


Figure 31 [SGF-1103]. Non-autonomous relative levels of gene activity. Pentagons, precursor cells; circles and squares are two differentiated cell types derived from the pentagon.

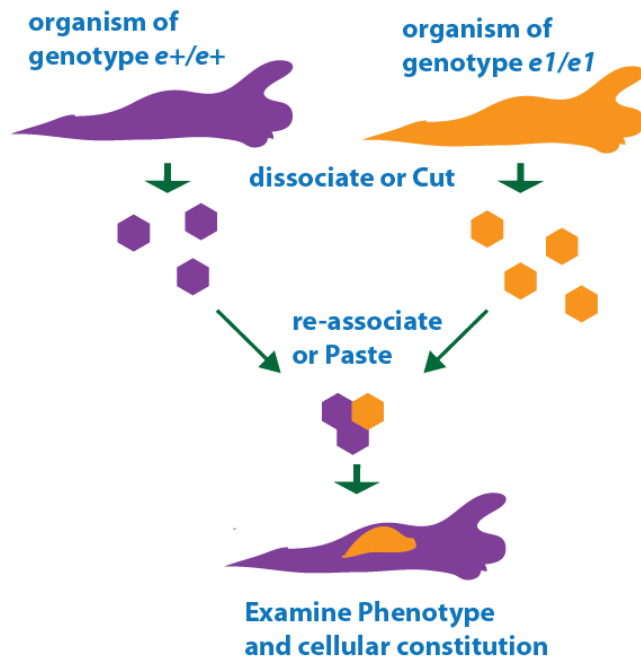


Figure 32 [SGF-1129]. Chimeric organisms constructed by physical means (re-association or surgery).

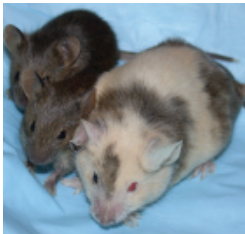


Figure 33 [SGF-1159]. Chimeric mice The chimera has produced offspring that carry the agouti coat color gene from the embryonic stem cells [<http://intramural.nimh.nih.gov/tgc/photogallery.html>] get SOMETHING BETTER

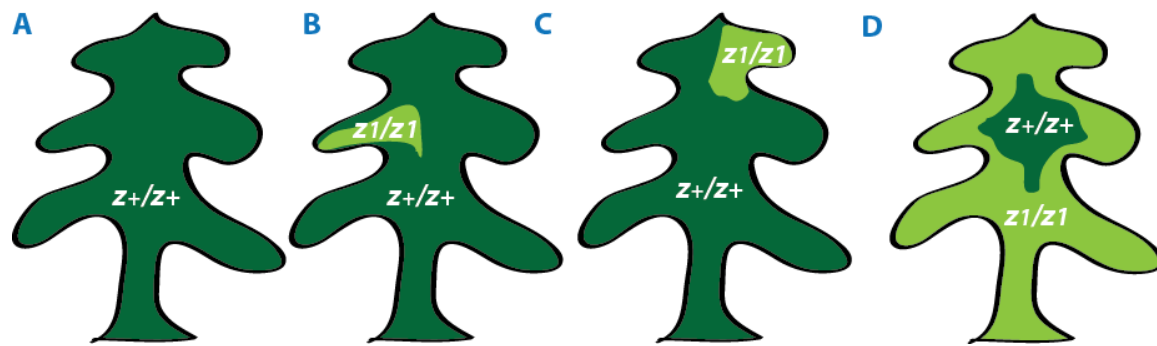


Figure 34 [SGF-1134]. Mosaic organisms. A. wild-type for gene *Z*. B. Patch of z_1/z_1 in an otherwise z_+/z_+ organism. C. A different patch of z_1/z_1 in an otherwise z_+/z_+ organism. D. A patch of z_+/z_+ in an otherwise z_1/z_1 organism.

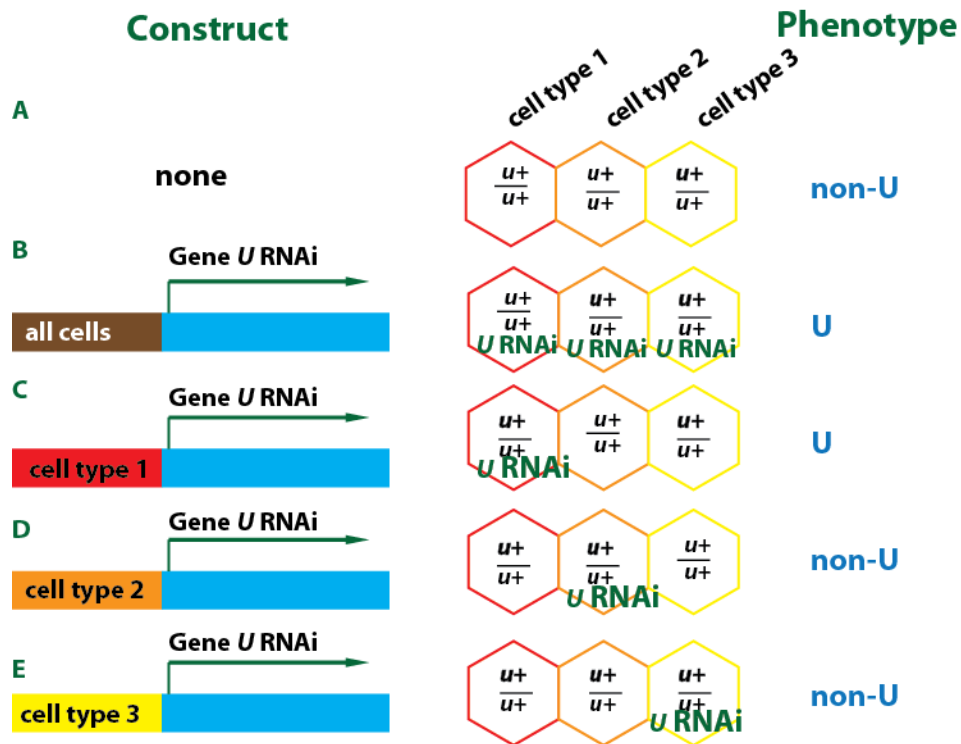


Figure 35 [SGF-1131]. Site of action analysis by expression in wild-type background of an RNAi-inducing construct. A. control organism has the wild-type (**non-U**) phenotype (Negative Control). B. Expression of *gf U* RNAi construct in all cells results in a **U** phenotype (Positive Control). D. Expression in cell type 1 results in a **U** phenotype. E, F. Expression of RNAi in cell types 2 or 3 has no effect.

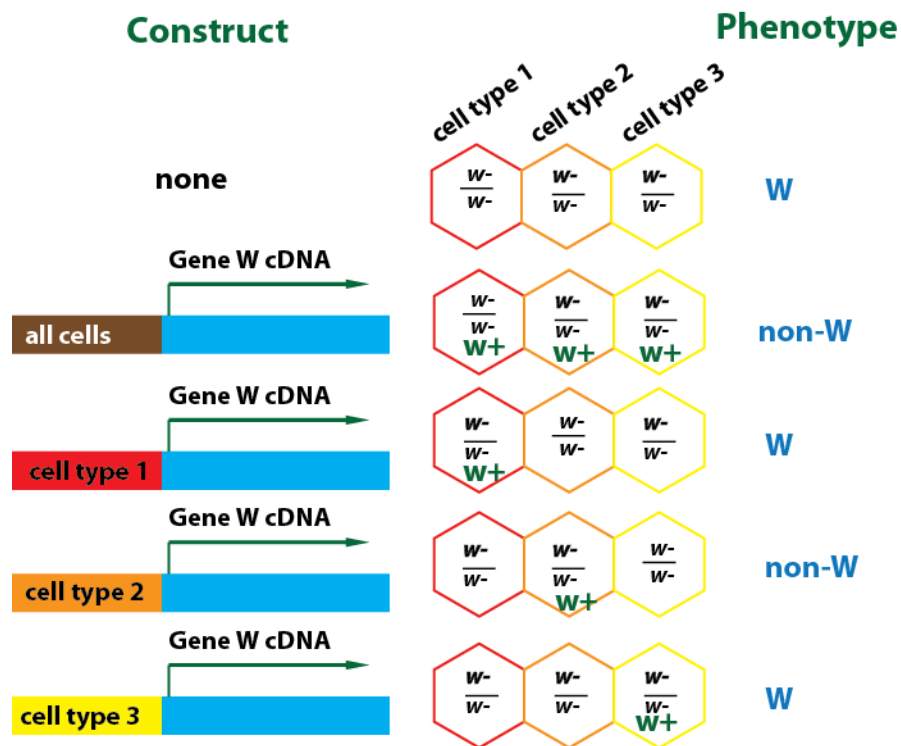


Figure 36 [SGF-1116]. Site of action analysis by expression of wild-type gene in mutant background implies sufficiency.

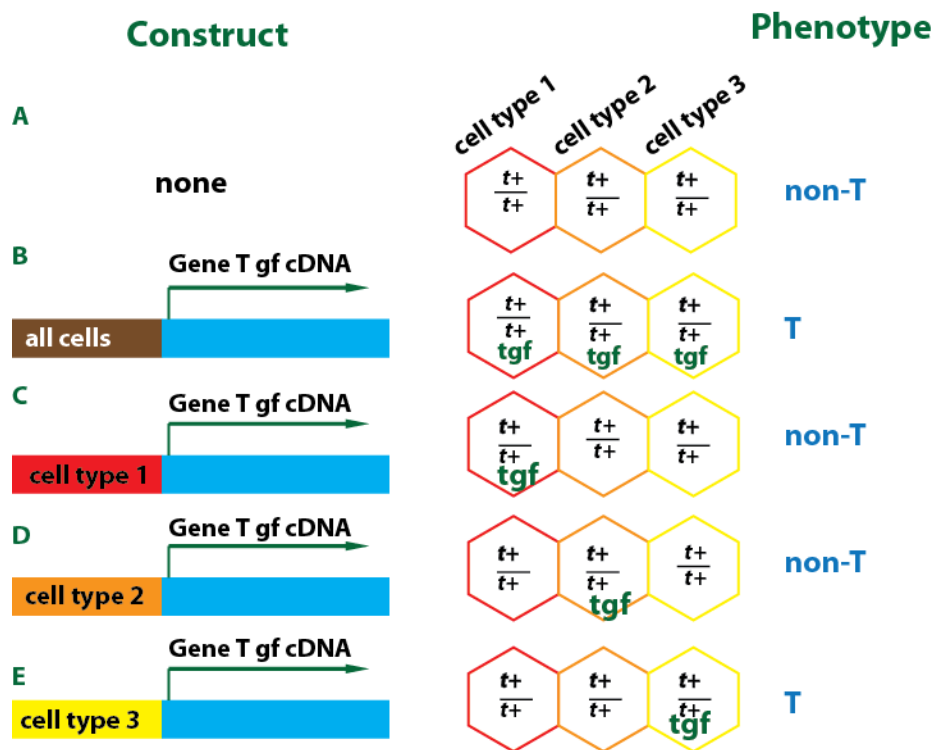


Figure 37 [SGF-1130]. Site of action analysis by expression in wild-type background of a gain of function allele. A. control organism has the wild-type (**non-T**) phenotype (Negative Control). B. Expression of gf *T* cDNA in all cells results in a **T** phenotype (Positive Control). C, D. Expression of gf in cell types 1 or 2 has no effect. E. Expression in cell type 3 results in a **T** phenotype.

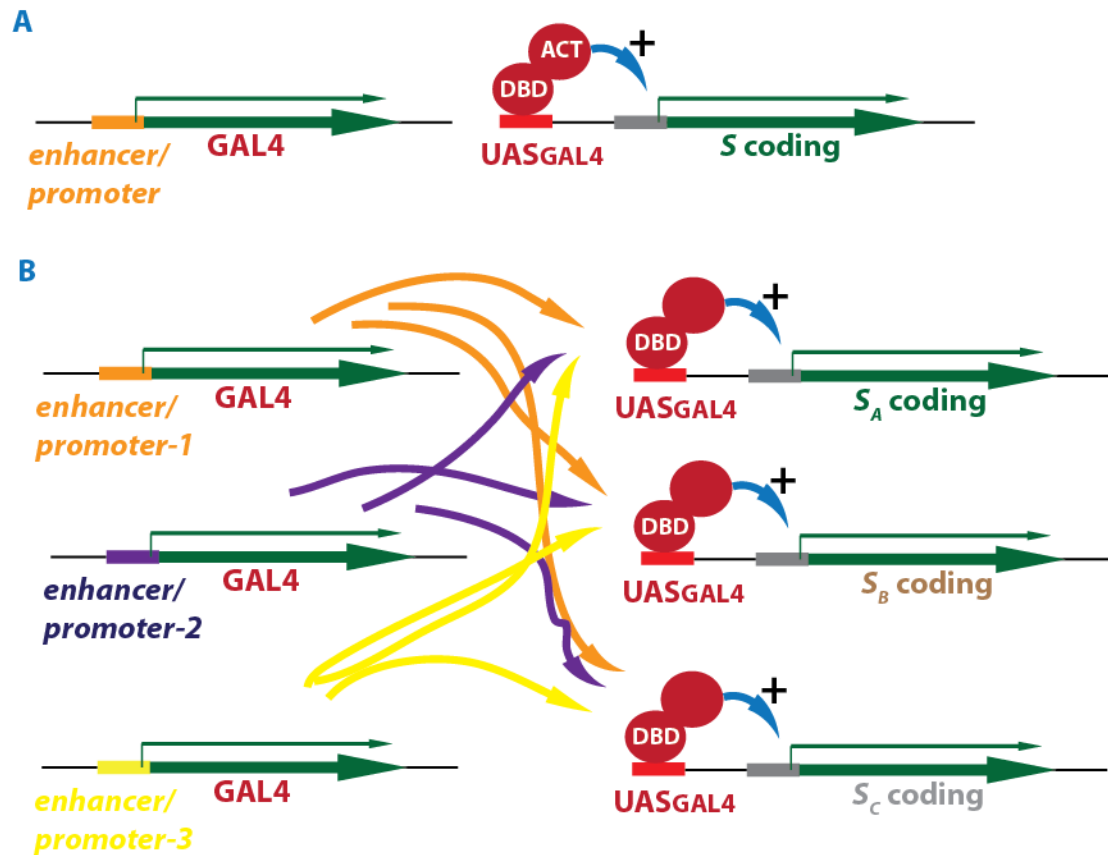


Figure 37.5 [SGF-1137]. GAL4-UAS system for gene expression. A. The GAL4 gene encode a bipartite activator comprising a DNA-binding domain (DBD) and an activation (ACT) domain. It binds a specific DNA sequence (UAS_{GAL4}) and activates transcription. **B.** Combinatorial use of the GAL4 system. Each enhancer/promoter (1, 2, and 3) directs GAL4 in a different expression pattern. Each target gene (S_A, S_B and S_C) responds to GAL4. These six constructs give nine potential combinations.

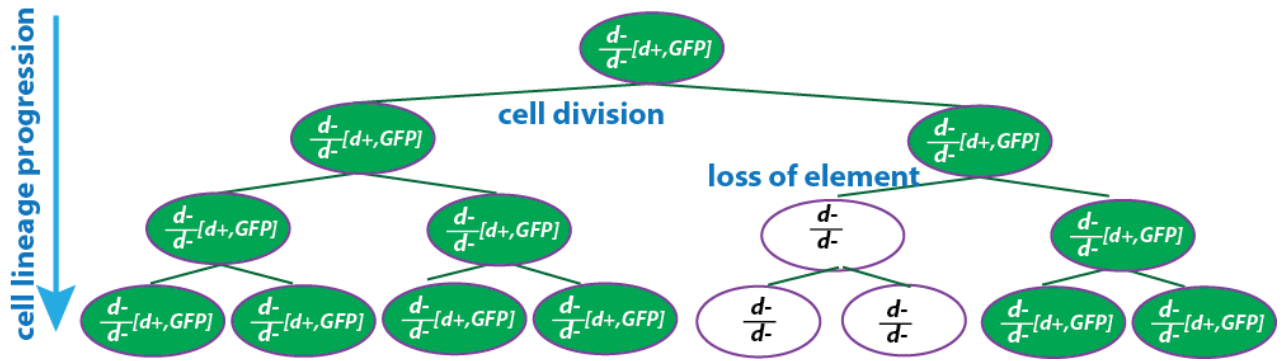


Figure 38 [SGF-1133]. loss of element leads to mosaicism.

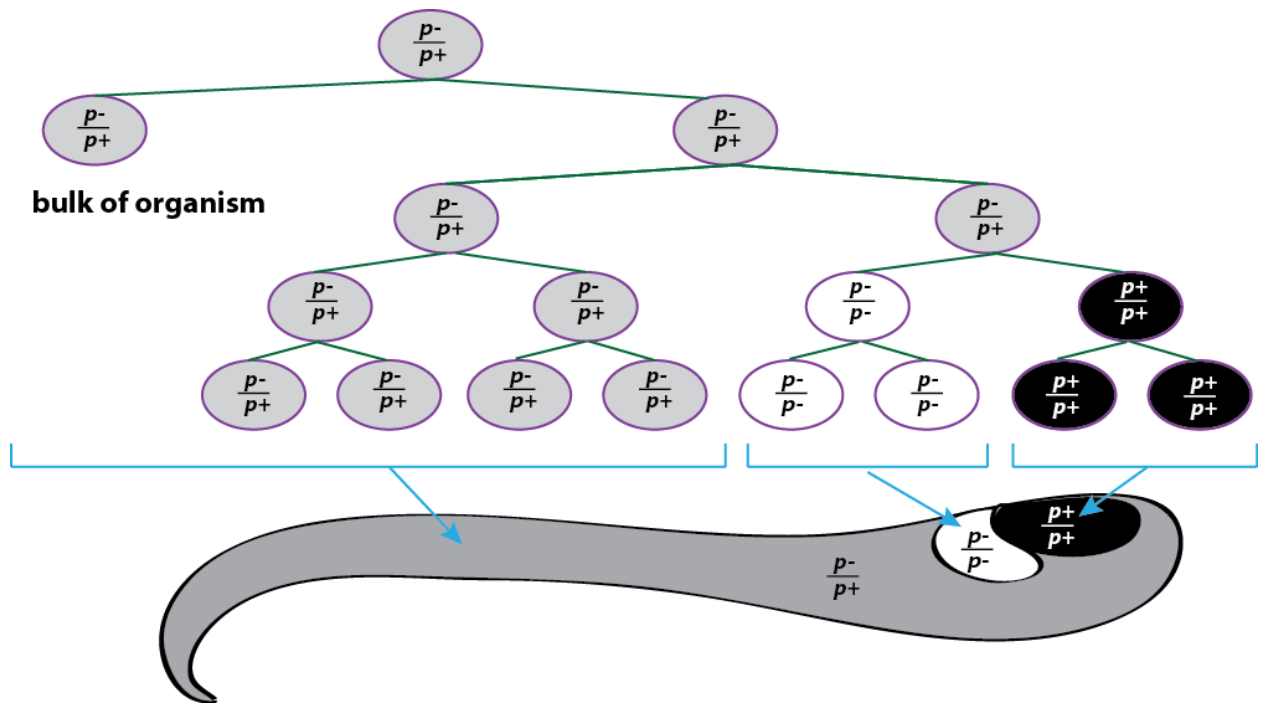


Figure 39 [SGF-1132]. Mitotic recombination leads to mosaicism.

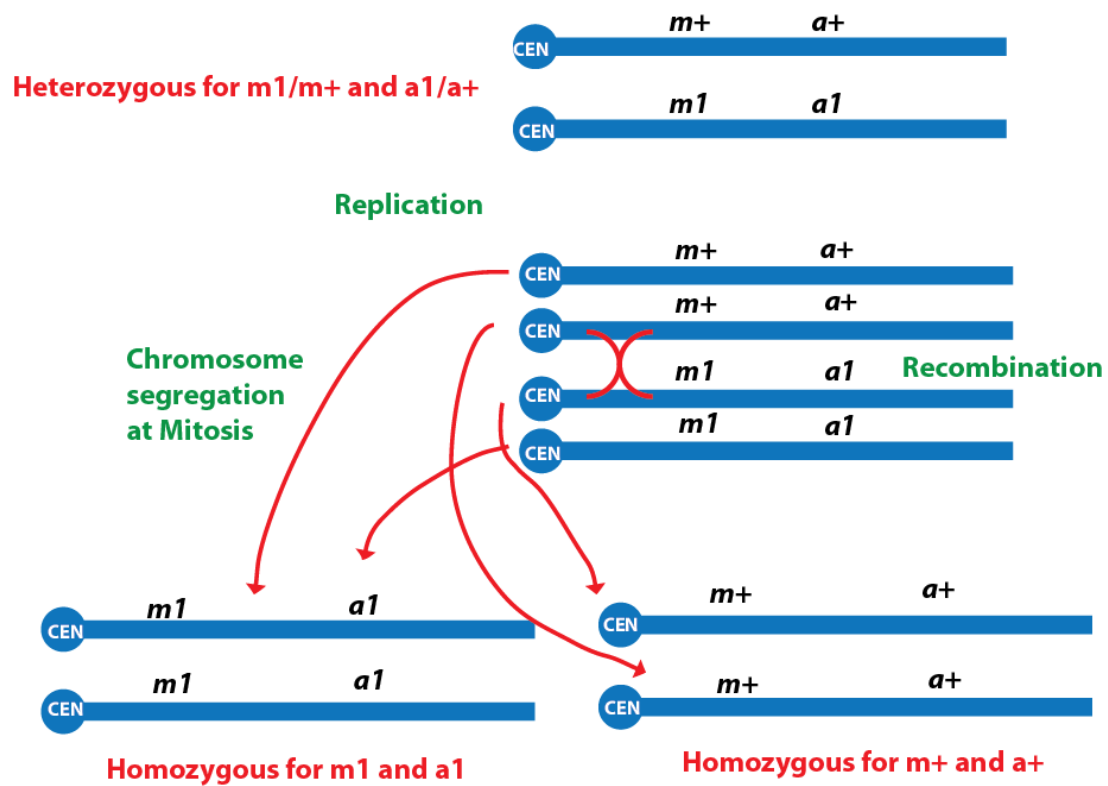
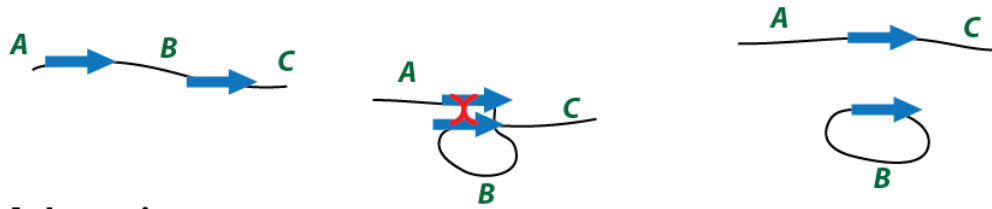


Figure 40 [SGF-1093]. Mitotic recombination can create mosaic organisms.

A. Deletion



A. Inversion

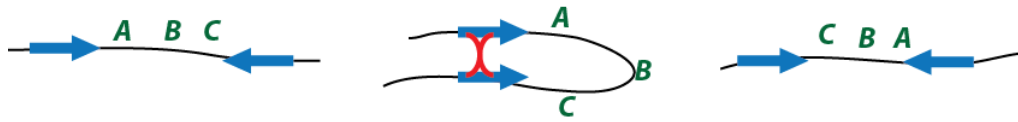


Figure 41 [SGF-1097]. Site-specific recombination can delete or invert chromosomal segments.

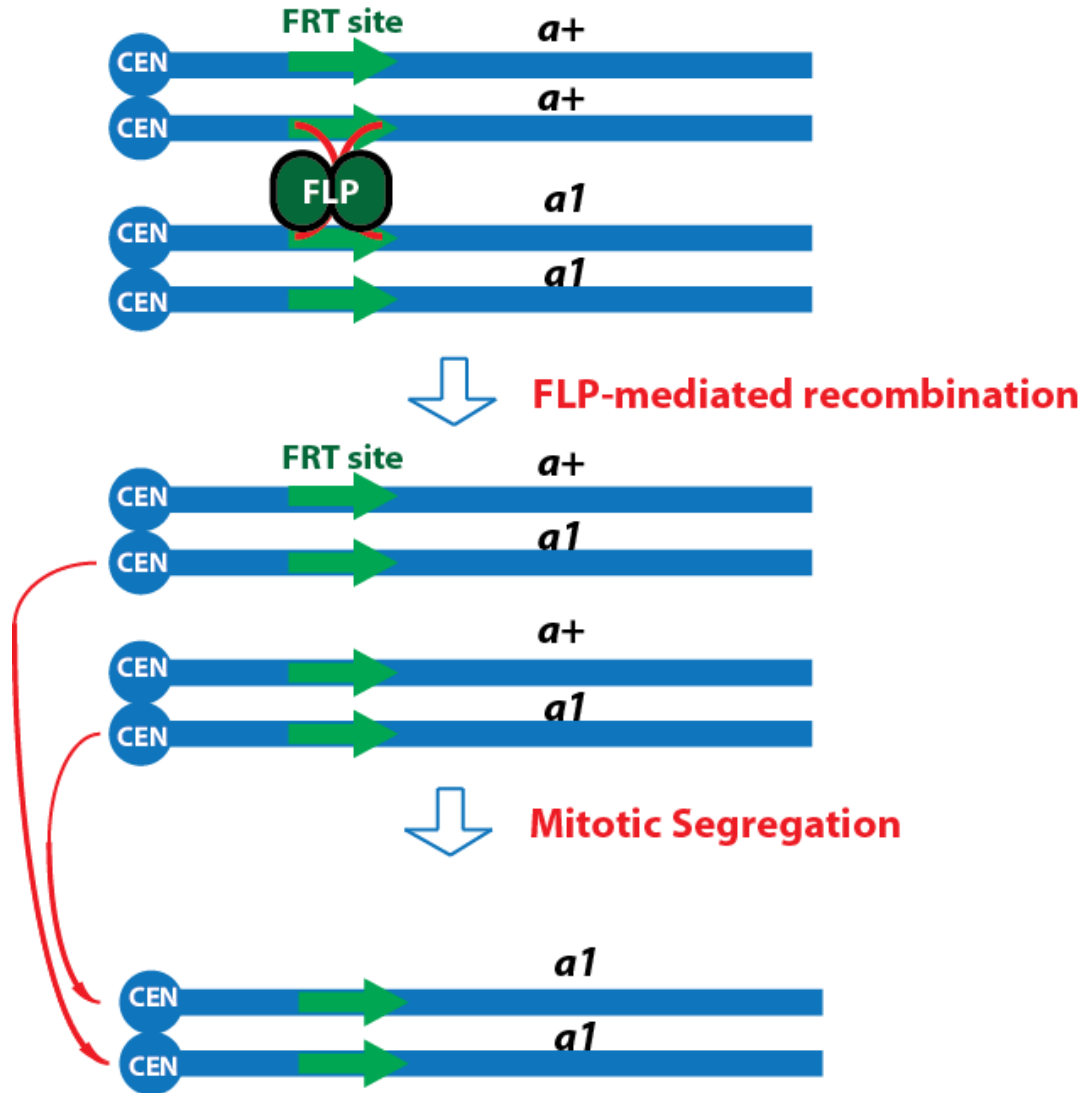


Figure 42 [SGF-1095]. FRT mosaics. Green arrows, FLP-Recognition Target (FRT). CEN, centromere. FLP, FLP recombinase.

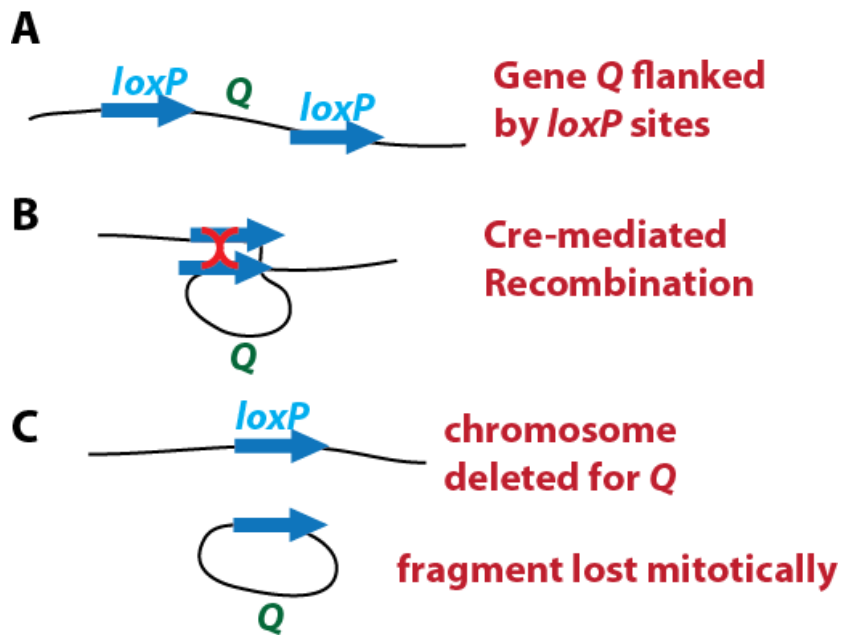


Figure 43 [SGF-1098]. Cre-mediated recombination at *loxP* sites deletes a gene in cells in which Cre is expressed. A. Gene Q is flanked by *loxP* sites. B. Cre mediates recombination between the *loxP* sites leading to a deletion of the intervening sequences: a single *loxP* site in the chromosome and the Q gene on a circular fragment, which is then lost during mitosis as it is unable to replicate.

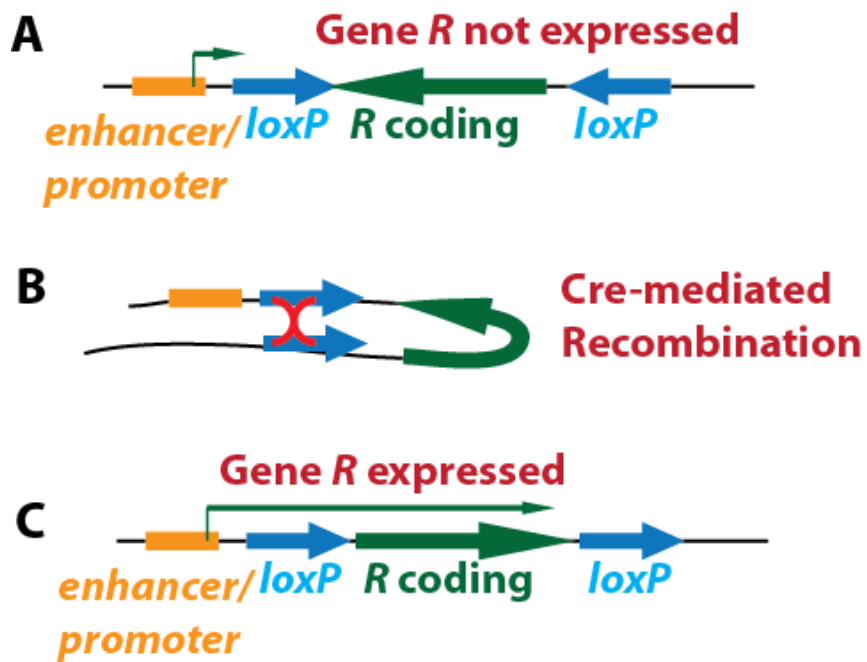


Figure 44 [SGF-1135]. Activation of gene expression by Cre-mediated inversion.

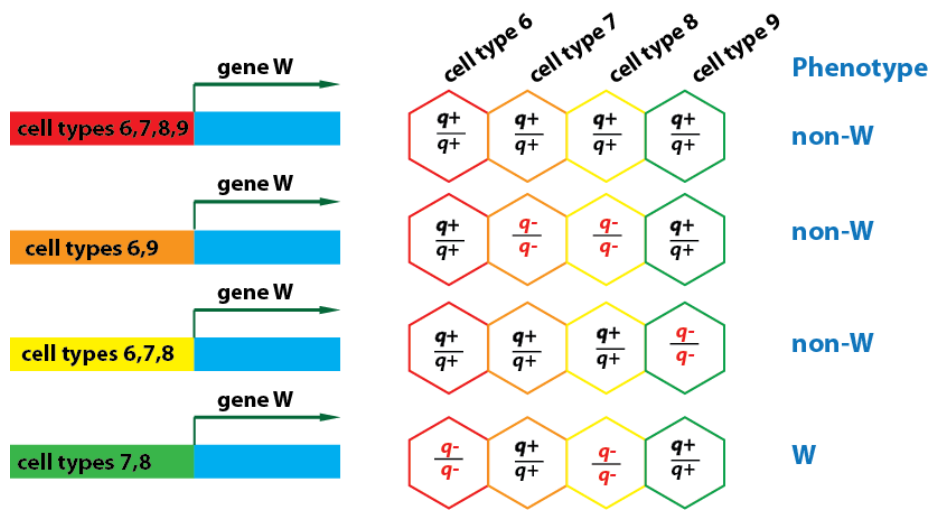


Figure 45 [SGF-1195]. Inferring site of action from multiple constructs

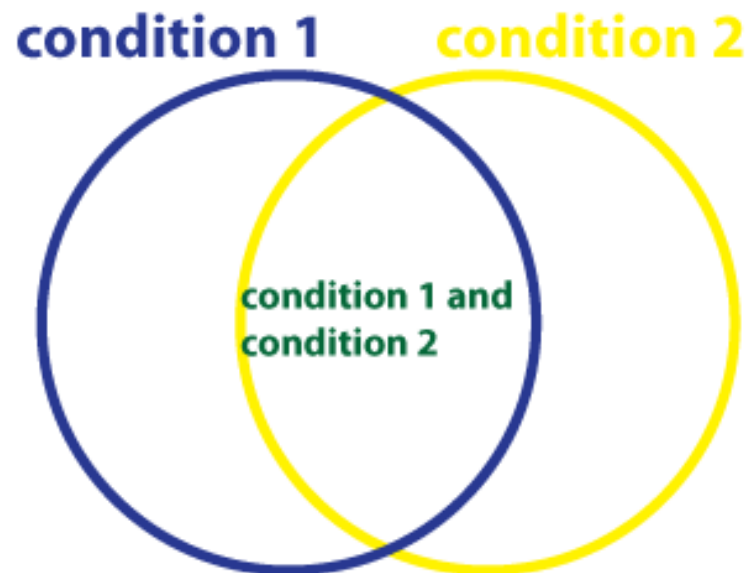


Figure 45.3 [SGF-1087]. Intersectional strategies result in an effect only if two conditions are met.

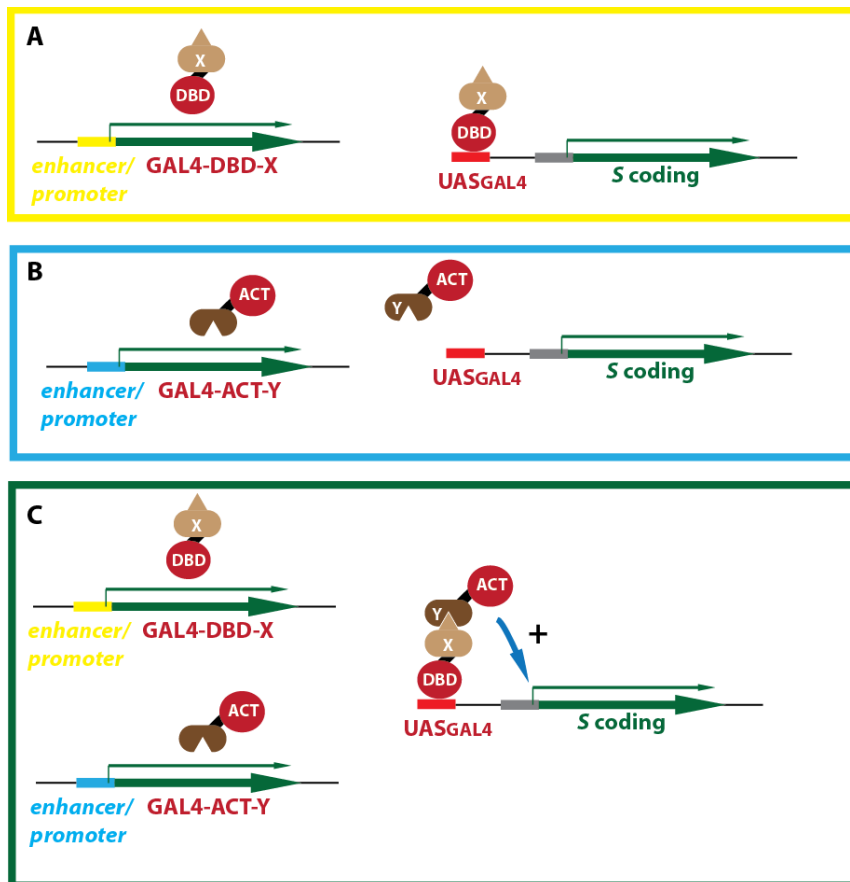


Figure 45.7 SGF-1138. Split GAL4-UAS system for gene expression. A. In some cells (yellow) GAL4-DBD-X is expressed. This lacks an activation domain and thus fails to activate transcription. In other cells (blue), GAL4-ACT-Y is expressed. This protein lacks a DNA-binding domain and thus fails to activate transcription. C. However, in other cells (green), both GAL4-DBD-X and GAL4-ACT-Y are expressed. If X and Y domains can bind tightly enough, a bipartite transcriptional activator activates transcription because the DBD binds DNA (namely GAL4-UAS) and recruits ACT via the interaction of X and Y.

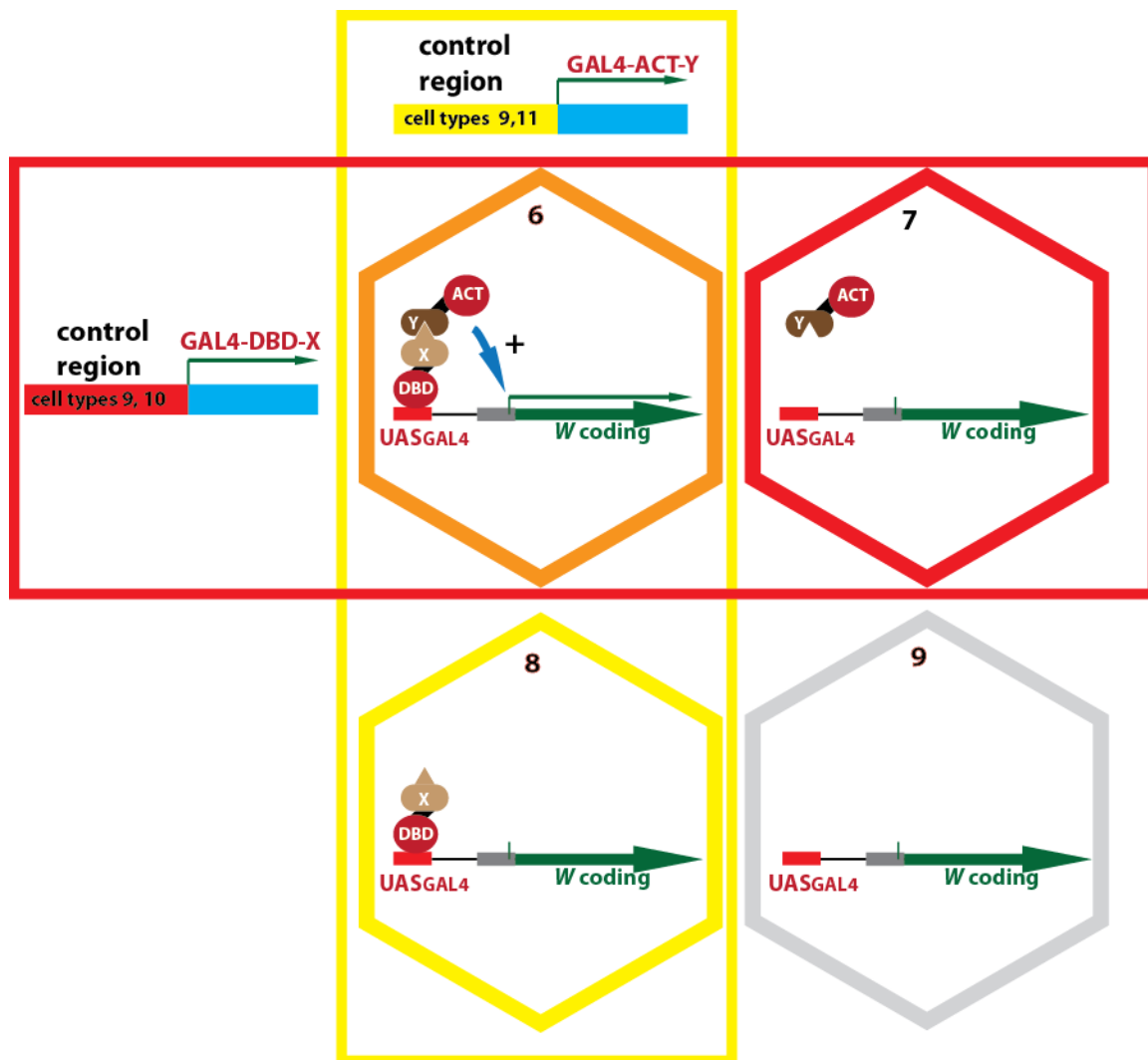


Figure 46 [SGF1196]. Intersectional strategy using split GAL4

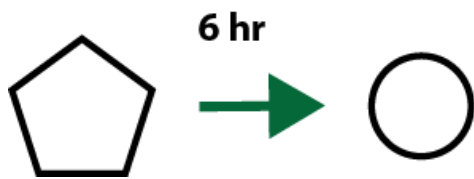


Figure 47 [SGF-1128]. Differentiation of generic cell.

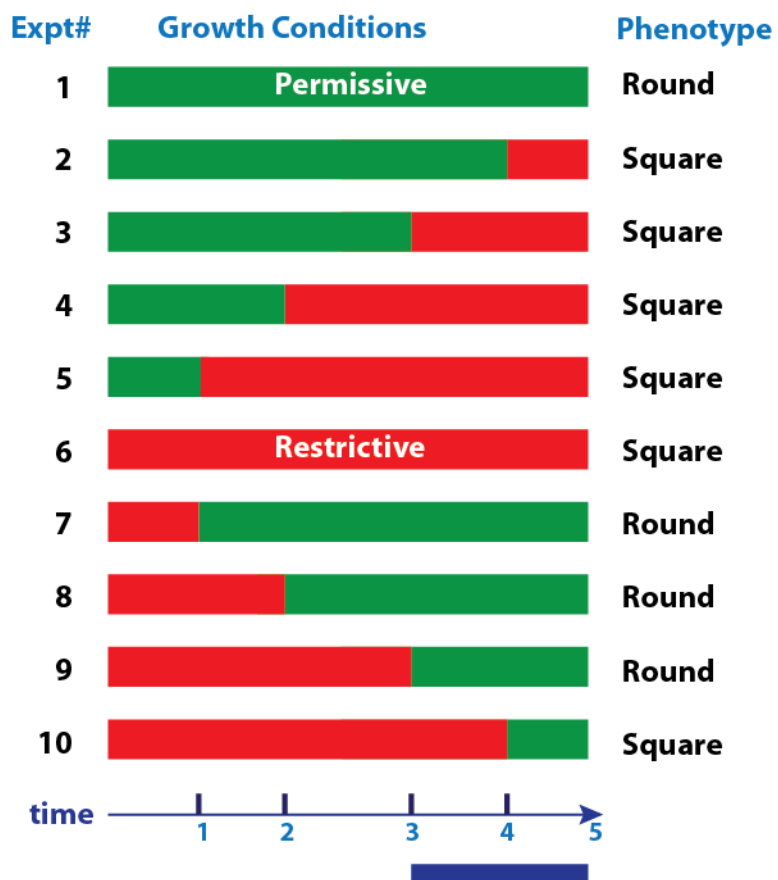


Figure 48 [SGF-1144]. Time of action experiments using b_2 . p, permissive conditions; r, restrictive conditions.

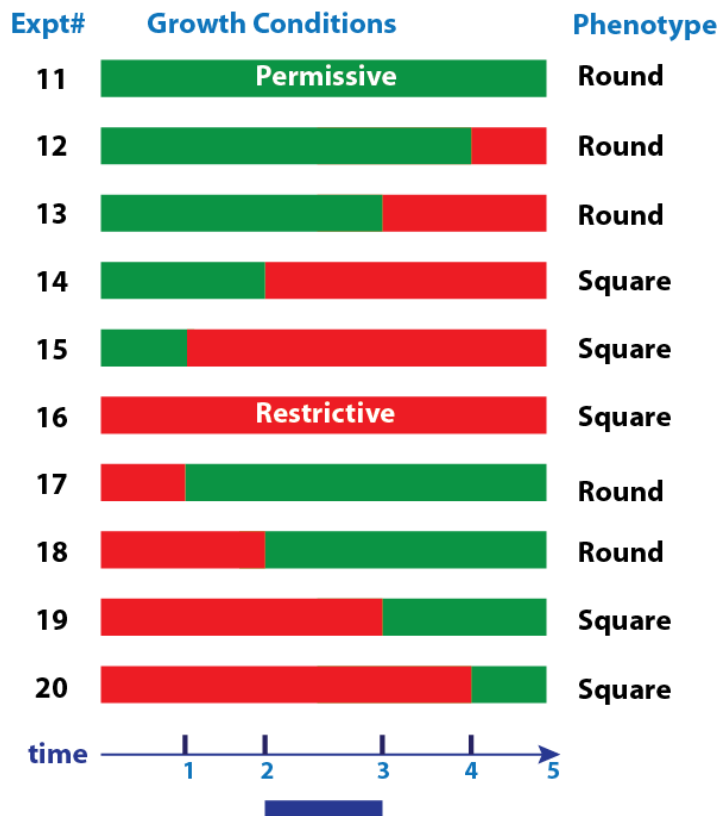


Figure 49 [SGF-1145]. Time of action experiments using c_2 . p, permissive conditions; r, restrictive conditions

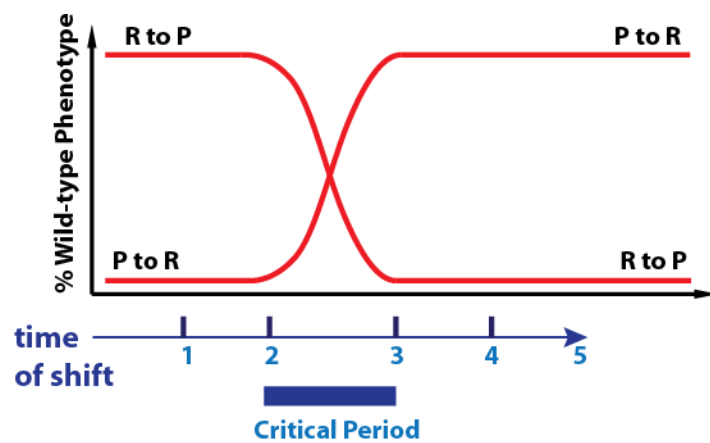


Figure 50 [SGF-1106]. Condition-sensitive period. A series of experiments as in Figure 1145 are carried out and the proportion of organisms with wild-type phenotype are plotted as a function of the time of shift from permissive (P) to restrictive (R) conditions, and from restrictive to permissive conditions.

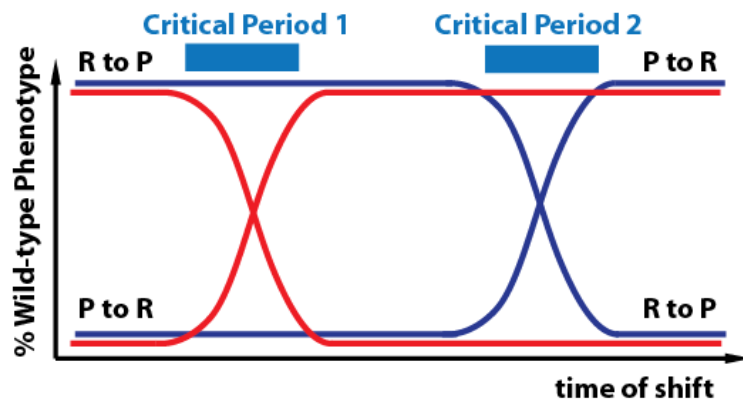


Figure 51 [SGF-1111]. Two distinct critical time periods.

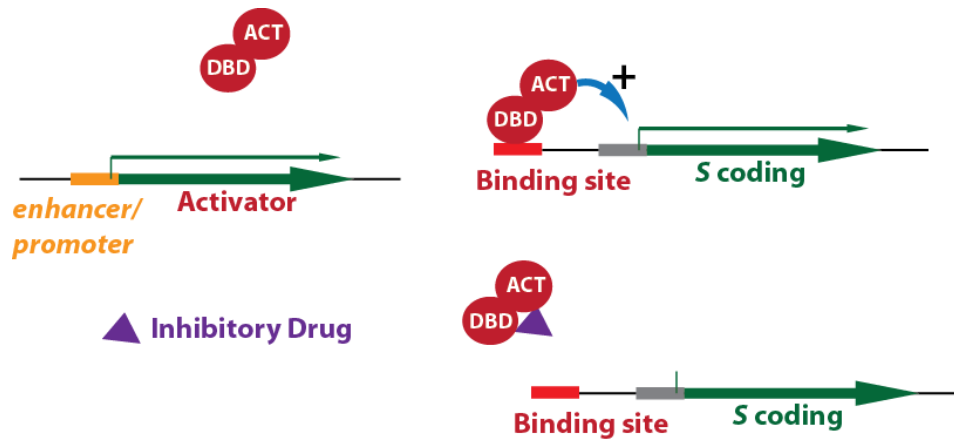


Figure 52 [SGF-1156]. Drug can add temporal control. Drug binds transcriptional activator and prevents activation of S gene expression.

no heat shock



heat shock



Figure 53 [SGF-1139]. Heat shock induction of gene expression to obtain temporal control.

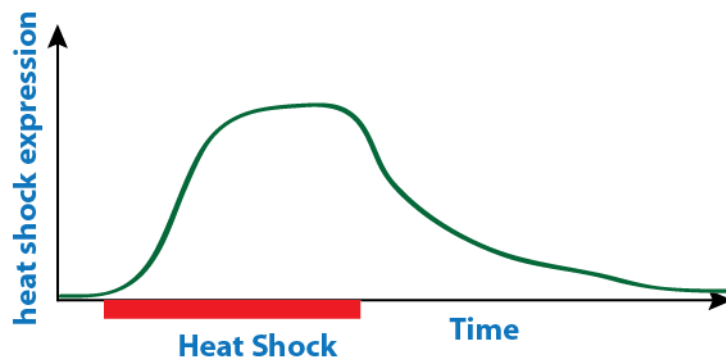


Figure 54 [SGF-1160]. Time course of heat shock induction of gene expression.

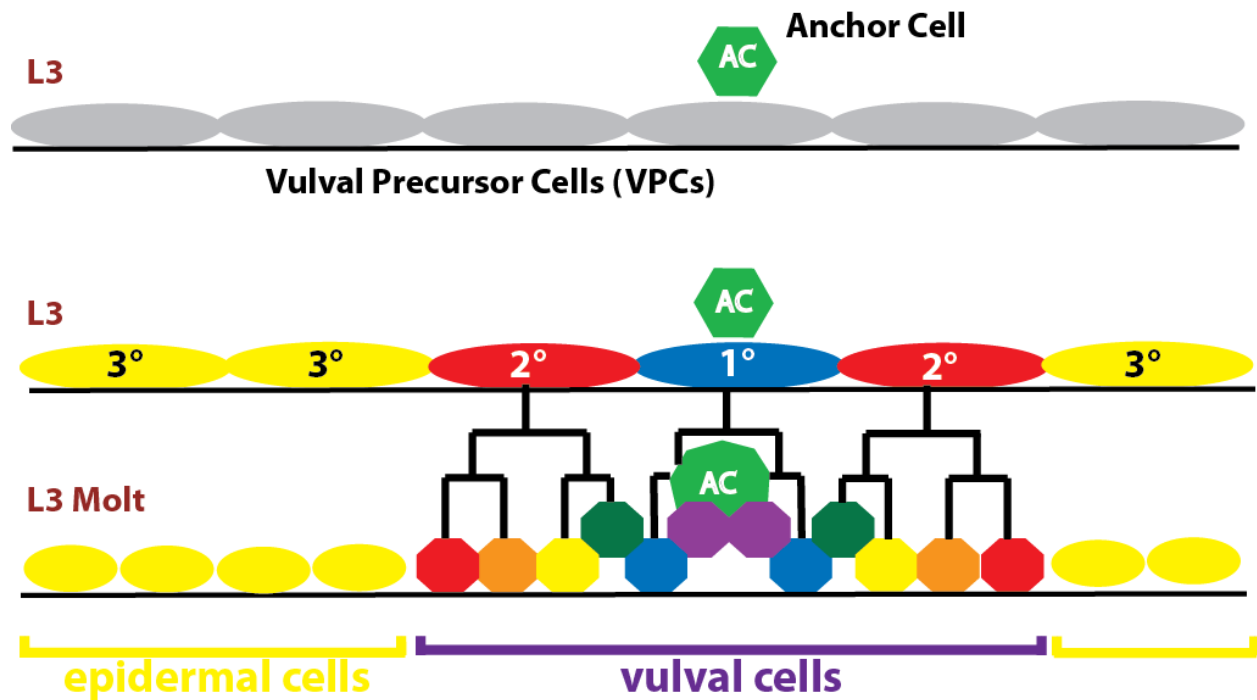


Figure 55 [SGF-1149]. Wild-type vulval Induction

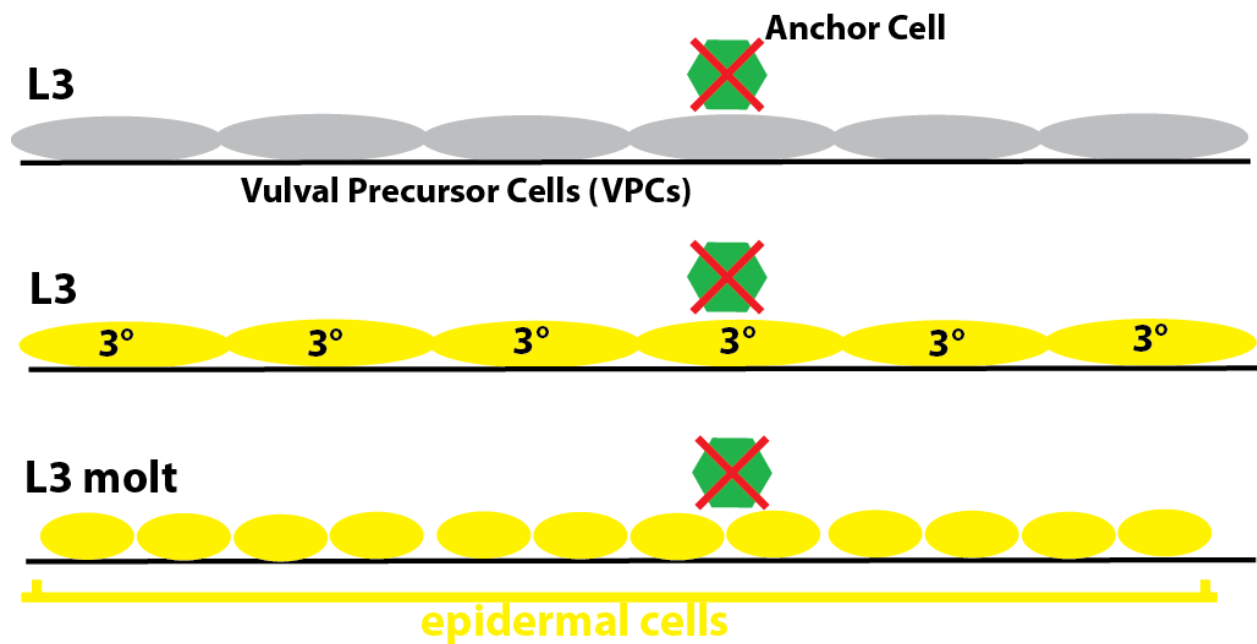


Figure 5- [SGF-1228]

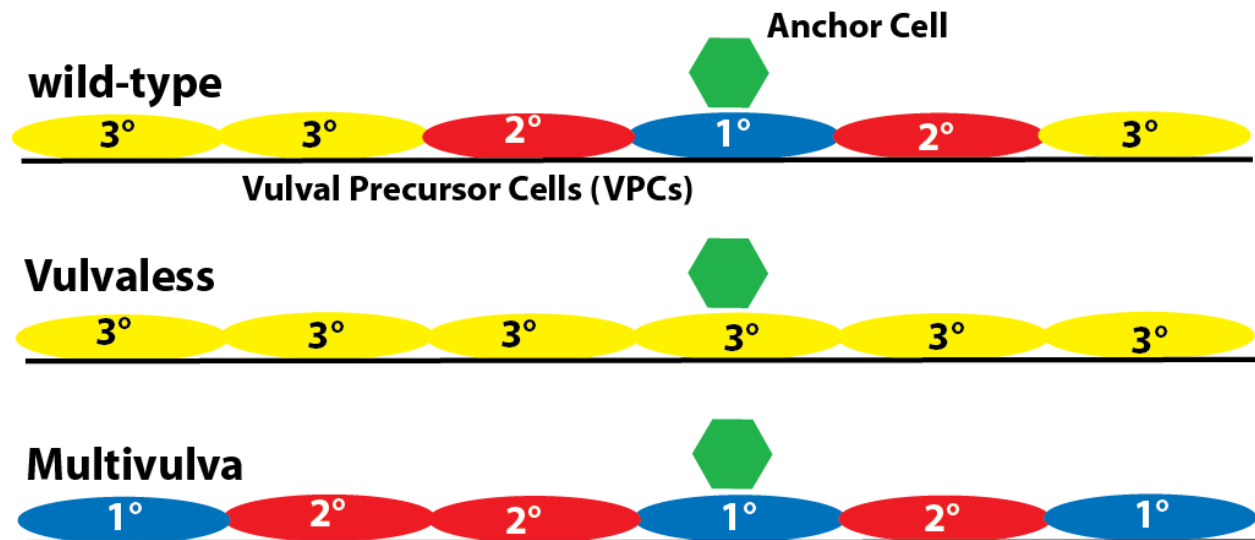


Figure 57 [SGF-1229]. Vulvaless and Multivulva Phenotypes. Yellow, non-induced epidermal fates; blue and red, induced vulval fates. Vulvaless, *let-23(rf)*, *let-60(rf)* and *lin-3(rf)*. Multivulva, *let-23(gf)*, *let-60(gf)*.

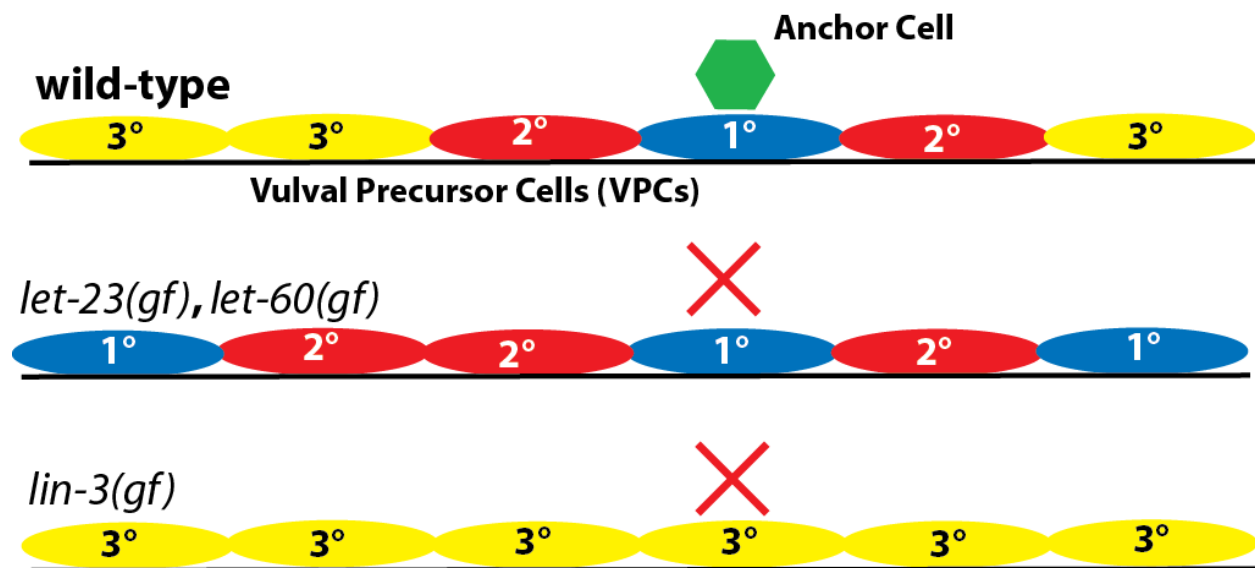


Figure 58 SGF-1230. *lin-3(gf)*, overexpressed transgenic.

BOX 309: RNAi

RNA interference (RNAi) is a broadly applicable technique to reduce gene function based on sequence information. There is no ambiguity over the direction of the effect: gene expression is only reduced, but the extent of reduction can be quite variable and sometimes is not effective.

RNAi is a cellular response to double stranded (ds) RNA in which dsRNA is processed into short interfering RNAs (siRNA), assembled into a ribonuclearprotein complex, and acts catalytically to destroy RNA recognized by the siRNA in the complex. The complex is called the RNA-induced silencing complex (RISC) (Figure SGF-1155). In some cases, such as *C. elegans* and *A. thaliana*, duplex RNA is amplified by RNA-dependent RNA polymerase (RdRP) after the target transcript recognized by siRNA. The details of the mechanisms are different.

dsRNA can be generated by combining sense and anti-sense transcripts, or by transcription of sequences with an inverted repeat (Figure SGF-1154). The dsRNA is cleaved by enzyme Dicer into 21-24 nucleotide siRNA. The siRNA duplex has a two base overhang at the 3' ends.

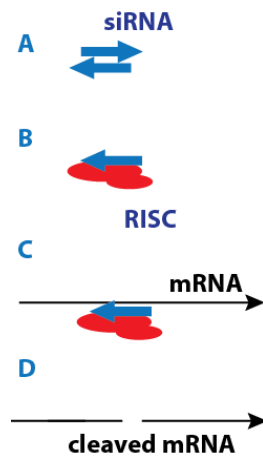


Figure 1155. siRNA targeting of mRNA.

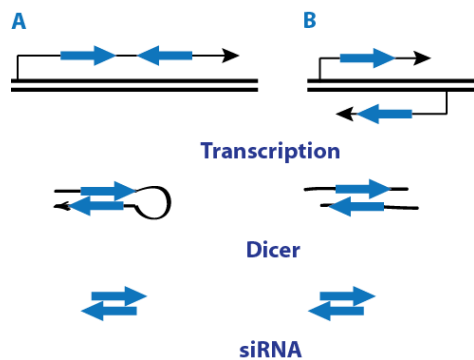


Figure 1154. siRNA Generation. A. Synthesis of a fold-back RNA. B. Synthesis of two complementary RNAs.

BOX 315: Example of allelic series application: *C. elegans let-23*

An allelic series supported the hypothesis that the loss of function phenotype of *let-23* is larval lethal, and was major support for the hypothesis that the vulvaless phenotype of *let-23* is a loss-of-function phenotype, even though the vulvaless alleles are hypomorphs.

Lethality is the likely null phenotype of *let-23* because recessive lethal alleles are the most common. This hypothesis is bolstered by an allelic series argument (Table 420). For example, the hypomorphic allele *sy12* is enhanced by a deficiency: *sy12/sy12* is 19% viable but *sy12/Df* is 0% viable. For lethality, the alleles can be ranked from most to least activities as follows: + > *sy12* > *Df*.

The phenotype of most interest is the vulvaless phenotype. Given that null phenotype is lethal prior to the time of vulval development, how can we determine whether the vulval phenotype is a loss of function phenotype? One solution is a mosaic analysis (Table 419). Another way is by complementation and an allelic series argument (Table 421). *n1045/n1045* has 44% vulval induction, while *n1045/sy12* has 18% and *n1045/Df* has 4.4 % vulval induction. Similarly, *sy1/sy1* has 14% vulval induction and *sy1/Df* has 0% vulval induction. For vulval induction, the alleles can be ranked from most to least activities as follows: + > *n1045* > *sy1* > *Df*.

Table 420. Allelic series can define loss of function phenotype: *let-23* lethality. *n1045*, *sy97* and *sy12* are hypomorphic alleles; *sy15* is a putative amorphic null allele. *mnDf68* is a deficiency for the locus. Data from Aroian and Sternberg (1991 Genetics).

<i>let-23</i> genotype	% viable
+/+	100
+/n1045	100
+/sy97	100
+/sy12	100
+/sy15	100
+/mnDf68	100
n1045/n1045	42
n1045/sy97	37
n1045/sy12	49
n1045/sy15	18
n1045/Df	25
sy97/sy12	19
sy12/sy12	19
sy12/sy15	0.3
sy97/sy97	11
sy12/Df	0
sy15/sy15	0

Table 421. Allelic series defines loss-of-function for one tissue: *let-23* and the vulva.

<i>let-23</i> genotype	% vulval induction
+/+	100
+/n1045	100
+/sy1	100
+/sy97	100
+/sy12	100
+/sy15	100
+/mnDf68	100
n1045/n1045	44
sy1/sy1	14

<i>sy12/sy12</i>	0.6
<i>sy97/sy97</i>	0
<i>sy1/n1045</i>	33
<i>sy1/sy12</i>	6.5
<i>sy1/sy97</i>	36
<i>n1045/sy12</i>	18
<i>n1045/sy97</i>	8.7
<i>sy97/sy12</i>	0
<i>n1045/Df</i>	4.4
<i>n1045/sy15</i>	1.9
<i>sy1/Df</i>	0

BOX 316: TALENs and ZFNs: sequence specific nucleases.

A breakthrough in genome engineering (called **genome editing**) uses engineered site-specific DNA binding proteins that have nuclease activity. The key steps are:

1. introduction of a site-specific nuclease;
2. generation of a double-strand break;
3. repair of the break can introduce a mutation by
 - a. non-homologous end joining, which is error-prone (for example, DNA is chewed back by endonucleases.
 - b. Homology directed repair converts the allele that has been cut (as in **Figure 1091**).

These nucleases use different ways of engineering specificity into the DNA binding.

ZFNs (Zinc-Finger Nucleases) have a DNA-binding domain composed of zinc finger units, each of which binds three nucleotides.

TALENs (Transcription Activator Like Effector Nucleases) have a DNA-binding domain from TAL effector proteins of a bacterial pathogen of plants. A 34 amino acid repeat unit in the protein specifies binding to a single base-pair. There are two amino acids in the repeat that are critical for DNA binding specificity: Asn Gly (NG) recognizes T; His Asp (HD) recognizes C; Asn Ile (NI) recognizes A and AsnAsn (NN) recognizes G or A. The TALEN works as a dimer. The nuclease is Fok1.

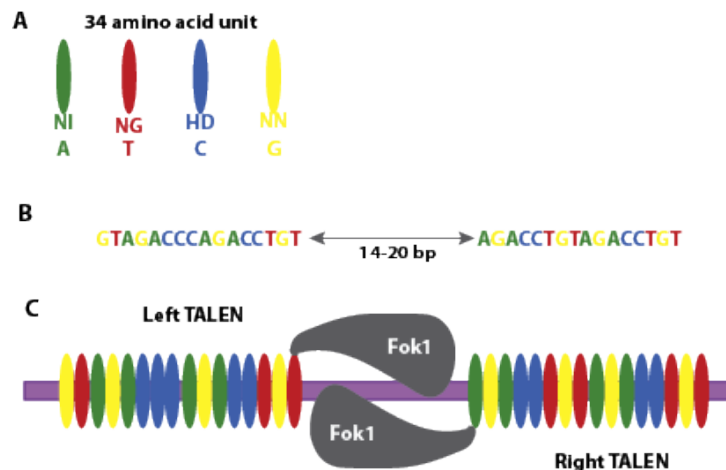


Figure [SGF-1146]. TALENs.

BOX 313: *let-23* and *lin-3* mosaics

Mosaic analysis was carried out for both *let-23* and for *lin-3*. *let-23* is cell autonomous and acts in the vulval precursor cells (VPCs) as opposed to the anchor cell. *lin-3* is cell non-autonomous and acts in the anchor cell (Table 419). For the *let-23* mosaics, *ncl-1(+)* rescue of a *ncl-1(lf)* mutant was used as a cell autonomous marker (Figure SGF-1198).

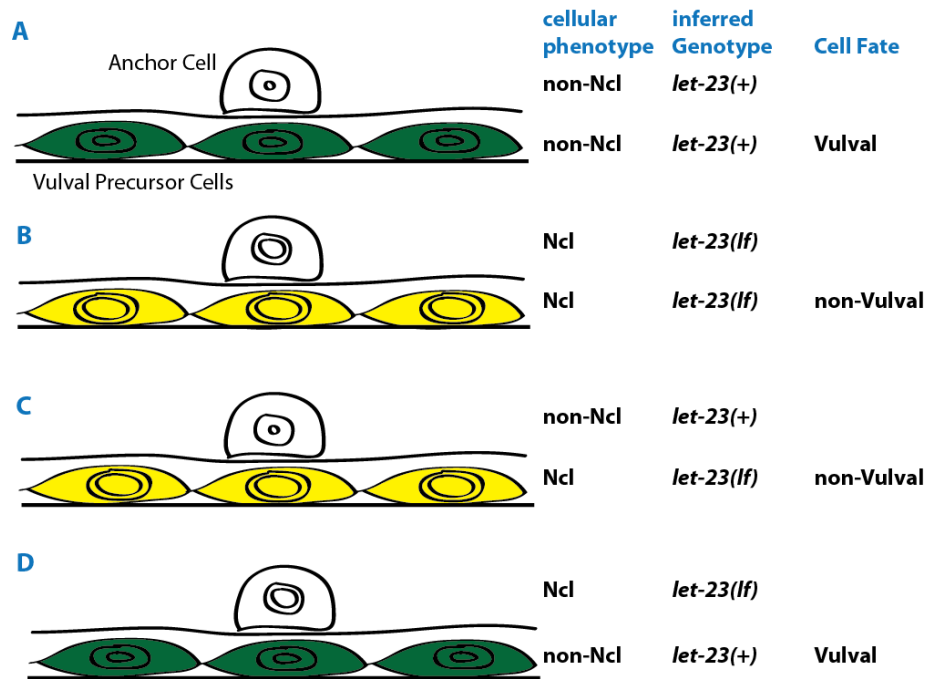


Figure [SGF-1198]. Mosaic analysis of *let-23*.

Table 419. *lin-3* is cell non-autonomous

<i>lin-3</i> Genotype of cell		Phenotype
VPCs	anchor cell	
+	+	wild-type
+	-	vulvaless
-	+	wild-type
-	-	vulvaless

BOX 310: dn and gf RAS in *C. elegans*

Dominant negative alleles can interfere with either activation of protein or its effect. The best example is ras proteins. In *C. elegans*, *let-60* encodes a Ras protein, and there are both gain-of-function and dominant negative alleles (Table 432). The gf alleles are semidominant for a multivulva phenotype. The dominant negative alleles are dominant for a vulvaless phenotype. The heterozygote carrying both a gf and a dominant negative allele is multivulva. This observation suggests that the gf encoded protein is insensitive to the dn encoded protein.

There is a nice molecular explanation for this interaction (Figure SGF-1157). The RAS protein undergoes a typical cycle of guanine nucleotide exchange and hydrolysis (Fig. SGF-1157A). An Exchange Factor is necessary for RAS activation as it catalyzes the exchange of GTP for GDP by helping release GDP from the inactive RAS•GDP complex (Figure SGF-1157B). The dominant negative RAS alleles bind to exchange factor unproductively, thereby preventing it from acting on wild-type RAS (Figure SGF-1157C).

Table 432. Interaction of *let-60* gf and dn alleles

<i>let-60</i> genotype	Phenotype
+/+	Wild type
gf/gf	Multivulva
gf/+	Multivulva
dn/dn	Vulvaless
dn/+	Vulvaless
dn/gf	Multivulva

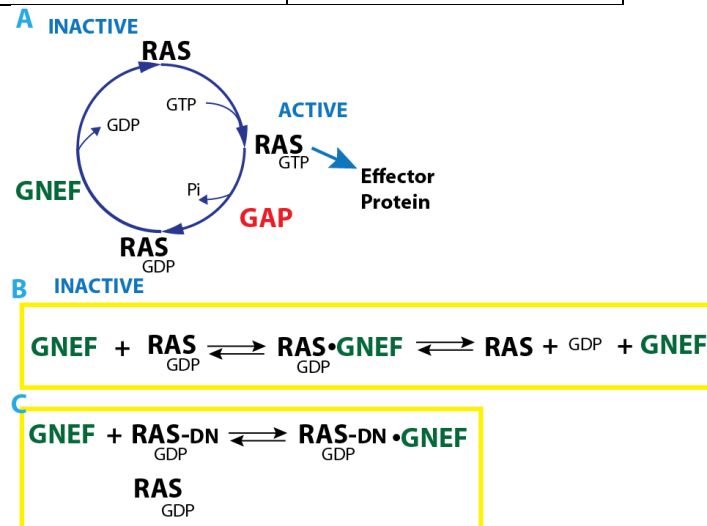


Figure SGF-1157. A. RAS GTPase cycle. B. Activation reaction for wild-type RAS. C. Reaction for dominant negative (DN) RAS.

BOX 317: Genetic definition of deficiencies.

Deficiencies and deletions. A genetic deficiency lacks function for one or more genes. In most cases, a deficiency is caused by deletion of the DNA corresponding to those genes. However, a deficiency might have other causes, such as mutation or rearrangement that causes lack of function of multiple genes due to chromatin effects, or removal of a positively-acting regulatory element. By using the term “deficiency” we are more conservative about the underlying molecular event.

A caveat to these genetic experiments to examine gene activity is that although loss-of-function alleles are typically recessive to the wild-type allele, this is not always the case. In most cases, having a wild-type copy of the gene is sufficient to make up for the loss of gene activity due to the loss-of-function allele. However, for some genes, the loss-of-function allele is dominant because the cell is particularly sensitive to the dosage of the gene, and thus halving gene dosage causes defects. Such dominant or semi-dominant loss-of-function mutations are referred to as **haploinsufficient**. Haplo-insufficiency is discussed further when we discuss manipulating the dosage of genes within cells.

The definition of a genetic Deficiency (*Df*) is its failure to complement multiple alleles that themselves complement. These *Df*s are typically due to a deletion of a region of the chromosome. A *Df* is shown in A by a deletion of part of a chromosome. The *Df* complements *a1*, *d1* and *e1* (panels B, E and F, respectively) but fails to complement *b1* and *c1* (panels C and D, respectively). However, *b1* and *c1* complement (panel G), indicating that the *Df* is deficient in loci *B* and *C* ([Figure SGF-1158]).

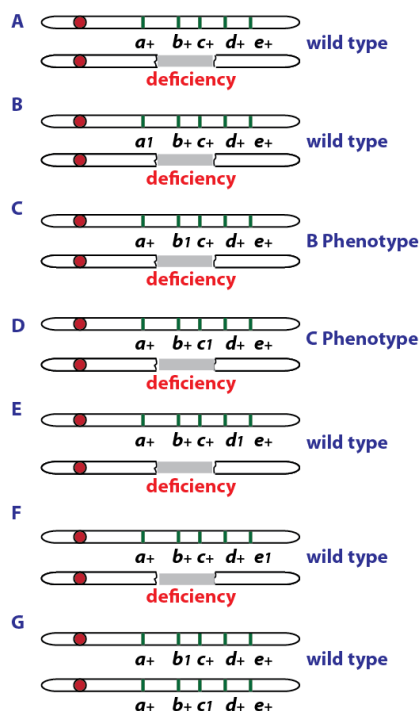


Figure SGF-1158. Genetic definition of a deficiency by complementation analysis.