The relationship between mutation and phenotype

Introduction

Initially, most genes were identified by the abnormalities produced when the gene was mutated. This is the classical or forward genetic approach- identifying genes responsible for a mutant phenotype. Conversely, in reverse genetics, a specific gene or gene product is disrupted or modified and then the phenotype is measured. This approach was only possible after whole genome sequencing had started and the gene sequences of a given organism was available, such that genes could be easily identified by their sequence, cloned or mutated.

Here, we will focus on how forward-genetic screens helps to understand how mutations can cause a phenotype using yeast as a model. Therefore, we will first review the types of mutations that can occur and become familiar with the terms geneticist use to describe the effect of mutations on gene function. Later, we will discuss the strategies applied to perform genetic screens in yeast and will see how these strategies have served to identify genes involved in fundamental biological processes.

Types of mutations

Mutations can be classified in different ways. First, mutations can be defined based on the effect on the DNA sequence. If mutations affect only a single base pair in the genome, this is called a point mutation. This can be either a nucleotide exchange (e.g., A instead of G) or a deletion or insertion of a nucleotide (e.g., a missing A or an additional G). Mutations can also affect larger pieces of DNA or even whole genes. Examples are deletions, inversions or translocations (see figure 2-1).

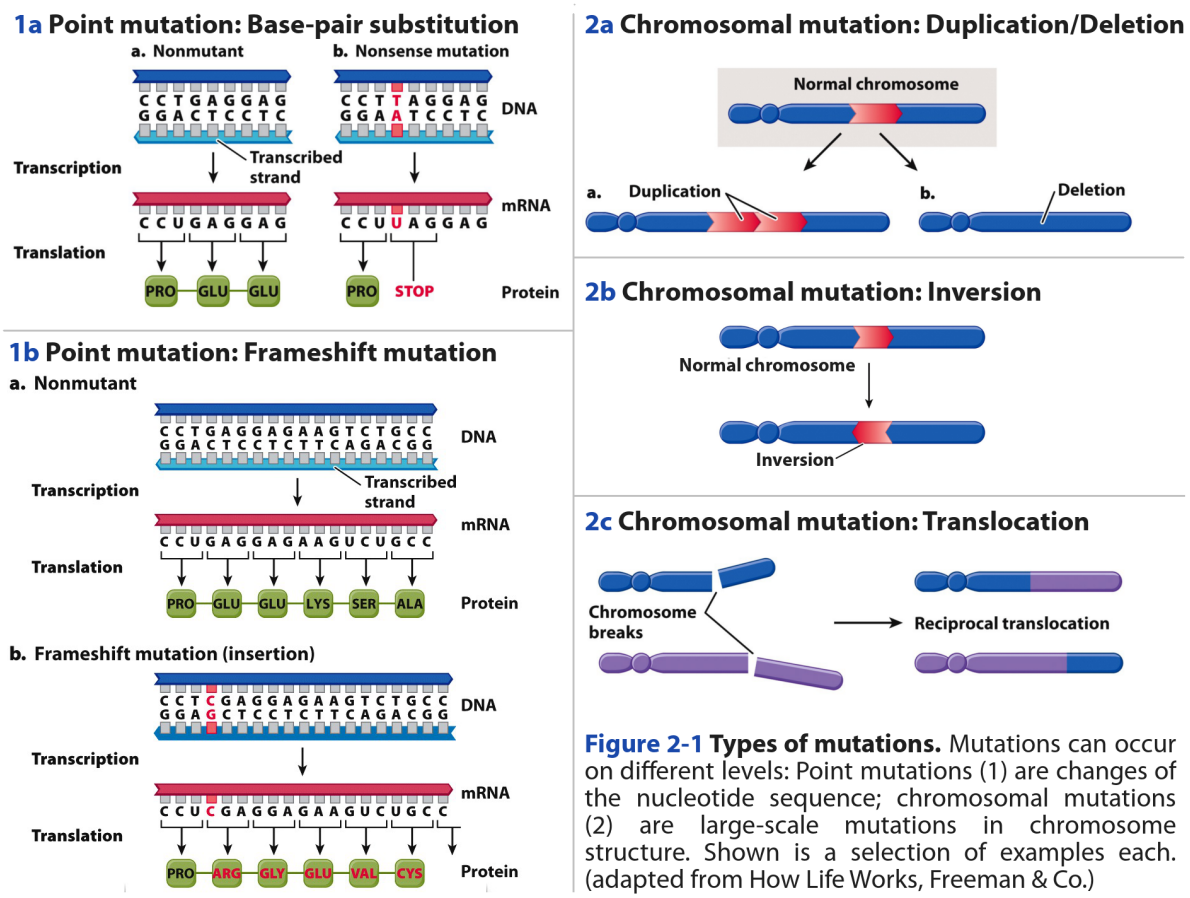

Figure 2-1 Types of mutations. Mutations can occur on different levels: Point mutations (1) are changes of the nucleotide sequence; chromosomal mutations (2) are large-scale mutations in chromosome structure. Shown is a selection of examples each. (adapted from How Life Works, Freeman & Co.)

Second, we can define mutations based on their effect on the function of a gene. For a genetic

screen, only those mutations are of interest because they are likely to cause a detectable phenotype. Geneticists use the following terms to describe mutations on a functional level:

Conditional mutation

A conditional mutation produces its phenotypic effect only under certain conditions, called the restrictive conditions. Under other (permissive) conditions, the effect is not seen. For a temperature-sensitive mutation, the restrictive condition is typically high temperature, while at low temperature, the mutation does not show an effect. Temperature-sensitive mutations generally destabilize the fold of the mutated protein, such that it loses function when exposed to higher thermal vibration (see below). Other conditional mutations can be mutations required for the synthesis of amino acids. In that case, the mutation has no consequence as long as the medium is a source of the amino acid. However, it becomes lethal when the amino acid is depleted from the medium. Such mutations are called auxotrophy mutations, meaning that unlike a prototroph, which does not need much more than a carbon and a nitrogen source (for example sugar and ammonium) to synthesize all other components needed to grow, these mutations create an auxotroph strain, i.e., a strain that requires the addition of at least one additional nutrient. Such mutations are not limited to amino acid auxotrophy, but exist also for strains that cannot synthesize the bases for DNA and RNA synthesis, or key lipids. For example, a strain that cannot synthesize uracil will be called uracil auxotroph. In the wild, yeast strains are generally complete prototrophs.

Loss-of-function mutation

A loss-of-function mutation either reduces or abolishes the activity of the gene. These are the most common class of mutations. Loss-of-function mutations are usually recessive, meaning that the (diploid) organism can usually function normally as long as it retains at least one normal copy of the respective gene. This situation is called heterozygous. If a diploid is homozygous for a loss-of-function allele it normally shows the same phenotype as the haploid mutant.

Null mutation

A null mutation a loss-of-function mutation that completely abolishes the activity of a gene.

Gain-of-function mutation

A gain-of-function mutation increases the activity of the gene or makes it active in appropriate or inappropriate circumstances; these mutations are usually dominant. We call this a positive dominant allele. The phenotype caused by the positive dominant mutation is identical in the haploid and in the heterozygous diploid (with the other allele corresponding to the wild type version of the gene) and is distinct (generally opposite) from the phenotype caused by the null mutation of the gene in haploids and homozygous diploids.

Dominant-negative mutation

A dominant-negative mutation is a dominant-acting mutation that blocks gene activity, causing a loss-of-function phenotype even in the presence of a normal copy of the gene. This occurs when the mutant gene product interferes with the function of the normal gene product. In this case, the mutation gives the same phenotype as that of a null mutation in the haploid, and this phenotype is still observed in the heterozygous diploid (one allele carrying the dominant mutation and the other corresponding to the wild type version of the gene).

Suppressor mutation

A suppressor mutation suppresses the phenotypic effect of another mutation so that the double mutant seems normal. This can be due to a second mutation in the gene that is affected by the first mutation or due to a mutation in another gene, whose product interacts with the product of the first gene.

All these types of mutations can, in principle, be due to any of the mentioned molecular changes (point mutation, deletion, translocation or inversion). However, deletions often cause lethal or null mutations, whereas point mutations are the main cause for conditional, dominant-negative or suppressor mutations.

How mutations can cause a phenotype

Now that we know the nomenclature of phenotypic effects of mutations, let's look at how mutations can cause a phenotype on a molecular level. Here are two examples:

Temperature sensitive (ts) mutations

Ts mutations are functional at low (permissive) temperatures, yet nonfunctional at high (restrictive) temperatures, and thus a rise in temperature quickly abolishes protein function. The mutant phenotype is usually due to a destabilization of the protein and the subsequent loss of protein function. This could be due to decreased melting temperature from the loss of a hydrophobic amino acid or due to a decreased ability to interact with DNA or other proteins, because the mutation removed an amino acid essential for interaction. Eventually, the destabilized or misfolded protein is degraded by the cellular unfolded protein response machinery.

Dominant-negative (dn) mutations

If a *dn* mutation affects a transcriptional activator, the altered protein retains DNA binding activity, but lacks the ability to transactivate. It can complex with the DNA binding sites and displace the WT protein.

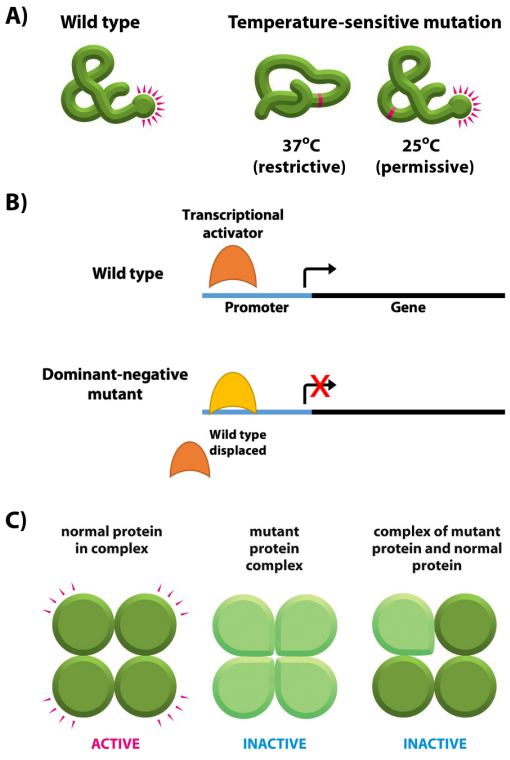

Figure 2-2 Examples of temperature-sensitive (A) and dominant-negative (B and C) mutations. **(A)** Temperature-sensitive mutants are functional at low (permissive), but nonfuncational at high (restrictive) temperatures. **(B)** A dominant-negative mutation in a transcriptional activator might abolish its transactivation activity, but allow it to complex with the promoter, such that the wild type protein is permanently displaced. **(C)** A dominant-negative mutation in a subunit of an oligomeric protein complex might render the whole complex functionally inactive. (A and C: adapted from Molecular Biology of the Cell, Alberts, 6th edition, Garland Science)

Gene interactions: The concept of epistasis

Genes rarely operate in isolation from each other. Epistasis is the term that refers to the action of one gene upon another to affect a phenotype. Gene A is epistatic over gene B if the phenotype of the double mutant is the same as the phenotype of mutants for gene A.

Epistasis occurs in the following scenarios:

- Whenever two or more loci interact to create new phenotypes
- Whenever an allele at one locus masks the effects of alleles at one or more other loci
- Whenever an allele at one locus modifies the effects of alleles at one or more other loci

Epistasis can be seen on genes that act in the same pathway, but also for genes acting in very different biological processes.

One example for the first scenario are genes implicated in the adenine synthesis pathway in yeast. Mutations in *ade2* result in cells unable to synthesize adenine (they are adenine auxotrophs). Mutations in one gene, *ade2*, show a specific phenotype: *ade2* mutant these cells accumulate a red pigment (the substance AIR) and appear as red colonies, while wild-type colonies are white. AIR accumulates because the Ade2 enzyme is normally required to carboxylate AIR on the way to adenine biosynthesis (see figure 2-3). However, this red color phenotype can be reverted by an additional mutation in the *ade3* gene. Both the *ade2*/*ade3* double mutants as well as *ade3* single mutants form wild-type like white colonies. Thus, the *ade3* mutation suppresses the red color phenotype of *ade2*. In genetics, we say that *ade3* is epistatic to *ade2* (see figure 2-3). The reason why the *ade3* mutation does not lead to the formation of red colonies even if combined with the *ade2* mutation is because the Ade3 enzyme is acting very upstream in the adenine biogenesis pathway and is therefore required for the formation of AIR. Thus, in this case, the *ade3* gene is epistatic over the *ade2* gene, which means that the Ade3 protein acts before Ade2 in the same biogenesis pathway. In this situation, once a strain has lost the *ade3* gene, the *ade2* gene becomes irrelevant (the substrate for the Ade2 protein is anyhow not present) and can be mutated without the mutation having any phenotypic consequence.

The shift in color from red to white (or vice versa) is often used as a non-selective phenotype to visualize the occurrence of various genetic events (e.g., to determine the efficiency of plasmid transformations or recombination), as we will see later.

Importantly, the epistatic interaction between *ade2* and *ade3* only acts to reverse the red color phenotype. The adenine auxotrophy is not rescued in *ade2*/*ade3* double mutants, because both genes act in the same biosynthetic pathway to produce adenine.

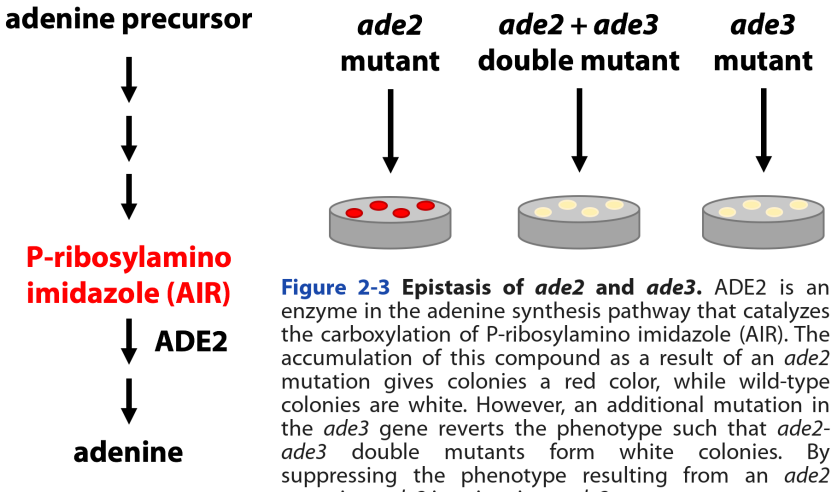

Figure 2-3 Epistasis of *ade2* and *ade3*. ADE2 is an enzyme in the adenine synthesis pathway that catalyzes the carboxylation of P-ribosylamino imidazole (AIR). The accumulation of this compound as a result of an *ade2* mutation gives colonies a red color, while wild-type colonies are white. However, an additional mutation in the *ade3* gene reverts the phenotype such that *ade2*-*ade3* double mutants form white colonies. By suppressing the phenotype resulting from an *ade2* mutation, *ade3* is epistatic to *ade2*.

An example where epistasis occurs between genes of different pathways is the interaction between genes acting in the secretory pathway and the ER-associated degradation pathway (ERAD). Randy Schekmann and colleagues searched for genes associated with the secretory pathway in yeast. Since secretion mutations would be lethal, they developed a screening procedure to identify temperature-sensitive (TS) mutants that accumulate secretory vesicles intracellularly. We will discuss the approach of this fundamental screen later on. Protein secretion starts with protein translocation into the endoplasmic reticulum (ER) where secretory proteins mature to adapt a functional three-dimensional conformation before they are packaged into ER-to-Golgi transport vesicles. Proteins that fail to fold in the ER are not allowed to enter these vesicles, and are initially retained in the ER. Most of these misfolded proteins are subsequently exported to the cytosol and degraded by proteasomes, a process called ER-associated degradation (ERAD).

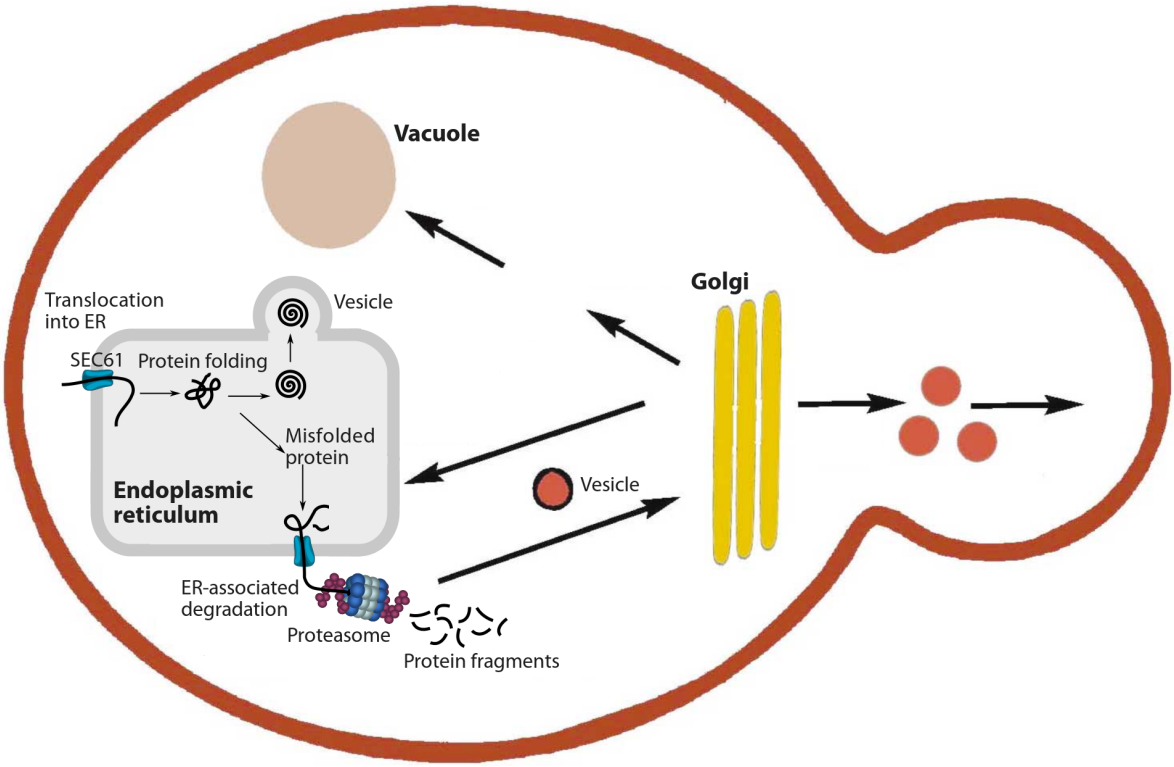

Figure 2-4 Schematic representation of the yeast secretory pathway. Secretory proteins enter the endoplasmic reticulum (ER) by translocation through the SEC61 translocation channel. The accumulation of misfolded proteins leads to the induction of the ER-associated degradation (ERAD). Correctly folded proteins are transported to the Golgi in a vesicle for further processing including additional glycosylation. Proteins exiting the Golgi may be secreted extracellularly or targeted to vacuoles for storage or degradation. (adapted from figures by Randy Schekman and from Wikipedia: ERAD)

In their screen, Randy Schekmann and colleagues identified a temperature-sensitive mutation in the gene *SEC61* (the *sec61-1* temperature-sensitive allele) that encodes the ER-embedded translocation channel mediating protein import into the ER (see figure X). At 37°C, part of the Sec61 protein is misfolded, although the full protein still shows activity. Nevertheless, the misfolded domain of the protein is recognized and leads the degradation of the entire protein by the ERAD machinery. Thus, *sec61* mutant yeast cells do not grow at 37°C. In this case, it is not the mutation per se (the misfolded protein would still function, although at a limited rate), but the degradation of the misfolded protein that eventually kills the cell. Interestingly, the viability of *sec61-1* mutant cells could be restored by introducing a secondary mutation that disrupts *UBC6*, a gene encoding an ubiquitin-conjugating enzyme that functions in the ERAD pathway. As a consequence of this second mutation, the Sec61-1 protein is no-longer degraded and can fulfill its function. Therefore, secondary mutations in ERAD

suppress the phenotype of the *sec61-1* mutant strains. This observation has allowed screening for additional mutations affecting the ERAD pathway and has led to the identification of many genes encoding components of the ERAD machinery. Thus, here we can say that the lethality due to the *sec61-1* mutation at 37°C is not caused by the fact that the mutation would inactivate the Sec61 protein but to how the cells react to the change in Sec61 organization, a change that they perceive as a stress.

Gene interactions can lead to lethality

An extreme case of how genes interact to produce a phenotype is called synthetic lethality. Synthetic lethality arises when the combination of mutations in two or more genes leads to cell death, whereas a mutation in only one of these genes does not. In a study, one can use a mutation that confers a certain phenotype, but does not kill the cell, and systematically test other mutations to determine which ones confer lethality. Thus, synthetic lethality indicates functional relationships between genes.

Synthetic lethality is a consequence of the tendency of organisms to maintain buffering schemes that allow phenotypic stability despite genetic variation, environmental changes and random events such as mutations. This robustness is the result of parallel redundant pathways that mask the effects of mutations so that important cellular processes do not depend on any individual component.

Synthetic lethality can reveal interactions between genes that function in the same biochemical process or between genes that act in pathways that appear to be unrelated, but affect the same ultimate process. In the latter, the interacting genes exhibit overlapping functions in the cell. One example for such a case is arginine availability. Arginine can either be imported through the CAN1 transporter, or synthesized by a set of enzymes *arg* genes). The CAN1 transporter received its name from the fact that it also transports the arginine analog canavanine, a component that is toxic for the cell. Thus, mutations in CAN1 render cells canavanine-resistant. The resistance to canavanine of CAN1 mutant cells is often used for selection in genetic analyses, as we will see later. CAN1 mutant cells can survive as long as they can synthesize arginine. Mutations in any of the *arg* genes render cells auxotrophic for arginine, thus, these cells rely on the import of arginine from the medium to survive. However, when a mutation in any of the *arg* genes is combined with the *can1* mutation, this leads to lethality even in medium containing arginine, since both pathways that assure arginine availability are disrupted.

Summary

In this lesson, we have seen how mutations can generate a phenotype and how genes can interact to affect a phenotype. We have also learned that it is not always the mutation per se that causes the phenotype. We will have to keep these findings in mind when moving on to study the methods used for isolation and identification of mutants in a genetic screen.