5 Molecular Genetic Techniques

Review the Concepts

1. A recessive mutation must be present in both alleles of a diploid organism in order for the mutant phenotype to be observed: that is, the individual must be homozygous for the mutation for the mutation to be expressed. Recessive alleles usually result from a mutation that inactivates the affected gene, leading to *loss of function*. If the inactivated gene is an essential gene, loss of function would be lethal in homozygotes with both mutated alleles. Lethal recessive mutations can be maintained in heterozygotes. In contrast, a dominant mutation produces a mutant phenotype even in the presence of one mutant and one wild type allele. Dominant alleles often result from a mutation that causes some kind of *gain of function*. Dominant alleles that affect the function of essential genes can be lethal even in heterozygotes. Thus researchers may use conditional mutations such as temperature-sensitive mutations to study the effects of dominant lethal alleles.

2. A temperature-sensitive mutation is a mutation that is expressed conditionally, i.e., only at certain temperatures. For example, the wild-type phenotype would be expressed at the permissive temperature, whereas the mutant phenotype would be expressed at the nonpermissive temperature. Presumably, a conformational change occurs in the protein at the nonpermissive temperature that results in the mutant phenotype. A temperature-sensitive mutation is particularly useful for identifying and studying genes that are essential for survival.

3. Complementation analysis can be used to determine whether two recessive mutations are present in the same or different genes. If a heterozygous organism containing both mutations shows the mutant phenotype, then the two mutations are in the same gene because neither allele provides a functional copy of the gene. In contrast, if a heterozygous organism shows a wild type phenotype, then the two mutations are in different genes because a wild type allele of each gene is present. Dominant mutations cannot be tested by complementation analysis because they will display a mutant phenotype even in the presence of a wild-type allele of the gene.

4. A suppressor mutation is a compensatory mutation in another or the same gene, which leads to suppression of the original mutant phenotype. A synthetic lethal mutation is a mutation that enhances rather than suppresses the deleterious effect of another mutation.

5. Bacteria that synthesize restriction enzymes also synthesize a DNA modifying enzyme to protect its own DNA. The modifying enzyme is a methylase, which methylates the host DNA. Methylated DNA is no longer a substrate for the encoded restriction enzyme. Restriction enzyme sites commonly consist of 4–8 base pair, palindromic sequences. After being cut with a restriction enzyme, the ends of the cut DNA molecule can exist as single-stranded tails (sticky ends) with either 5′ or 3′ overhangs, or as blunt (flush) ends. DNA ligase is an enzyme that catalyzes the reformation of the phosphodiester bond between nucleotides in the presence of ATP.

6. A plasmid is a circular, extrachromosomal DNA molecule that contains an origin of replication, a marker gene that permits selection, and a region into which foreign DNA can be inserted (cloning site). A plasmid is useful for cloning DNA fragments up to approximately 20 kb. Specialized plasmid vectors such as BACs (bacterial artificial chromosomes) have been developed that can accommodate DNA fragments as large as several million nucleotides.

7. DNA libraries are collections of randomly cloned DNA fragments. A cDNA (complementary DNA) library is a collection of DNA molecules that are copied from messenger RNA molecules using the enzyme reverse transcriptase. A cDNA

library does not contain every gene, only those that are expressed as mRNA at the time of RNA isolation. In contrast, a genomic DNA library consists of random fragments of the total genome. This would include not only genes but also areas of the genome that do not encode for genes. For hybridization screening, the *E. coli* clones are plated in petri dishes and allowed to grow. The resultant colonies are transferred to nitrocellulose filters where the cells are lysed and the DNA fixed to the filter. A radioactive probe is then hybridized to the filters and the location of plasmid DNA that hybridizes to the radiaoactive probe is determined by autoradiography. For expression screening, the protein encoded in the cloned DNA is expressed in the bacteria, which is then fixed to the membrane. The presence of the protein can be detected by the presence of an antibody or an activity stain. Because of the degeneracy of the genetic code there is some ambiguity in the codons used to direct the amino acid sequence. For the peptide Met-Pro-Glu-Phe-Tyr, the nucleotide sequence could be 5′-ATG, CC(A, T, G or C), GA(A or G), TT(T or C), and TA(T or C)-3′ for Met, Pro, Glu, Phe, and Tyr, respectively. Thus there would be 32 possible 15 nucleotide sequences that could encode for the peptide Met-Pro-Glu-Phe-Tyr.

8. The PCR reaction is performed as multiple cycles of a three-step process. The first step involves heat denaturation of a target DNA molecule. The second step involves cooling the DNA solution to allow annealing of short single-stranded oligonucleotide primers that are complementary to the target DNA molecule. In the final step, the hybridized oligonucleotides serve as primers for DNA synthesis. The resultant double-stranded DNA molecules are then subjected to further rounds of denaturation, annealing, and DNA synthesis (extension). A thermostable DNA polymerase was essential for automation of the PCR process. A nonthermostable DNA polymerase would be inactivated by heat denaturation during each cycle of the PCR process and would necessitate the addition of new enzyme prior to each DNA synthesis step.

9. Southern blotting is a technique in which DNA fragments are separated by size in a gel and then transferred to a solid support such as a nitrocellulose or nylon membrane. The DNA is fixed to the nylon membrane and hybridized to a labeled DNA or RNA probe. The hybridized probe is then detected by some technique such as autoradiography. Northern blotting is similar to Southern blotting, except that RNA instead of DNA is denatured and then separated on the gel. Southern blotting can be used to identify a DNA fragment that contains a DNA sequence of interest. Northern blotting can be used to determine the steady-state levels of a specific RNA.

10. In order to express a foreign gene, a recombinant plasmid would require a promoter for efficient transcription of the foreign gene. A promoter that is inducible would provide even higher expression levels of the foreign gene product. To facilitate purification of the foreign protein, a molecular tag can be added to the recombinant protein. An example of this type of molecular tag is a short sequence of histidine residues (a polyhistidine sequence). The resultant His-tagged protein will bind specifically to a bead that has bound nickel atoms. Other proteins can be washed out and the His-tagged protein can be released from the nickel atoms by lowering the pH of the solution. Bacterial cells are limited in their capacity to synthesize complex proteins because of their inability to perform many post-translational modifications, such as glycosylation, that mammalian cells can perform. These post-translational modifications are essential for the biological activity of the recombinant protein.

11. A DNA microarray consists of hundreds or thousands of individual, closely packed gene-specific sequences attached to the surface of a glass microscope slide. The expression of each of these genes can be analyzed globally following hybridization of the array with labeled cDNA prepared from RNA. Microarrays allow the simultaneous analysis of the expression of thousands of genes, whereas Northern blotting allows the analysis of gene expression a single gene at a time. Because Northern blotting involves the separation of RNA on a gel, the different sizes of a given mRNA can be observed, whereas microarray analysis does not allow the examination of different mRNA sizes.

12. The expression of mRNA in individual cells can be determined by in situ hybridization in whole cells or tissue sections. Fixed cells are exposed to

labeled DNA probes that are complementary to the mRNA of interest. After washing to remove excess probe, the cells can be examined microscopically to detect the locations of labeled mRNA. This process can also be used to identify mRNA locations in embryos.

13. Restriction fragment length polymorphisms (RFLPs) result from mutations that create or destroy restriction enzyme sites. As a result, DNA molecules with or without the restriction enzyme sites are cleaved into different sized fragments. Single nucleotide polymorphisms (SNPs) are changes in a single nucleotide between two individuals. Simple sequence repeats (SSRs), also known as microsatellites, consist of a variable number of repeating one-, two-, or three-base sequences. The number of these repeat units at a specific genetic locus varies between individuals. All of these types of polymorphisms can be used as molecular markers for mapping studies. The recombination frequency between two polymorphisms can be determined and can serve as the basis for development of a genetic map. In general, the farther two markers are separated on a chromosome, the greater the recombination frequency between those two markers, and vice versa.

14. Linkage disequilibrium mapping can sometimes be used in cases where a genetic disease commonly found in a particular population results from a single mutation that occurred many generations in the past. In such cases, most of the individuals with the disease would have inherited the disease from the same ancestral chromosome. The closer genetic markers are to each other, the less likely they will be recombined by crossing over during meiosis. Thus, individuals inherit sections of DNA from their parents, not just individual genes. DNA polymorphisms on part of a chromosome that are inherited together are called haplotypes. If geneticists can identify a haplotype common to all the affected individuals in a particular population, DNA markers associated with the disease haplotype might help localize the disease-associated gene to a relatively small chromosomal region.

15. Once a gene is roughly located along a chromosome by genetic linkage studies, further analysis is required to identify the "disease" gene. One strategy for identifying a disease gene involves gene expression analysis. Comparison of gene expression in tissues from normal and affected individuals by Northern blot analysis may indicate a gene that is involved in the disease process. Northern blot analysis allows a comparison of both the level of expression and the size of the transcripts between normal and disease tissues. Sometimes expression levels and/or size of the transcripts do not alter between the normal and the disease states. In this case, DNA sequencing of a potential disease gene from tissues of a normal and a disease state could reveal a single nucleotide change that results in the disease phenotype.

16. Dicer is an RNA endonuclease that cleaves double-stranded RNA into short (23-nt) double-stranded segments known as small inhibitory RNA (siRNA). RISC (RNA induced silencing complex) is a protein complex that cleaves or blocks translation of mRNA molecules that match the sequence of the short double-stranded segments produced by Dicer. The normal function of both Dicer and RISC is to allow for gene regulation by small endogenous RNA molecules known as micro RNAs.

17. To generate a knockout mouse, mouse embryonic stem cells are first transfected with a disrupted allele of the target gene. Through a process known as homologous recombination the disrupted allele replaces the functional homologous gene in the chromosome, resulting in a nonfunctional chromosomal gene. The ES cells, which now contain a mutant gene, are injected into a blastocyst. The blastocyst is transferred into a recipient mouse. Pups that are born will be chimeras. The *loxP*-Cre system can be used to conditionally knock out a gene. Using the above technology, *loxP* sites can be engineered to flank the gene of interest. Expression of the recombinase, Cre, in a specific tissue will result in loss of the flanked gene in that tissue. Knockout mice serve as models for human diseases. For example, if a human disease is known to result from a mutation in gene X, a knockout mouse can be generated that lacks gene X.

18. A dominant negative mutation is a mutation that produces a mutant phenotype even in cells carrying a wild type copy of the gene. This type

of mutation produces a loss of function phenotype. RNA interference (RNAi) is a method of inactivating gene expression by selectively destroying RNA. In this method, a short double-stranded RNA molecule is introduced into cells. This double stranded RNA base pairs with its target mRNA, promoting degradation of the mRNA by specific nucleases.

Analyze the Data

a. X and Y cells do not grow at the elevated temperature, indicating mutations X and Y are temperature sensitive mutations that are present in essential genes. Furthermore, the X and Y mutations complement each other, as shown by the growth of the X-Y diploid at the elevated temperature. This indicates that mutations X and Y are in different genes. It also tells us that each mutation is recessive, as a single copy of the wild-type gene is sufficient to overcome the growth defect.

b. The plate on the left lacks uracil; clones growing on this plate must be able to synthesize their own uracil (they must have a wild-type copy of the *URA3* gene). The yeast themselves are defective in uracil synthesis, so each clone on the left plate must be transformed with a plasmid containing the *URA3* gene. The single clone on the plate at the right grew in the absence of uracil at the restrictive temperature. Therefore it must harbor a plasmid that contains the wild-type copy of the X gene, allowing mutant X cells to grow at 32° C. If the plasmid were to be re-isolated from this clone and the yeast cDNA insert analyzed and sequenced, the X gene will have been identified.

c. The Southern data reveal that the DNA restriction fragment that contains gene X is larger in mutant X than in the parental strain or in mutant Y, which both contain wild-type copies of gene X. This larger restriction fragment could arise in a number of ways. An alteration in the fragment size could be due to a mutation that affects the sequence of a restriction site. Alternatively, an insertion of DNA sequence into gene X could lead to a larger restriction fragment. However, the PCR data indicate that gene X is the same size in mutant X as it is in the parental strain. Accordingly, these data suggest that the mutation in X is a result of a base change(s) in gene X which, fortuitously, has resulted in the loss of a restriction site.

d. At 32° C, mutant X cells containing the wild-type X-GFP fusion construct do not grow — in other words, presence of X-GFP protein is unable to rescue growth at the restrictive temperature. Even though it is synthesized, X-GFP protein cannot compensate for the mutation in X. The GFP-X construct, on the other hand, is able to rescue growth of X cells at the restrictive temperature. The fluorescent microscopy data show that the GFP-X protein is localized to the nucleus, while the X-GFP protein is localized to the cytoplasm. The presence of GFP at the C terminus of X seems to interfere with protein X's ability to be localized to the nucleus. Furthermore, it appears that X must be localized to the nucleus in order to function.

e. The gene product of X is needed earlier in the process of bud formation (earlier in the cell cycle) than the gene product of Y.