S1.When a cell experiences a change in its environment, it may activate one or more genes as a way to respond to the changes. For example, when confronted with a toxic substance, a cell may turn on genes that encode proteins that can degrade or export the toxin. Let's suppose a researcher is interested in the ability of mouse liver cells to protect themselves against toxic heavy metals, such as mercury. To understand the cellular response, the researcher grows mouse liver cells in the laboratory and divides them into two samples. One of the samples is exposed to mercury while the other is not. Samples of cDNA were made from these cells, and then a subtractive cDNA library was made as described in figure 21.1. In this case, the cDNAs derived from the cells that were not exposed to mercury were the cDNAs bound to the column. The results showed that the subtractive library contained seven different types of cDNAs.

A. What do these results mean?

B. What would you do next, assuming that the entire mouse genome is contained within a database?

Answer:

A. The results mean that seven different genes are turned on when liver cells are exposed to mercury.

B. Take the seven different cDNAs and subject them to DNA sequencing. Next, the cDNA sequences are used (via computer programs) to find their genomic matches. This will make it possible to determine the entire gene sequences. Then take the entire coding sequence, for each of the seven genes, and look for homology between each mouse gene sequence and sequences with a database. In other words, you would do a database search, using each of the seven sequences as the query sequence. If you discovered that any of the seven genes were homologous to other genes whose function is already known, this would give you direct information regarding the probable function of particular genes. For example, if one of the genes turned out to be homologous to a gene in bacteria that encodes a protein that is already known to function in the export of metal ions, this would suggest that one way that mouse liver cells try to avoid the toxic effects of mercury is to try to pump it out.

S2. To answer this question, you will need to look back at the evolution of the globin gene family, which is shown in chapter 8, figure 8.10. Throughout the evolution of this gene family, mutations have occurred, and these mutations have caused the modern-day globin polypeptides to have similar but significantly different amino acid sequences. If we look at the sequence alignment in figure 21.10, we can make logical guesses regarding the timing of mutations, based on a comparison of the amino acid sequences of family members. What is/are the most probable time(s) that mutations occurred to produce the following amino acid differences? Note: You will have to examine the alignment in figure 21.10 and the evolutionary timescale in figure 8.10 to answer this question.

A. Val-111 and Cys-111

B. Met-112 and Leu-112

C. Ser-141, Asn-141, Ile-141, and Thr-141

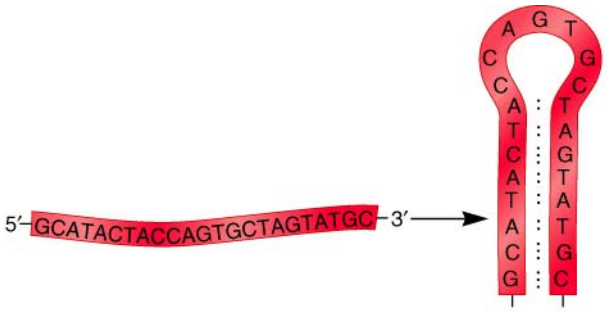
Answer:

A. We do not know if the original globin gene encoded a cysteine or valine at codon 111. The mutation could have changed cysteine to valine or valine to cysteine. The mutation probably occurred after the duplication that produced the α-globin family and *β*-globin family (about 300 million years ago) but before the gene duplications that occurred in the last 200 millions years to produce the multiple copies of the globin genes on chromosome 11 and chromosome 16. Therefore, all of the globin genes on chromosome 11 have a valine at codon 111 while all of the globin genes on chromosome 16 have a cysteine.

B. Met-112 occurs only in the *ε*-globin polypeptide; all of the other globin polypeptides contain a leucine at position 112. Therefore, the primordial globin gene probably contained a leucine codon at position 112. After the gene duplication that produced the
ε-globin gene, a mutation occurred that changed this leucine codon into a methionine codon. This would have occurred since the evolution of primates (i.e., within the last 10 or 20 million years).

C. When we look at the possible codons at position 141 (i.e., Ser-141, Asn-141, Ile-141, and Thr-141), we notice that a serine codon is found in *θ*-globin, ζ-globin, and γ-globin. Since the *θ*- and ζ-globin genes are found on chromosome 16 and the *γ*-globin genes are found on chromosome 11, it is probable that serine is the primordial codon, and that the other codons (asparagine, isoleucine, and threonine) arose later by mutation of the serine codon. If this is correct, the Thr-141 codon arose after the gene duplications that produced the *θ*- and ζ-globin genes. And the Asn-141 and Ile-141 mutations arose after the gene duplications that produced the γ-globin genes. Therefore, the Thr-141, Asn-141, and Ile-141 arose since the evolution of primates (i.e., within the last 10 or 20 million years).

S3. Using a comparative sequence analysis, the secondary structures of rRNAs have been predicted. Among many homologous rRNAs, one stem-loop usually has the following structure:



You have sequenced a homologous rRNA from a new species and have obtained most of its sequence, but you cannot read the last five bases on your sequencing gel.

5′–GCATTCTACCAGTGCTAG?????–3′

Of course, you will eventually repeat this experiment to determine the last five bases. However, before you get around to doing this, what do you expect will be the sequence of the last five bases?

Answer: AATGC–3′

This will also form a similar stem-loop structure.

S4. How can codon bias be used to search for structural genes within uncharacterized genetic sequences?

Answer: Most species exhibit a bias in the codons they use within the coding sequence of structural genes. This causes the base content within coding sequences to differ significantly from that of noncoding DNA regions. By knowing the codon bias for a particular species, researchers can use a computer to locate regions that display this bias and thereby identify what are likely to be the coding regions of structural genes.