7

GENES, CHROMATIN, AND CHROMOSOMES

REVIEW THE CONCEPTS

1. A gene is commonly defined as the entire nucleic acid sequence that is necessary for the synthesis of a functional gene product (RNA or polypeptide). This definition includes introns and the regulatory regions (e.g., promoters and enhancers of the gene).

   (i) a, b

   (ii) a, b

   (iii) b

2. Single or solitary genes are present once in the haploid genome. In multicellular organisms, roughly 25–50% of the protein-coding genes are solitary genes. Gene families are sets of duplicated genes that encode proteins with similar but nonidentical amino acid sequences. An example of a gene family is the β-like globin genes. Pseudogenes are copies of genes that are nonfunctional even though they seem to have the same exon-intron structure as a functional gene. Pseudogenes probably arise from a gene duplication followed by the accumulation of mutations that render the gene nonfunctional. Tandemly repeated genes are present in a head-to-tail array of exact or almost exact copies of genes. Examples of tandemly repeated genes include ribosomal RNAs and RNAs involved in RNA splicing.

3. Satellite, or simple sequence, DNA can be categorized as microsatellite or minisatellite DNA depending upon the size of the repeated DNA sequence. Microsatellites contain repeats that contain 1–13 base pairs. Minisatellites consist of repeating units

of 14–100 base pairs, present in relatively short regions of 1 to 5 kb made up of 20–50 repeat units. The number of copies of the tandemly repeated DNA sequences varies widely between individuals. DNA fingerprinting is a technique that examines the number of repetitive units at a specific genetic locus for several separate loci. This technique can distinguish and identify individuals on the basis of differences in the number of repeats in their micro- and minisatellite simple-sequence DNA.

4. A bacterial insertion sequence, or IS element, is a member of the class of mobile DNA elements. The IS element usually contains inverted repeats at the end of the insertion sequence. Between the inverted repeats is a region that encodes the enzyme transposase. Transposition of the IS element is a three-step process. First, transposase excises the IS element in the donor DNA; second, it makes staggered cuts in a short sequence in the target DNA; and, third, it ligates the 3′ termini of the IS element to the 5′ ends of the cut donor DNA. The final step involves a host-cell DNA polymerase, which fills in the single-stranded gaps, generating 5–11 base pair short direct repeats that flank the IS element before DNA ligase joins the free ends.

5. Retrotransposons transpose through an RNA intermediate. One class of retrotransposons, the LTR retrotransposons, contain long terminal repeats (LTRs) at their ends and a central protein coding region that encodes the enzymes reverse transcriptase and integrase. The retrotransposon is first transcribed into RNA by host RNA polymerase. This RNA intermediate is then converted into DNA by the action of reverse transcriptase primed with a cellular tRNA in the cytoplasm. The double-stranded DNA copy generated is then imported into the nucleus and inserted into chromosomal DNA by the action of an integrase that is similar to the transposases of DNA transposons. Retrotransposons that lack LTRs transpose by a different mechanism. Non-LTR retrotransposons, of which LINES are an example, consist of direct repeats that flank a region that encodes two proteins: ORF1, an RNA-binding protein, and ORF2, which is similar to reverse transcriptase. The LINE element is first transcribed by host RNA polymerase and exported to the cytoplasm, resulting in the translation of ORF1 and ORF2. These proteins bind the LINE RNA and import it into the nucleus, where ORF2 makes staggered nicks in A/T-rich target DNA. The resulting T-rich strand of chromosomal DNA then hybridizes to the poly (A) tail at the 3′-end of the LINE RNA and primes reverse transcription of the RNA by ORF2 protein. The RNA strand of the resulting RNA/DNA hybrid is replaced by DNA and both resulting DNA strands are ligated to the chromosome ends generated by the original ORF2 cut through the action of host-cell enzymes that normally replace the RNA primers of Okazaki fragments and ligate them together during cellular DNA synthesis.

6. Insertion of transposons can generate spontaneous mutations that may influence evolution. In addition, unequal crossing over between homologous mobile elements at different chromosomal locations leads to exon duplications, gene duplications, and chromosomal rearrangements that can generate new combinations of exons. Subsequent divergence of duplicated genes leads to members of gene families with distinct functions. The inclusion of flanking DNA during transposition also results in the movement of genomic DNA to another region of the

genome. This can result in new combinations of exons, an evolutionary process known as exon shuffling, as well as new combinations of transcriptional control regions.

7. A nucleosome consists of a protein core of histones with DNA wound around its surface. The protein core consists of an octamer, containing two copies of histones H2A, H2B, H3, and H4. Approximately 150 base pairs of DNA are wrapped less than two complete turns around the octameric histone core. The histone H1 binds to the linker region, which varies in length from 10 to 90 base pairs and is located between nucleosomes. The nucleosomes are folded into a two-start helix (see Figure 6-30) to form a 30-nm fi ber.

8. Chromatin modifications affect whether the DNA is tightly or loosely compacted. As transcription factors and RNA polymerase must be able to access the DNA, transcription will only occur when the DNA is loosely compacted. When a gene is actively transcribed, the chromatin would likely be acetylated. When a gene is not actively transcribed, the chromatin would likely be methylated or deacetylated.

9. The technique now commonly used for chromosome painting is a variation of fluorescence in situ hybridization (FISH). It use probes specific for sites scattered along the length of each chromosome and labeled with several different fluorescent dyes with distinct excitation and emission wavelengths. The probes are hybridized to chromosomes and the sample is observed with a fluorescence microscope. The information is conveyed to a computer, and a special program assigns a false-color image to each type of chromosome and displays the cell's karyotype.

10. Polytene chromosomes are present in larval salivary glands of the fruit fl y Droso phila melanogaster , and are also present in cells in other dipteran insects and in plants. These enlarged interphase chromosomes, which can be observed with a light microscope, form as a result of multiple rounds of DNA replication (poly-tenization) without chromosome separation or cell division. Polytene chromo-somes consist of multiple gene copies, which when transcribed provide the cells with an abundance of mRNA encoding proteins required for larval growth and development.

11. Replication origins are the points at which DNA synthesis is initiated. The centromere is the region to which the mitotic spindle attaches. The telomeres are specialized structures located at the ends of linear chromosomes. (a) The chromosome would not be capable of being duplicated during S phase. (b) The chromosome could be replicated, but it may not be segregated evenly to the two daughter cells. The centromere is responsible for proper segregation of the duplicated chromosomes; without it the chromosomes will be distributed to the daughter cells by chance.

12. Because DNA polymerase is unable to initiate synthesis of a nucleotide strand, RNA primers must first be introduced to synthesize both leading and lagging strands. Eventually, the RNA region is degraded and DNA polymerase fills in the missing nucleotides. In the case of the lagging strand, DNA polymerase will fill in this gap using an upstream Okazaki fragment. At the very ends of the chromosome, there is no upstream Okazaki fragment, and thus it is not possible to replace the RNA nucleotides with DNA. Thus the chromosome shortens after each round of replication, this being known as the end replication problem. While telomeres do not halt this problem, they ensure that the lost DNA is non-coding DNA. Rather than lose protein or RNA encoding DNA, the genome instead loses non-coding telomere DNA.