Cheat Sheet

Chapter 21 Questions

1. **Who is our closest living relative on the evolutionary tree of life?**
2. What are the scientific names of corn, California two-spot octopus, and elephant shark?
3. **What is genomics?**
4. **What is bioinformatics?**
5. **What is the whole-genome shotgun approach?**
6. **What is metagenomics?**
7. **What is gene annotation?**
8. **What is WD40?**
9. **What does epigenetic mean?**
10. **What is proteomics?**
11. **What is systems biology?**
12. **How many genes do humans have?**
13. **What is the size of most bacterial, archaeal, and eukaryotic genomes?**
14. **How many genes do bacteria, archaea, and eukaryotes have?**
15. **What are pseudogenes?**
16. **What is repetitive DNA?**
17. **What are transposable genetic elements (transposable elements)?**
18. **What are transposons?**
19. **What are retrotransposons?**
20. **What are *Alu elements*?**
21. **What is LINE-1 (L1)?**
22. **What is simple sequence DNA?**
23. **What is a short tandem repeat (STR)?**
24. **How much of the genome do coding genes compose?**
25. **What are multigene families?**
26. **Where are globin gene families found in humans?**
27. **What are the three biggest rRNAs and what are the four rRNAs found in ribosomes?**
28. **How did the human chromosome 2 evolve?**
29. **What can cause the divergences of gene-size regions of DNA?**
30. **What are lysozyme and α-lactalbumin?**
31. **What is collagen?**
32. **What is exon shuffling?**
33. **What is tissue plasminogen activator (TPA)?**
34. **What are genes that have remained similar called?**
35. **What is the *FOXP2* gene?**
36. **What are copy-number variants (CNVs)?**
37. **What is evolutionary developmental biology (evo-devo)?**
38. **What is a homeobox?**

Chapter 21 Answers

1. Chimpanzee (*Pan troglodytes*)
2. Zea mays, Octopus bimaculoides, Callorhinchus milii
3. Study of whole sets of genes and their interactions
4. Application of computational methods to store/analyze biological data
5. Starts with cloning and sequencing of DNA fragments from randomly cut DNA. Powerful computer programs assemble overlapping short sequences into single continuous sequence
6. Where DNA from entire community (metagenome) is collected and sequenced
7. Process by which scientists identify protein-coding genes in a sequence and their functions. Computers scan search for start and stop signals, RNA splicing sites, and sequences that specify known mRNAs (expressed sequence tags or ESTs). Suspected genes compared to known genes from other organism. RNA-seq or similar method used to confirm that relevant RNA is expressed from proposed gene
8. Protein domain in many eukaryotic proteins, known to function in signal transduction pathways.
9. Alters gene expression without changing the DNA sequence (epigenome is epigenetic features of the genome)
10. Study of sets of proteins and their properties (proteome = set of proteins expressed by group of cells)
11. Approach that aims to model dynamic behavior of whole biological systems based on the vast amounts of data generated in these types of studies
12. slightly less than 21000
13. between 1 and 6 Mb, same range as bacteria, larger (most animals/plants have atleast 100 Mb)
14. 1,500-7500 for bacteria/archaea, 5000 in unicellular eukaryotes, atleast 40,000 in some multicellular eukaryotes
15. Former genes that have accumulated mutations and no longer produce functional proteins
16. Sequences that are present in multiple copies in the genome, make up most of DNA between function genes (14%)
17. Stretches of DNA that can move from one location to another within the genome. In transposition, element moves from one site in a cell's DNA to a target site by a recombination process (proteins bring sites together by bending DNA). (75% of repetitive DNA, 44% of genome), considered non-coding even though they code for proteins (do not carry out normal cell functions)
18. Move within genome via DNA intermediate, type of transposable element. Enzyme transposase coded by transposon cuts/paste or copies/pastes transposon
19. Move via RNA intermediate that is a transcript of retrotransposon DNA, always leave copy at original site (RNA intermediate converted to DNA by reverse transcriptase, enzyme encoded by retrotransposon)

20. family of similar sequences in humans and other primates (10% of genome), about 300 nucleotides long, many are transcribed into RNA, some of RNAs thought to regulate gene expression, transposable
21. type of retrotransposon, 17% of genome, 6,500 base pairs, typically low rate of transposition, more active in developing brain
22. DNA containing many copies of tandemly repeated short sequence, 3% of genome, most located at telomeres and centromeres
23. simple sequence DNA where repeating unit contains 2-5 nucleotides
24. 1.5%, 25% including introns and regulatory sequences
25. Collections of two or more identical or very similar genes
26. various types of α-globin found on chromosome 16, ß-globin on chromosome 11. Forms of globins expressed at different times in development to suit conditions
27. 18S, 5.8S, 28S found in same transcriptional unit, primary transcript cleaved to form the three rRNAs. 5S rRNA added to form ribosomes
28. An ancestors 12 and 13 fused end to end (chimpanzees still have these separated)
29. Unequal crossing over, slippage during DNA replication
30. Enzyme that helps protect animals against bacterial infection by hydrolyzing bacterial cell walls

    Nonenzymatic protein that plays a role in milk production in mammals, not present in birds, evolved from duplicated lysozyme gene
31. Structural protein with highly repetitive amino acid sequence
32. mixing and matching of exons within a gene or between different genes, can lead to new proteins
33. Extracellular protein that helps control blood clotting. Four domains of three types (each encoded by exon, one of which is present in two copies)
34. highly conserved
35. product regulates genes that function in vocalization in vertebrates
36. Loci where some individuals have one or multiple copies of a particular gene/region rather than two copies
37. Compares developmental processes of different multicellular organisms.
38. 180-nucleotide sequence in homeotic genes in *Drosophila*, codes for 60-amino-acid homeodomain in encoded proteins, similar sequence found in many animals, *Hox* short for (**h**omeob**ox**-containing genes)