MAT 394 - PS2: Molecular and Human Genetics

1.) Describe three differences between DNA and RNA.

Answer:

(a) The backbone of a DNA molecule includes the sugar deoxyribose, which has one less hydroxyl (-OH) group than the sugar ribose found in RNA.

(b) Uracil (U) occurs in RNA molecules whereas thymine (T) occurs in DNA molecules.

(c) Under physiological conditions, DNA is usually double-stranded, whereas RNA is usually single-stranded.

2.) Briefly explain the central dogma of molecular biology and then give one example that shows that this is not strictly true.

Answer: The central dogma of molecular biology asserts that information flows from DNA through RNA (in the guise of mRNA) to proteins. James Watson, who somewhat casually originated the phrase 'the central dogma', later publicly regretted his choice of words, as the term dogma suggests that this is something more than just a very broadly applicable, but informal rule. In fact, many counterexamples are now known, such as the use of RNA by retroviruses to encode hereditary information, which is then transmitted to progeny viruses via a DNA intermediate within infected host cells.

3.) Approximately how many base pairs in the human genome are contained in the exons of protein-coding genes?

Answer: The human genome contains approximately 3×10^9 base pairs (per haploid complement), of which approximately 2% are contained in exons. Thus the number of such base pairs is approximately 6×10^7 or sixty million.

4.) CFTRΔF508 is a mutation in the CFTR gene that is the leading cause of cystic fibrosis. Identify the location of the CFTR gene in the human genome and describe how ΔF508 alters the protein encoded by this gene. What is the frequency of this mutation in Caucasians? Provide a reference for your answers.

Answer: The CFTR gene is found in region 7q31.2 of the long arm of human chromosome 7 and the ΔF508 mutation is a three base pair deletion in the gene which results in a loss of the amino acid phenylalanine from position 508 of the protein. As a result of this deletion, the mutant CFTR protein, which is a chloride transporter, is degraded before it reaches the cell membrane. The frequency of this mutation in Caucasians is reported to be in the range of $0.025 - 0.03$.

Additional information can be found at the address http://ghr.nlm.nih.gov/condition/cystic-fibrosis.

5.) Suppose that a population contains 100 haploid individuals and that 90 of these carry the nucleotide A at a particular position in the genome, while the remaining 10 carry the nucleotide G at that position. Calculate the nucleotide diversity π at this position.

Answer: The answer depends on whether we sample with replacement or without replacement. In the former case, the probability of sampling two individuals with different nucleotides is

$$\pi = 2 \times 0.1 \times 0.9 = 0.18,$$

where the factor of two reflects the fact that two nucleotides can be sampled in either order: A and then G or G and then A. On the other hand, if we sample without replacement, then the probability of sampling two individuals with different nucleotides is equal to

$$\pi = \frac{90}{100} \times \frac{10}{99} + \frac{10}{100} \times \frac{90}{99} = \frac{2}{11} = 0.18\overline{18}.$$

Thus, in this case, the population is large enough that the two probabilities are nearly identical. In practice, nucleotide diversities are usually calculated under the assumption that we sample with replacement.

6.) Why are STRs generally preferred to SNPs in forensic genetic analyses?

Answer: STRs have several advantages over SNPs in forensic analyses, the most important being that they usually segregate many more than just two alleles (as is typical at most SNPs) which makes one STR marker more informative of identity than a single SNP marker.