Review of Central Dogma; Simple Mendelian Inheritance

Problem Set #1 (not for credit):

Problems 1-18 test your knowledge of the central dogma. The remaining problems are related to the topic of Mendelian inheritance patterns.

This table of the genetic code may be useful for some of the questions in this problem set.

First letter	Second letter: U	Second letter: C	Second letter: A	Second letter: G	Third letter
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	U
	UUC Phe	UCC Ser	UAC Tyr	UGC Cys	C
	UUA Leu	UCA Ser	UAA Stop	UGA Stop	A
	UUG Leu	UCG Ser	UAG Stop	UGG Trp	G
C	CUU Leu	CCU Pro	CAU His	CGU Arg	U
	CUC Leu	CCC Pro	CAC His	CGC Arg	C
	CUA Leu	CCA Pro	CAA Gln	CGA Arg	A
	CUG Leu	CCG Pro	CAG Gln	CGG Arg	G
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser	U
	AUC Ile	ACC Thr	AAC Asn	AGC Ser	C
	AUA Ile	ACA Thr	AAA Lys	AGA Arg	A
	AUG Met	ACG Thr	AAG Lys	AGG Arg	G
G	GUU Val	GCU Ala	GAU Asp	GGU Gly	U
	GUC Val	GCC Ala	GAC Asp	GGC Gly	C
	GUA Val	GCA Ala	GAA Glu	GGA Gly	A
	GUG Val	GCG Ala	GAG Glu	GGG Gly	G

1. If one strand of DNA in a double helix has a base composition of 5'-GTCATAACGGT-3', what is the base sequence composition of the complementary strand?

2. DNA was isolated from the organism *Staphylococcus afermentans* and analyzed for base composition. Results of this analysis indicate that 22% of the bases are cytosine. From this information can you predict what percentage of the bases are adenine? If so, what percentage? If not, why not?

3. RNA extracted from tobacco mosaic virus particles was analyzed for base composition and found to contain 30 percent cytosine. From this information can you predict what percentage of the bases are adenine? If so, what percentage? If not, why not?

4. The nucleic acids from various viruses were isolated and examined for base composition. Given the results shown below, what can you hypothesize about the physical nature of the nucleic acids from these viruses?

(a). 31% A, 31% T, 19% G, 19% C.
(b). 34% A, 34% U, 16% G, 16% C.
(c). 35% A, 30% U, 30% G, 5% C.

5. Indicate whether each of the following statements about the structure of DNA is true or false (each letter below refers to the concentration of that base in DNA).
(a). A+T = G+C
(b). A = G; C = T
(c). A/T = C/G
(d). A/C = T/G
(e). A+G = C+T
(f). G/C = 1
(g). A = T
(h). Hydrogen bonding between bases stabilizes the double helix
(i). Hydrophobic interactions provide stability to the double helix.
(j). The stability of the DNA double helix is unaffected by pH.
(k). When seperated, the two strands of the double helix are identical.
(l). If the sequence of one strand of a double helix is known, the complementary strand can be deduced.
(m). Each nucleotide pair contains two phosphate groups, two deoxyribose molecules, and two bases.

6. DNA double helices can be melted (denatured) into single-stranded molecules by heating. For denaturation to occur, molecular forces holding the single strands together must be overcome. Which of the following two DNA molecules will have the higher melting temperature (T_m)? Why? What are the relevant forces?

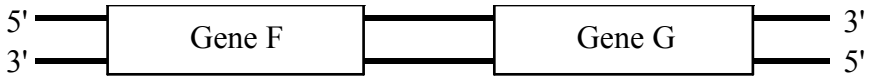
(I)	(II)
ATATAGACTATGCAT	CGACGTGAGGCATGG
TATATCTGATACGTA	GCTGCACTCCGTACC

7. For each of the terms in the left column, choose the best matching phrase in the right column.

(a). codon	1. removing base sequences corresponding to introns from the primary transcript.
(b). colinearity	2. a group of three mRNA bases signifying one amino acid.
(c). reading frame	3. the strand of DNA that has the same base sequence as the primary transcript.
(d). frameshift mutation	4. a transfer RNA molecule to which the appropriate amino acid has been attached.

(e). degeneracy of the genetic code	5. UAA, UGA, or UAG
(f). nonsense codon	6. most amino acids are not specified by a single codon.
(g). initiation codon	7. using the information in the nucleotide sequence of a strand of DNA to specify the nucleotide sequence in a strand of RNA.
(h). template strand	8. the grouping of mRNA bases in threes to be read as codons.
(i). RNA-like strand	9. AUG in a particular content.
(j). intron	10. the linear sequence of amino acids in the polypeptide corresponds to the linear sequence of nucleotide pairs in the gene.
(k). RNA splicing	11. produces different mature mRNAs from the same primary transcript.
(l). transcription	12. addition or deletion of a number of base pairs other than three into the coding sequence.
(m). translation	13. a sequence of base pairs within a gene that is not represented by any bases in the mature mRNA.
(n). alternative splicing	14. the strand of DNA having the same base sequence complementary to that of the primary transcript.
(o). charged tRNA	15. using the information encoded in the nucleotide sequence of an mRNA molecule to specify the amino-acid sequence of a polypeptide molecule.

8. The coding sequence for gene F is read from left to right. The coding sequence for gene G is read from right to left. Which strand of DNA serves as the template for transcription of each gene?



9. How many possible open reading frames (frames without stop codons) are there that extend through the sequence below?

```
5'CTTACAGTTTATTGATACGGAGAAGG3'
3'GAATGTCAAATAACTATGCCTCTTCC5'
```

10. Describe the steps in transcription and translation that require complementary base pairing.

11. The yeast gene encoding a protein found in the mitotic spindle was cloned by a laboratory studying mitosis. The gene encodes a protein of 477 amino acids.
(a). What is the minimum length in nucleotides of the protein-coding part of this yeast gene?

(b). A partial sequence of one DNA strand in an exon containing the middle of the coding region of the yeast gene is given below. What is the sequence of amino acids in that part of the yeast mitotic spindle protein?

(c). What is the sequence of nucleotides of the mRNA in this region of the gene? Show the 5′ and 3′ directionality of your strand.

5′GTAAGTTAACTTTCGACTAGTCCAGGGT3′

12. (a) In how many cases in the genetic code would you *fail* to know the amino acid specified by a codon if you know only the first two nucleotides of the codon?

(b). In how many cases would you *fail* to know the first two nucleotides of the codon if you know which amino acid is specified by it?

13. (a). Synthetic RNAs with random sequences can be made and were used to help crack the genetic code. The nucleotide composition of the RNA product is determined by the relative concentration of the nucleoside diphosphates (NDP) that are added to the reaction mixture. Consider reactions that contain only GDP and UDP. If the RNA is produced from a mixture that contains an excess of UDP, more valine than glycine is incorporated into protein, but when GDP is in excess, more glycine than valine is incorporated. WITHOUT looking at the Table of the genetic code, list the codons containing G *and* U that could possibly be coding for Val and Gly.

(b). Determine the relative frequencies of all possible triplet codons in the RNA produced by a synthesis reaction mixture consisting of 60% UDP and 40% CDP?

(c). Suppose the RNA made in question **(b)** is used as a messenger RNA in a protein synthesizing system. Using the genetic code in the Table, determine what percentages of the various amino acids are expected to be incorporated into polypeptides under the direction of this RNA.

14. The average mass of the 20 common amino acids is about 137 daltons. Estimate the approximate length of an mRNA molecule that encodes a polypeptide with mass of 65,760 daltons. Assume that the protein contains equal amounts of all 20 amino acids.

15. The thymine analog 5-bromouracil is a chemical mutagen that induces single base-pair substitutions in DNA. Specifically, 5-bromouracil causes transition mutations, which are substitutions of one purine for another purine, or one pyrimidine for another pyrimidine. Which of the following amino acid substitutions should you expect to be induced by 5-bromouracil with the highest frequency: (a) Met-->Val (b) Met-->Leu (c) Lys-->Thr (d) Lys-->Gln (e) Pro-->Arg (f) Pro-->Gln ?

16. Mutations resulting in an absence of enzymatic activity in a particular E. coli gene have been identified. The enzymes produced by these mutations have been purified and sequenced, and the amino acid sequences from these defective enzymes are shown below along with the corresponding sequence of the wild type enzyme:

wild type: - Ala - Pro - Trp - Ser - Glu - Lys - Cys - His -

Mutant 1: - Ala - Pro - Trp - Arg - Glu - Lys - Cys - His -

Mutant 2: - Ala - Pro-

Mutant 3: - Ala - Pro - Gly - Val - Lys - Asn - Cys - His -

(a). What is the most likely molecular basis of each mutation? (*Hint:* start by writing out the sequence of the RNAs expected to code for these proteins).

(b). What is the sequence of the DNA strand that is used as template for the mRNA that specifies this part of the protein?

17. A single nucleotide addition and a single nucleotide deletion approximately 15 sites apart in the DNA cause a change in protein sequence from:

- Lys - Ser - Pro - Ser - Leu - Asn - Ala - Ala - Lys -

to:

- Lys - Val - His - His - Leu - Met - Ala - Ala - Lys -

What are the old and the new mRNA nucleotide sequences? (Use the Table with the genetic code). Which nucleotide has been added and which has been deleted?

18. The Yeast Genome is approximately 12,000 kb in length and contains about 6000 open reading frames encoding an average of 500 amino acids.

(a). Using these numbers, calculate the proportion of the yeast genome that is protein coding DNA.

(b). The human genome is approximately 3,000,000 kb and contains about 35,000 open reading frames each encoding an average of 500 amino acids. What proportion of the human genome is devoted to protein coding?

19. You have two different derivatives of fibroblast cultures: one grows as completely flat cells, and the other grows as round, ragged cells. When you mix flat cells with extracts prepared from live round ragged cells, all the flat cells become round and ragged. These cells continue to grow as round ragged cells.

(a). What could explain the flat cells becoming round, ragged cells?

(b). You decide to heat up an extract of the round, ragged cells and mix some of this extract with a culture of live, flat fibroblasts. These cells fail to become round, ragged and remain flat. What could explain why the flat cells are not transformed into round, ragged cells following heat treatment of the extract?

(c). Since the heated extract did not transform the flat cells, you decide to blend the flat cells with a round, ragged cell extract that you have split into three fractions: one fraction is exposed to RNAse, one fraction is exposed to DNAse, and the last fraction is exposed to protease. The DNAse and RNAse treatments do not seem to affect the transformation, but protease treatment destroys the ability of the extract to transform flat cells into round, ragged cells. What is your conclustion of the nature of this genetic material? Is this unusual?

20. Mendel verified his principle of independent segregation by performing testcrosses. Describe how you would use testcrosses to determine the genotypes of F2 peas that result from a self-cross of a yellow, smooth dihybrid produced from a cross of pure-breeding green, wrinkled peas and yellow, smooth peas. Be explicit in stating the ratios that you would expect. This problem is easier if you draw the Punnett square first.

21. In flies, the *ebony* mutation (symbol *e*, produces dark bodies) is recessive to normal body color (use the symbol *E*); the *white*-eyed mutation (symbol *w*) is recessive to normal eye color (use *W*), and *scute* (some of the bristles are gone, use the symbol *sc*) is recessive to normal bristles (use *Sc*). You obtain a fly of unknown genotype that has normal body color, normal eye color, and normal bristle color.

(a). Describe how you would use a testcross to determine the genotype of the mystery fly. Be sure to state the genotype of the fly you will use for the cross.

(b). After performing your cross, you tally the following phenotypes:

430 normal body color, normal eye color, normal bristle pattern
416 ebony, normal eye color, normal bristle pattern
422 normal body color, white eyes, normal bristle pattern
433 ebony, white eyes, normal bristle pattern

What is the genotype of the mystery fly?

22. While at the Skagit Valley Tulip festival, you were chatting with a kindly farmer about some of his tulip varieties. He tells you of two true-breeding varieties he grows, one gives red flowers, and the other gives yellow flowers. He tells you that when he crosses a true-breeding red with a true breeding yellow, he only gets red flowers. However, if he then self-fertilizes these monohybrids, he observes 286 red flowers, and 108 yellow flowers.

(a). What do you think are the genotypes of the P, F1 and F2 plants? Define and use your own symbols.

(b). Does the observed phenotypic ratio fit with a 1:1 segregation, or with a 3:1 segregation? Test each possibility using a χ^2 test, and state the genetic hypothesis you will propose to the farmer.

(c). Impressed with your knowledge, the farmer then tells you of a different variety he is growing. The petals of these flowers have serrated, fringed edges. He tells you of another cross he performed, using a red, serrated parent and a yellow, normal parent. All of the offspring had red, normal-shaped petals. He then self-crossed these dihybrids, and observed the following progeny: 56 red, plain; 20 red, serrated; 24 yellow, plain; and 12 yellow, serrated. Propose a genetic hypothesis and test it with a χ^2 test, defining your genetic symbols.

23. Mendel studied several other traits in his peas. One was the position of flowers and seedpods (Axial, *A*, dominant to terminal, *a*), another was stem length (Long, *L*, dominant to dwarf, *l*) and another was flower color (Purple, *P*, dominant to white, *p*). Suppose you have a trihybrid that is heterozygous for all of these genes (*AaLlPp*) that you self fertilize.

(a). What proportion of the offspring do you expect to be axial, dwarf and purple?

(b). What proportion can you expect to have exactly one of the recessive traits?

(c). What proportion can you expect to have at least two recessive traits?

24. In Drosophila, the allele dp^+ determines long wings, and *dp* determines short (dumpy) wings, with dp^+ dominant to *dp*. A separate locus e^+ determines gray body color, *e* determines ebony body, and e^+ is dominant to *e*. The follwoing crosses were made, starting with pure-breeding parents:

P long, ebony X short, gray

F1 long, gray X short ebony (pure-breeding).

F2	long, ebony	54
	long, gray	47
	short, gray	52
	short, ebony	47

Propose a genetic hypothesis for these observation and test using χ^2 test.

25. A geneticist, in assessing data that fell into two phenotypic classes, observed values of 250:150. She decided to perform χ^2 analyses to test two different hypotheses: 1) the data fit into 3:1 ratio, and 2) the data fit into 1:1 ratio. Calculate the χ^2values and *p* to verify her results. What can be concluded about her hypotheses? How would your answer change if she only counted 40, and observed 25:15?

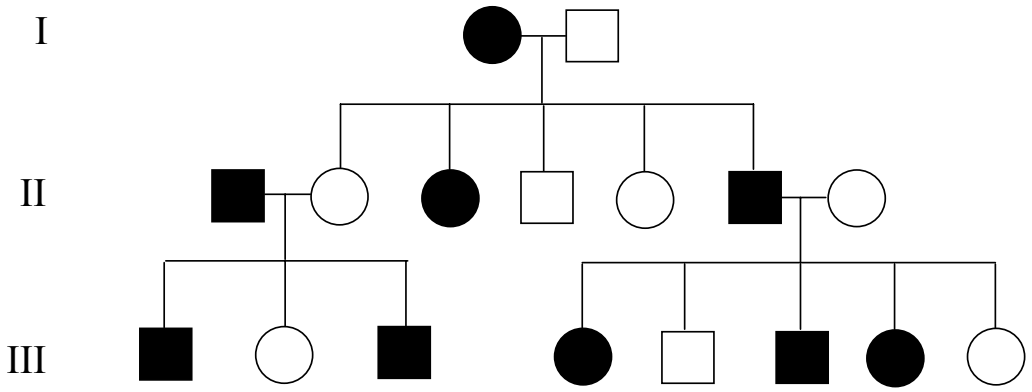
26. *A_B_* flies have black abdomens, *A_bb* flies have tan abdomens, *aaB_* flies have yellow abdomens and *aabb* flies have white abdomens.

(a). What proportion of progeny from the doubly heterozygote cross are yellow abdomens? Draw the Punnett square and give the phenotypic and genotypic ratios of all the progeny.

(b). If a black abdomen fly (*AaBb*) is crossed to a yellow abdomen fly (*aaBb*) then what is the probability of getting a black or a yellow fly in the F1?

(c). If a black abdomen fly mates with a tan abdomen fly and they have the following progeny, what are their genotypes?
376 black, 421 tan, 197 yellow, 98 white
Calculate the χ^2 value and determine whether the control of abdomen color can really be explained by two genes.

27. Consider the following pedigree for a hypothetical disease:



(a). Could the allele for the disease be dominant? If so, describe the genotypes of as many individuals as you can.

(b). Could the allele for the disease be recessive? If so, describe the genotypes of as many individuals as you can.

(c). If the disease were dominant, what proportion of offspring of a mating between III-3 and III-4 would you expect to be afflicted? What proportion would not be afflicted? How would your answer change if the disease was recessive?

28. Some disease causing mutations can be identified with a simple DNA test. For example, the autosomal dominant disorder Huntingtin's disease is caused by a mutational event in which a small number of CAG codons (typically ~10) expand, resulting in 40 or more CAG codons (see figure below). DNA from an individual at risk of developing Huntingtin's disease can be collected from a blood sample and the DNA can be digested with restriction enzymes, which cut at specific sites along the DNA (designated by the R in the figure below). The digested DNA fragments can then be subjected to agarose gel electrophoresis, which separates DNA by size (small DNA fragments migrate quickly on the gel and large fragments migrate more slowly). The gel can then be subjected to Southern blot analysis, to allow specific detection of the DNA fragments bearing the Huntingtin's disease gene. Results of such an analysis are shown for a family segregating Huntingtin's disease on the following page (the results obtained for each individual are shown directly below that individual). Is the child in the pedigree at risk of developing Huntingtin's disease?