Review of Central Dogma; Simple Mendelian Inheritance

Problem Set #1 Answers:

1. 5’-ACCGTTATGAC-3’

2. No. You would also need to know if this organism has a double stranded DNA genome. Assuming that it does have a double stranded genome, G+C = 44%, and A+T = 56%. Thus, 28% of the bases would be adenine.

3. No. tobacco mosaic virus has a single stranded RNA genome and thus, the base-pair stoichiometry of DNA does not apply.

4. (a). double-stranded DNA  
   (b). Double-stranded RNA  
   (c). Single-stranded RNA

5. (a). False  
    (b). False  
    (c). True  
    (d). True  
    (e). True  
    (f). True  
    (g). True  
    (h). True  
    (i). True  
    (j). False  
    (k). False  
    (l). True  
    (m). True

6. Molecule II will have the higher Tₘ. Thermal vibrations from heat can disrupt hydrogen bonds and base stacking interaction, the major forces that hold the two strands of a DNA duplex together. Molecule II has more GC base pairs than molecule I; 10 of 15 compared to 4 of 15. GC base pairs have three H-bonds compared to the two of AT base pairs. In addition, adjacent GC base pairs have greater stacking energy than adjacent AT base pairs. Therefore, more thermal energy will be required to separate the strands of a DNA molecule with a higher percent of GC.

7. (a). 2  
    (b). 10  
    (c). 8
(d). 12
(e). 6
(f). 5
(g). 9
(h). 14
(i). 3
(j). 13
(k). 1
(l). 7
(m). 15
(n). 11
(o). 4

8.

<table>
<thead>
<tr>
<th>5'</th>
<th>3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene F</td>
<td>Gene G</td>
</tr>
</tbody>
</table>


5' C) (TTA) (CAG) (TTT) (ATT) (GAT) (ACG) (GAG) (AAG) (G3' 5'CT) (TAC) (AGT) (TTA) (TTG) (ATA) (CGG) (AGA) (AGG) 3' 3' (GAA) (TGT) (CAA) (ATA) (ACT) (ATG) (CCT) (CTT) (CC5'

10. Transcription: Adding the appropriate ribonucleotide complementary to the template. Translation: codon in mRNA and anticodon in tRNA are complementary.
11. (a). Each amino acid is specified by a codon. A codon consists of three nucleotides. Therefore, there must be at least 477x3=1431 nucleotides in the coding part of this yeast gene.

(b). Since this sequence derives from the middle of the coding region it must contain an open reading frame that extends through the entire sequence. There are six possible reading frames (including the complementary strand, which we can easily derive) and only one contains an open reading frame (the complementary strand) that extends through the entire sequence:


The corresponding amino acid sequence would be: ?-Pro-Trp-Thr-Ser-Arg-Lys-Leu-Thr-Tyr.

(c). The sequence must correspond to that in which the complete open reading frame exists:

\[ 5'ACCCUGGACUAGUGCGAAGUUAACUUAC3' \]

12. (a). Using the genetic code in the Table, there are eight cases in which knowing the first two nucleotides does not tell you the specific amino acid.

(b). If you knew the amino acid, you would not know the first two nucleotides in the cases of Arg, Ser, and Leu.

13. (a). Val: UUU, UUG, UGU, GUU (however, among these, only GUU actually codes for Val)

Gly: GGG, GGU, GUG, UGG (however, among these, only GGG and GGU actually code for Gly)

(b). UUU = (0.6)^3 = 21.6%; UUC, UCU, and CUU = (0.6)^2(0.4) = 14.4% each; UCC, CUC, and CCU = (0.6)(0.4)^2 = 9.6% each; CCC = (0.4)^3 = 6.4%.

(c). The actual percentage of Phe would be 36% (UUU = 14.4%, UUC = 21.6%). The percentages of other amino acids would be:

Leu = 14.4% (CUU) + 9.6% (CUC) = 24%
Ser = 14.4% (UCU) + 9.6% (UCC) = 24%
Pro = 9.6% (CCU) + 6.4% (CCC) = 16%

14. The polypeptide would contain \(\frac{65,760}{137} = 480\) amino acids. Since each codon contains three nucleotides, the coding region of the RNA would have to be \(480 \times 3 = 1440\) nucleotides long.

15. Met-->Val is the only substitution that involves a transition. All others require a transversion; that is, a substitution of purine for pyrimidine or vice versa.

16. From the genetic code table, the wild type amino acid sequence must be coded for as follows:
- Ala - Pro - Trp - Ser - Glu - Lys - Cys - His - 
5' GCN CCN UGG $^{A/G,U/G,C}$ GA$^{A/G}$ AA$^{A/G}$ UG$^{U/C}$ CA$^{U/C}$

where: $N = A, G, C,$ or $U$ is present at the position. 
$A/G = \text{either A or G is present at the position.}$ 
$U/C = \text{either U or C is present at the position.}$

**Mutant 1:**
- Ala - Pro - Trp - Arg - Glu - Lys - Cys - His -
5'GCN CCN UGG AG$^{A/G}$ GA$^{A/G}$ AA$^{A/G}$ UG$^{U/C}$ CA$^{U/C}$

Ser has been changed to Arg. Assuming the mutation is the simplest possible change, a single bp change, it could have occurred in either of two Ser codons, AGU or AGC, converting the triplet to the Arg codon AGA or AGG.

**Mutant 2:**
Ala - Pro - STOP  
5' GCN CCN (UGA or UAG)

The protein is truncated compared to normal indicating that the codon following the one for Pro has been changed to a STOP codon. The Trp codon (UGG) can be changed by a single bp mutation at position 2 or position 3 to one of STOP codons (UAG or UGA).

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

Two frame-shift mutations occurred as shown below.

- Ala - Pro - Trp - Ser - Glu - Lys - Cys - His - 
5' GCN CCN UGG AG$^{U/C}$ GA$^{A/G}$ AA$^{A/G}$ UG$^{U/C}$ CA$^{U/C}$

\[ (-U) \quad \quad (+U \text{ or } C) \]

- Ala - Pro - Gly - Val - Lys - Asn - Cys - His -
5' GCN CCN GGA GUG AAA AA$^{U/C}$ UG$^{U/C}$ CA$^{U/C}$

Any Ser of the 6 codons (UCN and AG$^{U/C}$) could be present in the original sequence and allow the formation of a GCU or GCA Gly codon in place of Trp in the frameshift region. However, the next amino acid in the frameshift region is Val whose codons are GUN. To create a GUN codon at that position the original Ser codon must be an AGU codon, not AGC or one of the four UCN codons.

The Glu codon in the original protein must be GAA, not GAG, because only with the GAA codon will the frameshift create a codon in for Lys (which must be AAA, not AAG) in the frameshift region. The Lys in the original sequence must be coded for by AAA, not AAG, because only with the AAA codon will the frameshift create a codon for Asn (AA$^{U/C}$).
So, the original mRNA sequence is:

- Ala - Pro - Trp - Ser - Glu - Lys - Cys - His -
5'GCN CCN UGG AGU GAA AAA UG\textsubscript{U/C} CA\textsubscript{U/C} 3'

And, the original DNA sequence is (written in the unconventional 3'-5' direction to show it's complementarity to the mRNA sequence:

3' CGN GGN ACC TCA CTT TTT AC\textsubscript{A/G} GT\textsubscript{A/G} 5'

17. The old mRNA (top) converted to the new mRNA by the indicated nucleotide deletion and nucleotide addition:

- Lys - Ser - Pro - Ser - Leu - Asn - Ala - Ala - Lys -
  AA\textsubscript{A/G} AGU CCA UCA CUU AAU GCN GCN AA\textsubscript{A/G} \\
  \downarrow (- A) \downarrow (+ G)
  AAA\textsubscript{G} GUC CAU CAC UUA AUG GCN GCN AA\textsubscript{A/G}
- Lys - Val - His - His - Leu - Met - Ala - Ala - Lys -

N = A, G, C, or U

18. Each amino acid is specified by a codon. Each codon consists of three nucleotides. 
(a). Therefore 6000 open reading frames of 500 amino acids each would correspond to: 
3X500X6000 = 9,000,000 nucleotides, or 9,000 kb. Since the total size of the yeast genome is 
12,000 kb, the fraction of the yeast genome devoted to protein coding sequences is: 
9,000 kb/12,000 kb = 0.75, or 75% of the yeast genome. Thus, the vast majority of the yeast 
genome is devoted to protein coding sequences.

(b). 35,000 open reading frames of 500 amino acids each would correspond to: 
3X500X35,000 = 52,500,000 nucleotides, or 52,500 kb. Since the total size of the human 
genome is 3,000,000 kb, the fraction of the human genome devoted to protein coding 
sequences is:
52,500 kb/3,000,000 kb = 0.0175, or 1.75% of the human genome. Thus, the vast majority of 
human DNA does NOT code for protein. About half of the non-protein coding sequences are 
thought to represent introns and the other half is thought to represent sequences between 
genes.

19. (a). This experiment is similar to the Avery and McCleod experiment discussed in class. 
In this case, something in the round, ragged extract "transformed" the flat cells into round, 
ragged cells. Although it was not necessary for credit, you could think of the round, ragged 
trait as dominant to the flat trait, since the flat trait was masked. However, this was not set
up as a traditional cross - it was simply the addition of everything present when you break open round, ragged cells - including DNA, RNA and proteins - to intact flat cells.

(b). Heating destroyed the activity responsible for the "transformation". Since the ability of DNA and RNA to transform cells is mostly resistant to heat, and proteins are mostly destroyed by heat, this result suggests that the transforming activity could be protein.

(c). Since the DNAse and RNAse treatment had no effect, you can conclude that the transforming activity is neither DNA nor RNA. The fact that protease destroyed the transforming activity suggests that protein is responsible. This is very unusual, since we think of traits as being determined by DNA and sometimes RNA. However, there are rare exceptions to this rule. Certain diseases, for example sheep scrapie, Mad Cow disease, or Creutzfeldt-Jakob spongiform encephalitis, are spread by "infectious" proteins, known as prions that are completely free of a DNA or RNA component. The infectious protein is an extremely stable variant of a normal protein that acts like a crystal seed for the rest of the normal protein in the cell. When this variant is ingested by normal cells, it forces normal protein to misfold into the variant form, which is then stable and can go on to do the same to other cells. If it does not kill the cell, the protein crystal could cause the cell to adopt a different shape, which could then be transmitted to its daughter cells. It would therefore behave like an inherited trait. However, the prions that cause diseases wind up killing the cells, causing holes to form in brains of those infected. The prion diseases are spread by unknowingly feeding extracts of infected animals to normal animals.

20. A testcross is defined as a cross with an individual that is recessive for all traits, and therefore will not contribute any dominant traits. Therefore, you would cross each of the F2 individuals with a wrinkled, green pea plant (yyww), and analyze the offspring.

Of the yellow, smooth parents:
1/9 give only yellow, smooth peas. Since one of the parents could only contribute green or wrinkled alleles, these yellow smooth parents are YYWW.
2/9 give yellow, smooth and yellow, wrinkled offspring in a 1:1 ratio. These yellow, smooth parents are therefore YYWw.
2/9 give yellow, smooth and green, smooth offspring in a 1:1 ratio. These yellow, smooth parents are therefore YyWW.
4/9 give yellow, smooth and yellow, wrinkled, and green, smooth and green, wrinkled offspring in 1:1:1:1 ratios. These yellow, smooth parents are therefore YyWw.

Of the green, smooth parents:
1/3 give only green, smooth peas. These parents are therefore yyWW.
2/3 give green, smooth and green, wrinkled offspring in a 1:1 ratio. These green, smooth parents are therefore yyWw.

Of the yellow, wrinkled parents:
1/3 give only yellow, wrinkled peas. These parents are therefore YYww.
2/3 give yellow, wrinkled and green, wrinkled offspring in a 1:1 ratio. These yellow, wrinkled parents are therefore Yyww.
Of the green, wrinkled parents:
   All of the offspring give only green, wrinkled offspring. These parents are \( yyww \).

21. (a). You would cross the mystery fly with a fly that is homozygous recessive at all loci. In this case, it would be \( eewwscsc \).

(b). First, look at the distribution of phenotypes; there is a 1:1:1:1 segregation of body color and eye color traits, but all have normal bristles. This suggests independent assortment (at least for two of the traits). Second, examine each trait individually. Since the tester fly only has recessive alleles, the alleles present in the mystery fly are revealed. Half of the flies show the dominant body color trait, half show the recessive trait; the mystery fly is therefore \( Ee \). Similarly, the genotype at the eye color locus is \( Ww \). However, since all the bristles are normal, the mystery fly must be homozygous for the dominant bristle trait; \( ScSc \). Putting these all together, the genotype of the mystery fly was \( EeWwScSc \).

22. (a). For example, \( R \)=red, dominant, then \( r \)=yellow, recessive. The parents (P) are \( RR \) and \( rr \); the F1s are \( Rr \); the F2s are \( RR, Rr \) and \( rr \).

(b). For 3:1 ratio \( \chi^2=(286-295.5)^2/295.5 + (108-98.5)^2/98.5 = 1.22; p =0.26 \).

For 1:1 ratio \( \chi^2=(286-197)^2/197 + (108-197)^2/197 = 80.2; p \text{ is } <<0.01 \).

Therefore, 1:1 hypothesis not likely, 3:1 hypothesis is more consistent with the observations. Tell the farmer he has two alleles of one gene; Red is dominant to yellow, and his monohybrids were heterozygous.

(c). Hypothesis: Parents are homozygous for two different genes: \( R \) or \( r \) for petal color and \( N \) or \( n \) for normal or serrated petal shape (\( N \) dominant to \( n \)); they are \( RRnn \) and \( rrNN \). F1s should then be \( RrNn \), and selfing these dihybrids should produce progeny in the ratio of 9:3:3:1. If so, then expect 63:21:21:7 in F2s. Set up equation:
\[ \chi^2=(56-63)^2/63 + (20-21)^2/21 + (24-21)^2/21 + (7-12)^2/7 = 4.8. \]
Look up in table, with 3 DF, \( p =0.2 \). Do not discard the hypothesis of two independently assorting genes.

23. (a). 3/4 will be Axial; 1/4 will be dwarf; 3/4 will be Purple. Since each of these genes assort independent ly, use multiplication rule for independent events. \( P = (3/4)(1/4)(3/4) = 9/64 \)

(b). \( (3/4)(1/4)(3/4) = 9/64 \) will be Axial, dwarf and Purple.
\( (1/4)(3/4)(3/4) = 9/64 \) will be terminal, Long and Purple
\( (3/4)(3/4)(1/4) = 9/64 \) will be Axial, Long and white.

Since each of these events is mutually exclusive, add the individual probabilities together; \( P = 27/64. \)

(c). \( (1/4)(1/4)(3/4) = 3/64 \) will be terminal, dwarf and Purple
\( (1/4)(3/4)(1/4) = 3/64 \) will be terminal, Long and white
\( (3/4)(1/4)(1/4) = 3/64 \) will be Axial, dwarf and white
(1/4)(1/4)(1/4) = 1/64 will be terminal, dwarf and white.
Add individual probabilities together: \( P = 10/64 \).

24. Long winged, gray flies in F1 are heterozygous for two genes (genotype \( dp^+/dp \); \( e^+/e \)). F1 flies were testcrossed, so you should expect a 1:1:1:1 phenotypic ratio in offspring. Total number of flies is 200, each phenotypic class will be 1/4 of total, thus there should be 50 of each class. Set up \( \chi^2 \):
\[
\chi^2 = (54-50)^2/50 + (47-50)^2/50 + (52-50)^2/50 + (47-50)^2/50 = 0.76.
\]
With 3 degrees of freedom, \( p = 0.5-0.9 \); so deviation from perfect 1:1:1:1 ratio can be explained by chance.

25. Hypothesis 1) \( \chi^2 = (250-300)^2/300 + (150-100)^2/100 = 33.33 \)
Hypothesis 2) \( \chi^2 = (250-200)^2/200 + (150-200)^2/200 = 25 \)
In each case, \( p << 0.01 \), so she should reject both hypotheses and start over. If she had only analyzed 1/10 as many progeny from her crosses, she would have calculated \( \chi^2 = 3.33 \) for 1), and 2.5 for 2). Based on these numbers, both hypotheses appear to be valid. This shows how sampling a larger number of offspring from a cross allows for more rigorous hypothesis testing.

26. (a). The cross is \( AaBb \times AaBb \). 3/16 of the progeny of this cross are of genotype \( aaB_\) and will have yellow bellies. Genotypic ratios: 9/16 are \( A_B_ \), 3/16 are \( aaB_\), and 1/16 are \( aabb \). Phenotypic ratios: 9/16 are black abdomen, 3/16 are tan abdomen, 3/16 are yellow abdomen, and 1/16 are white abdomen.

(b). Black flies are \( A_B_ \), and must be the result of a union of the proper gametes. 1/4 of the gametes of the black fly will be \( AB \), the gametes of the yellow fly will be \( a_\), this combination equals 1/4 \( AaB_\). Another possible combination is \( Ab \) (1/4) from the first fly, with \( aB \) (1/2) from the second fly, \( P \) of this combination is 1/8. For yellow abdomens (\( aaB_\)), you can have \( Ab \) from fly1 (1/4) and \( a_ \) (1) from second; thus, the probability of yellow is 1/4. In addition, the combination of \( ab \) (1/4) with \( aB \) (1/2) will also give yellow abdomens. The probability of black or yellow is 1/4+1/8+1/4+1/4=6/8=3/4. Or draw the Punnett square and count.

(c). Since \( aabb \) flies are produced, the black fly has to be \( AaBb \), and the tan fly has to be \( Aabb \). The phenotypic ratios of the progeny are 3/8 black, 3/8 tan, 1/8 yellow, 1/8 white. Expect 409.5 black, 409.5 tan, 136.5 yellow, 136.5 white. Plug into equation: \( \chi^2 = (376-409.5)^2/409.5 + (421-409.5)^2/409.5 + (197-136.5)^2/136.5 + (98-136.5)^2/136.5 = 40.72 \). This is quite large, with 3 DOF \( p \) = very small, so the theory of two genes must be rejected.

27. (a). Yes, it could be dominant. If so, then
I-1 Aa, I-2 aa
II-1 Aa, II-2 aa, II-3 Aa, II-4 aa, II-5 aa, II-6 Aa, II-7 aa
III-1 Aa, III-2 aa, III-3 Aa, III-4 Aa, III-5 aa, III-6 Aa, III-7 Aa, III-8 aa
(b). Yes, it could also be recessive. However, this would require that the recessive allele is common, and the chances of finding homozygous recessive in the general population are therefore relatively high. If the recessive allele was common, the resulting genotypes would be:
I-1 aa, I-2 Aa
II-1 aa, II-2 Aa, II-3 aa, II-4 Aa, II-5 Aa, II-6 aa, II-7 Aa
III-1 aa, III-2 Aa, III-3 aa, III-4 aa, III-5 Aa, III-6 aa, III-7 aa, III-8 Aa

(c). If it were dominant, then 3/4 would be affected, and 1/4 would not be affected. If it were recessive, then all would be affected.

28. No. The larger DNA fragment with the CAG expansion mutation is not seen in the child.