Biology 321 Spring 2013 Assignment Set #6

NOTE new feature: if a problem number is underlined, a detailed answer will be available. Otherwise the answer will (or may) be brief.

Required Reading: Genomes Unzipped Blog Post from 4/25/2013

The complicated relationship between genotype and phenotype: why predicting the phenotypic effect of mutations is hard

http://www.genomesunzipped.org/2013/04/why-predicting-the-phenotypic-effect-of-mutations-is-hard.php

Required Reading: Genomes Unzipped Blog Post from 2/16/2012

All genomes are dysfunctional: broken genes in healthy individuals

http://www.genomesunzipped.org/2012/02/all-genomes-are-dysfunctional-broken-genes-in-healthy-individuals.php

---

➔➔ ***Have a look at these Web Sites***

DNA *from the Beginning*:
http://vector.cshl.org/dnaftb/
Review DNA structure here:
http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/D/DoubleHelix.html

---

➔➔ **Stuff to Memorize about nucleic acid structure**

- Terminology: purine pyrimidine adenine, guanine, thymine, cytosine
- Size of human genome and the E.coli genome
- dNTP – ribose structure; carbon numbering and where purine or pyrimidine base is attached

➔➔ **Stuff to Review/Learn on your own:**

- features of the genetic code: tracking the conversion of a sequence of three bases in DNA to the presence of a specific amino acid in protein
- definitions (and implications) of nonsense, frameshift, silent, synonymous & missense mutations

---

➔➔ *Required Reading and Problem Assignments Text 10th edition*

Warm up with Chapter 1*: Sections 1.1 (The nature of biological information) and 1.2 (How information becomes biological form) and pg 20 (paragraph on PCR). Look at figures 1.7 (Representative Chromosomal Landscapes) & 1.8 (A specific human chromosomal landscape) carefully.*

Chapter 7*: DNA structure and replication: Browse through pgs. 251-268 and carefully read sections 7.2 (The DNA structure) & 7.3 (Semiconservative replication) & 7.4 (Overview of DNA replication).* ***Read through solved problems 1 & 2. Work problems 6, 8, 21, 23, & 32***

Chapter 8: *Review RNA structure and Transcription: read pgs 283-290 (Sections 8.1 RNA and 8.2 Transcription) including figure 8.7 (promoter sequences in E. coli) & 296-297 (about splicing). Look carefully at figures 8-2 through 8-7 & 8-11.* ***Work problems 19, 20 & 21***

Chapter 9: *Review nuts and bolts about proteins and the genetic code: read pgs. 309-319.* ***Read through Solved problems 1& 2. Work problems 12, 13, 14, 18, 33, 38***

Chapter 16: Mutation, Repair and Recombination *Read pages 553-558 (Section 16.1 The phenotypic consequences of DNA mutations); 560-567 (Mechanisms of spontaneous mutations &*

Section 16.3 The molecular basis of induced mutations). Browse through ***Section 16.4 on DNA repair mechanisms.*** *You should be aware that DNA repair systems exist and that normal function is critical to limiting the rate of spontaneous and induced mutation but we will not cover this material in lecture.]* ***Work problems 8, 9, 10, 11, 13, 21, 22, 25, 31***

Cancer

- ***Chapter 16*** Section 16.5 *Cancer: an important phenotypic consequence of somatic mutation*
- ***Chapter 17*** Figure 17-2 *(Types of chromosome mutations)* & pgs. 589-590 *(Intro to chapter)*
- ***Chapter 7*** *Browse through section 7.7 on telomeres*

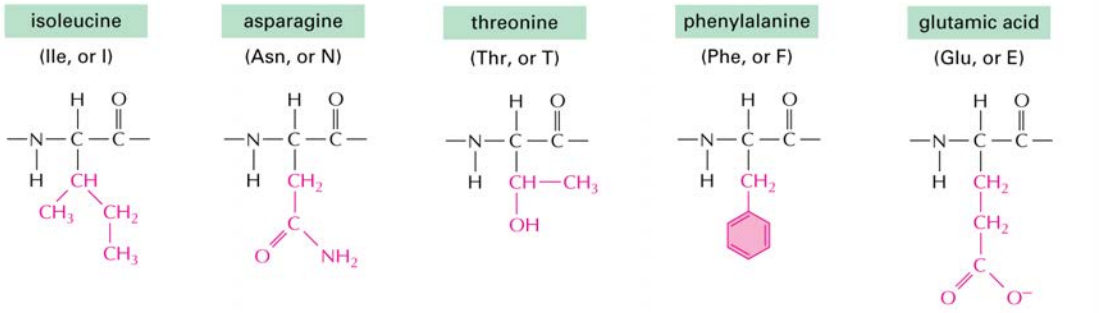
Required Reading and Problem Assignments in 9th edition of text
http://fire.biol.wwu.edu/trent/trent/assignmentset6.9.pdf

The next 8 pages of questions reviews material from Biology 205.

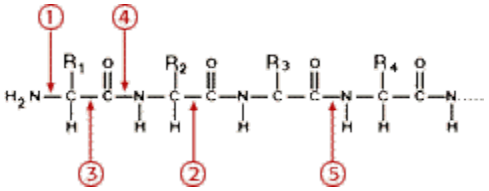
1. The bonds that link amino acids at the primary level of protein structure are called:
a). noncovalent bonds b). ionic bonds c). peptide bonds d). disulfide bonds

2. Which of the levels of protein folding depend on amino acid side chain interactions?
a). primary b). secondary c). tertiary d). all of the above

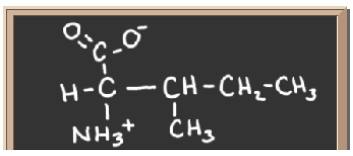
3A. Which amino acid(s) has/have hydrophobic side chain(s)? Circle the relevant amino acids

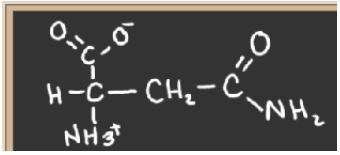


3B. Which arrow(s) point(s) to a peptide bond?
a). 1
b). 2
c). 2 and 3
d). 3
e). 4 and 5

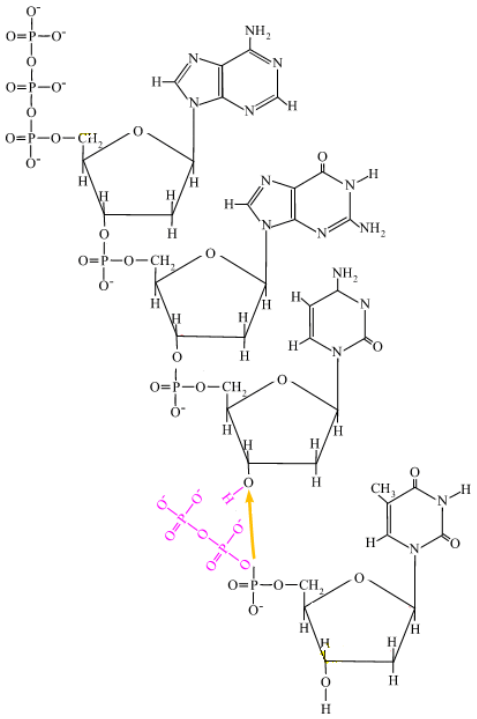


4. A mutation in the DNA sequence of a gene causes an isoleucine residue to be replaced by an asparagine residue in the protein coded for by the gene. With this single amino acid substitution, this globular protein is no longer functional. ***Briefly propose an explanation for the inactivation of the protein. A couple of sentences and/or drawing will suffice. Use proper terminology.*** **Isoleucine** **asparagine**





5. Examine the structure below. What is the name of this molecule?



a). single stranded RNA
b). single stranded DNA
c). Neither
d). Both

6. Examine the structure in the previous question. Indicate the 5' and the 3' carbons of this molecule. Write them in the figure.

7. From 5' to 3', what is the sequence represented above (assume the reaction depicted has been completed; i.e., 4 units joined together):
a). 5'-AGCT-3'
b). 5'-TCGA-3'
c). 5'-UCGA-3'
d). 5'-AGCU-3'

8. Examine the following DNA sequence. Which of the strands below represents the complementary RNA strand that would be transcribed from this DNA strand? 5' AATTGCGCTTAG 3'

a). 3' UUAACGCGAAUC 5'
b). 5' UUAACGCGAAUC 3'
c). 5' AAUUGCGCUUAG 3'
d). 3' TTAACGCGAATC 5'

9. A codon consists of:
a). three consecutive amino acids in a polypeptide
b). three consecutive deoxyribonucleotides in a DNA molecular
c). three consecutive ribonucleotides in an mRNA
d). three consecutive RNA molecules

10. What is meant by a reading frame?
a). A consecutive sequence of non-overlapping anti-codons
b). A consecutive sequence of overlapping anticodons
c). A consecutive sequence of non-overlapping codons
d). A consecutive sequence of overlapping codons

There may be more than one correct answer per question. <u>Circle all correct answers.</u>

11. The site where RNA polymerase associates with the DNA template during transcription initiation is called…

a. primer **b.** promoter **c.** cis-element **d.** AAUAAA site **e.** trans-element

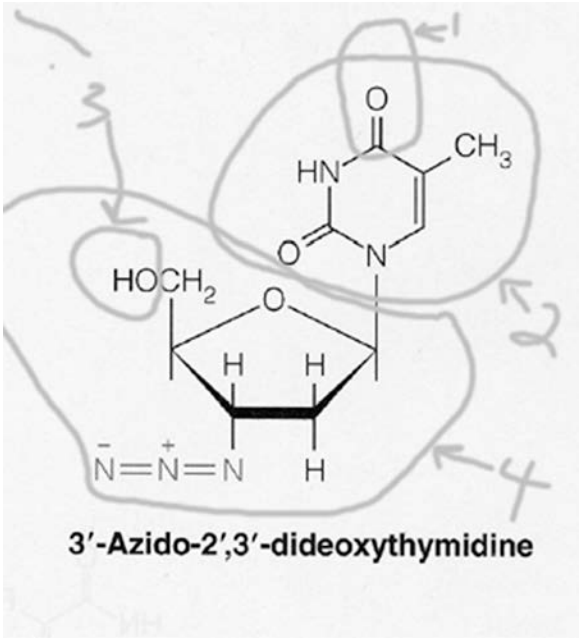
12. Primary transcripts in eukaryotes may be alternatively …

a. guanine-capped **b.** poly-adenylated **c.** spliced **d.** replicated

…to yield different mRNA sequences.

14 ***Which of the following is not true about hydrogen bonds?***
a. They contribute to the secondary and tertiary structure of tRNA molecules
b. They contribute to the secondary and tertiary structure of polypeptides
c. they form the basis for storage, replication and chemical conversion of genetic information
d. they represent the sharing of a hydrogen atom between two electronegative atoms
e. they connect a given amino acid to its tRNA

15. Structure Recognition: The compound shown below is known by the acronym **AZT.** It is used to treat HIV (human immunodeficiency virus) infections.
a. This compound is remarkably similar (but not identical) to what structure? **(a couple of words -- no explanation necessary)**
b. Examine the chemical structure of this compound. Name the functional groups/structures that are circled. Choose from the list below the figure.



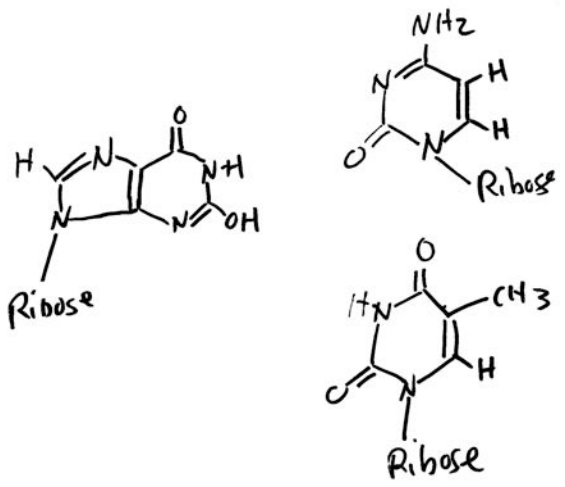
3′-Azido-2′,3′-dideoxythymidine

1. _____
2. _____
3. _____
4. _____

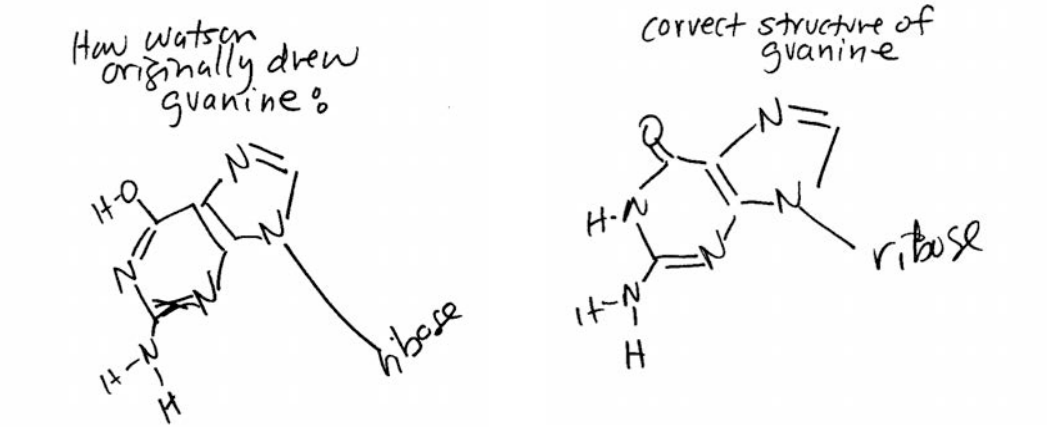
purine pyrimidine methyl hydroxyl carbonyl carboxyl ribose deoxyribose phosphate modified-deoxyribose modified ribose

16. Before James Watson and Francis Crick proposed the correct structure of DNA in 1952, there was much incorrect speculation about the structural nature of the DNA polymer. For example, one prominent scientist (Linus Pauling) proposed that DNA was a ***triple helix***: three polymers wound around each other with ***the bases on the outside*** and not interacting with each other. (Yikes!) One particular observation that was critical to Watson and Crick was that in DNA the molar amount of A always equalled the molar amount of T and likewise for G and C. They used this information to propose two specific aspects of DNA structure. **What aspects of the structure of DNA explain these data? List 2.**

17 Examine the structure on the left. Which base on the right would it form the ***most*** hydrogen bonds with? Indicate the hydrogen bonds with a dotted line. *No explanation necessary*.



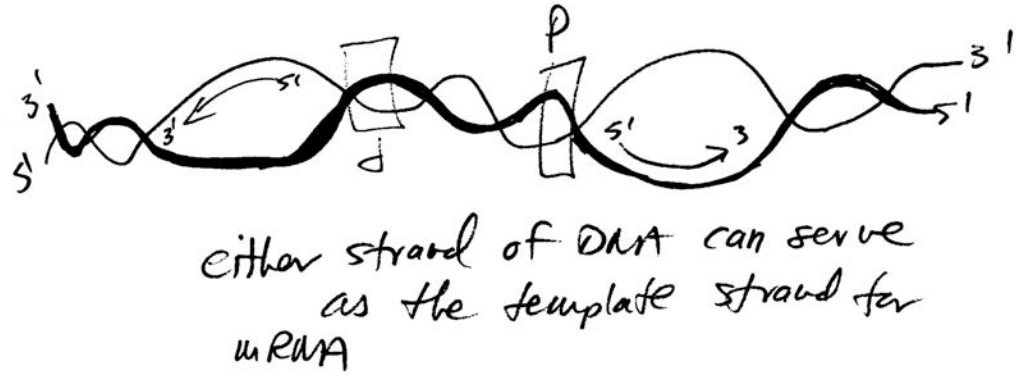
18. As James Watson admits in his book *The Double Helix*, when he and Crick first started to explore various models for the structure of DNA, his knowledge of chemistry was a bit "weaker" than it should have been. By mistake, he originally drew the structure of guanine as shown below. More importantly, though, he could not visualize the correct structure of DNA until the mistake was corrected by an organic chemist (who no doubt gave him some grief about it). ***Examine the structures shown below, and then explain why this was a critical mistake. Be sure to use proper terminology in your answer. One or two sentences here.***



19. A tRNA anticodon reads 5' GUU 3'. Consult the code table in your text. What amino acid should be covalently linked to the tRNA? ***Show your work***.

20. The DNA sequence adjacent to a gene that binds RNA polymerase and directs transcription to a specific strand of DNA is called a _____.

21. Transcription issues:
a. What does the P stand for? b. What is the function of P? c. What defines the P site?



22. The following DNA sequence shows a "gene" encoding a small peptide. The three "stop" codons are UAA, UAG, and UGA. AUG is the start codon. The bottom strand is the template strand.

promoter
5'(tgacgtatatta)TGACCGTACATGAGTAATACAAGATAAA 3'
3'(actgcatataat)ACTGGCATGTACTCATTATGTTCTATTT 5'

How many animo acids long will the small protein encoded by this "gene" be?
a. 3 b. 4 c. 5 d. 6 e.7

24 Consult the genetic code table in your textbook. If the sequence on the template strand of DNA reads 3' ATG 5', which amino acid will the resulting mRNA codon specify?
a. Tyr b. Lys c. His d. Val e. Met
When you work this problem, don't forget that a codon sequence is always written 5' to 3' and that the mRNA polymer is synthesized anti-parallel to the DNA template.

25. List the types of RNA involved in translation and the specific role of each of the RNAs in this process.

26. Which of the following is **not** a feature of the triplet genetic code used for translation? Note, there *may* be more than one correct answer to this question.
a. it is very ancient
b. it differs in eukaryotes and prokaryotes, reflecting basic differences in cell structure and function
c. it refers only to sequence information involved in specifying protein and is not relevant to sequence information involved in the specification of promoters or tRNA.
d. a few of the triplets specify stop codons, which act to terminate translation

27. Using the genetic code, identify a possible sequence of nucleotides in the ***DNA template strand*** for an mRNA that codes for the polypeptide sequence:
met-trp-phe. All sequences are written 5' to 3'
a. AUG-UGG-UUU b. AAA-CCA-CAU c. UUU-GGU-GUA
d. AAA-CCA-CAT e. ATG-TGG-TTT f. TAC-ACC-AAA

29. *With respect to a codon, indicate if the statement is true or false:*
___ It may code for the same amino acid as another codon does.
___ It never codes for more than one amino acid.
___ The genetic code used for translating the monomer sequence of mRNA into protein is different in bacterial and human cells, reflecting our vast differences in complexity.

30. HIV = human immunodeficiency virus = causative agent of AIDS
The genome of this *retro*virus is single-stranded RNA. After infecting the cell a virus enzyme called *reverse transcriptase* uses the RNA genome as a template to make a DNA copy of the genome.
Take the name of this protein apart and explain what it means:
a. ase:
(one word)
b. transcript (or to transcribe):
(a few words or one sentence)
c. why is it reverse (or retro)? relative to what?
(one sentence)

31. A short mRNA is synthesized by RNA polymerase using the sequence shown below as a template. The mRNA will code for which of the polypeptides shown to below? **Be sure to show your work.**
AUG = start codon
template strand :
5' C T A G G T T T T A C G C A T G A T T A C G 3'

a. met arg lys thr	b. met arg thr asn	c. met ile thr
d. leu gly phe thr his asp tyr		e. arg asp his ala

END of REVIEW QUESTIONS

Additional Study Problems

Problem -2 A gene may be viewed as a series of binding sites for RNA and protein. What binds to each of these sites? Check your answers in the text: pg 498.

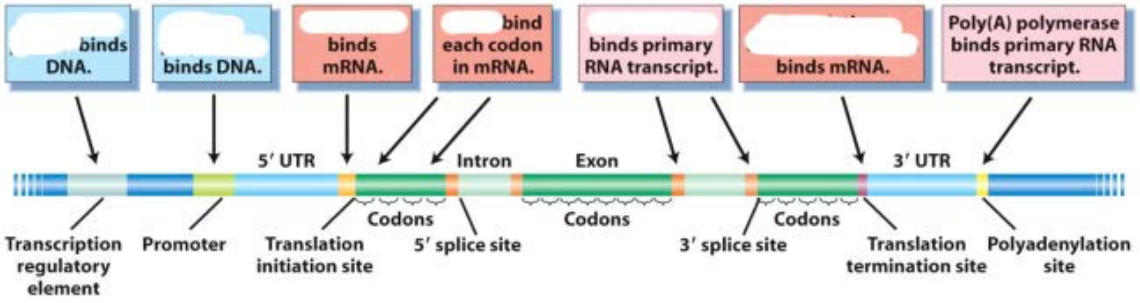


Figure 14-9
Introduction to Genetic Analysis, Tenth Edition

Problem -1

1. Here is a cartoon of a generic eukaryotic gene.

a. In the blanks by the arrows label the gene parts. Use proper terminology

b. Label the 5' and 3' ends of the gene

c. Draw **an arrow to** and, using proper terminology, **label the site** to which RNA polymerase would bind to initiate mRNA synthesis.

d. What is the name of the process described in c? _____

e. Based on your understanding of scale in higher euk genes, what do the // marks signify? *One sentence*

f. Why are some boxes black and some white and what determines the boundary? 1-2 sentences

Problem 0 Here is a quote from a Science paper published a couple of years ago. *"We analyzed the whole genome sequences of a family of four, consisting of two siblings and their parents. Family-based sequencing allowed us to ……directly estimate a human intergeneration mutation rate of ~1 x 10^{-8} per position per haploid genome copy."*
How many total base pair differences were detected between a particular child and its parents? *Show your work and circle your answer. You must track units to get full credit for this calculation.*

Problem 1 Spontaneous mutation rates have been determined for a variety of model organisms. In *E. coli* the forward mutation rate for a typical gene is 2 X 10^{-6} mutant copies per cell division (1 mutant copy per 500,000). In contrast, the forward mutation rate for a typical Drosophila gene is ~4 X 10^{-5} per gamete (1 mutant copy per 25,000 gametes).

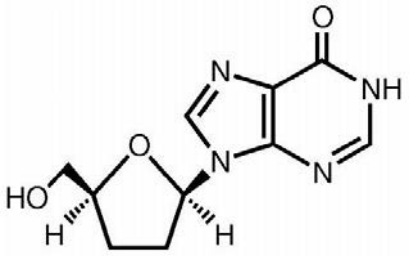
Assuming the the accuracy of DNA replication is the same in both species, briefly explain the apparent difference in mutation rate. (Hint: think about the factors that are important in mutagenesis).

Problem 2
For both DNA and RNA polymerases the error rate (due to tautomeric shifts) for initial nucleotide selection is estimated to be one mistake in every 10^4-10^6 bases added. Accordingly, the overall mistake rate for transcription ranges between 1 X 10^{-6} to 1 X 10^{-4} mistakes per base copied, but the overall mistake rate (errors per base per round of copying) for DNA replication is estimated to be ~1 X 10^{-10}. Explain these observations.

Problem 3
Which statement is false?
a. GC base pairs are harder to denature than AT base pairs
b. The phosphodiester bonds in DNA connect the 3' and 5' carbons of adjacent nucleotide residues.
c. For double-stranded DNA, the ratio of A+T/G + C is always 1.0.
d. The deoxy in DNA refers to the 2' carbon.
e. The chemical difference in the ribose of RNA and DNA makes RNA biochemically more unstable than DNA.

Problem 4 Shown below is the structure of *didanosine* (trade name Videx) which is used as part of the combination anti-HIV treatment.

a. Label the carbons on the ribose.



b. (This compound is also known as *dideoxy*I. Briefly explain the underlined portion. Be sure to refer to a normal DNA nucleoside in your answer. **Two sentences.**

Problem 4 continued

c. After human enzymes convert this molecule to its "TP form", the HIV reverse transcriptase will accept is as a substrate.
What does TP stand for? **(two words; no explanation)**

d. Addition of this molecule to the 3'end of a DNA polymer results in termination of DNA synthesis. Briefly explain why. **One sentence.**

Problem 5A (See also next problem)

a. One of the most common spontaneous lesions that occurs in DNA under physiological conditions is the hydrolysis of the amino group of cytosine, converting it to uracil. If this lesion is left unrepaired, what type of mutation would eventually result? Be sure to use proper terminology.

b. Cells have evolved a DNA repair surveillance system that (1) recognizes uracil bases as abnormal (not belonging in DNA) and (2) removes them. Use this information to provide an "rationale" for the fact that thymine is "used" instead of uracil in DNA.

Problem 5B (See also previous problem)

a. In some organisms, cytosine is methylated at carbon 5 of the pyrimidine ring after it is incorporated into DNA. If a 5-methyl cytosine is hydrolyzed as described above for cytosine, what base will be generated?

b. Why is this spontaneous lesion mutagenic? If left unrepaired, what type of mutation will eventually result from hydrolysis of the amino group of 5-methyl cytosine?

Problem 6

Aflatoxin B (AFB) is a powerful carcinogen that forms a covalent bulky addition product with guanine. This results in the loss of the modified guanine from the nucleotide and the generation of an apurinic site. Studies have shown that adenine is inserted preferentially **across** from an apurinic site during DNA replication. Aflatoxin must induce:

a. GC ---> AT transition
b. GC ---> CG transversion
c. GC ---> TA transversion
d. single base frameshift mutation
e. none of the above

Problem 7 Achondroplasia is a *completely penetrant*, autosomal dominant disorder characterized by disproportionate short stature -- the arms and legs are short compared with the head and trunk. More than 80% of the people who have achondroplasia are born to parents of *normal* stature. Findings from molecular studies have shown that these sporadic cases of acondroplasia are almost always
caused by mutations inherited from the father. The occurrence of achondroplasia is higher for older fathers: 50% of children with achondroplasia are born to father older than 35 years old. There is no association with maternal age.

a. Based on your understanding of the molecular basis of mutation, ***briefly*** speculate as to why

- most new mutations in this gene are paternal rather than maternal
- and why the occurrence of this disease is higher among older fathers

b. Draw a diagram to show how a GGU→ AGU codon change could result from a *spontanous* mutation that occurs during DNA replication and *involves a C in its rare tautomeric form*.

- *Indidate the rare form of C as C**
- *label 5' and 3' ends of all DNA strands*
- *indicate the strand that serves as the template for RNA polymerase during transcription*
- *be sure to establish the mutation in both strands of DNA*

Problem 10 The following amino acid sequence is a portion of an enzyme:

met trp tyr arg gly ser pro thr

Various mutant forms of this protein have been characterized. Three are shown below.

Mutant #1 **met trp**

Mutant #2 **met cys ile val val leu gln**

Mutant #3 **met trp his arg gly ser pro thr**

For each mutant:
a. State the general class of mutation
b. Consult a table showing the genetic code. Indicate the specific base pair change that occurred for each mutation. Assume these mutations resulted from single base pair changes. One correct answer is sufficient (even if there is more than one possibility).
c. State which of the following classes of mutagens could have caused the mutation: base analog, base modifying agent, intercalating agent
d. One of these mutant enzymes has 50% of the normal enzymatic activity. Which one is it?

Problem 11 (For additional practice: this question is like #10. See also question 15.5 in text)

Three mutations have been identified in the gene coding for the following short polypeptide. Each mutant polypeptide results from a **single** base pair addition, deletion or substitution. Note: There may be more than one possible mutagenic event.

Wild-type: met lys asn tyr met lys met trp

Mutant 1: met lys asn

Mutant 2: met lys asn tyr ile lys met trp

Mutant 3: met lys asn tyr met lys ile val val thr arg gly glu

(i) Examine mutant #1 carefully. Which of the following mutagens could have induced this mutation?
a. base analog
b. mutagen that induces transversions
c. intercalating agent.
d. all three mutagens (a, b and c)
e. two of these mutagens

(ii) Examine mutant #2 carefully. Which of the following mutagens could have induced this mutation?
a. base analog b. base modifying agent c. intercalating agent.
d. all three mutagens (a, b and c)
e. two of these mutagens

(iii) Mutant #3 must result from: a. insertion of a base pair b. deletion of a base pair c. a single base-pair substitution d. none of the above

Problem 12

Familial hypercholesterolemia (FH) is an autosomal disease state resulting in high levels of cholesterol. fh^+ = wild-type fh = mutant

	cholesterol level
fh^+ fh^+	150-250 mg/dl
fh^+ fh	200-400
fh fh	> 500

The protein specified by the FH gene is an LDL receptor. The diagram below indicates three mutant LDL receptor proteins. The polypeptide on the left shows the wild-type amino acid sequence and the change that occurred in individual FH380. Note: the + and - signs on the figure refer to the charge on the amino acid and are not relevant to this question.

a. For each mutant protein indicate the type of mutation (at the **protein** level).

b. For FH380, in mRNA language indicate the single base pair change that occurred to generate this mutation:

c. Draw a diagram of the wild-type DNA sequence of the val-tyr-gln and show how the FH380 mutation could have occurred (diagram the specific steps):

- Assign GUG to val, UAC to tyr and CAA to gln
- Assume that the mutation resulted from a tautomeric shift in a thymine base during DNA replication.
- Show both strands of DNA and label the 5' and 3' ends.
- Indicate which strand is the template strand for RNA polymerase.

d. Individual FH380 is heterozygous for the FH mutation. He is 25 years old and has already suffered one heart attack. As an industrial worker, his father was exposed to high levels of a compound called hydroxylamine. This compound acts as a mutagen by adding a hydroxyl group to cytosine. This modified form of cytosine always base-pairs with adenine.

The father sues the company that he works for claiming his son's disease was directly the result of his exposure to hydroxylamine. You are called as an expert witness.

- Are you going to testify for the defense or the prosecution?
- Is there additional information that needs to be collected? If so, be very explicit about what you need to know before you appear in court. The more airtight your testimony, the more you will get paid.

e. The parents of FH683 are 80 years old and in good health. They have a similar complaint (as the family above) and they want to get in on the legal action. The mother worked in a peanut processing plant for many years and was exposed to highly levels of aflatoxin and proflavin. Aflatoxin causes AT → CG mutations and proflavin is an intercalating agent. The stop codon in FH683 is UAG. Do they have a stronger or a weaker case than FH380's family? **Be sure to examine the codon table carefully.**

DiAGRAM on next page

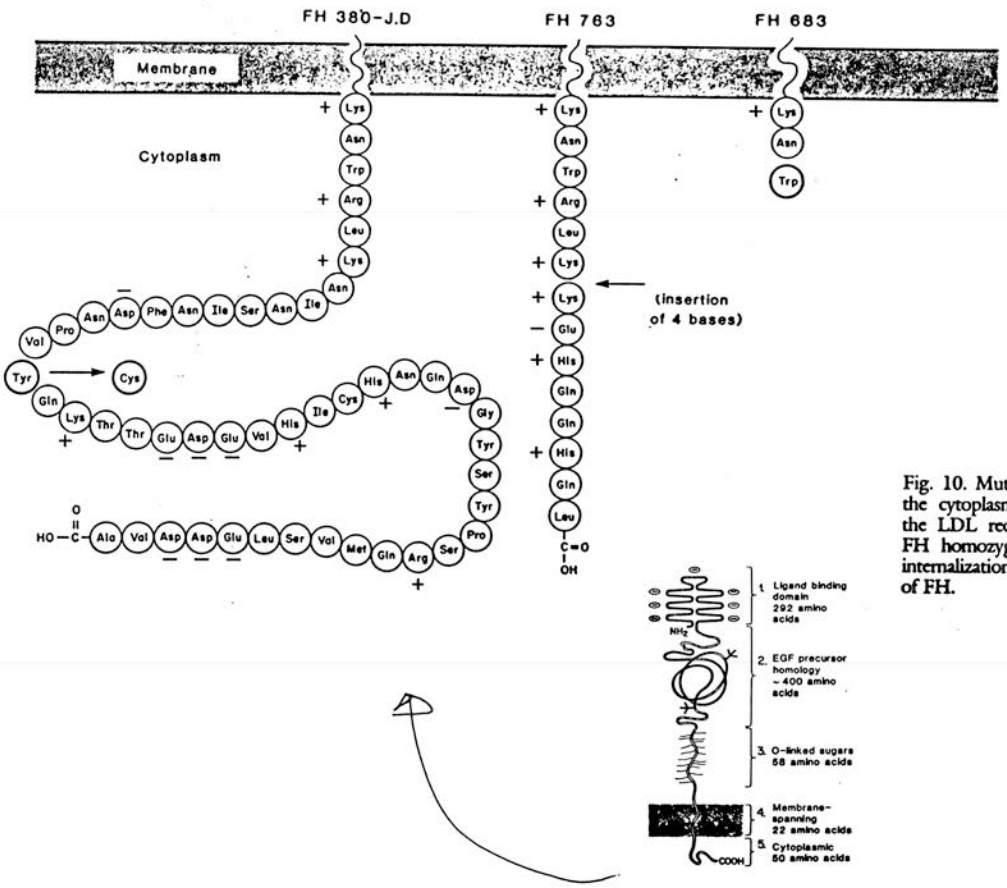


Fig. 10. Mutations affecting the cytoplasmic domain of the LDL receptor in three FH homozygotes with the internalization-defective form of FH.

Problem 18 About 5% of breast cancer cases are caused by an inherited susceptibility allele. A mutation in a gene called BRCA1 is thought to account for approximately 80% of families with a high incidence of both early-onset breast and ovarian cancer.

a. Examine the table 1 below. What is meant by a neutral sequence variation? One sentence maximum.

b. Examine the data on PM7: Is this sequence variation a *polymorphism*?
Circle: YES or NO. One sentence explanation.

c. You want to set up a genetic screen to identify individuals at high risk for early onset breast cancer. What is the significance of the data in this table in the context of this goal? (Two sentences maximum.)

d. Is the sequence variation PM2 likely to result in a neutral missense mutation or silent (same sense) mutation? Briefly explain (one sentence).

e. Is sequence variation PM3 likely to result in a neutral missense mutation or a silent (same sense) mutation? Briefly explain (one sentence).

Table 1. Neutral sequence variation in BRCA1. For the frequency in control chromosomes, the number of chromosomes with a particular base at the indicated site is shown (A,C,G, or T).

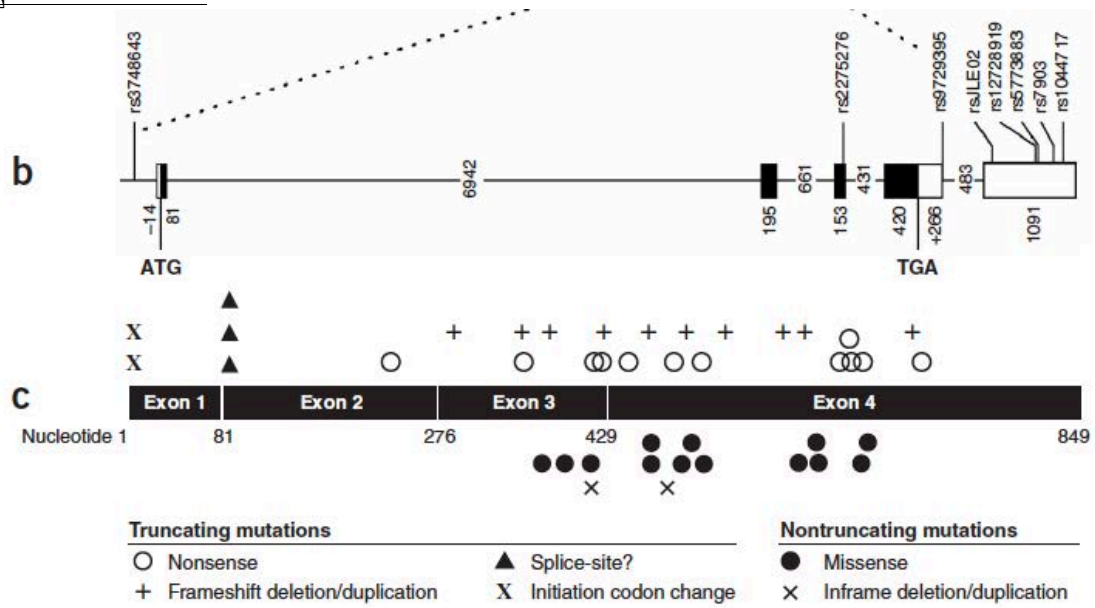
			Frequency in control chromosomes			
Name	Codon location	Base in codon*	A	C	G	T
PM1	317	2	152	0	10	0
PM6	878	2	0	55	0	100
PM7	1190	2	109	0	53	0
PM2	1443	3	0	115	0	58
PM3	1619	1	116	0	52	0

*That is, position 1,2 or 3 of the codon.

Table 2. Predisposing mutations in BRCA1. Science 266: 66 1994
NA indicates not applicable. ND = not determined

		Mutation		Frequency
Kindred	Codon #	Nucleotide change	Coding effect	in control chromosomes
1901	24	-11 bp	frameshift	0/180
2082	1313		Gln --> Stop	0/170
1910	1756	Extra C	frameshift	0/162
2099	1775	T-->G	Met --> Arg	0/120
2035	NA	ND	Loss of transcript	ND

Problem 19



NATURE GENETICS VOLUME 38 [NUMBER 1 [JANUARY 2006

1. Inspect panel B and identify exons, introns, translational start and stop sites.
2. Inspect panel C. For each type of mutation indicate why is it/could be truncating or not
3. Why would changing the initiation codon lead to a truncation?

Problem 20 Nature 384: 457 12/5/96 Non-insulin-dependent diabetes mellitus (NIDDM) affects about 2% of the world's population. Genetic factors are important in the etiology of this disease, and specific genes that influence the development of NIDDM (such as the HNF-1 gene) have been identified. Carefully examine the table shown below which was included in a paper describing genetic variaiton in the HNF-1 gene.

a. Briefly explain why the changes listed in this paper are called polymorphisms. **One sentence.**

b. Refer to Table 2 codon 17: ***At the protein level*, what term(s) (from the mutation jargon) best describe(s) this polymorphism?** No explanation necessary

c. Refer to Table 2: codon 27: ***At the protein level, what term(s) (from the mutation jargon) best describe(s) this polymorphism?*** Briefly explain in one sentence.

d. Briefly explain the importance/significance of the data in Table 2. Why was it included in this study of mutations in the HNF-1 gene? **Two sentences.**

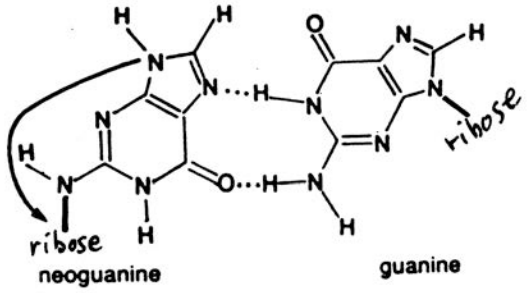
e. Why are there no codons/amino acids indicated for the intron polymorphisms? One sentence.

TABLE 2 DNA polymorphisms in human HNF-1α gene

Location			
Exon	Codon	Nucleotide change	Frequency
1	17	CTC(Leu) → CTG(Leu)	C 0.57, G 0.43
1	27	ATC(Ile) → CTC(Leu)	A 0.63, C 0.37
1	98	GCC(Ala) → GTC(Val)	C 0.98, T 0.02
4	288	GGG(Gly) → GGC(Gly)	G 0.67, C 0.33
7	459	CTG(Leu) → TTG(Leu)	C 0.63, T 0.37
7	487	AGC(Ser) → AAC(Asn)	G 0.72, A 0.28
8	515	ACG(Thr) → ACA(Th)	G 0.79, A 0.21
Intron 1	nt – 91	A → G	A 0.88, G 0.12
Intron 1	nt – 42	G → A	G 0.66, A 0.34
Intron 2	nt – 51	T → A	T 0.85, A 0.15
Intron 2	nt – 23	C → T	C 0.88, T 0.12
Intron 5	nt – 47	C → T	C 0.99, T 0.01
Intron 7	nt 7	G → A	G 0.57, A 0.43
Intron 9	nt 44	C → T	C 0.96, T 0.04
Intron 9	nt – 24	T → C	T 0.59, C 0.41

The frequency of each polymorphism was determined by genotyping 20–56 unrelated normal healthy white subjects. DNA polymorphisms found in introns are noted with respect to the splice donor or acceptor site. nt, Nucleotide.

Problem 21 DNA is subject to constant physical attack from environmental and endogenous sources. Heat-induced DNA lesions occur at a significant rate even at ordinary mammalian body temperatures. One example is shown in the figure below. This lesion is induced by a heat-induced rupture of the normal bond between guanine and ribose and subsequent reattachment of the guanine to the ribose via a novel bond. This so called "neoguanine" base-pairs with regular guanine.



a neoguanine • guanine mispair

➔ **Show how this type of spontaneous lesion could cause a missense mutation where a histidine (his) replaces an aspartic acid (asp). Note you can start with either asp codon**.

FIRST: indicate the mutational event in codon language:

__ __ __ ➔ __ __ __

SECOND indicate the type of mutation at the DNA level (use proper jargon)
one word:

THIRD: On a separate sheet of paper: draw a diagram to show how the mutation occurs.
Indicate neoguanine as G*

- *label 5' and 3' ends of all DNA strands*
- *indicate the strand that serves as the template for RNA polymerase during transcription*
- *be sure to establish the mutation in both strands of DNA*

Problem 22 The small crucifer *Arabidopsis* has become a valuable model system for studying various aspects of plant development including flower morphogenesis. Mutations in an *Arabidopsis* gene called *apetala* fall into one of the following two classes:
Class 1: Consists of common mutations that are recessive and result in excessive production of petals and sepals. There are frameshift and nonsense mutations in this class.
Class 2: Consists of a much rarer type of mutation that is dominant and results in the failure to make petals and sepals.

Choose the best conclusion:

a. As indicated by Class 2 mutations, this gene is haploinsufficient and normally acts to inhibit the formation of petals and sepals

b. Class 1 mutations are loss-of-function and Class 2 mutations are gain-of-function. The wild-type function of this gene is to stimulate formation of petals and sepals

c. Class 1 mutations are loss-of-function and Class 2 mutations are gain-of-function. The wild-type function of this gene is to inhibit formation of petals and sepals

d. Class 1 and 2 mutations are both loss-of-function mutations with Class 2 having a more severe effect on protein function.

e. None of these conclusion are consistent with the data

Problem 23 The mouse *Mus musculus* has become a valuable model system for studying various aspects of mammalian development and physiology. A geneticist is interested in studying thyroid function by identifying mutant mouse strains that exhibit either hyperthyroidism (as indicated by an excessive production of thyroid hormone) or hypothyroidism (decreased production of thyroid hormone). Animals from a very large colony of mice are treated with a mutagen and twenty mutant progeny strains exhibiting abnormal thyroid function are identified. The strains fall into two classes:

- **Class 1: Hyperthyroidism.** All but one of the mutant strains fall into this category. All nineteen mutations are autosomal recessive and fail to complement each other.
- **Class 2: Hypothyroidism.** Only one mutant strain shows this phenotype, which is inherited as an autosomal dominant.
- Using various genetic and molecular techniques, the experimenter demonstrates that the same gene is mutated in the Class 1 and Class 2 strains. One class 1 mutant gene is sequenced and shown to have a mutation in a splice site consensus sequence. The class 2 mutant allele results from a missense mutation.
- *How can opposite phenotypes be generated by mutations in the same gene? Be sure to use proper genetic terminology*
- *Briefly explain how the information given above supports your interpretation. Which observations are most important?*
- *Does the normal (unmutated) gene product stimulate or inhibit production of thyroid hormone?*

Problem 24 On 11/7/05 Colgate-Palmolive announced an alliance with Introgen Therapeutics (a biotechnology company). The commercial goal of their alliance will be to incorporate gene therapy (an experimental and, so far, largely unsuccessful form of medicine) into mouth washes, gels and similar products to treat and prevent oral cancers. Introgen has put a *wild-type* copy of a type of gene found mutated in cancers into a disarmed cold virus (called adenovirus), which infects the cells and directs transport of the gene into the nucleus, where it can be transcribed. In answering this question, assume the mouth cells acquire at least two copies of the wild-type gene. [NYT 11/7/05 Business: *Gene therapy in a bottle of Mouthwash*]

- ***List the three major classes of genes mutated in cancer cells (two of these classes control the rate of cell division) and in one sentence indicate the wild-type function of each category. Use proper terminology.***
- ***For each category (categories) of genes indicate whether is would be a good candidate for this type of gene therapy. Briefly defend your answer in 1 sentence using proper terminology.***

NAME OF GENE CLASS:
General function of wild-type product:
Good candidate for gene therapy Circle YES or NO *One sentence explanation:*

NAME OF GENE CLASS:
General function of wild-type product:
Good candidate for gene therapy Circle YES or NO *One sentence explanation:*

NAME OF GENE CLASS:
General function of wild-type product:
Good candidate for gene therapy Circle YES or NO *One sentence explanation:*