Virginia State Science Olympiad Regional Tournament

2013 – Division C
Designer Genes

- You may write your name on this sheet before the event examination has begun.
- The reverse of this exam cover contains the Genetic Code; you are welcome to use your own copy.
- Two (2) non-programmable non-graphing calculators, one (1) standard-size (8.5"x11") double-sided sheet (*not two one-side sheets*) of paper containing any information, and writing implements may be used.
- Any other electronic devices are ***not*** allowed for this event; please consult the proctor about a safe location to store these devices for the duration of the event if you happen to carry one; *any team caught with an electronic device during the event will be immediately disqualified.*

Student Names: _____

_____

School: _____

Team #: KEY

PART I - TATA

As far as we are aware, all genes require promoters. In many eukaryotic genes, these promoters contain a *cis*-regulatory element called a TATA box. The transcription factor binds to this protein is called a TATA-binding protein (TBP).

1. TBP forms part of a complex protein complex generally called *Transcription Factor II D* (TFIID). What is the general purpose of these transcription factors in the first place? [1pt]

   recruitment of RNA Polymerase II to initiate transcript

2. While TATA boxes are found in many eukaroytic genes, you seldom find evidence of TATA box schemes in isolated mRNA extracts. Why is this so? [2pt]

   TATA box is not transcribed, so is not present on mRNA extract

3. There are relatively few known TBP mutations in nature. Why is this so? [2pt]

   TBP is an important, conserved gene – mutations to this protein are most likely harmful so are avoided through selection

4. The TATA box is called a "*cis*-regulatory element", which the TBP is called a "*trans*-regulatory element". What's the difference between the two? [3pt]

   cis-regulatory = on same molecule (i.e. a DNA sequence)
   trans-regulatory = on different molecule (i.e. transcription factor, which is a protein, even DNA on other chromosome)

5. The following is a fragment of the cDNA sequence of the DNA-binding domain (DBD) of the TBP gene sequence. Fill in the sequence data for the mRNA transcript, the tRNA anti-codon sequence, and the polypeptide primary sequence. [6pt]

cDNA sequence (5′ -> 3′)
AGAATTGTTCTCCTT ATTTTTGTTTCTGGA AAAGTTGTATTAACA

mRNA sequence (5′ -> 3′)
AGAAUUGUUCUCCUU AUUUUUGUUUCUGGA AAAGUUGUAUUAACA

tRNA anti-codon sequence
UCU AAU AAC GAG AAG AAU AAA AAC AGA UCC UUU AAC UAC UAA UGU

polypeptide primary sequence
R I V L L I F V S G F V V L T

14

Points:_____

6. Another known mutation of the TBP gene produces the following sequence over the same domain:

cDNA sequence (5′ -> 3′)

AGATTTGTTCTCCTT ATTTTTGTTTCTGGA AAGACTGTAGTAACA

A AGT T

polypeptide primary sequence

R F V L L I F V S G F T V V T

Annotate the five (5) sequence changes above and write in the new polypeptide sequence. [5pt]

7. Most of the changes in question 6 are probably an example of which kind of mutation? [1pt]
   a. frameshift (in/del)
   b. point substitution
   c. polynucleotide expansion
   d. inversion
   e. double-stranded break

8. One of the changes indicated in question 6 does not alter the primary sequence of the polypeptide. Of what phenomenon is this an example? [1pt]

codon redundancy → nucleotide degeneracy

9. The formation of the preinitiation complex requires the recruitment of multiple transcription activating factors (TAFs), which help bind RNA polymerase to the transcription origin. Explain briefly how this process helps speed up the rate of transcription. [3pt]

The binding of the multiple TAFs helps increase the binding affinity of the PIC to RNA polymersase, speeding up the process of initiation. Moreover, by selectively associating certain proteins to one location, it reduces the effect of stochastic dispersion, which slows the process of transcription considerably.

10. The mutation indicated in question 6 alters the TBP structure sufficiently to enhance DNA binding. What effects on transcription might result from this change? [2pt]

If DNA binding is enhanced, transcription might be made uncontrollably faster – a pretext for proto-tumor development in some cells

11. Part of the preinitiation complex includes multiple domains that influence helicase and ATPase activity. Briefly explain the purpose of both: [4pt]

helicase - separates the DNA double-strand to allow RNA polymerase access to the template strand

ATPase - helps phosphorylate appropriate proteins to modify their structure to either repress or enhance TAF affinity, consumes ATP in the process

Points: 16

PART II - Smorgasbord!

The following statements compare/contrast replication and transcription. Determine whether each statement is true or false, and briefly explain your choice if the statement is false. [2pt each]

12. RNA polymerase is required to form primers for replication and transcription.

F – primase is used in replication

13. Topoisomerase is only used to unkink DNA strands during replication, not during transcription.

T

14. Okazaki fragments are produced by the lagging strand during transcription.

F – Okazaki fragments are produced during replication

15. DNA ligase is required to bind together the leading and lagging strand after replication is complete.

T

16. In replication, only DNA molecules are synthesized.

F – primase creates a RNA primer, which is later destroyed

17. In transcription, only RNA molecules are synthesized.

T

18. Deoxyribonucleoside-triphosphates are used for replication, while ribonucleoside-triphosphates are used for transcription.

T

19. Termination of replication usually involves the formation of a poly-adenylated tail on the lagging strand.

F – Termination of transcription usually involves some mRNA alterations

20. Termination of transcription usually involves the formation of a poly-adenylated tail on the mRNA transcript.

T

21. Transcription of DNA is required for cells undergoing mitosis.

T – transcription is almost always required

22. Replication of DNA is required for the proper expression of genes.

F – terminal cells never replicate, but generally function fine (as long as DNA repair apparatus is okay

Points: **20**

23. Match the following scientists to their contributions to the field of genetics [1pt each]:

Person		Contributions	
C	Oswald Avery	**A**	Helped discover that DNA replication occurs via semi-conservative methods
F	Erwin Chargaff	**B**	Developed one of the first successful methods for in vitro sequencing of DNA
K	Francis Crick	**C**	One of the first researchers to suggest that DNA was the "transforming principle" after confirming the results of prior experiments
M	Frederick Griffith	**D**	One of the first researchers to propose the "one gene, one protein (enzyme)" hypothesis
L	Alfred Hershey	**E**	First to associate genes to chromosomes, and identified some of the first few mutations associated to chromosomal defects
H	Gregor Mendel	**F**	Discovered that the proportion of adenine/thymine and cytosine/guanine residues are equal to one another, and that A+G=C+T
A	Matthew Meselson	**G**	First to extract and characterize DNA
G	Friedrich Miescher	**H**	Coined "Father of Genetics" - discovered that genes assort separately and independently, leading to the famous 3:1 ratio of observable traits in a monohybrid cross
E	Thomas Morgan	**I**	Helped discover the structure of DNA
N	Kary Mullis	**J**	Helped discover the genetic code using synthetic ribonucleotide sequences
J	Marshall W. Nirenberg	**K**	In addition to discovering the structure of DNA, helped discover that DNA is transcribed via a three-nucleotide set called "codons"
B	Frederick Sanger	**L**	Helped determine that DNA - as opposed to proteins - were the molecules responsible for heredity
D	Edward Tatum	**M**	Discovered that bacteria can absorb a "transforming factor" that gave it characteristics of recently killed bacteria
I	James Watson	**N**	Helped develop the process of in vitro replication and amplification of DNA sequences

PART III – Forensics!

You are working at a forensics lab, and unlike all of those people masquerading as "forensics experts", you're going to do all of your work in the lab. Your first case is to determine the paternity of an infant child. A multi-billionaire heiress has suddenly died, leaving her infant daughter in the capable hands of the woman's husband (A). However, five other men have come forward claiming paternity of the child: a famous celebrity (B); a random paparazzo photographer (C); a foreign prince (D); a former bodyguard of the late heiress (E); and the butler (F).

You first decide to test the paternity of the child using the ABO blood test. You obtain the following results:

	Child	**Mother**	**Person A**	**Person B**	**Person C**	**Person D**	**Person E**	**Person F**
Bloodtype	O	O	A	A	O	B	AB	O

24. Explain briefly why it is *not* necessary to include the heiress's blood type as a standard for comparison or a reference. [2pt]

Although extremely useful, the maternity of the child is (hopefully) not in question. The use of the mothers as a reference is a tool for resolving the paternity.

25. Based on the data above, calculate the probability that Person A cannot be excluded as the actual father of the child; be sure you include *all* possibilities in your calculations. [4pt]

father must donate an "O" allele; father has "A" phenotype

*there are two genotype possibilities for the father: AA and AO

if father is AA, probability of paternity = 0%
if father is AO, probability of paternity = 50%

Since there are *two* possibilities, the probability of paternity = $\frac{0\%+50\%}{2}$ = **25%**

26. Based on the data above, who is the only person who can be *absolutely* excluded as a potential father? [1pt]

Person E – AB is only genotype that cannot give an "O" allele.

27. Based on the data above, which individuals *cannot* be absolutely excluded as a potential father? [2pt]

Persons C and F – both have only "O" alleles

28. Explain briefly why using ABO markers is a very *bad* system for distinguishing individuals apart in forensic cases. [2pt]

ABO markers are not very discriminating: too few different alleles, moreover each polymorphism is very common

You decide replicate the previous experiment, including the Rh antigens this time. With this setup, you obtain the following data:

	Child	Mother	Person A	Person B	Person C	Person D	Person E	Person F
Bloodtype	O+	O-	A+	A-	O+	B+	AB+	O+

29. Assuming that Person D is heterozygous at the Rh locus (+/-), construct a Punnett square showing a cross between Person D and the heiress. What is the probability that the Person D could have fathered the child? [4pt]

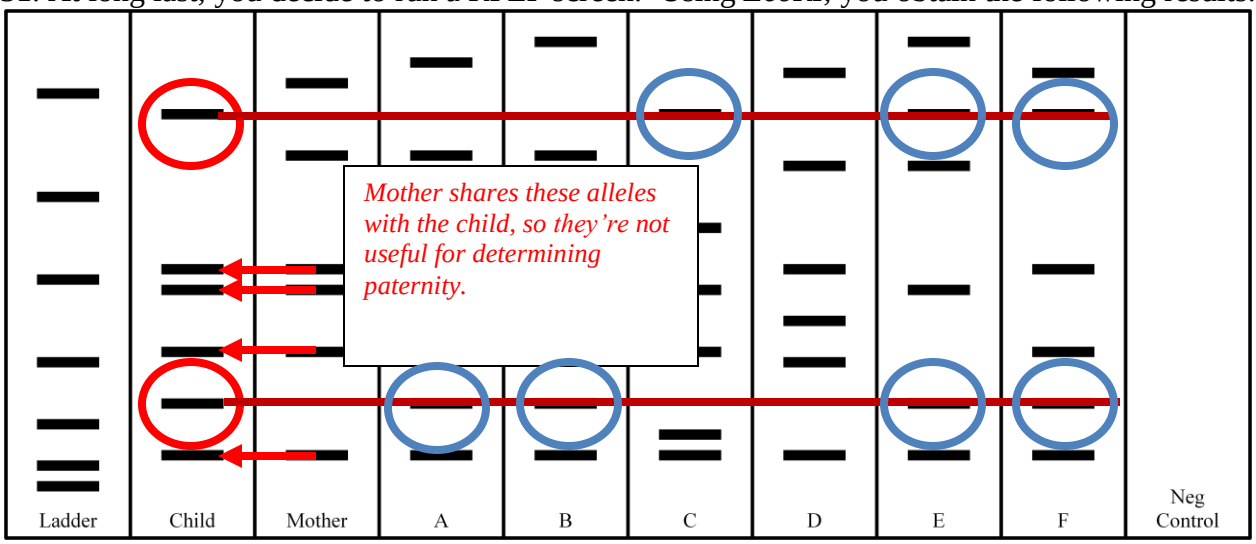
Child is O+, only one (1/4)
There are *two* possible genotypes for person D, one which is impossible for paternity, so half the probability = 1/8 = **12.5%**

♀\♂	B+	B-	O+	O-
O-	BO+/-	BO-/-	**OO+/-**	OO-/-

30. Based on this information, which individuals can be *absolutely* excluded as possible fathers (including those you'd determined for question 26)? [2 pt]

Person E – AB cannot give "O" allele
Person B – -/- cannot give "+" allele

31. At long last, you decide to run a RFLP screen. Using *EcoRI*, you obtain the following results:



On the results, circle the bands that are most informative for determining paternity [2pt].

TIE-BREAKERS

(1) What does RFLP mean?

Restriction (Endonuclease) **F**ragment **L**ength **P**olymorphism

(2) Determine the probable identity of the father.

Person F – the Butler : The only person whose genetic profiles were capable of giving the necessary missing alleles for the child (Person E had been already eliminated based on ABO profile)

Points:__8___

PART IV – Super Mice!

In an attempt to create knock-out mice, you have accidentally created a strain of mice that actually knock-out researchers with super-mouse strength (which is sufficient to knock out people). After cordoning the area off, you start to use your knowledge of genetics to see if there is anything that you can do.

32. The genes that you have mutated appear to be responsible for a transcription factor. How can a mutation on this type of gene have such a drastic effect on the mouse phenotype and physiology? [3pt]

Transcription factors can alter the expression of multiple genes – a potential pleiotropic effect. In this case, it is possible the mutated transcription factor is responsible for muscle development (or lack thereof).

33. Fortunately, it only appears that a minority of your mutated mice carry the knock-out effect. Moreover, the knock-out phenotype is most prevalent is male mice, whereas female mice can *only* have the phenotype if the father exhibits the knock-out effect. Based on this limited information, which of the following best describes the inheritance pattern? [1pt]

a. Autosomal dominant
b. Autosomal recessive
c. X-linked dominant
d. X-linked recessive
e. Y-linked

34. Assuming that one of the male knock-out mice escapes into the general mouse population and starts breeding, calculate the percentage of knock-out (phenotype) mice that would be created in the **first** generation. *Use a Punnett square to calculate your work.* [4pt]

♀\♂	X*	Y
X	XX*	XY
X	XX*	XY

no knock-out mice are possible in the first (F1) generation

35. Fortunately, all of the mice in your lab are labeled with a protein that causes the mice to glow in the dark. Unfortunately, the glowing and knock-out genes are *not* linked. Construct a Punnett Square in the space below to help calculate the probability a mouse with the knock-out allele (genotype) *also* carries the glow-in-the-dark gene, which is an autosomal dominant trait. [5pt]

♀\♂	X*G	YG	X*g	Yg
Xg	XX*Gg	XYGg	XX*gg	XYgg
Xg	XX*Gg	XYGg	XX*gg	XYgg

G = glowing
g = not glowing

Of the mice with the mutant knock-out allele, 50% also have the glowing gene
If glowing genotype is presumed unknown, actual probability is 75%.
Maximum actual probability of 100% if knock-out mice are homozygous for glowing.
FOR FULL CREDIT – must explain the probabilities

Points:___**13**___

36. In question 35, why is it important to distinguish whether the two genes are linked? [3pt]

Linkage affects the assortment of alleles during meiosis. If alleles are linked, a fewer number of allele combinations are possible for the gametes.

37. Unfortunately, some of the mice have escaped – a total of 20 male and 20 female knock-out (phenotype) mice. Assuming that the normal mouse population contains 1000 male and 1000 female mice, calculate the *allele frequency* of the knock-out gene compared to the wildtype (non-mutated) gene. [4pt]

1000 XX (2000 normal)
1000 XY (1000 normal)
40 X*X* (80 mutant)
40 X*Y (40 mutant)
3000 normal (p) + 120 mutant (q) = 3120 total alleles
p = 3000/3120 = **96.2% = 0.962**
q = 120/3120 = **3.8% = 0.038**

38. Fortunately, it appears that the knock-out mice selectively mate only with other knock-out mice. Why is this advantageous for you? [3pt]

Selective mating reduces the effective population size. Instead of a potential population of 2040 mice, you are limited to just 40 mice. It also avoids spreading the mutant alleles in the native population.

39. It is discovered that the mutated gene in the knock-out mice no longer works, and by no longer working produces the phenotype. In an unusual "fix" to your genetics problem, you propose to "rescue" the mutant phenotype with a normal, functioning gene. Propose briefly how you intend to do this; include a brief discussion about the costs and benefits of the methodology that you specifically choose. [6pt]

Answers will vary:

Viral vectors

Non-viral vectors (direct-injection *not* acceptable)

TIE-BREAKERS

(1) In genetics parlance, explain what "knock-out" means.

knock-out: a gene whose function has been eliminated (loss of function = LOF mutation)

(2) Explain briefly why "rescuing" the mutant phenotype in adult mice would be futile.

Adults already have the phenotype – it's much easier to kill the mice than to attempt rescuing the phenotype would take a lot of time for muscle tissue to lose their tone.

PART V – Create a gene!

You want to create a synthetic gene that produces a final pentapeptide product with the following sequence: Met-Ala-Pro-Lys-Glu. However, you want to be sure that the peptide product is produced only if enough lysine is present. Moreover, you want to ensure that the peptide product is synthesized only in specific cell lines. Using your knowledge of gene expression, your task is to design a simple template for creating your gene.

40. Using the space below, draw a simple sketch showing the relevant sequences of DNA (no actual DNA sequences – use boxes) for your gene. Be sure to include *cis*- and *trans*-regulatory elements and any necessary elements. Do *not* include any introns. [4pt]

cis-factors:
Promoters! –must include one that mentions specific targets for particular (any) cell line.
5'UTR-*Coding Sequence*-3'UTR (promoters may be upstream or downstream)
trans-factors:
Transcription factors; lysine repressor (which binds to a promoter sequence)

41. Assuming that you've managed to sequence all of the relevant regulatory elements, your next task is to design the actual DNA sequence to code for the pentapeptide product. Be sure to include the untranslated regions (UTR), the coding sequence (CDS), and the complementary strand. [4pt]

TATA[any short 10-12 sequence is fine]
ORIGIN[start of 5' UTR] – any short sequence is fine
Coding sequence: 3' –|TAC CG(N) GG(N) [AA(T/C)]OR[GA(N)] CT(T/C) [AT(T/C)]OR[ACT]|–5'
3'UTR

42. Using the sequence above that you designed, transcribe the sequence into a polished mRNA transcript – do not forget the poly-A tail and the 5'-cap. Underline the coding sequence of your transcript as well. [4pt]

SEQUENCE MUST BE COMPLEMENTARY TO Q41
5'cap (7-methylguanosine triphosophate)
ORIGIN[start of 5' UTR]
Coding sequence: 3' –|AUG GC(N) CC(N) [UU(A/G)]OR[CU(N)] GA(A/G) [UA(A/G)]OR[UGA]|–5'
3'UTR (poly-A tail)

43. Assuming that you've managed to get everything all synthesized properly, you now have a functioning gene, but no way to verify that it works. Moreover, you don't have enough copies to work with, so you decide to clone your whole gene template into a plasmid. What is the main benefit of using plasmids for gene cloning? [2pt]

Plasmids can selectively amplify sequences that you want with high fidelity, while PCR may amplify sequence artifacts with lower stringency, polluting your intended sequence for study.

44. You choose a plasmid that was engineered to include the entire *lac* operon. What is the purpose of including the *lac* operon on this plasmid? [2pt]

the *lac* operon will be used in a blue-white assay to determine the success of transformation and ligation

45. The plasmid also contains a gene that confers to the host bacterium resistance to antibiotics. What is the purpose including this gene on the plasmid? [2pt]

the antibioitic-resistance gene will preserve bacteria that have the plasmid, while bacteria without the plasmid will die in a medium containing antibiotics

46. In the middle of the *lac*-Z gene is a restriction that is recognized by the *HinDIII* restriction endonuclease. When the restriction site is cleaved, *lac*-Z is broken. What is the net effect of this action? [2pt]

without a functioning *lac* Z gene, the bacteria will be unable to digest lactose

47. After digesting the plasmid lightly in *HinDIII*, you add in your own sequence and attempt to transform competent *E.coli* bacteria to clone the gene. It ends in failure. What was your mistake? [2pt]

cloned sequence needs to be ligated into the plasmid; add ligase before attempting to transform competent bacteria.

48. Having fixed the mistake in the previous question, you successfully transform bacteria and try to streak plates coated in X-gal and ampicillin. What is the purpose of each substance? [4pt]

X-gal: **lactose analog – when digested, surrounding medium turns blue (release of indole dyes)**

Ampicillin: **antibiotic – bacteria without plasmid will die/not reproduce; bacteria with plasmid will be immune to antibiotic and grow**

49. When you come back to select your colonies, you discover that most colonies are blue, while a few are white. What is the meaning of each colored colony and which one(s) do you want to select for further analysis? [3pt]

blue: X-gal has been digested, *lac-Z* activitiy has been repaired, gene not successfully ligated
white: X-gal not digested; *lac*-Z activitiy has been disrupted, gene successfully (maybe) ligated.

pick a white colony (not white with blue in the middle!)

50. Having picked the right colony, you discover the gene that you have constructed was successfully cloned. The next big step is to transfect the synthetic gene into a new cell line and to integrate the gene into the genome. Without using a viral vector, identify three common methods of transfection to get your gene into the cells. [3pt]

non-viral methods of transfection:
liposome-assisted transfection (using lipid-bound membranes to carry DNA into cell)
direct injection (injection of native DNA molecule into cell, and hope for successful integration)
electropolation (shocking the cell membrane lightly, DNA will be absorbed into the cell)
cationic polymers (using positively charged particles to shuttle DNA into cell)
gene guns (using heavy atoms to shoot (literally) DNA into the cell)

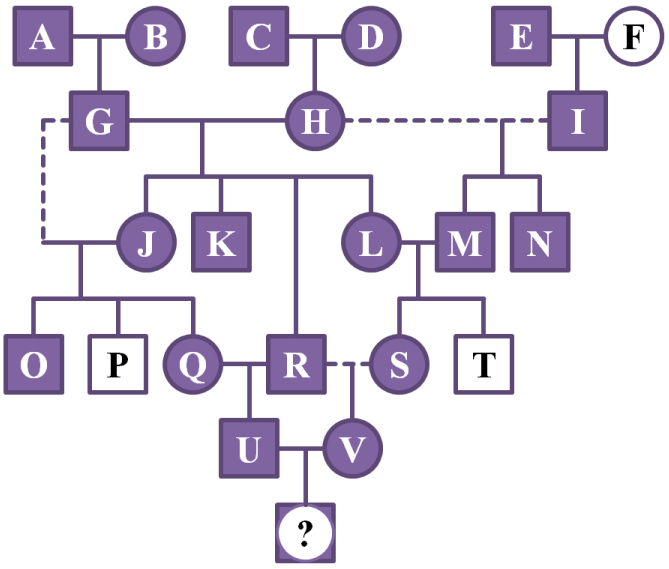
TIE-BREAKER

You notice – belatedly – that your gene acquired a mutation, such that the 15th nucleotide in your coding sequence is now a "G". How will this change affect your gene product?

GA(A/G) [E] → GAC [D]

Aspartic acid is sometimes analogous to glutamic acid.

PART VI – Smorgasbord Again!



Using the pedigree to the right, answer questions 51 to 55

51. Which of the following is <u>not</u> an inbred (intrafamilial) mating? [1pt]
    a. G+J
    b. I+H
    c. R+Q
    d. R+S
    e. U+V

52. Assuming that "shaded" boxes show a particular trait, while "non-shaded" boxes do not show the trait, which of the following explains the inheritance pattern? [1pt]
    a. Autosomal dominant
    b. Autosomal recessive
    c. X-linked dominant
    d. X-linked recessive
    e. Y-linked

53. Given the information in the pedigree, which of the following best describes genotype of <u>Individual I</u>? [1pt]
    a. homozygous dominant
    b. homozygous recessive
    c. codominant
    d. hemizygous recessive
    e. heterozygous

54. Which of the following statements are true? [1pt]
    I. The genotype of Individual G must be heterozygous.
    II. The genotype of Individual H must be heterozygous.
    III. The genotype of Individual M must be heterozygous.

    a. I only
    b. II only
    c. I & II
    d. I & III
    e. I, II, & III

55. Putative Individual "X" (I know I skipped "W") is the resulting offspring of Individuals U and V. Calculate the probability that Individual X will be homozygous recessive for the trait indicated on the pedigree (*hint*: calculate each parent's genotype probabilities first!) [6pt]

parents have *two* genotypes possible: AA and Aa; fortunately, the parents both have a similar familial background:
Q/S has probability of 1/3 AA and 2/3Aa – 2/3 probability of giving "A" and 1/3 probability of giving "a"
R has probability of 3/8 AA and 5/8 Aa (due to ambiguous nature of mother H) – 11/16 probability of giving "A" and 5/16 probability of giving "a"
Probability that both parents are Aa = (21/48)^2 = (7/16)^2 = 49/256 = 19.14%

Probability that both parents give two "a" alleles = 25% * 19.14% = <u>4.785%</u>
1/4 * 49/256 = <u>49/1024</u>

56. In a special culture of paramecia, "AA" individuals are red, "BB" individuals are blue, while "AB" individuals are purple. Which of the following best describes this system? [1pt]

a. complete dominance
b. incomplete dominance
c. codominance
d. epistasis
e. all of the above

57. In a strict cross between only red paramecia and only blue paramecia, what is the expected ratio of phenotypes? [1pt]

a. 50% red; 50% purple; 50% blue
b. 100% red; 0% purple; 0% blue
c. 0% red; 0% purple; 100% blue
d. 0% red; 100% purple; 0% blue
e. 25% red; 50% purple; 25% blue

58. You observe a population of paramecia and count the numbers of individuals as follows:

Red	Purple	Blue
2314	2246	7293

Calculate the proportion of "A" and "B" alleles in the population. [3pt]

AA = 2314 (A=2314)
AB = 2246 (A =1123; B = 1123)
BB = 7293 (B=7293)
Total = 11853
A = 3437/11853 = 29.0% = 0.290 || B = 8416/11853 = 71.0% = 0.710

59. Is this population currently in Hardy-Weinberg equilibrium? [2pt]

A + B = 1 | [0.290 + 0.710 = 1]
$(A + B)^2 = 1$ | $A^2 + 2AB + B^2 = 1$| 0.0841 + 0.4118 + 0.5041 = 1 (11853)
AA = 997 | AB = 4881 | BB = 5975
The population is not in HWE.

60. After counting all of the paramecia, you realize that you accidentally spilled a culture of another, more aggressive paramecia (which are not red, blue, or purple). You observe changes to your existing culture of paramecia after the first two generation of changes:

	Red	Purple	Blue
1st	1157	1797	3647
2nd	579	2396	1824

Calculate the proportion of "A" and "B" alleles in the population for *each* generation. [6pt]

AA = 1157 (A=1157)
AB = 1797 (A =898.5; B = 898.5)
BB = 3647 (B=3647)
Total = 6601
A = 2055.5/6601 = 31.1% = 0.311 || B = 4545.5/6601 = 68.9% = 0.689

AA = 579 (A=579)
AB = 2396 (A =1198; B = 1198)
BB = 1824 (B=1824)
Total = 4799
A = 1777/4799 = 37.0% = 0.370 || B = 3022/4799 = 63.0% = 0.630

61. It is evident after two generations that the purple paramecia have some sort of advantage that is not evident in either the red or blue paramecia. This is a very explicit example of what genetics phenomenon? [1pt]

heterozygote advantage