2025 SOAPS Invitational

Designer Genes C

Science Olympiad at Penn State Invitational

Saturday, January 17th, 2026

The Pennsylvania State University, University Park, PA

Team Name:_____

Participant Names:_____

Total Score:____/318 **Rank:_____**

This test was written by Krish Shah (Carnegie Mellon University '29)

Please direct any questions to krishshah@cmu.edu

1. (1 point) An organism with both a dominant and recessive allele for a specific gene (i.e. has two different alleles for one gene) would best be described as which of the following?

   A. Hemizygous **B. Heterozygous** C. Homozygous D. Nullizygous

Tay-Sachs disease is inherited through an autosomal recessive pattern of inheritance. Let R represent the dominant allele and r represent the recessive allele.

2. (1 point) What is the genotype of an individual with Tay-Sachs disease?

   A. RR B. Rr **C. rr** D. There are multiple possible genotypes

3. (1 point) What is the genotype of a carrier for Tay-Sachs disease?

   A. RR **B. Rr** C. rr D. There are multiple possible genotypes

4. (1 point) If a homozygous dominant parent has a child with an individual with Tay-Sachs disease, what will the most probable genotype of their child?

   A. RR **B. Rr** C. rr D. They are all equally likely

5. (2 points) If two unaffected parents have a child with Tay-Sachs disease, what is the probability of their second child also having Tay-Sachs disease?

   A. 1/4 B. 1/3 C. 1/2 D. 2/3 E. 3/4

6. (5 points) If one parent is known to be a carrier, and the other parent does not have Tay-Sachs disease and had one parent with no family history of Tay-Sachs disease and another that was a carrier, what is the probability of their second child having Tay-Sachs disease given that their first child does not have Tay-Sachs disease? The second parent does not know their genotype. **(TB)**

   A. 1/16 B. 1/8 **C.** 3/28 D. 1/7 E. 1/6 F. 2/15 G. 1/5 H. 5/24 I. 3/16 J. 3/8

Snapdragon flowers can be red, pink, or white, and this flower color is controlled by a gene. Plants with two red alleles have red flowers, plants with one of each allele have pink flowers, and plants with two white alleles have white flowers.

7. (1 point) The color of the flower is best described as which of the following?

   A. Allotype B. Genotype **C. Phenotype** D. Karyotype

8. (2 points) What pattern of inheritance is this most emblematic of?

   A. Codominance **B. Incomplete dominance** C. Incomplete penetrance D. Overdominance

9. (2 points) If a red plant and a white plant are mated, what proportion of offspring would be expected to be white?

   A. 1/4 B. 1/3 C. 1/2 D. 2/3 E. 3/4

10. (2 points) In the same cross (red with white), what proportion of offspring would be expected to have pink flowers?

    A. 1/4 B. 1/3 **C.** 1/2 D. 2/3 E. 3/4

11. (2 points) If two plants have a red flowered offspring, which of the following could have been the phenotypes of the parents? (select all that apply)

    A. Red/Red **B. Red/Pink** **C. Pink/Pink** D. Pink/White E. White/White

12. (2 points) In humans, the effects of mutations in *BRCA2* only appear in the phenotype about half of the time. What pattern of inheritance is this most emblematic of?

    A. Codominance B. Incomplete dominance **C. Incomplete penetrance** D. Overdominance

13. (1 point) Which of the following is the first stage of mitosis?

A. Anaphase B. Metaphase **C. Prophase** D. Telophase

14. (1 point) A human cell with 45 chromosomes would most likely be an example of which of the following?

A. Aneuploidy B. Pentaploidy C. Polyploidy D. Trisomy

15. (2 points) How many centrioles are in a typically eukaryotic centrosome?

A. 0 B. 1 **C. 2** D. 3 E. 4 F. 5 G. 6 H. 7 I. 8 J. 9 K. 10

16. (2 points) How many daughter cells are typically produced from one mother cell in meiosis?

A. 1 B. 2 C. 3 **D. 4** E. 5 F. 6 G. 7 H. 8

17. (1 point) If a somatic cell in a specific species has $2n$ chromosomes, then how many chromosomes will the daughter cells have?

A. $n/2$ B. $n-1$ **C. n** D. $2n-1$ E. $2n$ F. $4n$

18. (2 points) In which phase of mitosis does the mitotic spindle pull the chromosomes to be in the center of the spindle poles?

A. Anaphase **B. Metaphase** C. Prophase D. Telophase

19. (2 points) In which phase of mitosis does the nuclear envelope reform around the two sets of chromosomes?

A. Anaphase B. Metaphase C. Prophase **D. Telophase**

20. (3 points) Which phase of mitosis starts with the ubiquitylation of securins?

A. Anaphase B. Metaphase C. Prophase D. Telophase

21. (2 points) A cell treated with an inhibitor of separases would most likely be arrested at the start of or before which phase of mitosis?

A. Anaphase B. Metaphase C. Prophase D. Telophase

22. (2 points) When are sister chromatids separated in meiosis?

A. Metaphase I B. Anaphase I C. Telophase I D. Metaphase II **E. Anaphase II**

23. (2 points) In which phase of meiosis do homologs align on a plane between the two centrosomes?

A. Metaphase I B. Anaphase I C. Telophase I D. Metaphase II E. Anaphase II

24. (3 points) In which phase of prophase I does crossing over occur in?

A. Diakinesis B. Diplotene C. Leptotene **D. Pachytene** E. Zygotene

25. (2 points) If two daughter cells contain n chromosomes and the other two contain $n+1$ and $n-1$ chromosomes, this nondisjunction likely resulted from an error in which phase of meiosis?

A. Metaphase I B. Anaphase I C. Telophase I D. Metaphase II **E. Anaphase II**

26. (2 points) If a drought occurs and individuals of a specific plant species with deeper root systems are favored and more easily able to survive, then (assuming depth of roots is inherited), what type of selection is this most emblematic of?

A. Directional B. Disruptive C. Equalizing D. Stabilizing

27. (1 point) If a finch is born with a mutant allele that increases its fitness, then how would the frequency of this allele be expected to change over time assuming the allele is not lost by genetic drift?

A. Decrease B. Stay the same **C. Increase**

28. (1 point) In which of the following population sizes would genetic drift have the most significant effects?

A. 10 B. 100 C. 1000 D. 10000

29. (1 point) If 100 birds are blown off to a faraway island, and they cannot easily leave, how would the genetic diversity of the population on this island compare to the original population (assuming the original population is significantly larger)?

A. Lower B. Same C. Higher D. Not enough information

30. (1 point) According to the central dogma, which of the following stores genetic information?

A. Carbohydrates **B. DNA** C. Proteins D. RNA

31. (1 point) Which of the following is not a stop codon?

A. UAA B. UAG C. UGA **D. UGG**

32. (1 point) A mutation which changes a codon from tyrosine to a stop codon would best be described as which of the following?

A. Frameshift B. Missense **C. Nonsense** D. Silent

33. (1 point) A mutation that inserts a new nucleotide within a prokaryotic open reading frame would most likely be described as which of the following?

A. Frameshift B. Missense C. Nonsense D. Silent

34. (1 point) In which direction(s) can DNA polymerase III add nucleotides? (select all that apply)

A. 3' to 3' B. 3' to 5' **C. 5' to 3'** D. 5' to 5'

35. (2 points) Which of the following best describes the structure of a typical bacterial and eukaryotic chromosome, respectively?

A. Circular, circular **B. Circular, linear** C. Linear, circular D. Linear, linear

36. (2 points) Which of the following is the linker histone? **(TB)**

A. H1 B. H2A C. H2B D. H3 E. H4

37. (2 points) What is the expected ratio between the amount of the histone in question 36 and the other histones in a typical eukaryotic chromosome?

A. $\frac{1}{4}$ **B.** $\frac{1}{2}$ C. 1 D. 2 E. 4

38. (3 points) Which of the following is the most common form of DNA?

A. A-form DNA **B. B-form DNA** C. Y-form DNA D. Z-form DNA

39. (2 points) In bacteria, which of the following is used to replace the RNA primers with DNA nucleotides?

A. Helicase **B. DNA polymerase I** C. DNA polymerase II D. DNA polymerase III

40. (4 points) Which of the following is used to induce negative supercoiling in DNA?

A. DNA gyrase B. DNA topoisomerase IIα C. DNA topoisomerase IIβ D. Topoisomerase IV

41. (2 points) In its natural form inside a eukaryotic nucleus, chromosomes are most commonly in what form?

A. Achromatin **B. Euchromatin** C. Heterochromatin D. Homochromatin

42. (2 points) In a eukaryotic chromosome, what is the shorter arm referred to as?

A. m B. n **C. p** D. q

43. (3 points) Which of the following best describes the y chromosome in humans? Hint: there are pedigrees later in the test you can use to look at the y chromosome :)

A. Acrocentric B. Metacentric C. Paracentric D. Submetacentric

44. (3 points) Which of the following DNA repair methods would most likely be used to repair a thymine dimer?

A. Base excision repair B. HRR C. NHEJ **D. Nucleotide excision repair**

45. (1 point) Which of the following is the principle enzyme in transcription?

A. DNA polymerase III B. Ribozymes **C. RNA polymerase** D. Topoisomerase

46. (2 points) A mutation that substituted a arginine for which of the following would likely have the smallest effect on the function of the protein?

A. Aspartic acid B. Glycine **C. Histidine** D. Serine E. Tryptophan

47. (3 points) An aminoacyl-tRNA bound to the A site in a ribosome would move to which site after one amino acid is added?

A. A B. E **C. P** D. W

48. (2 points) Regulation of gene expression by a repressor is best described as which of the following?

A. Ambivalent control **B. Negative control** C. Neutral control D. Positive control

49. (2 points) Presence of glucose would have what effect on the expression of the lac operon?

A. Decrease B. Stay the same C. Increase

50. (2 points) In a 10^6 bp long sequence of DNA, about how many cuts would you expect a restriction enzyme to make if it recognized the sequence TGCTAA and all nucleotides appear randomly?

A. 100 **B. 250** C. 500 D. 1000 E. 2500

51. (a) (1 point) How many X chromosomes do human males have?

(b) (2 points) Color blindness is a X-linked recessive trait. Fill in the pedigree on your answer sheet for the children given that the father is not color blind, but the mother is by shading in the children that will have the disease and shading in half for carriers.

(c) (2 points) Fill in the pedigree on your answer sheet for a mitochondrial inherited disease as in the previous question (affected children shaded, carriers half shaded).

(d) (4 points) In some dogs, fur color is controlled by two genes. When the second gene has at least one dominant allele, if the first gene has a dominant allele, the fur is black, wheras the recessive phenotype is brown fur. However, if the second gene has both recessive alleles, the fur will be yellow, regardless of what alleles are present for the first gene. What proportion/ratio of each phenotype would be expected for the offspring of a dihybrid cross?

(e) (4 points) In some plants, flower color is controlled by two genes. If either gene expresses the recessive phenotype, the flowers are white. When both express the dominant phenotype, the flowers are pigmented. What proportion/ratio of each phenotype would be expected for the offspring of a dihybrid cross?

Assume that in this mystery plant, flower color and size are linked with blue and large dominant, and that one parent was heterozygous and the other was a true breeding purple small plant.

Phenotype	Frequency
Blue, large	231
Blue, small	16
Purple, large	10
Purple, small	243

(f) (2 points) Which of Mendel's laws is violated by linkage?

(g) (3 points) Estimate the distance between these loci in cM.

52. (a) (2 points) What are the 3 steps to PCR?

(b) (2 points) Which step in PCR has the highest temperature? Why?

(c) (3 points) In PCR, there is typically a GC clamp, a guanosine or cytosine near the 3' end of the sequence, in the primer. Why is it preferred to have a G or C (i.e. why is G/C better than the other bases)? What molecular quality makes them better? **(TB)**

(d) (3 points) What is the melting temperature of a primer/oligonucleotide, and how does the melting temperature change with G/C content?

(e) (2 points) Approximately what is the optimal length for a PCR primer?

(f) (4 points) What is the primary problem with having primers that are too long? What about primers that are too short?

(g) (2 points) What is the biggest problem with a primer of an adequate length that is a palindrome? You may assume that it is in the optimal length range.

(h) (4 points) What is the principal enzyme used in PCR, and which domain is the species it originates from. Why is it used?

(i) (4 points) What are dNTPs and rNTPs, and which (if any) are included in PCR master mix (the solution containing required components such as the enzyme from the previous part)? **(TB)**

(j) (4 points) What would be the effects of using master mix contaminated with a wide assortment of nucleases for PCR? If you visualized it with gel electrophoresis, what would it look like (how many bands, how bright, etc. relative to if you did it properly)?

(k) (4 points) What does the RT in RT-PCR stand for, and what is the primary difference between this process and normal PCR?

53. (a) (4 points) What are four conditions to Hardy-Weinberg Equilibrium?

Consider a population of pea plants with purple and white flowers, with purple flowers being dominant to white flowers, where 11% of plants have white flowers. Assume this population is in Hardy-Weinberg Equilibrium

(b) (3 points) What is the frequency of each allele?

(c) (3 points) What is the frequency of purple flowers in the population, and what proportion of purple flowered plants are heterozygous? **(TB)**

(d) (2 points) If pollinators favored the less common phenotypes (e.g. when there are more purple flowers, they pollinate white flowers more frequently and vice versa), what type of selection is this most emblematic of?

(e) (3 points) If a serious drought occurred and only 5 individuals in the population survived, then how would the genetic diversity of the population change? What effect is this an example of?

Let's now consider how the population might change if the different genotypes had different fitnesses. In our scenario, purple flowered heterozygous plants are the most fit (relative fitness of 1.0), while white flowered plants are less fit (relative fitness of 0.6) and purple flowered homozygous plants even less so (relative fitness of 0.4).

(f) (2 points) Which genotype has the highest proportion of its individuals survive and reproduce? What is this phenomenon called where this genotype has the highest proportion of survival and reproduction?

(g) (4 points) What is the mean fitness of the population given that there are still 11% of individuals with white flowers?

(h) (3 points) After one generation, what proportion of the population is heterozygous? **(TB)**

(i) (1 point) After an indefinitely long amount of time, which genotype would be most common?

(j) (2 points) Why isn't the long run mean fitness equal to 1.0?

Image A

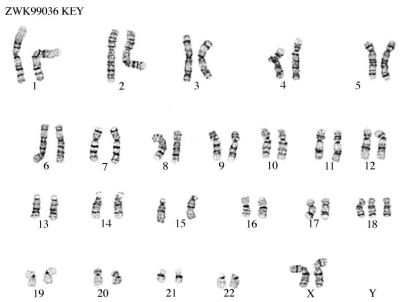

Image B

Image C

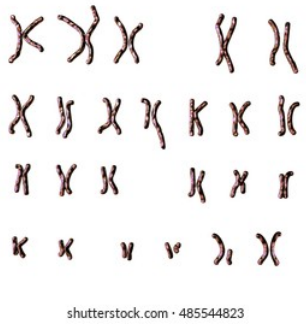

Image D

54. (a) (2 points) What is the name for pictures of this form?
    (b) (1 point) Where are the sex chromosomes located on images of this type?
    (c) (1 point) In image A, which chromosome has the abnormality?
    (d) (2 points) What is the name for the disorder caused by the abnormality in image A?
    (e) (3 points) In image B, which chromosome has the abnormality? What is the name for this type abnormality?
    (f) (2 points) In one sentence or less, how could nondisjunction in meiosis II have led to the abnormality in image B?
    (g) (3 points) What is the name for the disorder caused by the abnormality in image B?
    (h) (2 points) The abnormality in image C is the full version of that type of abnormality that is not lethal in humans. Why might this be the case?
    (i) (2 points) What is the name for the disorder caused by the abnormality in image C?
    (j) (4 points) What is the name for the abnormality in image D? **(TB)**
    (k) (3 points) Is image D an example of a Robertsonian translocation? If so, why? If not, what type is it?

55. (a) (2 points) How does an up promoter mutation affect transcription?
    (b) (2 points) How would the the region affected by a somatic mutation differ between an individual where the mutation occurred in the first month of life vs an individual where the mutation occurred after ten years?
    (c) (4 points) What is the product of deamination of cytosine, and how can this lead to mutations?
    (d) (2 points) A mutagen would likely produce what output when used in the Ames test?
    (e) (3 points) What does NHEJ stand for, and when is it used?

56. (a) (1 point) What does DNA stand for?

    (b) (5 points) In two sentences or less, what is the function of and primary difference between endonucleases and exonucleases?

    (c) (2 points) In a single stranded RNA strand, if exactly 25% of the strand is composed of nucleotides with adenine, what is the approximate proportion of the other bases? How do you know?

    (d) (3 points) If it takes bacterial enzymes 1 second to replicate 500 nucleotides, how long would it take to duplicate a 100000 nucleotide plasmid? Disregard the amount of time needed for proofreading and assume that the rate of replication is a constant 500 bp/s and the replication is bidirectional.

    (e) (2 points) What are the discontinuous segments of DNA in the lagging strand known as?

    (f) (3 points) Let's say you start with DNA with nucleotides with a heavier isotope of nitrogen, and add all necessary components for DNA replication with normal nitrogen. After exactly one round of replication, predict the result of using gel electrophoresis on the products. What model of DNA replication would this provide evidence for? **(TB)**

    (g) (2 points) What is the purpose of single-strand binding proteins?

    (h) (6 points) What is the primary difference between the process of eukaryotic and prokaryotic removal of primers? Include the primary enzyme used in both processes, but the difference should relate to the process and not the specific proteins used in either process.

57. (a) (1 point) In what order do transcription and translation occur?

    (b) (2 points) How does the energy usage of ρ (rho) dependent and ρ independent termination differ?

    (c) (3 points) What does it mean for mRNA to be polycistronic? Is this more common in bacteria or eukaryotes?

    (d) (2 points) In bacterial mRNA, what is the name of the special sequence upstream of the start codon that assists in ribosomal binding and initiation of translation?

    (e) (2 points) A mutation in the Pribnow box would likely lead to what effect in bacteria?

    (f) (2 points) In one sentence or less, what is wobble in regards to codons?

58. (a) (1 point) The *lac* operon is used to regulate what sugar?

    (b) (3 points) Is the *lac* operon inducible or repressible? Why might this be beneficial given its function?

    (c) (3 points) In high glucose conditions, is CAP activated? What does this say about the expression of the *lac* operon?

    (d) (1 point) What molecule binds to the lac repressor?

    (e) (2 points) What is the name for the region of DNA to which the repressor binds?

    (f) (4 points) Ignoring all of the other adverse effects of this mutation, a mutation that causes overactive adenylyl cyclase would likely cause what effect on the *lac* operon? Why? Not changing anything is also an effect.

    (g) (3 points) Is the *trp* operon inducible or repressible? Why might this be beneficial given its function?

    (h) (2 points) What molecule binds to the trp repressor? Is this positive or negative feedback?

59. (a) (5 points) What is the difference between ddNTPs and dNTPs? Which is used in Sanger sequencing, and what purpose do they serve?

    (b) (2 points) How does Sanger sequencing compare with Illumina sequencing for long sequences of DNA or sequencing whole genoomes?

    (c) (1 point) Are restriction enzymes needed for the Gibson process?

    (d) (2 points) What sugar is used in blue-white screens, and what color does it turn after being cleaved?

60. For this question, treat capital letters as dominant alleles and lowercase letters as recessive alleles. Let one parent plant have the genotype AaBbCcddEE and the other have the genotype AaBbccDdEE. Assume that these genes are not linked, independent, and inherited through simple Mendelian genetics.
    (a) (2 points) What is the probability of offspring having the genotype AabbCCddEe?
    (b) (2 points) What is the probability of offspring having the genotype AABBccddEE?
    (c) (2 points) What is the probability of offspring having the genotype AaBbCcddEE?

    Now, imagine that all of the genes contribute to the same trait, the height of offspring, with each gene expressing the dominant phenotype adding a certain amount to the height. For example, aabbccddee will have the minimum height, and AaBBCCDdEe would have the maximum height.
    (d) (2 points) What is the name for this type of inheritance?
    (e) (2 points) Given that there are 5 different genes, how many different possible heights exist if each allele contributes the same amount? This includes offspring that cannot be produced in a cross between these parents.
    (f) (3 points) If each gene contributes 1 unit of height, what is the probability that a cross between two parents with genotypes AaBbCcDdEe and aabbccddee results in a child with height 3 units.
    (g) (5 points) Now, again assuming that each gene contributes 1 unit of height, what is the probability that the two parents from the start of this problem have a child with the lowest possible height for these parents? What is this lowest possible height, assuming that the minimum possible height is 0 and each dominant gene adds 1 to the height?
    (h) (3 points) Imagine (just for this part) that the amount each gene contributes to the height is different by an order of magnitude. For example, this could mean that the gene with A as an allele contributes 10 times more than the gene with B as an allele, B 10 times more than C, and so on. In this case, how many different heights are possible?

    Now, let's say that the genes still contribute to the height of offspring, but now every dominant allele contributes some amount to the height. Yes, this means that it's now technically not dominant, but we'll still use the capital letter to distinguish it from the other allele. This means that, for example, aabbccddee will have the minimum height, and AABBCCDDEE would have the maximum height.
    (i) (2 points) How many different possible heights are there if each allele contributes the same amount? As before, this includes offspring that cannot be produced in a cross between these parents.
    (j) (4 points) Assume that the minimum height (for a plant with aabbccddee) is 0 units, and each dominant allele contributes 1 unit of height to the offspring's height. What is the minimum and maximum possible height for the offspring of the two parents from the start of this question?
    (k) (6 points) What is the probability of the child plant having a height of 1 unit? 3 units? **(TB)**

The table below show the number of nucleotide that are different between each pair of species in a certain sequence of DNA.

	A	**B**	**C**	**D**	**E**
A	0	11	15	15	15
B	11	0	15	16	15
C	15	15	0	8	13
D	15	16	8	0	14
E	15	15	13	14	0

61. (a) (1 point) Which two species are most closely related?
    (b) (5 points) Construct a phylogenetic tree for these species using the template on your answer sheet.
    (c) (3 points) Using the neutral theory of molecular evolution, how does the length of time since A and B diverged compare to the time since A and C diverged? Why?