Started on	Tuesday, 20 April 2021, 3:39 PM
State	Finished
Completed on	Tuesday, 20 April 2021, 4:36 PM
Time taken	56 mins 24 secs
Grade	**38.00** out of 50.00 (**76**%)

In PCR sequencing by Sanger's method, the sequence of nucleotides reaching the other end of the gel electrophoresis reads 5' TGACGATTA 3'. Give the sequence of the template strand in the correct orientation.

5' ACTGCTAAT 3'

Comment:

Give approximate values of window size and stringency (for matches) you would use for dot plot analysis to identify multiple copies of domains.

Window Size: 15

Stringency: 10

Comment:

In double-stranded DNA, which of the following base ratios always equals 1?

○ a. G/C

○ b. C/T

○ c. (A+T)/(G+C)

○ d. G/A

◉ e. (A+G)/(C+T) ✔

Your answer is correct.

The correct answer is:
(A+G)/(C+T)

If two sequences are homologous, then which of the following statements is true?

- a. The two sequences are identical

- b. The two sequences share full-length sequence similarity

- c. The two sequences are partially conserved

- d. The two sequences share a common ancestor

- ● e. The two sequences share a common ancestor and they either share full length similarity or are partially conserved. ✔

Your answer is correct.

The correct answer is: The two sequences share a common ancestor and they either share full length similarity or are partially conserved.

Construct a restriction map of a linear fragment of DNA, using the following data. Give the fragment lengths and their positions on a linear scale.

DNA	Sizes of Fragments (bp)
uncut DNA	10,000
DNA cut with EcoRI	8000, 2000
DNA cut with BamHI	5000, 5000
DNA cut with EcoRI + BamHI	5000, 3000, 2000

When cut with EcoRI:

location : length

1-8000 : 8000

8001-10000 : 2000

When cut with BamHI:

location : length

1-5000 : 5000

5001-10000 : 5000

When cut with Both:

location : length

1-5000 : 5000

5001-8000 : 3000

8001 - 10000 : 2000

Thus we can conclude that the restriction site for EcoRI is at 8000 and the restrictions site for BamHI is at 5000.

Comment:

Which of the two scoring matrices, PAM250 and PAM125 would you use for comparing (i) closely related sequences and (ii) evolutionarily distant sequences?

(i) PAM125

(ii) PAM250

Comment:

The E-value in the BLAST program depends on database size. True/False.

Select one:

○ True ✔

○ False

The correct answer is 'True'.

If a circular bacterial genome has N sites for a particular restriction enzyme, into how many fragments will that restriction enzyme cut the genome?

○ a. N-1

○ b. N+2

○ c. N+1

◉ d. N   ✔

Your answer is correct.

The correct answer is:
N

Computing the log odds score (match and mismatch) for comparing DNA sequences that exhibit 75% identity. Assume background frequency to be 0.25 and also assume equiprobability for mismatch.

log((0.75*0.25)/(0.25*0.25)) = log(3)

Comment:

Which variant of the dynamic programming algorithm would you use if the two sequences are closely related:

⦿ a. does not matter, as both will give same results ✔

○ b. global alignment

○ c. local alignment

Your answer is correct.

The correct answer is:
does not matter, as both will give same results

Protein sequence database search is better than DNA sequence search for finding remote homologies. What are the possible reasons? [Note: List all the correct options]

☑ a. Because of degeneracy in the code, a change at DNA level may not always lead to a change at the amino acid level, making it possible to find distant homologies using protein database search ✔

☑ b. Scoring matrices for proteins are more sensitive compared to those used for comparing DNA sequences ✔

☑ c. Only 4 characters in DNA compared to 20 in proteins, this leads to more hits by random chance at DNA level. ✔

☐ d. Computationally, protein sequence is one-third the size of the corresponding DNA sequence.

Your answer is correct.

The correct answers are:

Because of degeneracy in the code, a change at DNA level may not always lead to a change at the amino acid level, making it possible to find distant homologies using protein database search,

Only 4 characters in DNA compared to 20 in proteins, this leads to more hits by random chance at DNA level.,

Scoring matrices for proteins are more sensitive compared to those used for comparing DNA sequences

Base pairing between the two strands of DNA is formed by -----------

Answer: Hydrogen Bonds ✔

The correct answer is: H-bonds

Give the expected frequency of observing GC boxes, GCCGCC in a genome of length 10MB and G = 0.2?

Answer: 640 ✔

The correct answer is: 640

Question **14**

Partially correct

Mark 1.50 out of 2.00

The SARS-CoV-2 virus genome is being continuously sequenced across the world from infected people. What is the advantage of re-sequencing? [Note: List all the correct options]

- ☑ a. demographic distribution of various strains of the virus ✔
- ☑ b. to understand the migration pattern ✔
- ☑ c. to know if the virus is mutating ✔
- ☐ d. to know if the virus is becoming more virulent or loosing its virulence

Your answer is partially correct.

You have correctly selected 3.
The correct answers are:
to know if the virus is mutating,

to understand the migration pattern,

to know if the virus is becoming more virulent or loosing its virulence,

demographic distribution of various strains of the virus

Homology is implied if

○ a. both sequence and structure are similar

○ b. two sequences are > 50% identical

◉ c. two sequences are similar   ✖

○ d. two structures are similar

Your answer is incorrect.

The correct answer is:
both sequence and structure are similar

Which of the following is not a difference between DNA and RNA?

- a. DNA contains alternating sugar-phosphate molecules whereas RNA does not contain sugars ✔

- b. RNA is single stranded and DNA is double stranded

- c. DNA contains thymine whereas RNA contains uracil

- d. DNA contains deoxyribose and RNA contains ribose

Your answer is correct.

The correct answer is:
DNA contains alternating sugar-phosphate molecules whereas RNA does not contain sugars

Which of the following two alignments is likely to evolutionarily more plausible? Why?

(i)   ACAAT   (ii)   ACAAT
       A -A - T   A - - AT

The second option is more likely to occur in nature as it is more common in nature for a series of k indels to occur together rather than separately.

Comment:

Compute the score of the following pairwise alignments:

Global alignment: Match = +1, Mismatch = 0, Gap: opening = -5, extension = -1

GAGG - - - TGAC - TGGGCA

GTGGATGCGGCAAGGG - -

1+0+1+1-5-1-1+0+1+0+1-5+0+1+1+1-5-1

= -10

Comment:

Question **19**

Correct

Mark 1.00 out of 1.00

Increasing the Gap Costs will result in shorter local alignments?

Select one:

- ⦿ True ✔
- ○ False

The correct answer is 'True'.

BLAST program misses some good biological homologies below the accepted statistical cut off value. How would you identify these distant homologies?

I would use PSI-BLAST to identify these distant homologies. It is useful because it is specifically made to find sequences that are similar and have 15-20% similarity.

Comment:

FASTA uses exact word match, while BLAST uses a scoring scheme to look for similar words as seeds in the hashing step of the algorithm. So which is more sensitive program for database search:

- ○ a. FASTA   ✖
- ○ b. BLAST

Your answer is incorrect.

The correct answer is:
BLAST

The two strands of DNA molecule are:

○ a. run antiparallel, i.e., are reverse of each other

○ b. mirror images of each other

○ c. reverse complement of each other

◉ d. complement each other   ✗

Your answer is incorrect.

The correct answer is:

reverse complement of each other

An organism's phenotype is influenced by its:

- ○ a. genotype

- ○ b. environmental factors

- ◉ c. both genotype and environmental factors ✔

Your answer is correct.

The correct answer is: both genotype and environmental factors

If the length of sequence read is 200bases, and 10x coverage is desired, give the number of reads that will be required to sequence a human genome.

Answer: 150,000,000 ✗

The correct answer is: No. of Reads required = GenomeLength DesiredCoverage / ReadLength = 3 x 109 x 10/200 = 1.5 x 108 = 150Million reads

Comment:

Polymerase Chain Reaction can be used for

○ a. amplification of DNA sequences

◉ b. both amplification and sequencing of DNA sequences ✔

○ c. sequencing of DNA sequences

Your answer is correct.

The correct answer is:
both amplification and sequencing of DNA sequences

Cloning can be used for ___ of an unknown DNA sequence

○ a. storing

○ b. expression

○ c. storing and amplification

○ d. amplification

◉ e. storing, amplification and expression ✔

Your answer is correct.

The correct answer is: storing, amplification and expression

Which approach would you use for amplification of an unknown DNA sequence: (i) cloning, or (ii) PCR? Give reasons.

Answer: I would use PCR because PCR is faster, doesn't require an RE, requires a smaller sa ✖

The correct answer is: Cloning. For an unknown sequence we do not know the primers.

The recognition site for Sau3A I is GATC and is contained in the recognition site of BamH I, GGATCC. Will the two REs give the same results? If not, which one will give larger number of fragments?

No. Sau3A will give a higher number of fragments since wherever BamHI cuts, Sau3A will also cut, but Sau3A can cut where BamHI cannot.

Comment:

Sequences in DNA that restriction enzymes bind to and cut are mostly:

○ a. Antiparallel

○ b. Not symmetrical about the midpoint

○ c. Random sequences

◉ d. Symmetrical about the midpoint ✔

Your answer is correct.

The correct answer is:

Symmetrical about the midpoint

Homologous sequences are not always highly similar. **True / False**

Select one:

◉ True ✔

○ False

The correct answer is 'True'.

Which one of the following statements is true:

○ a. The end products of genes are proteins

○ b. Chromosomes consist of RNA and proteins

○ c. The end products of genes are proteins and RNA

◉ d. A gene codes for only one protein product   ✗

○ e. One triplet code codes only for one amino acid

Your answer is incorrect.

The correct answer is:

The end products of genes are proteins and RNA

A single scoring matrix can be used for aligning sequences at different evolutionary distances.

Select one:

○ True

◉ False ✔

The correct answer is 'False'.

If ATATATAT is the sequence in the forward strand, give the sequence of the other strand in the correct orientation.

Answer: TATATATA ✖

The correct answer is: ATATATAT

Genes are stable entities and inherited from generation to generation. Are mutated genes also stable entities and inherited in the same way as normal genes? **[Yes/No]**

Answer: Yes ✔

The correct answer is: Yes

The bond linking adjacent nucleotide building blocks in DNA is called:

○ a. An ether bond

○ b. A polynuclieotide bond

○ c. A phosphate bond

◉ d. A phosphodiester bond ✔

Your answer is correct.

The correct answer is:
A phosphodiester bond

Jump to...