

## SOLUTION KEY- SECTION 10

1. How is a cDNA library different from a genomic library?

A **genomic library** is a population of host bacteria, each of which carries a DNA fragment that was inserted into a cloning vector, such that the collection of cloned DNA fragment represents the entire genome of the source organism. The DNA fragment of an organism of interest are isolated, partially digested by an endonuclease restriction enzyme and ligated into vectors. The vectors can then be taken up by suitable hosts. The hosts are kept in liquid media and can be frozen at  $-80^{\circ}\text{C}$  for a long period of time. Usually the hosts are bacteria that do not contain any plasmids, so as to be sensitive to antibiotics. A **cDNA library** is a collection of cloned complementary DNA (cDNA) fragments inserted into a collection of host cells. A cDNA is produced from fully transcribed mRNA found in the nucleus and therefore contains only the expressed genes of an organism. In eukaryotic cells the mature mRNA is already spliced, hence the cDNA produced lacks introns and can be readily expressed in a bacterial cell. While information in cDNA libraries is a powerful and useful tool since gene products are easily identified, the libraries lack information about introns, and other regulatory elements found in a genomic DNA library.

2. You are interested in making many copies of a specific DNA sequence. The sequence to be amplified is flanked by regions whose sequence is given below. Which primer set would you use to amplify the DNA sequence.

Primer 1 should bind in this region

5' CGCGCGAATTCGATCGA  
3' GCGCGCTTAAGCTAGCT

Primer 2 should bind in this region

TTAATACGTACTAG 3'  
AATTATGCATGATC 5'

Primer 1

Primer 2

Set A: 5' TCGATCGAATTC 3'

AND

5' TAATACGTACTA 3'

Set B: 5' GCTTAAGCTAGC 3'

AND

5' GATCATGCATAA 3'

Set C: 5' GAATTCGATCGA 3'

AND

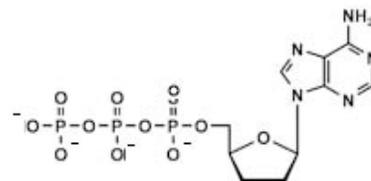
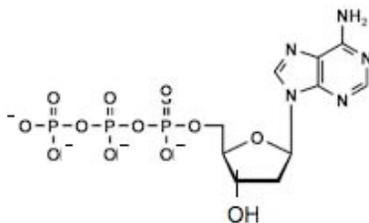
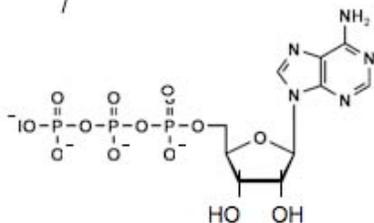
5' CTAGTACGTATT 3'

3. Take a look at the molecules drawn below

#1

#2

#3



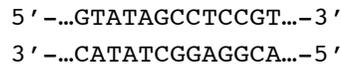
a) Which of the above molecules is non-physiological and is only used in DNA sequencing? #3 since it only has H atoms at the 2'C and 3'C position of the sugar (i.e. it has dideoxyribose sugar).

b) Which one of the above molecules is used in RNA? #1 since it has OH groups at the 2'C and 3'C position of the sugar (i.e. it has ribose sugar).

c) Which one of the above molecules is used in DNA? #2 since it has OH groups at the 3'C and H atom at 2'C position of the sugar (i.e. it has deoxyribose sugar).

d) Which one of the above molecules is used as the major source of energy in cells? #1 since it is ATP, which serves as the energy currency of the cell.

4. You decide to sequence the following piece of DNA using the following primer:



The primer you use is 5'-ACG-3' (Note: real primers are usually 18-22 nts long). You set up a DNA sequencing reaction with this DNA, this primer, DNA polymerase, dATP, dTTP, dGTP, dCTP, and ddATP.

a) How many different products would form from this reaction?

Since you are using ddATP in the reaction mixture halts whenever there is an incorporation of a ddATP instead of dATP. Since there are 3A's in the sequence (excluding the A that is a part of the primer) and therefore you may have 3 reaction products of different lengths. There is always a probability that you may have a reaction product where on dATP is incorporated and NO ddATP is incorporated at all. Combining these two together you may have 4 reaction products (5b, 10b, 12b and 13bases long)

b) How many nucleotides long would each different product of this reaction be?

They would be 5b, 10b, 12b and 13bases long).

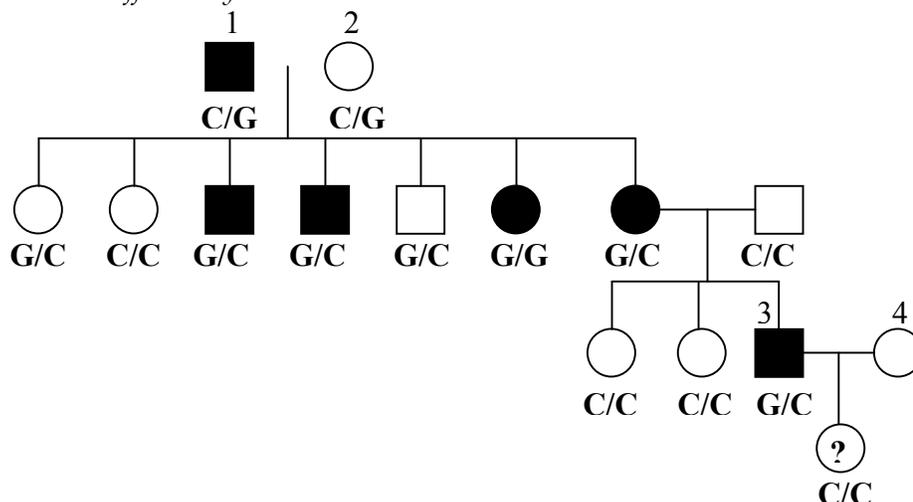
c) Write the sequence of DNA that you read from this gel below. Label the 5' and 3' ends of the molecule.



d) Would you be able to tell the sequence of the molecule if you had loaded into a single lane a reaction in which all four ddNTPs had been added from a lane?

Yes, but only if each ddNTP was labeled with a fluorophore of different color.

5. You identify a hypothetical Gene R in humans that encodes for protein "R" which is involved in maintaining low blood cholesterol level. This gene shows an autosomal dominant mode of inheritance and the affected individuals are at a higher risk of developing a cardiac disorder (CD). You come across a SNP (Single nucleotide polymorphism) that is tightly linked to Gene R. You decide to use this SNP as a marker for CD. The two alleles of SNP (C and G) are shown for each individual in the following pedigree. Individuals affected by CD are shaded in black.



a) Assuming no recombination between this SNP and Gene R, which allele of the SNP is linked to the disease associated **R** allele in individual #1?

*It is the "G" allele.*

b) Assuming no recombination between this SNP and Gene R, what is the probability that the offspring of individual #3 and individual #4 has the disease?

*This offspring has the C/C alleles for this SNP. Therefore he is obtaining the good allele "C" from #3 and another allele, also "C" from #4. Therefore this offspring will be normal and will not have the disease i.e. the probability is 0.*

---

---

MIT OpenCourseWare  
<http://ocw.mit.edu>

7.013 Introductory Biology  
Spring 2013

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.