

BIOL3104 **Midterm 2**, 2011, 6 pages, Total marks possible = 57

1) Define each of the following (/1 each)

a) transcriptome

the sum total of all the messenger RNA molecules expressed from the genes of an organism.

b) ortholog

Same gene function found in various organism

c) synteny

/5 When 2 species have the same gene order within a chromosome.

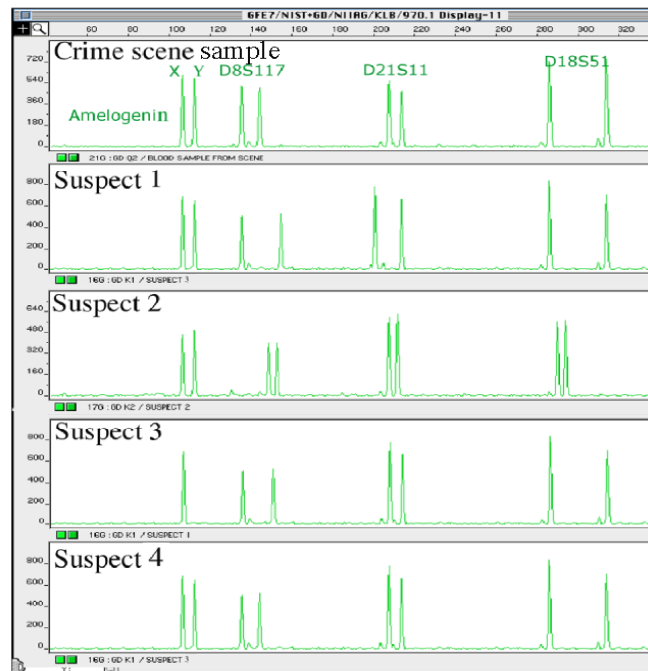
d) selective sweep

When a trait sweeps over a population over time and becomes fixed.

e) STS

Sequence tagged site

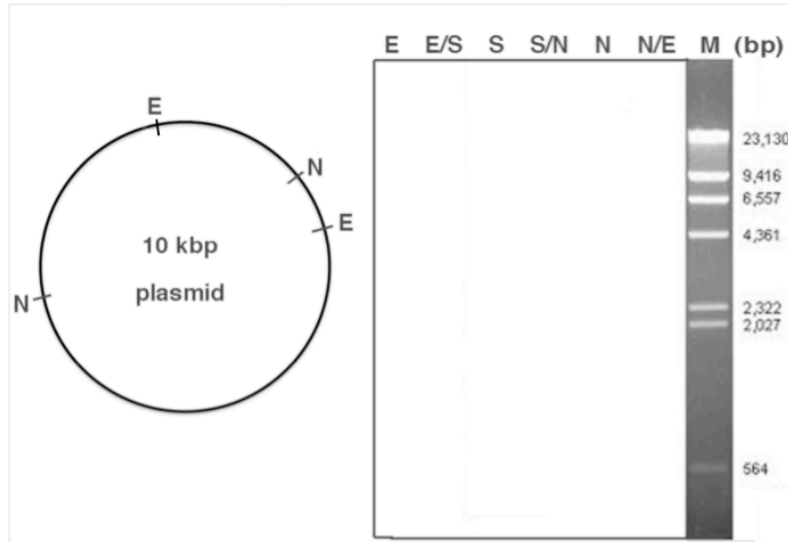
/4 2) a) Based on the figure at right, briefly explain whether each suspect is guilty or innocent (answer carefully). (/3)



b) The markers D8S1179, D21S11, and D18S51 refer to (circle one or more) (/1)

- a) AFLPs b) PCR amplicons c) minisatellites d) RAPDs e) STRs

3) You digest DNA of the plasmid shown below left with the restriction enzymes *EcoRI* (E), *EcoRI* + *SalI* (E/S), *SalI* (S), *SalI* + *NotI* (S/N), *NotI* (N), *NotI* + *EcoRI* (N/E). You then load the digested DNA into separate wells of a 0.7% agarose gel with TAE buffer and use electrophoresis to separate the fragments. Lambda DNA digested with *HindIII* serves as a size standard in your gel (far right, size in bp for each fragment is given). Draw the fragments you expect to see for each digest in the box below.



4) A cross is performed between two true breeding strains of a diploid organism to get F1 progeny. The phenotypic ratios at right result from a test cross with the F1 progeny. Draw the genetic map and include map distances.

A	+	+	312
+	B	C	313
A	B	C	94
+	+	+	89
+	B	+	9
A	+	C	6
+	+	C	92
A	B	+	<u>85</u>
			1000

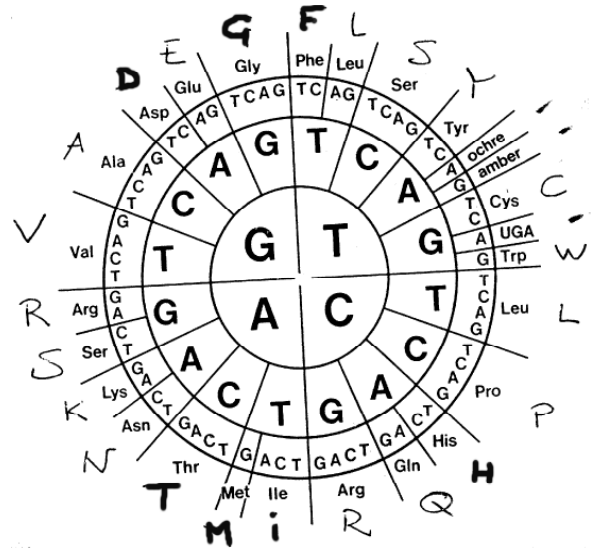
5) WARNING: SAVE THIS QUESTION FOR LAST – IT IS COMPUTATIONALLY EXPENSIVE. THE FRAGMENTS OVERLAP BY AT LEAST 6 NUCLEOTIDES

Shotgun sequencing was used to obtain the 5 sequences shown below.

/8

- 1) TGAAAAAGCCTGAA-3'
- 2) CACCGCGACGTCTGT-3'
- 3) TCGCGGTGAGTTCAGGCTTTT-3'
- 4) AACTATCTAGTTAG-3'
- 5) ACGTCTGTCGAGAACTATCT-3'

a) In the space below, arrange the sequences into a single contig. (/5)



b) Give the amino acid sequence of the longest predicted ORF. (/2)

c) Briefly state what is meant by the term “shotgun sequencing”? (/1)

6) Draw the key stages of “bridge amplification” as used in the SOLEXA sequencing platform.

/4

/6 7) The terms a – f (below) are ways that Green et al. tried to ensure that sequences analyzed were from Neandertal. Briefly provide a reason why each was used. (/1 each)

a) clean room

To ensure that nothing would contaminate the sample during the preparation and extraction.

b) “tagged” DNA fragments

A clean room tag was added so that when the dna was sequenced, only DNA from the clean room would be used

c) restriction enzyme digests

A CpG digest is used to cut out microbial that are found in the bone

d) comparisons of sequences obtained to genome sequences and GenBank entries.

e) mtDNA analysis

Allowed to calculate the error rate by doing a depth analysis.

f) Y chromosome analysis

because the three individuals were found to be females, any Y-specific DNA would be contamination.

/2 8) Are spontaneous transition mutations more common than transversions? (/1) Provide a brief explanation of why or why not. (/1)

/1 9) An enzyme found in association with retroviral activity and involved in the synthesis of DNA from an RNA template is: A. Reverse polymerase I B. RNA polymerase
C. DNA ligase D. Reverse transcriptase E. DNA polymerase I

10) Introns are

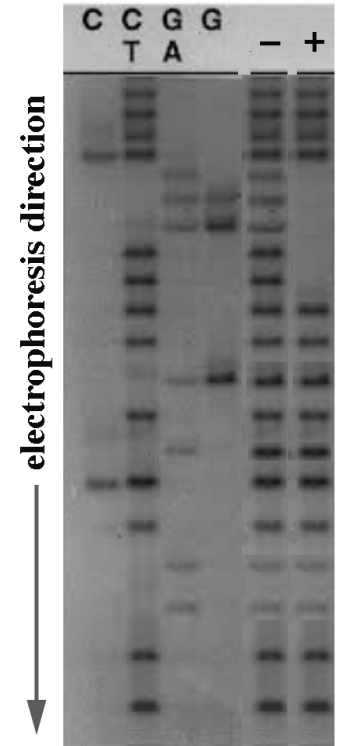
/1 A. noncoding regions spliced out of transcripts. B. external regions of a gene
C. coding regions joined together during splicing. D. 5' and 3' UTRs.
E. multiple genes transcribed together.

11) The sequencing method used at right involves 5' end-labelling the template DNA and chemically inducing breaks at C, C&T, G&A and G (left 4 lanes). The method can also be used for "DNA footprinting" as shown in the 2 right lanes.

/4

- a. The method used is called _____ (1)
 b. Provide the DNA sequence (5' > 3') based on the results shown. (2)

c. Indicate along your sequence where the protein binds based on the footprint results. (1)



12) You wish to design PCR primers with a Tm of 55 – 58 °C to amplify a DNA molecule with the following sequence

/5

AGAGTGCTCGATCGATCATCATCTTTGGACGCAGCGGGCGATCATACTACGGCTACGAC-3'

a) Provide forward and reverse primer sequences for this purpose (1), indicate the 5' and 3' ends of your primers (1), and give approximate Tm for each primer using the formula discussed in class. (1)

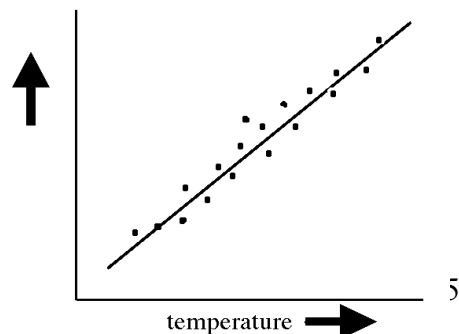
b) What does "Tm" stand for and what is the formal definition of Tm? (2)

13) Thep rob lemwit hfram esh iftsi stha the yar eh ar dtore ad. TRUE or FALSE? (bonus mark)

/1

- 14) In the figure at right relating to DNA denaturation, the Y axis would be
- a. % purines
 - b. % pyrimidines
 - c. % A + T
 - d. % G + C
 - e. DNA concentration

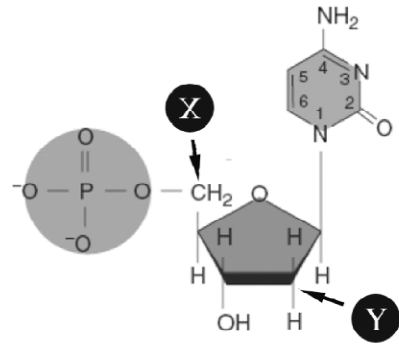
/1



5

/1 **15)** The diagram at right depicts a (i) found in (ii)

- a. (i) purine nucleoside, (ii) DNA.
- b. (i) purine nucleotide, (ii) RNA.
- c. (i) pyrimidine nucleoside, (ii) DNA.
- d. (i) pyrimidine nucleotide, (ii) RNA.
- e. none of the above.



/1 **16)** X and Y in the diagram at right represent the (i) and (ii) carbons, respectively, of the (iii).

- a. (i) 3', (ii) 5', (iii) sugar b. (i) 5', (ii) 3', (iii) sugar
- c. (i) 3', (ii) 5', (iii) base d. (i) 5', (ii) 3', (iii) base
- e. none of above

/3 **17)** What technique of gene finding can be used to identify very short exons (microexons)? (/1)
Briefly describe the technique. (/2)

/2 **18)** A researcher sequences a region of DNA that encodes a 100 amino acid protein. When a BLAST analysis is done, the DNA sequence matches to a GenBank entry with 79 % identity, whereas a BLAST search with the protein sequence matches to a GenBank entry at 90% identity. Briefly, explain why these identity values differ.

/2 **19)** What “counting experiment” was done in the neandertal genome study by Green et al. and what was the purpose of this counting experiment?