BCH3356 *Molecular Biology Laboratory*

Final exam: This exam is worth 40 % of the final mark for BCH3356

Name:_____
Student number:_____

Instructions

All questions should be answered within the space provided on THIS COPY EXAM.

You may use the endorsement of pages for your draft calculations.

Students are allowed to use a calculator.

Part I (/20 marks): *Basic Molecular Biology Techniques*

1. **RT-PCR** You have to amplify the full length cDNA sequence coding for Taq polymerase (see appendix II for sequence detail). You should assume that the 5' UTR and 3' UTR are not necessary for the intended experiment.
    a. Assuming that you have access to the appropriate mRNA template, describe a pair of primers that could be used to generate the full length cDNA sequence. In you answer, indicate the sequence for each of the two cDNA strands at and around the binding positions of the two primers. Also include the sequence for each of the two primers and their annealing strand.
    Primers in blue

    b. Which of the two primers should be used for the RT step?
    The negative primer depicted at the top right above is to be used for RT

    c. How the cDNA or RT-PCR products would be affected if the two primers were added for the RT step?
    Addition of both + and – primers for the RT would not affect the overall RT-PCR product at all as the + primer would not affect RT reaction.

2. **PROBE SYNTHESIS FOR NORTHERN AND SOUTHERN ANALYSES** Describe the expected result for each of the following conditions used to synthesize a digoxygenin-labelled probe complementary to the fragment 961-1821 of the Taq polymerase cDNA sequence (see sequence detail in appendix).
    a. All optimal conditions with primer (1) caccctggccaagaaggcgg and primer (2) tctcctcccgggccacccgc.

    + primer is OK
    - Primer is in the wrong orientation and cannot anneal in parallel with the template stretch of DNA. Will be no amplification and no probe synthesis.

    b. All optimal conditions but only one DNA polymerase: Taq.
    Taq has a high processivity, but cannot add the digoxygenin-U's. The probe will be synthesized, but will not be labeled with digoxygenin.

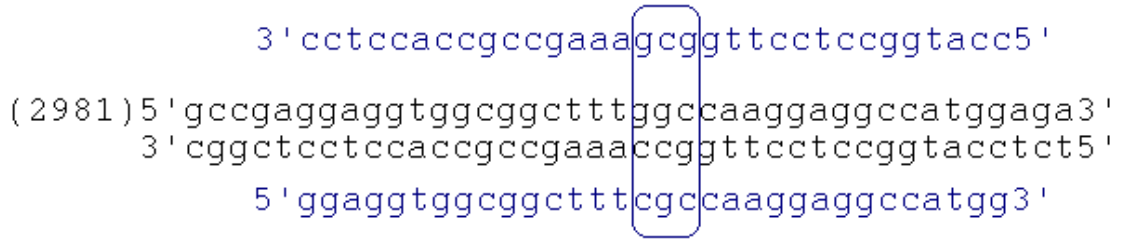
    c. All optimal conditions but only one DNA polymerase: Pwo.
    DNA polymerase can incorporate digoxygenin-U's, but has a low processivity (it takes off from the template). A mixture of digoxygenin-labeled fragments with different lengths will be amplified. Not a good amplification yield is expected given the low processivity of Pwo polymerase.

    d. How many digoxygenin molecules would you expect per molecule of cDNA probe?

This is an open question and different answers are OK if properly justified. Assuming 25% of T's in the 961-1821 fragment, the maximum number of dig-labeled U's could be estimated as 0.25x(1821-961) or 215. Obviously the actual number of dig-U's per molecule of probe has to be lower than 215.

3. **MUTAGENESIS** It has been suggested that the codon at positions 3001-3003 is essential for the thermal stability of the encoded enzyme. You want to test that hypothesis by substituting the amino acid encoded by nucleotides GGC at positions 3001-3003 with an arginine residue that is encoded by the CGC codon.

   a. **2 MARKS** Indicate the sequence of each mutagenic primers you would use.



   b. **2 MARKS** What enzyme or type of enzyme would you recommend for the mutagenesis PCR reaction?

A non displacing or non dislodging DNA polymerase is crucial for site-directed mutagenesis.

   c. **2 MARKS** Mutagenesis PCR amplification is considered as linear. What does linear amplification mean? How is it that the mutagenesis reaction can not yield an exponential PCR amplification as it is with conventional reactions?

Linear amplification means that the number of amplified strands is linearly related with the number of amplification cycles. For instance, if 10 initial templates are amplified 20 times, the number of mutated product will be 10 times 20 or 200.

The amplified, mutated product cannot be used as a suitable template for priming further amplification. This is because of the nicks that remain after amplification between the 5' end of each primer and the very last nucleotide added during PCR amplification.

4. **3 MARKS** For the research project, you used the pET30a vector containing the DNA sequence coding for minE and a purification column.
    a. Explain the underlying molecular principle of the affinity column chromatography you used to selectively purify the protein encoded by the pET30a vector. Use a diagram.

    b. How did the fusion protein purified with the affinity column chromatography described in your answer to question A differ from the protein encoded by the DNA fragment that was ligated in the pET30a vector?

    c. Describe the underlying molecular principle for two different elution buffers that could be used to elute the fusion protein encoded by the pET30a vector from the affinity column chromatography described in your answer to question a.

5. **3 MARKS** pBluescript
    a. Describe the genetic feature of pBluescript for specific screening of transformant cells.
    Transformed cells are screened based on the pBluescript gene coding for resistance to ampicillin.

    b. Describe the genetic feature of pBluescript for specific screening of colonies containing a recombinant vector with an insert.
    The a complementation screening gives blue and white colonies for empty and recombinant plasmids, respectively.
    c. Describe the genetic feature of pBluescript for sequencing DNA inserts ligated within the MCS.
    Presence of a unique DNA motif upward and downward the multiple cloning site (T7 and T3 promoters) allows the use of universal primers (T7/T3 primer binding sites) to initiate sequencing from either end of the insert.

    d. Describe the genetic feature of pBluescript for regulating the transcription of DNA inserts.
    T7 and T3 promoters control transcription from either end of the DNA insert.

6. **5 MARKS** Define the following terms or expressions:
    a. Copy number: Number of plasmid copies per transformed cell

    b. Transformation yield: Number of transformant colonies per ug of plasmid DNA used for a transformation

c. T7 primer binding site: DNA motif that can be located at either end of a plasmid that is recognized by a primer usually used to initiate DNA sequencing.

d. Microcarrier gel: Polysaccharide gel used to improve precipitation yield of RNA or DNA during centrifugation, especially when found at low concentrations.

e. Semi-dry transfer: Electrophoretic transfer in which the transfer sandwich is kept wet and placed between two plate electrodes (the transfer sandwich is not directly soaked in a transfer buffer).

f. Monoclonal antibody: Antibody preparation prepared from a single clone of B cells and that recognizes one single antigen or epitope.

g. Epitope: Part of an antigen that is recognized by an antibody.

h. Chemoluminescence: Emission of light as the result of a chemical reaction

i. Displacing DNA polymerase: DNA polymerase that can pull apart or melt two complementary strands so elongation can be pursued at the 3' end of the elongating fragment.

j. Positive primer: Primer that anneals onto the negative or antisense strand to initiate the elongation of the + or sense strand.

k. Sense strand: DNA strand that corresponds to is NOT transcribed into mRNA; it is identical to the mRNA sequence (except that DNA contains T where RNA has U).

l. 3' UTR: Part of the messenger RNA that is located downward the stop codon and is NOT translated.

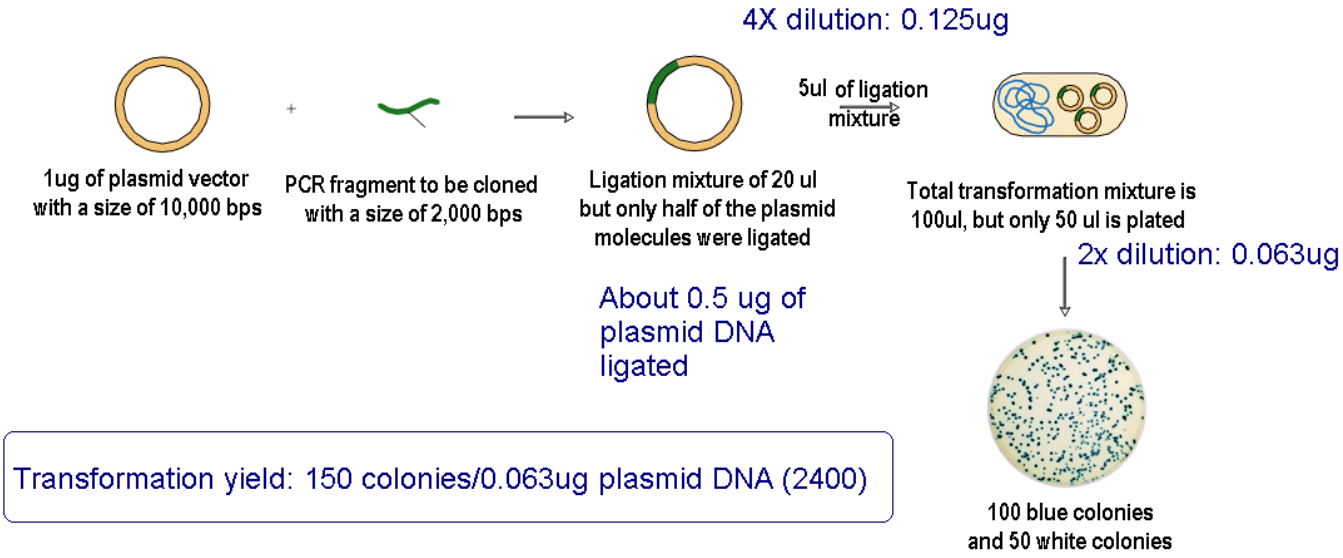
7. Presume that (1) the full length coding cDNA sequence for DNA polymerase (Appendix 2) was ligated at the EcoRV site of the Pinpoint Xa plasmid (Appendix I) using the T/A cloning procedure and (2) transformation of competent cells was further completed with that ligation product. Your challenge is now to screen for the orientation of the cDNA insert.
WE DID NOT PERFORM ORIENTATION SCREENING AS A DIRECTIONAL CLONING STRATEGY HAD BEEN USED DURING THE SEMESTER. YOU ARE NOT EXPECTED TO BE FAMILIAR WITH T/A CLONING, THE STRATEGY USED LAST YEAR. IGNORE THIS QUESTION.

a. ~~Describe one restriction enzyme that could be used to discriminate between inserts in the T7 orientation versus those in the opposite orientation. You should refer to appendix 3 to access the list of restriction sites within the cDNA insert.~~

b. ~~Bacterial clones can also be analyzed directly by PCR to determine insert orientation. Describe one vector-specific and one insert-specific primers that could be used to determine insert orientation. Also predict the expected amplicon size for the DNA polymerase cDNA insert in (A) the T7 orientation versus (B) the opposite orientation.~~

Part II (/30 marks): *Analytical Skills*

1. **4 MARKS** The figure below displays an overview for a ligation and transformation procedure.



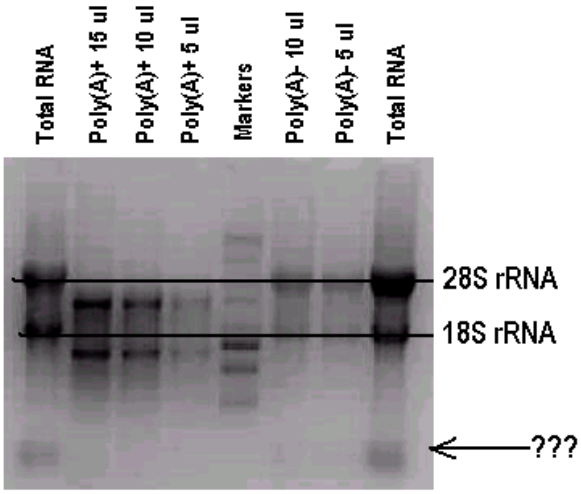
a. **3 MARKS** What's the transformation yield for that experiment? Make sure to use the appropriate units so that you can compare your experimental value with the expected yield. Show your calculations.

See above. 2400 colony forming units per ug of plasmid DNA

b. **1 MARK** Is the experimental transformation yield acceptable? In other words, is it too low or too high relative to the expected transformation yield assuming that competent cells had been treated with calcium ions and heat-shocked?

For heat shock based-transformation, yield should be about 10^{+6} cfu/ug of plasmid DNA. This is roughly 500 times higher that the experimental result.

2. **6 MARKS** RNA gel is displayed below for results obtained by one group of students during the semester. The rat liver RNA fractions were obtained by using the oligo dT-column.

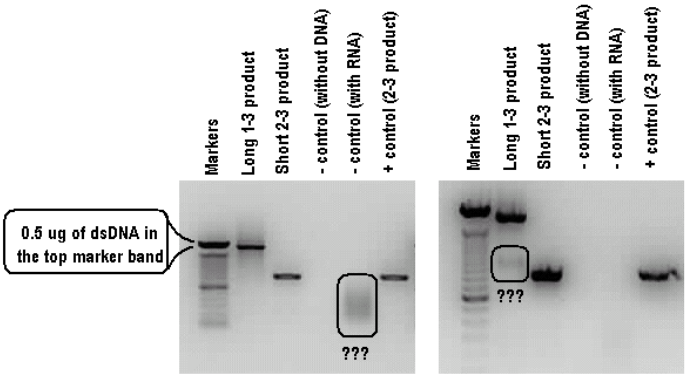


a. Could you provide an explanation why the two rRNA bands visible in the poly(A)+ lanes are located at a lower position relative to the total and poly(A)- RNA samples?
That was a problem that had been noticed last year. I suspect there is a conformational effect that could relate to contamination of poly(A)+ fraction with some of the binding buffer that has a very high Li+ concentration.

b. Two faint bands can be seen at the bottom of the gel on both total RNA fractions (see arrow with ???). What does that band represent?
That could be the 4/5S rRNA, tRNA or degradation product.

c. Apart from the position of the rRNA bands, what's different between poly(A)+ and total RNA bands?
Compared to total RNA, poly(A)+ has a relatively more intense smear in the 1000-2000 bases range, which likely corresponds to a collection of different mRNA's. That suggests that the poly(A)+ RNA fraction was enriched in its relative mRNA content compared to the total RNA fraction.

3. **MARKS** The figure below displays the RT-PCR results for two groups as performed in lab 2 during the semester.



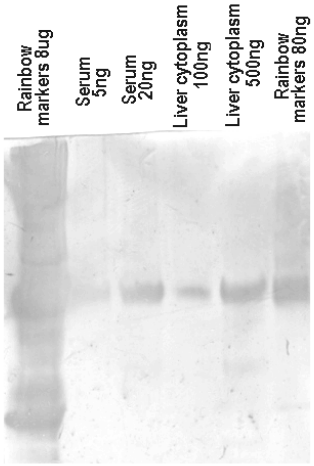
a. What would you expect to be located inside the box shown on the gel displayed at the left? Possibly some of the RNA template that had been used for that negative PCR control.

b. The gel displayed at the right shows one of the very few groups who amplified non-specific bands, i.e., extra bands that do not have the expected length. Explain two experimental conditions that could lead to non specific PCR amplification.
Temperature for the annealing PCR step might be low; primers can partially hybridize at different sites than the targeted annealing position; cross-contamination of the DNA template to be used for amplification with another source of DNA; …

c. Estimate the number of dsDNA molecules that were amplified in the PCR amplification for the short 2-3 fragment on the gel picture provided at the left. Assume that (1) the band intensity of the 2-3 fragment is about the same as the one for the top molecular weight standard and (2) 5ul out of 20ul of the PCR reaction were loaded on the gel.
Assuming that the band is about 1000 bps long (that was indicated to students during the exam last year):

0.5ug of DNA x(1umol bps/660ug DNA) x (1umol of DNA molecules/1000umol bps) = $0.8x10^{-6}$umol of DNA molecules or $4.6x10^{+11}$ molecules of DNA

4. **MARKS** Western analysis result obtained by one group is displayed below.



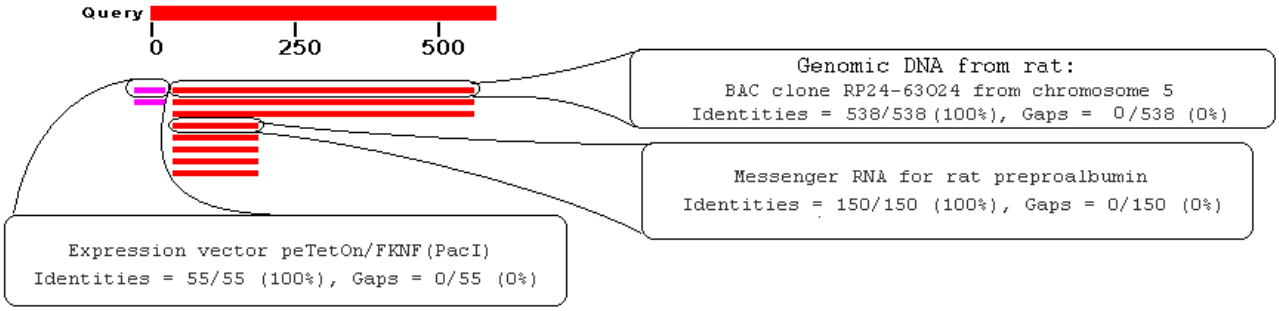
Relative intensity:	1	5	3	10
Amount (ng):	5	20	100	500
Int/amount:	0.2	0.25	0.03	0.02
Average:	0.22		0.025	

a. Refer to the results displayed on the figure just above to discuss the relative abundance of rat albumin between the two fractions that were investigated, i.e., serum and liver cytoplasm. Show your calculations and justify your reasoning.
Rat serum albumin is about 10-times more abundant in the serum compared to liver cytoplasm

b. What was the purpose to include two different amounts of the Rainbow markers for the Western analysis?
That's a question that you can ignore since different protein markers had been used last year.

c. Were preproalbumin and proalbumine simultaneously detected along with albumin by the Western analysis procedure that was done during the semester?
Preproalbumin differs from proalbumin and albumin only by a few extra amino acids. In other words, all three peptides share a fairly long common stretch and all three will be co-simultaneously recognized by the preparation of polyclonal antibodies.

4. One set of sequencing result obtained during the semester were submitted to nucleotide BLAST and the NCBI website returned the information displayed below.



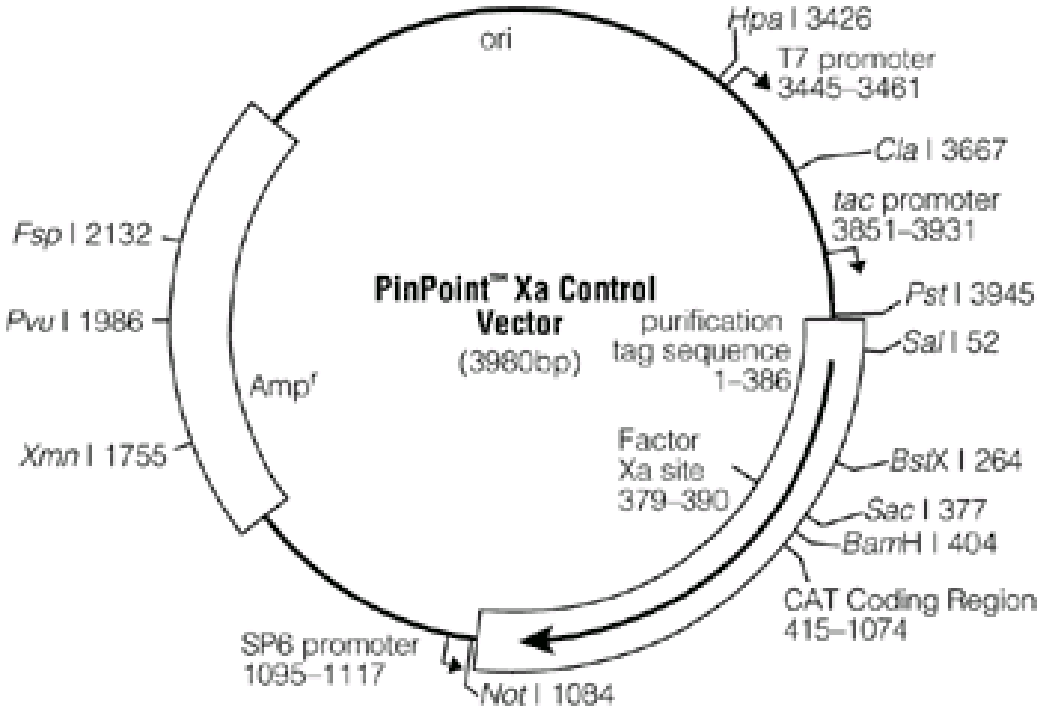
a. **1 MARK** How is it that a short fragment of the sequencing result matches an expression vector?
That short alignment corresponds to the DNA sequence located between the T7 (or T3) promoter that is targeted by T7 primer and the insertion point within the multiple cloning site.

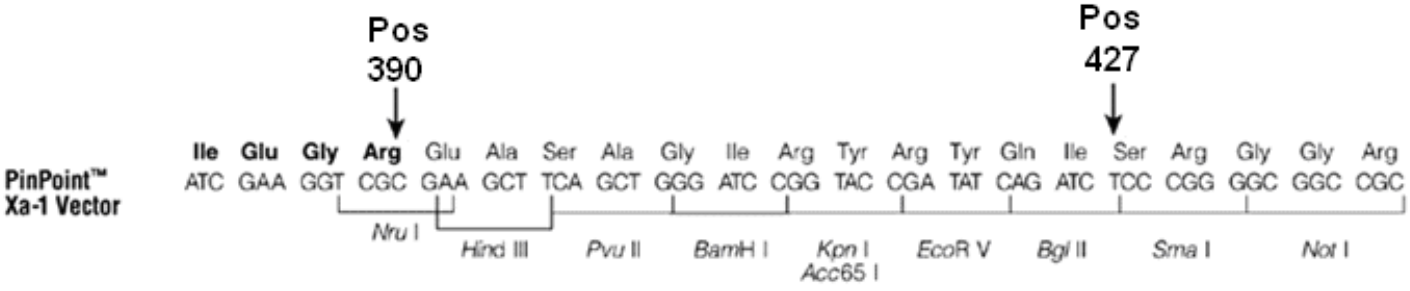
b. **1 MARK** Describe the nature of the DNA insert based on the information provided at the top.
A fragment of genomic DNA coding for rat serum albumin and containing part of an exon at the 5'end and part of an intro at the 3' end.

c. **1 MARK** One sequencing reaction can read up to 1200 nucleotides. Discuss the technical complications related with sample preparation for DNA sequencing that might explain why the result displayed above only contains about 550 nucleotides.
PCR reaction didn't work well and that could be due to pipetting or dilution errors in the preparation of the DNA template or any technical problem with PCR such as an improper ratio of regular nucleotides versus terminator dyes, improper temperatures of amplification, …

Appendix I: Informative Relative to the Pinpoint Xa Cloning Vector

The PinPoint™ Xa Protein Purification System is designed for the production and purification of fusion proteins that are biotinylated in vivo. Biotinylated fusion proteins are produced in *E. coli* and are affinity-purified using the SoftLink™ Soft Release Avidin Resin. This proprietary resin allows elution of the fusion protein under nondenaturing conditions. The PinPoint™ Vectors feature the encoded endoproteinase Factor Xa (pronounced “ten a”) proteolytic site that provides a way to separate the purification tag from the native protein. The PinPoint™ Xa Control Vector contains the chloramphenicol acetyltransferase (CAT) gene and is provided as a means of monitoring protein expression, purification and processing conditions.





Appendix 2: Thermotolerant DNA polymerase from *Thermus thermophilus*

```
     FT   5'UTR           1..589
     FT   CDS             590..3094
     FT   3'UTR           3095..3221
   1 tctagaggaa gcatgagcct caccctggca gacaaggtgg tctacgagga ggagatccag
  61 aaaagccgct tcatcgccaa ggcggccccc gtggcctcgg aggaggaggc cttggcgttt
 121 ttggccgaga accgggagcc tgaggccacc cacaacggcc acgcctacaa gatcggcctc
 181 ctctaccgct tctctgacga cggggagccc tcgggcaccg caggcaggcc catcctccac
 241 gccatagagg cccagggcct ggaccgggtg gcggtcctgg tggtgcgcta cttcggcggg
 301 gtgaagctcg gggccggggg gcttgtgcgg gcctacgggg gggtggcggc ggaggcctta
 361 aggcgggcgc ccaaggtccc cttggtggag cgggtggggc tcgccttcct cgtgcccttc
 421 gccgaggtgg gccgggtcta cgccctcctg gaggcccgcg cccctgaagg ccgaggagac
 481 ctacaccccg gaggcgtgcg cttcgccctc ctcctcccca agcccgagcg ggaaggtttc
 541 ctcagggcgc tcctggacgc cacccgggga caggtggccc tggagtagca tggaggcgat
 601 gcttccgctc tttgaaccca aaggccgggt cctcctggtg gacggccacc acctggccta
 661 ccgcaccttc ttcgccctga agggcctcac cacgagccgg ggcgaaccgg tgcaggcggt
 721 ctacggcttc gccaagagcc tcctcaaggc cctgaaggag gacgggtaca aggccgtctt
 781 cgtggtcttt gacgccaagg ccccctcctt ccgccacgag gcctacgagg cctacaaggc
 841 ggggagggcc ccgacccccg aggacttccc ccggcagctc gccctcatca aggagctggt
 901 ggacctcctg gggtttaccc gcctcgaggt ccccggctac gaggcggacg acgttctcgc
 961 caccctggcc aagaaggcgg aaaaggaggg gtacgaggtg cgcatcctca ccgccgaccg
1021 cgacctctac caactcgtct ccgaccgcgt cgccgtcctc caccccgagg gccacctcat
1081 caccccggag tggctttggg agaagtacgg cctcaggccg gagcagtggg tggacttccg
1141 cgccctcgtg ggggacccct ccgacaacct ccccggggtc aagggcatcg gggagaagac
1201 cgccctcaag ctcctcaagg agtggggaag cctggaaaac ctcctcaaga acctggaccg
1261 ggtaaagcca gaaaacgtcc gggagaagat caaggcccac ctggaagacc tcaggctctc
1321 cttggagctc tcccgggtgc gcaccgacct ccccctggag gtggacctcg cccaggggcg
1381 ggagcccgac cgggaggggc ttagggcctt cctggagagg ctggagttcg gcagcctcct
1441 ccacgagttc ggcctcctgg aggcccccgc ccccctggag gaggcccccc ggcccccgcc
1501 ggaaggggcc ttcgtgggct tcgtcctctc ccgccccgag cccatgtggg cggagcttaa
1561 agccctggcc gcctgcaggg acggccgggt gcaccgggca gcagacccct tggcggggct
1621 aaaggacctc aaggaggtcc ggggcctcct cgccaaggac ctcgccgtct tggcctcgag
1681 ggaggggcta gacctcgtgc ccggggacga ccccatgctc ctcgcctacc tcctggaccc
1741 ctccaacacc acccccgagg gggtggcgcg gcgctacggg ggggagtgga cggaggacgc
1801 cgcccaccgg gccctcctct cggagaggct ccatcggaac ctccttaagc gcctcgaggg
1861 ggaggagaag ctcctttggc tctaccacga ggtggaaaag cccctctccc gggtcctggc
1921 ccacatggag gccaccgggg tacggcggga cgtggcctac cttcaggccc tttccctgga
1981 gcttgcggag gagatccgcc gcctcgagga ggaggtcttc cgcttggcgg gccacccctt
2041 caacctcaac tcccgggacc agctggaaag ggtgctcttt gacgagctta ggcttcccgc
2101 cttggggaag acgcaaaaga caggcaagcg ctccaccagc gccgcggtgc tggaggccct
2161 acgggaggcc caccccatcg tggagaagat cctccagcac cgggagctca ccaagctcaa
2221 gaacacctac gtggaccccc tcccaagcct cgtccacccg aggacgggcc gcctccacac
2281 ccgcttcaac cagacggcca cggccacggg gaggcttagt agctccgacc ccaacctgca
2341 gaacatcccc gtccgcaccc ccttgggcca gaggatccgc cgggccttcg tggccgaggc
2401 gggttgggcg ttggtggccc tggactatag ccagatagag ctccgcgtcc tcgcccacct
2461 ctccggggac gaaaacctga tcagggtctt ccaggagggg aaggacatcc acacccagac
2521 cgcaagctgg atgttcggcg tccccccgga ggccgtggac cccctgatgc gccgggcggc
2581 caagacggtg aacttcggcg tcctctacgg catgtccgcc cataggctct cccaggagct
2641 tgccatcccc tacgaggagg cggtggcctt tatagagcgc tacttccaaa gcttccccaa
2701 ggtgcgggcc tggatagaaa agaccctgga ggaggggagg aagcggggct acgtggaaac
2761 cctcttcgga agaaggcgct acgtgcccga cctcaacgcc cgggtgaaga gcgtcaggga
2821 ggccgcggag cgcatggcct tcaacatgcc cgtccagggc accgccgccg acctcatgaa
2881 gctcgccatg gtgaagctct tcccccgcct ccgggagatg ggggcccgca tgctcctcca
2941 ggtccacgac gagctcctcc tggaggcccc ccaagcgcgg gccgaggagg tggcggcttt
3001 ggccaaggag gccatggaga aggcctatcc cctcgccgtg cccctggagg tggaggtggg
3061 gatgggggag gactggcttt ccgccaaggg ttaggggggc cctgccgttt agaggaagtt
3121 caaggggttg tccctcagaa acgcctccag gggaacgccc tctgggccta ccaggaggcc
3181 tttagcccca aaggtgcggg tgaaggcttc caggccctgg g
//
```

Appendix 3: Restriction Sites within the DNA sequence provided in appendix II

#	Enzyme	Specificity	Sites & flanks	Cut positions (blunt - 5' ext. - 3' ext.)
1	AgeI	A^CCGG_T	list	*706/710
2	AleI	CACNN^NNGTG	list	2228
3	ApaLI	G^TGCA_C	list	1589/1593
4	BamHI	G^GATC_C	list	2373/2377
5	BcgI	_NN^(N)$_{10}$CGA(N)$_6$TGC(N)$_{10}$_NN^	list	*652/650+686/684
6	BclI	T^GATC_A	list	#2478/2482
7	BfuAI	ACCTGCNNNN^NNNN_	list	2343/2347
8	BmgBI	CAC^GTC	list	*1951
9	BsaI	GGTCTCN^NNNN_	list	470/474
10	BsaWI	W^CCGG_W	list	706/710
11	BsgI	GTGCAG(N)$_{14}$_NN^	list	731/729
12	BsmBI	CGTCTCN^NNNN_	list	*1042/1046
13	BspHI	T^CATG_A	list	2874/2878
14	BspMI	ACCTGCNNNN^NNNN_	list	2343/2347
15	BsrFI	R^CCGG_Y	list	*706/710
16	BsrI	ACTG_GN^	list	3077/3075
17	BtsI	GCAGTG_NN^	list	1130/1128
18	DraIII	CAC_NNN^GTG	list	*1891/1888
19	EagI	C^GGCC_G	list	*1582/1586
20	EcoP15I	CAGCAG(N)$_{25}$^NN_	list	1629/1631
21	HindIII	A^AGCT_T	list	2689/2693
22	HpyCH4III	AC_N^GT	list	2587/2586
23	KasI	G^GCGC_C	list	*366/370
24	MslI	CAYNN^NNRTG	list	2228
25	NarI	GG^CG_CC	list	*367/369
26	PflFI	GACN^N_NGTC	list	1046/1047
27	PvuII	CAG^CTG	list	2062
28	SbfI	CC_TGCA^GG	list	1577/1573
29	SfoI	GGC^GCC	list	*368
30	SphI	G_CATG^C	list	2932/2928
31	TspRI	_NNCASTGNN^	list	1130/1121
32	Tth111I	GACN^N_NGTC	list	1046/1047
33	XbaI	T^CTAG_A	list	1/5
34	XmnI	GAANN^NNTTC	list	2897