Biology 113 Closed Book Take-Home Exam #1 – Information

There is no time limit on this test, though I have tried to design one that you should be able to complete within 3 hours. There are 6 pages in this test, including this cover sheet and the data gallery. You are not allowed to look at someone else's test, nor use your notes, old tests, the internet, any books, nor are you allowed to discuss the test with anyone until all exams are turned in no later than **8:30 am on Monday Sept. 19**. If you turn in your exam late, you will lose a letter grade for each day you are late. The **answers to the questions must be typed in this Word file** unless you are asked to draw on a separate page, or you want to use scratch paper. If you do not write your answers in the appropriate location, I may not find them. Tell me where to look if you put your answer at the back of your test.

I have provided you with a "Data Gallery" in the form of figures and tables. To choose a figure in support of your answer, simply state Figure #x. Do not assume how many of the data images you will use, or not use. **Simply choosing the data is not sufficient support for your answer. You must explain the significance of the data and how they support your answer.** *I have given you word limits so be concise.*

-3 pts if you do not follow this direction.
Please do not write or type your name on any page other than this cover page.
Staple all your pages (INCLUDING THE TEST PAGES) together when finished with the exam.

Name (please print):

Read the pledge and sign if you can do so with honor:

_____

On my honor I have neither given nor received unauthorized information regarding this work, I have followed and will continue to observe all regulations regarding it, and I am unaware of any violation of the Honor Code by others.

How long did this exam take you to complete?

Lab blended with lecture Questions:

4 pts.

1) How did you generate sticky ends on the promoter DNA that you cloned using GGA? *Answer Limit: 30 words.*

Added 4 bases to 5' end of both oligos to complement sticky ends on cut plasmid.

6 pts.

2) In biology, it is important to know how to set up dilutions so that your research will work. If you boiled and cooled your oligos at a concentration of 7.5 μM but you need them to be 30 nM in GGA, how much of each reagent would you pipet into the GGA reaction tube?

oligos: __1 μL of diluted oligos (see comment below)

10X buffer: __1μL___

water: __6μL___

ligase: 1 μL

BsaI: 1 μL

Final volume: 10 μL

Additional comments (maximum of 40 words):

It was acceptable to dilute either 250 fold as we did in lab, or 25 fold for final concentration of 30 nM.

Lecture Questions:

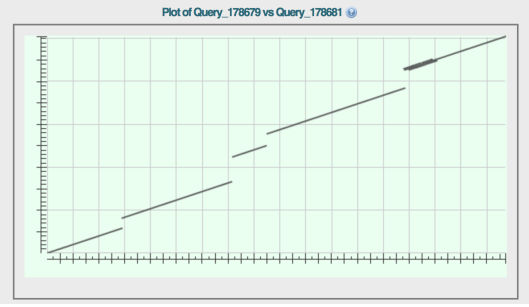
10 pts.

3) For this question, you will need to use two web tools from NCBI:

BLAST2 (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&BLAST_SPEC=&LINK_LOC=blasttab&LAST_PAGE=blastn&BLAST_INIT)

ORF finder (http://www.ncbi.nlm.nih.gov/orffinder)

You will also need to use the additional data file called wrinkled_smooth.docx. The data file contains an mRNA and a gene.



a) Tell me which sequence is the gene. Support your answer with data. *Answer Limit: 30 words.*

Sequence_2 is the gene. It is longer than the mRNA.

b) How many introns does the gene have? Support your answer with data. *Answer Limit: 30 words.*

Four introns (gaps).

c) Tell me the second amino acid and the last amino acid in the encoded protein. Support your answer with data.

Answer Limit: 30 words.
2nd = V, last = Q see sequences located after the data gallery.

15 pts.
4) Here are some more sequence-based questions
a) Translate this ORF using the single letter amino acid code in the data gallery:
UUUCUAGAUGGAAGAGACCUGGAUACUGGACUGCACAUCAUGAAUGAGA

translation here: MEETWILDCTS (ORF begins with start codon)

b) In the sequence below, the +1 base is under the number 1 and underlined. Label these seven aspects of a gene in the sequence below: start codon; start transcription site; -10 box; -35 box; 3' end; direction RNA polymerase would move; signal sequence coding region.

RNA polymerase direction → -35
CACCCGGTACTGGGCAGAAAATATATGAGATTGATCCACTTTTGCAAGCTCACCGTCAGCATCTTGATTT

-10 +1=start transcription start codon
CCGTTATGGCTATATGATGAGAATTCGGGAGGAAATTGACAAATAAGAAGGTGGTCTTGATGCATTTTCG

signal sequence
CGTGGGTACGAAAAGTTTGGCTTCACACGCAGTGCTACAGGCATTACTTACAGGGAGTGGGCACCTGGAG
3'

15 pts.
5) There is a some irony and a pun in the fact that a celibate monk is the "father" of genetics. Please answer these not so funny, but interesting, questions.
a) What two very important aspects of meiosis that contribute to the random nature of genetic crosses did Mendel discover when he conducted his dihybrid crosses? Name these two and provide the evidence Mendel used to reach his insights into meiosis. *Answer Limit: 35 words for each one.*
1. Independent assortment: pea color and texture gave every possible combination 9:3:3:1 ratios
2. Segregation: 1 allele in each gamete and never 2 (blended colors)

b) Which one phase of meiosis generates the greatest amount of unpredictable diversity in progeny? Support your answer with data. *Answer Limit: 30 words.*
Prophase I when recombination happens. (Figure #2)

16 pts.
6) Here are some genetics problems I would like you to answer. Show your work if you want to be eligible for partial credit.

a) A man and a woman want to have children but they are concerned about what might happen. The man has an upside down nose with nostrils pointed up, a rare recessive trait, and he risks drowning every time it rains. The woman is a carrier for upside down nose. What is the probability that their first pregnancy will produce a boy at risk of drowning in the rain?
¼
b) If the same couple already has a boy who is at risk of drowning in the rain, what is the probability of the second pregnancy being a girl who is not at risk, given the trait is not sex-linked?
¼
c) The most common form of muscular dystrophy is a sex linked disease. A husband does not have muscular dystrophy but the woman is a carrier. What is the probability of them producing a girl with muscular dystrophy?
0% chance
d) What is the probability that the couple in question c could have 4 boys from 4 different pregnancies, all of whom have muscular dystrophy?
1/256

15 pts.
7) Today, there is no question that DNA is the heritable material, but this was not always the case.
a) In no more than six steps, outline the experimental design that best disproved protein was the heritable material. *Answer Limit: 20 words per step.*
1. grow phage so that label DNA with ^{32}P and protein with ^{35}S (separately)
2. infect E. coli cells for 5 minutes
3. blend cell/phage mixture for 2 minutes
4. centrifuge to pellet cells and anything in or on the cells
5. quantify both forms of radiation in cell-free media
6.

b) On a separate piece of paper, draw a picture that illustrates what it means when people say that DNA replication is semiconservative. Label neatly.
Various ways to do this, but needed to show old and new strands

c) Choose exactly two time points from any experiment shown in the data gallery that led to the discovery of semiconservative DNA replication. Explain why the precise data you chose led to this conclusion. *Answer Limit: 40 words.*
Figure 7, generations 1 (one band disproved conservative) and 1.9 (two bands disproved mosaic)

9 pts.

8) You have been working with a promoter in Bio113 lab. Answer these questions about promoters.

a) Do transcription factors and RNA polymerase all bind in a row to adjacent DNA with no extra DNA bases between them? Support your answer with data. *Answer Limit: 30 words.*

No, they can bind with gaps between them (Figure 5)

b) On a separate piece of paper, draw a picture that illustrates the approximate size comparison of a nucleotide and RNA polymerase. Label neatly.

RNA polymerase covered all bases on top row of Figure 19.

c) How does a transcription factor "know" where it should bind within a gene? Support your answer with data. *Answer Limit: 40 words.*

Protein feels base pairs in major groove (Figure 12)

10 pts.

9) DNA is one of the most famous TLA (three letter acronym). But most people do not know what you know about DNA.

a) What is sometimes referred to as the 5th DNA base? Support your answer with data. *Answer Limit: 30 words.*

Methylated cytosine (Figure #1)

b) Explain how this 5th base can affect gene activity. Support your answer with data. *Answer Limit: 30 words.*

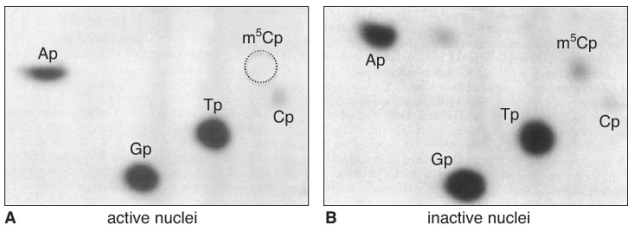
Methylation suppresses gene activity (Figure #25 un-suppressed fetal hemoglobin)

c) In order to use an epigenetic modifier to treat a human disease, what major improvement would be needed compared to the experiment shown in Figure 25? *Answer Limit: 30 words.*

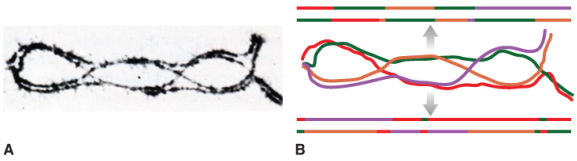
Needs to be more specific (cell type and/or genes)

Data Gallery

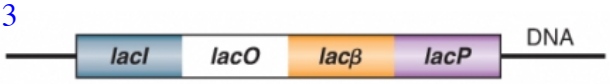
1



2



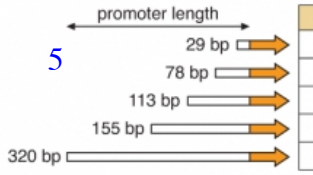
3



4

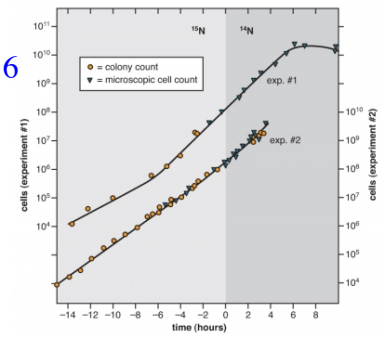
sample source	extracellular	intracellular
^{35}S-Protein Figure 1.8	~80%	~20%
^{32}P-DNA Figure 1.8	~30%	~70%
^{35}S-Protein refined experiment	~99%	~1%
^{32}P-DNA refined experiment	~30%	~70%

5

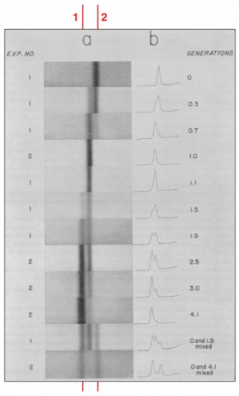


doubling time	drug resistant
no growth	none
5 hours	none
5 hours	none
3 hours	yes
3 hours	yes

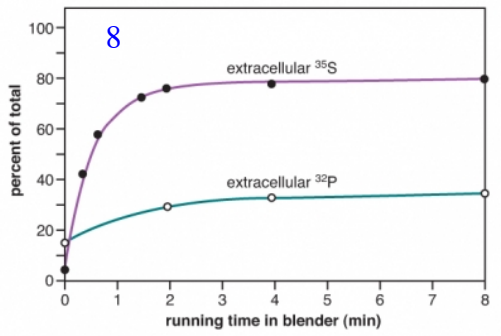
6



7



8



9

sample source	extracellular	intracellular
^{35}S-Protein Figure 1.8	~80%	~20%
^{32}P-DNA Figure 1.8	~30%	~70%
^{35}S-Protein refined experiment	~99%	~1%
^{32}P-DNA refined experiment	~30%	~70%

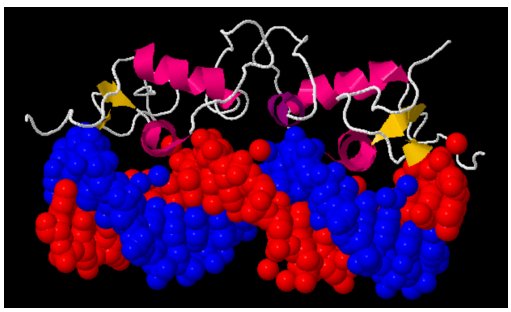
10

position #	1	2	3	4	5	6	7
A	−6.64	1.84	−6.64	0.84	1.26	−6.64	−0.72
C	−6.64	−6.64	−0.37	−6.64	−6.64	−6.64	−6.64
G	−0.37	−6.64	−6.64	1.18	−0.37	−6.64	1.92
T	1.57	−6.64	1.57	−6.64	−0.72	1.84	−6.64

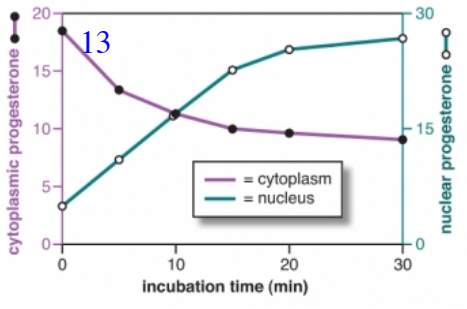
11

plant number	smooth pea	wrinkled pea	plant number	yellow pea	green pea
1	45	12	1	25	11
2	27	8	2	32	7
3	24	7	3	14	5
4	19	10	4	70	27
5	32	11	5	24	13
6	26	6	6	20	6
7	88	24	7	32	13
8	22	10	8	44	9
9	28	6	9	50	14
10	25	7	10	44	18
totals	336	101	totals	355	123

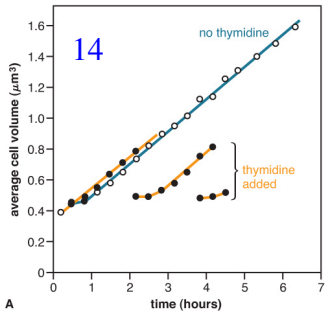
12



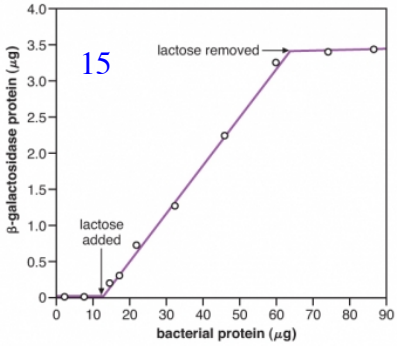
13



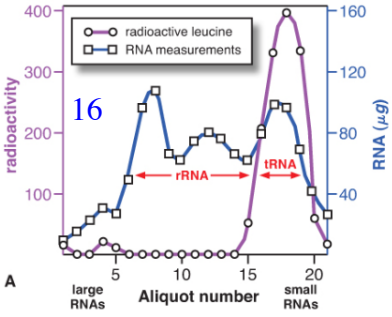
14



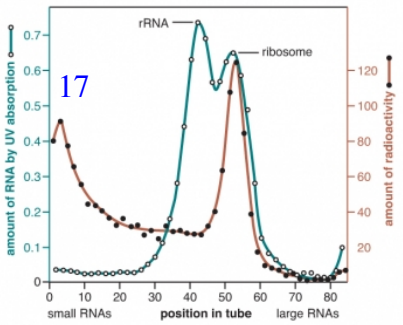
15



16



17



18

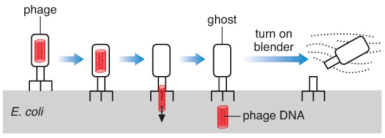
second base in codon

first base in codon	U	C	A	G
U	UUU phe F UUC phe F UUA leu L UUG leu L	UCU ser S UCC ser S UCA ser S UCG ser S	UAU tyr Y UAC tyr Y UAA stop UAG stop	UGU cys C UGC cys C UGA stop UGG trp W
C	CUU leu L CUC leu L CUA leu L CUG leu L	CCU pro P CCC pro P CCA pro P CCG pro P	CAU his H CAC his H CAA gln Q CAG gln Q	CGU arg R CGC arg R CGA arg R CGG arg R
A	AUU ile I AUC ile I AUA ile I AUG met M	ACU thr T ACC thr T ACA thr T ACG thr T	AAU asn N AAC asn N AAA lys K AAG lys K	AGU ser S AGC ser S AGA arg R AGG arg R
G	GUU val V GUC val V GUA val V GUG val V	GCU ala A GCC ala A GCA ala A GCG ala A	GAU asp D GAC asp D GAA glu E GAG glu E	GGU gly G GGC gly G GGA gly G GGG gly G

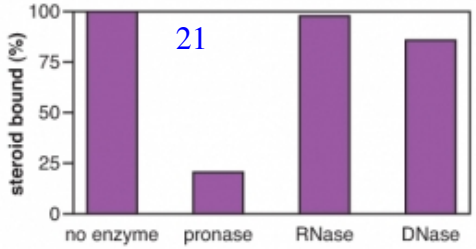
19

```
V-T7      5′...TAAACACGGTACGATGTACCACATGAAACGACAGTGAGTC... 3′
V-fd      5′...GCTTCTGACTATAATAGACAGGGTAAAGACCTGATTTTTG... 3′
V-SV40    5′...ATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCA..... 3′
V-1       5′...ACTGGCGGTGATACTGAGCACATCAGCAGGACGCACTGAC... 3′
B-tRNA    5′...GTCATTTGATATGATGCGCCCCGCTTCCCGATAAGGAGC  ... 3′
B-Lac     5′...TCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAA... 3′
```

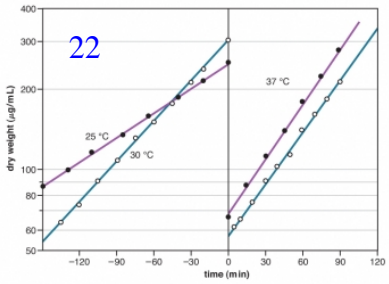
20



21

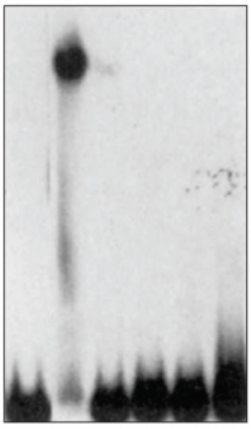


22

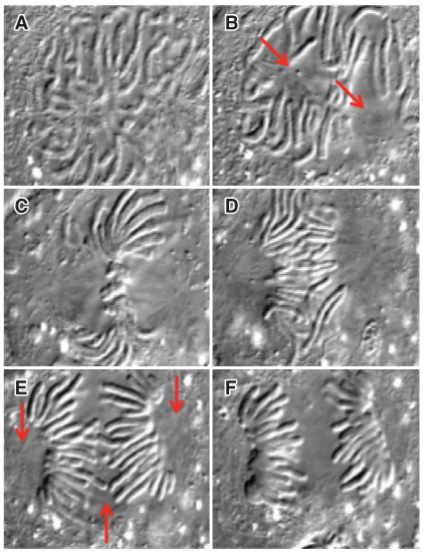


TBP	+		+	+	+
TFIIB	+	+		+	+
PAR 74	+	+	+		+
RNA pol	+	+	+	+	

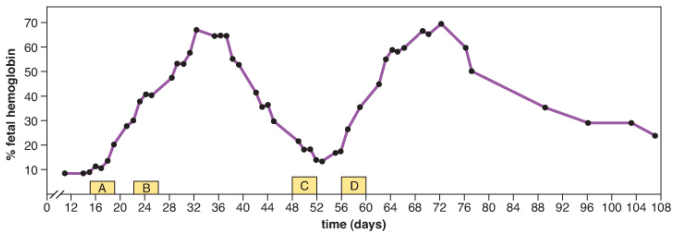
23



24



25



>Sequence_1 = mRNA and ORF is color coded
AAGAAGAAGATGGTTTATACAATCTCAGGAATTCGATTTCCCGTTCTTCCTTCTCTGCACAAGTCGACCT
TACGCTGCGATCGCAGAGCTTCATCTCACTCTTTTTTTCTCAAGAACAACTCGTCATCATTCTCTCGGAC
CTCACTTTATGCAAAGTTTTCCCGCGATTCTGAAACCAAATCTTCCACAATTGCTGAATCTGATAAAGTA
CTTATTCCTGAAGATCAAGATAACTCTGTATCCTTGGCAGATCAACTTGAAAATCCTGATATAACCTCAG
AGGATGCACAGAACTTGGAGGATTTAACCATGAAAGATGGGAATAAGTACAATATTGATGAATCAACTAG
CAGTTATAGAGAGGTTGGAGATGAGAAAGGTTCTGTTACGTCATCATCACTTGTAGATGTTAACACCGAC
ACTCAAGCCAAGAAGACATCAGTTCATTCAGACAAGAAAGTAAAAGTAGATAAACCTAAGATCATTCCTC
CACCCGGTACTGGGCAGAAAATATATGAGATTGATCCACTTTTGCAAGCTCACCGTCAGCATCTTGATTT
CCGTTATGGACAATACAAAAGAATTCGGGAGGAAATTGACAAATATGAAGGTGGTCTTGATGCATTTTCG
CGTGGGTACGAAAAGTTTGGCTTCACACGCAGTGCTACAGGCATTACTTACAGGGAGTGGGCACCTGGAG
CTAAGTCAGCAGCATTAGTTGGAGATTTCAACAATTGGAATCCAAATGCAGATGTAATGACTAAGGACGC
TTTTGGTGTATGGGAGATCTTCTTGCCAAACAATGCCGATGGTTCGCCACCAATTCCTCATGGTTCTCGA
GTCAAGATCCACATGGATACTCCCTCTGGGATCAAGGACTCGATTCCTGCTTGGATCAAATTCTCTGTAC
AGGCTCCTGGTGAAATTCCATACAATGGAATATACTATGATCCCCCAGAGGAGGAAAAGTATGTCTTCAA
ACATCCACAGCCAAAACGACCACAGTCAATTAGAATATATGAGTCACACATTGGAATGAGTAGCCCGGAG
CCAAAAATCAACACATATGCGAACTTTAGAGATGATGTACTACCTCGCATTAAAAAACTTGGCTATAATG
CTGTCCAGATTATGGCTATCCAAGAACATTCTTATTATGCTAGTTTTGGGTACCATGTTACTAATTTCTT
TGCACCTAGCAGTCGGTTTGGTACTCCAGAAGATCTCAAGTCTCTGATAGACAGAGCCCATGAACTAGGC
CTGCTTGTTCTGATGGATATTGTACATAGCCATTCCTCAAATAATACATTGGATGGCCTGAACATGTTTG
ATGGAACTGATGGTCATTACTTCCATCCTGGTTCACGGGGTTATCATTGGATGTGGGATTCTCGCCTTTT
TAACTATGGAAGCTGGGAAGTACTAAGGTACCTACTTTCAAATGCAAGATGGTGGCTGGATGAATATAAG
TTTGATGGGTTTCGATTTGATGGCGTCACATCAATGATGTACACTCATCATGGACTGCAGGTATCTTTTA
CCGGAAATTACAGTGAGTATTTTGGTTTAGCAACTGATGTTGAGGCTGTGGTTTACATGATGCTTGTTAA
TGATCTCATTCACGGGCTCTTCCCTGAAGCTGTTTCAATTGGTGAAGATGTGAGTGGAATGCCAACATTC
TGCCTTCCGACGCAAGATGGTGGGATTGGCTTTAACTACCGCTTGCATATGGCTGTTGCAGACAAGTGGA
TTGAGCTTCTTAAGAAACAAGATGAAGATTGGAGAATGGGTGATATAGTCCACACACTGACAAACAGAAG
GTGGCTGGAAAAATGTGTGGTTTATGCTGAAAGTCATGACCAGGCCTTGGTTGGTGATAAGACACTAGCA
TTTTGGTTGATGGACAAGGATATGTATGACTTCATGGCTTTAGACAGGCCATCTACTCCTCTAATAGATC
GTGGTATAGCATTACACAAAATGATTCGGCTTATTACTATGGGCCTTGGTGGTGAAGGGTATTTGAATTT
TATGGGGAATGAATTCGGACATCCTGAGTGGATCGATTTTCCAAGGGGTGAGCAACATCTTCCTAATGGC
AAAATAGTTCCTGGGAATAACAACAGTTATGATAAATGCAGACGTAGATTTGACTTGGGTGATGCAGATT
ATCTAAGATATCATGGAATGCAAGAATTTGATCGAGCTATGCAGCATCTAGAAGAGAGATATGGTTTCAT
GACTTCTGAGCACCAATATATTTCACGGAAAAATGAAGGTGACAGAGTTATAATCTTTGAAAGAGACAAC
CTGGTCTTTGTCTTCAATTTTCACTGGACCAACAGTTATTCAGATTACAAGGTTGGCTGTTTAAAGCCAG
GGAAATATAAGATTGTCTTGGATTCAGATGATACCTTGTTTGGTGGCTTCAATAGGCTCAATCATACTGC
TGAGTACTTCACCTCTGAAGGGTGGTATGATGACCGGCCTCGATCTTTCCTCGTTTATGCACCTTCTAGA
ACAGCAGTGGTGTATGCCCTTGCGGATGGAGTTGAATCAGAGCCAATTGAACTGTCAGATGGAGTAGAAT
CAGAGCCAATTGAGCTGTCAGTTGGAGTTGAATCAGAGCCAATTGAGCTGTCAGTTGAAGAAGCTGAATC
AGAGCCAATTGAGCGGTCAGTTGAAGAAGTTGAATCAGAGACAACTCAACAGTCAGTTGAAGTTGAATCA
GAGACAACTCAACAGTCAGTTGAAGTTGAATCAGAGACAACTCAATAGTCAGTTGAAGTTGAACCATGAT
GACAAGTTTTTCTGGAATGAATATATCAATTTCATAATGGCATCAGTGCAATCAAGTCAAAACCAGCCAA
CCAACTTAAGGTGATTAGAGCAGGCTCTGAAACTGCAAGGCGGAGCCCTTGTGTTAACTTAATCTTTAGT
GCGGCAGCAAACCAACGTTGAAGTGTTTGTACCCATTGCTTGCTAGGTCAATTATGTATAGCACGTCAAG
TTTTTTGTTTTTACTTGGGGTTTAATAGGCACAAGTTGCGTGGCGTGTGAATGTGAAACAACAAATGTAC
TGTCCTCAGATCCTCCGTTTCTGTGGACATATAAATCCTTTTTACCCATAAGTGCACAAACCTTAATTAA
ACCCTGGAAAGTTTTGCGGGAGGCCTTATATATGATTTCTCCTTTCAGAATTAAAAGTTGAGTCATAGCA
ATAGGGCCGTGCTGATGAAATAATTGTTTTGTACACGGTGAAAATAGTATTGCCTTACTATAAAATATAA
ATGGTGGAACCTTTGCCTCCAATTGCAAAAAAAAAAAAAAAAAGGAATTC

>lcl|ORF1_1:9:2777 unnamed protein product
MVYTISGIRFPVLPSLHKSTLRCDRRASSHSFFLKNNSSSFSRTSLYAKFSRDSETKSSTIAESDKVLIP
EDQDNSVSLADQLENPDITSEDAQNLEDLTMKDGNKYNIDESTSSYREVGDEKGSVTSSSLVDVNTDTQA
KKTSVHSDKKVKVDKPKIIPPPGTGQKIYEIDPLLQAHRQHLDFRYGQYKRIREEIDKYEGGLDAFSRGY
EKFGFTRSATGITYREWAPGAKSAALVGDFNNWNPNADVMTKDAFGVWEIFLPNNADGSPPIPHGSRVKI
HMDTPSGIKDSIPAWIKFSVQAPGEIPYNGIYYDPPEEEKYVFKHPQPKRPQSIRIYESHIGMSSPEPKI
NTYANFRDDVLPRIKKLGYNAVQIMAIQEHSYYASFGYHVTNFFAPSSRFGTPEDLKSLIDRAHELGLLV
LMDIVHSHSSNNTLDGLNMFDGTDGHYFHPGSRGYHWMWDSRLFNYGSWEVLRYLLSNARWWLDEYKFDG
FRFDGVTSMMYTHHGLQVSFTGNYSEYFGLATDVEAVVYMMLVNDLIHGLFPEAVSIGEDVSGMPTFCLP
TQDGGIGFNYRLHMAVADKWIELLKKQDEDWRMGDIVHTLTNRRWLEKCVVYAESHDQALVGDKTLAFWL
MDKDMYDFMALDRPSTPLIDRGIALHKMIRLITMGLGGEGYLNFMGNEFGHPEWIDFPRGEQHLPNGKIV
PGNNNSYDKCRRRFDLGDADYLRYHGMQEFDRAMQHLEERYGFMTSEHQYISRKNEGDRVIIFERDNLVF
VFNFHWTNSYSDYKVGCLKPGKYKIVLDSDDTLFGGFNRLNHTAEYFTSEGWYDDRPRSFLVYAPSRTAV
VYALADGVESEPIELSDGVESEPIELSVGVESEPIELSVEEAESEPIERSVEEVESETTQQSVEVESETT
QQSVEVESETTQ

>Sequence_2 introns are color coded
AAGAAGAAGATGGTTTATACAATCTCAGGAATTCGATTTCCCGTTCTTCCTTCTCTGCACAAGTCGACCT
TACGCTGCGATCGCAGAGCTTCATCTCACTCTTTTTTTCTCAAGAACAACTCGTCATCATTCTCTCGGAC
CTCACTTTATGCAAAGTTTTCCCGCGATTCTGAAACCAAATCTTCCACAATTGCTGAATCTGATAAAGTA
CTTATTCCTGAAGATCAAGATAACTCTGTATCCTTGGCAGATCAACTTGAAAATCCTGATATAACCTCAG
AGGATGCACAGAACTTGGAGGATTTAACCATGAAAGATGGGAATAAGTACAATATTGATGAATCAACTAG
CAGTTATAGAGAGGTTGGGGATGATGATATTGCTGCGCTCGTTGTTGACAATGGCTCCGGTATGTGCAAG

GCCGGTTTCGCCGGGGACGATGCCCCCCCGTGCTGTGTTCCCATCTATCGTGGGTCGCCCCAGACATCAGG
GTGTGATGGTTGGTATGGGCCAGAAAGACAGCTACGTTGGTGATGAAGCCCAGAGCAAAAGAGGTATCCT
GACCCTGAAGTACCCCATTGAACACGGTATTGTCACCAACTGGGAGATGAGAAAGGTTCTGTTACGTCAT
CATCACTTGTAGATGTTAACACCGACACTCAAGCCAAGAAGACATCAGTTCATTCAGACAAGAAAGTAAA
AGTAGATAAACCTAAGATCATTCCTCCACCCGGTACTGGGCAGAAAATATATGAGATTGATCCACTTTTG
CAAGCTCACCGTCAGCATCTTGATTTCCGTTATGGACAATACAAAAGAATTCGGGAGGAAATTGACAAAT
ATGAAGGTGGTCTTGATGCATTTTCGCGTGGGTACGAAAAGTTTGGCTTCACACGCAGTGCTACAGGCAT
TACTTACAGGGAGTGGGCACCTGGAGCTAAGTCAGCAGCATTAGTTGGAGATTTCAACAATTGGAATCCA
AATGCAGATGTAATGACTAAGGACGCTTTTGGTGTATGGGAGATCTTCTTGCCAAACAATGCCGATGGTT
CGCCACCAATTCCTCATGGTTCTCGAGTCAAGATCCACATGGATACTCCCTCTGGGATCAAGGACTCGAT
TCCTGCTTGGATCAAATTCTCTGTACAGGCTCCTGGTGAAATTCCATACAATGGAATATACTATGATCCC
CCAGAGGAGGAAAAGTATGTCTTCAAACATCCACAGCCAAAACGACCACAGTCAATTAGAATATATGAGT
CACACATTGGAATGAGTAGCCCGGAGCCAAAAATCAACACATATGCGAACTTTAGAGATGATGTACTACC
TCGCATTAAAAAACTTGGCTATAATGCTGTCCAGATTATGGCTATCCAAGAACATTCTTATTATGCTAGT
TTTGGGTACCATGTTACTAATTTCTTTGCACCTAGCAGTCGGTTTGGTACTCCAGAATATGGAGAAGATC
TGGCACCACACTTTCTACAATGAGCTGAGAGTAGCCCCTGAGGAGCACCCTGTGCTGCTCACAGAGGCTC
CCCTGAACCCCAAAGCCAACAGAGAGAAGATGACACAGATCATGTTTGAGACCTTCAACACCCCAGCCAT
GTATGTAGCCATCCAGGCTGTGCTGTCCCTGTATGCCTCTGGTCGTACCACTGGTATTGTGATGGACTCT
GGTGATGGTGTTACCCACACTGTGCCCATCTATGAAGGCTACGCCCTCCCCCATGCCATCCTCCGTCTGG
ATCTGGCTGGCCGTGACCTGACGGACTACCTCATGAAGATCCTGACAGAGAGAGGCTACAGCTTCACCAC
CACAGCCGAGAGAGAAATTGTGCGTGACATCAAGGAGAAGCTGTGCTACGTCGCACTGGATTTCGAGCAG
GAGATGGCCACAGCTGCCTCTAGCTCTTCCCTGGAGAAGAGCTATGAACTCCCTGATGGTCAGGTCATCA
CCATTGGCAATGAGAGGTTCAGGTGCCCCGAGGCCCTCTTCCAGCCATCTTTCTTGGGTATGGAGTCCGA
TCTCAAGTCTCTGATAGACAGAGCCCATGAACTAGGCCTGCTTGTTCTGATGGATATTGTACATAGCCAT
TCCTCAAATAATACATTGGATGGCCTGAACATGTTTGATGGAACTGATGGTCATTACTTCCATCCTGGTT
CACGGGGTTATCATTGGATGTGGGATTCTCGCCTTTTTAACTATGGAAGCTGGGAAGTACTAAGGTACCT
ACTTTCAAATGCAAGATGGTGGCTGGATGAATATAAGTTTGATGGGTTTCGATTTGGTATCCATGAAACT
ACCTTCAACTCCATCATGAAGTGTGATGTGGATATCCGTAAGGATCTGTATGCCAACACAGTGCTGTCTG
GTGGTACCACAATGTACCCTGGCATTGCTGACAGGATGCAGAAGGAGATCACAGCCCTGGCACCTAGCAC
AATGAAAATCAAGATCATTGCCCCACCTGAGCGCAAGTACTCTGTCTGGATTGGAGGCTCTATCCTGGCC
TCCCTGTCCACCTTCCAGCAGATGTGGATCAGCAAGCAGGAGTACGGATGGCGTCACATCAATGATGTAC
ACTCATCATGGACTGCAGGTATCTTTTACCGGAAATTACAGTGAGTATTTTGGTTTAGCAACTGATGTTG
AGGCTGTGGTTTACATGATGCTTGTTAATGATCTCATTCACGGGCTCTTCCCTGAAGCTGTTTCAATTGG
TGAAGATGTGAGTGGAATGCCAACATTCTGCCTTCCGACGCAAGATGGTGGGATTGGCTTTAACTACCGC
TTGCATATGGCTGTTGCAGACAAGTGGATTGAGCTTCTTAAGAAACAAGATGAAGATTGGAGAATGGGTG
ATATAGTCCACACACTGACAAACAGAAGGTGGCTGGAAAAATGTGTGGTTTATGCTGAAAGTCATGACCA
GGCCTTGGTTGGTGATAAGACACTAGCATTTTGGTTGATGGACAAGGATATGTATGACTTCATGGCTTTA
GACAGGCCATCTACTCCTCTAATAGATCGTGGTATAGCATTACACAAAATGATTCGGCTTATTACTATGG
GCCTTGGTGGTGAAGGGTATTTGAATTTTATGGGGAATGAATTCGGACATCCTGAGTGGATCGATTTTCC
AAGGGGTGAGCAACATCTTCCTAATGGCAAAATAGTTCCTGGGAATAACAACAGTTATGATAAATGCAGA
CGTAGATTTGACTTGGGTGATGCAGATTATCTAAGATATCATGGAATGCAAGAATTTGATCGAGCTATGC
AGCATCTAGAAGAGAGATATGGTTTCATGACTTCTGAGCACCAATATATTTCACGGAAAAATGAAGGTGA
CAGAGTTATAATCTTTGAAAGAGACAACCTGGTCTTTGTCTTCAATTTTCACTGGACCAACAGTTATTCA
GATTACAAGGTTGGCTGTTTAAAGCCAGGGAAATATAAGATTGTCTTGGATTCAGATGATACCTTGTTTG
GTGGCTTCAATAGGCTCAATCATACTGCTGAGTACTTCACCTCTGAAGGGTGGTATGATGACCGGCCTCG
ATCTTTCCTCGTTTATGCACCTTCTAGAACAGCAGTGGTGTATGCCCTTGCGGATGGAGTGAATCCGGAC
CCTCCATTGTCCACCGCAAATGCTTCTAAACCGGACTGTTACCAACACCCACACCCCTGTGATGAAACAA
AACCCATAAATGCGCATAAAACAAGACGAGATTGGCATGGCTTTATTTGTTTTTTCTTTTGGCGCTTGAC
TCAGGATTAAAAAACTGGAATGGTGAAGGTGTCAGCAGCAGTCTTAAAATGAAACATGTTGGAGCGAACG
CCCCCAAAGTTCTACAATGCATCTGAGGACTTTGATTGTACATTTGTTTCTTTTTTAATAGTCATTCCAA
ATATTGTTATAATGCATTGTTACAGGAAGTTACTCGCCTCTGTGAAGGCAACAGCCCAGCTGGGAGGAGC
CGGTACCAATTACTGGTGTTAGATGATAATTGCTTGTCTGTAAATTATGTAACCCAACAAGTGTCTGAAT
CAGAGCCAATTGAACTGTCAGATGGAGTAGAATCAGAGCCAATTGAGCTGTCAGTTGGAGTTGAATCAGA
GCCAATTGAGCTGTCAGTTGAAGAAGCTGAATCAGAGCCAATTGAGCGGTCAGTTGAAGAAGTTGAATCA
GAGACAACTCAACAGTCAGTTGAAGTTGAATCAGAGACAACTCAACAGTCAGTTGAAGTTGAATCAGAGA
CAACTCAATAGTCAGTTGAAGTTGAACCATGATGACAAGTTTTTCTGGAATGAATATATCAATTTCATAA
TGGCATCAGTGCAATCAAGTCAAAACCAGCCAACCAACTTAAGGTGATTAGAGCAGGCTCTGAAACTGCA
AGGCGGAGCCCTTGTGTTAACTTAATCTTTAGTGCGGCAGCAAACCAACGTTGAAGTGTTTGTACCCATT
GCTTGCTAGGTCAATTATGTATAGCACGTCAAGTTTTTTGTTTTTACTTGGGGTTTAATAGGCACAAGTT
GCGTGGCGTGTGAATGTGAAACAACAAATGTACTGTCCTCAGATCCTCCGTTTCTGTGGACATATAAATC
CTTTTTACCCATAAGTGCACAAACCTTAATTAAACCCTGGAAAGTTTTGCGGGAGGCCTTATATATGATT
TCTCCTTTCAGAATTAAAAGTTGAGTCATAGCAATAGGGCCGTGCTGATGAAATAATTGTTTTGTACACG
GTGAAAATAGTATTGCCTTACTATAAAATATAAATGGTGGAACCTTTGCCTCCAATTGCAAAAAAAAAAA
AAAAAGGAATTC