

Submission deadline: 29 Nov, 23:59

- Write your answers(images) and C++ code inside the *answers* folder in order to generate a single PDF file. Replace the images and cpp files that are already included in the project. Do not change the file name.
- Read the questions carefully and write your answers clearly. Answers that are not legible and that doesn't follow the format will not have any score.

Outcomes:

- a. Apply appropriate mathematical and related knowledge to computer science.
 - b. Analyze problems and identify the appropriate computational requirements for its solution.
-

A. Longest Common Subsequence

Problem 1 (Outcome b) - 4 points

Consider the problem presented in lectures to find the Longest Common Subsequence of the following strings:

String 1	CTGTCCAGC
String 2	ACCATGCT

- 4
- Write down the table, the recurrence and the procedure needed to solve the problem using dynamic programming
 - Once the table is constructed, explain how can we obtain the longest common subsequence between both strings? Write the final answer.
 - Explain the foundation behind the optimal substructure principle for the Longest Common Subsequence problem.

PROBLEM 1

a)

TABLE

"a ACC AT GCT
0 0 0 0 0 0 0 0 0
C 0 0 1 1 1 1 1 1 1
T 0 0 1 1 1 2 2 2 2
G 0 0 1 1 1 2 3 3 3
T 0 0 1 1 1 2 3 3 4
C 0 0 1 1 2 2 3 3 4
G 0 0 1 1 2 2 3 4 4
A 0 1 1 2 3 3 4 4
G 0 1 1 2 3 3 4 4
C 0 1 2 2 3 3 4 5

PROCEDURE

```

1 LCS(X, Y)
2 m = X.length
3 n = Y.length
4 let c[i][j] be a new table
5 for i=0 to m
6   for j=0 to n
7     c[i][j] = 0
8 for i=1 to m
9   for j=1 to n
10   if x[i] == y[j]
11     c[i][j] = 1 + c[i-1][j-1]
12   else
13     c[i][j] = max(c[i-1][j], c[i][j-1])
14 return c[m][n]

```

b)

We must traverse through table c, starting from $c[m,n]$, checking the following conditions to obtain the LCS:

Let S be the sequence than will contain the LCS.

If $x_i = y_j$, then we add this character to the first position of S (like a "push_front"), and move to the cell one row above and one column to the left.

Otherwise, we must consider moving to the cell to the left or to the cell above the current cell being traversed. We move to the cell containing the maximum value between these two.

Stop this procedure if we moved to an empty string cell (cell highlighted with ).

Else, go back to .

"a ACC AT GCT
0 0 0 0 0 0 0 0 0
C 0 0 1 1 1 1 1 1 1
T 0 0 1 1 1 2 2 2 2
G 0 0 1 1 1 2 3 3 3
T 0 0 1 1 1 2 3 3 4
C 0 0 1 1 2 2 3 3 4
G 0 0 1 1 2 2 3 4 4
A 0 1 1 2 3 3 4 4
G 0 1 1 2 3 3 4 4
C 0 1 2 2 3 3 4 5

Final answer: CCAGC

c)

The prefixes of an LCS for two given sequences X and Y are optimal solutions to the prefixes of X and Y .

For example:

CCAGC is the LCS of CTGTCCAGC and ACCATGCT.

If we take the prefix CCAG or CCAGC, we realize that this prefix is an LCS of CTGTCCAG and ACCATG, which are prefixes of X and Y , respectively.

Thus, this problem has an optimal-substructure property.

B. Knapsack Problem

Problem 2 (Outcome b) - 3 points

Consider the following items for the Unbounded Knapsack problem:

3

Weights (Kg)	1	3	8	2	5
Values (S.)	10	5	20	15	12

- Write down the table, the recurrence and the procedure needed to solve the problem using dynamic programming considering a knapsack that supports 10 Kg
- Once the table is constructed, explain how can we obtain the longest common subsequence between both strings? Write the final answer.

PROBLEM 2

a) TABLE

Note : THE ITEMS COLUMNS ARE USED TO APPLY THE INTUITION TO THE PROBLEM; BUT THE ANSWER FOR A GIVEN K_5 IS SAVED IN THE LAST COLUMN.

ITEM	ITEM 1	ITEM 2	ITEM 3	ITEM 4	ITEM 5	CAPACITY
1	0	10	30	80	10	0
2	10	0	10	10	10	10
3	30	10	0	10	10	20
4	80	10	10	0	10	30
5	10	10	10	10	0	40
6	0	50	50	50	50	50
7	50	0	50	50	50	60
8	50	50	0	50	50	70
9	50	50	50	0	50	80
10	50	50	50	50	0	90
11	100	100	100	100	100	100

FINAL TABLE

CAPACITY 0 10 20 30 40 50 60 70 80 90 100
MAX PROFIT 0 10 20 30 40 50 60 70 80 90 100

PROCEDURE

UNBOUNDED-KNAPSACK (weight, weights, values)
 1: let $K[0..weight]$ be a new array
 2: for $i=0$ to weight
 3: for $j=1$ to weights.length
 4: if $weights[j] \leq i$
 5: $K[i] = \max(K[i], K[i - weights[j]] + values[j])$
 6: return $K[weight]$

RECURSION

LET n REPRESENT THE WEIGHT AVAILABLE; v_i BE THE VALUE OF THE i -TH ITEM AND k_n BE THE MAXIMUM PROFIT WITH n Kg AVAILABLE.

$$k_n = \max_{1 \leq i \leq n} (v_i + k_{n-i})$$

b) How to obtain the items: Once we build a table like this one,

CAPACITY	0	10	20	30	40	50	60	70	80	90	100
MAX PROFIT	0	10	20	30	40	50	60	70	80	90	100

WE CAN OBTAIN THE ITEMS THAT MAKE UP THE SOLUTION, PERFORMING THE FOLLOWING PROCEDURE:

LET $K[0..W]$ BE OUR TABLE, W THE CAPACITY OF THE BAG, $w[i]$ AND $v[i]$ BE THE WEIGHT AND THE VALUE OF THE i -TH ITEM, RESPECTIVELY, FOR $1 \leq i \leq n$, WHERE n IS THE NUMBER OF AVAILABLE ITEMS. WE DECLARE A VARIABLE c , WHICH WILL REPRESENT THE CURRENT CAPACITY OF THE BAG. THE INITIAL VALUE OF c IS W . WE DECLARE ANOTHER VARIABLE S , THAT WILL CONTAIN THE ITEMS THAT MAKE UP THE SOLUTION. THE PROCEDURE WORKS AS FOLLOWS:

FROM ALL THE ITEMS AVAILABLE, WE MUST FIND ONE THAT SATISFIES THE FOLLOWING CONDITION:

$$c - w[i] \geq 0 \wedge K[c - w[i]] + v[i] = K[c] / \text{i.e., did we use the } i\text{-th item to get } K[c]?$$

WE ADD THE i -TH ITEM TO S , AND UPDATE c , SUBTRACTING $w[i]$ FROM IT.
WE REPEAT THIS PROCEDURE UNTIL c EQUALS 0.

FINAL ANSWER:	CAPACITY	0	10	20	30	40	50	60	70	80	90	100
	MAX PROFIT	0	10	20	30	40	50	60	70	80	90	100
	WE USE	#1 #1 #1 #1 #1 #1 #1 #1 #1 #1 #1 #1 #1										

START FROM 0
SOLUTION: USE ITEM #1 TEN TIMES.

C. Sequence Alignment

Problem 3 (Outcome b) - 4 points

Align the following strings using dynamic programming

String 1		DYNAMIC
<hr/>		
String 2		PROGRAMMING

- 4/
- Write down the table, the recurrence and the procedure to solve the problem using dynamic programming considering the gap penalty as -2 and the reward for a match as +1 and mismatch as -1.
 - Once the table is constructed, explain how can we obtain the alignment between both strings? Write the final alignment.
 - What is the relation between Dynamic Programming and a Directed Acyclic Graph (DAG)? Explain the relation between the nodes and the edges for the sequence alignment problem and plot the DAG for the strings above.

PROBLEM 3

a)

TABLE	
-	P R O G R A M M I N G
-	0 -2 -4 -6 -8 -10 -12 -14 -16 -18 -20 -22
D	-2 -1 -3 -5 -7 -9 -11 -13 -15 -17 -19 -21
Y	-4 -3 -2 -4 -6 -8 -10 -12 -14 -16 -18 -20
N	-6 -5 -4 -3 -5 -7 -9 -11 -13 -15 -17 -19
A	-8 -7 -6 -5 -4 -3 -2 -1 -0 -10 -12 -14 -16
M	-10 -9 -8 -7 -6 -5 -3 -2 -1 -9 -11 -13
C	-12 -11 -10 -9 -8 -7 -6 -5 -4 -3 -2 -1 -0
G	-14 -13 -12 -11 -10 -9 -8 -7 -6 -5 -4 -3 -2

RECURSION

Let X and Y be two sequences.
 Let $a_{i,j}$ be the maximum score for the alignment of $X[1..i]$ and $Y[1..j]$.

$$a_{i,j} = \begin{cases} 0, & i=0 \\ a_{i-1,j}, & i>0, j=0 \\ a_{i-1,j-1}, & i>0 \\ \max(a_{i-1,j} + c, a_{i-2,j-1}, a_{i,j-2}), & \text{if } X[i] = Y[j] \text{ and } i,j > 0 \\ \max(a_{i-2,j-1}, a_{i-1,j-2}, a_{i,j-1}), & \text{if } X[i] \neq Y[j] \text{ and } i,j > 0 \end{cases}$$

b)

Once we get the table a , we can perform the following procedure to get the alignments.

```
Get-Alignments(A, X, Y)
1: i := X.length
2: j := Y.length
3: If A and B have similar sequences
4: while i > 0 and j > 0
5:   if X[i] = Y[j]
6:     i := i - 1
7:     j := j - 1
8:   else if a[i-1,j-2] > a[i,j]
9:     i := i - 1
10:    j := j - 1
11:   else if a[i-2,j-1] > a[i,j]
12:     i := i - 2
13:   else if a[i-1,j-1] > a[i,j]
14:     i := i - 1
15:     j := j - 1
16:   Push '-' to the front of A
17:   Push Y[i] to the front of B
18:   Push X[i] to the front of A
19:   Push '-' to the front of B
20:   i := i - 1
21:   j := j - 1
22: else if a[i-1,j-2] > a[i,j]
23:   Push '-' to the front of A
24:   Push Y[i] to the front of B
25:   Push X[i] to the front of A
26:   Push '-' to the front of B
27:   i := i - 1
28: while i > 0
29:   Push '-' to the front of A
30:   Push X[i] to the front of B
31:   Push '-' to the front of A
32:   Push Y[i] to the front of B
33:   i := i - 1
34: return A and B
```

Procedure

```
Scoring-Algorithm(X, Y)
1: m = X.length
2: n = Y.length
3: let a[0..m, 0..n] be a new table
4: for i = 0 to m
5:   for j = 0 to n
6:     a[0, i] = i * -2
7:     a[i, 0] = j * -2
8:     for i = 1 to m
9:       for j = 1 to n
10:         c = 1 if X[i] = Y[j] else -1
11:         a[i, j] = max(a[i-1, j-1] + c, a[i-1, j] - 1, a[i, j-1] - 1)
12: procedure algoritmo
```

- P R O G R A M M I N G	
-	0 -2 -4 -6 -8 -10 -12 -14 -16 -18 -20 -22
D	-2 -1 -3 -5 -7 -9 -11 -13 -15 -17 -19 -21
Y	-4 -3 -2 -4 -6 -8 -10 -12 -14 -16 -18 -20
N	-6 -5 -4 -3 -5 -7 -9 -11 -13 -15 -17 -19
A	-8 -7 -6 -5 -4 -3 -2 -1 -0 -10 -12 -14 -16
M	-10 -9 -8 -7 -6 -5 -3 -2 -1 -9 -11 -13
C	-12 -11 -10 -9 -8 -7 -6 -5 -4 -3 -2 -1 -0
G	-14 -13 -12 -11 -10 -9 -8 -7 -6 -5 -4 -3 -2

A = - - D Y N A - M I - C
 B = P R O G R A M M I N G

c) SOMETIMES WHEN APPLYING DYNAMIC PROGRAMMING, DETERMINING AN OPTIMAL SOLUTION FOR A SUBPROBLEM INVOLVES CONSIDERING AN OPTIMAL SOLUTION TO ANOTHER SUBPROBLEM.

IT IS SAID THAT DYNAMIC PROGRAMMING IS RELATED TO DAGs, BECAUSE WE CAN REPRESENT THE SUBPROBLEMS AS NODES, AND USE DIRECTED EDGES IF ONE SUBPROBLEM USES THE SOLUTION OF ANOTHER ONE TO BUILD UP ITS SOLUTION. THE GRAPH IS ACYCLIC BECAUSE A SMALLER SUBPROBLEM WILL NEVER USE THE SOLUTION OF A BIGGER SUBPROBLEM.



Problem 4 (Outcomes a,b) - 4 points

Consider the Global Alignment problem and implement the Star Alignment Heuristic to align the 5 DNA sequences that are included in this question. For this question you should attach your C++ code, a .txt file with the final result and an image comparing your alignment.

Instructions

- Implement a C++ algorithm that receives 2 strings and returns its global alignment.
 - Research and implement a C++ algorithm with the Star Alignment Heuristic.
 - Use the Star Alignment Heuristic to align the 5 DNA sequences contained in the *.fasta files
 - Write the final alignment in the .txt file. Plot your alignment replacing each component by a color, generate a picture and attach it as part of your answer. You can use any tool and language just for the visualization part.

11/30/2020

Sequence Manipulation Suite

Color Align Properties results

G, A, V, L, I
F, Y, W
C, M
S, T
K, R, H
D, E
N, O
P

```

XM_025192633.1 PREDI AAAATGGTACAGGATTTCGAAGTGTGAGAAGGATTCTTGAGACTACTGTCACTACAGAAATCCCTCGCAAAGCAAT 160
XM_03101670.1 PREDI C-TTCTCATATCCCAATTGCCGGGATTTCGAGATTTGTCCAGAAAGGTTCTTGAGACTATTGTCACTACAGAAACCC 159
XM_037906217.1 PREDI ATGGATGAC-TCCAACGGCATGTATAAGGGGAAAGTGGCATTTC-CATATCAGAACCCGTCAGGATTAAGTCG 158
XM_002939080.5 PREDI CGGGCGGGCTCGGGGGCATGGGGAGCCGGCCGGCGCTCGGGGCCTCTGTATGCTGTCTGTGTTGAGCCG 160
XM_026853751.1 PREDI TATTAATCTACCCCTGGCGCATCTCTGCCTCTCGACTACGAGCTGTGGCCACCGAACCGCTCTGATAAGCCGACTT 160

```

XM_031616703.1 PRDEI ACCGTAAGACGGTAACTGGATACATTCATAGAAGCTGACTGAACTGCG
XM_031616703.1 PRDI TGTGAACGAGACTGGTCACTGGAGAACATGTCGAAACCTGACTGAGC
XM_037962172.1 PRDI GAAGGCTTCGGGACTGACTGGCACTGAGAACCCCTGTCGAC
XM_037962172.1 PRDI CCGGGAGCCTGGGGCTTACAGTGCCTGATGATGTTTAACTGTC
XM_0293908.0 PRDI GCGGGAGCCTGGGGCTTACAGTGCCTGATGATGAAAGAAATTGCA-TCTCTATCG
XM_0293908.0 PRDI GCGGGAGCCTGGGGCTTACAGTGCCTGATGATGAAAGAAATTGCA-TCTCTATCG

XM_025192633.1	PREDI	GCCAGGAAACCTACGAGTGTCTGTCCTCAGGTTACAGAGAAGGATTGCCAGTGATTGATGCCATCTCAG	400
XM_031061670.1	PREDI	TGGCAACTCTAACCGAACATGCTATAGTGCTGTCAGGAAACCGAGTACACAGGAAAAGAACATGCTGACTGGATTGATGCC	399
XM_037906217.1	PREDI	CTCTGCTAGTCTGCCCGGCCCTGCACGGGGCCTACTGTCACCCCCCGGCCGACACCTGACTGAGCTGCTCCCT	397
XM_002939080.5	PREDI	CCTCTGAATAGTGTGAAACCATCTGAAACCATCTGTCAGGAAAGGATTGATCTGCTGTCAGGAAACCTGAGTCTGATTGACAGGA	398
XM_026853751.1	PREDI	CTGGGAGCATCTTGGACCTACTGGAAATTCTGAT/GAT/GTCAGGCCGGGCCATTGACCTGAGCTTAAGGTC	400

```

XM_025192633.1 PREDI CCTTGTGCCAACGGAAGTACATGTA  
CTCTGGAAATAAGTCCTCGATCTGCTCTGCTGC  
TACACTGGCCAGAA 480
XM_031061670.1 PREDI TTGCATGCTCAGCGCGTGGCA  
ATGGAAACTTCTGATCGCATCTGGAGAG  
TTTTCTGTCATACATGCTCTGTT 478
XM_037906217.1 PREDI GCGCAGGCTCAGCGGAAGTAC  
GGCGTGGCAAGTCAGTGGATGTCG  
CATCCCTTCACATTGCACTT 477
XM_002939080.5 PREDI GAGGACTG-CAGATACTCTGA  
ATCGCACATCTGCTTGTGCT  
CACCCCTGCTTAAATGGAGG  
ACCTGGCAGCTGCTTC 477
XM_026853751.1 PREDI CCTATGGCAAAGCATTCGAGATCGGAT  
TGGGATTCTGGA  
TTTCTAACCCGTTGGCTGCTGGTTCTATCGGG 480

```

XM_021519623.1 PREDI GTGAGATAGATGAAATGAATGGCTTACCATGGGGCTTGTGACATTGTTGGCGATATGTCATGTTGGCTTGTGACAT 556
XM_031061670.1 PREDI ACATCTGGACAAATAAGTGTGAAATGTGTGATGTTGATGTCACCAACAGGAACTGGCAGATGTTGGCAACCTGTCTCAAT 557
XM_037906217.1 PREDI TGTCTGGCAATAAGTCTCTCCGACCTGGCCGGGCTACAGGGCCAGGGTTGAGACGGACGTGAGCTGGCC 557
XM_022939085.5 PREDI TCCAGGACACATATGACTGTTGCTTCCCACGAGTTCACAGGAAAGACTGGCCTGATTGACCTTGTCTGTC 557
XM_026853751.1 PREDI ATCTGGAGAGGGAGATGGGGGCCAGATGTGCTGGGCTGTGTCCTGCTGGGATGCCGGCTGTGACGGCTTAC 556

XM_02519233.1 PREDI ATCGATGCCAGTGGCAACGAGCTATACAGGCCCTGCCTGTAAGCTTATGTCGCTTCTCCCAACCTACCCGGTTATG 640
XM_031061670.1 PREDI GTCTGCCCTATCATCAGTCAGTGCAACGCCAGCTTACCGGCCAACCTGGCTGAAGAACGACTATGTCGCTTGTTC 638
XM_037906217.1 PREDI GCGCCGGGGCTGGCAGCACGGGGCCTGGCTCAAA---CCTGAGCTTCTCTGGCTGC-T-C-A---CAG---A---A 637
XM_02939080.5 PREDI CCATGTCCAATTTGAAACTTGTATCATCTGGAAGCTGGTTCTCTGGCCTGGCAGCAGGCTATATACTGGCCAAA 637
XM_026853751.1 PREDI AGTGTGAGGTGGCAGTGGCAGCTTAAATGAGGGAAATGATCACCAATGATGACGGCTTCGATCACTGCAATG 640

XM_025192633.1 PREDI AATGGAGGAACTTGTCCAGAACAGTGACTTGGGCTTGGAGGCAACTGTCCAGGGTTGGGGACCGCTGTCGTG 720
XM_031061670.1 PREDI TTTCACAGATGTAAGTGGAGGACATGTCTCACAGACTGGTATTGGCTGTATGAATGCACTGTCTGCCAGGGTTGGAG 685
XM_037906217.1 PREDI --G-CA-GCACAAATGC-C-T-T-T-CC-TACAGGGATTCTCA-GGAT-TA-GG-TGAACAGGGCA-AATCTTAAAG 685
XM_002939080.5 PREDI GTGGAGATGGATCTAAATGTCACCAAGGACATCTGGCAAGCCTGGTCAATGCTCAGGCTTGTCACTT 717
XM_026853751.1 PREDI ACAGATGCTTGTGGAAATACTGGCCAACCTGGCAAGAACCATCTGGCTCAATGGCTCAGGCTTGTCACTT 717

XM_025192633.1	PREDI	ACACAACTGTGACGACTGCCAAAT GAGATACCTGTGAGCACCAACAT A-AAAGGAGGAGGAAA-G-CTG-C- ACCATTGCGTGGAGGAGGAGGAAA TGACGCTTCTGACCTCTAACGGTCT	ACACAACTGTGACGACTGCCAAAT GAGATACCTGTGAGCACCAACAT A-AAAGGAGGAGGAAA-G-CTG-C- ACCATTGCGTGGAGGAGGAGGAAA TGACGCTTCTGACCTCTAACGGTCT	ACACAACTGTGACGACTGCCAAAT GAGATACCTGTGAGCACCAACAT A-AAAGGAGGAGGAAA-G-CTG-C- ACCATTGCGTGGAGGAGGAGGAAA TGACGCTTCTGACCTCTAACGGTCT	ACACAACTGTGACGACTGCCAAAT GAGATACCTGTGAGCACCAACAT A-AAAGGAGGAGGAAA-G-CTG-C- ACCATTGCGTGGAGGAGGAGGAAA TGACGCTTCTGACCTCTAACGGTCT	ACACAACTGTGACGACTGCCAAAT GAGATACCTGTGAGCACCAACAT A-AAAGGAGGAGGAAA-G-CTG-C- ACCATTGCGTGGAGGAGGAGGAAA TGACGCTTCTGACCTCTAACGGTCT
XM_031061670.1	PREDI					
XM_037906217.1	PREDI					
XM_02993080.5	PREDI					
XM_026853751.1	PREDI					

XM_025192633.1 PREDI GCTGCCACCGCAATGGACAGGCCAGTTTGACTGAAGACGTTGATGAATGTCAGGCTCAGGCCAACGCTTGTTCAGAAAT 880

XM_031061670.1 PREDI ACTTACAACGTGCCGCTGCCCTCACAGTGGACAGGTCAAGTCAGTTCTGTACTGAGGATGTGGATGAATGTCAGCTGCAGCCAA 878
 XM_037906217.1 PREDI CAAAGTGCCTGAACGGGGCGTGTGCGTGACGGGGTAACACCTACAACACTGCCGCTGCCGCCCAATGGCTGGCCAGT 837
 XM_002939080.5 PREDI AATGGAGGGACATGTCAATCAAC-CAATGACTTACCTATGAGTGTAACTGTCTGCCAGGTTGAAGGAAG-CACCTGC 875
 XM_026853751.1 PREDI AATGTCTGCTCCCTACTCGCCTGTATGAACGGGGCACCTGCCACCTGCTTACTCTGGAACTC-TATGACTGTATCT 878

XM_025192633.1 PREDI GGAGGTACCTGCACCAACCACAAATGGAGGCTATGCCCTGTGCTGTGCAATGGTTGGAGCGGGGACGACTGCAGTAAGAA 960
 XM_031061670.1 PREDI TGCCTGCTAAAATGGAGGTACCTGCACCAACCACAAATGGCGCTATGCCCTGTGCTGTCAACGGCTGGAGTGGGATG 958
 XM_037906217.1 PREDI TTTGCACCGAGGGATGTGGATGAGTGCAGCTCCAGCCAAATGCCAGAACGGGGCACCTGCCACCAACCAAACGGG 917
 XM_002939080.5 PREDI GAGCAGAAATTGTGACTGCCAAATCATATTGCCAGAATGGGGATTGTGGATGGTGAACACTTACAATTG 955
 XM_026853751.1 PREDI GCCCCTTGGATGGAGAGGAATGATTGCAAATTGATGCCCTCTCAGCCATGCCAACGGCAGTACCTGC 958

XM_025192633.1 PREDI TATTGATGACTGCTCACTGCCCTGTGCTATGGATCTACTTGCAATTGATCGAGTGGCTTCTTTTCATGCATTGTC 1040
 XM_031061670.1 PREDI ATTGCAAGAATATTGATGACTGTCTT-ACGGCTTCTGTGCCAATGGATCTACTTGCAATTGATCGTGTGGCTTCATT 1037
 XM_037906217.1 PREDI GGCTACGCCCTGCGCTGCGTCAACGGCTGGAGCGGGGACGACTGCACCAAGAACATCGACGACTG-CTTCACTGCCCTCCT 996
 XM_002939080.5 PREDI CCACTGCCCTCCCACAGTGGACAGGTCAAGTTGTACTGAGGATGTGGATGAATGTCAGCTCCAGCCAAACGCCGTG 1035
 XM_026853751.1 PREDI ATCCCCAATGAGAA-CCAGTATACTGTTCTGTCAAGGCTGGATACACCGCCTCAAGTGTGAAGTGGATGTGAACGAGT 1037

XM_025192633.1 PREDI CAGAAGGAAAGGCAGGTCTCTGCCACTTGGATGATGCATGTGTTAGCAATCCATGCCAGAAGGGTGTCTGTGAT 1120
 XM_031061670.1 PREDI TGCATGCATTGTCAGAAGGAAAGGCAGGCCCTGTGTCATCTGGATGATGCATGTGTTAGCAACCCCTGCCAGAAGG 1117
 XM_037906217.1 PREDI GCGCCAACGGCTCACCTGCACTGCCACAGGGGGCTCCCTCTCTGCAATTGTCAGAGGGGAAGGCAGGTCTCTGTG 1076
 XM_002939080.5 PREDI ACGGAGGGACCTGCAACCAACCACATGGAGGCCACAGCTGTGTCATGTCAATGGCTGGAGCGGGGAGGACTGCAGTAAG 1115
 XM_026853751.1 PREDI GTGGCCGCCAGGGATGTGCAAAATGGAGGAACATTGTGAGAATTTCCTGGTCTACAGGTGTCAGTGTACCCGGGT 1117

XM_025192633.1 PREDI ACCAATCCTGTGAATGGACACTATATCTGTACGTGTCCTCAAGGCCATAAGGGAGCAGACTGCACTGAGGATGTGGACGA 1200
 XM_031061670.1 PREDI AAGCTCTATGTGATACCAACCTATGAATGGACACTATATCTGCACTTGTGCCCCAAGGCTACAAAGGAGCTGACTGTACT 1197
 XM_037906217.1 PREDI CACTTGGATGATGCCCTGTGTTAGCAATCCGTGCCAGAAAGGTGCTCTGTGACACCAACCCCTGTGAATGGAAACTACAT 1156
 XM_002939080.5 PREDI AACATTGACGACTG-TTTCAACTACTCCCTGTGCTGCCGGTCCACCTGCAATTGATCGTGTGGCTCATTTACATGTATCT 1194
 XM_026853751.1 PREDI TACACTGGTCAGCGCTGTGAGACCCCTTCATCCCCCTGCTCTCCCCATGTATGAATGGTGGCACCTGCCGTAGAC 1197

XM_025192633.1 PREDI GTGTGCAATGGCAAACAGCAATCCATGTGAACATGCTGGAAATGCGTAAACACAGAAGGGCTTTCCACTGTGAGTGT 1280
 XM_031061670.1 PREDI GAAGATGTAGATGAGTGTGCTATGGCAAACAGTAATCCGTGTAACATGCTGGAAATGTGTAATACTGAAGGCTCCTT 1277
 XM_037906217.1 PREDI CTGCACCTGTCCCCAGGGCACAAGGGAGCAGACTGCACTGAGGATGTGGATGAATGTCATGGCAAACAGCAATCCGT 1236
 XM_002939080.5 PREDI GTCCAGAAGGAAAGACAGGTCTCTGTGCCACTTGGATGATGCAATGTGTTAGCAACCATGCTGATGGAGCTGTG 1274
 XM_026853751.1 PREDI -TAGTGAATTGCTATGAATGCAACTGCCCTGCCAGGATTGATGGGA-AAAATGTGAAACCAACATAGATGATTGTCC 1275

XM_025192633.1 PREDI TGAAGGGCTACACAGGACCTCGCTGTGAAATGGACATCAATGAATGCCATTCAAATCCATGCCAAACGATGCCACCTGC 1360
 XM_031061670.1 PREDI CCACTGTGAGTGTGCAAGGGATACACAGGACCAAGCCTGTGAAATGGACATTAATGAGTGTCAATCCAATCCATGTCAAA 1357
 XM_037906217.1 PREDI GTGAGCACGCTGGGAAATGCGTCAACACCGAGGGCTCTTCCACTGTGAGTGCCTGAAAGGGCTACACAGGGCTCGCTGT 1316
 XM_002939080.5 PREDI GATACCAACCCCTGCAATGGACACTACATCTGACCTGTGCCCCAAGGCTATAAAGGAGCAGACTGTACTGATGACGTG 1354
 XM_026853751.1 PREDI CAGTCATAAGTGCCAGAATGGTGGAACCTGTTGATGGGATGGGAACACCTACAACACTGCAGGTGCTCTCCC-CAGTGGACT 1354

XM_025192633.1 PREDI CTGGATAAAATTGGGGATTCACTTGTGCTGTGTTAGGCCAGGTTTAAAGGTTCACTGTGAGAAGATATTGATGAGTG 1440
 XM_031061670.1 PREDI ATGATGCCACTTGCCTGGATAAAATGGAGGATTCACTGTGTTAGGCCAGGTTAAAGGCATGCACTGTGAACTG 1437
 XM_037906217.1 PREDI GAGATGGACATCAATGAGTGCACCTCAACCCGTGCCAGAACGATGCCACCTGCTGGACAAAATTGGAGGTTACGTG 1396
 XM_002939080.5 PREDI TGAGTGTGCTATGGAAACAGCAACCCATGTGAGCACGCCGGGAAATGTGTAATACAGAAGGCTCTTCACTGTGAAT 1434
 XM_026853751.1 PREDI GGCCAGTTTGCACAGAGGATGTTGATGAGTGCACCTGCCAAATGCTTGCCTTGGACTGTGAC 1434

XM_025192633.1 PREDI TCTAAGTAATCCCTGTGTAACATGGAGAGTGTCTTGTGATAAGTCACCGCTTCTCTGTGTCCTGCTCTGGATTCA 1520
 XM_031061670.1 PREDI GAGATCAATGAATGCCCTAGTAGTCCCTGTGTAACATGGAGAGTGTCTGGATAAGGTCAATCACTTTTGTC 1517
 XM_037906217.1 PREDI TCTCTGCATGCCAGGTTTAAAGGAGTTCACTGTGAGAAGAAGATATTGATGAAATGCCCTGTGTAACATG 1476
 XM_002939080.5 PREDI GTCTGAAGGGGTACACAGGGCTCGTGTGAACTGGACATTAATGAAATGTCTTCCGACCATGTCAAGATG 1514
 XM_026853751.1 PREDI CAATGGAGGCTACAACACTGCCTGTGTAATGGATGGAGTGGGATGATTGAGGCCAGAATATCGATGACTG-TGCCACT 1513

XM_025192633.1 PREDI GTGGTCCGTATGTCAGATTGATAGATGACTGCTCTAGCACACCGCTCTCAACGGAGCAAATGTATTGACCATCCC 1600
 XM_031061670.1 PREDI CCCACCTGGTTTACTGGAGGCCGTGCCAGATTGACATCGATGACTGCTCCAGCACACCTTGCTAAATGGAGCAAAGT 1597
 XM_037906217.1 PREDI GAGTGTGCTTGATAAAAGTCACCGCTCTCTGTGCTGCTGCTCTGGGTTAGTGGCCAGATGACATA 1556
 XM_002939080.5 PREDI TGCCTGGATAAAATGGAGGTTACACAGGGCTCGTGTGAACTGGACATTAATGAAATGTCTTCCGACCATGTCAAGATG 1594
 XM_026853751.1 PREDI GCTGCCCTGTGCCAAATGGGCTACTGTGATTGATCGAGTTGCTCTTATGTGTTGCCAGAGGAAAGATTGACT 1593

XM_025192633.1 PREDI AATGGCTACGAGTGCCAGTGTGCCACAGGTTTACTGGCCTTCTGTGTTGAGGAAATATCAACAACTGTGATCTGT 1680
 XM_031061670.1 PREDI GCATTGATCATCCAATGGCTACAGAGTGCCTGCAAGGTTTACTGGCATCTTGATGAGGAAATATAATAAT 1677
 XM_037906217.1 PREDI GACGACTGTTCCAGCACGCCGTGCCACAGGAGCCAAGTGCATTGACCCACCCCAATGGCTACAGTGCAGTGTGCCAC 1636
 XM_002939080.5 PREDI ATGCCAGAGTAATCCCTGTGTAACATGGAGAATGTCTGGATAAAACTCAATGCTCCAGTGTATCTGCCACCTGGTT 1674
 XM_026853751.1 PREDI CCTGTGTCACAAAATGATGCCCTGTTAGTAATCCTGCCACATGGCTTTGTGATACCAACCCCTGTACTGGAC 1673

XM_025192633.1 PREDI TTGCCACCATGGGAATGTCAAGATGGAATTGATTGATACATACGTGTTATGCAATCCTGGTACATGGGTGCCATATGCA 1760
 XM_031061670.1 PREDI TGTGATCCTGCCACTATGGTGAATGCCAAGATGGTATTGATTGATACACATGTGTCGCAATCCTGGATACAT 1757
 XM_037906217.1 PREDI AGGTTTACTGGCACTC-TGTGTGAGGAGAACATCAACAACTGTGACCCGTGCCACACGGGGAAATGCCAGGAT 1715
 XM_002939080.5 PREDI TCACTGGAGCTGTGCGCAGATGGACATTGATGATTGCTCCAGCACACCTGTCTAAATGGAGCCAAGTGCATTGATT 1754
 XM_026853751.1 PREDI AAGCCATTGCACTGCCCTCTGGTACAAAGGACTGCCACAGAGGACATAGATGAATGCTTTGGTTAACAGC 1753

XM_025192633.1 PREDI GTGAGCAGATAGATGAATGCCACAGCAATCCTGCCCATCAAGGGCGTGCATTGATTGGTGAATGGCTACCAGTGC 1840

XM_031061670.1 PREDI GGGTGCCATTGTAGTGAACAGATTGATGAATGTTACAGCAATCCCTGCCCTAACAGGACGCTGCATTGACTGGTGA 1837
 XM_037906217.1 PREDI GGCATAGATTACATCACACGTGCATCTGCACACCCTGGCTACATGGGTGGCATATGCAGTGAGCAGATTGATGAGTGCACAG 1795
 XM_002939080.5 PREDI CCAAATGGCTATGAATGCCAGTGTGCCACAGGTTTACTGGCACT-TTGTGTGAAGAGAATATCAACAATGTGACCCCTG 1833
 XM_026853751.1 PREDI AATCCCTGTGAGCATGCAGGAAATGTGTCATAACTGAAGGTTCCCTTATTGTGAATGCTC-CAAAGGATTACCGGGC 1832

XM_025192633.1 PREDI AACTGTTGCTGGGCACCTCAGGAGTGAAGTGTGAAAATAACTTGTGACTGTGCAAGTAACCCATGCATTACGGGG 1920
 XM_031061670.1 PREDI ATGGTTACCAATGCAACTGTTACAGGCACATCAGGAATGAACTGTGAGAATAATTGTGATGACTGTGCAAATAACCC 1917
 XM_037906217.1 PREDI CAATCCCTGCCCTCAACCAAGGGCGCTGCATCGACCTGGTCAATGGCTACCGAGTGCACACTGCCCTGGCACCTCAGGAA 1875
 XM_002939080.5 PREDI ACCCTTGCCACCATGGTGAATGCCAGGATGAAATTGATTGTCACACTTGTGTGCAATTCTGGTACATGGGTGCCATA 1913
 XM_026853751.1 PREDI CCCGCTGTGAAATGGACATTAATGAGTGCCTATCTGAGCCATGCCAGAATGATGCCACCTGTCTGGATAAAATCGGAGGC 1912

XM_025192633.1 PREDI CTGTATTGATGGCATTAAATCGCTACAATTGTCCTGCAAACCTGGATTACAGGCCACGGTGTAAATGTGGACATTGATG 2000
 XM_031061670.1 PREDI TCGGTTCATGGGACCTGTATAGATGGCATTAAACCACTACAGTTGCATCTGCAACCCAGGAGTACCGGTCCAAGATGTAA 1997
 XM_037906217.1 PREDI TAAACTGTGAGAATAACTTGACGACTGTGCGAGTAACCCGTGCATTGATGGGACCTGTATCGATGGCATCACTGC 1955
 XM_002939080.5 PREDI TGCACTGTGAGCAGGTAGATGCCAGAGTAACCCCTGCCCTCAATGAAGGACGCTGCATTGACTGGTGAATGGTATCA 1993
 XM_026853751.1 PREDI TTCACATGTTGTGATGCCAGGTTACAAGGAAATACACTGTGAAATTGAGATCAATGAATGCCCTAAGCAACCCGTGT 1992

XM_025192633.1 PREDI AATGCACATCCAGCCCCGTCTATAATGGTGAACATGTATAATGAAGTCATGGTTTCGGTGTGTGTCCTGAAGGC 2080
 XM_031061670.1 PREDI TACTGATATTGATGAATGTATCTCCAGTCCCTGTCACATGGTGAACATGTGTCATGAAATTAAATGGATTCGGTGT 2077
 XM_037906217.1 PREDI AACTGTGCTGTGAACCTGGATTACAGGGCCCGGTGCAACGTGGACATGACGAGTGCACGTCAGCCGTGCCACAA 2035
 XM_002939080.5 PREDI GTGCAACTGTCTCTGGGACTTCAGGAGTGAAGTGTGAGAATAACTTGATGATTGTGCAAGTAACCCCTGATTTATG 2073
 XM_026853751.1 PREDI GAACAATGGAGAGTGTGTTGACAAAGTGAATCGTTGTGCACTGGGATTTCACTGGATCAGTGTGCCAGA 2072

XM_025192633.1 PREDI TACCATCACCCCTACTGTCAGTCTCAGGAGATGGCTGTTAGTAATCCCTGTGACATGCCAACTGCACACATATTAC 2160
 XM_031061670.1 PREDI TATGTCCTGAAGGCTACCATCACCC-CACCTGTCACTCACAGACAGATGAATGTTAATAACCCCTGTGTCATGGCAA 2156
 XM_037906217.1 PREDI CGGGGGCACGTGCATCAACGAGGTGAACGGGTTCCGCTGCGTGTGCCCCGAGGGCTCCACCACCGCA-CTGCCAGCG 2114
 XM_002939080.5 PREDI GAGTCTGTATCGATGGCATTAACTGCTACGATTGTCGACAGCTGGAGTTACAGGTCCAAGGGTGAATATTGATATT 2153
 XM_026853751.1 PREDI TTGATATAGACGACTGCTTAGCAACTTCATGTCAGAATGGTGCATGGTTGATCATCCATCTGGCTATGACTGTATA 2152

XM_025192633.1 PREDI CAGTGGGTACAAGTGTCTGTGATCCAGGCTGGTAGGAGATTACTGTTCAACAGAAGGGAAACGAATGCCAAATCAAAC 2240
 XM_031061670.1 PREDI CTGCACACATGACGTCAACGGGTACGTGTCCTGTGACCCAGGTTGGACCGGAACTGTGAGTAGAGGAAATG 2236
 XM_037906217.1 PREDI CAGGCCGACGGCTGCCCTCAGCAGCCCCGTGTCGACGGCAACTGCACGCCACACGCCACGGGGTACAAGTGTCTGT 2194
 XM_002939080.5 PREDI GACGAATGCATCTCCAGTCCCTGTCATAATGGTGAACGTGTGAAATGAAATGAATGGATTCCGGTGTCTGTCTG 2233
 XM_026853751.1 PREDI TGTGCAACAGGCTTCACAGG-TGTTTGTGCGAAGAAAATATCAACAATGTGAACCACCAACCC-TGTCATTATGGAACA 2230

XM_025192633.1 PREDI CATGCCAGAACGGAGGAACCTGTGAGGATCTGTTGAATGGATATAGATGTACCTGTAGGAAAGGATTCAAAGGTGTGAAC 2320
 XM_031061670.1 PREDI AGTGTAGGTCAAATCCGTGCCAGATGGAGGAACCTGTGACAATCTGTAATGGATATAGATGCACCTGCAAGGAGAGGA 2316
 XM_037906217.1 PREDI CCCAGGCTGGATAGGAGATTACTGTTCAACAGAAGGGAAATGAATGTAATCAAACCCATGCCAGAATGGAGGAACATGTG 2274
 XM_002939080.5 PREDI AGGCTTCCATCAACC-CAGCTGTCACTCACAGCTGATGAACTGTCTTAGTAACCCCTGTGACATGGAAACTGCCACAAAT 2312
 XM_026853751.1 PREDI TGCCAAGATGGCATTGATTCTTACGTGCACTGGCTGACAGGCTGGCTATATGGGTGCCATTGCAAGCAGCAGATCGATG 2310

XM_025192633.1 PREDI TGCCAGGTAGTGGTTGCTCCCTGATCCTTGTGAAAACCTGGGCAATTGCCAAATCTCTGACTCTGACTGAGG 2400
 XM_031061670.1 PREDI TTCAAAGGTGTGAACTGCCGGATAGTGCAGGCTCCTGTTCCCTGATCCTTGTGAGAATTCTGGAATCTGCCAAGAATC 2396
 XM_037906217.1 PREDI AGGATCTGCTGAATGCATACAGATGTACCTGTAGGAAGGGATTCAAAGGTGTGAACTGCCAGGTAGTGGTTCTCTTGT 2354
 XM_002939080.5 PREDI GATGTCAGTGGATACATATGTCTCTGTGACATTGGTTGGTGGACAGACTGTGAGTAGAAGAGAACGATGTCTGT 2392
 XM_026853751.1 PREDI ATGCCCTGAGTAACCCCTGTCTCAATGAAGGCCGTGTTGACCTTGTCAATGGCTATCAGTGAACGGCTACCT-GGA 2389

XM_025192633.1 PREDI CTACACCTGTCAGTGTGCCCTGGCTGGGAAGGAGAGAGATGTACAGTAGATATTGATGAATGTCCTCAAAGCCCTGCA 2480
 XM_031061670.1 PREDI ATCTGGCTCTGATGGCTATATTGTCTCTGTGTCCTGGCTGGGAAGGAGAGAGATGTACGGTGATATTGATGAGTGT 2476
 XM_037906217.1 PREDI TCCCCTAACTCTTGTGAGAACTCGGGAATCTGCCAAGAATCTCTGACTCCGAAAGGCTACACCTGCCAGTGTGCCCTGG 2434
 XM_002939080.5 PREDI AAACCCATGCCAGAATGGAGGAAATTGTGACAGTCTGGTGAATGGATACAGATGTACCTGCAGTAAAGGATTCAAAGG 2472
 XM_026853751.1 PREDI ACATCTGG-GCCAAGATGTGAGAAGAACATGGTGTGACTGTGCAAGCAACCCGTGCACCTATGGGAACATGTGTTGG 2468

XM_025192633.1 PREDI AAAATCATGCTCTGTGCCATAATATTCAAGGCAGTTACCTGTGCGAATGCGCCAGGCTTCAGTGGAGGAGACTGTGAC 2560
 XM_031061670.1 PREDI TGTCAAAGCCCTGCAAGGAATCTGCTGTGTCCTGCCACAAACATTCAAGGCAGCTACCTGTGTAATGCCAAACCCAGGCTTC 2556
 XM_037906217.1 PREDI CTGGGAAGGGAAAGATGTCACAGCTGGATATTGATGAGTGTCTCTCAAGGCCCTGCAAGAACCATGCCCTGTGCCACA 2514
 XM_002939080.5 PREDI TGAACCTGCCAGCTCGTGTCTCCCTGCTCTCCCTGATCCATGTGAGAACACTGTGGAATCTGCCAGAATCTCCAGAATC 2552
 XM_026853751.1 PREDI TAAACCGCTACGACTGTGTCGCCCCAGGATTACAGGCCCTCAGTGTAAAGATGACATTGATGAGTGTGCCCTAAC 2548

XM_025192633.1 PREDI AGTAACATTGATGACTGCCCTTCAAATCCCTGCCAAAACGGAGCTTCAGTGTGAGGATGAAACTCTTCTCATGCA 2640
 XM_031061670.1 PREDI GGGGGAGACTGTGAAAGCAACATTGATGACTGCCCTTCAAATCCCTGCCAGAATGGAGCTCATGTGTTGGAGGGGTGAA 2636
 XM_037906217.1 PREDI TCCAGGGCAGTTACCTGTGCGAGTGTGCTGGCTGGCTGGAGGTGAGAGATGTACAGTGGATATAGATGAGTGTGTC 2594
 XM_002939080.5 PREDI GAGGGCTACACTTGGCCAGTGTGCTGGCTGGAGGTGAGAGATGTACAGTGGATATAGATGAGTGTGTC 2632
 XM_026853751.1 PREDI CCCTGCCATAATGGAGGAAGTGTGCTGAAATGAACAGAACAGATTTCGCTGATCTGTCGGAGGGATTGCCAGCC-CA 2627

XM_025192633.1 PREDI CTGCCTACCTGGATTTCATGGAGATAAAATGTCACACAGATACCAATGAATGTCAGTGAACCATGCAGGAATGGAGG 2720
 XM_031061670.1 PREDI CTCTTCTTATGCATCTGTCTGCCCTGGTTCACAGGGAAACAGTGTGAGCAGACATACCAACAGAGTGTATGAATGAC 2716
 XM_037906217.1 PREDI AATCCCTGCCAGAACGGGGCCTGTCGTGTTGGATAAACTCCTCTGTCATGCCCTGCCGGATTTCACGGGG 2674
 XM_002939080.5 PREDI TTGCAAGAATCATGCTGTGCTGCAACTCAAGGCAGTTACTGTGAGTGTGCAATGCCGCCAGGCTTACTGGAGGAGACT 2712
 XM_026853751.1 PREDI GCTGTTTCGAGGTGGACATGTGTCATAACCCATGATGACATGCCACAGATGACGTCAATGGGTACAAG 2707

XM_025192633.1 PREDI CGTGCACCCACTATGTCAACAGCTACACATGCAAGTGTCAAGCTGCCCTGGTTGAGGGAAACCAACTGCGAGAACACATCGAT 2800

XM_031061670.1 PREDI GCAAGAATGGAGGCACCTGCTCTGACTATGTCACACAGCTACACGTCAAGTGCCAG-CAGGATTGAGGGGACCAATTG 2795
 XM_037906217.1 PREDI CAAAGTGTCAAGACAGACACCAACGAGTGCCTGAGTGAGGCCCTGCAGGAACGGGGCACCTGCACCCACTACGTCAACAGCT 2754
 XM_002939080.5 PREDI GTGATAGCAACATTAATGACTGCCTTCAATCCCTGCCAGAACATGGAGCTCGTGTGGATGGGTGAACCTTTCTTA 2792
 XM_026853751.1 PREDI TGTCTCTGCCAACCGAGCTGGACAGGCAAGA-AATGCGAGGTGGATAAAAACGAGTGTCTTCCAACCCCTGTCAGAATG 2786

XM_025192633.1 PREDI GAATGCACTGAGAGCCTGTTCAATGGAGGTACTTGTGTGGATGGATCAATTCTTCACTTGCCAGTGTCCAGTGGG 2880
 XM_031061670.1 PREDI TGAGAACACATCAATGAATGCCACTGAGACTTCGTGTTCAATGGAGGCACTTGTGTGGATGGATCAACTCCTTACAT 2875
 XM_037906217.1 PREDI ACACCTGCAGGTGCC-GCCGGCTTCAGGGCACCAACTGCGAGAGAACATCGACGAGTGACGGACAGCTCTGTT 2833
 XM_002939080.5 PREDI TGCACCTGTCTGCCCTGGGTTCAAGGGGATAAAGTGTCAAGAGACACATAATGAATGTCTGAGTGAAACCATGCAAGAATGG 2872
 XM_026853751.1 PREDI GGGGACATGCGATAACCTGGATACAGATGCTCTGCAAGAAGGGTTACAACGTGAGATTGAC 2866

XM_025192633.1 PREDI ATTTACAGGACCGTCTGCCACAGAAATTAAACGAGTGCATTACACATCCTGCCAGAACAAAGGCTTGTGCTGGACA 2960
 XM_031061670.1 PREDI GCCAATGCCGATTGGG-TTCACAGGAAACCTTTGTCTTATGGAAATTAAATGAATGTGATTCAACATCCTGCTGAA 2954
 XM_037906217.1 PREDI CAATGGAGGCACTTGTTGGATGGAATCAATTCCCTCACCTGCCAGTCCC-CCTGGCTTCACGGGCCCTCTGCTG 2912
 XM_002939080.5 PREDI AGGCACCTGCACGGACTACGTCAACAGCTACACCTGCAAGTGTCAAGGAGGG-TTGAGGGGACCACTGTGAGAACA 2951
 XM_026853751.1 PREDI ATCGATGAATGTGCTTCAATCCATGTTAAACCATGGGACGTGCAAGAGATGGGATCAATGGTACACGTGCCACTGCGG 2946

XM_025192633.1 PREDI GCTTAGGCACATACCGATGTATATGCTTTGGTTATACTGGGAAGAACTGTCAGGCACCTATGGATCTGTGAGCAAG 3040
 XM_031061670.1 PREDI AGGGACTTGTGTGGATAGCCTGGGACCTATCATTGTGTTTGTCCATGGTTACACTGGAAAGAACATGTCAGATACTTG 3034
 XM_037906217.1 PREDI ACAGAAATCAACGAGGTGTGACTCCCACCCCTGCCCTAACAGGGGACCTGTGTGGACAGCCTGGCACATA-CAGGTGA 2991
 XM_002939080.5 PREDI ATTGATGAGTGTACCGAGAGTTCTGTTCAATGGAGGCACTTGTGTGGATGGATCAACTCCTTCCAGCAGTGC 3031
 XM_026853751.1 PREDI GCTGCCGTATACTGGTACAAACTGCCAGAACATTCTGACCCCCCTGTAGCCCTGAACCGTGCCACAATGGTGAAG 3026

XM_025192633.1 PREDI TCTCCCTGTAAGAATAAGGCACGTGCCAAAATGGAGCCCAGACTCGATGTATCTGTCATCTGGCTGACTGGTGC 3120
 XM_031061670.1 PREDI TGGATCTGTGTAACAAGTTCCCTGTAAGAACAAAGGCACATGTGTCAGCA-GAATACTCACTTCATTGCAACTGTG 3113
 XM_037906217.1 PREDI TCTGTCCTGGCTACACTGGCAAGAACACTGCAAGACCCCTGGTGGATGTGCAAGTCTCCCTGCAAGAACAGGG 3071
 XM_002939080.5 PREDI -GGTGGGGTTACTGGGCTTCTGCTGTGGATGGAAATTAAATGAATGTGATTCCCATTCTGTTGAACAAAGGGACTGT 3110
 XM_026853751.1 PREDI AGGAATCTGAAGATTCAAGAGTTCAAGACTTGTGCAACTGGCTGCCAGGGTGTACTATTGATATTGAT 3106

XM_025192633.1 PREDI TTACTGCGATGTGCCAATGTCTCTGTCAGGTGGCAGCTTCACAACGGGAGTCACAGTGGACAGTTGTGCCAGCACT 3200
 XM_031061670.1 PREDI CCCTGGTTGGAGTGGTGCCTACTGTGATGTGCCAATGTCTCTGCCAACGGGCTGCTCACAGAGGGGGTCCCTGT 3193
 XM_037906217.1 PREDI ACGTGCGTGCAG-ACCCCTGGCCAGACGCGCTGCGTGTGCCGCCAGCTGGACCGGCCACTGCGACGTGCCAGCG 3150
 XM_002939080.5 PREDI GTGGATAGCCTGGGACCTATCA-GTGCATCTGTCCTGGGTTACACTGGAAAGACTGCCAACACTCAAGGATCTGT 3189
 XM_026853751.1 PREDI GAATGTATAACCAACCCCTGCAAAACTATGGTCGCTGCCAAACATACCAAGGGGATACATTGCCATTGCTTGAAGG 3186

XM_025192633.1 PREDI CTGGTCATTGCCCAATGTTGAAACACGCACCACTGTCAGTGCCAGGTGGTACACGGGAGTTACTGCGAGGTGCAG 3280
 XM_031061670.1 PREDI ACCAGTGTGCAACACTCTGCCAATTGCTTAACGTTGAAACAGGCCACATTGCAATGCCGACAGGCTACACTGGT 3273
 XM_037906217.1 PREDI TCTCGTGCCTGGCAGCTTCACAGAGAGGAGTCCCCGTGGACAGCCTGTGCGCAGCTCTGTCACTGCCCACTG 3230
 XM_002939080.5 PREDI GCAACAAGTCTCCCTGTAAGAACAAAGGCACATGCATCCAG-AGGGAAACTCAGTCTCAATGCATCTGTCATCTGGT 3268
 XM_026853751.1 PREDI CTTTAATGGAGTCACACTGTGAAAACAACATTGACGATTGCCCTTCTAATCCGTGTCAATGGGCTTCTGTAGATG 3266

XM_025192633.1 PREDI CTGGATGAGTGTGACTCCAGCCCTGCCAGAACGGTGCTACCTGCCGGGATCACCTGGGTTACACGGGAGTTACTGCGAGGTGCAG 3360
 XM_031061670.1 PREDI AGCTATTGTGAAAGAGCAGGTGGATGAATGCATCTCCAGTCTTGTCAAGATGGGCCACCTGCAAGGGACTATCTGGAGG 3353
 XM_037906217.1 PREDI GGGAACAGCCACCACTGCCAGTCCAGGTGGCTACACGGGAGCTACTGGAGGTGCAGCTGGATGAGTGCAGTCAG 3310
 XM_002939080.5 PREDI GACTGGCGCTATTGTGACGTGCCAACGTCCTGTCAAGGTGGCTGCTTCACAGAGGGGAGTCCCCTAGACCAATTGT 3348
 XM_026853751.1 PREDI GAATCAACACTTCTCTGTAACTGCCCTGGGGGTTCCATGGGACAAGTGCAGAAAGATGTGGATGAATGTGCAAGC 3346

XM_025192633.1 PREDI GCCTGGATACCAGGGTGTCACTGTGAATATGAAGTGGATGAATGCCAGTTCAGCCCTGCCAGATGGAGGAACATGTA 3440
 XM_031061670.1 PREDI ATA-TAGCTGTGAGTGTGTGCCAGGGTACCAAGGGTGTCAACTGCGAGTATGAAGTGAATGAGTGTCAAGTTCAGC 3432
 XM_037906217.1 PREDI CCCCTGCCAGAACGGGACACCTGCAAGGGACACCTGGCGGATACCA-CTGCCAGTGCCTGGCTGGATACCAAGGGG 3389
 XM_002939080.5 PREDI GTCAGCATTCCGG-CATTGCTTAATGTGGAAACAGTCATGCTGCCAGTGCAGTCAGTCTCAACTGGCAGTACTG 3427
 XM_026853751.1 PREDI AACCTGTAAAAATGGAGGCTCTGCACAGACTACGTCAACAGCTACACGTCAAGTGTCA-GCCGGGATTGAA 3425

XM_025192633.1 PREDI TAGACCTCGTAATCTAATTCAAGTCTCTGTCACAGGAAACTCGGGCGTTCTGTGAAGAGAAATGTGGATGACTGT 3520
 XM_031061670.1 PREDI CAAAATGGAGGAACCTGCAATTGATCTCATCAATCTTAAATGCTCTGTCACCCAGGAAACTCGGGGCTTCTGTGA 3512
 XM_037906217.1 PREDI AACTGCGAGTACAGGGTGGATGAATGCCAGTCCAGGCCAGAACGGAGGGACCTGCATCGACCTGGTCAATCACTT 3469
 XM_002939080.5 PREDI TGAAGAGCAGTTGGACGAATGCACCTCAATCCATGCCAGAACGGGACCTGCAAGTGCAGTACTATCTGGGGATATC-GG 3506
 XM_026853751.1 PREDI TCCACTGTGATAAACACATTGATGAATGCACAGACACGTCTGCTCAATGGTGTACCTGCGTGGATGGCATTATTCC 3505

XM_025192633.1 PREDI GTTTCAGACTCAGGAGTACCCCGCTGTTAATGGAGGGCAGTGCATTGACCAAGATGGGGCTACAGCTGTATTGTCT 3600
 XM_031061670.1 PREDI GGAGAACATTGATGACTGCATGCCAGATTCAAGGAGGGCTCGCTGCCCTTAAATGGAGGACAGTGCATAGACCAATTGGGG 3592
 XM_037906217.1 PREDI CAAGTGTCTGCCGCCAGAACCCGCGCCGGCTGTGTGAGGAGAACGTGGATGACTGCATCTCAGACTCAGGGC 3549
 XM_002939080.5 PREDI TGTGAGTGTGCGCCAGGATACCAAGGGAGTCACGTGAGTATGAAGTGGATGAGTGCCTGCCAGTCCAGCCGCCAGAATGG 3586
 XM_026853751.1 PREDI TTCACCTGCCAGTGCCTCAGGGACAATTGCTGTGTTGAGATCAATGAGTGTGGTCCAGGGGCCACCCCTGCA 3584

XM_025192633.1 PREDI CCCAGGCTTGCAAGCGAGGAGATATCAACGAATGCCCTCTAATCCTGTAACCCCTGTTGGAAAGTCTGG 3680
 XM_031061670.1 PREDI GCCACAGCTGTCTTGTCTTCAGGCTCGGAGGCAAGCAGTGTGAGGGGACATTAAATGAATGCCCTCTCCACCTTGT 3672
 XM_037906217.1 PREDI CGCGCTGCTTCAACGGGGGGCAGTCATGACCCAGATGGGGGTTCACTGCGCTGTG-CCCCGAGGCTTCCGGGG 3628
 XM_002939080.5 PREDI AGGAACCTGCATTGATCTTGTCAATCATTTCAATGTCAGTGTGCTCTGTGAGGAGAACACTGGGCCACCCCTG 3666
 XM_026853751.1 PREDI TTAATGGAGGCACTGTGTGGATGGAATGGGAAACATATCA-CTGCACTGTGCGGATACAGAGGCAAGAAC 3663

XM_025192633.1 PREDI ACTGTGTTGAGTGTGAGGAGATGAAATAGAT 3760

XM_031061670.1	PREDI	AGCCCTCGTGGTAGCCGAGACTGTGTTAGCTGGTCATAATTACATGCCCATCTGCCAGTGCCTTACTGGCCGTCA	3752
XM_037906217.1	PREDI	GCGCTCGAGGGCGACATCAACGAGTCCTCACCACCCCTGCAACCCGCGGGAGCCTGGACTGCGTCAGCTACCA	3708
XM_002939080.5	PREDI	TTGATGACTGTGCTTCAGA-TCTGGAGGACCCGTTGCTTAATGGAGGACAGTCATTGATCAAATTGGGGTACAG	3745
XM_026853751.1	PREDI	GGAGTTGGTGATCTGTGCACTGGCTTCCATGTAAAAACAGAGGCATCTGCA---AGCAAGTAAAACAGAGGAA	3740
XM_025192633.1	PREDI	GTGTGTCGGAGGAGCTGTGCAACACGGAGGTAGTTGTCGTTGCCAGCAATATGCCGATGGGTCATCTGCCAGTG	3840
XM_031061670.1	PREDI	CTGTGAAACTGTGATAGATGTGTCCTCAGAACGCTTGTCAAGAATGGAGGACTTGTGCTGTTGCCAGTAATATGCC	3832
XM_037906217.1	PREDI	ACGACTACAGGTGCACTCTGCAGAACGCGTTCACTGGCCGTCAGTGTGAGAGTGTCAATTGACGTGTCGCGGAAGCC	3788
XM_002939080.5	PREDI	GTGCCACTGTCTCCA-GGCTTCGAGCGAGCGCTGCGAGGGGATATTAATGAGTGCCTTCAACCCCTGAGCCCT	3824
XM_026853751.1	PREDI	GCATTGTCCAACGGTTGACTGGCATTTACTGTGATAT-CCCAGATGTCTTGTGAAGTGGCTGCGTCAGAGGG	3819
XM_025192633.1	PREDI	TCCTCCGGTTATTCTGGAGCAAATGTGAGTTAGCAGCACGCCACAGCACCTGTGAGACAAGTGAAGAAGGGGAGC	3920
XM_031061670.1	PREDI	ATGGATTCACTGCACTGTCCTCCAGGTTATTCTGGGCTAAATGTGAATT---CAG-CA-A-TACTTGTGGCAAATG	3906
XM_037906217.1	PREDI	TGCCAGAACGGCGGACCTGCGCGTGGCCAGCAACATGCCGAGGGCTCATCTGCCAGTGCCCCCGGGTTATTCTGG	3868
XM_002939080.5	PREDI	CATGGTAGCCTAGACTGCATTCACTGGTCAATGATTACATGCGTATGTCGAGTCGCTTCAGTGGCTGCGTGTGA	3904
XM_026853751.1	PREDI	TGTGGCAGTGCAAGATCTATGTGAGCATGCTGG-CATTGCAATCAACTGGCTCAGTCACCCTGCGCAGTGC	3897
XM_025192633.1	PREDI	AGTGCATCCAGACATCTTCTGGCACGGTGCTATTGCCAAGCTGTCTGAGGAGAGGAATGCCAGAACAAACAGGGT	4000
XM_031061670.1	PREDI	AGGTGCAAGAAAGGGGAACTGTGCACTCAGACATCCTCAGGGCCCCGTTACTGTCCAGAGCTCTTGGTGGTGGAGGA	3986
XM_037906217.1	PREDI	AGAAAATGTGAGTCAACAGCACAGTACCTGTGAGAACATGCAAGAAGGGGAAACAGTCATCCAAACGTCTT	3948
XM_002939080.5	PREDI	GACCGTAACAGATGTGTCCTCAGCAGCCTGTGCAAGATGGCGTACTTGTGCTGTCAGTCACGTGCGTGTGGT	3984
XM_026853751.1	PREDI	AGGCTACATGGCAGTTACTGTGAGGGGAACTGGATGAGTGTGCGTCCAATCCATGCCAGAACAGCGTACTTG-TGTG	3976
XM_025192633.1	PREDI	TGTGCCAGTGGCCCTGCCCAGAATGGGAAGCTGCCACCCCTCATTCTCAGCCTCTTACTATTATTGTCACTGCCC	4080
XM_031061670.1	PREDI	GTGCCGGCTACTGCACTGAGCTGACCCCTGCCAGAACATGGGGAAATTGCCATCCTAGGCCCTCAGGCTCTTATT	4066
XM_037906217.1	PREDI	CTGGTCTCGGTGCTATTGCCCTAACGTGTGTCGAGGAAATGCCAGAC-ACACCCAGGCTGCGCAGTGGCCCTG	4027
XM_002939080.5	PREDI	TCATCTGCCAGTGTCTCCAGGTTATTCTGGGTTAACATGTGAGTTGGCAGCCACAGTGCCTGTCGGCAAGTGA	4064
XM_026853751.1	PREDI	GATCGTCACTGGGGATA-CGAGTCAAGTGCCTGCCAGGGTACCAAGAGTAAACTGCGAGTATGAGATAGACGAGTGC	4055
XM_025192633.1	PREDI	TCACACCCAAGGCAGCCAGTGTGAAACAGCCGTAACCTTCAGCCCAGAGCCCCGAGGATGTTGGATTCCCAGTGTG	4160
XM_031061670.1	PREDI	ATTACTGCCAGTGTCCCAGTT-ACACTGAAGGCAGCCAGTGTGAGCATTGACAC-CCCAGCCGAGAACCTGAAGGGT	4144
XM_037906217.1	PREDI	CCAGAAATGGGGGAGCTGCCACCCCTCATCTCAGCCTCTTACTATTACTGCCAGTG-CCCTGCTCACACGCCAGG	4106
XM_002939080.5	PREDI	AAGAAAGGGGAGCAATGCATCCAGACATCCTCTGGGGCCGATGTTACTGCCAAAGCTCTCAGCTGGCA-CGGCGTGC	4143
XM_026853751.1	PREDI	AGATCCAGCCTGCCAACATGGGGACATGTGTTGA-CTTGGTAAATCAATTCCGATGCTCTGCCCTCAGGAACACG	4134
XM_025192633.1	PREDI	AAAAAGCAAGAGATGGTATTGTGATGAAAGACTGTAACACTCATGCCCTGAGGGATGGAGGCAGTGTCCCTGAGA	4240
XM_031061670.1	PREDI	GTCTGGTCAACACTGCATGGAAACGAGCAAAGATGGCTACTGTGATGAAAGACTGCAACACGCACTGGGAT	4224
XM_037906217.1	PREDI	CACTGTGAGCAGCCGT-CACGTTCAGCCGGAGCCCCGAA-GATGCCCTGGATTCCAGTGTGCAAGAGAACCCAGGG	4184
XM_002939080.5	PREDI	AGAC-TGGTACAGGTGTCAGTGCAGCGTGCCAGAATGGGAAGCTGCCATCC-CAGGACTCAGCCACCTACTACT	4221
XM_026853751.1	PREDI	AGGTCCCTCTGTGAAAGAAAATAGATGACTGTGCTCAAGTTCT-GAAGGGCCGACTGTTCAATGGTGGAAAGGTG	4213
XM_025192633.1	PREDI	ATGGAAGATCTTGGCCAACGTCTCTCTTCAATTGCGCTGTTGGATGCTTTCAATGGACAGTGTGACGAGTCTGCA	4320
XM_031061670.1	PREDI	GGAGGGACTGTTCCCTGATGATAGAAGATCCCTGGGCAACTGTTCTCGCTGCGCTGCTGGAGCATTCAATGG	4304
XM_037906217.1	PREDI	GGCTACTGTGAGGGACTGTAACACCCACGCCCTGCCAGTGGGATGGAGGACTGCTCCCTGACTG-TGGAGGATC	4263
XM_002939080.5	PREDI	ACTGCCAGTG-TCCAGATCACTGCAAGGCAAGCCAGTGTGAAACACTCACTGCGC-CCCAGCCAACAGGCCAC	4298
XM_026853751.1	PREDI	TTAGATAAAATTGGCGCTACAGCTGTGAAATGCCCTCC-TGGGTTGCCAGGGAGAGGTGTGAAGGCGATGTGAAT	4292
XM_025192633.1	PREDI	CACACCCGAGTGCCTGTTGACAACTTGAATGTCAGAAAATAGCAAGTATGACAAGTATTGTGCAAGTC	4400
XM_031061670.1	PREDI	CTTGTGCGATGAGGCCCTGTAATACGCCCTGAGTGCTGTTGACAACTTGAGTGTCAAGCAGAACATGCAAGT	4384
XM_037906217.1	PREDI	GGGCCAACTGCTCCCTCCCTGCCCTGCTGGATGCTTTCAATGGACAGTGTGACGAGTTCTGCAACACGCCAGTG	4343
XM_002939080.5	PREDI	CTGGGCCAACACTGTGCAAGAAAAGAGCAAAAGATGGCTACTGCGATGAGGACTGTAACACGCCACGCCAGTGG	4378
XM_026853751.1	PREDI	GCTTGTCAATCCTTGTAGTGCAAGGGCAGCCTGGACTGCATTCACTGCGATAACAACTATTTGCAAGTGAAGC	4371
XM_025192633.1	PREDI	ACTACGCCGATGGCCCTGTGACCCAGGGCTGTAATAGTGAAGGAGCTGGCTGGGACGGTGGACTGTGCTGTC	4480
XM_031061670.1	PREDI	ATGGTAAATATTGCGAGGATCACTATGCAAATGGACTCTGTGACAAAGGATGTAACACGAGGAGTGTGGTGGAG	4464
XM_037906217.1	PREDI	CTGTTGACAACTTGAGTGTGCAAGAAAATAGCAGGACATGCAAGTGTGACGAGTTCTGCTGCGGATCACTAC	4423
XM_002939080.5	PREDI	AGGTGACTGTTCCCTGACAG-TGGAGGATCCCTGGGCAATGCTCTCGCTGAACTGCTGGCAGCACTTCAATGG	4457
XM_026853751.1	PREDI	TGGATTACAGGTGCCACTGTGAGACGATAACAGATTGTG-CCCAACTCGACCTGTCTGAATGGGGACATGCC	4450
XM_025192633.1	PREDI	GCTGAGAAGCTGGCAGAACGGGCTAATCATTGTTGCTGATGCCCTGACGAGCTGCTGGGTGATGTTGCGAGCT	4560
XM_031061670.1	PREDI	CTGGACTGTGCCGTGACAAAGCTGAGAAACCTGGCGGAAGGTACCCCTATAATTGTTGCTGATGCCCTGAA	4544
XM_037906217.1	PREDI	GCGCTGCGACCAGGGCTGCAACAGCGAGGAGTGCAGGACTGCGCTGGGACGGCTGGACTGCGCCGGGACAAG	4503
XM_002939080.5	PREDI	CAATGCGATGAGCTGTAACACACCCGAGTGCCTTTCGACAACACTCAAGTGTGCAAGCAGAACATAGCA	4536
XM_026853751.1	PREDI	GTTG-CTGTGAATATGCCAGGAAACTTACCTGTGAGCAGGCTTTCTGTTCAATGGGACCAAGGTGCTACTGT	4526
XM_025192633.1	PREDI	CTTGCCTACTTGGAACTCTCTGCACACCAATCTGAGAATCAAGCTGGACTCCCAGGGAAACCCCATGGTATT	4640
XM_031061670.1	PREDI	GCTTCAGATGTCCATAGCTTCTATGGACTCTAGGAACCTCTACACACCAACTGAGAATAAAACGAGATTCCCAG	4624
XM_037906217.1	PREDI	CGGAGGGCACCCCTCATCGTGGCTCTGATGCCCTGGACAGCTGCGGAGTGTCCGCAGCTTCTGCGCACCTG	4583
XM_002939080.5	PREDI	ATGATAAGTACTGTGCAAGTCACTATGCAAGACGGGCGTGTGACCGAGGACTGAACTGAGTGTGGAGTGG	4616
XM_026853751.1	PREDI	ATAATACATGTGGACACTATAAGTGCAGGAAACTGTGGGAAACTGTATTAAAC-TCCATTGGGACCAAGGTGCTACTGT	4605
XM_025192633.1	PREDI	ACTACGGGGAGAAATCAGCAGCACGGAGTCGGGATCACTGTGTTGCAAGCACAGAGAACTAGAGCAAGAGG	4720

XM_031061670.1 PREDI GAAAGCTGATGGTATAACCCATACTATGGAGAGAAGT---CAGCACGGAGCCGGCAGTC-CGTGTTGGCTTTCGCAAGGC 4700
 XM_037906217.1 PREDI GGCACGCTGCTGCACACCAACCTGAGGATTAAAGCTGGACTCCCAGGGAACCCCATGGTGTCCCCACTATGGGAGAA 4663
 XM_002939080.5 PREDI CTGGACTCGCTGGACAAAGCTGAGAGACTGGCAGAAGGGACCTCATTATTGTGTTATTGATGCCCGATGAGCT 4696
 XM_026853751.1 PREDI CACGGGAC-CAGCTGGGA-ATGCATGCCA-ATCTAAATTAGGCTGTGCCAGCCAGCCTGTAAAAACAAATGGAGTTGCC 4682

XM_025192633.1 PREDI ATTGGCACCGGGTTCCTGGAGATTGACAACCGTCAGTGTGCAGAGGATTCAAGGACAATGCTTCAGAACACAGAAC 4800
 XM_031061670.1 PREDI CAGAGAGCTTGAACAAGAGGCCATTGGCACAAAGGTGTTCTGGAAATTGACAATCGACAGTGCCTGAAGAGTCGGACC 4780
 XM_037906217.1 PREDI ATCGGGCGCGCGAGCCGTCGGTGC-TCGTGTGGTTCGAAGCACAGAGACTGGAGCAAGAGGTCAATTGCCAAC 4742
 XM_002939080.5 PREDI GCT-CAGTGTGCCCCAGTTCCTGGACTTAAAGGATCCCTTCAACCAACCTGAGAATAAAACTGGACTCACAG 4775
 XM_026853751.1 PREDI AGAGAAAGGA-TCAGCCTCCCTATTA-TACATGCAATTG-CCCTAATGGAAT-TGTTGGACTCCAGTGTGAA-A-T---T 4753

XM_025192633.1 PREDI TGCTGCAGCTCTTAGCTGCTCAGGCCATCAAGGGCATGGCCATTCTCGTGTCTGTCAGAACAGTGAACCCCTGT 4880
 XM_031061670.1 PREDI AGTGCTCAAAAATACAGAACGCTGCAGCAGCTTTAGCAGCTCAGGCCATTAAAGGGCATGGTGCCTTATCCTTTGTG 4860
 XM_037906217.1 PREDI GTGTTCTGGAGATCGACAACGCCAGTGTGCAGAGGACTCAGAGCAATGCTTCACAACACAGAACGCTGCTGCAGCTCT 4822
 XM_002939080.5 PREDI GGAAACCCATGGGTACCCGACTATGGGAGAGCCTGGCATCACGGAGTCGGCAGTCG-CCTCGCAAGC 4854
 XM_026853751.1 PREDI -----CTA-TTGATTG--A-TGGG-TGTCACCTCCACCACTGTCCTGATAAAAGCAAGAGACGGGACTGCGATCAGGA 4821

XM_025192633.1 PREDI TACCACCGAACACAGTTGCTTACCTACTTGCTGTGGCAGCTTGATCATCCTTTAATTCTCTCTGGGTGTGATG 4960
 XM_031061670.1 PREDI TCTGTTGAAAGTGAACGTGTATCACGAGGGAGACCCAGATGCTTACCTGCTGTGGTAGCTTGTGATCGGCTTAT 4940
 XM_037906217.1 PREDI CTTAGCTGCTCAGGCCATCAAGGGCATGCTGCCATTCTTTGTGTCAGTCCAGAGTGTGAGCCCTGTTACCAAC 4902
 XM_002939080.5 PREDI CCAGAGAACTTGAACGGGAGGTACATGGCACGAAGGTGTCTGGAAATTGACAACCGCAATGTATTGAAGATTGAGAC 4934
 XM_026853751.1 PREDI GTGTAACACGATGAATGCCCTTGGATGGTGGAGACTGCTTTAAC-ATGGAAGACCCATGGCCAACGTTCCTCC 4900

XM_025192633.1 PREDI ATGGCAAAGCGTAACGCCAACATGGTCCGTGGCTCCCAGAGGGCTTTATTCTGCCAGGGATCCTAGCAATCATAA 5040
 XM_031061670.1 PREDI AATCCTCTTTGGGTGTGATGATGGCAAAGCGAAACCGAAGCATGGCTCCCTGGCTCCCTGGGCTTTATTTGC 5020
 XM_037906217.1 PREDI CCCAGCTGCTTACCTGCTGTGGCAGCTTGATCATCCTGCTGATCTGCTGTGGGGGTGATGATGGCAAGCGC 4982
 XM_002939080.5 PREDI CAGTCTTAAGAGCACAGAACGCCCTGCAGCTCTTAGCAGCTCAGGCCATCAAGGGCATGCTGCCCTACCCCTTG 5014
 XM_026853751.1 PREDI TCATTGAAATGCTGGAAATTAAATGGCAATGTGATGAGCTGTAAACACCTGAGTGCTTTGACAACTTGA 4980

XM_025192633.1 PREDI GCGCAGAGAGCCAGTTGGGAAAGATGCTGTGGACTCAAAAATTGCTGTCAGATACCAGAGGGTAACCTGCAGATT 5120
 XM_031061670.1 PREDI ACAGAGACTCC-AGTAACCACAAACGAGAACCGAGTTGGGAAAGACGAGCTGGATTGAAAATCTGCACTGAAAGT 5099
 XM_037906217.1 PREDI AAGCGCAAGCAGTGGCTCCCTGTGGCTCCCGAGGGCTTCATCCTGCGCAGGGAC-CCCAGCAACACAAACGAGAGAC 5061
 XM_002939080.5 PREDI GTCTGTCAAAGTGAACCGTGGTACCAAAGAAAACCCAGTTGCTTACCTACTGCTGTGGCAGCTGTGATCATCTT 5094
 XM_026853751.1 PREDI GTGTAAGAAAAGAG-TCAA-TCTGCAAGTATGATAAGTACTGTGAGATCACTATGCAAATGGTCACTGTGACAAAG 5058

XM_025192633.1 PREDI CTGGACCAACTGAACACTGGCAGGTGATGGTGGACCCAGCCAAAGAGAGCAAAGGCTGAGGATCAAGCCTGCTCCG 5200
 XM_031061670.1 PREDI TTCTGACGGCAATCTGGCAGATTGACTGCAAGGGGAAACATTGGTCACTGAGTGGACCTCAGCCAAGAGAGCTAAGT 5179
 XM_037906217.1 PREDI CCGTCGGGGAGGACGCTGTGGACTCAAAAATCTGTCAGTCCAGATAACAGAGGGTAACCTGGCCGACCCCTGGGAC 5141
 XM_002939080.5 PREDI TCATCTGTTGGTGTCTTGCAAGGAAAGGCAAGCATGGCTCCCTGTCGCTTCCGGAGGGCTTGTCTA 5174
 XM_026853751.1 PREDI CTGTAACACAAAAGATGTGGCTGGATGGACTCGACTGCTACAGACAAACCTGAGAATCTGCTGAGGGTACCTGG 5138

XM_025192633.1 PREDI GAAGGTGATGAGCAGATCGATCAGGCCAGTGGACCCAGCAGCACTTGGAGGCAGCAGATGTCTGGAGCACGTCACT 5280
 XM_031061670.1 PREDI CTGAAGATGAAGCATTGCTCCAGAACAGATGAGCAGGTAGATCAATGCCAGTGGACACAACACCTCAGGGCAGCT 5259
 XM_037906217.1 PREDI GAGCACTGGCAGGAGATGGTGGACCTCAGCCAAGAGAGCTAAGGCTGAGGACCAAGGCCCTGTCGCCAGGGGAC 5221
 XM_002939080.5 PREDI CGCCGGGAC-CCCAGTAACCAACAGCGCAGAGAGGCCAGTGGGAAGATGCAAGGACTGAAGAATCTGCACTTCAG 5253
 XM_026853751.1 PREDI TCATAGTAGTGTCAATGCCCTAAAGGAACTGCT-AAAAAAATTCCAGAAATTCCGGAGTTAGGCAGCAATTCTCG 5217

XM_025192633.1 PREDI AGCCCTACACCACCTCAAGCAGATCAAGAACAGTGGACCTCTAGATGTCAGAGGGCCAGATGGATGACTCCAC 5360
 XM_031061670.1 PREDI GACATAGATGGGACACATCCCTGCTTACACCACCTCAGGCAGATCAAGAACAGTAGATGTCCTGGATGTCAATGTTAG 5339
 XM_037906217.1 PREDI GCAGATTGACCAGGCCAGTGGACGACGCAGCACCTGGAGGCGCGACGTCTGCCAGCACCTCCCTGCCCTCACCC 5301
 XM_002939080.5 PREDI CTA-CAGAGGGTAATTGACTGACTGCAGTCCAGCGAGCACTGGACTGGCAATGGAGGACCTGCTGCTAAGAGAGCAA 5332
 XM_026853751.1 PREDI CACCAACCTAAGGATCAAA-AGGAATCCAGATGGGAGCCCATGGTCTTCCATATTAGAGAAAATCAAGCGCA-CA 5295

XM_025192633.1 PREDI TTATGTTGCATCTGCGCTGGCAGCTCCGATATCAGCGAGGATGATGAGGATGAGAACAGATTCTCAGCAATATA 5440
 XM_031061670.1 PREDI AGGACCAAGATGGGTCACCCACTCATGTTGGCATTCTGCGTGTGGCAGCTCTGACATTAGTGACGAAGATGAAGATG 5419
 XM_037906217.1 PREDI CGCCCCAGGCTGATCAGGAGGTGATGTCGGATGTCAGCTGGGGCCAGACGGCTGCCACCTGGCTGCTGGCA 5381
 XM_002939080.5 PREDI GGCTGAAGACCAAGCTTGCTGTCAGACGGAGGGAGGAGCAGGTAGATCAGGCCAGTGGACACAGCAACACCTAGAAGC 5412
 XM_026853751.1 PREDI GCGTTCACGTAGATCTC-TCT-TAGAATTACGGTAGTCC-GGGAGCTGAGGAGCAGCAATTGGCTCAAAGTCTCAT 5372

XM_025192633.1 PREDI ATAACAGATTGATCTATCAAGGTGCCAACCTGCAGGCCAAACAGACCGCACTGGGAGATGGCACTGCACTTAGCAGC 5520
 XM_031061670.1 PREDI GGGAGGATTCTCAGCAAATAATAACAGACACTGATCTATCAAGGGCTAACCTGCAAGCTAACACAGACCGTACTGG 5499
 XM_037906217.1 PREDI TCTCTGCGTGCAGGGGGCTCCGACATCAGCGAGGATGATGAGGACGGGGAGGATTCTCTGAGCAAAACATCATCACAGATCT 5461
 XM_002939080.5 PREDI CTGACATCTGTTGGACCCATCTCTAGCCCTACACCACCGCAGGAGGATCAGGAAGTGGATGGCTGGATGTCAATGTC 5492
 XM_026853751.1 PREDI TAGAAATTGACAACCAACAGTGTGCTGGATTCACTGAGTTAAAGACTACAGACTCCGCTGCTGCCATGCTGGCT 5452

XM_025192633.1 PREDI CCGTTACTCAAGGAGCAGATGCTGCAAAACGTTGCTGGATGCTGGCAGATGCTAATGCACGGACAACATGGGAC 5600
 XM_031061670.1 PREDI GAGATGGCACTGCACTTAGCAGCCGCTATTCAAGAGCAGATGCAAGGAAACAGCTGTTGGATGCTGGCTGATGCTAA 5579
 XM_037906217.1 PREDI GATCTACCAGGGGCCAACCTGCAAGCGCAGACAGACCCGACTGGGGAGATGGCCCTGCACTGCCAGCCGCTACTCCC 5541
 XM_002939080.5 PREDI AGAGGGCCAGATGGGTCACTCCACTCATGTTGGCATCTCCGAGCTGGCAGCTCCG----A-CGA--A-GATGAAGA 5563
 XM_026853751.1 PREDI GCCCATGCTGTACAGGGTACACTTAATTATGCCCTGGTCTCAGTGGGGTTGAAAAGGCGCACCATGCTCCAGACA-C 5531

XM_025192633.1 PREDI CTCCCCCTGCATGCACTGCTGAGCTGCTGACGCCAGGGTGTCTCCAGATCCTGATCCGAAACAGAGTGAACAGACTGGAT 5680

XM_031061670.1 PREDI CGCCCGAGACAACATGGGGAGGACTCCTCTTCATGCAGCTGTGGCAGCTGATGCACAGGGTGTCTCCAGATCTTGATCC 5659
 XM_037906217.1 PREDI GAGCCGACGCCAAGCGCTCTGGATGCCAATGCCAGGGACACATGGGACGGACCCCTGATGCAT 5621
 XM_002939080.5 PREDI AATGGAAGATTCTTCAGCAAATAATAACAGACTTGATTATCAAGGTGCCAACCTGCAGGCCAAACAGACCGTACTG 5643
 XM_026853751.1 PREDI TGATGTACATA-TTGGCCGTTATATTGGCCATCATTATTGTCATTCTGCTGCTGGTTGTGGTACTGGCCAAAAG-AAGC 5609

XM_025192633.1 PREDI GCTCGGATGAATGATGGAACAACCCCCCTTAATTCTGGCTGCACGTCTGGCTGGAAGGAATGGTGGCTGAGCTAATTAA 5760
 XM_031061670.1 PREDI GGAACAGAGTGCACAGACCTGGATGCTCGCATGAATGATGGAACAACCTCGTTAATTCTGGCAGCTCGTGGCTGAG 5739
 XM_037906217.1 PREDI CGGGCCGTGGCTGCTGACGCCAGGGCGCTCCAGATCTGATCCGTAACAGAGTGCACAGACCTGGACGCTGGATGAA 5701
 XM_002939080.5 PREDI GGGAGATGGCACTGCATTAGCAGCCGCTATTCAAGAGCAGATGCAGCAAACGCTTACTAGACGCTGGTGTGATGCT 5723
 XM_026853751.1 PREDI GCAAACATGGCACCCCTTGGTCCCTGAAGGTTCCCTGCACCGGGAC-ACCAGCAACCAGAAGCGCAGAGCCTGT 5688

XM_025192633.1 PREDI TTGCCAGGGCAGATGTGAACGCCCTAGATGACCATGGTAAATCTGCTCTTCACTGGGCTGCTGTCATAAATGTGAAAG 5840
 XM_031061670.1 PREDI GGAATGGTGGCTGAGCTGATCAACTGCCAGGAGATGTCAATGCAGTGGATGACCAAGGTAATCTGCTCTTCAATTGGGC 5819
 XM_037906217.1 PREDI CGACGGGACGACCCCGTGTGATCTGGCTGCACGTCTGGCTGGAAGGGATGGTGGCGAGCTCATCAATTGCCAGGCTG 5781
 XM_002939080.5 PREDI AATGCACGAGACAACATGGGCCGTCTCCCTCATGCAGCTGTGGCAGCTGATGCACAGGGTGTCTTCAGATCCTGAT 5803
 XM_026853751.1 PREDI TGGGCAAGGATGCTGTTGAATGAAAAATTGTCATGTGC-CCATAGTGGATGAAACTT-ACTGGATTCTAATCAGAATGA 5766

XM_025192633.1 PREDI CAACACTAGTGCTGCTGAAAAATGGGCCAACAGAGACATGCAGGATAACAAGGAAGAGACACCCTTCTGCTGCT 5920
 XM_031061670.1 PREDI TGCAGCTGTTAACATGTGGAGGCAACGTTAGTGTGTTGAAAAATGGGCCAACCGAGACATGCAGGACAACAAGGAGG 5899
 XM_037906217.1 PREDI ATGTCACGCCGTGGATGACCCACGGTAAATCTGCTCTTCACTGGGCTGCTGTCACAATGTGAAAGCAACACTAGTA 5861
 XM_002939080.5 PREDI CCGTAACAGAGTGCACGGACCTGATGTGCGGATGAACGATGGCACAAACCCCACTAATTCTGGCTGCCGACTGGCTGTGG 5883
 XM_026853751.1 PREDI TCACTGGTCAGATGATCAGGACACAGCCAAGAAATCCAAGTCTGAAGGTCAGGCTCTCTGCGGGAGGAGATGACC 5846

XM_025192633.1 PREDI CGGGAAAGGCAGTTTGAAGCAGCAAAATCTGTTGGACCATTTGCCAATCGAGACATCACAGACCACATGGACCCCT 6000
 XM_031061670.1 PREDI AGACACCACTTCTGCGAGCTGGGAGGGAGTTTGAAGCAGCTAAATTCTGTTGGACCATTTGCCAATCGGGAC 5979
 XM_037906217.1 PREDI CTGCTGAAAATGGGCCAACAGAGACATGCAGGACAACAAGGAGGAGACTTCACTTCTCGTGTGCGGAAAGGAG 5941
 XM_002939080.5 PREDI AAGGAATGGTGGCCGAGCTGATCAACTGTCAGGCTGATGTCAATGCAGTGGATGATCATGGTAAATGGCTCTTCACTGG 5963
 XM_026853751.1 PREDI AAGTTGAT-AGCCGTCACTGGACACAGCAGCACCTGGTGGCTGCAGATATTGGATGACCCCATATTGGCTCTCACACC 5925

XM_025192633.1 PREDI GCCCCGAGATGTGGCCCAGGACCGATGCACCATGACATAGTGCAGCTCTCAACGAGTACAACGTGGCGCACAGCCCG 6080
 XM_031061670.1 PREDI ATCACAGACCACATGGACCGTCTCCCCGTGATGTGGCACAGGACCGATGCACCATGACATTGTCAGCTCTCAATGA 6059
 XM_037906217.1 PREDI CTTTGAAGCAGCAAAATCTGTCGGACCATTTGCCAACCGAGACATCACGGACACATGGACCGCTGCCGGGACG 6021
 XM_002939080.5 PREDI GCTGCTGTTAACATGTGGAGGCAACATTAGTGTGTCGAAAAACGGAGCAACAGAGACATGCAGGACAACAAGGA 6043
 XM_026853751.1 PREDI TCCTCAAGCAGAGCAAGACATGGATGTCAAGATGTCAACGTGCGAGGGCAGATGGCTGCACACCCTTATGTTGGCCT 6005

XM_025192633.1 PREDI CTGGGCACCCCTGGTGCATGCTGAGCTCTGCCCTCTCCCGTCATTGTCGCCCCAAATAGGTCTTCTCAACCTGAAG 6160
 XM_031061670.1 PREDI GTACAATGTGGCCACAGCCCATCTGGCACCTGGCACCATACTAAACTCTGCCCTCTCCCCAGTGTGATGGCCCA 6139
 XM_037906217.1 PREDI TGGCCCAGGACGGATGCACCCACGACATCGTGCAGCTCTCAACGAGTACAACGTGGCCACAGGCCCTGCCGGCACCC 6101
 XM_002939080.5 PREDI GGAGACCCGTTTCTGGCAGCACGGGAAGGGAGTTTGAAGCAGCTAAATCTGTTGGACCATTTGCCAATCGAG 6123
 XM_026853751.1 PREDI CCCTGCGAATGGGAAGC-CCTCACATAGAGGAAGATGATGAGGAGGCGAGAAGATCCATGCTCAACGTCTACTGACCT 6084

XM_025192633.1 PREDI CATGCTCATTAAAGCAAGAAGTCAAGAAAACGAATGCCAAGGCATCATGCCAACAAACCTCACCACAAAGATGCAAAAGAG 6240
 XM_031061670.1 PREDI ATAGATCTTCTCAATTGAAACATGCTCTCTAAGCAAGAAGTCGAGGAAACCAATGCCAAGAGTGTGATGGCCCA 6219
 XM_037906217.1 PREDI GGGGCCATGCTGAACTCTGCCCTCTCCCTGTGATCTGTCGCCCCAACAGTCCTCTCAACCTGAAACACGGCATCACT 6181
 XM_002939080.5 PREDI ACATCACAGACCATGGATGCCCTCTCCCGAGATGTGGCACAGGACCGATGCACCATGACATCGTGCACCTCTGAAT 6203
 XM_026853751.1 PREDI GATTTATCAGGGAGCTGTTGCAGGCCAGACAGATGCCACAGGAGAAAATGCCCTGCACTGGCAGCTGCTACTCAA 6164

XM_025192633.1 PREDI GAGAAGGAAGAAGTCCCTCAGTGAGAGCAAGGGCAGTTTCAGAGAGCTCAGTGACCCCTGTCACCAGTTGATTCCCTAG 6320
 XM_031061670.1 PREDI AATCTTGCCTAACAGAGCAAGAGAAGAAGGAAAGAAATCTGACTGAGAGCAAGGGTCAGTTCTGAGAGCTCAGTGC 6299
 XM_037906217.1 PREDI AAGCAAGAAGTCACGGAAACCGAATGCCAA-AG-G-CATGCCCACTAGCTCACCACAGATGCCAAGAGGAGGAGGAAGA 6258
 XM_002939080.5 PREDI GAGTACACCTGACCCATAGCCCTCTGGCATTCTGGCACCATGTCAGCTCAGCCCTCTCACCAGTGTGATCTGGCC 6283
 XM_026853751.1 PREDI GGGCCGATGCTGCCAACGCAAGGAGCTGATGTCAGGACAGTATGGGCCGACCCCTCAT 6244

XM_025192633.1 PREDI AATCACCCATGCTTGGCTCCGAGCCACATCGTCTCTGATGCCATCACCAGGGGTGTTACATTGTCGCCAGC 6400
 XM_031061670.1 PREDI CCTCTCACAGTCGACTCCCTGGAAATCTCCCTGGCAGCTTGTGATCTGACAAACGTCCTCTCCCTGATGCCCTCC 6379
 XM_037906217.1 PREDI AGTCCCCTCAGCGAGGGCAAAGGGCAGTTTCCGAGAGCTCCGTGACCTGGATTCCTGGAGTCTCCCAT 6338
 XM_002939080.5 PREDI AAATAGATCTTCTCAATTGAAACATGCTCTTGAGCAAGAAGCAGAAACCAAGTGCCTGGGCTGTGCCA 6363
 XM_026853751.1 PREDI GCAGCGGTCTGCTGATGCACAGGGGTTTCCAGATTGATCGTAACAGAGTCACAGACCTCGATGCGAGGATGAA 6324

XM_025192633.1 PREDI TCACTGCTGACTGCTCCAGCTGTGCAGGCAGCACACAGATGTCTCTCCAAACCTCCATGAGATGCAAGCCCTTAGGACG 6480
 XM_031061670.1 PREDI GGATTCTGCATTCTCGGCCAGCTCACTGCTACGGCCTCAGTTGTCAGGCCGCACACCCCATGTCCTCTGCTGCTC 6459
 XM_037906217.1 PREDI GCTTTGCACTGTAACCAACCCCTCTCTGATGCCGTCACCAGGGTGTGCACTCATGCCAGCTCGTGC 6418
 XM_002939080.5 PREDI CCAATCTCCAAGGATGCAAAGAGGAAGGGAGAAATCTGAGTGAAGCAAGGGCAGTTCTGAGAGCTCAGTG 6443
 XM_026853751.1 PREDI TGATGGTACAACCCCTCATCTTAGCTGCCGCTTGCTGAGGATGGTGCAGACTAATAAACTGCCAGGCTG 6404

XM_025192633.1 PREDI TGGATCCAGCACGGTCTTCCATCTGTCAGGCCAGCTGCTTCAGCACAGAACCCCCCTCCAGTAGCGGGCTGGCA 6560
 XM_031061670.1 PREDI CATGAGATGCAGCCGCTGGGGCAGGGATCTAACACCGTGTCTCCATCTGTCAGGCCAGCTGCTTCACAGCACCA 6539
 XM_037906217.1 PREDI GG-CTGCGTCGGTGACGGCGCGCACACCATGTCCTCTCAACCTGACAGAGATGCCAGCCCTGGCGTGGCTCCAGC 6497
 XM_002939080.5 PREDI ACCCTCTCACCGAGTTGATTCCCTGGAATCCCCCATGCTTTGTATCTGAAACCAACCCATCTCCGTGATGCC 6523
 XM_026853751.1 PREDI ATGTTAATGCACTGGATGATCTGAAATCTGCACTGGCTGCTGAGTAAACAAATGTTGAGGCTACAATTGTT 6484

XM_025192633.1 PREDI GGCTCCGGCCAGCCAACATTCTACTGAGTGGATGAACCGTATGGAGGTGAACGAGTCCCAGTACAATGAGATGTTGGC 6640

XM_031061670.1 PREDI AGCTTCAGCAATGGGCTCGTAGACTCCAGGCCAGTAAACATTTCCACCGAATGGATGAACCACATAGATGTGAATGAAT 6619
 XM_037906217.1 PREDI ACGGTGCTGCCCTCCGTCAGCCAGCTGCTCAGCACAGAACCCCTCCAACAGCGGGCTGGGCAGGCTGCGGCC 6577
 XM_002939080.5 PREDI AGGGGTTTACCTTCGTC-CCCACACACGCAACGCTG-CCACATCAGTACAGGCTGCGCACACCATGCTTCTCCAA 6601
 XM_026853751.1 PREDI CTGCTAAAGAATGGCAAATAGGGATATGCAAGACAACAAGGAAGAACACC-CTTGTCTGCTAGAGGGTA 6563

XM_025192633.1 PREDI ATGGTCCAGGTGATGCATTGCACTGGCACCCAGTGTTCAGCAGAGCAGTTGGTCAAGCAGATGCCAAGCACATGGG 6720
 XM_031061670.1 PREDI CCCAATACAATGAGATGTTGGATGATCCCACAGGTAGTCCATGCTACCCAGGGCTCTCAGCAAGGAAGTCTGAGT 6699
 XM_037906217.1 PREDI AGGCAACGTCTCCAGCAGTGGATGAAACCGCATGGAGGTGAACGAGTCCCAGTACAACGAGATGTTGGCATGGTCCCC 6657
 XM_002939080.5 PREDI CCTCCATGAGCTGCAAGCCCTGGGATGTGGATCCAGCACTATGCTTCTCAGTCAGCCAGCTGCTTCTCAGCACAGAA 6681
 XM_026853751.1 PREDI GTTTGAAGCAGCAAATGCTTTGGATTACTTGCAAAACGAGAAATCAGTACATGGATCGCTTCCCGTGC 6643

XM_025192633.1 PREDI AGTGTCCAGGGAGTCTCTGCCCCCTATCATGACTTTCCAGCTGGTCCCAAGGGTGGCATAAACCAGCAAGCAGTCCGC 6800
 XM_031061670.1 PREDI CAAACAGATGTGAAGCATGTGGGAATCTCCAGGGAGACACTGCCCCCTATTATGACTTTCCAGCTGGTCCCAAAGTGG 6779
 XM_037906217.1 PREDI AGGTGATGCATTCCACCCAGCAGTGGCTCGCAGCAGAGCGGCTGGTCAAGCGGACGCGAAGCACCTGGGGTGTCCAGG 6737
 XM_002939080.5 PREDI TCACCCCTCCAAACTGTGGGCTGATCGGCTCAGGCCGTTAACAT-TTCTGCTGAATGGATGAGTCGTATGAAATGAA 6760
 XM_026853751.1 PREDI ATTGCCAAGGACAGAAATGCATCATGACATTGCGCCTGTTGGACGAGTATAACTTGGTCACAGCCCACAAGGACAGCC 6723

XM_025192633.1 PREDI AGCAGACCCAGTCAAACAGCAGCTCAGAACATGGCTGGTCTTAGCTCCATGTACCAGATACCAGATATAGCTCAGTG 6880
 XM_031061670.1 PREDI CATGAACCAGCAGGTCTGGCACACAGGCCAGTCAACTGCCCTCAGAACATGTAGCTGGTCTCTAGTTCAATGTACC 6859
 XM_037906217.1 PREDI GAGTCTGCCCCCATCATGACTTTCCAGCTGGTCCAAAGGGCAGCAAAACAGCAAGCAGTCCACAGCAGGCCA 6817
 XM_002939080.5 PREDI TGAATCCCAGTATAATGAGATGTTGGATGGCTGGTCCAGGTAGTGCATTCACCCAGCATGCCAGCAGTGGCA 6840
 XM_026853751.1 PREDI TGGTACAATGCT-TTCAATTCCCTGTCGCCAGCAGTGTGGCTCAACAGATTTCTGAACCTGAAACATG---CC 6799

XM_025192633.1 PREDI CCCAGTGCCTCTTCTCATGGCTGCCATCCCTCAGCAGGATGGCAGGTGGCTCAGACGGTCTCCGGCTACACCACCA 6960
 XM_031061670.1 PREDI AGATAACGCCAGATCTAGCTCAGCTGCCAAGTGCCTCCTTCCATGGCAGTAATCCCTCAGCAGGATGGCAGGTGCT 6939
 XM_037906217.1 PREDI GTCCAACCTGCCTCAGAACATGGCTGGCCCTTCTAGCTCCATGTACCGAGT-C-CAGATCTAGCCCAGTCCCAGCG 6894
 XM_002939080.5 PREDI TGATTCAAACCGATGTGAAGCACGGGAGTCTCAGAGTGTCACTGCCACCTATTATGACTTTCCAGCTGGTCCCAA 6920
 XM_026853751.1 PREDI ---GC---A---AAA-CAA-AG---AA-----G-G-CCA-GA---GGCAA-G-TGC-----CA 6833

XM_025192633.1 PREDI GTTTAGAGCTCGATGGGAAGTACCCACGCCCTCGCACAGCACAGCTGCTACTCCTCGGCTACAGAACGGACTCCCTA 7040
 XM_031061670.1 PREDI CAGACAATCCTCCCAGCCTATCACCAAGTCCAGAGCTCGATGGAAAGTATCCACTCCGCCATCACAGCACAGCTACAC 7019
 XM_037906217.1 PREDI CCTCGTCTCCATGGCTGCCATCCCTCAGCAGGACGGCAGGTGGCTCAGACGGCTCCGGCTACCACCGATTCCAG 6974
 XM_002939080.5 PREDI GGTGGCATAAACACAGCAGCAATGCCAGCAGGACAGTAAACTGTGCTCAGAAATATAGCTGGCCTCTTAGTCCAT 7000
 XM_026853751.1 PREDI -AAA---GT---GT---AGC-A---C---CTC-GA-AG-TC---T---TC-CA-AACTT-A-GCCAAAGA---CC 6876

XM_025192633.1 PREDI GCCACAACCGGACTTACAAGGGGAGCACCCATACCTGACTCCGTACCCGAATCTCAGACCAGTGGTCAAGCTCCTCC 7120
 XM_031061670.1 PREDI CTCTCTCTGCCACGGAGAGGACTCCTAGCATAACAGCAGCAGCTGCAAGGGGAGCACCCATATCTGACGCCATCTCTG 7099
 XM_037906217.1 PREDI AGTCGATGGGAAGTACCCACGCCCTCCAGCACAGT---GCTACTCCTGGCCACAGAACGGACTCCTAGGCCA 7051
 XM_002939080.5 PREDI GTACCAAGT---ACCAGATCTAGCTCAGTTGCCAAGTGCCTCCATGGCAGTGATCCCTCAGCAGGATGGACAAG 7077
 XM_026853751.1 PREDI A---AAAG-A---T---ACCAA-A---ACAG-GA-G---G---A---GG---A-----A-AAAGTCAG-A 6912

XM_025192633.1 PREDI CCACACTCTGCCCTCTGACTGGCTGATGTGGCAACCAGTCCAGTAACAAACAGCAGGGCCGGCGCTCACATGAC 7200
 XM_031061670.1 PREDI AATCTCCGATCAGGGTCAAGTCTCTCCCGACTCTGCCATGAGTTGGCTGATGTGGCAGCAGCCCCAGCCATGGT 7179
 XM_037906217.1 PREDI CAACCGGACTTACAAGGGGAGACCGTACCTGACTCCGTGCCCCGAATCTCGGACCAGTGGTCAGCTTCCCG 7131
 XM_002939080.5 PREDI TAGCTCAGACAAACCTCCAGCCTATCACCAAGTTTCAGAGCTCAGTGGAAAGTACCCACTCCGCATCTCAGCACAGC 7157
 XM_026853751.1 PREDI TACTG---A-CAGTA-A---AGCG---CA-G---C-T-T---T-C-TGA---AAG---CT---CA-GT---A--- 6949

XM_025192633.1 PREDI AGAGCAGCAGCGAACACATGCAAGGTGATGCCTGAAGGCTGGCAAGGTGAAGCCGATTAAGCGGACTCCAGAACT 7280
 XM_031061670.1 PREDI AGTACCCAACGTGGTCCGGTAGCACGAATGTCAGAGCAGCAGCGGAACACATGCAAGGTTATGCATGAAGAGCTTGAA 7259
 XM_037906217.1 PREDI ACTCTGCCCT---T-GACTG---GT-----C---T-G---ACGTGG-CC---A-C-C-A-GTC- 7167
 XM_002939080.5 PREDI T---ACTCCTCCTCTGCCACAGAGGACTCCAGGCCATACCAGGACTTGCAGGGAAACCCATATCTGACTCCATC 7234
 XM_026853751.1 PREDI A-CATTAT-C---CC---C-T---G-TGGATTCT---C-TGAAT---CCCC-CC---A-TG-CTACTTCATCCT-A- 7001

XM_025192633.1 PREDI CATGAAGGTCTCAGGACAGAAATGAAGAACCATTTAGTATCCAAGACTCTTAATGAGCCAGGCTGTTCTTCTTG 7360
 XM_031061670.1 PREDI TGGTGGCTGAACCTAAGTGGACTCCAGTACCAATGAAGGTCTTCCAGGAGGGAAATGAAGAACCCCTTTCTTTAGT 7339
 XM_037906217.1 PREDI ---C---A---G---C---A---CA---A---C---C---A---GC---G---C---G---GG---C---CC---G---C- 7192
 XM_002939080.5 PREDI ACCTGAGTCACCTGATCAATGGTCAAGCTGCCCCACACTCTGCCCT---AGATTG---GT-----C 7291
 XM_026853751.1 PREDI -CC-ATT-ACA---TCT-C---CT-G---GAG-CG-CT-G-CAT---ACG---TCCC-CT-GGA-CAGATGATT-GCA---G-C 7053

XM_025192633.1 PREDI GAGGGACTGTGGTCCGCTGTAGTCTGTATGCTTATGCTAAACAGAGAACACTGTAGAAGCTGGTGTGGAGAACCGAGT 7440
 XM_031061670.1 PREDI GCTCAAGAGTCTTAATGAGCCCAGTCTGTTGGAGGACCTTGTACCTAACAGTCTCCAGTACTACATGATTGCTCCAAACACT 7419
 XM_037906217.1 PREDI ---GG---CC-----G---CCC---AC-----A-TGC-C-GGAG-C-----AG---CAG---C----- 7218
 XM_002939080.5 PREDI ---T-G-----ATGTGG-CC---A-C-C-A-G-C---C-CCA---G---C-CTCGGCA---GCA-AC-----C- 7326
 XM_026853751.1 PREDI CTCTACGT--TGCA-GC-CA---CCA-C---G-CACAACA-A-TGTC-T---TT---AACTAATCTCAGC-CGTTG-GT-T- 7112

XM_025192633.1 PREDI TGGGTCACTCTAACCAACTCCATGCACCTGGGTGTGATTGCTGACTGGATTCTCCAGTCTGCTTTAAATAT 7520
 XM_031061670.1 PREDI GTGGAAGCTAACATGGGAAGACCTGTTGGTGTACCTAACAGTCTCCAGTACTACATGATTGCTCCAAACACT 7499
 XM_037906217.1 PREDI ---GGA---A---C---A---ACA---T-GC---AG-G---TGT---ACGCT---T---GA-C---GGC 7248
 XM_002939080.5 PREDI ---A---GC-G---AG-GG---C---C---TG---C-----AG---C---AC--- 7344
 XM_026853751.1 PREDI ---GTG-CA---GCCA-ATAATATG-C---T---GC---CTT---C---TGTGG-----A-CA-G---TTG---TTATCA- 7156

XM_025192633.1 PREDI GTTACACCATTCTTGGGGAGACAGCTTCAAGCTTAATGCATCTGTTGGCCTTCTCCAGAAATTAGCCAAACCCA 7600

XM_031061670.1 PREDI AGATTCCCCAGGTGTTGCCAGCCTTGCATTCAGTCATTCCACCATTTCCCTGGGACACTATCCCTCAGGAGCCTC 7579
 XM_037906217.1 PREDI TG---C---CCA-----CAG--AG--C---A---G---CCC--CACTA-A-----A-GCG--G---A---C---T-C--- 7280
 XM_002939080.5 PREDI ATA-TG-TC-AGAG-C-----AG--CAG--C-----GG-----A-GC---A-----ACA-T-GC---AG- 7376
 XM_026853751.1 PREDI CAC---T---C-G-GCTCA-AA-ATA-AT-GT--TCTTA---CA-GCA-CA-----A-AGA-C-AGCTCA---CT 7201

XM_025192633.1 PREDI GCTTTGTATCCTTATTTGTTACTCCTGTTACCCATGTAACCTCTCTGAAGCACCTCTACTAGTTCTCTTAGTTC 7680
 XM_031061670.1 PREDI ACAGTTAACGTGTTTCTGAAATTCTAACATCTAACACTAGTTATGTAATTATTGTTACTCATCTAACCCCTTGTACCGC 7659
 XM_037906217.1 PREDI CA---G-A---A-C---T---CAT---G---A-----AGG-TCTTC-CAG---GA-TG-----G-AA---A---T-GA- 7314
 XM_002939080.5 PREDI G---T-T---T---AT-----G-C-----CTG-----A-----AG-AG-GTTAAAATG---GC- 7403
 XM_026853751.1 PREDI A-TG---T-C-AT---C-T-G-A-T---TGGCTGAGCCGATGGA-AAGGA-AT---GAC---A---CCCAG---TAT-A-A--- 7251

XM_025192633.1 PREDI AAGTTGCCCTGTCTGACACAGCTTAGAGCTGCACCCACATTCCCTAGACCCCTCCTCCCTCAGCCACTGTTGCTCAGCA 7760
 XM_031061670.1 PREDI TCTTAAATTCCCTTTCTCCCTGCTCTGCTTAGTCTCCAGGAGGGCAAGCGTCTGACTTTTTAACGGCTCCG 7739
 XM_037906217.1 PREDI -AG---A-ACCATTAA-G-TA-TC---C---A-A-----G---A---G-TCTT-AA-----TG--- 7345
 XM_002939080.5 PREDI -TGCT-GA-----A-GC---TG---A-GC---T-----G-A---CTGCAGTCCCCAT---G---A---CA-G- 7438
 XM_026853751.1 PREDI T---GAAAT---GTTG---GT-----A-T-----A-GTC---CAC---C---A-----TG---ACACA-A---A-T--- 7285

XM_025192633.1 PREDI AGCAAGCCAGACTTTCAAAACACATGCAGGTCAAGTGACACCTCTGCGCGAGCTCACCCCTTCCTCAAGAGTCTT 7840
 XM_031061670.1 PREDI AGATGGCACCCACAAGTACCTCTCCTACTCTCCACTTGGTGGGCCCCCTTGTCTAGTAACCAAACAGGCCCCAT 7819
 XM_037906217.1 PREDI --AG---C-CC---A-GG---A-----TG---TTCTT---C---G-T-TG-GA-GG-----G----- 7372
 XM_002939080.5 PREDI TC-ACCCAG---GA-TG-----G-AA---A---T-GA---AG---A-ACCATTAA-G-T-GT---TC---A-A----- 7477
 XM_026853751.1 PREDI ---A-A-C---CAC-T---CA---TG-----TA-T---TCC-TCA---A---CA---A---AATGGT---A---T---T- 7318

XM_025192633.1 PREDI TGTCTTCAGCAGACCCACTAACACCCCTTACGATTCTACTGCCTGACCCCCAGGTATTATGAGATTATATACATATA 7920
 XM_031061670.1 PREDI GCTCATGTGAGACAAGTGTCCGGAGCTGCCCTCTCTAAACCCCTACTGCACAAGACCCCTGTTAAGGTGCTGGTCCC 7899
 XM_037906217.1 PREDI --AC-T-GT-GG-----TC---GC-T-GT-----A-G-----A-CT---TGT-ATGC---T-TTC---A---TG--- 7405
 XM_002939080.5 PREDI A-AG-TC-TGAA-----TG---CG---C-CC---A---TCCACCACCCG---CCTT---C---CG 7512
 XM_026853751.1 PREDI ----A---T---C-C---A-A---G-----C-----AG-----A---T-----A---TC---A- 7334

XM_025192633.1 PREDI GTCTTCCTCTAGGTTTTGTTGCGTTCCCCCCCACACATTCTGAAATGCTAGTCTGACCTGGTCTTCCCCCGTGG 8000
 XM_031061670.1 PREDI TCATACATGAATGTGTACATTTCAGCTGCCAGCCCCATAAAATCTCTCAGATAAAATCTGCATGTCCTTGACATCT 7979
 XM_037906217.1 PREDI -CT-AA---A-TG---GA-----G-----AA-----CT-G-----T-AGA-A-G-CTG-GT-----A---T---G---G- 7436
 XM_002939080.5 PREDI -T-TG-GA-----AG-----AAC-T-GT---G-----AT---TGC-T-GT-----A-G- 7536
 XM_026853751.1 PREDI -----A-----GC-----C---A-----A-----C-----A-----A- 7342

XM_025192633.1 PREDI TGACCTGTAGGCACATCTCCTGTGTTGATGTTCTTGGACACACCTGACCATTGACACACTACCCCTCCGTGCC 8080
 XM_031061670.1 PREDI CTCTAATTCTATGACCAAAGTACTGGTCTGACATGGTCCCTCAGCTGCCCTCAGATAATGGTCTTCAAACCTTCC 8059
 XM_037906217.1 PREDI -AG---A-AGA---C-CAG-----CTG---G-----G-----TC-ACCTC---TGACC---A---ACT-G---G---A 7470
 XM_002939080.5 PREDI -----GT---T-TTACGC---TTC-----C-CGA---A-T---CAGA---G---AA---GCTG 7565
 XM_026853751.1 PREDI -C-----T-----G-----C-----ATA-TG-----G-G-----TG---TG-GC--- 7359

XM_025192633.1 PREDI CCCGTCTGCCAGGTGTCCTGACCCACCTTACCCACCTCATGGCTCCCGTGACAGAAGAGATGCCCTGCAACCTCTG 8160
 XM_031061670.1 PREDI TTGAGATGTCTGACCTCTTGACACACTAGCACACTAACATGTCCCCCCCACCCCTCATGCAAGTGCACATCTATCTG 8139
 XM_037906217.1 PREDI T-G-TGACT---GCTGT---CT---G-CA-CTGGATT-----C---TCC-----A-GTCTG---CTTT-----TA-A- 7511
 XM_002939080.5 PREDI GAGATTGGGA-A-GAC---GTG-----T-TG-GGCCA---CCTCTAG-----C-CAG-TTCCCCAG---G-----G- 7611
 XM_026853751.1 PREDI ---TCGG-----G-----A-T-GC-----T-----T-----T-----T-----G-C--- 7373

XM_025192633.1 PREDI AAGGCCTTTGGGAGCTTCTAGGATGAGCATGAGTGTTCAAAACCTGGTAACACAAGCTGCCAGCGTTCTGCACAGT 8240
 XM_031061670.1 PREDI GCTAGCCGTCTGAGGTACCAATTGTTAAGAGGAGTTCTGGAAATGAAACCATCTCTACAGAAACTCTCAGTTAACGG 8219
 XM_037906217.1 PREDI AC---AC---GT---TCCA---C-CAT-----TG-----C---T-T-TG---G---G---GA-G-A-C-AGCTCTT---C 7547
 XM_002939080.5 PREDI -----AC-T-AGT---C-GTGA-TTGCTCTAT---ATGA-T---AT-T---CTCTAG---A-CT---TT----- 7653
 XM_026853751.1 PREDI ---CT-----CCTA---TT---ATG---AC---T-----T-----T-CCA-GC---T 7395

XM_025192633.1 PREDI ACTCCCCCTCCCTCCATCCGAGGGTGGCTGTTCTGCTCTCCAGACATGCTGCCCTTTATGGTGTATCAA 8320
 XM_031061670.1 PREDI TGAGGAACCTTTGACCTAATACCGAACCTCCAGCATTACTACAGTGTCCCTCAGTAAGAGTAATGCACCTTGAATATAG 8299
 XM_037906217.1 PREDI AG---CT---TAAT---A-CATC-TGT---G-----GTCA-TTT-----CTCCAGAAATTAG-CC---A-A---AA---CCA 7594
 XM_002939080.5 PREDI -GTCC---A-GTCTG---CTT-----TA---A-ACTTA---T---TCC---GC-CA---CT---TTC--- 7688
 XM_026853751.1 PREDI -----T---A---TCCCA-----A-----A-G-----A---G---TG-----G-A---A---T 7413

XM_025192633.1 PREDI GAAGTGTGAGGACGTCGTTACATGTTGGCAGTGTGTCCTCTGCGTCAGGAATGGCATTCCAGCCCTTGGAAAGTG 8400
 XM_031061670.1 PREDI TTGTCCTGAAACCTCTCAAATTAGCATCTGTCCTCTATTAAACGGGACCCCTCAGTCCTGGTAGAGTACACTGCCCT 8379
 XM_037906217.1 PREDI G-----C---T-T-----T---GT-AT-C-C---TT---AT-T---TTGT-T---A-C-TC-----C-TCCTG--- 7625
 XM_002939080.5 PREDI T-T-T-----GA-G-A---GA-TTGT---GCTCAG---CT---TAAT---CCTTC-T-T-CTG---GTTACAGCTT 7734
 XM_026853751.1 PREDI -----G---A-ACC---A-----G-----C-----AG-TC----- 7425

XM_025192633.1 PREDI TCCCTTTGCTCCAGCCCTGGGCTCAGGGCTGCCCTCAACTTCCCTCCACCAAGGAGGCTGCAGCACTGACTCCCCAGC 8480
 XM_031061670.1 PREDI ATAGCCACTGAAGGTGATTAGGCTCACCTGATGCTGCACATGTGGTGTGCTGCTGCTGAACCTCAGTTAATGG 8459
 XM_037906217.1 PREDI -----TT-----A-CC-CA---C---GT---A---AC-CT-----C-T-----CT---CTGA---A-G---CA-CCT--- 7655
 XM_002939080.5 PREDI ---CCT-C---ATAGTCA-ACCTCA-A-ATAA---CTG-----TAAT-T-----T---T-A-C-C-CT---A 7772
 XM_026853751.1 PREDI TG---G---C---A---TCCA-GA---C-----ACTT-----G-----T-----G----- 7444

XM_025192633.1 PREDI AGGCTTCCAGCAGCAGTGGCATGCCTGAAGCCAGGAGCTGGCGCAAGTGGCAGACCAAATAGCTCAGAAACCACTGTCT 8560

XM_031061670.1 PREDI CGCCGTATCTAAAACCGGCAGCTGGATTCAAGTAGCTGACAAGCAGTTCAAAACCATTCCCTGACACAGCTAGCAAG 8539
 XM_037906217.1 PREDI T-CT--AC---T--AGTT---CT-C-G-AGTTC---AA-GTTGCC-T-TGT---CTTGACA-C-AG-CTCAG- 7704
 XM_002939080.5 PREDI -----TTG-----C-T-----C-TTCTG-----ACA-CC-C-----CCTTGGT-----A-C--- 7798
 XM_026853751.1 PREDI -----C-----A-----GC-----A-G-----G-----CTA-G-----T----- 7456

XM_025192633.1 PREDI AATGCAGTTGCTTAACAGGTTGGGGCTGTCCTCTGCTTGTGCTCAGCCTGTTAAAGATTCCCAACTGAGGTGCC 8640
 XM_031061670.1 PREDI TGACATTGGGGCCAGAGTCTGTGATGCACCTCAGGACATATCAGAAACTCCTCTGTTGGGCTGTTGCAGGACTCTAGTG 8619
 XM_037906217.1 PREDI AGCTG-C-A-C----C-CA-CATT-C-CT----AGC-----C-----C-T-G-C---TC-CT---T-C-CCTC 7740
 XM_002939080.5 PREDI ---C-AT----C-C-----CA-G-----A-T-T-CT-TC-CCT---TT-T-C-C-C-CA-C--- 7825
 XM_026853751.1 PREDI -----C-----A-T-----G-----G-T-----CA-----G-A-A-----TG-----T----- 7470

XM_025192633.1 PREDI CCACGTGCTCTGTGGGACGGTAGTGCCTTGTACTGCAGCTCCCCGATATTTGGAGTTACAGAACGCCAGCTGTCGC 8720
 XM_031061670.1 PREDI GGATTTTATTGCAGTTGCCCTACATACGAGTAGGTTACAGACCCCAATTTCATGGCTTCACTGGATGAGAGAACT 8699
 XM_037906217.1 PREDI --G-----TC----C-ACTCT-T-GC---T-CA-G---C-A-AGCG-----A-GC---TGGACCT-C---C---C 7775
 XM_002939080.5 PREDI AA-----C-TAT-TGGTTC-TGAC---TGTG-ATCTG-AGCTATC-A-C-----CCTC-CCA-T-GGCT-----C--- 7872
 XM_026853751.1 PREDI -GG-----CT-G-----C-CA---T-----G-----T-A-T---C---AGA-TT-C-C 7491

XM_025192633.1 PREDI TGAGCTTAGACAAACAGATGCTTTCTGCTTGTCTGAAAGAGTAGTGATTCCCTGAAGGGCTCTGAACCAGTCAGCAG 8800
 XM_031061670.1 PREDI TTTTGCCCTGTGTCATGAACTGTGCACTGTCTGCTAGAACAAATGCTTGTAGGGGCTCTGGACAAATCAAGGGAGA 8779
 XM_037906217.1 PREDI A---A-A-ACACC---TGCA-GG---TCA---A-GTGA-CAC---CT-CC---TGTC---A-C---GCA-G-CT 7816
 XM_002939080.5 PREDI -C-----C-CT-AGAC---CAAT-GT---CCT---G---AACTC---C-C-C-CT---G---TCCA---T 7906
 XM_026853751.1 PREDI T-GAT---A-G-----A-----GCAC-G-GA-----T-C-----A-----CTG---GC-----T----- 7514

XM_025192633.1 PREDI CTGAAGTGAACATTCTGGGATGGTTGCTCCTACCTCTGGGGAACTCTGCAGACACACGTTCCAAGAGATTGGGG 8880
 XM_031061670.1 PREDI AGTGTAAATGTTTAGGAACGTCTTAAATTCTACCTCTTGTAGGAGAACCTGGCAAATGAACATTCAAAGAGACTTGAG 8859
 XM_037906217.1 PREDI --T---C-A---CCCCT-T---C-CT---C-TTC---A-G-----G-----A-G---T-C-----TT-TG 7843
 XM_002939080.5 PREDI TA-A-C-----CAAATCG-CC---C-C-----ATAGACA-----TG-A-GGACAGCTGTCTA-GAG- 7948
 XM_026853751.1 PREDI -----C-----T-----T-----CT-----GTTC---CTAT-----A-GCA---A-TGAT--- 7536

XM_025192633.1 PREDI CTGCGGCTGCCCAAAGGACAGAGAACGGCACTCCCTCCAAGCAGCGGCCATTCAAGCATCTTACCATGTACAGTATT 8960
 XM_031061670.1 PREDI CGTTTATACTGCCACAAAAGACAAGGATCCCGCTCCTCTGAGCAGCCATGTCTGGCAACTAACCATATGCAGTAT 8939
 XM_037906217.1 PREDI CCT-TGTA-GAG-ACC---C-ACT-AACA-TCC-----C-TT-AGCG---T-T-CT-----C---AC---TG---CCT 7886
 XM_002939080.5 PREDI CCAC---CT-CC---CG-C---CA-C---TGC-----C-T-----A-GCTCT-T-----C-C---C---T 7978
 XM_026853751.1 PREDI G---CA---T-CAA-----C-ATGATG-G-ACAG-----T-T---A-TC-----T----- 7561

XM_025192633.1 PREDI CTCTCTGATGTTCTCCCATTCTAATCAAGAGTAAAGCTTTCTTTGATTCAACTCAATGGTACATGAAGGCAGGG 9040
 XM_031061670.1 PREDI TTCTTCAGTGTGATGCTTTAATTCTAGACGAGAGTAAAGCTGATTCTGACTCAATGGTACATGAA 9019
 XM_037906217.1 PREDI GCT-C-C-C---CAG---G---T---A-T---TATG-AG-A---TT---A-T-AT-A---T-A-C---A-T 7918
 XM_002939080.5 PREDI -----A-----A-----TGC-----AGTGCC---TA-G-G-CC---CTCCTGAAC---CCTCGCCTAC- 8014
 XM_026853751.1 PREDI --C-----AA-----A-----CT-A-----TG-----C-TC 7573

XM_025192633.1 PREDI CCTCAGGAGGTGCTGGTCCGGCGCTGCTGAGTTAGCTGCAAGTTTCACTGAAAGTTTCTGTTGACTGCTTTTATGACGCAC 9120
 XM_031061670.1 PREDI GGCAGGGCCTTGCACGAATCATTCCCTGGGCCACATTGGGCTGAAGAATTTCCTCTTTCTCTTTTT 9099
 XM_037906217.1 PREDI ATAGT---CT---TTTTAGTTG-TT-TG---T---T---T-----G-----TT-----AT-TTCCT---CC 7957
 XM_002939080.5 PREDI ATGAGTG---CT-T-GT---AC---AATGTTG---ACT-CT---TC-C---AGCAG---G-----CA---GCCCT--- 8057
 XM_026853751.1 PREDI -----ACA-----GC-T---T-----A-----C-----C---A--- 7584

XM_025192633.1 PREDI ACAACGTTCCCGGACGCTCGCTGAGTGTCTGTGAAGGTTAGCTGCAACCGACTCAGATCTCAGGGCAGGGGCC 9200
 XM_031061670.1 PREDI GATTCTCTAGTCCTGCATTTCGATGCCACAGCATTCCCTGCCGACAAATGAGGTGTACTGTGAAGTAGAT 9179
 XM_037906217.1 PREDI -CA-----AA-AA-----CGC-----T-GGTC-C-----GA-C-C-T-G-----G-----TCTT-C---C-C 7986
 XM_002939080.5 PREDI -G-A---CTT---A-T-T-----CT---C-AGA-CA-AGT---C-----TA-----CACG- 8083
 XM_026853751.1 PREDI -----T-C---C-CC-----T-----T-----CA-----GAA-----CCCA-ATGG---GA-----A- 7607

XM_025192633.1 PREDI TGCAGGCTCTGGAGGTCTGTCAGGTACATGTTCACTGCAAATAATGAAGTATTGCTACATCGATGCTAAAGAAAAAA 9280
 XM_031061670.1 PREDI GCAACTGACTCAGATTCCAGGCAGAACGCCCTCAAACCTCGTGGAGGTCTGGGACTTGTATTTATCGCAGATAATC 9259
 XM_037906217.1 PREDI AC---CA---CAGTG---GC-CC---AT---G-G-G-C-----CCA-T-CTC---TTCCCTG-AG-A-CAC-C 8026
 XM_002939080.5 PREDI -----G-----CAT-TCT-CCT---CCTC-----CGAA-TA-----TC---T-GGT-C--- 8110
 XM_026853751.1 PREDI ---A-ATATCCAAC---CC-CACCA-----T-----C-AC---A-AC---AC-A-G--- 7635

XM_025192633.1 PREDI AGGGCTTCTCCTAGCCCCAAATAATCAAGTGCCTTTTATATATGTTGCTACCTTTACAAGATGAGTATTCTG 9360
 XM_031061670.1 PREDI AAGGGTAGAAAATTATTACTGAAGAAACCAAGGGCGTTTACCTTCTGCCCAAATATCCAAGTGCCTTTAATATT 9339
 XM_037906217.1 PREDI ---C---C-T-GA---C-----C---A---TTT-T---G---A-C-----A-C-----ACT-----ACC--- 8049
 XM_002939080.5 PREDI ---AGA---C-T-----TC-TAC---CTCA---GC-TTC---T---AC-CC---T---CCACG---C----- 8141
 XM_026853751.1 PREDI -CT-----A-----CACT-----TCCA---CTGATA-----AAA-C---CC---CAAAC 7663

XM_025192633.1 PREDI TGGAAATGAAGCCAGTTTGTGCCCTGGCGTGTAAATACGATGCCCTCTGCCCTTTTTTTAACTATTGAAAGATA 9440
 XM_031061670.1 PREDI GGTGGCCATGGCAAGGCAAGCATGCTTTCACAGAATGAAGCGGTTTGACCTGGAGGGAGAGGGTAATATGATG 9419
 XM_037906217.1 PREDI --C---T-CC-CCGTGCC---CATCC---TGCC-----A-GG-AG-CAT---CCATG-A-C-TGC-C---T-C- 8092
 XM_002939080.5 PREDI ---TCCCA---C-C---TT-CC-----A-CA-GC---C-AAGT-CA-GAC---C-----CCTCCT--- 8174
 XM_026853751.1 PREDI ---C---AT---A-A---T---GGACA---T-T-T-----A-----CC-CG-----G----- 7683

XM_025192633.1 PREDI ATATCTCTTAAGGTGATATTCTCTTAAGTTGAGAGTAACCATTACACAGCAGCATTTCCTATATTAATAATG 9520

XM_031061670.1 PREDI GCCCTTCTGCCGTTTCGTACAACCTTGAAAGGTGATGATACTCCTGGTGGTATTGATTGCCTATGCCTGT 9499
 XM_037906217.1 PREDI C---AT---G---GG---CT-C-CC---A-TGCA---G---A-A-GA---G---ATG-CAC---T-G---CAAC--- 8127
 XM_002939080.5 PREDI ---TC---A-C---ACT---A---C---T-CTGCCGTG---A---ATGCC--- 8198
 XM_026853751.1 PREDI -----A-G-----A---GC--- 7688

XM_025192633.1 PREDI GAAAATATTCGTTATTTTACATTTTATAAAAGTCTTTATGAAAATAGTGTGCAAATGTATAAGCAATTTTA 9600
 XM_031061670.1 PREDI AGAGTAACATTACACAGCAGCATTTCTGTATTAATAATGAAAATATGTGCTTATTTTACATTTTATAAAATC 9579
 XM_037906217.1 PREDI -----TGATA---C---AGGCT---T-TTG---G---G---AT-GAG-----CAC-G-AG-----TG- 8158
 XM_002939080.5 PREDI -A---GG---G-C-T---CA-G---AGCCAGC-C---CAC---A---G-AGG---CACC---TTTCTCA---A-A 8237
 XM_026853751.1 PREDI A---CC---CTTA---CT---TG-A---C-CC---C-T---T-C---AC---CA---G- 7712

XM_025192633.1 PREDI TCCCCAGAATAAAGAGTATGTGCCGG----- 9627
 XM_031061670.1 PREDI TATTATGAAAATAGCTGTTGCACCATGTATAAGCATTTTATCCCCAAAATAAACAGTACATGTCCAGCCTGGGTG 9659
 XM_037906217.1 PREDI --T---T-CAAAA-A-CTGC-TA---AC-----A-C-----A-AGC---T-C-CC-----AG 8187
 XM_002939080.5 PREDI -GA---G-----GAAAC---T-G-C-C-----CAGAAAAGA-A---GC---CA---TAGT-TTAG---G---AAA 8275
 XM_026853751.1 PREDI -AAT-C---TC-C---AG-A-----TCA---G---T---G-----GT-----C---C-A 7733

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI CAACATCTCACTTCCTACCATAGGCATATGCTAGAAGGTACCTCTGACTTGCTTGCCAATTCCCTCGTGGCGTTGTC 9739
 XM_037906217.1 PREDI C-GTTCT-T---GC-----ACAGT---A-GTCC---CCTC---CC---C-TC---CA---T-CCA-AGG-----G-T 8226
 XM_002939080.5 PREDI ACCCT-CAG---CTTAAGCA-TG-AG---GAAGT---T---T-C---CTGACCT-C---A---CAC----- 8315
 XM_026853751.1 PREDI G---C---T---C---CT-----CAC-----C-TAACT---CT- 7750

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI TCAGCATGAAGAACCCATTAGCGTCCCTAGACTAACAGGCTACACAAATGGCCCTCCCATAGTGAAGCCCTCAGCTCCG 9819
 XM_037906217.1 PREDI -----GG---CTG---TG---TG---TCCT-GT---C---CTC-CAG-AT---A-CC---A---T-A-T 8257
 XM_002939080.5 PREDI -CA-C---A-A---AT---TCC---TG---ATCCCAT---TAC-----ACAGT---A---TC--- 8345
 XM_026853751.1 PREDI -GCCTC---TG---A---T---T-GGT---C---AGA-----T-GTA---T-GTA--- 7771

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI CCTGATGTGAAGAAATCTCTAACGGGTTCCCTCAGACTTATTTGAGCTGTTAGGAAGGAAATTGCCACAGTGCA 9899
 XM_037906217.1 PREDI ---G---C---T-G---CCT---C-T---T---T-A-----TG---GT-----A---T-T-A-----TCA 8280
 XM_002939080.5 PREDI ---ACAGC-T---A-AGAGT---CAAAGG---TG-T---TTG---ACTG---TG---G-----T-G---C--- 8379
 XM_026853751.1 PREDI ---A---C---T---A-----CAAG-----TCCGA---C-C---CC---TA--- 7790

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI GAACAAAGCTATGAGACTCATCTGGCTGAACAGGAGCTAAAGGATCACCATGGGCTGTGGCTTGTGAAGTCCCTTG 9979
 XM_037906217.1 PREDI -AC---A---AG-TG---C-----CAA---A-C-GT---G---G-T-T---A---CA---T---G-T--- 8305
 XM_002939080.5 PREDI -TCTC-GAG---T---CTTGCA---GGT---T---G-----GGT-G-TTGTGTC-T-CT---CT--- 8414
 XM_026853751.1 PREDI ATG---G---CA---A-----TC-----A-----A-----A 7802

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI TCCCTAGAGAAGGTCACTGCCTCTGGAGATCCCTGTCTGGTGGAGTGCTGTTCCCTGTTCCCAAGTAGCCTCAGTA 10059
 XM_037906217.1 PREDI -----T---T---GC---G---GT---G---TGT---GT-CC---T---T---CT-----G-CA-----TCAGGA-A--- 8334
 XM_002939080.5 PREDI -----GTG-CG-----T-A-----T---C---A---TGAGGTG-----GC-A-----A--- 8435
 XM_026853751.1 PREDI -----G-----A-----ATG-C-----C-T-----GCCA----- 7814

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI GGAGTTGGAGGGAAAAGGGATAAATAGGCCTACCTGCCAGCTCAGCTTGCAAGGACATCACCTGATGATCTGTAGA 10139
 XM_037906217.1 PREDI ---C---A---AC---ACCC---AGCC-C---T---TT---GA---G---AA---G---T---G--- 8358
 XM_002939080.5 PREDI ---G---G-T-T-GTAATCA-T---G-T---AT---AT---GC---AG---CAT---AT-T---TGT-CC 8468
 XM_026853751.1 PREDI --CAC---A---TATG---C---CT-----G-AACAAT---C--- 7833

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI AAATGTGGTAAGGGAGACATGATGCAGGGAGATGCCCTGGTGGTGTGATGCCCTGACTGCTTAGTCTCACATCCTGTAT 10219
 XM_037906217.1 PREDI --T-CC---CT---TT-TGC---TG-CA---GCC-----C---TG-GGT-----C---A---CA---GCG- 8388
 XM_002939080.5 PREDI --T---T---TG-TA---CCA-TTCA---CCTGTGTTA-----A---TGTGACCC---TCAATC- 8506
 XM_026853751.1 PREDI ---ACA-CAATG---G---CA---T-G---CA---G-----A---T-----TTATGCA-T---AG--- 7861

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI CCATCACACTTACACTTCACACAGGTTCTTTATTTGTTGGTGTGATGCCCTGACTGCTTAGTCTCACATCCTGTAT 10299
 XM_037906217.1 PREDI -GCCT-GT-G-C---TCCC---C-----TCG-CAGCGC-A-C-GG-----A-GGC-CGCA-G---C--- 8424
 XM_002939080.5 PREDI C-----T-G---G-----GATAGGGCATATTGCC---C---CTT-T---T---A---TGCC-----C- 8539
 XM_026853751.1 PREDI --G-A-----C-AC-G---GT-----GC-----A-----AT---T----- 7875

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI GAGAGATGCAAGAACATCAGTGCACCAAGACAAAGGCCATCCTGGAATAACTCAGAAGGGAGCCTGGGGTGGTGC 10379
 XM_037906217.1 PREDI A-----CTG---ACTG-CC---CTGC---AG-C-C-C-TCC---A-G-GAGCTG---G---GC-T---GTG-----G-C 8463
 XM_002939080.5 PREDI -----TTTGGAGGT---GATC-AGA-ACA---GCCTAATATAGTG-CTGT-CC---AC-----TAG-AAG-GC-A--- 8588
 XM_026853751.1 PREDI --T---A---CGT-TTA-----ATCGT-----T-----GA-A-C-----A-A-A-T---GAA-G--- 7901

XM_025192633.1 PREDI ----- 9627

XM_031061670.1 PREDI CAGGATGGTACTGCACCTCTGGCATGAGATTCGTGACTGCTGAAACACAGGCCATCTCTGTAAATGCCATTT 10459
 XM_037906217.1 PREDI --A-GCCGACC--AAATAGC-TCAG-A---A-----ACC---A---T--TGC-CTC--CT-GC---AG--- 8501
 XM_002939080.5 PREDI --AG-----AATGCTAG-A-G-TC-TTCTTCCTAGGCT-TTA--G--TTAGTTGC--A--C-CTCAT-C- 8635
 XM_026853751.1 PREDI -G-----CT---A-T---TG---G-----AG--GC--AAT-T--TA---T---ACTTTCA----- 7927

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI GTTCACCTCTGGGGCTGCATTTCACTTGATTCCTGTGCTCTCACAGTCATTACTTAAGAGCAAGGGATTCTAAA 10539
 XM_037906217.1 PREDI C---CTGGT---C-ACA-G-----C---TCTG---GGC---TG-T---C-CT---CTGC---C-C---T- 8535
 XM_002939080.5 PREDI --TA---AAATTG---GAAG---T---TG-----GTTAAGTTGGAGACC--AAATAG-TTCAG-A---A--- 8678
 XM_026853751.1 PREDI -----GAAG---T---TG-----CT-----T-C-----A-----A 7939

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI GCAAGACCTTAGGGTGGCAGGCTGGTGGTACCCCTGGTATGCTGGCATAACCGAACTACTGTGCTTTAGTAACGT 10619
 XM_037906217.1 PREDI C---AG---C-TC-CAG-----C---C-T---GAGC-AGAG-CTCT---C---CC---A---G-CCCAGG-GG--- 8573
 XM_002939080.5 PREDI ---ACC---A---T-C-AT---CT-G---A---AAC---ATCT-GAG---C-AGATG-----GAT-T-----C 8711
 XM_026853751.1 PREDI -T---C-----T-----GC-----A-----G-CATC---T---G-T--- 7953

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI CAGACTTGCAGGATAGGGCCTTGAGGATCTGAAGAGCTCCTGTAGCACAGTGTACCAGGTATATAATTACCGTATTAA 10699
 XM_037906217.1 PREDI CC-----T-CT---G---T---G---CT-C-TG---T-G---G---GA---CA-G-TAGGGCC-----GCTC-G- 8606
 XM_002939080.5 PREDI --ATTGAT-----T---CTG-----C---AAT-----A-TGC-T-TCAG-----C-ACAT-----A-CCA-A-AA 8746
 XM_026853751.1 PREDI -----T---TAT-----G---T---A---T---T-T-CC-----T-----TG-TA-----CCA---C 7974

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI TAGGATAATTCAACCCCTCTATGATGTACGATCAATAACAGTAAAAATGTTCAAAATGTATGATAATTGGAGGCAGC 10779
 XM_037906217.1 PREDI TG-G-TGC-G---G---C---TGCT-CC-----A-T---CG---TT-T-GG-----A---GC---TG-TG---G- 8638
 XM_002939080.5 PREDI CT-----G-CACA---CTTG---TG-----ATGCT---G---T---TG---C-CCATG-----T-GAT 8778
 XM_026853751.1 PREDI A-----CAG-G-TA-----CCTA---T-T-----TCAAT-----T-----G-CA-G 7997

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI TAATTGATGCAGCATATGCAAACACTCAAGTCACTCTCAGGGAGTCTTACTGACTCTGGGTTTGCTCTTATGACTG 10859
 XM_037906217.1 PREDI -----C-----A-G---C-----CA-G-----TT---GGCA-C-----TGA---C---C-T-TG-----G-CA- 8663
 XM_002939080.5 PREDI CCCACAAGA---CAAG-TA-GG-----AAT---T---ATAC---TG---C---AGCTTC-----T-CTA 8817
 XM_026853751.1 PREDI G-----TG---C---TCTT-T-----T-----TTG-----TT---A---T-----AT---A-ACA--- 8020

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI CTTGCCTCAAAATGAAGAAATGCTCAACTTGATAACTCAGTTCTCACTCATTCACTGGTAGTTACAGTGAGCAATAT 10939
 XM_037906217.1 PREDI -CACA-GCAG-C---A-C---CTTC-----TG-----C-T-T---TG---T---C-C---A-CAAG---A-G---C-A---G 8698
 XM_002939080.5 PREDI AATT-T-----TTA-----G---TG---G---G---TTTAC-----A-AAAC-----CCA-G-----CTT--- 8847
 XM_026853751.1 PREDI --C---A-AGTG-TTGT-----AC-----AG---A---C-----TG-----TACTA---AAG--- 8046

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI AAATTAAATAAGAAGACTTGTGTTTCAGACTTATTGATATTTCAGTGTACCATAGTTTCAGCAAATACAATAAT 11019
 XM_037906217.1 PREDI -----C-G---CT---TC---C---G---A-AG-G-GC-----T-----C-----CAG-----A 8720
 XM_002939080.5 PREDI TCATC-----GA-GC---C-T-T-----AACTG-GA---CA-A---CAG---AGC-TGCTTC-----T-----CCTCT 8887
 XM_026853751.1 PREDI ---CTT-C-----C---T---G---T---CTG---A-----A-C-----AAA---G---G-A---T---CG---AG 8071

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI GCCATTACGCTGTTGACTTGGGATAGTTGTTACTGAAACTAAAGTTCTGTATCAGTGGGCCCGCAGCAAGGAA 11099
 XM_037906217.1 PREDI ---C-----CA-G---CC---AG-----CAGC-----C-----G-A---A---G-TG-AACA-TT-G-T-G---G 8749
 XM_002939080.5 PREDI ---TG---T---C-C---A---GATCTCTG-TACTGCCCTGTGAGAAC---ACT---TC---C-C---TG---CAG-G-GC 8933
 XM_026853751.1 PREDI ACTGAA-TT-CATA-C---AAT-G-TCTAGTA---A-----A-----A-----A-C-----A 8101

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI CTTTGATCTGGCCACTGTTGGCAGTAACCAACAAATTGTTGAGGAATAAAACTGAATAAAACCCACCACTTAATCA 11179
 XM_037906217.1 PREDI -----G-GA-T---G-GT-TTG---C---T---CC---TA-C-----C-----TCTT---CA---G---GGG-AG---T- 8780
 XM_002939080.5 PREDI -----T-----CTG---GA-----CA-AG---TC-A-----AG----- 8948
 XM_026853751.1 PREDI -GAA---A-CT---G---C---T-----CAGA---T-----C-AATT---T-----T-----C--- 8124

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI ACCAGAAAGAAATGTAAATAAAATGTGCGTGTGCTTAGGATATTAAATGCAGCCCCAATGTTGGGCTCCTATGGAAGTTTC 11259
 XM_037906217.1 PREDI GC-----TG---CA-G---ACAC-G---T-----C---T---T---CC---AAAG---A-G-----CT-CG---G 8809
 XM_002939080.5 PREDI -----G---A---TA-GAAG-TGTAAT-TT-CCTAGAAG---TCG-CA-T-----TAAT-----T-----C-TA-C- 8987
 XM_026853751.1 PREDI AA-----CA-T-C---T---G-----GGC-----GC---A---AA-----G-AACA---TT---T- 8148

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI TGAGACTGGGAAGACAGCCAAGTCTGCTGTGAGAGAGAAGCAGGTCACTGGAGAGTGTGCTTGTGGGAAGGACATGG 11339
 XM_037906217.1 PREDI -----GG-----T---G-CC-A-GCT---GCTGCCA---AA-----G-GACT---GAAG---GCC---G---CTGCCCT--- 8848
 XM_002939080.5 PREDI -----C-----TCTT---CA-----GG-AG---AT---C---CCTG---CA-----A-A---T-----GAAC-AT 9016
 XM_026853751.1 PREDI -----T---T-AG-----TG---CC-T-----T---AT-TTA-GAGA---CAAG---GG----- 8173

XM_025192633.1 PREDI ----- 9627

XM_031061670.1 PREDI CAAGTGGCCTCCAAGTCATGCTTCAAAGAGCGTGGCCTCAAGTCATTGGCTAAAGGGCTTGGGCAGGGCTAAG 11419
 XM_037906217.1 PREDI --GCG--AGC-CG-C-C-----A---T-C-----CA---ACA---C-----CCT---A-C-CG-TGT---ACAGTA 8883
 XM_002939080.5 PREDI --T---CC---AAG-GA-G-----AT-----TTTGAGCACTA-TTTTGCTGCCA-ACAA 9054
 XM_026853751.1 PREDI -----T---C---AG---T-C-----C-AAA-----T---G---T----- 8186

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI TGGTGCCTGGGAAGGGTCTGAGTCCAAGGGTTAGGAGAGTGTGTTACTGTGGGGTCACCTAACCTGTG 11499
 XM_037906217.1 PREDI TTG-C-C---TC---T---GC-T-----CTT---C---CCTCCCA-T---T-CT---AA---T---CAAGA-GTC-A--- 8921
 XM_002939080.5 PREDI -----AAGAC---AGGGAGTC-TG---CT---CCT---CTCTGAGC-AG---C-----A-GTTAC---TGC-A--- 9095
 XM_026853751.1 PREDI -G-----AGA-----TA-G-----CCTT---TCA-----G-G-----A----- 8203

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI CTTGAGACTTCAGTAAGGCTGTACAACCTCTGAGAATGCCCTCCATGTTGGCTCCACTCTGCTTTGAGTTGCCCTC 11579
 XM_037906217.1 PREDI -AA---G-T-----TGC---T---T-T-----G-----CT-TG-TT-----TT---T---G---T-T-----T-T- 8946
 XM_002939080.5 PREDI CA-----TTTTA-C-C-AGGT-----ACAGTATT-----TCGAGT---G-AT-----ACTT-----T 9130
 XM_026853751.1 PREDI -GA---C---A---TTG---CA-C-TGC-TC-----ACAC---TG-----GCGC-----G---TCTG-TACT--- 8237

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI TTTCTGGGAACCTCCCAGTCTGGTTACCATCCAGCAGCAAAGTTAAAGCTCCCCACAGCAGGGTCCATTAATGCCAT 11659
 XM_037906217.1 PREDI -G---T---T---T-TGTTT---TCCA-----A-----TTC---C---A-C-CCCA---TGG-TA-CAT--- 8978
 XM_002939080.5 PREDI -----T-GTCCTAGTAA-AGT---TAGGA-G-C-A---CAA-----TGC-----AT-T----- 9161
 XM_026853751.1 PREDI --T-T-----TCT---T-----GC-----T---CA---C-----CTG--- 8252

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI ACCTTTCTCTGGAGGGTTCACAGTAAGCTGATCCTCAAGACCATTCAAGAAGTGTGATTGAGATGGTGCTGGAAACC 11739
 XM_037906217.1 PREDI -----G---AAG---GCAGGG-CC-----T-C---A-----GGA---G-----G---AGCTG--- 9003
 XM_002939080.5 PREDI -T-T---CT---TT---T---G-----AT-----T---T---T-TG-----A----- 9177
 XM_026853751.1 PREDI -GA-----ATA---TC-----T---GAAGAT-CAA-A---T---C-A 8273

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI TCTGAGTTTTGCCTTCCACTTCAGTTGTTCTAAGCATCAGTTAGCACCTTAATGGCACTGACACTGATCT 11819
 XM_037906217.1 PREDI -GTT---CC---C-----G---GG-GCC---G-----TC---G---CTTG---GGCT-GA---C-G---T 9033
 XM_002939080.5 PREDI ---C-TC---A---A---AGG-TA-CGT-----G---AAG-----GCAGGG-CC---CTG 9205
 XM_026853751.1 PREDI -ATG---T-G-CA-----AAT-----T---C-A-----C---ATTCA-TCCA-----GTAG--- 8301

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI CCGTGAACCGAGTCATGCTGAGTTGTGCTTGTGCCATCAGGGCCAGGCTTCAAGAACATTGATCTTG 11899
 XM_037906217.1 PREDI T-TT-T-AG---T---TTT-----T-A---G---CTG---TG---T---GTG---GA---GC-TGC---CGC---T 9066
 XM_002939080.5 PREDI GC---A---GGA-----A-T---CATT---C---C-----G-TGG-GCC--- 9226
 XM_026853751.1 PREDI AT-----GTG---C-----AAC---T---T-----G----- 8313

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI TAGATTGAGAAGGGAGGTTCTGTCGGGTGTTCATCTGTTAGGAAAGGTGGGTTGTCACCATGACAAGGTTAGCAA 11979
 XM_037906217.1 PREDI TT-T---CC---C---GGCGCA-----C---AC-A---G---C-----GTTC---C-CG-G---C 9093
 XM_002939080.5 PREDI -----AC-----CTTT---CTGGGCT-GA---AG---TT-GT-T---G-AT---TCT-----T-----T-----T 9257
 XM_026853751.1 PREDI -----TT---A---A---TT---AGCC---AC-----AGGC-----GGG-G---AG----- 8335

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI CATTTAGCTTGCATATGTAACCTTTGTCGTAAGCTATGGTAAAACACTAAACAAATCAATCCACCGGCCATAAGGAT 12059
 XM_037906217.1 PREDI A-C-----GCT---G---G---C-TG-C-----AG---C-----G---G-----G---G-C--- 9112
 XM_002939080.5 PREDI G---T---T---TA---AT-----GC---T-TT-T-----TGATGCA-----C---AC-A---G--- 9282
 XM_026853751.1 PREDI C---AC-CT---G-TC-----A-G-T-C-----C---CA-----TTTA---CAAGG---GA 8361

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI GGTATGAGTCATCCCTAGGATTTGATTAATTGCTCTATTAAATTTCAATAGTCAAAAAACAGTGGAAAAGAGCTGACAT 12139
 XM_037906217.1 PREDI AGC---T-G---ACT-C-AGAC---C---T-----CC---AGG---GGTCT---G---T-CC-AGG---T-A---CA---C 9149
 XM_002939080.5 PREDI C-----CTCTCTTC---C-C-CGCCCA-C-----ACT---G-AG---ATG-----T 9310
 XM_026853751.1 PREDI TGTA-----T---GT---G---T-----A-----T-AT-ATA-T-ATA-T-ATA---A---C-T 8388

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI TTTCTGAGGGGTGCCTTGTCTGGTGGATGCCTTGACTGTAATACCAAGGTTGAAATCACTGGGCTAACCTTT 12219
 XM_037906217.1 PREDI -C-----TTCACT-G-C-----A-AA-T-----A-A---TG---A-----A---GTG-----TTG-CTAC-----A- 9180
 XM_002939080.5 PREDI G---TAC---TG---TG---AAG-----G---AG---ATAG---AACT-C-AG-C---C---T-----TT 9340
 XM_026853751.1 PREDI A---C-A-----TG---TTTAAATGT-----AAAATAT---C-----ATT---T---G----- 8415

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI CTGTTCCAGGCCTTAAATCAATGAGCTTAAGTAATTAGGGTGAGACAGCATAGGGAGATGGAACGTGATCACCTCTAG 12299
 XM_037906217.1 PREDI ACC---A-TG---CTA---AA---G---AA---A---C-AAA-----G-----GG-CTT---T---T---GC---T-CG- 9212
 XM_002939080.5 PREDI TGG---GGGCT-----T---GG-----GACT-G-----A-AA-T-----A-A---T 9363
 XM_026853751.1 PREDI AAT---GT---T---T-TG-----T-----A-C-T-C 8429

XM_025192633.1 PREDI ----- 9627

XM_031061670.1 PREDI CTGCCTGCTCCTGGCCTGGATAACTGCAGGTGATGGGTTGAGAAATCCCAGCTGAGTGGTTGTCAGGCCATGGC 12379
 XM_037906217.1 PREDI --TTA-GCT----C-C-A---A-AT--AAC---CA---AG-TGCCT--T--T-T--AT-A-T--A--TG 9247
 XM_002939080.5 PREDI --CA-----A--G-G---ATTG-CT-----TA-A---ATAATG-CCA---AA---G-AA-A---C-CAA 9396
 XM_026853751.1 PREDI -T-----CA-G-C---CT----- 8436

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI ACTGGTGAATGATATAGATAACTGTTTCCTCCTACTGAAGCTCGTCTGCAATTGAGTCAGCTCTACAGTTGTTGC 12459
 XM_037906217.1 PREDI -T-GGCT----A---C-C-----TT-TT-A---CA---A---G-ATG---AG---TG---CTT- 9274
 XM_002939080.5 PREDI -----G---GG-CTT---T---T-GC---C---TTTA-GC---C-C-TA---A-AT--AAC---CCCA--- 9428
 XM_026853751.1 PREDI T-----TT-----AAT---ACC-----A-----GTTT---AT---G-T-CA---A 8457

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI TCTTGACGTGCTCTTCCATAATTAATCTGAGACCTAGAAACACAGCTGGTCTCAAAGGTAGGGTGCCCTGTGGGGAG 12539
 XM_037906217.1 PREDI -T-CT---G-T-G---G-AA---TGA---AG-C-C---G-G---TT-C-C-G-----T-G- 9299
 XM_002939080.5 PREDI ---AG-TGCC---T-T-T-A---A-CT---TTGGC-----A---GC-C---AT-TT-----C 9457
 XM_026853751.1 PREDI -----T---AT---TC-----CT-----T---T---ATA---A-----C 8472

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI TTGTAAAAGCAGGCCAGAGCGAAGAGGATGCAGGGAGGGGATATGAAGTGGTCAGATCTGCCTCCCTCCCTTGGCC 12619
 XM_037906217.1 PREDI CCT-G-----GCT-GT-T---AA---TA-----C---GA---T-----G-GC-----C- 9321
 XM_002939080.5 PREDI A---A---G---G---CAG---GTATG-----T-T-CT-----T-----T-G---C--- 9477
 XM_026853751.1 PREDI -----T---T-AT-----C---C-A---G-----CTC-ATT----- 8486

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI CTGATATCCATGCTGGAGACGACCAGGTTCTGCCA-ATGCAGGGCATGCTATGCACCCCTGGACCCCTAGAGCAGAGTCAG 12698
 XM_037906217.1 PREDI -----C---TCTG-----CCTTT---C---T-T-C-T-CT-----A---T-T- 9341
 XM_002939080.5 PREDI -----AT---C-----T-G---G-----TG-T-GT-T-CGA---A-----C- 9494
 XM_026853751.1 PREDI ---C---C-----ACT---TTTA-----ATG-----CTG---TA-----T---TTAAAAAAAG-----G- 8515

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI CAACACCCGGCATGCGTACCCCTTGGCACATGACCACAGACTGGCTCTAGGCAGCTGCTGCCAGCCACGAGCCCA 12778
 XM_037906217.1 PREDI ---T-----T-T-T-----T-----T-----A-----A-C-TA-C---TT-G--- 9357
 XM_002939080.5 PREDI --GA---T-----G-----TC-----CCTT 9504
 XM_026853751.1 PREDI ---A---A-A---AAGT---TG---AC---T-----TTT-----G-T-----T---TG----- 8535

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI GGCTGCTCCCAGCAGCGGCTGAAGGCTGCAAGCCCTGCTGCAAGCACCCGGGCACCATGGCTGACAGCACGCTACCCG 12858
 XM_037906217.1 PREDI T-A---A---AGA-----T---A---T-CT-----C---CT---T---A---A-GG-T---GG-T---A 9381
 XM_002939080.5 PREDI T---C---TGC---C-----TCT-----T-----T 9515
 XM_026853751.1 PREDI -----G-C-AG-----CA-GA-----A-A-T-----G---TT-C----- 8550

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI GAGCTGGGCCGGCTGCAGGCGGAGACTGCAAACAGCAGCTGCCTCTAGCCTCAGCCCCGACTGCCCTCAGC 12938
 XM_037906217.1 PREDI -----T-T-----T-----T-----A-T-----C-T-----C-G-T---A-----AG-T--- 9396
 XM_002939080.5 PREDI -T-----A---TA---C-T---TC-----T-G---GT-A---A---AG-----CT---GAGAT-CT----- 9540
 XM_026853751.1 PREDI -----T-----A---G-----CAA-----C-----T---T----- 8559

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI CCTGGCTCTGGGGAGGCAGCAGGTGCTGGCGATGCTCCCCACCGTGCTGCCCTCCCGTCAGCTCTGCAGAGCAGCGTG 13018
 XM_037906217.1 PREDI -----TG---TAG---AG---T-A---A-C-CA---T---TC-----A---CA-CA-----GC-A- 9420
 XM_002939080.5 PREDI ---C---CT---T---A---A-GG-T---GG-T---A---T-T-----T-----T---A-T 9559
 XM_026853751.1 PREDI -C-----T-G-----T-----GCC-----AG-TG-----GAT-T----- 8575

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI TGCAACCCGGGGAGGCAGCAGCTGCTGCCGCTGCAAGCTCCCGCTGCCCTGGTCACACTTCCCTGCGTGCACA 13098
 XM_037906217.1 PREDI --T---CA---T---TT---CC-T-A---T-----A---T-T-AA-----TA---ATG-GAAA----- 9446
 XM_002939080.5 PREDI -----C-T---C-----A-----T-----AAG---CAG---T-A-----A-C 9574
 XM_026853751.1 PREDI -CT-GTT-----A---C-----AAT-GT-----C-----A-----C-G-TG-ATG----- 8596

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI GGGGCAGCGCGGGCAGCAGCTGCAAGGTTGTCAGCTGCCCTCCCTGCGTGCCTGCCGCCGGCTCCACAGCAGCTGCTGCAGAG 13178
 XM_037906217.1 PREDI --AT---A---T---TTG---CT---T---A-TTT---T-----A-----A-C-A---TT-T--- 9468
 XM_002939080.5 PREDI -CA---T---TC---A---CA-CA-----GC-A-----GCA-----T---T-----TT---CC-T-A-----CA 9600
 XM_026853751.1 PREDI -----T-----T-----T-----T-----T-----T 8600

XM_025192633.1 PREDI ----- 9627
 XM_031061670.1 PREDI GGAGCTCAGAGCAGCAGGCAGCTGCTGCCCTCCCTGCGTGCCTGCCGCCGGCTCCACAGCAGCTGCTGCAGAG 13258
 XM_037906217.1 PREDI --T---T---T-----AT---AAA---A-GT-----C---T-TTA---T---GA---A---A---A---T---AG--- 9494
 XM_002939080.5 PREDI ---GT---AA---TA---ACA---G-GAAA-----AT---A-T-----GTG---CT---T---A-TTT---T---T 9630
 XM_026853751.1 PREDI -----TG---G-----GA---A-----G-----G----- 8608

XM_025192633.1 PREDI ----- 9627

11/30/2020

Sequence Manipulation Suite

XM_031061670.1	PREDI	CCGGCGGGAGCACGGGCCACAGTGTGCCCCCTGCCCTGCCCTGGCACAGATCTGGGGAGTGCTGGTCCCCCTCT	13338
XM_037906217.1	PREDI	----T---TG--CT-----G--CCA-A--ATG--T---A-----TA--A--G--CA-A-----	9516
XM_002939080.5	PREDI	-----A--C-A-----T--T--T-----A--AAA-AAAAGGT-----C	9648
XM_026853751.1	PREDI	-----TTT-----ATG-----	8614
XM_025192633.1	PREDI	-----	9627
XM_031061670.1	PREDI	GCCCCCCGGGAGTGGCGCTGGGCCAGGAGCCGGCCCACCCGACCTGAGCCCCGAGCAGGCTGCAGCCGTGG-----	13413
XM_037906217.1	PREDI	---T---T---T---T-T-A-----T-----CCCC--A-----G--A--A-----	9531
XM_002939080.5	PREDI	--T-ATTA--T-GA--A--A--A---T-A-----AT--TG--CT-----ACA-A-AATG--T-	9676
XM_026853751.1	PREDI	-----T-----A-----TA-----T-GC-----G	8622
XM_025192633.1	PREDI	-----	9627
XM_031061670.1	PREDI	-----	13413
XM_037906217.1	PREDI	-----TA-----A-----A-----AAGTATG-TGC--CCAT-----	9549
XM_002939080.5	PREDI	-----A-----TA-----A-----G-C-A-----T-----T-----T-T-----TA-----	9690
XM_026853751.1	PREDI	---G--A-CA--TG-----T-CGTGA-----C-GC-----G-----GG-----G	8640
XM_025192633.1	PREDI	-----	9627
XM_031061670.1	PREDI	-----	13413
XM_037906217.1	PREDI	-----	9549
XM_002939080.5	PREDI	-----T-----CCCC--A-----A-----ATA-----A-----A-----GAGTA-----TACGTCCAGTG	9719
XM_026853751.1	PREDI	-----A-----A-----T-----C	8644
XM_025192633.1	PREDI	-----	9627
XM_031061670.1	PREDI	-----	13413
XM_037906217.1	PREDI	-----	9549
XM_002939080.5	PREDI	-----	9719
XM_026853751.1	PREDI	-----T-----AA-----C-G-----A-----GCA-----GC-----CATC-----G	8660
XM_025192633.1	PREDI	-----	9627
XM_031061670.1	PREDI	-----	13413
XM_037906217.1	PREDI	-----	9549
XM_002939080.5	PREDI	-----	9719
XM_026853751.1	PREDI	-----A-----C-----T-----T-GTATT-----T-A-----A-----T-----A-----T-----A	8676
XM_025192633.1	PREDI	-----	9627
XM_031061670.1	PREDI	-----	13413
XM_037906217.1	PREDI	-----	9549
XM_002939080.5	PREDI	-----	9719
XM_026853751.1	PREDI	-----A-----A-----CGT-----CTGAT-----CTTT-----T-----CT-----A-----C-----	8695
XM_025192633.1	PREDI	-----	9627
XM_031061670.1	PREDI	-----	13413
XM_037906217.1	PREDI	-----	9549
XM_002939080.5	PREDI	-----	9719
XM_026853751.1	PREDI	-----CT-----T-----T-----T-----T-----T-----TT-----T-----G-----	8705
XM_025192633.1	PREDI	-----	9627
XM_031061670.1	PREDI	-----	13413
XM_037906217.1	PREDI	-----	9549
XM_002939080.5	PREDI	-----	9719
XM_026853751.1	PREDI	-----T-----CTG-----T-----TAC--A-AATG--CT-----CA-----T-----TA-TTG-----	8728
XM_025192633.1	PREDI	-----	9627
XM_031061670.1	PREDI	-----	13413
XM_037906217.1	PREDI	-----	9549
XM_002939080.5	PREDI	-----	9719
XM_026853751.1	PREDI	-----T-----GT-----GT-----T-----T-----TA-----T-----A-----GA-----	8741
XM_025192633.1	PREDI	-----	9627
XM_031061670.1	PREDI	-----	13413
XM_037906217.1	PREDI	-----	9549
XM_002939080.5	PREDI	-----	9719
XM_026853751.1	PREDI	A-----AGTA-----A-----A-----GA-----A-----CT-----TTTCCAGTA	8762

Sequence Alignment

```

1 #include <iostream>
2 #include <vector>
3 #include <iomanip>
4
5 void print_matrix_salg(std::vector<std::vector<int>> m, std::
6     string s1, std::string s2) {
7     std::cout << std::setw(10) << "-";
8     for (char &c : s2) {
9         std::cout << std::setw(5) << c;
10    }
11    putchar('\n');
12    int i = 0;
13    for (std::vector<int> row : m) {
14        std::cout << std::setw(5) << (i == 0 ? '-' : s1[i - 1]);
15        for (int num : row) {
16            std::cout << std::setw(5) << num;
17        }
18        std::cout << std::endl;
19        ++i;
20    }
21
22 std::vector<std::vector<int>> sequence_alignment(std::string s1,
23     std::string s2) {
24     std::vector<std::vector<int>> dp(s1.length() + 1, std::vector(
25         s2.length() + 1, 0));
26     for (int i = 0; i <= s1.length(); ++i) {
27         dp[i][0] = i * -2;
28     }
29     for (int i = 0; i <= s2.length(); ++i) {
30         dp[0][i] = i * -2;
31     }
32     for (int i = 1; i <= s1.length(); ++i) {
33         for (int j = 1; j <= s2.length(); ++j) {
34             int c = s1[i - 1] == s2[j - 1] ? 1 : -1;
35             dp[i][j] = std::max(dp[i - 1][j - 1] + c, std::max(dp[i
36             ][j - 1] - 2, dp[i - 1][j] - 2));
37         }
38     }
39     return dp;
40 }
41
42 std::pair<std::string, std::string> get_alignments(std::vector<std
43     ::vector<int>> dp, std::string s1, std::string s2) {
44     int i = s1.length();
45     int j = s2.length();
46     std::string ans1 = "";
47     std::string ans2 = "";
48     while (i > 0 && j > 0) {
49         int c = s1[i - 1] == s2[j - 1] ? 1 : -1;
50         if (dp[i - 1][j - 1] + c == dp[i][j]) {
51             ans1 = s1[i - 1] + ans1;
52             ans2 = s2[j - 1] + ans2;
53             i--;
54             j--;
55         } else if (dp[i][j - 1] - 1 == dp[i][j]) {
56             ans1 = "-" + ans1;
57             ans2 = s2[j - 1] + ans2;
58             j--;
59         } else {
60             ans1 = s1[i - 1] + ans1;
61             ans2 = "-" + ans2;
62             i--;
63         }
64     }
65     return {ans1, ans2};
66 }
```

```

47         ans1 = s1[--i] + ans1;
48         ans2 = s2[--j] + ans2;
49     } else if (dp[i - 1][j] - 2 == dp[i][j]) {
50         ans1 = s1[--i] + ans1;
51         ans2 = "-" + ans2;
52     } else if (dp[i][j - 1] - 2 == dp[i][j]) {
53         ans1 = "-" + ans1;
54         ans2 = s2[--j] + ans2;
55     }
56 }
57 while (i > 0) {
58     ans1 = s1[--i] + ans1;
59     ans2 = "-" + ans2;
60 }
61 while (j > 0) {
62     ans1 = "-" + ans1;
63     ans2 = s2[--j] + ans2;
64 }
65 return std::make_pair(ans1, ans2);
66 }
67
68 int main() {
69     std::cout << "Sequences alignment\n";
70
71     std::string s1 = "DYNAMIC", s2 = "PROGRAMMING";
72     s1 = "ABCC";
73     s2 = "BCCD";
74
75     // Table:
76     //print_matrix_salg(sequence_alignment(s1, s2), s1, s2);
77
78     // Alignment solution
79     auto [ans1, ans2] = get_alignments(sequence_alignment(s1, s2),
80                                         s1, s2);
81     std::cout << ans1 << std::endl << ans2 << std::endl;
}

```

Star Heuristic Alignment

Reference:

Singh, M. (2000). Multiple Sequence Alignments I. Retrieved from <https://www.cs.princeton.edu/~mona/Lecture/msa1.pdf>

```

1 #include <iostream>
2 #include <vector>
3 #include <climits>
4 #include <filesystem>
5 #include <fstream>
6
7 std::vector<std::vector<int> > sequence_alignment(std::string s1,
8         std::string s2) {
9     std::vector<std::vector<int> > dp(s1.length() + 1, std::vector(
10        s2.length() + 1, 0));
11    for (int i = 0; i <= s1.length(); ++i) {

```

```

10         dp[i][0] = i * -2;
11     }
12     for (int i = 0; i <= s2.length(); ++i) {
13         dp[0][i] = i * -2;
14     }
15     for (int i = 1; i <= s1.length(); ++i) {
16         for (int j = 1; j <= s2.length(); ++j) {
17             int c = s1[i - 1] == s2[j - 1] ? 1 : -1;
18             dp[i][j] = std::max(dp[i - 1][j - 1] + c, std::max(dp[i]
19             ][j - 1] - 2, dp[i - 1][j] - 2));
20         }
21     }
22     return dp;
23 }
24 std::pair<std::string, std::string> get_alignments(std::vector<std
25   ::vector<int> > dp, std::string s1, std::string s2) {
26     int i = s1.length();
27     int j = s2.length();
28     std::string ans1 = "";
29     std::string ans2 = "";
30     while (i > 0 && j > 0) {
31         int c = s1[i - 1] == s2[j - 1] ? 1 : -1;
32         if (dp[i - 1][j - 1] + c == dp[i][j]) {
33             ans1 = s1[--i] + ans1;
34             ans2 = s2[--j] + ans2;
35         } else if (dp[i - 1][j] - 2 == dp[i][j]) {
36             ans1 = s1[--i] + ans1;
37             ans2 = "-" + ans2;
38         } else if (dp[i][j - 1] - 2 == dp[i][j]) {
39             ans1 = "-" + ans1;
40             ans2 = s2[--j] + ans2;
41         }
42     while (i > 0) {
43         ans1 = s1[--i] + ans1;
44         ans2 = "-" + ans2;
45     }
46     while (j > 0) {
47         ans1 = "-" + ans1;
48         ans2 = s2[--j] + ans2;
49     }
50     return std::make_pair(ans1, ans2);
51 }
52
53
54
55 std::vector<int> calculate_scores(std::vector<std::string>
56   sequences) {
57     std::vector<std::vector<int> > pairwise_alignments(sequences.
58     size(), std::vector<int>(sequences.size(), 0));
59     for (int i = 0; i < sequences.size() - 1; ++i) {

```



```

58     for (int j = i + 1; j < sequences.size(); ++j) {
59         std::string s1 = sequences[i], s2 = sequences[j];
60         pairwise_alignments[i][j] = sequence_alignment(s1, s2)[
61             s1.length()][s2.length()];
62     }
63     for (int i = 1; i < sequences.size(); ++i) {
64         for (int j = 0; j < i; ++j) {
65             pairwise_alignments[i][j] = pairwise_alignments[j][i];
66         }
67     }
68     std::vector<int> scores;
69     for (int i = 0; i < pairwise_alignments.size(); ++i) {
70         int current_score = 0;
71         for (int j = 0; j < pairwise_alignments.size(); ++j) {
72             current_score += pairwise_alignments[i][j];
73         }
74         scores.push_back(current_score);
75     }
76     return scores;
77 }
78
79 int max_length_among_strings(std::vector<std::string> strings) {
80     int max_length = INT_MIN;
81     for (std::string &s : strings) {
82         int current_length = s.length();
83         if (current_length > max_length) {
84             max_length = s.length();
85         }
86     }
87     return max_length;
88 }
89
90 std::string fill_remaining_spaces(std::string s, int
91     remaining_spaces) {
92     for (int i = 0; i < remaining_spaces; ++i) {
93         s += "-";
94     }
95     return s;
96 }
97 std::pair<std::vector<std::string>, std::vector<std::string> >
98 build_alignments(std::vector<std::string> sequences, std::vector<int>
99     scores, std::vector<std::string> proteins={}) {
100    int max_score = INT_MIN;
101    int max_score_index = -1;
102    for (int i = 0; i < scores.size(); ++i) {
103        if (scores[i] > max_score) {
104            max_score = scores[i];
105            max_score_index = i;
106        }
107    }

```

```

106     std::string max_score_sequence = sequences[max_score_index];
107     sequences.erase(sequences.begin() + max_score_index);
108
109     if (proteins.size()) {
110         std::string first_protein = sequences[max_score_index];
111         proteins.erase(proteins.begin() + max_score_index);
112         proteins.insert(proteins.begin(), first_protein);
113     }
114
115
116     std::string current_string = max_score_sequence;
117     std::vector<std::string> result;
118     result.push_back(current_string);
119
120     for (int i = 0; i < sequences.size(); ++i) {
121         auto [s1, s2] = get_alignments(sequence_alignment(
122             current_string, sequences[i]), current_string, sequences[i]);
123         result.push_back(s2);
124         current_string = s2;
125     }
126
127     int max_length = max_length_among_strings(result);
128
129     for (int i = 0; i < result.size(); ++i) {
130         int current_size = result[i].length();
131         result[i] = fill_remaining_spaces(result[i], max_length -
132                                         current_size);
133     }
134
135     return std::make_pair(result, proteins);
136 }
137
138 std::pair<std::vector<std::string>, std::vector<std::string> >
139 read_sequences() {
140     std::vector<std::string> sequences, proteins;
141     std::string path = "genes/";
142     for (const auto & entry : std::filesystem::directory_iterator(
143         path)) {
144         std::ifstream genes_file(entry.path());
145         std::string line = "";
146         if (!genes_file.eof()) {
147             std::getline(genes_file, line);
148             proteins.push_back(line);
149             std::string gene = "";
150             while (std::getline(genes_file, line)) {
151                 gene += line;
152             }
153             genes_file.close();
154             gene.erase(std::remove(gene.begin(), gene.end(), '\r'),
155                       gene.end());
156             sequences.push_back(gene);
157         }
158     }

```



```

153     }
154     return std::make_pair(sequences, proteins);
155 }
156
157 void write_sequences(std::vector<std::string> sequences, std::vector<std::string> proteins) {
158     std::ofstream output("sequences.txt");
159     int i = 0;
160     for (std::string &sequence : sequences) {
161         output << proteins[i++] << std::endl;
162         output << sequence << std::endl;
163     }
164     output.close();
165 }
166
167 int main() {
168     std::vector<std::string> sequences = {"ATTGCCATT", "ATGGCCATT",
169                                         "ATCCAATTTC", "ATCTTCTT", "ACTGACC"};
170     std::vector<int> scores = calculate_scores(sequences);
171     for (std::string sequence : build_alignments(sequences, scores)
172          .first) {
173         std::cout << sequence << std::endl;
174     }
175
176     auto [gene_sequences, proteins] = read_sequences();
177     write_sequences(build_alignments(gene_sequences,
178                               calculate_scores(gene_sequences)).first, proteins);
179     return 0;
180 }
```