Electronic Supplementary Material (ESI) for ChemComm.


## Supplementary information

# A Peptide-based Synthetic Transcription Factor Selectively Activates Transcription in a Mammalian Cell

**Koushik Roy,[a] Abhishek Mazumder[a], Piya Ghosh[b], Gitashri Naiya[b], Basusree Ghosh[b] and Siddhartha Roy[b]***

## Experimental Section

### Materials

All amino acids and Rink amide PEGA resin were either purchased from Novabiochem (Switzerland) or Anaspec (China). TFA, phenol and thioanisole were purchased from Merck (Germany). Alexa-488 succinimidyl ester and lipofectamine$^{TM}$ 2000 were purchased from Invitrogen (NY, USA). Luciferase reagent was purchased from Promega (Madison, Wisconsin, USA). C6-aminolink oligonucleotides were purchased from Trilink (USA) and $O_R3$-$O_R3$ oligonucleotide was purchased from IDT (USA). The sequences of the DNA used in this study are shown in Table S1.

#### *Peptide Synthesis*

Peptide synthesis were followed as described in Dhar et al.,[1] Briefly, the peptides were synthesized in solid phase on Rink Amide PEGA Resin in an automated peptide synthesizer (Protein Technology International, AZ, USA). The N terminal amino group of amino acids was Fmoc protected and other reactive side chains were protected with appropriate protecting groups. Standard Fmoc chemistry was used to synthesize the peptides. DMF and piperidine were kept overnight in molecular sieve to produce peptide synthesis grade anhydrous reagent before use. Rink Amide PEGA Resin were swelled in DMF for 45 min. Amino acids were dissolved in DMF followed by addition of 1.7x molar excess of N,N-Diisopropylethylamine (DIPEA) and then transferred to the reaction vessel containing the resin. The coupling reactions were performed under nitrogen gas. Coupling of each amino acid was performed twice with 3-10 fold molar excess of the amino acid in each reaction and coupling times were 20-45 min. The amount of amino acid and the time of coupling reaction depended on the length of the peptide. The resin was washed 3 times with DMF to ensure removal of excess uncoupled amino acids. Fmoc group was removed by mixing with 20% piperidine for 3-10

min and Fmoc removal were repeated twice. The Fmoc deprotection reaction was performed under nitrogen gas pressure. The resin was washed 3 times with DMF to ensure removal of the excess piperidine. Fmoc-Lys(Fmoc)-OH was used as the branching point for the dimer. The dimeric chains were grown simultaneously through α and ε amine of Lys. The solid phase resin and side chain protections were removed from peptide by a cocktail of 90% trifluoro acetic acid (TFA), 5% water, 5% phenol and 5% thioanisole for 3 h followed by precipitation with cold diethyl ether. The peptides were purified by RP-HPLC using acetonitrile-water mixture (10-60 % gradient of water to acetonitrile with 0.1% TFA).

***Labeling oligonucleotides***

The oligonucleotide bearing 5'-C6 amino link was labeled with 10 times molar excess of alexa-488 succinimidyl ester as described previously [2].

***Measurement of Fluorescence Anisotropy***

The fluorescence anisotropy measurements were performed with fluorescein labeled 1 nM oligodeoxynucleotide duplexes in 50 mM Tris buffer pH 8.0 in a PTI Quantimaster T-geometry fluorimeter at 25°C. The experiments were carried out in a 0.5cm path length quartz cuvette. The excitation wavelength was at 480 nm and emission was monitored at 530 nm, the excitation and emission band passes were 5 nm each. The anisotropy measurement data were fitted to an equation considering 1:1 binding stoichiometry. The anisotropy values were fitted to the following single site binding equation using Kyplot as described previously[3].

$$A_{obs} = A_0 + \frac{[(A_\infty - A_0)*\{(K_d + X + [DNA]) - \sqrt{((K_d + X + [DNA])^2 - 4.X.[DNA])}\}]}{2*[DNA]}$$

Where, $A_{obs}$ is the observed anisotropy, $A_0$ is the initial anisotropy, $A_\infty$ is the final limiting anisotropy, [DNA] is the fluorescein end-labeled DNA concentration, X is the total peptide concentration and $K_d$ is the dissociation constant. All concentrations were expressed in Molar units.

***Isothermal Titration Calorimetry***

All ITC experiments were done in a VPITC instrument from Microcal.Inc. The titrations were done in 50 mM Tris-HCl buffer (pH 8.0). The peptides (8 μM) was taken in the cell (volume 1.47 mL), and either $O_R3$ or C12-$O_R3$ DNA (80 μM) was added from a syringe with continuous stirring in 50 mM Tris-HCl buffer, pH 8.0. The data were fitted to a single site binding isotherm using MS Origin 7.0 software as described previously[4]

***Cell viability assay***

$10^4$ B16F10 cells/ well were plated in a 96 well plate in 200 μl volume. Non adherent cells were removed by washing after 16 h. Adherent cells were treated with 5 μM, 10 μM, 20 μM, 30 μM, 40 μM and 50 μM peptides for 36 h. 20 μl of MTT (stock 5 mg/ml) was added per well and kept for 4 h. Cells were centrifuged and the resulting precipitate was dissolved in 100 μl of 0.1 (N) HCl in isopropanol. The absorbance was measured at 595 nm.

***Confocal microscopy***

$10^4$ B16F10 cell/ well were plated in a 4 well chamber slide (Nunc, UK). 2 μM FITC labelled peptide was added and incubated for 2 h. Then cells were washed thoroughly and fixed with Ultracruz mounting media containing DAPI (Santacruz Biotechnology, sc-24941). The confocal images were captured in Nikon (A1R). The images were captured at DIC channel (cell morphology), green channel (peptide uptake) and blue channel (nuclear marker).

***Cloning of $O_R3$-$O_R3$ variants in pGL3***

The plasmid pGL3 containing the mouse IL-10 promoter (pGL3-IL-10) was a kind gift from Dr. Syamal Roy. The sequence of the $O_R3$-$O_R3$ used in this study is as follows, TATATCCCTTGCGGTGATAGATATTTATCCCTTGCGGTGATAT, with KpnI overhang on both the 5' and 3' side. This $O_R3$-$O_R3$ sequence was inserted in the KpnI restriction site in the pGL3-IL-10 plasmid (Figure S7A). The plasmid containing multiple $O_R3$-$O_R3$ sites is represented as pGL3-nX $O_R3$ where "n" represents 2, 4 and 6 (Figure S7B).

***Transfection and luciferase assay***

The plasmids used for transfection were pGL3 (blank), pGL3-IL-10 (control), pGL3-2X $O_R3$, pGL3-4X $O_R3$, pGL3-6X $O_R3$. B16F10 cells were grown in a 6 well plate with $4X10^5$ cells/ well. The cells were allowed to adhere for 16 h and non-adherent cells were removed by washing. The transfections were performed with 1 μg of plasmid using lipofectamine according to manufacturer protocol. Transfected cells were then kept in resting condition for 12-14 h in DMEM with 10% FCS followed by treated with ATF and incubated for 36 h. Floating cells were removed by washing with PBS. Attached cells were removed from the plate using 200 μL trypsin-EDTA followed by quench with 20 μL FCS with addition of 500 μL media. The cells were centrifuged at 400 rcf for 5 min at room temperature. The cell

pellet was resuspended in 1 mL PBS. 20 µL were kept for cell counting and the rest of the cells were centrifuged for 400 rcf for 5 min at room temperature. The cells were counted by staining with trypan blue. Trypan blue staining showed that > 95% of the cells were viable. The cell count was very similar among the different conditions tested. The total cell count was approximately $2.2X10^6$ cells/ well. Thus, neither the cell viability nor the cell count changes significantly due to transfection with different plasmids. The cells were lysed with 200 µl of passive lysis buffer and 20 µl of lysate were used to assay luciferase activity using luciferase kit/reagent as manufacturer protocol (Promega). The luciferase activity was normalized with respect to the total cellular protein concentration[5]. To further confirm and reduce the effect of experimental variability, luciferase activity was normalized with respect to viable cell number as well. The luciferase activity obtained by cell number normalization and protein normalization produces similar results.

***Whole genome microarray***

To study the effect of the wild type peptide on the global transcription of B16F10 DNA, microarray experiments were performed. The treatments were done as described previously. After 36 h cells were resuspended in RNA later solution. The microarray experiments were conducted by Genotypic Inc. (Bangalore, India) on payment for service basis. c-DNA was prepared from extracted RNA and then labeled c-RNA was synthesised by T7-promoter based amplification. Labeled RNA were hybridized *in situ* with 60 mer oligonucleotides. The hybridisation was carried at 65°C for 17 h. The samples were run on slide, Mouse 8x60K Array AMADID, and the data were collected and analysed.

***Promoter Analysis***

The sequence of the promoter (region near the transcription starts site) of the upregulated (3 fold or more) gene following treatment with Kix-Br-$Ala_{\beta}$-Cro were obtained using UCSC

genome browser. The database provides the information regarding the localization of the gene in the genome such as chromosome number and position within the chromosome. Though promoter sequence of the all gene is yet not characterized. We looked for the $O_R3$ sequence within the promoter with maximum 2 bp mismatch using NCBI sequence alignment tool.

## References


1. A. Dhar, S. Mallick, P. Ghosh, A. Maiti, I. Ahmed, S. Bhattacharya, T. Mandal, A. Manna, K. Roy and S. Singh, *Pep Sci*, 2014, **102**, 344-358.
2. S. Chatterjee, Y. N. Zhou, S. Roy and S. Adhya, *Proc Natl Acad Sci U S A*, 1997, **94**, 2957-2962.
3. S. Polley, S. Guha, N. S. Roy, S. Kar, K. Sakaguchi, Y. Chuman, V. Swaminathan, T. Kundu and S. Roy, *J Mol Biol*, 2008, **376**, 8-12.
4. A. Mazumder, A. Maiti, K. Roy and S. Roy, *ACS Chem Biol*, 2012, **7**, 1084-1094.
5. S. Sen, K. Roy, S. Mukherjee, R. Mukhopadhyay and S. Roy, *PLoS pathogens*, 2011, **7**, e1002229.

## Figures

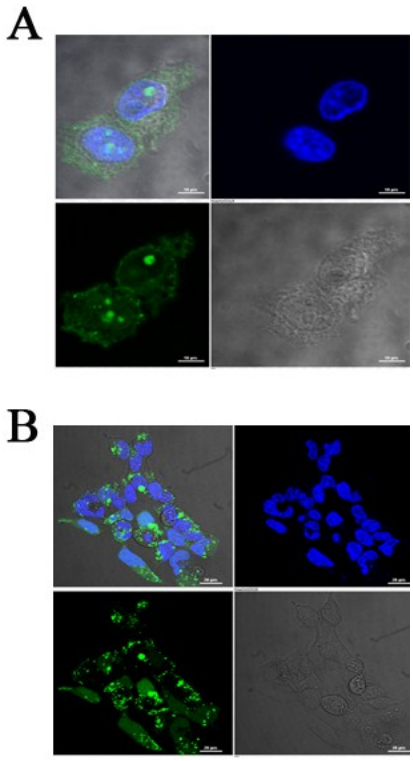



**Figure S1.** Nuclear localisation of the fluorescein label (A) Kix-Br-Ala$_\beta$-Cro and (B) No-NLS- Kix-Br-Ala$_\beta$-Cro. Cells were treated with 2 μM of FITC labelled peptide (green) for 2 h and nucleus was stained with DAPI (blue). Right lower corner represent DIC image, left lower corner represent green fluorescence of fluorescein, right upper corner represent DAPI fluorescence and left upper corner represent overlay of all the three channel.

**Figure S2.** Helical wheel diagram for a single peptide chain of the branched peptide used in the study. The output presents the hydrophilic residues as circles, hydrophobic residues as diamonds, and potentially positively charged as pentagons. The red starred residues correspond to residues which make contacts with the Operator DNA ($O_R3$) in the crystal structure (PDB id: 6cro). The blue starred residues were replaced with Aib. The residues here are numbered starting from residue 26 in the Cro (residue 1 here and so on). Thus, residue 30 is I5, residue 34 is I9 and residue 36 is A11 here.

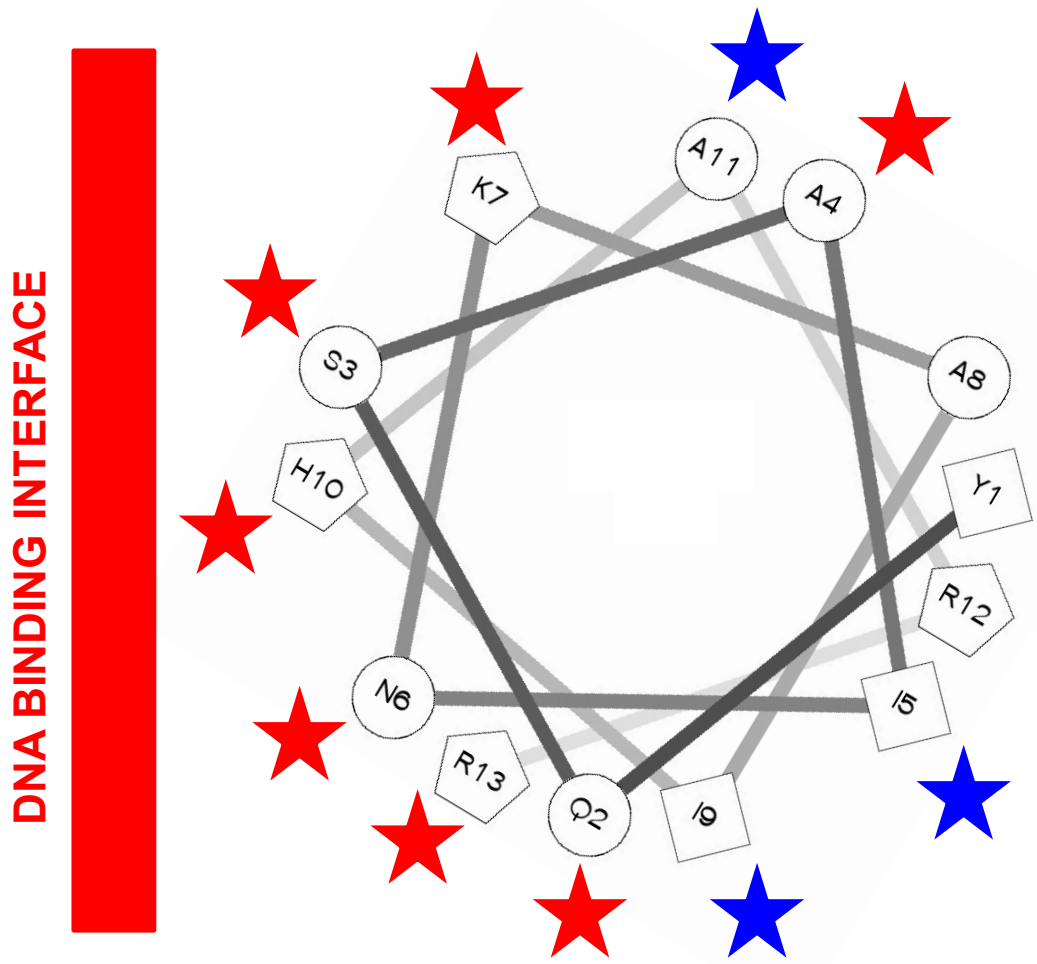
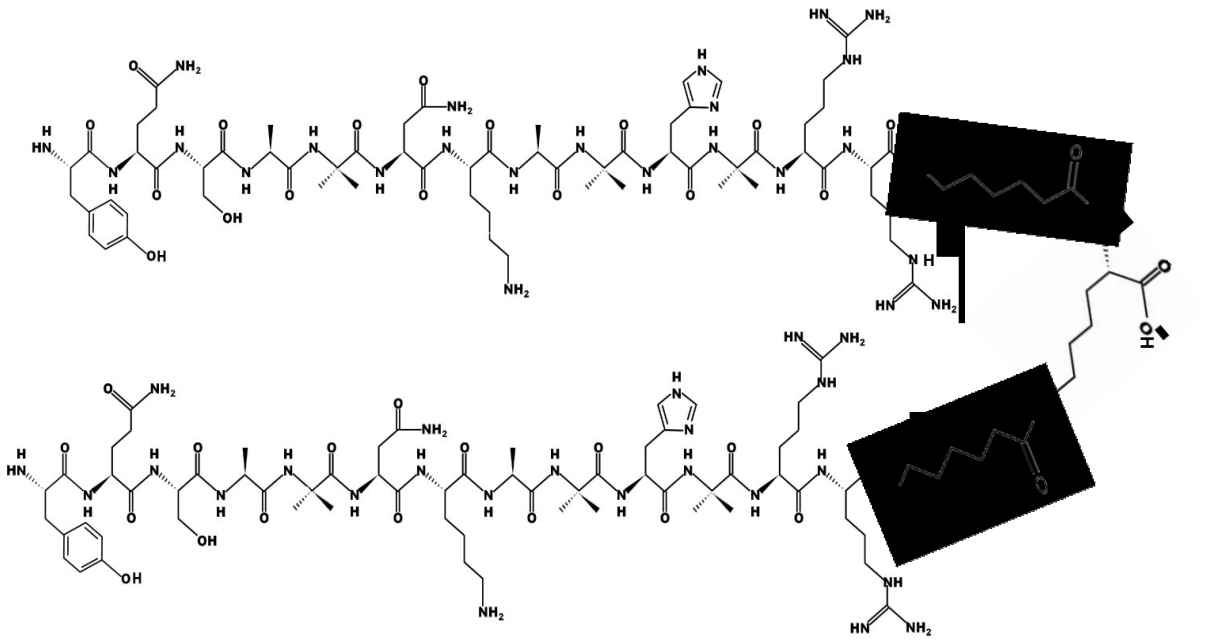


**Figure S3.** Design and sequence of the peptides used as DBD. The red color amino acids and terminal 'K' was used as linker. Using amino hexanoic acid (abbreviated as $A_h$ in the diagram) as linker the distance is ~24 Å and the peptide is designated as Br-Ahx-Cro. Using β-alanine (abbreviated as $A_b$ in the diagram) as linker the distance is ~19 Å and the peptide is designated as Br-$Ala_β$-Cro. Using glycine (G) as linker the distance is ~16 Å and the peptide is designated as Br-G-Cro. The scramble Br-$Ala_β$-Cro peptide contains scramble sequence of the Br-$Ala_β$-Cro and the peptide is designated as S-Br-$Ala_β$-Cro. α-amino isobutyric acid is represented as B.

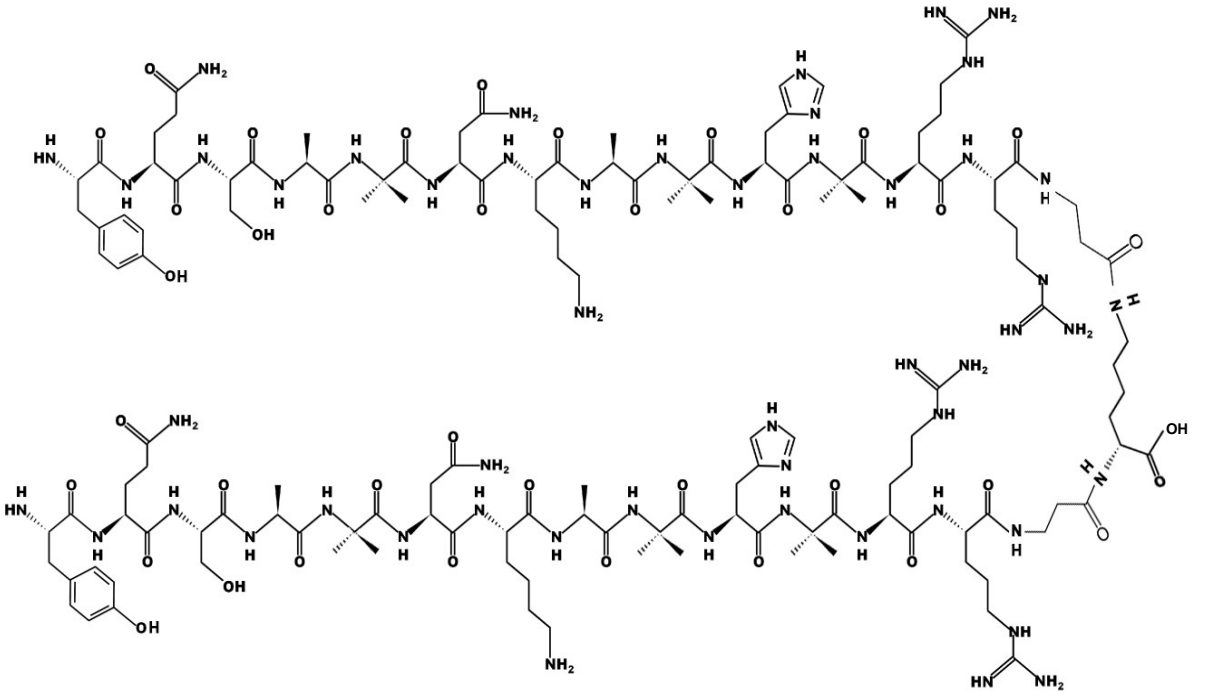
| Peptide | Sequence | Linker length |
|---|---|---|
| Br-Ahx-Cro | YQSABNKABHBRRA$_h$ > K<br>YQSABNKABHBRRA$_h$ | ~ 24 Å |
| Br-$Ala_β$-Cro | YQSABNKABHBRRA$_b$ > K<br>YQSABNKABHBRRA$_b$ | ~ 19 Å |
| Br-G-Cro | YQSABNKABHBRRG > K<br>YQSABNKABHBRRG | ~ 16 Å |
| S- Br-$Ala_β$-Cro | RBSRNKHQAABBYA$_b$ > K<br>RBSRNKHQAABBYA$_b$ | ~ 19 Å |

**Figure S4.** Chemical structures of (A) Br-Ahx-Cro (B) Br-Ala$_\beta$-Cro (C) Br-G-Cro (D) S-Br-Ala$_\beta$-Cro (E) Kix-Br-Ala$_\beta$-Cro (F) S-Kix-Br-Ala$_\beta$-Cro (G) No-NLS-Kix-Br-Ala$_\beta$-Cro (H) VP1-Br-Ala$_\beta$-Cro and (I) VP2-Br-Ala$_\beta$-Cro.

**A**

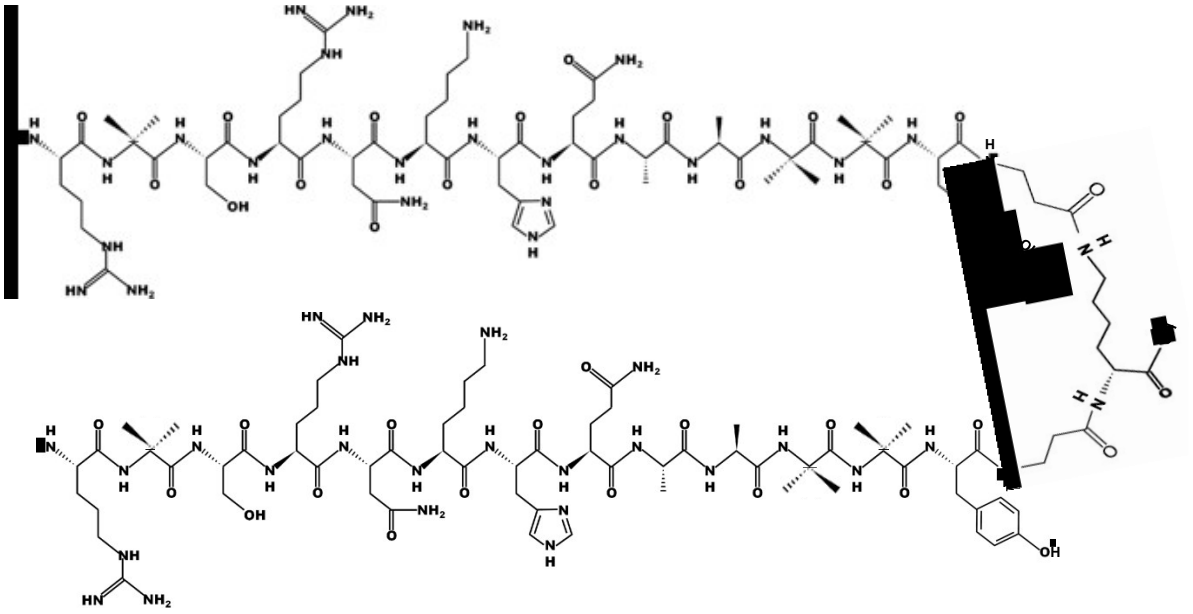


B

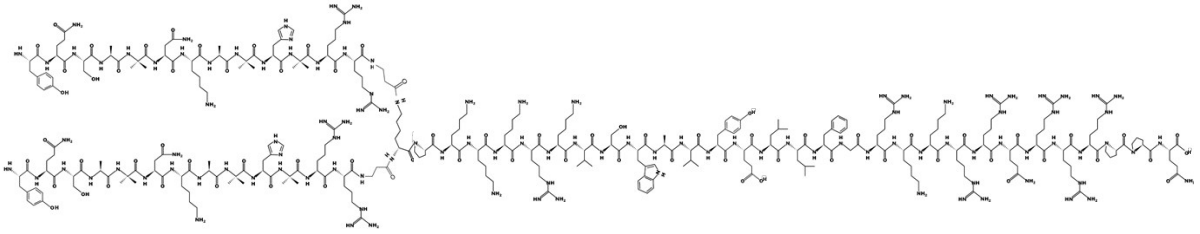


C

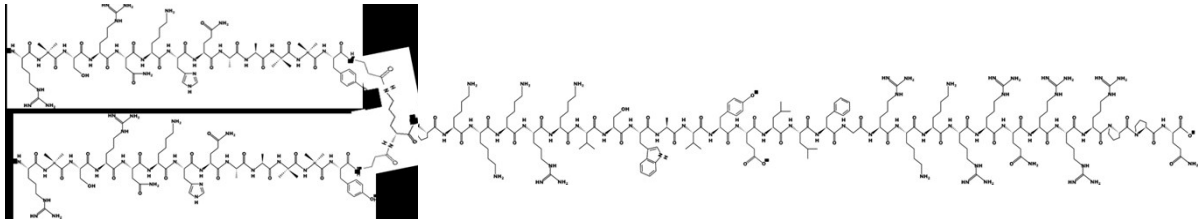
D



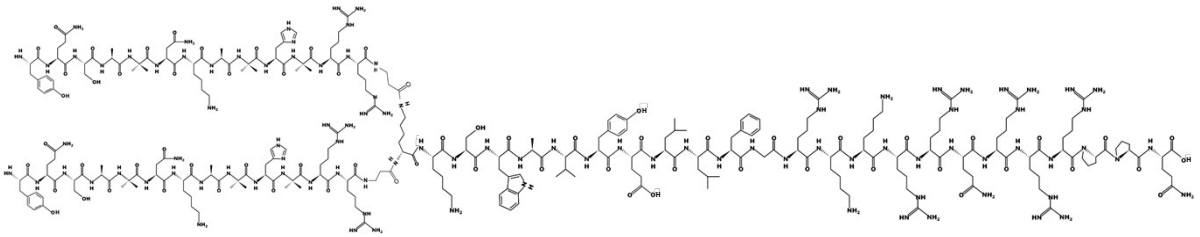
E



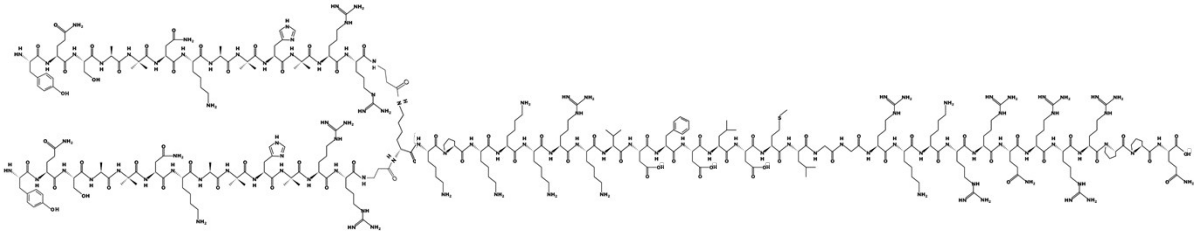
F



G



H



I

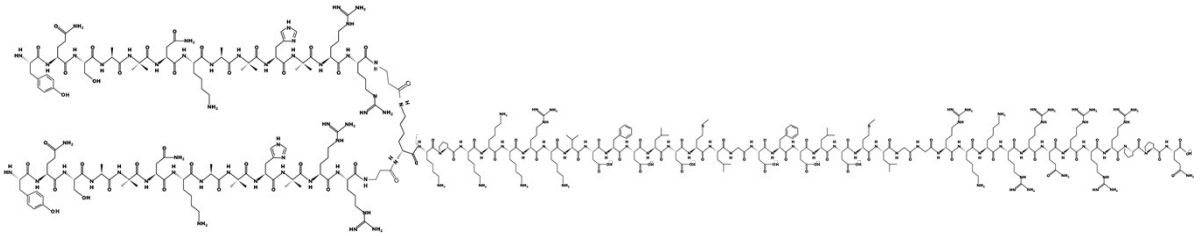
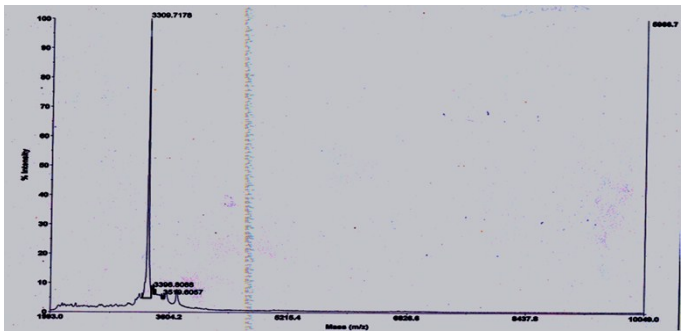
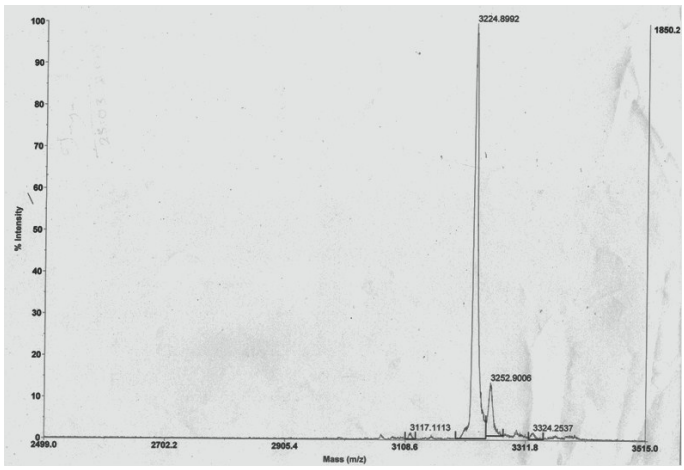


**Figure S5.** Mass spectrra of the A) Br-Ahx-Cro. (B) Br-Ala$_\beta$-Cro. (C) Br-G-Cro (D) S-Br-Ala$_\beta$-Cro (E) No-NLS-Kix-Br-Ala$_\beta$-Cro (F) VP1-Br-Ala$_\beta$-Cro (G) VP2-Br-Ala$_\beta$-Cro (H) Kix-Br-Ala$_\beta$-Cro.

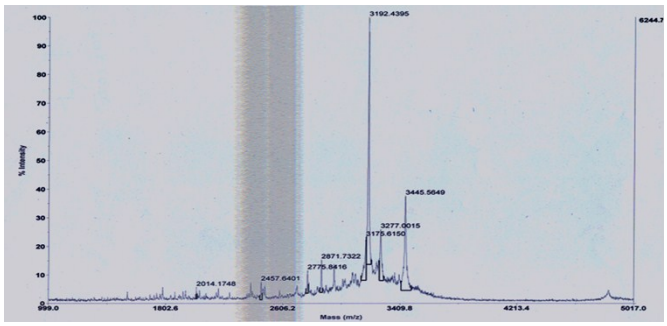
**A**



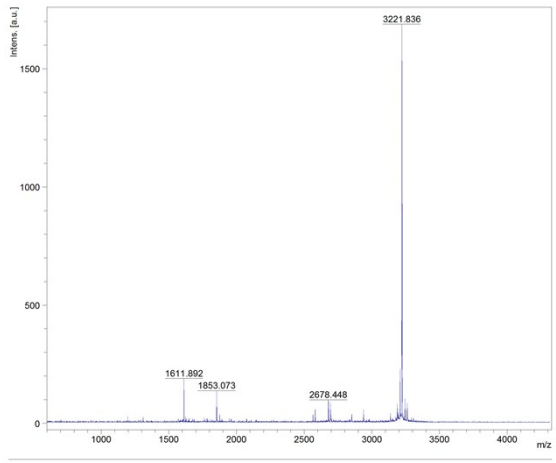

**B**



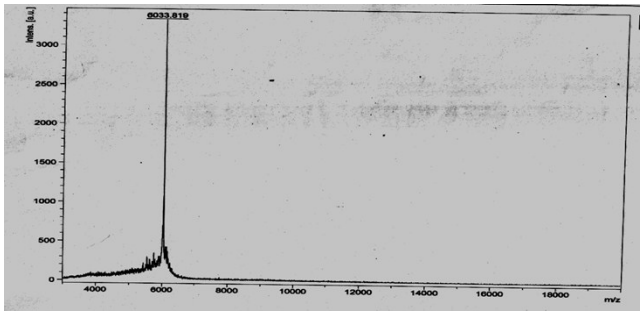

**C**



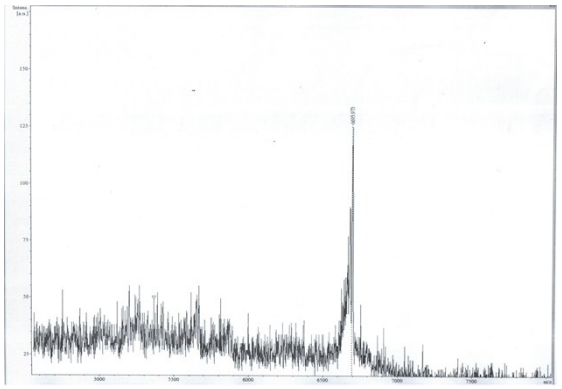
D


Intens. [a.u.]
1500
1000
500
0
3221.836
1611.892
1853.073
2678.448
1000
1500
2000
2500
3000
3500
4000
m/z

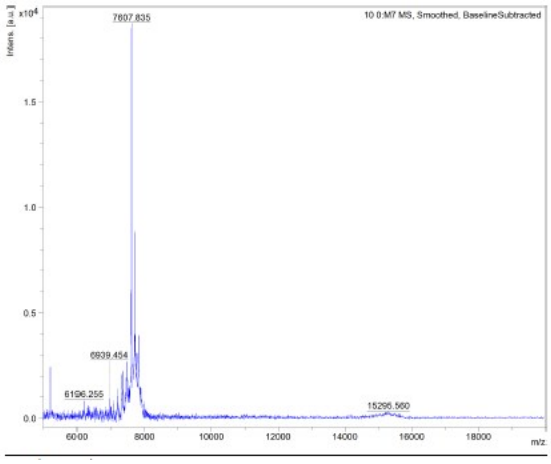
E


6033.819
Intens. [a.u.]
3000
2500
2000
1500
1000
500
0
4000
6000
8000
10000
12000
14000
16000
18000
m/z

F



G


Intens. [a.u.]
x10^4
10 0.M7 MS, Smoothed, BaselineSubtracted
7607.835
6939.454
6196.255
15296.560
1.5
1.0
0.5
0.0
6000
8000
10000
12000
14000
16000
18000
m/z

H

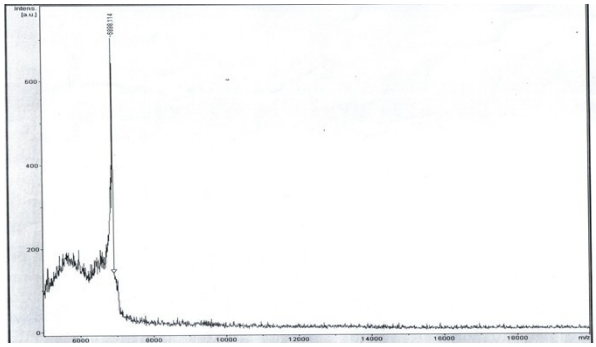
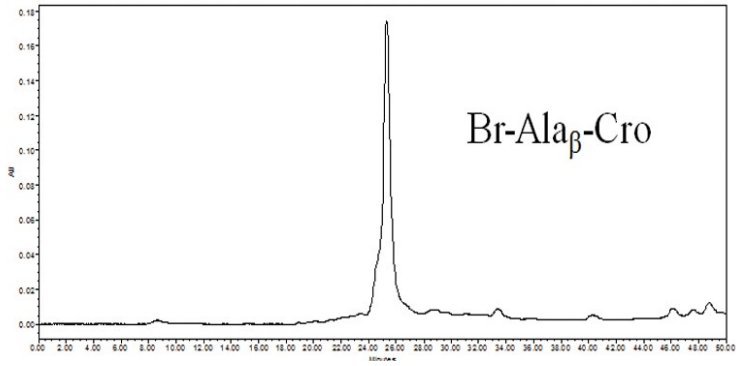
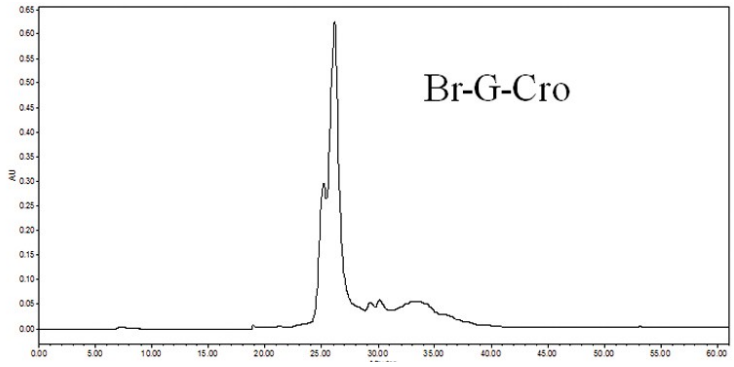

intens.
[a.u.]
8886.114
600
400
200
0
6000
8000
10000
12000
14000
16000
18000
m/z

**Figure S6.** HPLC profiles of some representative peptides to show quality of peptides used. (A) Crude Br-Ala$_\beta$-Cro (B) Crude Br-G-Cro (C) Crude Br-Ahx-Cro (D) Crude S-Br-Ala$_\beta$-Cro (E) Purified S-Br-Ala$_\beta$-Cro (F) Purified VP1-Br-Alaβ-Cro (G) Purified Kix- Br-Ala$_\beta$-Cro. (H) Purified S-Kix-Br-Ala$_\beta$-Cro.

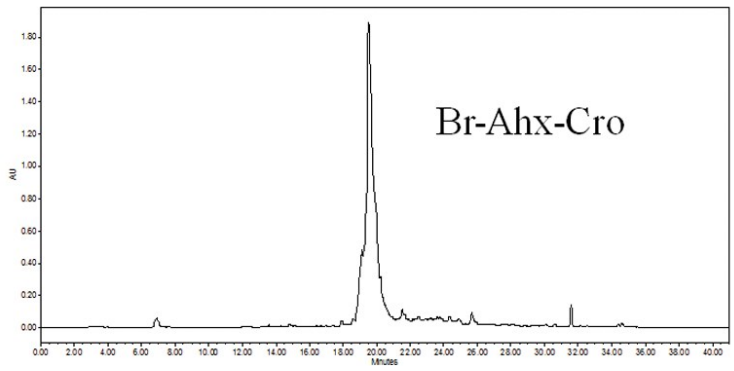
**A**



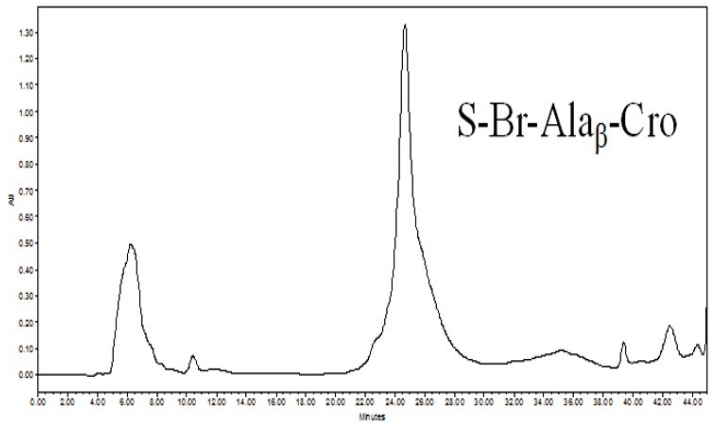

**B**



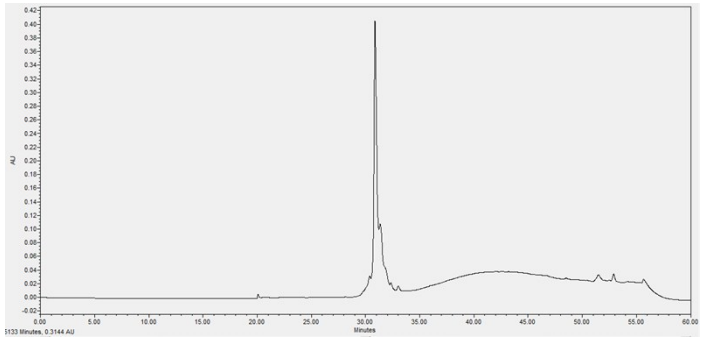

**C**



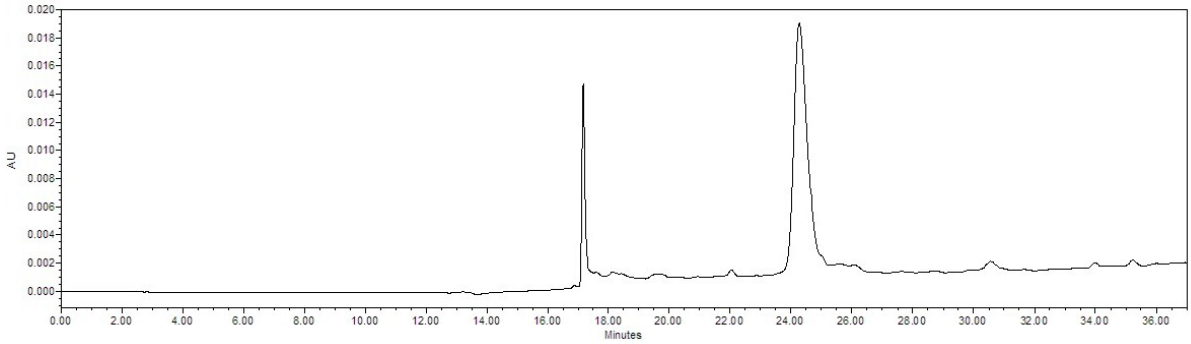
D



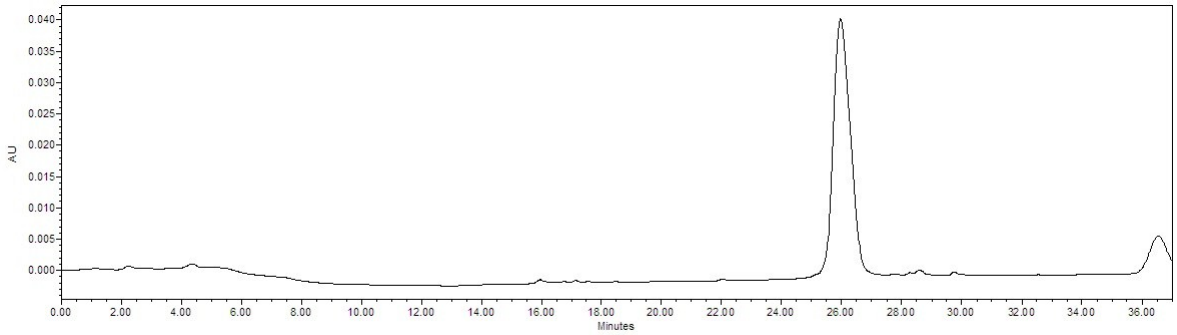

(E)




F



G


AU
0.040
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.000
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00 26.00 28.00 30.00 32.00 34.00 36.00
Minutes

H

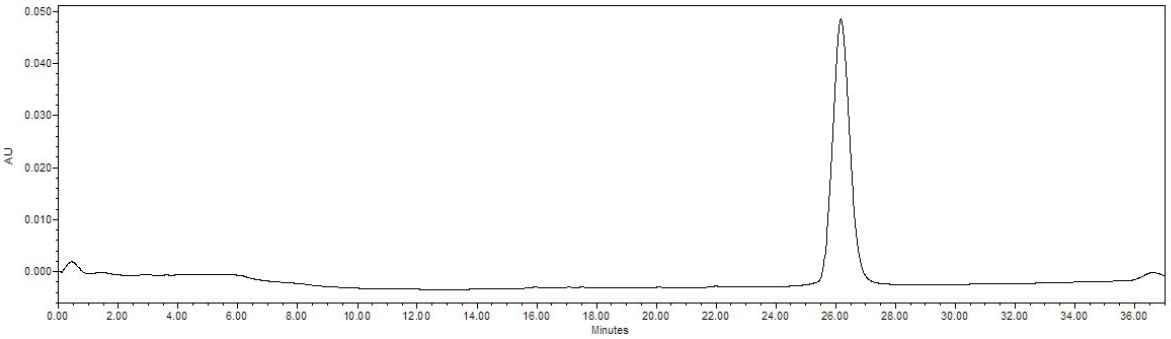
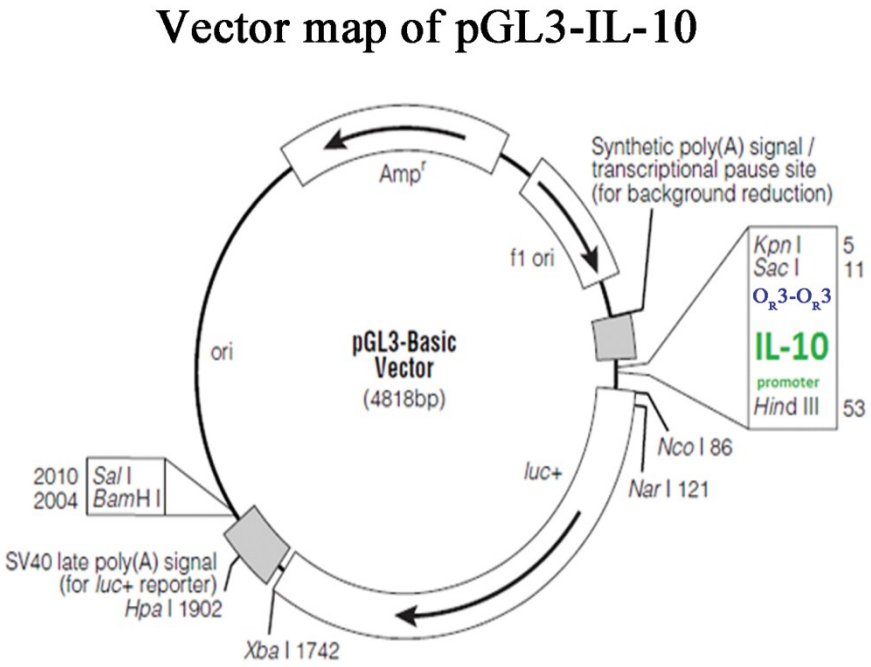

AU
0.050
0.040
0.030
0.020
0.010
0.000
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00 26.00 28.00 30.00 32.00 34.00 36.00
Minutes

**Figure S7.** (A) Vector map of the pGL3 containing $(O_R3\text{-}O_R3)_n$ and mouse IL-10 promoter. (B) Schematic representation of the reporter plasmid. n represents number of $O_R3$ in the upstream of IL-10 promoter, n=1 represent 2X $O_R3$; n=2 represent 4X $O_R3$ and n=3 represent 6X $O_R3$.

**A**




**B**

**Figure S8.** The cell viability was by measured by MTT assay. The cells were treated with increasing concentrations of (5, 10, 20, 30, 40 and 50 μM) Kix-Br-Ala$_\beta$-Cro. The cell viability was measured at 36 h. The data are plotted as % viable cell against concentration of the peptide. Untreated cells were considered as 100% viable.

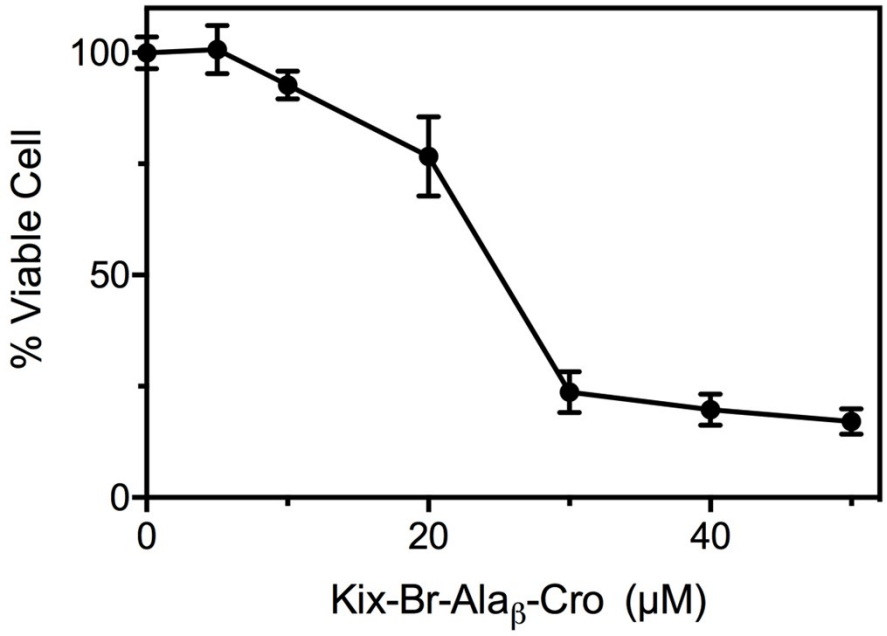


**Table S1:** Name and sequence of the oligonucleotide used in the study

| Oligonucleotide | sequence |
|---|---|
| **$O_R3$** | 5'-TATCACCGCAAGGGATA-3'<br>3'-ATAGTGGCGTTCCCT AT-5' |
| **C12-$O_R3$** | 5'-TATCACCGCAAGGGGTA-3'<br>3'-ATAGTGGCGTTCCC CAT-5' |

**Table S2:** Sequences of activation domain and artificial transcription factor (ATF) used in this study

| | |
|---|---|
| KBP | SWAVYELLF |
| VP1 | DFDLDMLG |
| VP2 | DFDLDMLGDFDLDMLG |
| Kix-Br-Ala$_\beta$-Cro | (YQSABNKABHBRRA$_b$)$_2$KPKKKRKVSWAVYELLFGRKKRRQRRRPPQ |
| S-Kix-Br-Ala$_\beta$-Cro | (RBSRNKHQAABBYA$_b$)$_2$KKKKRKVSWAVYELLFGRKKRRQRRRPPQ |
| VP1-Br-Ala$_\beta$-Cro | (YQSABNKABHBRRA$_b$)$_2$KPKKKRKVDFDLDMLGGRKKRRQRRRPPQ |
| VP2-Br-Ala$_\beta$-Cro | (YQSABNKABHBRRA$_b$)$_2$KPKKKRKVDFDLDMLGDFDLDMLGGRKKRRQRRRPPQ |

NLS and CPP sequence are marked in red.

**Table S3: Gene upregulated more than 3 fold upon treatment with Kix-Br-Ala$_\beta$-Cro. Gene name with fold upregulation. The location of the promoter in the genome with availability of $O_R3$ like sequence within the promoter region.**

| Gene Name | Fold Change | Location of promoter in the genome | Or3 site in promoter |
|---|---|---|---|
| Endou | 16 | Chromosome 15: 97,717,800-97,718,201 | absent |
| Golph3l | 15 | Chromosome 3:95587401-95588001 | absent |
| Nlrp4a | 12 | Promoter not characterized yet | - |
| Pi4k2b | 11 | Chromosome 5:52741200-52743600 | absent |
| Cwc22 | 10 | Chromosome 2:77945200-77947000 | absent |
| Dym | 6 | Chromosome 18:75018201-75020001 | absent |
| Pdpk1 | 7 | Chromosome 17:24140200-24142600 | absent |
| LOC100040563 | 7 | Chromosome 4:102986001-102986401 | absent |
| Slc5a12 | 6 | Chromosome 2:110597200-110603400 | absent |
| LOC665635 | 5 | Chromosome 2:30677000-30677801 | absent |
| Tmx3 | 4 | Chromosome 18:90508400-90511601 | absent |
| Man2a2 | 4 | Chromosome 7:80369000-80369601 | absent |
| 1700013B16Rik | 4 | Promoter not characterized yet | - |
| Col9a2 | 4 | Promoter not characterized yet | - |
| LOC100047593 | 4 | Promoter not characterized yet | - |
| Rps18 | 4 | Chromosome 17:33954400-33957801 | absent |
| LOC668861 | 4 | Promoter not characterized yet | - |
| XM_001480358 | 4 | Promoter not characterized yet | - |
| Zfp397 | 4 | Chromosome 18:23953801-23956201 | absent |
| ENSMUST00000067307 | 4 | Promoter not characterized yet | - |
| Ints9 | 4 | Chromosome 14:64948001-64951201 | absent |
| 0610009D07Rik | 3 | Chromosome 12:4816600-4818800 | absent |

## Table S4: The location of the promoter in the mouse genome with chromosome number and position in the genome. Promoter sequences

**Endou promoter:**
Location: chromosome 15:97717800-97718201
Sequence:
TAGCCCAGTGGTTCTCAACCTTACCTCTTTGGCAAACATCATCTCTAAAAACATGTTTAC
ATTATGATTCATAACAGTAGTAAAATTACTGTTATGAAGTAGCAAGGAAAATAATTTTAT
GGTTGGGGGTTGACACAACATAAGGAACGTATTTAAGGGCAGCAGCGTTCCGAAGGTTGA
GAAACACCGGCATGCCTGTCTCTTGTGTGTTTACTGCAAAGCAGGCACCAGGGGGCATCC
CAGGATTGCGTGTATCTTAAAGAATCCCTCCCGCGGTTCACCTCTAACAACAGTTAGTAA
CTGTTGTTGTCTAAGGTTGCAGACAGGTTTGCCTGCATGCAGTCCCCAGGGCCAGCCACG
GAGACCCATACTGATCCGCGTCTGAGACAGAATCCAGCTTGC

**Golphl3 promoter**

Location: chromosome 3:95587401-95588001

Sequence:
TAGTCAAGAGTCAGTGCTGAAGAAAAGGAGAGTTAGTGAACCAAAGGAACATACCTAAAA
ATTATTTTTCTTATGCAAATTTGCAGTTCTCAACATTTAAACTAATATGCTGTATTATTT
GAAGGTGTGCCTTCTGTACCTCTGAAAATTATTTTGTAGTTCATAGTTATTAAACTAATA
AAAGTTTTAGCTATTATATTGTGTTTACCTGAGATAAATCCAAGGGAACTTTTGACTAAA
CAAAAATACAAATGTCATTGATTCTGCTCATATCCACTGAATACAACCTGCCTTCTATTT
TAAAGAAAGCTTTTTTTTTAAATGTGTTAAGTTTTAAGGCCCTCTAAAACCATTTTTGAA
GTGTTTATTCTGACTTTAAAATTCTGTAAGGATACACTTTTTACTTATGTTAAAACTGTA
CCATTTACTTGGTGCTATTAAAACAATGTTTGCTATTTTGATATAGTGTGTTATTAACAA
CTTTTTTTTTTTAATATTTATTTATTTATTATATGTAAGTACACTGTAGCTGTCTTCAGA
CACACCAGAAGAGGGCATCAGATTTCATTACAGATGGTTATGAGCCACCATGTGGTTGCT
G

**Pi4k2b promoter**

Location: chromosome 5:52741200-52743600

Sequence:

TTAGTAGCAAATAAATGCATATAATGCACAGATTATGATTACAAAACTATATATTTAACA
TATGGACTATAGTAATAGAAGTAGCTCCCAGCCTTCATGAACTTTTAGAGAAAAGTTAGT
GAAATAAATGTATATTATGTATATAGTAATAAGAGAGGACTATTTAGGTTGGAAGGGATG
CACCTCCCGCGGAAATGACGCTGGTCCTGGCTTTCTGCTGGATCCACTGCAATGGCTGGC
AACGCTCTGGGAGTTGAGACACAGGTGGCCATCAATGCCCTACCCCGCCCCGCCCCGCCC
CACCCTCGCTTGCCCCACCCCCTGTCCTTCTGCGCAGGCGTAAGGAAGGCGTCCCGGGGC
TCGAGAGGCGGGGCGTGGCCGGCGGCTGGGCGTGGCCGGCGCGGGCTGGTGGACTGGTGA
GGTGGCGTCGGTGGGAGCAGTCCCCCGCGGTTTGGCGTGATGCTGGGGTCCCAGTCCGGA
GCCCGGCGGCTCTGAGCGCCGGAGCCGCCCCGGGACATCATGGCGGAGGCCTGCGAGCCC
ACCCGCCCCTCGGAGGACGAGGACGAGGAGCGGGAGCCGCTGCTGCCTCGCGTTGCCTGG
GCCCAGCCGCGGAGGGTCGCGCCCGGGAGCGCCGTGAGAATGCAGGCGGACGAGGGCGCG
GATGTCCTCCGCGAGCCCGCTACCGACGAGCCGCCGGCGGTGTCCGGGGAAGGGTCGATC
TCCGCGAGCTTGTCCACCGAGCTGGATCGGACCCGCACTACGAGCTCAGGTGCGACCCAT
GCGCGCGCGCAACCCGCCCTTCCCTGCCTCAGTTTCCTCTCACCTTTCCCTCTGTGGCAG
GAGAACGGGGACCACTAAAATCTCACTCTGTAGGCTTCCCCAAAGAGTCAGTCCAAGTGG
GAACCCAACACTGAGGGAAACTTCATCCAGGACTGTGACGTAAGTTGTTTGTTCAAGTTT
GGACTTAAGTTCTGAATTAAGAACGCACAACGATCCAGAAAGTTCTCGTCGGAGTCATGA
GTTTACAGGGCGGATGATAAGCCGAGAGATAATAGGAAGGGATGACCATTTGCCGCTGGG
ACAAGTTTCAAAACTTTCTATTAAATGTTGCTATTGTGACGTTAGCCAGAGGGGGCGGAG
GCGGGCGAACTCCTGCCGCCAGCTCTCTGCAAAGTATATTTGCGCGCTTAAAAAGAAAAA
CACCTTTTCATTATCTCCTGGGCTGATTTTAAGTGGACCTTTGAAAGAGAGGCTAACTAA
CAATGTTTAAACTAAAAATTAAGGCATGCTATTGCAGGCCTGGGTGTCTGTGTGGTGAGT
TTGGACCGGTGGAGTGGAGTGTGGCTGAAATGGAGTAGATTTGTATGAGATCCATCTGGA
GTCTAGACTGGCTTTTTTTTTTTTTTTTTTTTCCCTTTTTCTCCTACTTGATGATGATA
GGATGGATCCAAAGAAGTATCATATGCACTATTCCACGAGCATCATTCCATATGACTCCT
TGTGGAAAAGCCACAGGGAAGGTAGTTGTTATTAATACTGCATTTCTTAGAAGTTAAGTT

TATTTGACCTACCAGTGTGCTAATGTGACTTGAAATGATTTGAACAGGGCAAGTCAAGAG
TTCTAACAACTTTTATTCAAACACAAATCCTCTAAATAAGAGACCTAAAGGAAGTCGTCT
CTGAAACACATTAACTACATAGGGGTCTGCCCTTACCACGACGTTAGCTGATGGATGAGT
TTGCCATGCAACACTGAATAGAGCAAAGGAATGTGAAGTTTTCCACTTTCACAAATCTCT
GTTCAGCCAGCTGCAGTGGACACTTTGGCCCTATCTGAGGTAGAATCTAGTCTGTGAGGT
AGCTAAGATCCATATGGCTGCTACTTGCCCCAGAGTTACCTCGTGTGTTAGATTGAAATA
GAGCCTTGAAATAGAGTCTTGAAAAACTTTTACTGTTTGAGTGGGTTGGATTGGTTTGTG
TGTTTTCTTTGTTGTCAACACTTATATTAAGGCTAAAGTCAAAAGGACGAGAATGTGGTA
GCTAAAATGCAAATGGATTATAAGAACTCTTAGCTCCTCGGTGGTGATAGCTAGGTCAGA
GTTCACAGTTTAAGATGAGTTTTGACATTGGTTCTCACAGCAGCACAATGGGTGAGTTGT
CGCTGACCCTCTTTGTGAAAACTGAAACTTAGTAAAACGGTAATAATGGTGGCTCCTATT
CCAAGATGTTTGCACAGTCACAGATCCAGGAAGTGGTCAAGTTATGACTGAGCCACTTCT
TTTCTCTTCCAGGTCCTGAACTTTGTCCGTTTTACTCTAGGAGTCTGTGTTACAAAACAA
CTACTGCTTTGGGGTATTATGTAGGTTCTTTTTAAAAGGATTTTTTTTTCTGACTCTTTA
T

**Cwc22 promoter**

Location: chromosome 2:77945200-77947000
Sequence:

AAATAAACAAAAGGCAGTTTAAATACAGAGCACAATTAACAGAAAAGGTAGCTCATTTTC
ACATGTATGGATCTTTTTTGCCTTCATATACATATAGGGACAGTTCAAGTTGCTGGTGCA
AGTGGCTATCAGCAAGCAGGTATCCAGTCCCCTGAATCTGATATTATAATTTGTTGTGAT
CCATCTTGTGGGTGCTTGGATCGAACCTAGATCCCTCTGAAAGAGTGCTTTTAACCACTG
GGCAAACTCACCAACGGCTTGCCGGGGCTCAGTTACTGTTTCTATTAATACCTTTCCCCA
AATTATTCTTCCGGGGGGCGGGGGGGCGAGTTAAGTGGCTACAATGTATCAAAAAAGTCT
TGCATTTAGAGCAGGTCAAACAGCTGGCACAGTGCTGTTTGTACTGTAGGGCCTGCAGCT
CGCTAGGGGCAGGAACTGCAAACCCATGTTTGGAAGTTGCAAATAGTAGAGGCTTTTAGA
GTTAAGGTTAACTACGTATGCGACCTAAACACAGAAGATGCGAAACATGGATCACTAGGG
CTAAACATGACAGGGGTTCAGTCAAGTCAATGAAAACGATCCATAAATGAAAATGTAATA
GAATTGGCTACTGGGATAAGCTGGAATGAGGCAAACGCGTTACCAGAATTTATTAATGCC
AGCCCCGGGACGCCGACTGGGGACGCGACCCAGGGTGCTCCCAGGAGAGGTGTTCACCGA
CAACCCAGACCAAGGCCCAGACACGCACGCCCCGCCGCCGGCGCTAGGCCCAGCCGGCTC
ACACGGCACCTCGCGGGCCGCCGGTGGAGACGGCGCAGCCACATGAGGAAAGCAAGGCCC
ACCGGGCAAGCGCCCACTCGCAAGTCCCTGCCCTCTCCACGCGACCTAGCAACCCGCCGC
CCCCACTCTGGCTCCTAGTCTATCCCGAGTCTTCAGAGTTCCGTAATTACCTAAAGCTCC
ACCGCGCCGCCATCATACCGGAAGAGCGTCAGTTCCCTTTCCCGGGTGAAAGTAGCGCGA
GAAATTAGTGAAAAACTGGTAGTGAATTGGCAGGTATGTGATACGATTAATGATAATGCG
GTATAGTAACACCTTGAGTCAAAGCAATCCAGCTAGTAACACTGTCATTCCTTACAAAGA
ACACGAGAAAAGAACATGTGAGTCCGAGTGAACTACAAAGTCCAGCATGCTCTGGCAGCG
CGGACGGAAATGAGTTGTAGTTGGACCGCAGTGCATGACGGGAATTGTAGTTCGTTCGGA
TAGGAGGTCACACTAGAGACACTTTTGAGAGCTAGGTTTTCGAGGCATGCTTCTTATTCT
AACACTACTCATCTTCTAAAGCCATAAAAAGTAGCTCTTAGTAGACACCGTAGCTTTTCT
TTCTTTTTTTTTTTTTTCAGTGTAAAAGCTTAACACTAAGATACATACTATGAAACTACC
GCAGAGGCTTACTATATAATACCTTGCCTGAAAATAATAAAAAAAAGTAGCGATCCCTTTC
AGATTTTTGACAGAAGCTTTACCAAAAGTCTGTTTGGGAGAAGAGGGGAAAGTGACAAAA
TGTTTAGTATCTATTAAGTTCCAAGATTTTAAATTCAAGTGAATGGAAGACTATCTGATT
ATAAATGATATGTTGTAAGGACAGTGAGAAGTAAATTGGAGGCCCAACATGGTGGTGTCT
GCCTTTACTCCAAATATTTTGGAGGCTGAGACTGGCACTGTGTGTTCCAAGACATCCAGA
GCTACATAATGAGCCCCTGTCTTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
A

**Dym promoter**

Location: chromosome 18:75018201-75020001
Sequence:
TACGAAGCAGAATGGTCTTTGCCATTTGAGTACTTAGATTATGGGCTCAAATTCCATCTT
TTTGTTGTTGTTGTTGTTTTGAGACAGGGCTTTGTTATGCGACTGTGGCGGGCTTCT
GATTCAAATATGCAGCCCATACTGATCTTGAACTCTGTACAATCCTCCTTTAGCTTCTCA
GGTAAATACCATACTTTGACGCTACACAGTATATGGCTTTCATTATGGTGCAACTAATCC
TGATTATATTGCTGAGATCTTTCTGTTAAAAATATTTTATTTTTTAATTCGACAAGGCCC
ACAAAAGACTTGTATGACATTGGTGTGATTGTCCAGCAGGAATTAACAGGCACCTTAGTC
TCAATTAAGTTACCCATGCGCAAGAGCTCTCAGTCCCAGGAGCCTAGAGGTTTTGGTTTC

TACACTTTGCCACGTGGCTATGTGGTTGCCCGGGAGACGCTAGGAGGATGGAACGCCGGA
AGGACTACATTTCCCAGAAGCTCGCGTGCTCGGCGTGCCACGGGACCGCCCGCTGCTGGC
GCGCGCTTAGCCGTTCGGTCCCCAGACGCGTCCTCCGTCTCTTCCCCCTTCTCACCCCAA
CTCCTCCTGCGTCTCCGCACGGACAGCCGCTGGGGCTGGAACGCGGTGGGAACCGGAGCC
CCTGACCTACGGAACCATGGAGGCCTAAGCCGCCGTGCACACCCGGGCCCCTGTGCGCCC
GCGCTCCCGTTCCCTTGCTGCCTGCCGGCCGCGGCTGGAAGCGGAGTCCGCGCCGCTGAC
CCAGGTGGGCTTCCCCGGCGCGGTACCGCCCCGCGGGGTCTGAGCCGGAGGTGACCGGCC
GCGCGCCGAAGGGCTGCGGGCCCGGAAGGTTGGGCCGTCCGGAGCGGTGCAGAGCTCGGC
GGGCCTCGTCCGCCTGCTCCCTCGGCGCCGGGGTGTGCAGAACCAAGGGAACCGAACCCC
TGGCCGGCCCGCCGGGACCCAAGGGTTGCCGGAACGGTTGGAGCTTCCGGACACTCCGGA
CAGCCGAGGTTGGCTGGACAAATGGGCGTGTGATGGGCAGGGCGAGCCTTAGTGTGTTGC
GTGTCTGACACACACGTTTTTATTGCGATATCATCCTATACTGTAAAAACGACAAGTACC
CCTGAAGTACAGCTCACTAGCTTTTAGTCTCGTTCACAGCTCGTACCTCCTTACGGCAAT
CCCTCAACATTTTTATCTCCTCGAAAGAGGTCTGTACCCATTTGCAACCGCTCCTACTCC
CTCCTTCACGCCCTGCAACCATTGGTCTATTTTTTGCCTCGGTGGATTTGCTTGAGTCAT
TGGGTGTTAGCTGAAGGATGGAAAGGGCCACATTGGGTGTGAGGGAAGTAGAGCGATAAA
GCAGGTGGAAGAAAGTCCTGGACAACAGCGAAGGGATGTAACGGGGCTAGCAAAGCCCTA
CCCTGAGACTGGGGATAGGAGGAAGTGCAATCCTTGAAGGTCTGGAGAGGAGGTTTTGGG
CAGGAAGATGAGGAAATGGGAGTGGAGTTGGTGAGGAGCAGTGTAGCCTTCCCAGAACCA
GGTGCTCCTGCCTGAGTTGAGGGCTCGGTCCTGGAAGGTAGTTTGTAAGTGAATGGTTGC
TTAATAGGATGGACTTCATGGGTTTCTAGAACGACTATCCTGCTCTGGAGTTTTTAGATA
ATCTTTTTGGCTCAGAGAAACTGGACAAGGGTAACTATCCCCATCTAGAAAACGAAGAAC
CGGCCTTTTAGAGGTTGGGGTACTTGGCTAAGCTCCCAAAGTTAGTAAGTAGCTGTGAGT
T

**Pdpk1 promoter**

Location: chromosome 17:24140200-24142600

Sequence:
GGTAAAGACACAAGTGTTGCAACACCCTTTTTTGCCTTTCACCATTTTTCTCAGCTGGAT
AGGTCCTCCTCCTTTTAGAATTTGGATTAATTACCATCCATCTCACAGTAGAGGGTTATA
ACTTCAAGTCCATGGATGAGATGCAGAAGGCTCTTGAGCCTCTGGTAGCAACAAAATTCC
GTGCAACTGTGTATGCTTCTCAAAAGGAGCCCATACTTTGATTCTTAAAGGGGCTCATGA
CCCTCCAAAAACCACTACCACAGACACATCCTAGACCATACTGAACCAAGTGAGTTTCTT
CAGTCCCCTCGAATAAAGAGACCAACAACCCTCTTCCCATCCCGCCCCACTCCCTGAAAG
CTGGAGTGAGGATCGAACCCAAGAGCCCTGAACATGTGAAGCAAGCAGATTACCACTGAA
CCACAGCCCTAGCCCAAGGAGGTTTGCCTTTTGAACCAAATCTTGCTCGGCGTCTGACAT
CTTCTCATCAGCCACTTCATATCTCATACAAGGCAGACACGCTGATCTAAGAACTCTAAA
ACTTACCAGAAAATACAACTATGGATACTTATTTAACAACACAGAGAAAACAAACATTTC
ACCCAACAAGTGGTTGGTTTTGTCGTCCTCGTCCTTCCGCCCGCCCCCCCCAACCCCTCC
TCCCTGTCAAACTTAGGCTGCCCTCGCTGGGGCCTCTAAGCCCGGCCTCCCAGGGCTTCC
CACCTCACTGAGCGTAGAACTGCAGGCAAGAGGAGGACCGAAATGCCCCGCTGACCAAGA
ATCTGGAAAGGAGCCTCCAGCATCAGGTGCTACAAGCAGGAACTCACATCTCCTGCTTCC
CAGGAACCTCCTCCAGGTAAAGGATGGAGAATGCTCCTCCTCGGGCAGGACCTCAAGGTA
AGCCCAGAGGCTCGGCTCCAGGCTGCACAGAGCCGAAAACAAACTTTCCAGGCACTTCCA
CCCCGCCCGACCGGGTGACACCTCCCGGCCTCCCCCGCCCTCCGTGCCCTGTCAGCGCCC
CCGGCTGCCGCCGACAGCCCGAGCCGGGGCCCGCATTCCCTGCCGCAGCCGAGAGGAACC
CTCGGCCGGTGCCAACAGTCTTGCCCGCCACCCAGCCGCAGGTGAGGCGCCGCCCGGCCC
GCGACACCGGGCCGCCCCGCCTCGCCCAGCAGAACCCAGGCGACCGCCACGCTCCCGGGT
AACGGACCGGTGCGCCCTACAGCAGCCGCGCGCTCACCAGCTGGCTGGTGGTCCTGGCCA
TGAGCCCCAAGTCGGCGCGGGCGCCTCCCCCGCAGCGTTGCGCTGCTCGGCGCCTCCTCA
GCGTCCTCCTCCCCGCGGCGGAGCCCCAGCAATGGCCGCCCTCATCCTGCGCCCGCCCCG
CCGCCTGCCGCCCCCAGCGCCACCCCGCCCCCGCGCGACGCTAAGCAGGACTTCCGCGAC
GTAGCCGCCGCTCCTCTCGCCCATTGGTCAGTATAGCTGTCATTGGCGGCGGGCCGTGAG
GGCCCCGCCCCCGACCAAGAGCCGTATTTCCGGGTTGGGGCCCGGCGCGCTTGTCGAGGG
AGGCTGAGCGAGCCAATCAGAAGTGACGGGGAGGGGCCGGGGCCGGAGCGCCGAGCTGTC
ATCCGCGTTCCCTGTCAGGTAAACTGAGGCTCCGTAGCCAGCGGGCCAAGGTCACGCCTG
GGGCAGCAGCAAGGCTGAGCTCGAACCTACTGCCAGGACAACTTGAATCTGATTCACAGG
TCACGAGTAATCTGGGAGAGAAATTGGGCTATGATTTGAAGCAGAAAACAAGCCGTCCTG
TCACTCCTGATGAACCTGTTTTTGCCAATATTTTGATGGACGTCTTTGTGAGCTTGCTTT
TTTTTTTTTTTTCCTTTCCCAGCAGGATCTCACTATCGAGAACAGACTGGCCTCCTACTT
GCGGCAATCCTGTCTCAGCCTCCCAGCTGCTGAAATGACAGGTATGCCACATCCTCCTTT

AGTTGTCATTATGACAACCGCTGTAGATGTTTTTTTAAATTATGGTGAAACGCATACAAT
TTTTGTAAACTTTTTGCGCTGTGTGTACAGTGTATGTGTGTGTCCACTTTGTGGGTTTGT
CAGATGTCTTCCTATACCAGTCGCTACCTGGAACTTGCTGATCCCTCTTGGCTGGCTGGC
TGGCCAGTGAGATGCAGTGGTCCCCCTGTCTCTGTTTTCCCAGGGGTGAAGGTACAGACA
CACCCCCACCATCTCCAGCTTTCTTTTACAGGAGTGCTGAGGACTTGCCTGCATAGCAGG
CATCATACTGAGCCATCTCTACAGACCTTCCCCATTTTTTTTTCAGTGTTGGGAATTAAC
TCAAGACCTCACACATGCTAGATCATCTCAAAATTAAATGTTTCTAAATTACACCCATAG
C

**LOC100040563 (predicted Tctex1d1) promoter**

Location: chromosome 4:102986001-102986401

Sequence:
GAGTGAGCAGTTATTATTGCATTATGTGAATTCTTAAGTAATTAAGTCTTTTTTGCTCTA
AAATAAATTCTGCTTACAAGTCCATTATATAGTAAGTCCGGAGTGACTTGTAAGATGAAT
TTACAACCTAGGAAAAGTGTTTATTCAAGTGTTTGGTACAAAGCAAGCTTTATTTTAGCT
TTCTTATTGAAAAAGTCCATTTGCATCTATTCCAAAAACGAAAGGAAGGGAGAAGAAGGA
AGGAAGAACAAGGAGGAGAGTAAGGGTTTCCTCTGCACTGGAAATATGGCAGGGCTGGCC
ACACCCCTCGTTAATGACGTCACAGTCATGTGACATAGCCAGGCCTCAAACACATCCAGC
ATCACCTCTCCTGATTGCAGTGTGGCTTTGAAACCATGCCT

**Slc5a12 promoter**

Location: chromosome 2:110597200-110603400
Sequence:
CTGAAGTAGTCAAATTTTATCACCTCCATGGTTCCTAATAATGTTTAACTTTCATTTTAA
AAAAATCTAGGAGATAATAAAAGCATGCCATCGGGCCAGCCTACTCCCTCTACCCCCCAG
AAACTTCAGGAGTCTTCTGTCAGATGTTGCAGAAGAAACGCAGACAGCTAGTGGCCAGAA
TCTGATCCCAATTTGTCATTTTTTCCCCCTTGGTGTCAGTTTCACAGCCAGCTTTAGAGA
ACTTTGAGGACTGTTGAGTATCTGAAAACTTAGACATGGATGGATGATCTTCAACTCTTC
AGAGAACCAAATTCTTCCTTCCATCTGAAAAAAGTTTTGGCTGTTGGAAATATCTCAGTG
AGCCCTGCTTGAAAGAGACTCTCTTGTGTTCCACTTCCAAGCATGAGGGTGAAGAACTTT
GAAGCTTGGGATTATGTTGTGTTTGCAGGTCTCTTCGTCATTTCCTCTGGAATTGGCGTG
TTCTTTGCCATAAAGGAGAGAAAAAAGACAACATCTCGGGAATTCTTAGTAGGAGGAAGG
CAAATGAGCTTTGGCCCTGTAGCTTTGTCTCTGACAGCCAGCTTCATGTCAGCTGTCACT
GTCCTGGGGACTCCTGCTGAAGTCTACCGCTTTGGGGCATCCTTCTTCCTTTTCCTCATT
TCATATGTATTTGTGGTCTTTTTCACATCTGAGCTCTTTCTTCCTGTGTTCTACAGGTCT
GGCATCACAAGCACCTATGAGGTAAGAGACCACCTCCCTGCCCCCATCACAGCAAGGAGA
TGGCACTTGCTTTCCACACAACGAGTCTTCCTTTTTCTTGCTTTGTTGTTTTTCTATGAC
ATCTTGACATTAGTAGTAGATGACTATTTTCAAGGCAATTCCCCACTGTTTAAAAAAAGA
CAATATGTAGTATATGGGGTGGGAGGTATAGTGCTAGGTTTTGGCTGTAGACACATCTGT
TTGTCACTCTGTAGAAGGTAACTAGTAATGTCACTCCCGCTGGCTAGTTCCATGATTTAA
TAAGGATCCATGATAAAAAAAAATGGATGCACTATTTAATTGATGAAAGATAAGTAAGAT
TAAAAGAGATAATGCTATTTTTATTGTCATTAGAGTGACCACTCTGAAAAAGTAAACGGA
ATACAGGGTACTAGTATTATAATTTTCAGCAGAGAAAAGTAAGATCTAGATAGGTATGTT
AGTCCTGTTTTAAAACTAGATAAATATCTACTGAGAAGCCAATATTTATATATGATCATT
TAAAGGCCTGTATATCAAACAGTAGCTCCCTAATGCTAATTATTGCTTCATTCCAGTTTT
AAGTAGCAGTTGGGAATATTTTGGAAAAGTGAGAAATATTTTAGTAGAAAGAAATTCCCA
AATAATAAAAATAGGAAAATTATTGTAATGAAGGGGAAATAAGACTATGAAATTTTGGCC
TTTAAACAATGCTAAATAAAACCAACATGTTCAAAACTGAATTTTCATGCCCAGTATTTA
AGTGTGCTTATAAATCACTTTTTTGTGAAGAAAAAAATATTGAACTTCATTGCCAAAAAC
CCTGAAAGTCATAGCAGAAGACTCATAAAAAAAACTTACTTTAAAAAAATCTGAGATAAA
TGATCACAAAATGTTTCTCTAGTCTTCCAAAGCTGTTCATGCAGATACATATTCTCCAGT
AGGTCTTAAGACATTCAGTTGCATGGACTATCTTATCCCAAATGTTCAACATTAATTATC
TTGTGCTGCTAAGAAATTAATCACAGTTAGGCAATTCCACAGCCATTGCTTGAGACACAA
CCCCAAGAGAAAAACAAGTAAATGTAGGCTTAGCTCTACAAACTGTGCACAGATGCTATT
CAATTGTAATGTACTTCTGAAAACCAAAATGTGGCCGGGATTTTCTCATCCATGCGTGAA
GATTTGACAGTAGAAGACATTGATTGCTGAGTTTTCTAGTTTTGTTCTGGAAATATTTTT
AACAATTTCCACTCTTATGCCTATGTTTCCTTTTGGAAAACATGTTGTGGTCAGCTTTTA
ATTTCATAAATACCCATTCGGCACTTACTTTGTCCTATGCACACCTTCAGGTCTTGAGGA

CATAAGTAAAGTGCTGTACTGCCCTCATGGGGGCAGCATTCATACAAAAGGAAACAAACA
ATAAATGCATAATACATATAGCACCAGGCTATATGTATTTAGCTGTGTGTTAAGTGCTGT
TCTAAGTATTATCTGTATTAACTCGTTTGGTGTTCGTCAGCACAAAAATCGAAGCTCTTG
CCTCATGGAATAAGAGATAGCTTATTCTGTAGCCAAATTTAAGTGAATATAGTCTTGGAA
ATACATAATGAGGTTAACAAATGTTCTATGTTTACTTAAAATTAATCGAGTAATATATTC
ACACGAAGAAATCTGCCTTCTCTAAACCTTGTAACTTGTATCTAACTTGAGTAAAGTGAA
TGAGAATTGCAAAAGCAATTCAGCAAATTCCAAATACATAATTTGCAGGAAACCAAGGAT
TCAAACCTGAAAAACCTCTGACACATGGCTCTCAGTGCCGCTAGGCCAGGTCAACTCTGA
ATAGATTCAAACTCTTTTCAAATTTACATGTTAGATTTAGTATACAAACCTATTTTAAAA
GTACTATTGCTTACCCGAAGAATTTTTTTAAATCCTTCTTAGGATATAAGGATTGTGTAG
AGTGCTTGTAAGGTTATTTTGAGAAAGGCTTCTTTTTTATAAGAATTGACAGTGTGGCCT
AGACATGGAGAGCAGGAACTCTAAGTGTGAGCCTCCTTGGCTTTGAATGATGGTTCTGCC
ACATGCTAGCTGCACAGCAAGAAAAAGTTCTCTATATCTGGTTTTCACAGTCCAAATATT
GCTCATAATACTTCTGTCTCGAAGGATTTCTAGAGTGGATGAGATAGTATATGTAGTTAT
TTTTATAGACCTGGGCTAGAATAGTATGTCGGCATTTCTATTATTTTCTCTCTCGTTTTT
TGATTTTTATTCCCCACCCTGCCAAAATTCCCTATCCCATCCTCCCTCCATCTGTTTCTA
TGAGATTGTTTCTCTACCCATCCACCCACTCCTGCTTCCCCACACTGACATTCCCCTACA
CTTTGGCATTGAGCCTTCACAGGACCAAGTGCCTCTCCTCCCGTTGATGCCCGACAAGGC
CATCCTCTGCTACATACGCAGCTGGAGCCATGGGTCCCTCCATGGGTACACTTTGGTTGG
TGGTTTAGTTCCTGGGAGCTTTGGGGGAGAGCGGTGTCTGGTAGGTTGATATTGTTGTTC
TTCCTACGGGGTTGCAAACATTGTTTCTTCAAATGATGTGGCATTTGAGTTTTTTCTTTA
TGCTCCATTATAGAGTAAAATATATGACTTTTGGGATCTTTTCAGGTGTGAAAGAATGAA
ATAAGTGTACTGTATTTTCCATTCTTTGAACATCAACCAGTTTTGTCATTTGCTTTGATA
AATAAAAAATTTCAGACAATTAGTGAGTTTTGACTCTAGACCTCTTGAAATTTATAGCTT
CTGTGTTTGTCCCTTGGAATATTGCCATTATATTGCTATGAAAGCAAGCTGGCACAACTG
ATCTAGCTCTGGAACAAGAATAGACACATACCACTACCACCACCACCATCATCACCACCA
CCACTACCACCACCACCACCACTACCACCACCACCACCACTACCACCACCACCACCACCA
CCACCACCACCACCACCAGAAGGGGGGCAACCTGGCTGAATGCCCACTTCCACATATGTA
TGTAGGCTGAGCTAGTCCATTCCTCAAGCCATTGAGTGTACAGTTGACTAGAGCTGCACA
AGAGCTTGGAGCAAGAACAGAGGAGTAACCATTAACCTCAGTGAGATGAAGAATAGTCAT
TCTTGAGTTCTGATCCCATCACTTTTGGGATGAATATTCTAGTTGTTTTCAGTTGCAGTG
ATTTAAAAACTGCCTGAAGCAAAGTTAGGGAGGAAAGTGGTTGTTTCATCTTATAGATAC
CAATTCAATAGAGGAGGAAGCTCAAGCAGGAACTCAGCTTAGACAGGAATTGAAACAGAG
ATCATGAATACACACTAGTTACATGGCTTGCTCTGCCTGCTTTTTTTTTTTTTTTTTTTT
TTTTTTATTTAACCCAGGACCACCTTTCCAAGGGTGGCACTATCTCCAGTGGCCTAGTCC
ATCCTTTATCATTAATTCAGAAAATGTCCCACAGACTCAATGGAGCTATTCCTCAATTAA
AGTTCCTTTTTTGCTTGATTCTACTTTGTTTCAAGTTGACAGAAACTAATCAACACAGTG
GATAAATTACAGGAATAAGCAACTGTTTGGGTACTTGGATAGGCAGTGGGATTTCAAGAA
AAAACTTCATCTGTGCAACTTTATCACCAGCTCATCCCATGCCAAAGCATAAGAATTATT
TTCATCCATGTGTTCTAAGTTAGGTTTATCTCTCTCTGCCTTTGCCTCTTGAGAGCTGGG
ATTAAAGGGACACAAACCACATCCAGCATAGTTCAGTTTCAATAGCATACATTTTCATTA
TCTCTAATTTGATTCTATTTTCATTTCAGTTTAGTTTATATTCATTTTCTGTGCTAATAA
GCCCATCTTATTATTTACTCTTAAACTGTTTAAAATATGTTTCATATTCTATCATTATGT
CTAGCTTCTTTGACCTTCCTAGTTCTTTCAATCTCAAGTTTTCATAACCAAACCACTTAG
ATTATAATGTCTCACACAATGATCTCTGAAAATTTATGTATTACATTTTAGTAAGTTAGG
TTTAGTTCTTTGGAGAATTAGATTCCATGGCATGATTTTGATAATACTTTGGTAAGTAAT
ATGTCTAAATGTTATAGAATACTACTTCATCAGAATAAATTATGTTATAAAATATCTGTC
CTAAAATGAGTATTTAATATCTGATAAGATACAACTCAATATCATGAGGCTTTTAGTTTA
GTCTATGAAGGAATTAGCTTTTGATACTGTTTGGTTAAGGAGTGGCATATTTTATGTGGT
TTCTATTTTAGGGTTATAGTCTAGCTGAATGTGGAGAATAATATTAGGGTTTTAACCATG
GAAGCAATTGGACTAAGTAGGAGTCTGTTGCCAAAATAAAGAAGGAAAGTTAATTTTAAA
GCAGCATTAAAAGAAACAAACCAAGTGTTTCTGTTTGAAAGTGTGGTGTACCAGGCTCTT
TTCTTGAATTGGATGTGTTATGATGGCAGACAATGAGGCAAGATTAGCTTCAGATTATTA
CTATTTCAGTGTGAGTAACCATGGAAAGTAGATTTAATGTGTCTCGAGTTAGGAAATAAT
GTGGCTGGGTCAAGTCTAGTGTTCTAAGAAAAGAAACAATGAAATGATATTAAATCATTA
CCGTTGGGTCTTTTAAAAACAGACATGAAATAATGTGATTTTTCTGTTAGACTCGGGTTT
TGCTCTCTCTTGAGCTTAAAAGGTTATATATTTATCAAAGATAACCAGTAAGTTCAGGAA
AACACTTGGAAGGAGGAGCATATACAGTCTTTCATATGTAGACCTGTGTGATGTCATGCA
GGAGAAAATAGTCTTAGATCATTGAAAGAGAGAACATCCTCCAGGGCGATGGAGCCAAAA
GGAAAGGACTGGCAACTAAGAGTCCAGAGAAGGTCAAAGGGGGAAAACTTAGGCACTTCT
GTGGGAATGTGCCACTGCAGCAGTGATTTGGCCTAGGTGAATTCAGCTTGAACATTTGAT
CTTCAGAAATGAAAGGTTATAAGTTTCTTTTCTTTTAAGCCACCCATATTCGTACATACC
CAACAGGGACTGATGCCCTTAAAAGTATTTAGTATATCTGAGTAACAGGAGAAAGAATAT
GCATTGAAATATGGGTGGCAGCTGGATCAATCTCCCTGAGAGATATATAACTAGAAGTGT
TGGGAACAAATATGAAACTTTAAATTTGAAGAGAATCTCTCTCTCTCTCTCTCTCTCTCT
CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCAGTTTTTTAATTAGGTATTTTCCTCAT
TTACATTTCCAGTGCTATCCCAAAAGTCCCCCATACCCTCCCCCTCACTCCCCTACCCAT

CCACTCCCACTATTTGGCCCTGATGTTCCCCTGTACTGGGGCATATAAAGTTTGCCTGTC
CAATGGGCCTCTCTTTCCAGT

**LOC665635 (Gm14486) promoter**

Location: chromosome 2:30677000-30677801
Sequence:
CGCCCGTGTGAGTGAGTGAGTGCCCGCGGGCGCGCGTGCAGGCCGGGCAGGGATTTGAAG
CGGCGTGAGCGCCGCTGATAAACAGCCCGGAGCGGCGCGCTGGCGGCTGCTATCGGAGCT
GCTGGCAGGCGCGCAAGAAGCTGTCCCGGCCGCCAGCGCTATCAGCCCACGGAGGGAGGG
CGGCAGGAGGGCGGGAGGGAGGGAAAGTGTTTGCTTCAGCCCTTAATCACGGAATTTGAA
ATCCCAGCGAGTCGGTGCCATTTATTTTCAGAGCCTGTCAAGGCTGGTTTCGGCGTTTAG
AAGAAAGGAGGCTGTGCGCATTTACATCTCCGGCTCCCACCGCCCTCCTGGATCCGGTGG
GGAGAGGCCAGCGGTCGGCTCCTGGGACTTCCCCCTCCACGGTCAGGGCTCTGGGTGGCA
GATGGGCCTGCTGAGTGCAGGGTGCTGGCGGGGGCATCACATTTTATCCCATAAAGCAAC
CCTGATCTCCCCACCCAAAATTTGCCCCGGTTCCGGGGGACTCCGGAGACTTCTGCCAAC
CTCTTACCCCCCTCAAGCCTAAGGGGGGCTGAGGGCACCGCAGGACAAGCCCCTTCCTGT
GCAGATATGGACACTGAGCCCTGAGACTGAGTTCTGGAGATCTGTTAGGCCAAGGTCACA
CAATGGCAGACCGAGACCAGGAGGGAGGCTGGGGATGAGGGTACACAGGGAATGGAGAGG
GAGCTGTCAGGCTGAAAGCTCTCTTGCTCAGTGACTTTACCCACTGTCCCTTCTCTGACC
TTGATGTGGAAAGTTAGGGACC

**Tmx3 promoter**

Location: chromosome 18:90508400-90511601
Sequence: CCAGCTGGACGGAAATCTGTGCCAGGCAGTAAACAAGCCCATATTTGGTGGATGGCTCGC
CCTTAATCCCTGATTGGCTGAGCAAGCCTGCCTGGTGAGGTGATGTCACCTGTGGTGATT
GGTTGAGGCGTGGTTGTGGCTTGAGTGGATAAAAAAACCGCTTGTACCCCGAGGCGCGGG
GGATTTGAAGAGAGAAGCTGCCTGAGTAAACTGCTTTAAGAGGAACCGGTTGTTGCGTTT
TCCTTGCTGGTCGAGTTGGACGTGACAAAGGCTAATTCTAAAATCAATTTTATTCTTGAA
TGTGTTTTCCCTATTTTTAAGGTTTGAACTTTAAACTTTTACATACCTATTTTCAGAGTT
GTTCATCTTATTAACTGCTGCTGCATATATAGTTGTATTGGATAATTTTCTCATAACTGT
GATAAAATACTTGAAAGAAGCCAGTGAATTTTTTTTAAATAGTTTGAGAGGATATAGTAC
ATCATGACTGGGGAAGGTACTGAGGCAGAGGCAGGTATAGGGCAATTGGTGACACTGCAT
TGGCAGTGAGAGAACAGAGAGCCCAGGTGCAGTGAAATTTGGAAGTCTTCATCTGAAGAC
TGTTTCAGCCAGGTAGGCATCATGTTCAGGTTTTAGAACTTCCCTAAAGAGCACTACCAG
CTAAAGAATAGCTGCTCGAATACATGAACCTGTGAGGGACACTTCTCATGTATTGCCTTG
TCCCCTTTCTCCCTCCCTTCTCCTTCTACTTCTCTGCCTTTGCATGCTCTGTTTTCATGA
CATAAATTCTAAGACCTAATGTAATCCAAATCTCCTGGCACTTCATTATTATTATTTTTT
AAAAATCTCATCTAAGTCAATGACTGCCACAGGGACCTACTTATAAGAAGTGCACTATTT
TTATAAGTCACAATGTAAGAAAGGATGAAGGATATACAATACAAAACTTTTATACTAAGG
ACTTCAGTCCACAAATAATCAGGAAACAGTAAAAACATAACCAAGATGTTATAATTAACT
TCCCCTTTTTTAAGTTTCTGAAGAAGCAATCCTTTTTATTCTTTAACATTTTAAAATATG
TTTATGATGAGCTTGGATCACTAATATACACATATAGAGACACATTCGTTCCCCCTTATC
ATACTGAAACTATTCTTGAAGACCTGCGTATACTTTCGTGTGCGTGTGTGTGTGTGTGTG
TGTGATGCAGTTTCATTTTTGCTGCTTGCATTAGCATGGCAACCTATCACTTAAGCAAAT
GTCACATACACCACCAGAGTAAACTGTTAATTGTCCATACTTCTCTCATTCATCATAGTA
ATGGACCAATTTATATGATTCTAAACAGGTTTTTACACAGACGGTGTCCAGAGCTCAGCA
ATCCGGTCTTTTATTTTAAAGCACGATTTCCCATCTATTAGGGTACCATCCCGGTGTACA
AGTGCGTTCACTGGACTAAGTAGCAATCTTAAAACATGTGACAAAACTTAGGCTTGCAGT
GCATATCGTACCCTACCTCACCTCAAGCACTAAGCTCGCTTGCAAGCCATACGTGCATTC
CGGGAAGAGTCCTGCAGCACACTTGCGCCGTCACAGGAAGAACCCCGCACGGATTCCCAC
GCTCTTCCGAGCTCTTTCTCAGCTTTAATTACAAAAAAAAAAAAAAAAAAAAAAAAACC
CTCTTCGAGAACTCAAATGCGCGTGCGCACTGCTTGGGCGCATGCGTCACTTCTTGCGCA
TGCTCACAGGCTACCTTCGCTGGGCTACCGGGATCCTGAAGGAGGCCGGCCTACCCCGCC
CCGCCCCGCCCCGTTCCTGGTGGGCGCTGCGTCCCAGGTCTTCCTCACTTTCCTTTCGGG
GGAGGCGGGAGCCTGCATTATTTGTCTCTGCTCGCGCTCATCGCGGGATTCACACCGCGC
TCCATGGCGAATGCAGTGGGGCGGAGGAGCTGGGCAGCCCTCCGGCTCTGCGCTGCGGGT
GAGGAGCACTCGGGGTGGTGGGGAACGGGTGGGAAGTTGGTCCGGAATGAAGCCGGGCCG
TGTGGGGCCGCGCCGCGGGGTACTTAGGATTTCTACTCTGCCTCCTTGGTGGGTCAGGCT
TCCTCCGCGGGACTCTGCGTGATCTACCGCCCAATGCGGAGAGGGCACTGCTATCTTCTC
TGCTATCAGGAAGGTTCCTTGACTCCACTGATGGAAGGACAAGGATTTTGAGCCCCAATT

TGGCATAGGTTCTGGGTCTTGTTACAGTGGAAGCACTTTTCTTGATTAACCTAAGAGAGG
GATGATGGCCAAGGGCAGATACTGGTAAGATATCCAAGACTGGTTTGAAGAGACCAAGCA
TCTTTAAAGCTCTCAGCCATTTACTTTCAGCAGGGAGTATTTGGGTTTTCTGTGGTTCAC
CCCTGACTCAGCCACCAATTTAGAACCCTCTGTGAACCAGGCACTGCCCTGGAATAAGTG
TTAGAACTGGAGTGCAGTATTTTCTTTTTGTACCTATTTTGGGGAATTACATTTGGACTT
TGTATCGAGGGGATAAATGCCACTAGAGAAAAAGGAAGTTATTTTTATCTCTAGGCCAGA
AAAATGGTGAAAAGGAAAAATTAAAACAAGTGGAAATATGTAGAATTATGTAGAAATATG
TCTTGTGTGTATTCGCCTACACTGCGCTTGCCTTACGAAACATGATTGCATGAAGAGTAA
ATTTTTACTGAGTGGATGAACTACTTTTGCTTTTGTGTCACTTACAAACCTTATAATGGA
TCATTGGTGAATTTTTTGGTAGATACTTTAAAAAGTTCAAACATTGTATAGGCAATTTAG
ACTTAAGGTGAAAATCAAAGTAATAAATCATTTATTTTCTACATTTTTTTCCTTTCCAGT
TATTTTACTTGATCTGGCTGTTTGTAAAGGATTTGTAGAAGATCTAAATGAATCGTGAGT
ATACATATATGTGTAAAATACTAGTAAATTCTATTTAACTTTGTCTAATTCAGATGCCAT
CAAAATGTCCACACATGAATCTAATCTTTTAATAATAGGTGCAGTGGCAACTGATATTAA
CATTGCCAATTTTGTTTCTGGAGTTGTTTAACATGCCTCCATTTACATGCAAAATGGAGG
AGTGCCCCAGAGAGTTTCCCTAATAGTTCTAAAATGAAAAACCAGCAAATGTTTCCATAT
TGTTTCTTTCTTTTGGTTTAAT

**Man2a2 promoter**

Location: chromosome 7:80369000-80369601
Sequence:
AAAAAAAAACAAGAACAGAAAGGCATCATCACTCAGACCACGGAAGGCAAACATTTCCTC
TCACATGCCCGTGCTCAGCGACCTACATGTTCCTCCTGTCTCAAGCACGGTGCTGAGGAG
GACTGCGTTCAGCTTAATGGAAAAGCAATACGGTGGTCAATCAGCAGTGTCTGCCAGAAA
CTTAAGTCAGGCTATTATTACACAAGGCAAATTTCTTGTTCCCCTCCTCCCCCTTCTCCT
CCTCCTCTTCCTCCTTCTCCTCCTCCTCCTCTCCCTCCTCCTTCAGCATTGCTGCTTTTC
ACTTCCTGTTACAGCCAGGAGCATAGCTCCGTGGTAGAGCATGTGTTTAGCATGCGCCAG
TCCTTGGGCATCCATCTCCACAAGAAAGACAAATCTGCCTAGGAATGTCCCTGGGCTACA
GCTACATAACTCAGTTGTAGATCGTTAGTGTTATATGCACGCATGCCCATGAGGAGGCTC
TAGGTCCAATCTCTAAGACAATGGGGGTGGGGGGGTGCTCTTGTCCTGTGCTGACAAGAC
TCTCCTGAGCCCAGGTGACTTGCATTTGAAGAAGAGAACAAGGTGTGACTCACCCTGGGG
AA

**1700013B16Rik (Spata31d1a) promoter**

Sequence: no sequence found

**Col9a2 promoter**

Sequence: no sequence found

**LOC100047593 promoter**

Sequence: no sequence found

**Rps18 promoter**

Location: chromosome 17:33954400-33957801
Sequence:
CCAGGCACCTCCTGCTGTGCTATCATTGTTTGCCAAAGTGCAGAATTAAAAAAAGTCAAC
ATGGCAAGGACTTAAAAGTGAGACTTCACAAGGTTAAAAAAATTTACTCCAACTTTTATG
TCCACCTGACATTACCTCGGCTCCCCAGGCCTCCTTCAAACCTCTCAGCAGCCACCCAGT
TCCATCTATACCCTAGCATCCCACAAGGTCCCACCACTGACAGGTCTCATGGCTGAACAC

ACCTGACCAGCAACTCCTCAAATCACCAACTTCAAGTCCTGTCCCTGTCTACAAGACTTT
TAGCTCTGCAGAGAAACATCCCAGCTTCCTCCTACGACCTCCCTTTTTGCCATAATGGGG
AAAACTGCCAGCTTCTGTGTCACCTTCCACTGTGGCATCTGCCAATAGCTTCTGCAGACC
ACCCCCAACCATCACCGCTAAGCCTGGACACCTCTTAGGATGTTAGTTTACAGGAATTCT
ACTCAAGGTTTGCTGACAGCAGAGACTGCTCTCGGAAGCTGAAGATCTCGGGAAACCTCT
GCAAGGTGGGACATGACTACACTCACACCCCACTACCCCTGGAGCAATGTTTTGGTTTAA
CCCTGGGTCATCATCAGGCTTTTATTGAAACGGGTTGGGTGAGTTTAGCTGCTTCTCTTT
TGCCCCTAACACTTTTACCTTAATGGCAGTGATGGCGAAGGCTATTTTCCGCCGCCCATC
GATGTTGGTGTTGAGTACTCGCAAAATGTGCTGGAACTTCTCAGGGATCACTAGAGACTG
AAAGAACACGGGCAGTTATCAAGACAGGTTACCCCCAAACCGGGAAGGTGGCCGAGGCAG
GGTTGGATTGTCTGCTAAGTTCTTACACAACTCCTCAGTCACCGGGGCTCACCCCTAACG
CGCTCAGGCGCGGGAAATCGCAACCTCTGACAAGAGGGTGGACGTCGGCCGCGGCCCCCA
ATTCCCTGGACAGGGGAAGCGGGAGGAGCCCACACACAGCCAAGCTCCTGGCAGCTTTGC
TGTGTAAGAGGAGCGAGCATTCCTCGGGGCCCCGATCCCCGGGATTTCACGTCTTAGGTC
GTCCCCTCTCGGGTCACACTCCATCGCTCCACAGAGGGCTGTGGGCCCGCAGCATGGGGC
ACTCTCTGAGTCCTGACGGAAATCTCTCCCTGCCCCGGCTCCGATGTCGCAGGATCCCGG
CGCGGATTCCGCTCTTACCATGGCGGCGGCACAGACGGCGGCGTGTAGGCCTCCTGTGGA
AGAGAGAGCGGAAGTGACGTATAAGTCTTATATAGATTCCTGGCGGAAGAGGCGGGGCTA
TCCCCGGAGGTCCACTTCCGGAAATTGCTTCTCTCCCCGTCTCCCCCTAAGGGAGGAGTT
GTATTGAGACAGTTTTGGAAGGGACCGGAGCTATAACTTCCGGTTCGGCCTTCTCTGTCG
TCGTACCGTTCTTGTGTGGGCGCGGAGCGAAGGTTCCGGAACAGAGTTGTGTTCGAGACG
TTCACCTTCCTGTTTGCACTTGACCGAAATGATCAAAGCCTAACCGGAGAGGCTGGAGCC
GATCGCCTTAGCTCAAATATCTGGGACCCTTTTTTGCTCCATAGAAGACATTCCAGGACT
TGTAGCTTTGCTTCGGGAAACGGACTTTCTGAGACTGACAGCTGCCGGAAGTCGGAAGTG
AAGCTGCGGGGGAATGGCAGCCGCAGCGACCATGGCGGCCGCTGCCCGGGAGCTGGTGTT
GCGGGCTGGGGCCTCAGATGTAGAGGAGGAGGAGGGCCCGCTGGTGAGAGCACGGGGCTG
TAGGTCCTCCGGAGTATCGTCAGGTTGTCCCTCTCCTCAGGCGTAGGCTCCATTTAGACA
AGAGGTTCGGAGCCGTCTTGCCACGGGTGATGGTGACCTAGCCTTCGTAGGGACCAAAGT
GGCTGAAACCCAAGCCCCAGGGACAGTGTCAGCTTCCTGTTCACGATTTTTAGTAGAAAA
AAATTTTAGTGTTTAGCGTAAAAGGTTAAGAACGTGGCCTCTGGAACCAATTTGCCCGGG
GTCCCACCTCCAGCCCCACCGCTTGAAAGCTATGCTGTGTTTTCGGTGCCAAAACTTCCC
ACACCTTTACAGTGCAAACGCTGTTAGGACATACCTTATAGGTGGTCGTTAGGATTAGCC
GAATGAGCATGTTAAATGGTGTACCAGACACAGGGGTAAACACTATGTAAGCTGGCGATT
TTAGCCTCTGGAATTCCCTCAGGGATCACCTGGAATAGCTGGAAACTGAGGTCCTAAATT
CAAAGGCAAGTTTGACTAGTCCTAGTTTTTTCTCATTTGAACAACTCCAGCAACTCCATG
AATTGCCTGATATTCTCAGTAGCCCAAGATTGTAGATTTGAACGCATCAGAATCTTTGAC
CCACTTACTCCGGCACCCCCCACCCCCCAAAGCACCCCCCACCTCACTTGGAGGATTAAG
GATACCCGATTATGTTTTGTTTCCAACATCTCTTTCTGTATCCCACATTCATTAACCTGT
CCCTCTTCAGGGGGGGGGTTCTGGACTCCAAGAACCCCTACAACTTGGAGAGTTGGACAT
CACCTCTGATGAATTCATCTTGGATGAAGTGGATGGTAAGTGCTGACATGGGGTGGGAGC
AAAGTAGACCCCTTTATTCTTTAGAAATATATTACATGACCCCAGTAGATTACCCTCTAT
GGTTTCTTTCTTTTGTATGTGTGTGTTTTTATTTTTAATTTATTTTTATTTTTTTGGT
TTTTTTTCGAGACAGGGTTTCTCTGTAGCTTTGGCTGTCCTGGAACTCACTCTGTAGACC
AGGCTGGCCTCGGACTCAGAAATCCGCCTGCCTCTGCCTCCCAAGTGCTGGGATTAAAGG
CGTGCACCACCACCCCCCGGCTGGCTTTTTTCTTAATGTACAGATAAAGTTTAATCCATG
AGTTGAGAAGGCTTAACTTCTAATTTAGGAATAGTAAGAAATGAACAGTACGAGTCGAAT
AGTCACAGCAATACACTATAATAAAATCGCAGCATCACTTTTTCCCCTTTGGGACTATTG
TTGAGTAAGAATACAGCCTGGTTGAACATACATACGCAGACAGCTAAGACTGAGTGACTG
GGAGGCAGGCACAGCACACACCATAGATTAACGGAGCAAAGATGTTTCAGATCCAAGCCC
AGGCGAGAGGTTCTCTAGCTTCTCAGAATGCCACACATTTTAGATCTTACAAGCCACACT
TTGGAAATCCCCTTTTAATATTTTCCATCTGTGGTCGAGTAGACCATGAGGAGCAGAAAC
AGCAGGAAGTCCCACCGGGCAAAGCAGCTGTGCCTATAACTGGAGAGGGTGGGGGCTGGG
AGGGAAGGGGAGATTGGTGTGGTGTGCGTGTGCCCAGTGTGG

**LOC668861 promoter**

Sequence: no sequence found

**XM_001480358 promoter**

Sequence: no sequence found

**Zfp397 promoter**

Location: chromosome 18:23953801-23956201
Sequence:
TTAGCTAACAAAAGAGAATTTAGTTGTTTCTATGTGATCGTCTTATTACCCTTTTACTAA
CTTTTTGAAATAATTACAACTAAATGTAAAACTATTAATAAGAAATAACTTGGTGTCCTT
AACAAATAAAATCGCTCACAAAGTAAAGTACCTTAGCTTTTTAAAGGGGAAAAGCTGGGG
CTGGAGAAAGGGCTCACCAATTAACATTTGCTGTTCTTCCAGAGAAACACTCACATATAT
CAGTAACTACAGTTCCAAGGGACCGATCTCCATAGGCACTGCACACTCACGCCATATAAA
CGAAGGCGCAGTTACACATAAACAAAAATAATAAAAGTAAATATTTTAAAATGGAGACAA
CAGGGCCTAAAAAGATGGTTCAGAGCATTTGATGCTTTTGTAGAGGACCAAGGTTCGGTT
CTCTCAGTACGTATAAGGTGGTTCACAACCATTGTCACTGGATTCAACGCCCTCTTCTAG
CCTCTACCAACAAAGAGTGGTGTGGCAAAGTATTTATACACCGAAAGAGTCAATATTTTT
AAAGGCGTGTCTGGAATCCGATCTGAGAGCAGCTTGGCTTTGTGCTAGGGCTGCAGAGGA
CGCTGCGGCTGCCGCCCAGAGCGTGAAGAGCGCGGTTGGACCGTTACTCCGTCCATTCAC
CCACAGGTTTACTCTCATCAGAGCTCACGAAAAACAGCCGTAAAAGAGGAAAATCCGATT
ATGGGACTGGGAAACTGAGCGAGAAGGGAGAGACCGGACAATGGGACACAGCGGGCTGCA
AAGCCGCGGGGCGGGGTGGGGCAAGAATTTTGTGCGCGCGCGAGCGGAAGGGGCGGGACA
GACGGGAGGTCGGACGCAGGCGAGCGGAAGGGGCGGGGACGCGCCTCAGGCCTTTGTGGC
TAATGTCCGCTACTCGGGCTGCTCCTCCGCTGTAGCTGGGGCTCCGTCCGGCGGGGAGCC
CTGCGGCAGATCCCGAAGTCGCCGCTGTCAGCTGGAGAGCCCCGTGGAGCCCCCTCGGGC
ACGCGCCTACGCTGTGCTTTAACCCGGAAGCTAAAAATAGGGTGCGTGCGGCGAGCCGCG
AGGAAGGCAGGTGTGGAGCTGGTGTCTCGCTAGCGGGCGCACGGAGCTGTGGTGTCGCCG
CTCGGGGTCCTTCCCCATGGGACTGAGGGAGAGGGAGAGAGCGGACCCGGCGACTGAGAA
GCGCAGAGGTGGTCAGCGTCTGGGTGAGAGCCGCGGGAACGCCGCCTGGGCCGCCCCAGG
CAGTACGGGGGCCAGGAGGTCCGGCGTGGCGAGGACGCAGGCTTAGTTAATGTGACCTAG
ATGGTTGTAGGGGAAGCAAGGCGTCCGGACTTGAGGCTGCTTGACAGGGCTCAGGTTTCT
CAGGTAGGGGAGTGGGGCTGGGTGAGTGTGTAACCTGGGAGACAAGACGTTATTAGTACT
GTCTCGCGCTGGCGGGTGGTGGGGAGGGCTCTCTCTACTCAGCCTTGTGAGCTCAGGGCC
CGGGTTTAACAGAAATTGAAGGAGCATACACATCTCTGTATACATCAGATCGTGTTTGTT
CCTCATCTCCAAATCTAGTAAATGGTTACTACCATATTTCGTTTGAATTTAGAAGACTCA
CGCTTTTGAGAGCTGAAGTTGACTTTTTGCGTCCATTGTATCATGTTTTTAATCCAAGCT
AAGCATTTCAGAGTTACAGAATTCCTCACTAAATTCATTCGTAAGCCAATCCTTTTATTA
ATACTGGCTCAGAACAAAGCCCACACAAACTAAGGGCCTCCATGTTTGCTGAATATTTCA
GTATATGAATGGGACCAATTGGCAAATACACAAAATTTTAAACAAATTTATGTTCAGTGC
AGATTTTAAAATGTTACTACCTTAAATGCACACCAGCACTACATGCTATAAACAGTAACT
AAAGAAAAAAAAAAATTAGAGCTTGCATTAAGTTCTGCCAAAAGCAGGGTTTGGTGCTGCA
TGCCTTTAGTCAAAGAACTGCAGGGGCAGGAAAGCTCTGAGCTCAAGGTCAGCCTACACT
ACATAGCAGGTTCCAGGACTACATAGAGGCCCTGTCTTAAAAACAAACAGGAACAATCTG
CAGCCTGTATTTTATTAAAATTTTCTGTGATTGGGGATGCGGCCCCGTGGTAGAATGCCT
GTCTAGCATGAACCTGAGTTCACTGTCAAGGAGTGGCAAAGATGTAGCTTACTAGCAAAG
TAAAAGTGTTACCTATCATTCCTAAGACCCTGAGTTCGTTCCCAGCACTGAGAAAAGTTC
CCGTTTACAGAGTTAGGAATTTAATGAGGACTGAAAGATAACTGAAAAGGAGATAATCTC
ACCATGTAATCATGTTAATATTGCATTCTTACACTTCTAAAATTTTTTCTCTCAGTGGTA
G

**ENSMUST00000067307 promoter**

Sequence: no sequence found

**Ints9 promoter**

Location: chromosome 14:64948001-64951201
Sequence:
GATAATGACTTGTGGCATAAATGCAAAATAACAACAAATTCGGGTCCTTTTCACTTAACG
ATGCATTTTATAGTGAAACCACAAAAAGCACAGCCCATATTAGCACAAACTAAGCAAAAA
AGTACAGGAGGTATTATGTCCATCGGGAAGAGAAATGTTCTGTGCCAAGGAAGGAAGGAA
GGAAGGAAGGAAGGAAGGAAGGAAGGGGGTGGGGGAGAGAGGAAGCTAAGGAAAGATAGG
AAAGAGAAGGCAAAGATTTCTATTATCTGCAGGTCCCCAGAAAAAGGGGCACGTGACTAT
ACCCTATATAACAGACGAGGGTTTTTTTAAGCTTCAGACAAAAATAAAAATAGCAATCAC
TTTTATCATTCAACTTTACGAAGCGTGTATCTTTACTGCTAAGTGCATTCTTTGCTTCAC

TCACCCAATGCTTCCTCACTCTTCCATACCAGTTTTCAAGCATTTGTGACACAGAACAGG
GTGGGGTAGAAGAAAAAACAAAACAAACAACAAAAAACAAAAAACAAAAAACAAAAAAAG
ACCACAGCGCCTTTCAGCGCACAGTCAACCCTCTTCTCCGCAAAAAATTTCCAACAGTTT
TTTCAACTACGAAAATAAGTCGCCTTATCTACCTCAGGATACAGAGCAGACGTTCTGTCT
CATTAGAAGACAGCGATGCGGATGGCATTTAGCACTCGAGATTCCCAGCCCCACGTTCCC
GAATTTTAAATAACTGGCAAGCTACTGAGCATGAAGATGACAAGATACCTCTACTTACTG
TTTCGGGGTGAAGGTGGGTGGAGGACTCCTTTCAAAGTAGCCCCAATAGTGTAACCCAAA
CACGGTTCCGGAATAAAATACTAAGAGACACTTTTCCTTTCCCACCTTTAGCTGCCGACT
TGAGGCAAAGATGCATCTCCACATCACATCATCTATATAGTCCTGCCGGGAGTTTGGGGT
GGAGACTGGCCCAGAGTAGGGCACCGAGTGTGAGCTGGGCGCCCAGACACTAACGAGTTA
CTGAGCCCCCCCCAAACCCTGCACAAGATCCTTGGACCAAAAGTCCAGCCCGGGACTCCTG
CGTGCTTCCCAAGCCCAGGCCGCTTATCTCTGACTCCAGCCATCTAACTTACCATCTACA
GACATTCCTTCAAGACGATCGCAGCACCCAAGTACACGTCACAGTTTAAAACACAGCGAG
ACACTCCCCAAACAAACCAGCACACCTCGAAACCACCTTACAGTTTCTCAAACTCCCAAG
CAGGAAGACCCCCCAACTCCAGGATTCAAACCTTCCGCCGGTCCTGAAGACCAGAGAGAA
GTTGGGAAGCCATCGACGCCTCTCCACGCCCGCCCCTCACAAACTGCCCCACCCCACAGC
CCGCGCACGGCCGCCCCCTCGGAAGTACCCCTACTTGCTCCCCTATGCTCAACCTGGCTA
CTCCAGAGACCCATCTGTGGCGCCTACTCACTTTCTCTCCCCCCCAAACATCTCGACATT
TTTTTTTCCTTAAAGTAATGAGAGAAGTGAGGAAAAGAGACTTGAGGCAGTCCAGCTTGC
AAACGACCCCACTCACCTTGGAAATCCAGCTGAGAGTTCTCTGCACCCTGCTCAGCCCCC
TCCCCAGCGTCCCAGTGCAAGGCTGAGGGAGGGAGGAGGAAGGCGGGAGGGGGAGGGGAT
CTTCACCGCGGGGCGCCCTGCGGGAGGGGCTGAGGAGGGCTCCAGGCCCGCGCAGGAAGA
GCCTAGGGGGGAGGGGAGGCCCCTATCTCCCAGTCTCCCCCTGAAAACCTCCCTTGTCTC
AATCACATCAGACTGACTGTCTTTGTCTGGCTGACAACCGCCATATTGATAGCAACAGCC
CCGCTCCCACCTGCTGCCGCTGGGCATCCCTGGGAGTTGTAGTTTCCGGTTCCCTAGAGG
AAAAGTGCAGAGTTGAGGGTCTGAGGCCCTAGGGAACTACAGCTTCCAGCAGGCACTGCG
CCCATACGCCTTCCCACAGGGAGGAGGGGAGCTGGGAGCTGAAATCGTGCGGACTACACA
TCCCAGAATACCATGGATGCTAGACCCAGATGAGGGACATGATATCATGCTGGGATGGAA
AGGAAGGATGAATTGGCGGCGGCGATGAATTTGAGCGGACGGAAGGAGGTGATTAGAGAG
AAAAGGGAGGAAAAGGAAAGACGCCCACAGTGCCCATGATACCTGCAGTAAAACTACATT
TCCCGTGAGGCTCTGGAAAATGCTTAAGGAGTGAAGACTCCAAGTCCCGCGATGCTCTGG
GCTATAGAGATTAAACCGGAAGATCCTTCCCTGTCTTTAGAAGTGAAGAGACTGGTACTG
AACAACCAAAAGTCCGATATGAAACTGGTGAGTTTCCACTACTAAGTCAGAGGGCAAAGG
TAATGAATGCTATATATCTTTGGAGTCTGTCGCCTCCACCATGATTTATCACGGAGGGAG
GTCAGCACTTAAAAATTATAAAGTGAGCTGGGTTTTGCGATTGGGATTTGAATTCCCGCT
TGCTCACATCCCACAGCTCCTGCATCTGTCAGGAGGTCTTCATTACTGTTTGGCAGATCC
CTTTGGCCGCACGTGAGGTTGTATGTTACATTTCTTCACCCTTGGGCTAAAGATAGGCTC
CAAAACTTTGGTTTGGTCATGTTTCTCAGCCTGCACCTGAGAGCTATAATAGATCATCAA
CCACTTTCCACTTAATGAGCACTTGAGTGGACTCCTACTGATATTGCGAGGCTGGATTCC
GATCTGAACAGAAAGTAGGAGCGATCAGACAGAAGGTCTAAATGTCCTGTTACTCCTAAC
TGCAGTTCGAATTACAAGAAAGAGATGCTAAAGGGTGGTGCCCTTGAAGGATTCCCGTCT
AAAGCCTGCTTGAAATAAACAGCTCTAGATAAAACATGCCTCCTGAAGAAGTAACTGTGT
GGCAAAACACTAAGCTGACAGGTGGCCTTTTCAGATATTTGCTTTGCCCCATTTTTCTGA
GCACACCGTTCTAACAAGTTGTTTTATTCAAGCTTAATGGAATAAGTAATAGAGTAATAA
GGAAATTGAACGCAGGTGTGAAATGCTTGTCGTCGCCCTCCTATTCCCTACATTCTCACT
GCCAACTTACTGTAGAATCTTCCTAATTGCTACGTGGAGAGAAATATCCTCAATAGAGTG
AAGTGTGTTTGGTATGTTCTG

**0610009D07Rik (0610009D07Rik) promoter**

Location: chromosome 12:4816600-4818800
Sequence:
TGTTTTCAGGGTAGACCTCTTCTCAGTATACCAAAAAAGCCTATGTAAGCTTGGACCTTC
GACAGGTTTATGTTCCTAAAATCCTCTGTAGCATCATAGAAGGAGCCAGCTTTGGGAGTA
TATCCCATCAGTTTCGCAGCGATAAGGCTTCAGCATTCCTTTTGTGTTAAGCTATGTAAG
TCTGTCACTTGCACTAGTATTTTACTGCAGCCTTCAGTCTTACCTGGTACAGAATGAGGG
TGCACATTTGAAAGCATGCTCTTTCCTGATGGGAAGTGAAAGGGTGTGCTGTTTTAGATA
CTCAAAACATTAGGGCTAGGCAGGATGGAGAATAACAGGGTTAAGGCCTTTCTTCAGGCT
GATAATCTTTTCCTTAGTACAAACGCCGGTGAAAACCTCCACGGAGCAGATTCTGCCTTT
CTCTGCAAACAGCCCCCAACGTCTGCTGCTGGTTCTTGATAGAGTCTACCCGGCAACTTA
TATTCAAGTGGCTAATATGTGGGGGCTAACAGGAGCGCGGGCAAAATCGTCTTGGAGCTG
CCCTAGTGAAGGAGGACCTCCGAGCAAACTCGGAAGTGGCCTCTCTTCCCATGTCCTGTA
AAGCTGGTCTAAGGAAATAAAAAGATGCATGTGACGGGTTGTGTGCTCTGTTTGCGTGGA
CAGGGGCCGCGGGAACCGGGTGAGGGAGTTCGACGGTCAGGTGAAGCGCGACCTGAGGCC
CAGGGCCCCGGGATCGGCCTTCCGAGCAAAGTTCAGATTGGGACCTCTGTTGGAAGCGCT

GGCGAGCCGGGTTCCTCTCCGAGACCTGGGCGCGCTAAATCGCAGGCTCTGGGTCTGATC
TCTCTCATGGACCCCGCCCCTCGCTCCTTAAGCCACACCCCCTCAAGCGGAACACCTCCC
TTCCCTTGGCATTCGCCCCGCCCCTTCCGGTCCGCAATCCGGAAGCGTTCTGGAGCTTGC
TAGTTCGCTTTCCGGCTTGAGAATGCGTTCTCTGTAGTGTGACTTGGAGCATCTTTTCTA
GGAAGTAAACGCGTCCTAATTTGAGACGCGGACTTTCCGGCATTGCTGTAAAAGCTTCCG
GCGCTACCTGGCTGTAAATACTCTGTATCTGACGCCGTGTCGCCCACGAAGGCGATCAGG
TTTAGAAATATCCAGAGGGATCTTACGGTAACCTACATCAGCCCGCCAAGATGGCGATGC
AAGCGGCTAAGAGGGCGAATGTGAGTACCCCGCTTTCTGCGTGACAGGGAAGCGATGTTG
TTGCTTGACTGTTGCTGCGTGCCTGGGGAGCGTTTTTTGTCCCTCCGCCCATAGTTGGGG
AAGCTTGTTATCAGTTCTGTCAAAGGAGGTTCTGTCCTGGGCTCGTCCTCAGGGGTGTGT
TCTTTGAGCGACTTTTCGCCCCACCCCTATTGAGAGCCCCCTGGCATCAGTGACCTCGTA
ACGCCTAGGACCAGGCGTTAGGCTCTTCTGAGTCCAAAAAGACATACGGAGTAACTATTT
GGTTCCTTGCCAGCTGGAGCCTGGGATTTTTCTCGAATGTAGAACTTTGGGAATCCTCTA
TTAATATTATGTAAGCACCTGCTTTGTCTCTTGTCTTACAAACGGGAAGCAATTCATCAG
TTGGGTTTATTCAACACTACTGTGTAAAGGACTAGTCGGGCTTATGTTCTACTGTTGTCA
GTGAGTGTTACCAAAACCTTTAAGGAAAAAAAAAAAACATCTTTCAGGAAATTAAGAAGTG
GTTAAAACTTAGAATCACACCAAAACAAAGACACATAATTGACCATTGGATATCCTGATG
AAGATTTTGGGCAACCGCGATTATGTGAATCATTTCAGTGAATGAAAACTTAGTTCCTGT
AAAATAAAAAGATGGTAGGAGTTAGGACAAAACTGTAGAGCTGATTCTATGATTTTTCTT
ATTAATAGAGAAATTGAATTGTTTTAAAGATACCTACAGTGCTTTTTATAGACCAACGAG
AGTGAGCCAACATAATGGTAGAGACCTGATATGTAAGATTCCAGAGTAAAACTGGAATGT
GGCTCAGTGATTGCCTAGTGGGCGTGAGTCTGGATTTGACTCCCAGCACCTCAAACCATG
ACAACCAAAAATAAGTAAAGAGGGGAGTGGTTGGTACACCAATAGAAGAAAAAAATGGTT
GGATGGACATAGATGGATAAAGTTTACAATAGGAATATATA