# Construction and application of a built-in dual luciferase reporter for microRNA functional analysis

Yanzhen Bi<sup>1</sup> 🖂 · Xinmin Zheng<sup>1</sup> · Changwei Shao<sup>2</sup> · Wen Pan<sup>3</sup> · Li Jiang<sup>2</sup> · Huiwu Ouyang<sup>2</sup>

1 Hubei Key Laboratory of Animal Embryo Engineering and Molecular Breeding, Institute of Animal Husbandry and Veterinary,

Hubei Academy of Agricultural Science, China

2 RNA Group, College of Life Science, Wuhan University, China

3 The Third Affiliated Hospital of Sun Yat-Sen University, Tianhe District, China

Corresponding author: biyanzhen@gmail.com Received October 27, 2010 / Accepted January 19, 2011 Published online: March 15, 2011 © 2011 by Pontificia Universidad Católica de Valparaíso, Chile

# ABSTRACT

**Background**: As key gene regulators, microRNAs post-transcriptionally modulate gene expression via binding to partially complementary sequence in the 3' UTR of target mRNA. An accurate, rapid and quantitative tool for sensing and validation of miRNA targets is of crucial significance to decipher the functional implications of miRNAs in cellular pathways.

**Results**: Taking advantage of an improved restriction-free cloning method, we engineered a novel built-in dual luciferase reporter plasmid where *Firefly* and *Renilla* luciferase genes were assembled in a single plasmid named "pFila". This design eliminates the transfection of a separate control plasmid and thus minimizes the time and labor required for miRNA-target sensing assays. pFila consistently produces *Firefly* and *Renilla* luciferase activities when transfected into human-, monkey- and mouse-derived mammalian cell systems. Moreover, pFila is capable of recapitulating the interaction of miR-16 and its known target CCNE1 in Hela cells. Additionally, pFila is shown to be a sensitive miR-biosensor by evaluating the inhibition efficiency of endogenous miRNA.

**Conclusions**: pFila would facilitate miRNA target identification and verification in a rapid and simplified manner. Also, pFila is a sensitive biosensor for active miRNA profiling *in vivo*.

Keywords: biosensor, luciferase, ligase-independent, miRNA, target

## INTRODUCTION

MicroRNAs are key post-transcriptional regulators of gene expression in a variety of cellular events. They mediate translational repression, and sometimes destabilization, of target mRNAs by directing miRISC (microRNA-induced silencing complex) to imperfect complementary sequences in 3' UTR (Bartel, 2009). It has been predicted that more than 60% of human genes are putative targets of one or more miRNAs, while it is also suggested that an individual miRNA is capable of regulating multiple target mRNAs (Backes et al. 2010). Consequently, a major challenge in miRNA study is the experimental identification and validation of its functional target(s). Various reporter systems have been developed to probe the interaction between individual miRNA and its target (Lee et al. 2008), of which dual luciferase assay is widely adopted to achieve this end (Brennecke et al. 2005). The current dual luciferase assay encompasses two separate plasmids, one containing the region of interest and the other serving as internal control to normalize transfection variation (Robertson et al. 2010). A major drawback of this system is the tedious steps for preparing and transfecting control vector. A more convenient reporter is desired to simplify the conventional protocol.

In addition, there are several methods, including northern blotting, real-time PCR, microarray and deep sequencing, to quantify miRNAs (Willenbrock et al. 2009). However, these techniques simply output the homeostasis of endogenous miRNAs other than active molecules. Moreover, these approaches are either time-consuming and laborious or cost-ineffective. Therefore, a convenient and quantitative miRNA biosensor is desired to measure functional miRNA *in vivo*.

In this study, by taking advantage of ligase-free homologous recombination in *E. coli*, we engineered a novel reporter that integrated *Firefly* and *Renilla* luciferase genes (*Fluc* and *Rluc*, same below) in a single plasmid. Its expressivity and applicability were further examined to demonstrate that this novel reporter will facilitate the screening and sensing of miRNAs and their targets in a simplified and precise manner.

## MATERIALS AND METHODS

## DNA and RNA oligos

Unless stated elsewhere, all DNA and RNA oligos are presented as 5'-3' direction. Primers for amplifying Fluc gene were Pf, AAGGATCCAGGTGGCACTTTTCG TGCGATCTGCATCTCAATTAG; Pr, GAAAAATAAACAAATAGGGGTTCCGCGCAC CTCACATGTTCTTTCCTGC (sequence annealing to Fluc gene was shown in boldface; sequence complementary to insertion site on pRL-TK was underlined). 3' outermost primer P2R was CGAAAAGTGCCACCTGGATCCTT. Sequencing primers for pFila was SF, GATGCACCTG ATGAAATGGG; SR, AGGACAGGTG CCGGCAGCGC. For creation of Apal site, see details in reference (Wang et al. 2009). RNA oligos were chemically synthesized and purified by Genepharma Co. Ltd., (Shanghai, P.R. China). Human miR16-1 was sense UAGCAGCACGUAAAUAUUGGCG and antisense CGCCAAUAUUUACGUGCUGCUA. Negative control for miRNA mimics was sense UUGUACUACAAAAGUACUG and antisense CAGUACUUUUGUGUAGUACAA. siRNA Rluc mRNA against sense was GUAGCGCGGUGUAUUAUACdTdT and antisense GUAUAAUACACCGCGCUACdTdT. Methylated anti-miR-16-1 inhibitor CGC CAA UAU UUACGU GCU GCU A, scramble anti-miR control UUG UAC UAC ACA AAA GUA CUG.

#### **Construction of plasmids**

pFila was fabricated by bridging-PCR coupled with homologous recombination in bacteria. A duplex bridging PCR was conducted in a 50 µl mixture: Pf 4 µl (250 nM), Pr 2 µl (125 nM), P2R primer 2 µl (125 nM), pGL3-promoter plasmid 1 µl (5 ng), modified pRL-TK plasmid 1 µl (10 ng or 50 ng), 2 mM dNTP 5 µl, 25 mM MgSO<sub>4</sub> 2 µl, 10 x KOD buffer 5 µl, KOD plus 1 µl (1 unit), PCR-grade water 27 µl. The condition was: 95°C 2 min, 30 cycles of (95°C 15 sec, 55°C 30 sec, 68°C 6.5 min). The PCR products were digested with *DpnI* (Fermentas, Lithuania) at 37°C for 2 hrs to destroy methylated plasmids while keeping the nascent DNA intact with the following reaction: PCR products, 26 µl; 10 x Tango buffer 3 µl; *Dpn I* 1 µl (1 unit). An aliquot of 5 µl digested PCR products were transformed into *E. coli* DH5α to generate recombinants that were subsequently sequenced to verify the integrity. Pf and Pr primer pair was used to amplify *Fluc* gene composed of SV40 promoter, *Fluc* coding region and SV40 late poly(A) signal. *Fluc* gene was designed to fuse into a modified pRL-TK plasmid downstream of 3'UTR of *Rluc* and upstream of beta-lactamase gene. For the sequence context of human CCNE1 target regions were sub-cloned into pFila with *Xba I* and *Apa I*.

#### Cell culture, transfection and dual luciferase assay

Human cervical carcinoma Hela cells, African green monkey kidney Vero cells and mouse myoblast C2C12 cells were maintained in high glucose DMEM (Invitrogen) supplemented with 10% fetal calf serum (Gibco) at 37°C and 5% CO<sub>2</sub>. 4 x 10<sup>4</sup> cells were seeded in a 24-well plate one day before transfection. For miRNA mimics and plasmid co-transfection, 1 µl 20 µM chemically synthesized miR16 mimics and 50 ng pFila (pFila-CCNE1-wildtype and pFila-CCNE1-mut1&2) or 50 ng pGL3-promoter (internal control) and pRL-ML plasmids (pRL-ML-CCNE1-wildtype and pRL-ML-CCNE1-mut1&2; 25 ng each) were mixed with 2 µl Lipofectamine2000 (Invitrogen) as transfection complex. For evaluation of endogenous miRNA inhibition with pFila, 20 nM 2'-O-methylated anti-human miR16-1 inhibitor was transfected into Hela cells by 0.5 µl Lipofectamine2000 (Invitrogen), RNAiMAX (Invitrogen), Sofast

(Sunmabio, China), Fugene (Roche), respectively. All transfections were performed in three independent experiments with each in triplicate. A DLR<sup>TM</sup> Assay (Promega) was adopted to measure luciferase activity in a Glomax luminometer essentially according to manufacturer's instruction.

#### Statistical analysis

Luciferase levels were reported as ratio over that observed in control transfections, where *Rluc* activities were normalized to *Fluc* activities. The data represented the mean  $\pm$  S.D. of three independent experiments and were analyzed by Student's *t*-test. Differences below *p* < 0.01 were regarded as significant.

#### Accession number

The sequence and annotation of pFila has been deposited in Genbank with accession number HQ425563. pFila is freely available upon request.



**Fig. 1 Concept and engineering of pFila.** (a) In this reaction, the primary process was the production of *Fluc* gene by Pf (Pf is composed of P1 and P2, which are complementary to upstream of *Fluc* gene and pRL-TK plasmid, respectively) and Pr (Pr is composed of P3 and P4, which are complementary to downstream of *Fluc* gene and pRL-TK plasmid, respectively) primers. In the secondary process, amplified *Fluc* gene bearing a deliberately designed region annealed with its homologous sequence in pRL-TK, thus generating a large linear fragment. In order to exponentially amplify the linear fragment, the outermost Pf and P2R (complementary to P2) primer pair initiated the ternary process to produce adequate fused fragment. (b) Duplex bridge PCR. The 2.4 kb fragment was *Fluc* gene, while the 6.4 kb fragment was the desired fusion product. M is 1 kb DNA ladder. (c) Map of pFila. The engineered pFila is 6486 bp in length. *Xbal* and *Apal* restriction sites are located in the 3'UTR of *Rluc* to facilitate the cloning of miRNA target for reporter assay. (d) 5' and 3' seaming sites of *Fluc* into pRL-TK. For detailed information, please refer to Gene sequence 1.

## **RESULTS AND DISCUSSION**

The complicated procedures of the current dual luciferase reporter assay for miRNA target screening prompted us to upgrade its practicality for simplified manipulations. Specifically, we aimed to integrate *Fluc and Rluc* genes in a single vector. As restriction sites were not available to sub-clone *Fluc* gene to pRL-TK plasmid with ligase-dependent method, we adopted an improved restriction-free gene fusion approach inspired by the principle of site-directed mutagenesis (Zheng et al. 2004). As shown in Figure 1a and Figure 1b, a duplex bridging PCR was carried out to produce the 6.4 kb linear fusion fragment

with homologous sequences at both 5' and 3' ends. The PCR products were digested by *Dpnl* and transformed into *E. coli* to achieve the circular plasmid based on homologous recombination. Sequencing of the recombinant (named pFila) revealed that *Fluc* gene had been successfully fused into pRL-TK plasmid at designed location (Figure 1c and Figure 1d; Gene sequence 1).

We then evaluated the expressivity of pFila in different mammalian systems. As presented in Figure 2a, luminescence of pFila was reported in a wide linear range when transfected into human-, mouseand monkey-sourced cell lines at gradient amounts, indicating that pFila consistently produces luciferases in vivo. This also implies that the ordered assembly of Fluc and Rluc luciferase genes in pRL-TK plasmid does not interfere with their individual expression. Next, the applicability and reproducibility of pFila were examined by recapitulating the regulation of human miR16-1 and its known target CCNE1 (Wang et al. 2009). miR16-1 mimics down-regulated the Rluc activity fused with wildtype CCNE1 3'UTR but not a mutant 3'UTR (Figure 2b); the latter carried altered residues that were introduced in the miR16-1 "seed-pairing" recognition site (Figure 2b). This observation perfectly photocopied the result that was achieved by traditional dual reporter assay (Figure 2b), indicating that pFila as a more convenient reporter is fully applicable to miRNA functional analysis. Finally, we applied pFila carrying the wild-type and mutant 3'TUR of CCNE1 to assessing the blockage efficiency of endogenous miR16-1. 2'-O-methylated anti-miRNA-16-1 RNA oligo was transfected by four types of transfection reagents, i.e. Lipofectamine2000, RNAiMAX, Fugene, Sofast. Inhibition efficacy of endogenous miR16-1 varied, of which RNAiMAX achieved the most potent blocking effect. This assay implies that pFila is a sensitive miRNA biosensor to reflect the level of functional miRNAs. It also suggests that the choice of delivery method is an important determinant when conducting loss-offunction analysis of miRNAs.

In summary, we have successfully engineered a novel dual luciferase plasmid that incorporated *Fluc* and *Rluc* genes in a single vector, allowing the simultaneous expression of both luciferase genes. This improvement maintains comparable reproducibility but minimizes the time and labor required in conventional dual luciferase protocol. Furthermore, several lines of evidence were presented to demonstrate its application in miRNA functional analysis. These results indicate that pFila will find its wide application in the screening, identification and validation of miRNA with its potential mRNA targets.



**Fig. 2 Expressivity and application of pFila.** (a) pFila consistently produces *Fluc* and *Rluc* activities in human, mouse- and monkey-sourced mammalian cell lines. pFila was gradientlytransfected into the cells stated above and *Fluc* and *Rluc* luciferase levels were determined by a DLR<sup>TM</sup> Assay (Promega). As shown in this graph, *Fluc* and *Rluc* activities were in a linear range for all the selected points. (b) pFila is capable of recapitulating the interaction of miR16 and its known target CCNE1. Wild-type and mutant CCNE1 3'UTRs were sub-cloned into pFila and co-transfected with miR16-1 mimics. Conventional dual luciferase reporter was conducted in parallel to compare their reproducibility. siRNA against *Rluc* serves as positive control. *Rluc* with mutant CCNE1 3'UTR was rescued in comparison with pFila-CCNE1-3'UTR-wildtype, indicating that CCNE1 is a direct target of miR16 as previously reported (Wang et al. 2009). (c) miR16-1 inhibitors were co-transfected with pFila-CCNE1-3'UTR-wildtype or mutant the selected agents, where RNAiMAX achieved the most potent inhibition. It implies that pFila is a sensitive biosensor for functional miRNA profiling.

Gene sequence 1. Sequencing of the seaming sites: sequence annealing to Fluc gene was shown in boldface; sequence complementary to insertion site on pRL-TK was underlined.

#### Seq.1 (sequencing primer SF: GATGCACCTGATGAAATGGG )

 ${\tt TTTTTAAATCCAATTCGGTTGAACGAGTTCTCAAAAATGAACAATAATTCTAGATTCCGGGAATATCGGTAATGGGC$  ${\tt CCTAGAGCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGA$ AAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTA TAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATCGCTGACT AATTTTTTTTTTTTTTTTTTGCAGAGGCCCGAGGCCGCCTCCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCCTTTTTT GGAGGCCTAGGCTTTTGCAAAAAGCTTGGCATTCCGGTACTGTTGGTAAAGCCACCATGGAAGACGCCAAAAAACAT AAAGAAAGGCCCGGCGCCATTCTATCCGCTGGAAGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGA TACGCCCTGGTTCCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGGACATCACTTACGCTGAGTACTTCG AAATGTCCGTTCGGTTGGCAGAAGCTATGAAACGATATGGGCTGAATACAAATCACAGAATCGTCGTATGCAGTGA TATAATGAACGTGAATTGCTCAACAGTATGGGCATTTCGCAGCCTACCGTGGTGTTCGTTTCCAAAAAGGGTTGCA AAAAATTTTTGAACGTGCAAAAAAGCTCCCAATCATCCAAAAAATTATTATCATGGATTCTAAAACGGATTACCAGG GGATTCAGTCGATGTACACGTCGTCACATCTCATCTACCTTCCCGGTTTATGATACGATTTTGTGCAAGAGTCTTC GATAGGACAAGACAATTGCACTGATCATGACCTCTCTGATCTACTGGACTGCTAGGTGCGACCTTGTCTCATAGAC TGCCTGGCCGGTAGAAATTCTCGCC

#### Seq.2 (sequencing primer SR: AGGACAGGTG CCGGCAGCGC)

#### Seq.3 pFila sequence

Base pair:	6486bp	
HSV TK promoter	7-759	
Chimeric intron	826-962	
T7 RNA polymerase Promoter (-17 1	to +2)	1006-1024
T7 RNA polymerase transcription ini	itiation site	1023
Rluc reporter gene	1034-196	69
Xbal restriction site	1971-197	76
Apal restriction site	1995-200	00
SV40 late polyadenylation signal(up	stream)	2039-2240
SV40 promoter	2279-2481	
Fluc reporter gene	2511-41	63
SV40 late polyadenylation signal(do	wnstream)	4195-4416
Beta-lactamase (AmpR)	4800-	5660
pBR322 plasmid replication origin	5	802-6445

1	AGATCTAAAT	GAGTCTTCGG	ACCTCGCGGG	GGCCGCTTAA	GCGGTGGTTA
51	GGGTTTGTCT	GACGCGGGGG	GAGGGGGAAG	GAACGAAACA	CTCTCATTCG
101	GAGGCGGCTC	GGGGTTTGGT	CTTGGTGGCC	ACGGGCACGC	AGAAGAGCGC
151	CGCGATCCTC	TTAAGCACCC	CCCCGCCCTC	CGTGGAGGCG	GGGGTTTGGT
201	CGGCGGGTGG	TAACTGGCGG	GCCGCTGACT	CGGGCGGGTC	GCGCGCCCCA
251	GAGTGTGACC	TTTTCGGTCT	GCTCGCAGAC	CCCCGGGCGG	CGCCGCCGCG
301	GCGGCGACGG	GCTCGCTGGG	TCCTAGGCTC	CATGGGGACC	GTATACGTGG
351	ACAGGCTCTG	GAGCATCCGC	ACGACTGCGG	TGATATTACC	GGAGACCTTC
401	TGCGGGACGA	GCCGGGTCAC	GCGGCTGACG	CGGAGCGTCC	GTTGGGCGAC

451	AAACACCAGG	ACGGGGCACA	GGTACACTAT	CTTGTCACCC	GGAGGCGCGA
501	GGGACTGCAG	GAGCTTCAGG	GAGTGGCGCA	GCTGCTTCAT	CCCCGTGGCC
	CGTTGCTCGC				
	CTTTAGTTCT				
	ATTCGAACAC				
	CATATTAAGG CCCGCTTAAA				
	CAGAAGTTGG				
	AGGTTTAAGG				
	TTGCGTTTCT				
951	TCTCTCCACA	GGTGTCCACT	CCCAGTTCAA	TTACAGCTCT	TAAGGCTAGA
1001	GTACTTAATA	CGACTCACTA	TAGGCTAGCC	ACCATGACTT	CGAAAGTTTA
	TGATCCAGAA				
	GATGTAAACA				
	GAAAAACATG				
	TTCTTATTTA				
	GTATTATACC GGTTCTTATA				
	ACTTCTTAAT				
	CTTGTTTGGC				
	ATAGTTCACG				
	GCCTGATATT				
1551	AAATGGTTTT	GGAGAATAAC	TTCTTCGTGG	AAACCATGTT	GCCATCAAAA
1601	ATCATGAGAA	AGTTAGAACC	AGAAGAATTT	GCAGCATATC	TTGAACCATT
1651	CAAAGAGAAA	GGTGAAGTTC	GTCGTCCAAC	ATTATCATGG	CCTCGTGAAA
	TCCCGTTAGT				
	TATAATGCTT				
	ATCGGACCCA				
	TTCCTAATAC GATGCACCTG				
	TCTCAAAAAT				
	TAGAGCGGCC				
	CAAACCACAA				
2101	TGATGCTATT	GCTTTATTG	TAACCATTAT	AAGCTGCAAT	AAACAAGTTA
ムエリエ					
	АСААСААСАА			AGGTTCAGGG	GGAGGTGTGG
2151		TTGCATTCAT	TTTATGTTTC		
2151 2201 2251	ACAACAACAA GAGGTTTTTT GGATCCAGGT	TTGCATTCAT AAAGCAAGTA GGCACTTTTC	TTTATGTTTC AAACCTCTAC GTGCGATCTG	AAATGTGGTA CATCTCAATT	AAATCGATAA AGTCAGCAAC
2151 2201 2251 2301	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCCG	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCTAACTC	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCCATCCC	AAATGTGGTA CATCTCAATT GCCCCTAACT	AAATCGATAA AGTCAGCAAC CCGCCCAGTT
2151 2201 2251 2301 2351	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCCG CCGCCCATTC	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCTAACTC TCCGCCCCAT	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCCATCCC CGCTGACTAA	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTTAT	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG
2151 2201 2251 2301 2351 2401	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCCG CCGCCCATTC GCCGAGGCCG	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCTAACTC TCCGCCCCAT CCTCGGCCTC	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCCATCCC CGCTGACTAA TGAGCTATTC	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTAT CAGAAGTAGT	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGGAGGCTT
2151 2201 2251 2301 2351 2401 2451	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCCG CCGCCCATTC GCCGAGGCCG TTTTGGAGGC	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCTAACTC TCCGCCCCAT CCTCGGCCTC CTAGGCTTTT	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCCATCCC CGCTGACTAA TGAGCTATTC GCAAAAAGCT	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTTAT CAGAAGTAGT TGGCATTCCG	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG
2151 2201 2251 2301 2351 2401 2451 2501	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCCG CCGCCCATTC GCCGAGGCCG TTTTGGAGGC TAAAGCCACC	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCTAACTC TCCGCCCCAT CCTCGGCCTC CTAGGCTTTT ATGGAAGACG	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCCATCCC CGCTGACTAA TGAGCTATTC GCAAAAAGCT CCAAAAACAT	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG CCGGCGCCAT
2151 2201 2251 2301 2351 2401 2451 2501 2551	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCCG CCGCCCATTC GCCGAGGCCG TTTTGGAGGC	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCTAACTC TCCGCCCCAT CCTCGGCCTC CTAGGCTTTT ATGGAAGACG GGAAGATGGA	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCCATCCC CGCTGACTAA TGAGCTATTC GCAAAAAGCT CCAAAAACAT ACCGCTGGAG	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC AGCAACTGCA	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG
2151 2201 2251 2301 2351 2401 2451 2501 2551 2601	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCCG CCGCCCATTC GCCGAGGCCG TTTTGGAGGC TAAAGCCACC TCTATCCGCT	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCTAACTC TCCGCCCCAT CCTCGGCCTC CTAGGCTTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCCATCCC CGCTGACTAA TGAGCTATTC GCAAAAACAT ACCGCTGGAG TGGAACAATT	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC AGCAACTGCA GCTTTTACAG	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT
2151 2201 2251 2301 2401 2451 2501 2551 2601 2651	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCG CCGCCCATTC GCCGAGGCCG TTTTGGAGGC TAAAGCCACC TCTATCCGCT AAGAGATACG	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCTAACTC TCCGCCCCAT CCTCGGCCTC CTAGGCTTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC GCAAAAAGCT CCCAAAACAT ACCGCTGGAG TGGAACAATT CTGAGTACTT	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC AGCAACTGCA GCTTTTACAG CGAAATGTCC	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG
2151 2201 2251 2301 2401 2451 2501 2551 2601 2651 2701 2751	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCCG CCGCCCATTC GCCGAGGCCG TTTTGGAGGC TAAAGCCACC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAAGCTAT TGCAGTGAAA	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCTAACTC TCCGCCCCAT CTAGGCTTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAAACGATAT ACTCTCTTCA	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC CCAAAAAGCT CCAAAAAGCT ACCGCTGGAG TGGAACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC AGCAACTGCA GCTTTTACAG CGAAATGTCC CAAATCACAG CCGGTGTTGG	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GCGCGTTATT
2151 2201 2251 2301 2401 2451 2501 2551 2601 2651 2701 2751 2801	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCCG CCGCCCATTC GCCGAGGCCG TTTTGGAGGC TAAAGCCACC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAAGCTAT TGCAGTGAAA TATCGGAGTT	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCTAACTC TCCGCCCAT CTTCGGCCTC CTAGGCTTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAAACGATAT ACTCTCTTCA GCAGTTGCGC	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC CCAAAAAGCT ACCGCTGGAG TGGAACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCGAACGA	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC AGCAACTGCA GCTTTTACAG CGAAATGTCC CAAATCACAG CCGGTGTTGG CATTTATAAT	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GCGCGTTATT GAACGTGAAT
2151 2201 2251 2301 2451 2451 2551 2601 2651 2701 2751 2801 2851	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCCG CCGCCCATTC GCCGAGGCCG TTATGGAGGC TCAAAGCCACC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAAGCTAT TGCAGTGAAA TATCGGAGTT TGCTCAACAG	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCTAACTC TCCGCCCCAT CCTCGGCCTC CTAGGCTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAAACGATAT ACTCTCTTCA GCAGTTGCGC	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC CCAAAAAACAT ACCGCTGGAG TGGAACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCGAACGA TCGCAGCCTA	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTTAT CAGAAGTAGT TGGCATTCG AAGAAAGGC AGCAACTGCA GCTTTTACAG CGAAATGTCC CAAATCACAG CCGGTGTTGG CATTTATAAT CCGTGGTGTT	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GCGCGTTATT GAACGTGAAT CGTTTCCAAA
2151 2201 2251 2301 2451 2551 2601 2551 2601 2751 2701 2751 2801 2851 2901	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCCG CCGCCCATTC GCCGAGGCCG TTATGGAGGC TCAAAGCCACC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAGCTGAA TGCCGAGGTT TGCTCAACAG AAGGGTTGC	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCTAACTC TCCGCCCCAT CCTCGGCCTC CTAGGCTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAACATAT ACTCTCTTCA GCAGTTGCGC TATGGGCATT AAAAAATTTT	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC CCAAAAAACAT ACCGCTGGAG TGGAACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCGAACGA TCGCAGCCTA GAACGTGCAA	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTTAT CAGAAGTAGT TGGCATTCCG AAGCAACTGCA GCTTTTACAG CGAAATGTCC CAAATCACAG CCGGTGTTGG CATTTATAAT CCGTGGTGTT AAAAAGCTCC	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GCGCGTTATT GAACGTGAAT CGTTTCCAAA CAATCATCCA
2151 2201 2251 2301 2451 2501 2551 2601 2651 2701 2751 2801 2851 2901 2951	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCAG CCGCCCATTC GCCGAGGCCG TTTTGGAGGCC TCTATCCGCT AAGAGATACG CGAGGGTGGAC CAGAAGCTAT TGCAGTGAGTT TGCTCAACAG AAGGGGTTGC AAAAATTATT	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCGCCCCAT CCTCGGCCTC CTAGGCTTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAAACGATAT ACTCTCTTCA GCAGTTGCGC TATGGGCATT AAAAAATTTT ATCATGGATT	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC GCAAAAACAT ACCGCTGGAG TGGAACAATT CTGAGTACTT GGCTGAATA ATTCTTTATG CCGCGAACGA CCGCGAACGA CGAACGTGCAA CTAAAACGGA	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC AGCAACTGCA GCTATTACAG CGAAATGTCC CAAATCACAG CCGTGTTGG CATTTATAAT CCGTGGTGTT AAAAAGCTCC TTACCAGGGA	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GGGCGTTATT GAACGTGAAT CGTTTCCAAA CAATCATCCA TTTCAGTCGA
2151 2201 2251 2351 2451 2551 2551 2601 2651 2701 2851 2801 2851 2901 2951 3001	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCGG CCGCCCATTC GCCGAGGCCG TTTTGGAGGCC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAAGCTAT TGCAGTGAAA TATCGGAGTT TGCTCAACAG AAGGGTTGC AAAAATTATT TGTACACGTT	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCGCCCCAT CCTCGGCCTC CTAGGCTTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAAACGATAT ACTCTCTTCA GCAGTTGCGC TATGGGCATT AAAAAATTTT ATCATGGATT CGTCACATCT	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAT GGACTATTC CCAAAAACAT ACCGCTGGAG TGGAACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCGAACGA TCGCAGCCTA GAACGTGCAA CTAAAACGGA CATCTACCTC	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC AGCAACTGCA GCTTTTACAG CGAAATGTCC CAAATCACAG CAGTGTTGG CATTTATAAT CCGTGGTGTT AAAAAGCTCC TTACCAGGGA	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GCGCGTTATT GAACGTGAAT CGTTTCCAAA CAATCATCCA TTCCAGTCGA
2151 2201 2251 2301 2401 2551 2551 2601 2651 2701 2751 2851 2851 2901 3001 3051	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCAG CCGCCCATTC GCCGAGGCCG TTTTGGAGGCC TCTATCCGCT AAGAGATACG CGAGGGTGGAC CAGAAGCTAT TGCAGTGAGTT TGCTCAACAG AAGGGGTTGC AAAAATTATT	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCTAACTC CCTCGGCCTC CTAGGCTTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAAACGATAT ACCATTGCGC TATGGCATT AAAAATTTT ATCATGGATT ACCATGGATT ACATGGATT	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC GCAAAAAGCT CCAAAAACAT ACCGCTGGAG TGGACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCGGAACGA CCGCGGACGA GAACGTGCAA CTAAAACGGA CATCTACCTC TAGGGACAAG	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC AGCAACTGCA GCTTTTACAG CGAAATGTCC CAAATCACAG CCGGTGTTGG CATTTATAAT CCGTGGTGTT AAAAAGCTCC TTACCAGGGA CCGGTTTTAA ACAATTGCAC	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GGGCGTTATT GAACGTGAAT CGTTTCCAAA CAATCATCCA TTCAGTCGA TGAATACGAT
2151 2201 2251 2301 2401 2551 2551 2551 2601 2751 2801 2751 2801 2901 2951 3001 3051 3101	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCCG CCGCCCATTC GCCGAGGCCG TTTTGGAGGCC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAAGCTAT TGCAGTGAAA TATCGGAGTT TGCTCAACAG AAGGGTTGC AAAAATTATT TGTACACGTT TTTGTGCCAG	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCCTAACTC CTCGGCCCCAT CTCGGCCTCC CTAGGCTTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAAACGATAT ACTCTCTTCA GCAGTTGCGC TATGGCCATT AAAAATTTT ATCATGGATT CGTCACATCT AGTCCTTCGA	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC CCAAAAACAT ACCGCTGGAG TGGAACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCGAACGA CACCAGCCTA GAACGTGCAA CTAAAACGGA CATCTACCTC TAGGGACAAG TGCCTAAAGG	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC AGCAACTGCA GCTTTTACAG CGAAATGTCC CAAATCACAG CCGGTGTTGG CATTTATAAT CCGGTGTTGT AAAAAGCTCC TTACCAGGA ACCATTGCAC	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GGGCGTTATT GAACGTGAAT CGATCATCAA TTCCAGTCGA TGAATACGAT TGATCATGAA CCTCATAGAA
2151 2201 2251 2301 2401 2451 2551 2601 2751 2601 2751 2801 2851 2901 2951 3001 3051 3101	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCCG CCGCCCATTC GCCGAGGCCG TTTTGGAGGC TAAAGCCACC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAAGCTAT TGCAGTGAAA TATCGGAGTT TGCTCAACAG AAGGGTTGC AAAATTATT TGTACACGT TTTGTGCCAG CTCCTCTGGA	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCCTAACTC TCCGCCCCAT CTAGGCTTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAAACGATAT ACTCTCTTCA GCAGTTGCGCC TATGGGCAT AAAAATTTT ATCATGGATT CGTCACATCT AGTCCTTCGA TCTACTGGTC GAGATTCTCG	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC CCAAAAAGCT CCAAAAACAT ACCGCTGGAG TGGAACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCGAACGA CCGCGAACGA CTAAAACGGA CTAAAACGGA CACTCACCTC TAGGGACAAG TGCCTAAAGG CATGCCAGAG	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC AGCAACTGCA GCTTTTACAG CGAAATGTCC CAAATCACAG CCGGTGTTGG CATTTATAAT CCGGTGTTGT AAAAAGCTCC TTACCAGGGA CCGGTTTTAA ACAATTGCAC TGTCGCTCTG ATCCTATTT	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GCGCGTTATT GAACGTGAAT CGATCATCAA TTCCAGTCGA TGAATACGAT TGATCATGAA CCTCATAGAA TGGCAATCAA
2151 2201 2251 2301 2451 2501 2551 2601 2651 2701 2751 2801 2901 2951 3001 3051 3101 3151 3201	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCCG CCGCCCATTC GCCGAGGCCG TTTTGGAGGC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAAGCTAT TGCAGTGAAA TATCGGAGTT TGCTCAACAG AAGGGTTGC AAAATTATT TGTACACGT TTTGTGCCAG CTCCTCTGGA	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCTAACTC TCCGCCCCAT CTCGGCCTC CTAGGCTTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAAACGATAT ACTCTCTTCA GCAGTTGCGC TATGGGCATT AAAAAATTT ATCATGGATT CGTCACATCT AGTCCTTCGA TCTACTGGTC GAGATTCTCG ATACTGCGAT	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC CCAAAAAGCT CCAAAAACAT ACCGCTGGAG TGGAACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCGAACGA CTAAAACGG CATGCCAAAG CACCTCACTC TAGGGACAAG TGCCTAAAG CATGCCAGAG TTTAAGTGTT	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC AGCAACTGCA GCTTTTACAG CGAAATGTCC CAAATCACAG CCGGTGTTGG CATTTATAAT CCGTGGTGTT AAAAAGCTCC TTACCAGGGA CCGGTTTTAA ACAATTGCAC GTCCTATTT GTTCCATTCC	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GCGCGTTATT GAACGTGAAT CGTTTCCAAA CAATCATCCA TTGATCAGAT TGATCATGAA CCTCATAGAA TGGCAATCAA ATCACGGTTT
2151 2201 2251 2301 2451 2501 2551 2601 2651 2701 2851 2801 2851 2801 2951 3001 3051 3101 3151 3201 3251 3301	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCAG CCGCCCATTC GCCGAGGCCG TTTTGGAGGCC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAGGTGGA TGCCAGCAGATT TGCTCAACAG AAGGGGTTGC AAAAATTATT TGTACACGTT TTTGTGCCAG CTCCTCTGGA CTGCCTGCGT ATCATTCCGG TGGAATGTTT TAATGTATAG	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCGCCCCAT CCTCGGCCTC CTAGGCTTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAAACGATAT ACTCTCTTCA GCAGTTGCGC TATGGGCATT AAAAAATTTT ATCATGGATT CGTCACATCT AGTCCTTCGA TCTACTGCGAT ACTACTGCGAT ACTACACTCG ATATTGAAGAA	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAT GGACTATTC GCAAAAACAT ACCGCTGGAG TGGAACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCGGACGA CCGCGAACGA CTAAAACGGA CATCTACCTC TAGGGACAAG TGCCTAAAG CATGCCAAGG CATGCCAAGAG TTTAAGTGTT GATATTGAT GAGCTGTTC	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC AGCAACTGCA GCTTTTACAG CGAAATGTCC CAAATCACAG CCGGTGTTGG CATTTATAAT CCGTGGTGTT AAAAAGCTCC TTACCAGGGA CGGTTTTAA ACAATTGCAC TGTCGCTCTG ATCCTATTT TGAGGAGCCT	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG CCGGCGCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA CGCGCTTATT GAACGTGAAT CGTTTCCAAA CTCATCAGAA TGAATACGAT TGATCATGAA CCTCATAGAA TGCAATCAA ATCACGGTTT CGAGTCGTCT TCAGGATTAC
2151 2201 2251 2301 2451 2501 2551 2601 2651 2601 2651 2801 2851 2801 2901 2951 3001 3051 3101 3151 3201 3251 3301 3351	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCAG CCGCCCATTC GCCGAGGCCG TTTTGGAGGCC TCTATCCGCT AAGAGATACG CGAGGGTGGAC CAGAAGCTAT TGCCAGCAGATT TGCTCAACAG AAGGGGTTGC AAAAATTATT TGTACACGTT TTTGTGCCAG CTCCTCTGGA CTGCCTGCGT ATCATTCCGG TGGAATGTTT TAATGTATAG AAGATTCAAA	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCGCCCCAT CCTCGGCCTC CTAGGCTTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAACGATAT ACTCTCTCA GCAGTTGCGC TATGGGCATT AAAAAATTTT ATCATGGATT CGTCACATCT AGTCCTTCGA TCTACTGGC ATACTGCGCT ATCACATCG ATACTGCGAT ACTACACCTCG ATACTGCGAT	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC GCAAAAACAT ACCGCTGGAG TGGAACAATT CTGAGTACTT GGCTGAATA ATTCTTTATG CCGCGAACGA CACCGTGCAA CAACGTGCAA CTAAAACGGA CATCTACCTC TAGGGACAAG TTCAAGGAT TTAAGTGTT GATATTGAT GACTGTTC GGTGCCAACC	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC GCTTTTACAG CGAAATGTCC CAAATCACAG CGGTGTTGG CATTTATAAT CCGTGGTGTT AAAAAGCTCC TTACCAGGGA CCGGTTTTAA ACAATTGCAC GTCCCTGT ATCCATTCT GTCCCATTCT TGAGGAGCCT CTATTCTCCT	AAATCGATAA AGTCAGCAAC CCGCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GCACGTCATA CGTTTCCAAA CGATCATCCA TGAATACGAT TGATCATGAA TGGCAATCAA ATCACGGTTT CGAGCCATCA ATCACGGTT TCAGGATTAC TCAGGATTAC TCTCGCCAA
2151 2201 2251 2301 2401 2551 2551 2601 2551 2601 2651 2701 2851 2901 3001 3051 3101 3151 3201 3351 33401	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCCG CCGCCCATTC GCCGAGGCCG TTTTGGAGGCC TAAGCCACC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAAGCTAT TGCCGAACAG AAGGGTTGC AAAAATTATT TGTACACGGT TTTGTGCCAG CTCCTCTGGA CTCCTCTGGA CTGCCTGCGT ATCATTCCGG TGGAATGTTT TAATGTATAG AAGATTCAAA AAGCACTCTG	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCCTAACTC CCCGGCCCCAT CCTCGGCCTC TATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAACGATAT ACTCTCTTCA GCAGTTGGGC TATGGGCATT ACTACTGGAT ACTACTGGAT ACTACTGGTC GAGATTCTCG ATACTGCGAT ACTACCGAT ACTACACTCG ATTGACAAAT	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC GCAAAAACAT ACCGCTGGAG TGGACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCGAACGA CCGCGCACCA CACCACCC TAAGAGCAAG TTTAAGTGT GACTGTTTC GAGCTGTTTC GACTGTTTC GACGCCAACC ACGATTTATC	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC GCTTTTACAG CGAAATGTCC CAAATCACAG CCGGTGTTGG CCGGGTGTTGG CCGGGTGTTAA ACAATGCAC TTACCAGGGA CCGGTTTTAA ACAATTGCAC TGTCGCTCTG ATCCTATTT GTTCCATTCC ATGTGGATTT TGAGGAGCCT CTATTCTCCT TAATTACAC	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GACGTGAAT CGTTTCCAAA CAATCATCCA TGAATACGAT TGAATACGAT TGATCATGAA CCTCATAGAA ATCACGGTTT CGAGTCGTCT TCAGGATTAC TCAGGCATCA CCTCCACAA ATCACGGTT
2151 2201 2251 2301 2401 2551 2551 2601 2551 2601 2751 2851 2901 3051 3001 3051 3101 3151 3201 3351 3301 3351 3401 3451	ACAACAACAA GAGGTTTTT GGATCCAGGT CATAGTCCGG CCGCCCATTC GCCGAGGCCG TTTTGGAGGCC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAAGCTAT TGCAGTGAACAG AAGGGGTTGC AAAAATTAT TGTACACGT ATTGTGCCAG CTCCTCTGGA CTGCCTCCGG TGGAATGTT TAATGTATAG AAGATTCAA AAGCACTCTG CTGGTGGCGC	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCTAACTC TCCGCCCCAT CTCGGCCTCC TATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAACGATAT ACCATTGCGC TATGGGCATT ACAAGGAT ACTACTGGAT ACTACTGGAT ACTACTGGAT ACTACTGGAT ACTACACTCG ATTGAAGAA GTGCGCTGCT ATTGACAAAT TCCCCTCTC	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC GCAAAAAGCT CCAAAAACAT ACCGCTGGAG TGGACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCGGACGA CACCAGCCTA GAACGTGCAA CTTAAACGGA CTTAAACGGA TTGAGGACAAG TTGAGGACAAG TTAAGTGTT GATGCTGTTC GAGCTGTTCC ACGCTGTTCC ACGATTTATC AAGGAAGTCG	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC GCTTTTACAG CGAAATGTCC CAAATCACAG CCGGTGTTGG CCGGTGTTGG CCGGTGTTGA ACAATGCCC TTACCAGGGA CCGGTTTTAA ACAATTGCAC GTCCATTCC ATGTCGATTT TGAGGAGCCT CTATTCTCCT TAATTTACAC GGGAAGCGGT	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GAGCGTTATT GAACGTGAAT CGTTTCCAAA CAATCATCCA TTCAGGCGA TGAATACGAT TGATCATGAA CCTCATAGAA TGGCAATCAA ATCACGGTTT CGAGTCGTCT TCAGGATTAC TCTCGCCAA GAAATTGCTT TGCCAAGAGG
2151 2201 2251 2301 2401 2551 2601 2551 2601 2751 2801 2951 2901 2951 3001 3051 3101 3151 3201 3301 3351 3401 3451 3501	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCGG CCGCCCATTC GCCGAGGCCG TTTTGGAGGC TCTAACCACC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAAGCTAT TGCAGTGAACA TATCGGAGTGC AAAAATTATT TGTACACGT TGTACACGT TTTGTGCCAG CTCCTCTGGA CTGCCTGCGT ACATTCCAG AAGATTCAAA AAGCACTCTG CTGGTGGCGC TTCCATCTGC	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCCTAACTC TCCGCCCCAT CTCGGCCTC CTAGGCTTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAAACGATAT ACTCTCTTCA GCAGTTGCGCATT AAAAATTTT ATCATGGATT CGTCACATCG AGTCTCTCG ATCACTGCGT ATCACACTCG ATACTGCGAT ACTACACTCG ATTTGAAGAA GTCCCTCTCT CAGGTATCAG	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC GCAAAAAGCT CCAAAAACAT ACCGCTGGAG TGGACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCGACGA CAACGTGCAA CAACGTGCAA CAACGTGCAA CATCTACTC TAGGGACAAG TTTAAGTGTT GATGCCAACG GCGCCAACC ACGATTTATC AAGGAATCTG	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC AGCAACTGCA GCTTTTACAG CGAAATGTCC CAAATCACAG CCGGTGTTGG CATTTATAAT CCGGTGTTGT AAAAAGCTCC TTACCAGGGA CCGGTTTTA ACAATTGCAC TGTCCATTCC ATGTGGATTT TGAGGAGCCT CTATTCTCCT TAATTACAC GGGAAGCGGT GGGCTCACTG	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GAACGTGAAT CGTTTCCAAA CAATCATCCA TTCAGTCGA TGAATACGAT TGATCATGAA CCTCATAGAA TGGCAATCAA ATCACGGTTT CGAGTCGTCT TCAGGATTAC TCTCGCCAA GAAATGCTT TGCCAAGAGG AGACTACATC
2151 2201 2251 2301 2401 2551 2601 2651 2701 2751 2801 2951 3001 3051 3001 3151 3201 3351 3401 3451 3551	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCGG CCGCCCATTC GCCGAGGCCG TTTTGGAGGC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAAGCTAT TGCAGTGAAA TATCGGAGTGC AAGAGTTACT TGTACACGT TTTGTGCCAG CTCCTCTGGA CTGCTCTGGA AGGATTCTT TAATGTATAG AAGATTCAA AAGCACTCTG CTGGTGGCGC TTCCATCTGC AGCTATTCG	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCCTAACTC CTCGGCCCCAT CTCGGCCTCC CTAGGCTTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAAACGATAT ACTCTCTTCA GCAGTTGCGCT TATGGCATTC AGTCCTTCGA TCTACTGGTC GAGATTCTCG ATACTGCGAT ATTGACAAAT TCCCCTCTCT CAGGTATCAG ATTACACCCG ATACACCCG	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC CCAAAAACAT ACCGCTGGAG TGGACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCGAACGA CTCACGCCAACGA CTCACGCCAACGA TTAAGTGTT GATCTTACT GAGCTGTTC CAGGCCAACG CTTAAGTGTT GAGCTGTTC GGTGCCAACC ACGATTTATC AAGGAAGTCG GCAACGATTA	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC AGCAACTGCA GCTTTTACAG CGAAATGTCC CAAATCACAG CCGGTGTTGG CATTTATAAT CCGGTGTTGA ACAATTGCAC TGTCCATTCC ATGTGGATTT GTTCCATTCC ATGTGGACCT CTATTCTCCT TAATTTACAC GGGAAGCGGT GGGCTCACTG TAAACCGGGC	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GCGCGTTATT GAACGTGAAT CGTTTCCAAA CAATCATCCA TTCAGTCGA TGCATACAA ATCATCAA ATCATCAA ATCATCAA ATCACGGTTT CGAGTCGTCT TCAGGATTAC TCTCGCCAA GAAATTGCTT TGCCAAGAGG AGACTACATC GCGGTCGGTA
2151 2201 2251 2301 2401 2451 2551 2601 2751 2601 2751 2801 2751 2801 2951 3001 3051 3001 3151 3201 3351 3401 3451 3501 3551 3601	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCGG CCGCCCATTC GCCGAGGCCG TTTTGGAGGC TCTAACCACC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAAGCTAT TGCAGTGAACA TATCGGAGTGC AAAAATTATT TGTACACGT TGTACACGT TTTGTGCCAG CTCCTCTGGA CTGCCTGCGT ACATTCCAG AAGATTCAAA AAGCACTCTG CTGGTGGCGC TTCCATCTGC	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCCTAACTC TCCGCCCCAT CTAGGCTTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAAACGATAT ACTCTCTTCA GCAGTTGCGC TATGGGCATT AAAAATTT ATCATGGAT CGTCACATCT GAGATTCTCG ATACTGCGAT ACTACTGCGAT ACTACACCG ATATTGACAAAT TCCCCTCTCT CAGGTATCAG ATTACACCCG ATTACACCCG ATATACACCCG ATATACACCCG ATATTTGACAAA	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC CCAAAAAGCT CCAAAAACAT ACCGCTGGAG TGGAACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCGACCAA CTAAACGGACAAG TCGCCAACGA CTTAAGTGTT GATATTGAT GACGTGTTC GGTGCCAACC ACGATTAC AAGGAAGTCG GCAAGGATAT AGGGGATGA	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC AGCAACTGCA CGAAATGTCC CAAATCACAG CCGGTGTTGG CATTTATAAT CCGTGGTGTT AAAAAGCTCC TTACCAGGGA CCGGTTTAA ACAATTGCAC TGTCCATTCC ATGTGGATTT TGAGGAGCCT CTATTCTCCT TAATTTACAC GGGACCGGG TGGATCTGGA	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GCGCGTTATT GAACGTGAAT CGTTTCCAAA CAATCATCCA TTCAGTCGA TGAATACGAT TGAACAGAA TGGCAATCAA ATCACGGTTT CGAGTCGTCT TCAGGATTAC TCTTCGCCAA GAAATTGCTT TGCCAAGAGG AGACTACATC GCGGTCGGTA TACCGGGAAA
2151 2201 2251 2301 2451 2551 2551 2551 2601 2751 2801 2751 2801 2751 2851 2901 2951 3001 3051 3051 3251 3301 3551 3551 3601 3651	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCCG CCGCCCATTC GCCGAGGCCG TTTTGGAGGC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAAGCTAT TGCAGTGAAA TATCGGAGTGC AAAATTATT TGTACACGT TTTGTGCCAG CTCCTCTGGA CTCCTCTGGA CTGCTCTCGG TGGAATGTTT TAATGTATAG AAGATTCAAC AAGATTCAAC CTGCTGCGCGC CTGCCACCTC ACAACTCTG AAGATTCAAC AGCACTCTG AGCTATTCG AGCTATTCG AGCTATTCG	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCTAACTC TCCGCCCCAT CTAGGCTTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAACGATAT ACTCTCTTCA GCAGTTGCGC TATGGGCATT AGTCCTTCGA AGTCCTTCGA AGTCCTTCG AGACTGCGAT AGTCCTCCG ATACTGCGAT ATTGACAAAT TCCCCTCCT CAGGGATCAG ATTACACCCG ATTACACCCG ATTACACCCG ATTACACCCG ATTACACCCG ATTACACCCG	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC CCAAAAAGCT CCAAAAACAT ACCGCTGGAG TGGAACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCGAACGA CTAAACGTGCAA CTAAAACGG CAACGTGCAAG TTAAGTGTT GATATTGAT GACCTAAAG CATGCCAAGG TTTAAGTGTT GATATTGAT GAGCTGTTTC GGTGCCAACC ACGATTATC AAGGAGATAT AGGGGATGA GCGAAGGTG AGGCGAACTG	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC AGCAACTGCA GCTTTTACAG CGAAATGTCC CAAATCACAG CCGGTGTTGG CATTTATAAT CCGTGGTGTT AAAAGCTCC TTACCAGGGA CCGGTTTTAA ACAATTGCAC TGTCGCTCTG ATCCTATTT GTTCCATTCC TGAGGAGCCT CTATTCTCCT TAATTTACAC GGGAAGCGGT GGGCTCACTG TGAACCGGGC TGGATCTGGA	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GCGCGTTATT GAACGTGAAT CGTTTCCAAA CAATCATCCA TTCAGTCGA TGAACATCATCA TGCAATACATC CCTCATAGAA ATCACGGTTT CCAGGATTAC TCAGGATTAC TCACAGAGG GAAATTGCTT TGCCAAGAGG AGACTACATC GCGGTCGGTA TACCGGGAAA GTCCTATGAT
2151 2201 2251 2301 2451 2501 2551 2601 2551 2601 2651 2701 2851 2801 2901 2951 3001 3051 3101 3151 3201 3251 3301 3401 3451 3551 3601 3551 3701 3751	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCGG CCGCCCATTC GCCGAGGCCG TTTGGAGGCC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAAGCTAT TGCCCAACAG AAGGGGTGC AAAAATTATT TGTACACGT ATCATCCGG CTCCTCTGGA CTGCCTGCGT ATCATTCCGG TGGAATGTTT TAATGTATAG AAGATTCAAA AAGCACTCTG CTGGTGGCGC TTCCATCTGC AGGTTGTCCG AGGTTGTCCG TATGTCCGGT AGGATGGATG	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCCTAACTC CCCGGCCCCAT CCTCGGCCTC CTAGGCTTTT ATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAACGATAT ACTCTCTTCA GCAGTTGCGC TATGGGCATT AATAAAATTTT ATCATGGATATC AGTCCTTCGA ATCACTGGT ATACACGGAT ACTACACTCG ATACTGCGAT ACTACACTCG ATTTGAAGAA GTGCGCTGCT ATTGACAAAT TCCCCTCTCT CAGGTATCAG ATTACACCCG ATTACACCCG ATTACAACCG ATTACAACCA TTAATCAAAG TATGTAAACA GCTACATTCT	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC GCAAAAACAT ACCGCTGGAG TGGACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCGAACGA CACCGTGCAA CACCGTGCAA CACCTACCTC TAGGGACAAG TTTAAGTGTT GACTGTTC GACGTGCTAC GACGTGTTC GGTGCCAACC ACGAAGATCG AGCGAACTG AGCGAACTG ATCCGGAACG GCAAGCACAG	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC GCTTTTACAG CGAAATGTCC CAAATCACAG CCGGTGTTG CCATTTATAAT CCGTGGTGTT AAAAAGCTCC TTACCAGGGA CCGGTTTTAA ACAATTGCAC TGTCGCTCTG ATCCTATTT GTTCCATTCC TGAGGAGCCT CTATTCTCCT TAACTACC GGGAAGCGGT GGGCTCACTG TGAACCGGC TGGATCTGGA GGCCAACGCC CTTACTGGA	AAATCGATAA AGTCAGCAAC CCGCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GCGCGTTATT GAACGTGAAT CGTTTCCAAA CGATCATCAA TTCAGTCGA TGATCATGAA TGACATGAA ATCACGGTT CCAGGGATACA CCTCATAGAA ATCACGGTCT TCAGGCATTAC TCAGGGATAC GCGCAGTA GAAATTGCTT TGCCAAGAGG AGACTACAT CGCGGAAA GCCCTATGAT
2151 2201 2251 2301 2401 2551 2551 2601 2551 2601 2751 2851 2901 3051 3001 3051 3101 3151 3201 3351 3301 3401 3551 3601 3551 3601 3751 3701 3751 3801	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCGG CCGCCCATTC GCCGAGGCCG TTTGGAGGCC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAGGTGGA CAGAGGTTGC AAGAGGTTGC AAGAGGTTGC AAAAATTATT TGTACACGGT ATGCCAGGG CTCCTCTGGA CTGCCTCCGG TGGAATGTTT AAAGATTCAAA AAGCACTCTG CTGGTGGCGC TTCCATCTGC AGGTAGTCC AGGTGGCGG TATGTCCGGT AGGATGGATG CACTTCTCA	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCCTAACTC CCCGGCCCCAT CCTCGGCCTC TATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAACGATAT ACTCTCTTCA GCAGTTGCGC TATGGCATT ACTACTGGAT ACTACTGGAT ACTACTGGAT ACTACTGGAT ACTACACCG ATTGACAAAT TCCCCTCTT CAGGTATCAG ATTACACCG ATTACAAAA	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC GCAAAAACAT ACCGCTGGAG TGGACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCGAACGA CACGCGCACGA CACCAGCCTA GACCGTGCAAG TTTAAGTGTT GATATTGAT GACTGTTTC GGTGCCAACC ACGATTATC AAGGAAGTCG ACGGAAGTTA AGGCGAACG ATCCGCAACG CCAAGGATAA	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC GCTTTTACAG CGAAATGTCC CAAATCACAG CCGGTGTTGG CCGGTGTTGG CCGGGTGTTGA ACAATGCAC TTACCAGGGA CCGGTTTTAA ACAATTGCAC TGTCGATTT GTCCATTCC ATGTGGATTT TGAGGAGCGT CTATTCTCCT TAATTTACAC GGGACGGC GGGCTCACTG GAACCGGGC TGGATCTGGA GACCAACGC CTTACTGGAA CTGATTAAGT	AAATCGATAA AGTCAGCAAC CCGCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG CCGCCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GACGTGAT CGTTTCCAAA CAATCATCCA TTCAGTCGA TGAATACGAT TGATCATGAA ATCACGGTT CGGCCATAGAA ATCACGGTT CCAGGATTAC TCTCGCCAA GAAATTGCTT TGCCAAGAGG AGACTACATC GCGGTCGTA TACCGGGAA ATCACGGAA ACCAAGACGAA ACCAAGGCTA
2151 2201 2251 2301 2401 2551 2551 2601 2551 2601 2751 2851 2901 3051 3001 3051 3001 3051 3101 3251 3301 3351 3401 3551 3651 3701 3751 3801 3751	ACAACAACAA GAGGTTTTT GGATCCAGGT CATAGTCCGG CCGCCCATTC GCCGAGGCCG TTTGGAGGCC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAGGTGGA CAGAGGTTGC AAGAGGTTGC AAAAATTATT TGTACACGT ATCGGCTGCGT ATCATCCGG TGGAATGTTT TAATGGAATG AAGATTCAAA AAGCACTCTG CTGCTGCGC TTCCATCTGC AGGTAGGCG TATGTCCGGT AGGATGGATG CACTTCTCA TCAGGTGGCT	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCCTAACTC CCCGGCCCCAT CCTCGGCCTC TATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAACGATAT ACCACTACG CATTGGGCATT ACTACTGGCT ATCATGGAT TCACTGGT ATCACTGGT AGTCCTTCG AGATTCTCG AGATTCTGAAGAA GTGCGCTGCT ATTGACAAAT TCCCCTCTCT CAGGTATCAG ATTACACCCG ATTACACCCG ATTACACCCG ATTACACCCG ATTACACCCG ATTACACCCG ATTACACCCG ATTACACCCG ATTACACCCG ATTACACCCG ATTACACCCG ATTACAACAC	TTTATGTTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC GCAAAAGCT CCAAAACAT ACCGCTGGAG TGGACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCGCACGA CTCAACGGCCAA CTAAAACGGA CATCTACCTC TAGGGACAAG TTTAAGTGTT GATCTGAT GACCTGTTC GGTGCCAACC ACGATTATC AAGGAAGTCG ACGCGAAGCTG ACGCGAACTA AGGCGAACTA AGGCGAACTA AGGCGAACTA ACCGGAAGCT ACCGGAACCT TTGAATCCAT	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC AGCAACTGCA GCTTTTACAG CGAAATGTCC CAAATCACAG CCGGTGTTGG CCGGTGTTGG CCGGTGTTGA ACAATGCAC TTACCAGGGA CCGGTTTTAA ACAATTGCAC GTCCATTCC ATGTGGATTT TGAGGAGCCT TAATTTACAC GGGACCACTG TAACCGGGC TGATCTGGA CTGATTAAGT CTTGCTCCAA	AAATCGATAA AGTCAGCAAC CCGCCAGTT TTATGCAGAG GAGGAGGCTT GTACTGTTGG CCGCCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GACGTGAT CGTTTCCAAA CAATCATCCA TTCAGTCGA TGAATACGAT TGATCATGAA ATCACGGTT CGAGCATCA GAAATTGCTT TCAGGATTAC TCTCGCCAA GAAATTGCTT TGCCAAGAGG AGCTATGAT GCGAGAGAA ACCAAGGCTA CCAAGGCTA CCAAGACGAA ACAAAGGCTA
2151 2201 2251 2301 2401 2551 2601 2551 2601 2751 2851 2901 3051 3001 3051 3001 3051 3001 3051 3001 3351 3401 3551 3601 3551 3601 3751 3801 3851 3801	ACAACAACAA GAGGTTTTTT GGATCCAGGT CATAGTCCGG CCGCCCATTC GCCGAGGCCG TTTGGAGGCC TCTATCCGCT AAGAGATACG CGAGGTGGAC CAGAGGTGGA CAGAGGTTGC AAGAGGTTGC AAGAGGTTGC AAAAATTATT TGTACACGGT ATGCCAGGG CTCCTCTGGA CTGCCTCCGG TGGAATGTTT AAAGATTCAAA AAGCACTCTG CTGGTGGCGC TTCCATCTGC AGGTAGTCC AGGTGGCGG TATGTCCGGT AGGATGGATG CACTTCTCA	TTGCATTCAT AAAGCAAGTA GGCACTTTTC CCCCCTAACTC CCCGGCCCCAT CTCGGCCTCC TATGGAAGACG GGAAGATGGA CCCTGGTTCC ATCACTTACG GAACGATAT ACCACTTCTCA GCAGTTGCGC TATGGGCATT AAAAAATTTT ATCATGGAT ACTACTGGAT ACTACTGCGAT ACTACACTCG ATTACACCG ATTTGAAGAA GTGCGCTGCT ATTGACAAAT TCACTCCTCTC CAGGTATCAG ATTACACCCG	TTTATGTTC AAACCTCTAC GTGCGATCTG CGCCATCCC CGCTGACTAA TGAGCTATTC GCAAAAACAT ACCGCTGGAG TGGACAATT CTGAGTACTT GGGCTGAATA ATTCTTTATG CCGCAGCCTA GAACGTGCAA CTTAAACGGA CTTAAACGGA CTTAAACGGA CATCTACCTC TAGGGACAAG TTTAAGTGTT GATGTCTTCC GGTGCCAACC ACGATTTATC AAGGAAGTCG ACGCGAAGGTG ACGCGAAGGTG ACCGGAAGGTG ACCGGAAGCTG ACCGGAAGCT GGAACCATAG CCTGAAGCCT TGGAATCCAT GGTCTTCCCG	AAATGTGGTA CATCTCAATT GCCCCTAACT TTTTTTTAT CAGAAGTAGT TGGCATTCCG AAAGAAAGGC GCTTTTACAG CGAAATGTCC CAAATCACAG CCGGTGTTGG CCGGTGTTGG CATTTATAAT CCGGGTGTTGA AAAAGCTCC TTACCAGGA CCGGTTTTA ACAATTGCAC TGTCGATCT GTTCCATTCC ATGTGGGATCT GGGCTCACTG GGGCTCACTG GGGCTCACTG TAAACCGGGC TGGATCTGGA CTGATTAAGT CTGCTCCAA CTGATTAAGT CTGCTCCAA	AAATCGATAA AGTCAGCAAC CCGCCCAGTT TTATGCAGAG GAGAGGCTT GTACTGTTGG CCGGCGCCAT TAAGGCTATG ATGCACATAT GTTCGGTTGG AATCGTCGTA GACGTGAAT CGTTCCAAA CAATCATCCA TTCAGTCGA TTCAGTCGA TGATCATGAA CCTCATAGAA TGCAATCAA ATCACGGTT TCAGGATCAT CGAGTCGTCT TCAGGATACCA GAAATTGCTT TGCCAAGAGG AGACTACATC GCGGTCGGTA TTCATGACA CCCCAACA CCACAGACGAA ACAAGGCTA

4001	GATCGTGGAT	TACGTCGCCA	GTCAAGTAAC	AACCGCGAAA	AAGTTGCGCG
4051	GAGGAGTTGT	GTTTGTGGAC	GAAGTACCGA	AAGGTCTTAC	CGGAAAACTC
4101	GACGCAAGAA	AAATCAGAGA	GATCCTCATA	AAGGCCAAGA	AGGGCGGAAA
4151	GATCGCCGTG	TAATTCTTGA	GTCGGGGCGG	CCGGCCGCTT	CGAGCAGACA
4201	TGATAAGATA	CATTGATGAG	TTTGGACAAA	CCACAACTAG	AATGCAGTGA
4251	AAAAATGCT	TTATTTGTGA	AATTTGTGAT	GCTATTGCTT	TATTTGTAAC
4301	CATTATAAGC	TGCAATAAAC	AAGTTAACAA	CAACAATTGC	ATTCATTTTA
4351	TGTTTCAGGT	TCAGGGGGAG	GTGTGGGAGG	TTTTTTAAAG	CAAGTAAAAC
4401	CTCTACAAAT	GTGGTAAAAT	CGATAAGGAT	CCGTCGACCG	ATGCCCTTGA
4451	GAGCCTTCAA	CCCAGTCAGC	TCCTTCCGGT	GGGCGCGGGG	CATGACTATC
4501	GTCGCCGCAC	TTATGACTGT	CTTCTTTATC	ATGCAACTCG	TAGGACAGGT
4551	GCCGGCAGCG	CTCTTCCGCT	TCCTCGCTCA	CTGACTCGCT	GCGCTCGGTC
4601	GTTCGGCTGC	GGCGAGCGGT	ATCAGCTCAC	TCAAAGGCGG	TAATACGGTT
4651	ATCCACAGAA	TCAGGGGATA	ACGCAGGAAA	GAACATGTGA	GGTGCGCGGA
4701	ACCCCTATTT	GTTTATTTT	CTAAATACAT	TCAAATATGT	ATCCGCTCAT
4751	GAGACAATAA	CCCTGATAAA	TGCTTCAATA	ATATTGAAAA	AGGAAGAGTA
4801	TGAGTATTCA	ACATTTCCGT	GTCGCCCTTA	TTCCCTTTTT	TGCGGCATTT
4851	TGCCTTCCTG	TTTTTGCTCA	CCCAGAAACG	CTGGTGAAAG	TAAAAGATGC
4901	TGAAGATCAG	TTGGGTGCAC	GAGTGGGTTA	CATCGAACTG	GATCTCAACA
4951	GCGGTAAGAT	CCTTGAGAGT	TTTCGCCCCG	AAGAACGTTT	TCCAATGATG
5001	AGCACTTTTA	AAGTTCTGCT	ATGTGGCGCG	GTATTATCCC	GTATTGACGC
5051	CGGGCAAGAG	CAACTCGGTC	GCCGCATACA	CTATTCTCAG	AATGACTTGG
5101	TTGAGTACTC	ACCAGTCACA	GAAAAGCATC	TTACGGATGG	CATGACAGTA
5151	AGAGAATTAT	GCAGTGCTGC	CATAACCATG	AGTGATAACA	CTGCGGCCAA
5201	CTTACTTCTG	ACAACGATCG	GAGGACCGAA	GGAGCTAACC	GCTTTTTTGC
5251	ACAACATGGG	GGATCATGTA	ACTCGCCTTG	ATCGTTGGGA	ACCGGAGCTG
	AATGAAGCCA			ACCACGATGC	
5351	GGCAACAACG	TTGCGCAAAC	TATTAACTGG	CGAACTACTT	ACTCTAGCTT
	CCCGGCAACA				TGCAGGACCA
	CTTCTGCGCT				
	AGCCGGTGAG			TGCAGCACTG	
	GTAAGCCCTC				
	ATGGATGAAC				
5651		CTGTCAGACC			TAGATTGATT
	TAAAACTTCA			AGGTGAAGAT	
	AATCTCATGA		TTAACGTGAG	TTTTCGTTCC	
	AGACCCCGTA			TTGAGATCCT	
5851		CTGCTTGCAA		CACCGCTACC	
5901		ATCAAGAGCT		TTTCCGAAGG	
5951		CAGATACCAA		TCTAGTGTAG	
6001	GCCACCACTT		GTAGCACCGC	CTACATACCT	CGCTCTGCTA
6051			TGCCAGTGGC	GATAAGTCGT	
	GTTGGACTCA			GGCGCAGCGG	
6151		GTGCACACAG			
6201		TACAGCGTGA			
6251		GACAGGTATC		CAGGGTCGGA	
6301		GCTTCCAGGG			
6351				TTTTTGTGAT	
6401		CTATGGAAAA			TTACGGTTCC
6451	TGGCCTTTTG	CTGGCCTTTT	GCTCACATGG	CTCGAC	

# ACKNOWLEDGMENTS

We appreciated the technical assistance from Cell Signaling Lab of Wuhan University.

**Financial support:** This work was funded by a research grant from Hubei Key Laboratory of Animal Embryo Engineering and Molecular Breeding to Y. Z. Bi (2010ZD163).

## REFERENCES

BACKES, C.; MEESE, E.; LENHOF, H.P. and KELLER, A. (2010). A dictionary on microRNAs and their putative target pathways. *Nucleic Acids Research*, vol. 38, no. 13, p. 4476-4486. [CrossRef]
BARTEL, D.P. (2009). MicroRNAs: Target recognition and regulatory functions. *Cell*, vol. 136, no. 2, p. 215-233.

BARTEL, D.P. (2009). MicroRNAs: Target recognition and regulatory functions. *Cell*, vol. 136, no. 2, p. 215-233. [CrossRef] Bi et al.

BRENNECKE, J.; STARK, A.; RUSSELL, R.B. and COHEN, S.M. (2005). Principles of microRNA-target recognition. *PLoS Biology*, vol. 3, no. 3, p. e85. [CrossRef]

LEE, J.Y.; KIM, S.; HWANG, D.W.; JEONG, J.M.; CHUNG, J.K.; LEE, M.C. and LEE, D.S. (2008). Development of a dual-luciferase reporter system for *in vivo* visualization of microRNA biogenesis and posttranscriptional regulation. *The Journal of Nuclear Medicine*, vol. 49, no. 2, p. 285-294. [CrossRef]

ROBERTSON, B.; DALBY, A.B.; KARPILOW, J.; KHVOROVA, A.; LEAKE, D. and VERMEULEN, A. (2010). Specificity and functionality of microRNA inhibitors. *Silence*, vol. 1, no. 1, p. 10. [CrossRef]

WANG, F.; FU, X.D.; ZHOU, Y. and ZHANG, Y. (2009). Down-regulation of the cyclin E1 oncogene expression by microRNA-16-1 induces cell cycle arrest in human cancer cells. *BMB Report*, vol. 42, no. 11, p. 725-730.

WILLENBROCK, H.; SALOMON, J.; SØKILDE, R.; BARKEN, K.B.; HANSEN, T.N.; NIELSEN, F.C.; MØLLER, S. and LITMAN, T. (2009). Quantitative miRNA expression analysis: Comparing microarrays with nextgeneration sequencing. RNA, vol. 15, no. 11, p. 2028-2034. [CrossRef]

ZHENG, L.; BAUMANN, U. and REYMOND, J.L. (2004). An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Research*, vol. 32, no. 14, p. e115. [CrossRef]

How to cite this article:

BI, Y.; ZHENG, X.; SHAO, C.; PAN, W.; JIANG, L. and OUYANG, H. (2011). Construction and application of a built-in dual luciferase reporter for microRNA functional analysis. *Electronic Journal of Biotechnology*, vol. 14, no. 2. <u>http://dx.doi.org/10.2225/vol14-issue2-fulltext-9</u>