Contents lists available at ScienceDirect

### Journal of Hazardous Materials

journal homepage: www.elsevier.com/locate/jhazmat

# Molecular cloning of *Bombyx mori* cytochrome P450 gene and its involvement in fluoride resistance

Hongliang Zhou, Keping Chen\*, Qin Yao, Lu Gao, Yong Wang

Institute of Life Sciences, Jiangsu University, No. 301 Xuefu Road, Zhenjiang 212013, PR China

### ARTICLE INFO

Article history: Received 7 June 2007 Received in revised form 26 December 2007 Accepted 3 March 2008 Available online 8 March 2008

Keywords: Bombyx mori Fluoride P450 FDD-PCR Quantitative real-time PCR (Q-PCR)

#### ABSTRACT

To investigate the effects of fluorosis on development and gene expression profiles of silkworm, highly resistant silkworm strain 441, and highly susceptible silkworm strain 440 were treated with 200 ppm fluoride (designated as 440F and 441F) and water (designated as 440DZ and 441DZ). Fluorotic silkworm showed body color and behavior changes. Statistical analysis indicated that growth index of 440F was lower than 440DZ, 441DZ, and 441F. The mortality of 440F was higher than others. Fluorescent differential display enabled us to obtain a differentially expressed cDNA. Bioinformatics analyses indicated that it belonged to cytochrome P450 family, denoted *Bmcyp306a1*, which contained seven exons and six introns. Phylogenetic tree showed BmCYP306A1 had high homology with *Manduca sexta* 'P450 protein. Expression analysis indicated that *Bmcyp306a1* was exclusively expressed in 441DZ and 441F and was down-regulated under fluoride treatment. The tissue-specific expression indicated *Bmcyp306a1* had high-expression level in midgut and ovary in 441F. The data revealed that there was obvious dose–effect and times relationship with the pathological changes and gene expression. Expression profiles of *Bmcyp306a1* suggested that P450 gene was crucial to physiological modification and might be involved in fluoride resistance.

© 2008 Elsevier B.V. All rights reserved.

### 1. Introduction

As a major pollutant in industrial areas, fluoride is present in varying amount in the air, water, and in some of the agricultural products produced in the polluted areas. Fluoride of high concentration has toxic effects on both humans and animals, causing toxification such as endemic fluorosis [1] and industrial fluorosis [2,3]. Fluorosis causes damage to many animal and human organs [1,4,5], predominately the skeletal systems and teeth. Simultaneously, the structures and functions of the nonskeletal systems such as brain, liver, kidney, and spinal cord are also damaged [6–8]. Previous studies have demonstrated that the damage caused by fluoride intoxication is mainly mediated through lipid peroxidation. Yet little is known about the mechanism of detoxification.

Many studies have proved that cytochrome P450 was involved in the production and metabolism of many molecules with important physiological functions and was believed to play important roles in detoxification. Currently, the cytochrome P450 superfamily consists of over 5500 designated sequences, in which approximately 1300 sequences were found in animals (http://drnelson.utmem.edu/cytochrome p450.html) [9]. Many P450 genes have multiple functions while a vast majority of cytochrome P450 have unknown functions. In human, 15 cytochrome P450 proteins can metabolize xenobiotics and all of them are from *CYP1*, *CYP2*, and *CYP3* families [10]. Due to involvement of the cytochrome P450 in physiological functions, many of them have been studied in detail in different tissues and cell types. The temporal and tissue-specific expression and regulation are very important to cytochrome P450 involved in steroid and eicosanoid biosynthesis and catabolism of vitamins [11]. It may be valuable to predict the function of cytochrome P450 on tissue- and cell-specific expression.

Silkworm is one of the economically important insects. Silk industry plays an important role in China. However, silk production has been seriously affected by fluoride pollution. In this study, highly resistant silkworm strain 441 [12] and highly susceptible silkworm strain 440 were used as materials to analyze differential expression of related genes and to find out whether the crucial genes were related to fluoride detoxification. Using FDD (fluorescent differential display) and *in silico* cloning methods, a full-length cDNA sequence encoding the putative P450 protein was obtained and named as *Bmcyp306a1*. Expression analysis suggested that *Bmcyp306a1* gene might play an important role in resistance to fluoride. Our study provides new information on the involvement of P450 protein in fluoride detoxification.



<sup>\*</sup> Corresponding author. Tel.: +86 511 8791923; fax: +86 511 8791923. *E-mail address*: kpchen@ujs.edu.cn (K. Chen).

<sup>0304-3894/\$ -</sup> see front matter © 2008 Elsevier B.V. All rights reserved. doi:10.1016/j.jhazmat.2008.03.001



**Fig. 1.** The change of body weight. According to data in Table 1, the growth index of body weight was analysed.

### 2. Materials and methods

### 2.1. Fluoride

Sodium fluoride (NaF): 200 ppm.

### 2.2. Silkworm strains

Two silkworm strains were used. Strain 440 was highly susceptible to fluoride. Strain 441 was highly resistant to fluoride. Strain 441 was maintained in our laboratory through adding fluoride in its fodder. Its fluoride-tolerance ability was 200 ppm at the time of experimentation. All larvae were fed with mulberry leaves three times a day at  $25 \pm 2$  °C under a 12 h light/12 h dark cycle and were maintained in two separate groups. All larvae were raised up to the fifth instar. Starting from the fifth instar, the larvae were fed with mulberry leaves treated with either clean water (control) or with 200 ppm NaF solution for 5 min. Then, 50 larvae of similar size were taken for investigating their physiological changes, respectively. Larvae of strains 440 and 441 feeding on leaves treated with water were designated as 440DZ and 441DZ, respectively. Those feeding on leaves treated with NaF were designated as 440F and 441F.

### 2.3. Pathological change investigation

Pathological changes were investigated from fifth instar to pupal stage. The observed items included body weight, body color, behavior, individual size, and mortality. The body weight was measured at 7:00 a.m. before they were fed with mulberry leaves. Growth index was calculated by using the body weight at the test time to divide the body weight at the beginning of experiment. For example, if the weight of a larva is 2 g at 7:00 a.m. of the first day of fifth instar and the same larva weighs 3 g at 7:00 a.m. of the second day, the growth index will be 1.50 which is 3 divided by 2.



**Fig. 2.** Statistics of silkworm mortality. Larvae in groups 440DZ and 441DZ grew normally and did not die. Those of group 440F were found to die from 96 to 168 h with the average number of death increasing from 2 to 7, while those of group 441F was found to die from 120 h, with the number of death being below 2.



**Fig. 3.** Result of the FDD analysis from only one pair of arbitrary primers. Lane 1, 440DZ; lane 2, 440F; lane 3, 441DZ; lane 4, 441F; lane 1 and lane 3, from midguts after fluoride treatment for 48 h; lane 2 and lane 4, from midguts after water treatment for 48 h. The differential expression band A9 is indicated by an arrow. A9 was expressed in fluoride resistant strain 441.

### 2.4. Total RNA extraction and reverse transcription

At 48 h of the fifth instar, 10 silkworms were sampled from each group for isolating total RNAs. Total RNAs were isolated from the midguts using Trizol (Invitrogen) reagent. The total RNAs were treated with RNase-free DNase (Promega, WI). The RNA precipitates were dissolved in 50 mL of 10 mM Tris/1 mM EDTA buffer (pH 8.5) and spectrophotometrically quantified at 260 nm. The integrity of the isolated RNA was examined on an ethidium bromide stained, denaturing 1% agarose gel. Single stranded cDNAs for PCR were synthesized from total RNAs ( $2 \mu g$  for all tissues) with MMLV reverse transcriptase (Promega, WI).

### 2.5. Fluorescent differential display analysis

For FDD–PCR (RNAspectra<sup>TM</sup> Kit, GenHunter), the total RNAs  $(2 \mu g)$  were reverse-transcribed with MMLV reverse transcriptase. 3'-Anchored oligo-T11A (FH-T11A) primer and arbitrary primers (HAP9-HAP21) were used. For each PCR, 2-µL first-strand cDNA was added into a mixture containing 1 unit of Taq polymerase (TaKaRa, Dalian), 50 µmol/L of each dNTP, 10 pmol of arbitrary primers,10 pmol of fluorescein iso-thiocyanate labeled 3'-anchored oligo-T11A (FH-T11A) primer, and 2.0  $\mu$ L of 10 $\times$  PCR buffer. FDD-PCR was performed using the following thermal cycling conditions: 94 °C for 3 min, then 40 cycles of 94 °C for 30 s, 40 °C for 2 min and 72 °C for 2 min, followed by a final extension at 72 °C for 5 min. Each PCR product was electrophoresed in a 6% denaturing polyacrylamide gel in the 1× TBE buffer. The FDD gel was scanned with the FMBIO II (Hitachii Genetic System). The differential bands of interest were cut and extracted to recover cDNAs. Re-amplified cDNA products were cloned into pM18-T vectors (Takara) and then transformed into *E. coli* strain DH5α followed by clone selection based  $\alpha$ -complementation. The plasmids were purified using the MiniBEST Plasmid Purification Kit (TaKaRa, Dalian) and were sequenced using CEQ8000 (Beckman).

### 2.6. Molecular cloning

Blast searches in NCBI were used to show that it had high similarity to ESTS (GenBank no. AV404609) of *B. mori*. Using *in silico* cloning method, its cDNA sequence was obtained. According to the sequence, specific primers were designed as follows: cross-primer pair 1 (sense primer 5'-GTCGACTATCAAGTAATATGG ACC-3' and anti-sense primer 5'-CTTCA GTCGGAATAAGAGTA-3') and cross-primer pair 2 (sense primer 5'-TGGCAATGGTTTCGTCAA-3' and anti-sense primer 5'-CTCGAGTTAA ATTGGTTCGCAAT AG-3'). Both primer pairs were used to re-amplify the target sequence.

### 2.7. Bioinformatic analysis

A homology search of the DNA sequences against Gen-Bank (http://www.ncbi.nlm.nih.gov/) was performed. The cDNA sequence was compared with the silkworm genomic



Fig. 4. Re-amplification of differential fragment A9 and verification of its existence in *Bmcyp306a1*. (a) M, DL2000; (b) M, Lambda DNA/EcoRI+HindIII. Band in lane 1 was about 920 bp. That in lane 2 was about 1050 bp. The two fragments were used to verify the existence of the A9-1617 (*Bmcyp306a1*) according to the two primer pairs designed.

sequences. SIM4 (http://pbil.univ-lyon1.fr/sim4.php) and Spidey (http://www.ncbi.nlm.nih.gov/IEB/Research/Ostell/Spidey) were used to align the cDNA with the genomic sequence. Translation into amino acid sequence was done with SwissProt database ExPASy Translate tool (http://au.expasy.org/tools/dna.html). Homology searches were performed using Blastx, SIB BLAST Network Service (http://au.expasy.org/tools/blast/) and DNAstar 5.0 software. The secondary structure prediction was carried out using the DSC secondary structure prediction method in PBIL (http://pbil.univ-lyon1.fr/).

### 2.8. Expression analysis of target gene

To compare the expression levels of target gene, semiquantitative PCR analysis was performed. The primers were designed according to the sequences of *B. mori cyp306a1*. The sense primer was 5'-TGGCAATGGTTTCGTCAA-3' and the anti-sense primer was 5'-CTTCAGTCGGAATAAGAGTA-3'. Silkworm housekeeping gene *actin A3* was amplified as control with the same templates. The upstream/downstream primers of *actin A3* (GenBank no. U49854) were 5'-GGATGTCCACGTCGCAC-3' and 5'-GCGCGGCTACTCGTTCACTACC-3', respectively. The thermal cycling profile consisted of initial denaturation at 95 °C for 3 min and 40 cycles at 95 °C for 30 s, at 52 °C for 30s, and at 72 °C for 45 s. The volume of PCR reaction mixture was 10 µL, containing 1.0 µL of buffer (Mg<sup>2+</sup>)(10×), 1 µL of each primer (1 µmol/L), 0.1 µL enzyme (TaKaRa) and 2.5 µL of 1:10 diluted cDNA templates.

To compare the expression levels of *Bmcyp306a1* gene between the larvae of 441F and 441DZ, real-time quantitative PCR was carried out using the Mx3000PTM PCR instrument (Stratagene). The thermal cycling profile was the same to that of semi-quantitative PCR, using the same primers and the PCR mixture (10  $\mu$ L) containing 1× SYBR GREEN PCR mix (Takara, Dalian), 0.1- $\mu$ L ROX reference Dye II (50×), 1  $\mu$ L of each primer (1  $\mu$ mol/L), and 2.5  $\mu$ L of 1:10 diluted cDNA templates, according to the manufacturer's instructions. The specificity of PCR amplification was determined by constructing a melting curve after the PCR amplification. Melting curve analysis was performed in the range of 55–95 °C by monitoring SYBR Green fluorescence with 0.5 °C increment. Each sample was conducted in triplicate. The relative expression ratios of gene were calculated relative to the housekeeping gene using the comparative Ct (2<sup>- $\Delta$ Ct</sup>) method [13].

### 2.9. Tissue distribution of target gene

To determine the tissue distribution of *Bmcyp306a1* transcripts, the tissues of midgut, fat body, ovary, hemocyte and silk gland were analyzed by semi-quantitative PCR, primed with the sense primer 5'-TGGCAATGGTTTCGTCAA-3' and the anti-sense primer 5'-CTTCAGTCGGAATAAGAGTA-3'. The thermal cycling profile and the PCR reaction system were the same as in Section 2.8.

### **Table 1**Body weight change at different time (unit: g)

Freatment	0 h	24 h	48 h	72 h	96 h	120 h	144 h
140DZ	41.74	68.59	95.77	125.63	154.45	169.44	176.04
140F	40.25	58.15	80.13	101.22	114.24	118.56	114.51
441DZ	34.67	63.69	86.71	111.56	132.96	144.21	145.61
441F	35.49	59.97	84.83	114.12	135.08	145.35	147.28

Each group was composed of 50 silkworms which were of similar size. Total body weight change was investigated. This experiment was conducted under the same condition except presence or absence of fluoride. Groups 440DZ and 441DZ were control. Groups 440F and 441F were fluoride-treated.

### 3. Results

### 3.1. Pathological changes

Growth data of the larvae were given in Table 1. At 24 h, growth index of 440F larvae was 0.445, the lowest among all treatments (Fig. 1). From 24 to 96 h, growth indexes of all treatments increased steadily. Yet, the increasing rate of 440F was lower than those of 440DZ, 441DZ and 441F. At 96 h, growth indexes of 440DZ, 440DZ, 440F, 441DZ, and 441F were 3.06, 1.95, 3.16, and 3.10, respectively. Again, the growth index of 440F was the lowest among all treatments (Fig. 1).

Mortality check revealed that no silkworm died in the control groups (440DZ and 441DZ). In groups 440F and 441F, first death of the larvae was seen at 96 and 120 h, respectively (Fig. 2). In group 440F, the larvae grew much unevenly and their movement was largely reduced. Most fluorosised larvae exhibited coarse and black-brown sickness spots on their abdomen which were inclined to be ulcerated. Among the 50 larvae observed, 2, 3, 6, and 7 larvae showed serious intoxication symptoms, spitted much fluid and died before 96, 120, 144, and 168 h, respectively. In group 441F, the larvae grew much more evenly than those of 440F. Most of the larvae did not show pathological changes seen on those of group 440F, though 2, 1, and 2 larvae died before 120, 144, and 168 h, respectively. Based on the fact that the mortality of 441F larvae (5 out of 50) was considerably lower than that of 440F (18 out of 50) and the first death in 441F was 24 h later that that in 440F (Fig. 2), we were confident that strain 441 possessed certain resistance to fluoride.

### 3.2. FDD analysis

To identify genes associated with the expressions of resistant genes, we compared the mRNA expression patterns in midguts of different silkworm strains using the FDD technique. The fragments were identified that were present in the profiles from resistant silkworm and absent from susceptible silkworm profiles, indicating a response unique to the resistant silkworm strain. Such differentially expressed bands were extracted from the gels and used for DNA clone and sequencing. In this experiment, silkworm was treated

1	ATGGACCTTTATTTTATTTGCCTGGTAACGTTCGTGGCTGGGTTTTGGATTTTCAAAAAA
1	M D L Y F I W L V T F V A G F W I F K K
61	ATAAAGGAATGGCAGAATTTGCCCCCCGGACCTTGGGGGGTTACCTATCGTCGGTTATTTG
21	TKEWQNI. PPGPWGI. PTVGYI.
121	CCTTTCATTGATCGCTATCATCCACATATCACCTTGACAAATTTGTCTAAAACATACGGA
/1	
101	
101	
01	
241	GTAGGAGATACGTTCTCAAAAGACAGTTTTTCTGGACGAGCACCTCTTTACTTAACACAT
81	V G D T F S K D S F S G R A P L Y L T H
301	GGCCTTATGAATGGAAATGGAATTATTTGTGCCGAAGGCGGTTTGTGGAGGGACCAAAGG
101	G L M N G N G I I C A E G G L W R D Q R
361	AAATTAATAACATCGTGGTTGAAAAGTTTTGGAATGAGTAAGCACAGTGTTTCCCGAGAA
121	K L I T S W L K S F G M S K H S V S R E
421	AAATTGGAAAAACGAATCGCTTCAGGAGTATACGAAATTTTGGAAAAATATCGAAAAAACT
141	K L E K R T A S G V Y E T L E N T E K T
481	TOTICATIONAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
161	
541	
101	
601	
001	
201	
001	ATACGCCATGTTCGCCATCAACACGAAAAACAATTGAAGTTCTAGTCCGTGGACAGGCA
221	I R H V S P S T R K T I E V L V R G Q A
721	CAGACGCATACTTTGTACGCAAGCATGATAGATAGACGAAGAAAAATGTTGGGCTTAGAG
241	Q T H T L Y A S M I D R R R K M L G L E
781	AAGCCTAAGGGAGCCGAATATGCTCCTCACGAAAACCTTTTAAAACTATATCCAAATGGC
261	K P K G A F Y A P H F N L L K L Y P N G
841	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCGAACACCGAGCATTTCTTTGATCCC
841 281	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S <b>P H T E H F F D P</b>
841 281 901	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S <b>P N T E H F F D P</b> AATACTCTTATTCCGACTGAAGGAGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG
841 281 901 301	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $\mathbf{P}$ H T E H F F D $\mathbf{P}$ AATACTCTTATTCCGACTGAAGGAGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG $\mathbf{M}$ T L I P T E G D C I L D N F L L E Q K
841 281 901 301 961	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $\mathbf{P}$ H T E H F F D $\mathbf{P}$ AATACTCTTATTCCGACTGAAGGAGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG $\mathbf{M}$ T L I P T E G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGATGAACAGTTACATTTT
841 281 901 301 961 321	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $\mathbf{P}$ H T E H F F D $\mathbf{P}$ AATACTCTTATTCCGACTGAAGGAGAGTTGCATACTGGATAATTTTCTGTTGGAGCAAAAG $\mathbf{M}$ T L I $\mathbf{P}$ T E G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGATGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F
841 281 901 301 961 321 1021	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $\mathbf{P}$ H T E H F F D $\mathbf{P}$ AATACTCTTATTCCGACTGAAGGAGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG $\mathbf{M}$ T L I P T E G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGATGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGATATGTTCGGTGCCGGACTGGATACTACATCGGTGACTTAGCTTGGTTT
841 281 901 301 961 321 1021 341	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $\mathbf{P}$ H T E H F F D $\mathbf{P}$ AATACTCTTATTCCGACTGAAGGAGGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG $\mathbf{M}$ T L I P T E G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGATGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGATATGTTCGGTGCCGGACTGGATACTACATCGGTGACCTTAGCTTGGTTT L L A D M F G A G L D T T S V T L A W F
841 281 901 301 961 321 1021 341 1081	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $\mathbf{P}$ H T E H F F D $\mathbf{P}$ AATACTCTTATTCCGACTGAAGGAGGAGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG $\mathbf{M}$ T L I P T E G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGATGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGATATGTTCGGTGCCGGGACTGGATACTACATCGGTGACTTTAGCTTGGTTT L L A D M F G A G L D T T S V T L A W F
841 281 901 301 961 321 1021 341 1081 361	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $\mathbf{P}$ <b>N T E H F F D</b> $\mathbf{P}$ AATACTCTTATTCCGACTGAAGGAGAGTGCATACTGGATAATTTTCTGTTGGAGCAAAAG $\mathbf{N}$ <b>T L I P T E</b> G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGATGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGATATGTTCGGTGCCGGACTGGATACTACATCGGTGACCTTAGCTTGGTTT L L A D M F G A G L D T T S V T L A W F TTGCTATACATGGCTTTGTTTCCAGAAGAACAGGAAGAAATACGTAAAGAAATCTTATCC L L Y M A L F P F F Q F F L R K F I L S
841 281 901 301 961 321 1021 341 1081 361 1141	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $\mathbf{P}$ <b>N T E H F F D P</b> AATACTCTTATTCCGACTGAAGGAGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG $\mathbf{M}$ <b>T L I P T E</b> G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGATGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGATATGTTCGGTGCCGGGACTGGATACTACATCGGTGACCTTTAGCTTGGTTT L L A D M F G A G L D T T S V T L A W F TTGCTATACATGGCTTTGTTTCCAGAAGAACAGGAAGAAATACGTAAAGAAATCTTATCC L L Y M A L F P E E Q E E I R K E I L S GTATATCCATGGCATGGATGGTGATACTTCAACCTCTTTTTATCCCAGCAGCTACATCGTTGATGTTCCAACTGCTTCAACCTCCACCACCTCTTTTTT
841 281 901 301 961 321 1021 341 1081 361 1141 381	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCGGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $\mathbf{P}$ <b>N T E H F F D P</b> AATACTCTTATTCCGACTGAAGGAGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG $\mathbf{M}$ <b>T L I P T E</b> G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGATGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGATATGTTCGGTGCCGGACTGGATACTACATCGGTGACCTTAGCTGGTTT L L A D M F G A G L D T T S V T L A W F TTGCTATACATGGCTTTGTTTCCAGAAGAACAGGAAGAAATCCTTATCC L L Y M A L F P E E Q E E I R K E I L S GTATATCCATATGACGATGATGTTGATAGTTCAAGGTTACCTCTTTTAGGCAGCAAATC
841 281 901 301 961 321 1021 341 1081 361 1141 381 1201	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $P$ N T E H F F D P AATACTCTTATTCCGACTGAAGGAGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG $\overline{M}$ T L I P T E G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGAGTGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGATATGTTCGGTGCCGGACTGGATACTACATCGGTGACCTTGGTTT L L A D M F G A G L D T T S V T L A W F TTGCTATACATGGCTTTGTTTCCAGAAGAACAGGAAGAAATCCTTATCC L L Y M A L F P E E Q E E I R K E I L S GTATATCCATATGACGATGATGTTGATAGTTCAAGGTTACCTCTTTTAGGCAGCAACTC V Y P Y D D D V D S S R L P L L M A A I
841 281 901 301 961 321 1021 341 1081 361 1141 381 1201 401	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $\mathbf{P}$ H T E H F F D $\mathbf{P}$ AATACTCTTATTCCGACTGAAGGAGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG $\mathbf{M}$ T $\mathbf{L}$ I $\mathbf{P}$ T E G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGAGTGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGATATGTTCGGTGCCGGACTGGATACTACATCGGTGACTTTAGCTTGGTTT L L A D M F G A G L D T T S V T L A W F TTGCTATACATGGCTTTGTTTCCAGAAGAACAGGAAGAAATCTTATCC L L Y M A L F P E E Q E E I R K E I L S GTATATCCATATGACGATGATGTTGATAGTTCAAGGTTACCTCTTCTTATGGCAGCAGCAATC V Y P Y D D D V D S S R L P L L M A A I TGTGAAACTCAGAGGATTCGATCGATGTTGTTCCAGTGGGAATACCCCATGGTTGTATAGAGG
841 281 901 301 961 321 1021 341 1081 361 1141 381 1201 401	CATATCAAATGCATAAAATACAGCAAAGTCTCTCTCGGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $\mathbf{P}$ H T E H F F D $\mathbf{P}$ AATACTCTTATTCCGACTGAAGGAGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG $\mathbf{P}$ T L I $\mathbf{P}$ T E G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGAGTGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGATATGTTCGGTGCCGGACTGGATACTACATCGGTGACTTTAGCTTGGTT L L A D M F G A G L D T T S V T L A W F TTGCTATACATGGCTTTGTTTCCAGAAGAACAGGAAGAAATACGTAAAGAAATCTTATCC L L Y M A L F P E E Q E E I R K E I L S GTATATCCATATGACGATGATGTTGATAGTTCAAGGTTACCTCTTCTTATGGCAGCAGATC V Y P Y D D D V D S S R L P L L M A A I TGTGAAACTCAGAGGATTCGATCGATGTTGTTCCAGTGGGGAATACCCCATGGTTGTATAGAGG C E T Q R I R S I V P V G I P H G C I E
841 281 901 301 961 321 1021 341 1081 361 1141 381 1201 401 1261	CATATCAAATGCATAAAATACAGCAAAGTCTCTCTCGGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $\mathbf{P}$ H T E H F F D $\mathbf{P}$ AATACTCTTATTCCGACTGAAGGAGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG $\mathbf{P}$ T L I $\mathbf{P}$ T E G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGATGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGATATGTTCGGTGCCGGACTGGATACTACATCGGTGACCTTAGCTTGGTT L L A D M F G A G L D T T S V T L A W F TTGCTATACATGGCTTTGTTTCCAGAAGAACAGGAAGAAATACGTAAAGAAATCTTATCC L L Y M A L F P E E Q E E I R K E I L S GTATATCCATATGACGATGATGTTGATAGTTCAAGGTTACCTCTTCTTATGGCAGCAGCAATC V Y P Y D D D V D S S R L P L L M A A I TGTGAAACTCAGAGGATTCGATCGATGTTGTTCCAGTGGGGAATACCCCATGGTTGTATAGAGG C E T Q R I R S I V P V G I P H G C I E GACGCTTACTTGGGTACTACAGAATCCCAATACGATCCCCATGGGTGGTGACTGGGGGAAT
841 281 901 301 961 321 1021 341 1081 361 1141 381 1201 401 1261 421	CATATCAAATGCATAAAATACAGCAAAGTCTCTCTCGGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $\mathbf{P}$ <b>H</b> T <b>E H F F D</b> $\mathbf{P}$ AATACTCTTATTCCGACTGAAGGAGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG $\mathbf{H}$ T L I $\mathbf{P}$ T E G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGATGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGATATGTTCGGTGCCGGACTGGATACTACATCGGTGACTTTAGCTTGGTT L L A D M F G A G L D T T S V T L A W F TTGCTATACATGGCTTTGTTCCAGAAGAACAGGAAGAAATACGTAAAGAAATCTTATCC L L Y M A L F P E E Q E E I R K E I L S GTATATCCATATGACGATGATGTTGATAGTTCAAGGTTACCTCTTCTTATGGCAGCAGCAATC V Y P Y D D D V D S S R L P L L M A A I TGTGAAACTCAGAGGATTCGATCGATGTTGTTCCAGTGGGGAATACCCCATGGTTGTATAGAGG C E T Q R I R S I V P V G I P H G C I E GACGCTTACTTGGGTAACTACAGAATCCCAAGAATCCCATGCAGTGG D A Y L G N Y R I P K N A M V I P L Q W
841 281 901 301 961 321 1021 341 1081 361 1141 381 1201 401 1261 421 1321	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $\mathbf{P}$ H T E H F F $\mathbf{D}$ P AATACTCTTATTCCGACTGAAGGAGAGTGCATACTGGATAATTTTCTGTTGGAGCAAAAG $\mathbf{M}$ T L I P T E G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGATGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGATATGTTCGGTGCCGGACTGGATACTACATCGGTGACTTTAGCTTGGTT L L A D M F G A G L D T T S V T L A W F TTGCTATACATGGCTTTGTTCCAGAAGAACAGGAAGAAATACGTAAAGAAATCTTATCC L L Y M A L F P E E Q E E I R K E I L S GTATATCCATATGACGATGATGTTGATAGTTCAAGGTTACCTCTTCTTATGGCAGCAAATC V Y P Y D D D V D S S R L P L L M A A I TGTGAAACTCCAGAGGATTCGATCGATGTTGTTCCAGTGGGAATACCCCATGGTTGTATAGAG C E T Q R I R S I V P V G I P H G C I E GACGCTTACTTGGGTAACTACAGAATCCCATGGTGATCCATTGCAGTGG D A Y L G N Y R I P K N A M V I P L Q W GCTATTCCATGGGATCCTAATGTTTGGGAAGAACCAGGAAAAATTCAAACCGCGAGTTT
841 281 901 301 961 321 1021 341 1081 361 1141 381 1201 401 1261 421 1321 441	CATATCAAATGCATAAAATACAGCAAAGTCTCTCTCGGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $\mathbf{P}$ <b>H</b> T <b>E H F F D P</b> AATACTCTTATTCCGACTGAAGGAGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG $\mathbf{M}$ T L I P T E G D C I L D N F L L E Q K AAAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGAGTGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGATATGTTCGGTGCCGGACTGGATACTACATCGGTGACCTTAGCTTGGTTT L L A D M F G A G L D T T S V T L A W F TTGCTATACATGGCTTTGTTTCCAGAAGAACAGGAAGAAATACGTAAAGAAATCTTATCC L L Y M A L F P E E Q E E I R K E I L S GTATATCCATATGACGATGATGTTGATAGTTCAAGGTTACCTCTTCTTATGGCAGCAAATC V Y P Y D D D V D S S R L P L L M A A I TGTGAAACTCAGAGGATTCGATCGATGTTGTTCCAGTGGGAATACCCCATGGTTGTATAGAG C E T Q R I R S I V P V G I P H G C I E GACGCTTACTTGGGTAACTACAGGATCCCAAGGAACAAGGAAATCCCATGGGTGG D A Y L G N Y R I P K N A M V I P L Q W GCTATTCACATGGATCTAATGTTGGGAAGAACCAGGAAAAATTCAAAACCGCGAGTATT A I H M D P N V W E <b>E P E K F K P R F</b>
841 281 901 301 961 321 1021 341 1081 361 1141 381 1201 401 1261 421 1321 441 1381	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $P$ M T E H F F D P AATACTCTTATTCCGACTGAAGGAGGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG M T L I P T E G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGAGA
841 281 901 301 961 321 1021 341 1081 361 1141 381 1201 401 1261 421 1321 441 1381 461	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $P$ M T E H F F D P AATACTCTTATTCCGACTGAAGGAGGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG M T L P T E G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGAGTGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGATATGTTCGGTGCCGGACTGGATACTACATCGGTGACTTTAGCTTGGTTT L L A D M F G A G L D T T S V T L A W F TTGCTATACATGGCTTTGTTTCCAGAAGAACAGGAAGAAATCCTTAGCTTGGTTT L L Y M A L F P E E Q E E I R K E I L S GTATATCCATATGACGATGATGATGATGATGATCAGGTTACCTCTTCTTATGGCAGCAATC V Y P Y D D D V D S S R L P L L M A A I TGTGAAACTCAGAGGATCGATCGATGGTTGTTCCAGGTGGGAATACCCCATGGTTGTATAGAG C E T Q R I R S I V P V G I P H G C I E GACGCTTACTTGGGTAACTACCAGGATCCAAAGAAATCCCCATGGTGGGA D A Y L G N Y R I P K N A M V I P L Q W GCTATTCACATGGATCCTAATGTTGGGAAGAACCAGGAAAAATTCAAACGGCAGCATTT A I H M D P N V W E E P E K F K P R R F TTGGCTCAGGATGGTGGTGATCTACTTAGCTCAAGGATCCATCGTTGTATCAAACTGGTAGTCTACTTAGGCGAAGAAATTCAAACTGGTAACTGCAGGATGTTGATAGTTGGAAGAAATTCCAAACTGGCGGAATTC A I H M D P N V W E F P E K F K P R R F
841 281 901 301 961 321 1021 341 1081 361 1141 381 1201 401 1261 421 1321 441 1381 461 1441	CATATCAAATGCATAAAATACAGCAAAGTCTCTCTCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $P$ H T E H F F D $P$ AATACTCTTATTCCGACTGAAGGAGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG H T L I P T E G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGAGA
841 281 901 301 961 321 1021 341 1081 361 1141 381 1201 401 1261 421 1321 441 1381 461 1441 481	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $P$ M T E H F F D P AATACTCTTATTCCGACTGAAGGAGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG M T L I P T E G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGATGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGATATGTTCGGTGCCGGACTGGATACTACATCGGTGACTTTAGCTTGGTT L L A D M F G A G L D T T S V T L A W F TTGCTATACATGGCTTTGTTTCCAGAAGAACAGGAAGAAATACGTAAAGGAAATCTTATCC L L Y M A L F P E E Q E E I R K E I L S GTATATCCATATGACGATGATGTTGATAGTTCAAGGTTACCTCTTCTTATGGCAGCAATC V Y P Y D D D V D S S R L P L L M A A I TGTGAAACTCAGAGGATTCGATCGATGTTGTTCCAGTGGGGAATACCCCATGGTTGTATAGAGG C E T Q R I R S I V P V G I P H G C I E GACGCTTACTTGGGTAACTACAGGAATCCCAAAAAATGCCAAGGAAATCCCATTGGAGGAT CAAACTCAGGGATCCTAATGTTTGGGAAGAACCAGGAAAAATCCCATGGTGGTGATATGGAG D A Y L G N Y R I P K N A M V I P L Q W GCTATTCCACATGGATCCTAATGTTTGGGAAGAACCAGGAAAAATTCCAAACCGGCGTAGATTT A I H M D P N V W E E P E K F K P R F TTGGCTCAGGATGGTGTGTACTACTTAAGGCCCAAGAAAATTCCATTCCGTTTCAAACTGGTAGTCCTAATGAGAACCAGGAAATCCCAAAAAATTCCAAACTGGTAGATTT A I H M D P N V W E F P E K F K P R F TTGGCTCAGGATGGTGGTGATGCTCATTAGGCCCAAGAAATTCCATTCCGTTTCAAACTGGTAGGTA
841 281 901 301 961 321 1021 341 1081 361 1141 381 1201 401 1261 421 1321 441 1381 461 1441 481 1501	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $\mathbf{P}$ H T E H F F D $\mathbf{P}$ AATACTCTTATTCCGACTGAAGGAGATGCATACTGGATAATTTTCTGTTGGAGCAAAAG $\mathbf{M}$ T L I P T E G D C I L D N F L L E Q K AAACGATTTGAGAGGGGGACCCAACGGCACTGTATATGAGAGAGA
841 281 901 301 961 321 1021 341 1081 361 1141 381 1201 401 1261 421 1321 441 1381 461 1441 481 1501 501	CATATCAAATGCATAAATACAGCAAAGTCTCTCCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $P$ H T E H F F D $P$ AATACTCTTATTCCGACTGAAGGAGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG M T L I P T E G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGAGTGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGATATGTTCGGTGCCGGACTGGATACTACACTGGTGACTTTAGCTTGGTTT L L A D M F G A G L D T T S V T L A W F TTGCTATACATGGCTTTGTTTCCAGAGAGAACAGGAAATACGTAAAGAAATCTTATCC L L Y M A L F P E E Q E E I R K E I L S GTATATCCATATGACGATGATGTTGATAGTTCAAGGTTACCTCTTTTAGGCAGCAATC V Y P Y D D D V D S S R L P L L M A A I TGTGAAACTCAGAGGATTCGATCGATGTTGTCCAGTGGGGAATACCCCATGGTTGTATAGAG C E T Q R I R S I V P V G I P H G C I E GACGCTTACTTGGTACATGATCTACAGGATGCCATGGTGGAACTCCATTGCAGTGGTG D A Y L G N Y R I P K N A M V I P L Q W GCTATTCACATGGATCCTAATGTTTGGGAAGAACCAGAAAAATTCAAACCGCGTAGATTT A I H M D P N V W E E P E K F K P R R F TTGGCTCAGGATGGTGTACTTACTTAAGCCTCAAGAATTCATTC
841 281 901 301 961 321 1021 341 1081 361 1141 381 1201 401 1261 421 1321 441 1381 461 1441 481 1501 501 1561	CATATCAAATGCATAAATACAGCAAAGTCTCTCCCGAACACCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $\mathbf{P}$ <b>N T E H F F D</b> $\mathbf{P}$ AATACTCTTATTCCGACTGAAGGAGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG <b>N T L I P T E G D</b> C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGATGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGATATGTTCGGTGCCGGACTGGATACTACACTCGGTGACTTTAGCTTGGTTT L L A D M F G A G L D T T S V T L A W F TTGCTATACATGGCTTTGTTTCCAGAAGAACAGGAAGAAATACGTAAAGAAATCTTATCC L L Y M A L F P E E Q E E I R K E I L S GTATATCCATATGACGATGATGTTGATAGTTCAAGGTTACCTCTTCTTATGGCAGCAATC V Y P Y D D D D V D S S R L P L L M A A I TGTGAAACTCAGAGGATTCGATCGATGTTGTTCCAGGGGAATACCCCATGGTGTTATAGAG C E T Q R I R S I V P V G I P H G C I E GACGCTTACTTGGTACATGATGTTGGAAGAACCCCAAAAAATGCCATGGTGTGTATAGAG C A Y L G N Y R I P K N A M V I P L Q W GCTATTCCACATGGATCCTAATGTTGGAGAGAACCGAAAAATTCAAACCGGCTGAGTTT A I H M D P N V W E <b>E P E K F K P R F</b> TTGGCTCAGGATGGTGTTCATGTTGCCAAGAAATTCCATTCGGTTGTAAAACTGGTAAGT C A Y L G N Y R I P K N A M V I P L Q W GCTATTCACATGGTTGTACTACTCCAAGGAATCCCCAAGAAATTCAAACTGGTAGATTT A I H M D P N V W E <b>E P E K F K P R F</b> TTGGCTCAGGATGGTGGTGCCACTACTGTTGCCGTATGTTGTGTAAACTGGTAAGTAGCAACTGTCCAATGTTTCAAGCGTGTGTCCGGTGGAACTACTGCTGTAGTGGCCTCGTAAGTAGGACTAA R M C P G D D E L S R M L S C G L V S R L TTCAGAAAGCAGCGGTATTCGACTGCCCTCCCGCATCCAAAAAATCCGACAGGAAGAGAGAG
841 281 901 301 961 321 1021 341 1081 361 1141 381 1201 401 1261 421 1321 441 1381 461 1441 481 1501 501 1561 561	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCGAACACCCGAGCATTTCTTGATCCC H I K C I K Y S K V S $P$ H T E H F F D P AATACTCTTATTCCGACTGAAGGAGATTGCATACTGGATAATTTTCTGTTGGAGCAAAAG H T L I P T E G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGAGTGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGGATATGTTCGGTGCCGGACTGGATACTACATCGGTGACCATTAGCTTGGTTT L L A D M F G A G L D T T S V T L A W F TTGCTATACATGGCTTTGTTTCCAGAAGAACAGGAAGAAATACGTAAAGAAATCTTATCC L L Y M A L F P E E Q E E I R K E I L S GTATATCCATATGACGATGATGTTGATAGTTCAAGGTTACCTCTTCTTATGGCAGCAATC V Y P Y D D D D V D S S R L P L L M A A I TGTGAAACTCAGAGGATTCGATCGATGTTCCAGTGGGGAATACCCCATGGTTGTATAGAG C E T Q R I R S I V P V G I P H G C I E GACGCTTACTTGGGTAACTACAGAATCCCAAAAAATCCCATGGTGGTGTGATAGAG C A Y L G N Y R I P K N A M V I P L Q W GCTATTCCACATGGATCCTACTACTGTTGGGAAGAACCGGAAAAATCCAATGCAGTGG D A Y L G N Y R I P K N A M V I P L Q W GCTATTCCACATGGATGCTACTACTACGAGAACAGCAGAAAAATTCAAACCGGTAGATT A I H M D P N V W E E P E K F K P R F F TTGGCTCAGGATGGTGGTACCTACTGTCCCGATGGTGGTGCCCGTTGGATAGAG C G G S L L K P Q E F I P F Q T G K CGGATGTTCCCGGTGGAGGAACTGCCCAAAAAATTCCAAACCGGTAGATTT A I H M D P N V W E E P E K F K P R R F TTGGCTCAGGATGGTAGTGTCACTTACGCCCCAAAAAATTCCAAACCGGCTGAAGAAGAACTCAAGAGAAGCAGCAGAAAATTCCAAACTGGTAGGAACTAC R M C P G D D E L S R M L S C G L V S R L TTCAGAAAGCAGCGTATTCGACTGCGCATCGATCGATGGTGGCACCGACGAGAAGAAGCAGCAGCAGAAAATCCGACAGAAGAAGCGGCGTGGAA R M C P G D E L S R M L S C G L V S R L TTCAGAAAGCAGCGTATTCGACTGCGCATCAAAAATACCGACAGCAGCAGAAGAGAGAG
841 281 901 301 961 321 1021 341 1081 361 1141 381 1201 401 1261 421 1321 441 1381 461 1441 481 1501 501 1561 561 1621	CATATCAAATGCATAAAATACAGCAAAGTCTCTCCGAACACCCGAGCATTTCTTTGATCCC H I K C I K Y S K V S $P$ H T E H F F D P AATACTCTTATTCCGACTGAAGGAGATGCATACTGGATAATTTTCTGTTGGAGCAAAAG H T L I P T E G D C I L D N F L L E Q K AAACGATTTGAGAGTGGAGACCCAACGGCACTGTATATGAGAGAATGAACAGTTACATTTT K R F E S G D P T A L Y M R D E Q L H F CTACTAGCGGATATGTTCGGTGCCGGACTGGATACTACATCGGTGACTTTAGCTTGGTT L L A D M F G A G L D T T S V T L A W F TTGCTATACATGGCTTTGTTCCAGAAGAACAGGAAGAAATACGTAAAGAAATCTTATCC L L Y M A L F P E E Q E E I R K E I L S GTATTCCATATGACGATGGATGTTGATAGTTCAAGGTTACCTCTTCTTATGGCAGCAATC V Y P Y D D D V D S S R L P L L M A A I TGTGAAACTCAGAGGATTCGATGTTGTGTCCAGTGGGAATACCCCATGGTGTATAGAG C E T Q R I R S I V P V G I P H G C I E GACGCTTACTTGGGTAACTACAGGATACCCCATGGTGATCCCATTGCAGTGG D A Y L G N Y R I P K N A M V I P L Q W GCTATTCACATGGATGCTACTACTGTTGTCGCAGAAAAATTCAAACCGCGTAGATTT A I H M D P N V W E E P E K F K P R F TTGGCTCAGGATGGTGGTGACTCCCATGGCGAAACCAGAAAATTCAAACCGCGTAGATT A I H M D P N V W E E P E K F K P R F TTGGCTCAGGATGGTGGTGACCTCCCATGGTGGACCCCATGGTGAGACTA R M C P G D E L S R M L S C G L V S R L TTCAGAAAGCAGCGGTATTCGACTGCCCATCGATGTGTGTCGGCAGAGAGAG

Fig. 5. Nucleotide sequence and amino acid sequence of the *Bmcyp306a1* gene. This gene contained an ORF of 1617-bp long and encoded 538 amino acids. The start and the stop codons were framed. Predicted polyA signal was underlined. Open boxes were important function sites (domain linker).

H. Zhou et al. / Journal of Hazardous Materials 160 (2008) 330-336



Fig. 6. Genomic structure of *Bmcyp306a1*. There are seven exons and six introns in this gene, and the 2nd and 7th exons are relatively short (146 and 147 bp, respectively), while the 4th exon is the longest (448 bp). The splicing signals are exon/WG-intron-GT/exon (W: T or A).



Fig. 7. Sequence distances of P450 BmCYP306A1 among eight species. The sequences are BAD23844 for BmCYP306A1, ABC96086 for Manduca sexta, XP\_391946 for Apis mellifera, XP\_318345 for Anopheles gambiae, XP\_968477 for Tribolium castaneum, XP\_684136 for Danio rerio, BAB85489 for Homo sapiens, and NP\_573319 for Drosophila melanogaster.

with fluoride for 48 h, band A9 was confirmed from the sample of midguts RNAs (Fig. 3).

### 3.3. Characterization of the Bmcyp306a1

Fragment A9 (Fig. 4a) was obtained and sequenced. Blast searches in NCBI showed that it had high similarity to P450 gene of insect. *In silico* cloning method showed that its full-length cDNA sequence obtained a 1617-bp ORF (Fig. 5). It was *Bmcyp306a1*-like gene. Then two fragments were obtained from the re-amplification by using specific primers. They were about 920 and 1050 bp long, respectively (Fig. 4b). Both fragments contained a 340 bp superposition sequence which was amplified by sense primer of cross-primer pair 2 and anti-sense primer of cross-primer pair 1, both of which confirmed the existence of *Bmcyp306a1*.

### 3.4. Bioinformatic analysis of Bmcyp306a1

The above obtained cDNA sequence was blasted against genomic sequences of *B. mori* in GenBank. It was found that this gene consisted seven exons and six introns. The 2nd and 7th exons are relatively short (146 and 147 bp, respectively), while the 1st and 4th exons are relatively longer (319 and 448 bp, respectively) (Fig. 6). The splicing signals are exon/WG-intron-GT/exon (W: T or A). The cDNA sequence had high similarity to only one contig (Ctg006538), suggesting that the *Bmcyp306a1* gene was a single-copy gene in the genome.

### 3.5. Analysis of amino acid sequence

SwissProt database (http://www.expasy.org/sprot/) was used to find out profiles of the BmCYP306A1 protein. The molecular weight of BmCYP306A1 was 61.58 kDa with 538 amino acids. The number of negatively charged and positively charged residues is 61 and 64, respectively. The theoretical pI is 8.22. The BmCYP306A1 has one *N*-glycosylation site, nine *N*-myristoylation sites, six protein kinase C phosphorylation sites, three casein kinase II phosphorylation sites and one amidation site. The relatively long 1st and 4th exons were found to encode such important function site as domain linker (PNTEHFFDPNTLIPT and EPEKFKPRRFLAQDGS LLKPQEFI PFQT) (indicated as open boxes in Fig. 5).

### 3.6. Phylogenetic analysis of BmCYP306A1

The multi-sequence alignment was performed using DNAStar 5.0 software to identify P450 protein sequence distances between eight species. The result indicated that the similarity of CYP306A1 reached as high as 80.1% between *B. mori* and *M. sexta*, which indicated they were homology protein and belonged to a protein family. The similarity of BmCYP306A1 with *Drosophila melanogaster*'s and *Apis mellifera*'s P450 protein reached 38.4% and 42.0%, respectively. The similarity between *Homo sapiens* and *Danio rerio* reached 40.3%. The result of sequence distances of P450 protein among eight species indicated that two generas of *B. mori* and *M. sexta* were close in genetic relationship (Fig. 7). Phylogenetic tree showed that *B. mori* has high genetic relationship with *M. sexta* but low genetic relationship to *H. sapiens*. *D. melanogaster* was closely related to *Anopheles gambiae* and *Tribolium castaneum*, *H. sapiens* showed high genetic relationship to *D. rerio* (Fig. 8).

#### 3.7. Bmcyp306a1 expression levels in midguts

Semi-quantitative PCR was used to find the expression levels of *Bmcyp306a1*. *Bmcyp306a1* was exclusively expressed in 441DZ and 441F (Fig. 9a). *Bmcyp306a1*'s expression in 441F was lower than in 441DZ, shown by real-time quantitative PCR (Fig. 9b). The mRNA comparative copies of *Bmcyp306a1* in 441DZ and 441F were 0.21 and 0.17, respectively. These indicated that expression of *Bmcyp306a1* was affected in direct or indirect ways under the condition of fluoride intoxication.







**Fig. 9.** Expression levels of *Bmcyp306a1* in the midguts in two strains, which were treated with different methods. (a) Lane 1, 440DZ; lane 2, 440F; lane 3, 441DZ; lane 4, 441F; Strain 440 did not express *Bmcyp306a1* gene. *Bmcyp306a1* was exclusively expressed in strain 441. (b) *Bmcyp306a1* mRNA levels in larvae of 441F and 441DZ. Real-time quantitative PCR data were evaluated using the formula  $2^{-\Delta Ct}$  ( $-\Delta Ct = Ct$  of *actin* – Ct of *Bmcyp306a1*) to calculate the comparative copies. Comparative copies of *Bmcyp306a1* in untreated sample 441DZ vs. treated sample 441F were 0.21 and 0.17, respectively. Each column and bar represented the average of three amplification reactions.

## 3.8. Distribution of Bmcyp306a1 gene expression in different tissues

Semi-quantitative PCR result indicated that the *Bmcyp306a1* gene expressed exclusively in midgut and ovary (Fig. 10). Indeed, the *cyp/actin* optical density ratios were much higher in the midgut and ovary than in other tissues studied (fat body, hemocyte, and silk gland). This result was consistent with the results of pathological changes and FDD result. Differentially expressed *Bmcyp306a1* in different tissues proved that *Bmcyp306a1* was related to fluoride resistance to some extent and had distinctive tissue distribution.

### 4. Discussion

In this study, data of pathological changes including the growth index and the mortality rate were obtained. These pathological changes may be related with fluoride treatment time and the accumulation and storage of fluoride in silkworm (Table 1 and Fig. 1). These results are also consistent with previous researches that fluoride intoxication might affect animals' growth, biochemistry, physiology, and pathology [14,15]. Resistance of silkworm strain 441 is a pre-adaptive phenomenon, so that even before exposure to fluoride, these individuals already have appropriate factors that allow them to survive from fluoride intoxication. A great number of researches have been accomplished in other animals and great progress has been made [16].

This report describes the use of FDD method to screen genes involved in fluoride resistance. Several differentially expressed fragments in fluoride resistant silkworms were obtained. The fragment A9 was analyzed, which had high similarity with P450 *cyp306a1* gene family. Cross-primers designed according with *cyp306a1* were used to prove the existence of A9 gene in midgut.



**Fig. 10.** Distribution of *Bmcyp306a1* in different tissues of the strain 441. Lane 1, midgut; lane 2, fat body; lane 3, ovary; lane 4, hemocyte; lane 5, silkgland. *Bmcyp306a1* is exclusively expressed in midgut and ovary but not in other tissues at 48 h in 5th larva of silkworm. However, the expression level was low.

Then we designated it *Bmcyp306a1*. Using bioinformatic tools, *B. mori* genomes were analyzed to establish *Bmcyp306a1* genomic structure. Only Ctg006538 showed high similarity to the cDNA sequence. It was also established that the *Bmcyp306a1* gene was a single-copy gene in the genome. This may indicate that it was relatively conservative in *B. mori*. Amino acid sequence distance and many studies also showed that P450 was different in many species, which indicated different kinds of P450 had different functions in species' evolving and development. *Bmcyp306a1* belonged to the *cyp306a1* gene family and might be related to fluoride resistance.

The role of P450 in insecticide resistance first appeared in early 1960s when it was reported that the resistance of house flies to carbaryl could be eliminated by the P450 inhibitor sesamex [17]. Since then, the evidence supporting P450 on mediating resistance has increased rapidly. P450 mediated detoxification is a very important resistance mechanism because it not only confers high level of resistance [18], but also provides cross-resistance to unrelated compounds due to the wide variety of substrates that P450 can metabolize [19]. The over-expression of P450 genes in resistant insect strains is a common phenomenon. Cytochrome P450 gene (CYP6a2) was expressed 20-30-fold higher in a malathion-resistant strain (91-R) of D. melanogaster than in a susceptible strain [20]. Lots of other evidence also proved these phenomena [21,22]. FDD-PCR and semi-quantitative PCR indicated that Bmcyp306a1 was exclusively expressed in midgut and ovary but not in other tissues in resistant strain 441. These results shed light on the possibility that *Bmcyp306a1* also had the relationship with fluoride resistance in midgut. Similarly, in early days of the fifth instar, quantitative-PCR indicated that expression level of Bmcyp306a1 was higher in 441DZ than in 441F, which indicated P450 may metabolize fluoride by a single- or multi-step process [19,23] and may be influenced by time and dose of fluoride treatment. P450 was down-regulated under the inducement of fluoride. Down-regulation of Bmcyp306a1's expression was consistent with pathological changes, which also might suggest Bmcyp306a1 has the function of enduring fluoride with the aid of other unknown factors.

Due to P450's pivotal function, tissue-specific expression was analyzed using semi-quantitative PCR. Tissue-specific expression of Bmcyp306a1 proved that Bmcyp306a1 had distinctive tissue distribution. Bmcyp306a1 was mainly distributed in midgut and ovary in strain 441. Previous researches proved that the resistance was inherited and survivors could pass the resistant gene(s) to their offspring [24]. So, the tissue-specific expression indicated that expression of Bmcyp306a1 in ovary was associated with its heredity function. Differential expression of Bmcyp306a1 between silkworm strain 441 treated with fluoride and the control group indicated that Bmcyp306a1 was indeed affected by fluoride intoxication in direct or indirect ways. This result was in accordance with the previous study [25]. In strain 440, Bmcyp306a1 was not obtained, which indicated that Bmcyp306a1 was not only a factor that facilitated highly resistant silkworm strain 441's normal growth, but also had further cooperation with multiple microeffect genes. The former and present studies also indicated that there was obvious dose-effect relationship between the fluoride concentration and the pathological changes. Fluoride also caused DNA damage and apoptosis and some other pathological changes [26]. Based on these, we supposed that *Bmcyp306a1* and other co-operation factors could affect or mediate fluoride metabolism to alleviate intoxication [27]. We also supposed that the expression level of *Bmcyp306a1* had significant difference in temporal expression. To determine the detailed functions of P450 gene and other factors in fluoride resistance, further researches on the functional assay and the catalytic activity are being carried out.

### 5. Conclusion

Fluorosis might affect animals' growth, biochemistry, physiology, and pathology. Pathological changes may be related to fluoride treatment time, the accumulation and storage of fluoride in the silkworm. Fluoride has toxic effects on silkworm body color, behavior, and can affect silkworm's growth and mortality rate, but fluoride's toxic effects were different in different strains. Highly susceptible silkworm strain 440 displayed a stronger pathological changes compared with highly resistant silkworm strain 441. Fluorosis relatively caused the body weight drop and mortality rate elevated in strain 440F. Meanwhile, strain 441 demonstrated certain resistance in delaying the time of first death. Moreover, fluoride might affect gene expression profiles. Expression analysis indicated Bmcyp306a1 was exclusively expressed in 441DZ and 441F and was down-regulated under fluoride treatment. The tissue-specific expression indicated that *Bmcvp306a1* had high-expression level in midgut and ovary in 441F. The data revealed that there was obvious dose-effect and times relationship with the pathological changes and gene expression. Expression profiles of *Bmcyp306a1* suggested that P450 gene was crucial to physiological modification and might be involved in fluoride resistance. All these might elucidate fluoride effect and provide a new method in the fluorosis research.

### Acknowledgements

This work was supported by National Basic Research Program of China (No. 2005CB121000), and National Natural Science Foundation of China (No. 30370773).

### References

- J.X. Li, S.R. Cao, Recent studies on endemic fluorosis in China, Fluoride 27 (1994) 125–128.
- [2] P. Grandjean, K. Juel, O.M. Jensen, Mortality and cancer morbidity after heavy occupational fluoride exposure, Am. J. Epidemiol. 121 (1985) 57–64.
- [3] E. Czerwinski, J. Nowak, D. Dabrowska, A. Skolarczyk, B. Kita, M. Ksiezyk, Bone and joint pathology in fluoride-exposed workers, Arch. Environ. Health 43 (1988) 340–343.
- [4] Z.D. Wei, Endemic food-borne fluorosis in Guizhou, China, Chin. J. Prev. Med. 13 (1979) 148–151.
- [5] R.B. Finkelman, H.E. Belkin, B. Zheng, Health impacts of domestic coal use in China, Proc. Natl. Acad. Sci. 96 (1999) 3427–3431.
- [6] Z.Z. Guan, Y.N. Wang, K.Q. Xiao, D.Y. Dai, Y.H. Chen, J.L. Liu, P. Sindelar, G. Dallenr, Influence of chronic fluorosis on membrane lipids in rat brain, Neurotoxicol. Teratol. 20 (1998) 537–542.
- [7] T. Dote, K. Kono, K. Usuda, H. Nishiura, T. Tagawa, K. Miyata, M. Shimahara, N. Hashiguchi, J. Senda, Y. Tanaka, Toxicokinetics of intravenous fluoride in rats with kidney damage caused by high-dose fluoride exposure, Int. Arch. Occup. Environ. Health 73 (2000) 90–92.

- [8] A.G. Wang, T. Xia, R.A. Ru, X.M. Chen, K.D. Yang, Antagonistic effects of selenium on oxidative stress and apoptosis induced by fluoride in human hepatocytes, Fluoride 37 (2004) 107–116.
- [9] M. Seliskar, D. Rozman, Mammalian cytochromes P450–importance of tissue specificity, Biochim. Biophys. Acta 1770 (2007) 458–466.
- [10] D.R. Nelson, D.C. Zeldin, S.M. Hoffman, L.J. Maltais, H.M. Wain, D.W. Nebert, Comparison of cytochrome P450 (CYP) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudo-genes and alternative-splice variants, Pharmacogenetics 14 (2004) 1–18.
- [11] K. Okuda, Liver mitochondrial P450 involved in cholesterol catabolism and vitamin D activation, J. Lipid Res. 35 (1994) 361–372.
- [12] C.Q. Lin, Q. Yao, D.X. Wu, Investigation and analysis on endurance fluoride of silkworm strain resources, Acta Sericolog. Sin. 22 (1996) 253–255.
- [13] M.L. Wong, J.F. Medrano, Real-time PCR for mRNA quantitation, Biotechniques 39 (2005) 75–85.
- [14] T.W. Burnell, E.R. Peo, A.J. Lewis, et al., Effect of dietary fluorine on growth, blood and bone characteristics of growing-finishing pigs, J. Anim. Sci. 63 (1986) 2053–2067.
- [15] A.M. Abdelhamid, T.M. Dorra, Effect of feedborne fluoride intoxication on broiler chicks' performance, biochemistry, physiology and pathology, Arch. Anim. Nutr. 42 (1993) 133–145.
- [16] A.W.A. Brown, P. Pal, The nature and characterization of resistance, in: A.W.A. Brown, R. Pal (Eds.), Insecticide Resistance in Arthropods, WHO, Geneva, 1971, pp. 9–52.
- [17] Eldefrawi, Methylenedioxyphenyl derivatives as synergists for carbamate insecticides in susceptible, DDT-and parathion-resistant house flies, J. Econ. Entomol. 53 (1960) 231–234.
- [18] J.G. Scott, G.P. Georghiou, Mechanisms responsible for high level of permethrin resistance in the house for fly, Pestic. Sci. 17 (1986) 195–206.
- [19] J.G. Scott, Insecticide resistance in insects, in: D. Pimental (Ed.), Handbook of Pest Management in Agriculture, CRC Press, Boca Raton, 1991, pp. 663–677.
- [20] L.C. Waters, A.C. Zelhof, B.J. Shaw, Possible involvement of the long terminal repeat of transposable element 17.6 in regulating expression of an insecticide resistance-associated P450 gene in *Drosophila*, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 4855–4859.
- [21] B. Pittendrigh, K. Aronstein, E. Zinkovsky, O. Andreev, B. Campbell, J. Daly, S. Trowell, R.H. French-Constant, Cytochrome P450 genes from *Helicoverpa* armigera: expression in a pyrethroid-susceptible and -resistant strain, Insect Biochem. Mol. Biol. 27 (1997) 507–512.
- [22] D. Nikou, H. Ranson, J. Hemingway, An adult-specific CYP6 P450 gene is over expressed in a pyrethroid-resistant strain of the malaria vector, *Anopheles gambiae*, Gene 318 (2003) 91–102.
- [23] K. Natsuhara, K. Shimada, T. Tanaka, T. Miyata, Phenobarbital induction of permethrin detoxification and Phenobarbital metabolism in susceptible and resistant strains of the beet armyworm *Spodoptera exigua* (Huebner), Pestic. Biochem. Physiol. 79 (2004) 33–41.
- [24] World Health Organization, Expert Committee on Insecticides, WHO Technical Report Series, 1957.
- [25] L.C. Waters, A.C. Zelhof, B.J. Shaw, Possible involvement of the long terminal repeat of transposable element 17.6 in regulating expression of an insecticide resistance-associated P450 gene in *Drosophila*, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 4855–4859.
- [26] Y.M. Shivarajashankara, A.R. Shivashankara, P.G. Bhat, Effect of fluoride intoxication on lipid peroxidation and antitoxidant systems in rats, Fluoride 34 (2001) 108–113.
- [27] J.M. Baron, D. Holler, R. Schiffer, S. Frankenberg, M. Neis, H.F. Merk, F.K. Jugert, Expression of multiple cytochrome p450 enzymes and multidrug resistanceassociated transport proteins in human skin keratinocytes, J. Invest. Dermatol. 116 (2001) 541–548.