

## Cloning, Expression and Sequence Analysis of A Luciferase Gene from the Chinese Firefly *Pyrocoelia pygidialis*

DONG Ping-xuan<sup>1, 2</sup>, HOU Qing-bai<sup>1</sup>, LI Xue-yan<sup>1</sup>, LIANG Xing-cai<sup>1,\*</sup>

(1. Kunming Institute of Zoology, the Chinese Academy of Sciences, Kunming Yunnan 650223, China;

2. Graduate University of the Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** The cDNA encoding the luciferase from lantern mRNA of one diurnal firefly *Pyrocoelia pygidialis* Pic, 1926 has been cloned, sequenced and functionally expressed. The cDNA sequence of *P. pygidialis* luciferase is 1647 base pairs in length, coding a protein of 548 amino acid residues. Sequence analysis of the deduced amino acid sequence showed that this luciferase had 97.8% resemblance to luciferases from the fireflies *Lampyrus noctiluca*, *Lampyrus turkestanicus* and *Nyctophila* cf. *caucasica*. Phylogenetic analysis using deduced amino acid sequence showed that *P. pygidialis* located at the base of *Lampyrus*+*Nyctophila* clade with robust support (BP=97%); but did not show a monophyletic relationship with its congeneric species *P. pectoralis*, *P. rufa* and *P. miyako*, all three are strong luminous and nocturnal species. The expression worked in recombinant *Escherichia coli*. Expression product had a 70kDa band and emitted yellow-green luminescence in the presence of luciferin. Five loops in the *P. pygidialis* luciferase, L1 (N198-G208), L2 (T240-G247), L3 (G317-K322), L4 (L343-I350) and L5 (G522-D532), were found from the structure modeling analysis in the cleft, where it was considered the active site for the substrate compound entering and binding. Different amino acid residues between the luciferases of *P. pygidialis* and the three other known strong luminous species can not explain the situation of weak or strong luminescence. Future study of these loops, residues or crystal structure analysis may be helpful in understanding the real differences between the luciferases between diurnal and nocturnal species.

**Keywords:** *Pyrocoelia*; Diurnal firefly; *Pyrocoelia pygidialis*; Luciferase; Homology modeling

## 中国萤火虫云南窗萤荧光素酶 cDNA 的克隆表达和序列分析

董平轩<sup>1, 2</sup>, 侯清柏<sup>1</sup>, 李学燕<sup>1</sup>, 梁醒财<sup>1,\*</sup>

(1. 中国科学院昆明动物研究所, 昆明 650223; 2. 中国科学院研究生院, 北京 100049)

**摘要:** 从一种来自中国日行性萤火虫(云南窗萤)发光器官 mRNA 中克隆、测序并表达了有功能的荧光素酶。云南窗萤荧光素酶的 cDNA 序列有 1 647 个碱基, 编码 548 个氨基酸残基。从推测得到的氨基酸序列的比对分析得出: 云南窗萤的荧光素酶与来自 *Lampyrus noctiluca*, *L. turkestanicus* 和 *Nyctophila* cf. *caucasica* 三种萤火虫的荧光素酶有 97.8% 的序列一致性。从推测得出的氨基酸序列进行系统发育分析, 其结果表明: 云南窗萤和 *Lampyrus*+*Nyctophila* 聚在一起, 与同属的发光强夜行性的萤火虫不形成的单系。云南窗萤荧光素酶在大肠杆菌中表达的条带大约 70 kDa, 并且在有荧光素存在时发出黄绿色荧光。对荧光素酶的结构模拟和分析表明, 云南窗萤荧光素酶基因的氨基端和羧基端结构域之间的裂沟处存在这 5 个多肽环, 这正是从其他荧光素酶推测得到的催化荧光反应时的底物结合位点。云南窗萤和窗萤属的其他 3 种萤火虫的荧光素酶相比, 有 13 个不同氨基酸位点, 位于模拟分子结构的表面。对于这些多肽环、不同氨基酸残基和晶体结构的进一步研究有利于解释日行和夜行性萤火虫荧光素酶的差异。

**关键词:** 窗萤属; 日行性萤火虫; 云南窗萤; 荧光素酶; 同源建模

中图分类号: Q969.48+6.2; Q951.3; Q754 文献标识码: A 文章编号: 0254-5853-(2008)05-0477-08

Received date: 2008-04-17; Accepted date: 2008-06-05

Fundation items: Supported partly by the Natural Foundation of Sciences of Yunnan Province (2006C0046Q); Partly by the Chinese Academy of Sciences (O70651141)

\*Corresponding author (通讯作者), Tel: 86-871-5190441 (Fax), E-mail: liangxc@mail.kiz.ac.cn

收稿日期: 2008-04-17; 接受日期: 2008-06-05

Luciferases are the enzymes that catalyze the reactions yielding visible light. A great number of bioluminescence reactions depend on various luciferases and widely different substrates (Alipour et al, 2004). Firefly luciferase (EC1.13.12.7) is responsible for the oxidation of firefly D-luciferin with molecular oxygen in the presence of ATP and  $Mg^{2+}$  to generate bioluminescence (McElroy, 1969; White et al, 1971; DeLuca, 1976; Wood, 1995; Emamzadeh et al, 2006). As the enzymatic luminescence assay is highly sensitive, rapid, nonradioactive and quantifiable (Gould et al, 1988), the firefly's bioluminescence has been widely applied in fields of life, environment and medical analysis techniques.

Since the cDNA and genomic DNA of the North American firefly *Photinus pyralis* luciferase were reported (De Wet et al, 1985, 1987), the luciferase cDNA and/or genomic DNA of more than 20 firefly species have been cloned, sequenced and/or expressed (Masuda et al, 1989; Tatsumi et al, 1992; Devine et al, 1993; Ohmiya et al, 1995; Sala-Newby et al, 1996; Ye et al, 1997; Viviani et al, 1999; Lee et al, 2001; Choi et al, 2002; Viviani et al, 2004; Alipour et al, 2004; Branchini et al, 2006; Emamzadeh et al, 2006; Li et al, 2006a). The crystal structures of *P. pyralis* and *Luciola cruciata* luciferases were discovered with X-ray diffraction (Conti et al, 1996; Nakatsu et al, 2006). And the substrate binding site and color determination were explained by structure and mutant analysis (Ohmiya et al, 1996; Viviani & Ohmiya, 2000; Viviani et al, 2001; Ugarova & Brovko, 2002; Nakatsu et al, 2006), though the AMP binding site and the structure-function relationship mechanism were still not very clear (Hirokawa et al, 2002; Emamzadeh et al, 2006).

Although fireflies are a well-known luminous beetle, not all species glow in their development from egg to adult. In the case of non-luminous or weak luminous fireflies, it is usually the male adult that has no luminous organs or only has vestiges of luminous organs. Such fireflies are sometimes thought of as diurnal species; analogically, species with both luminous male adults and luminous female adults are called nocturnal fireflies. Though male adults of all species in some firefly genera such as *Pristolycus* (Jeng et al, 2002) and *Cyphonocerus* (Jeng et al, 2006) are non-luminous, other genera, such as *Pyrocoelia*, include not only diurnal but also nocturnal species (Suzuki, 1997). The function of photic signals, undoubtedly, is believed to be for mating in firefly adults

(Branham & Wenzel, 2000, 2003). However, it has been reported that the diurnal fireflies, with weak luminescence or without luminescence, mainly use chemical signals (pheromone) for pair formation (Suzuki, 1997; Branham & Wenzel, 2000).

So far, *Pyrocoelia* luciferases have been cloned from three species: *P. rufa*; *P. miyako* and *P. pectoralis* (Ohmiya et al, 1995; Lee et al, 2001; Rong et al, 2007), all of which, with developed luminous organs, are nocturnal and glow strongly and continuously (Suzuki, 1997; Ohba, 2004; Wang et al, 2007). The luciferase from diurnal fireflies was studied very few, except for one case from a Brazilian twilight active firefly *Macrolampis* sp. (Viviani et al, 2005), which displays an unusual bimodal spectrum. In order to explore the characteristics of the diurnal firefly luciferase and its relationship with those of congeneric nocturnal species, we cloned the luciferase gene of the diurnal firefly *P. pygidialis*, which is distributed in Yunnan (Li et al, 2008), and described its expression and performed sequence analysis.

## 1 Materials and Methods

### 1.1 Insect samples

Larvae of *P. pygidialis* were collected from the gardens of the Kunming Institute of Botany (N25.0224° /E102.1225° ; 1968m), the Chinese Academy of Science, in the northern suburb of Kunming city, Yunnan Province, China, on October 19, 2006, by LI Xue-yan. Live fireflies were taken back to the laboratory and used in the experiment.

### 1.2 RNA extraction and RT-PCR

The lanterns of two larvae were dissected and pulverized under liquid nitrogen with a mortar and pestle. Total RNA was extracted using the RNAiso reagent (TaKaRa, #D312). The first strand of cDNA was synthesized at 42°C for 60min in the presence of 200 U/ $\mu$ L M-MuLV Reverse transcriptase (TaKaRa, #DRR019A), 20U RNase inhibitor, dNTP mixture (final concentration each at 1 mmol/L), and dT adaptor: 5'-GTTTTCCCAGTCACGACTTTTTTTTTTTTTTTT T-3' as the reverse primer. The specific primer sequences used for selective amplification of *P. pygidialis* luciferase gene were: 5'-ACGCGCTAATATCATTGCA-3' (sense primer), based on the luciferase gene of *Lampyrus noctiluca* (GenBankTM/EBI accession number X89479), and the antisense primer-M13: 5'-GTTTTCCCAGTCACGAC-3'. The RT-PCR amplification of cDNA was

carried out by use of the cDNA first strand and under the following condition: initial denaturation at 95°C for 5min, a 36-cycles amplification (94°C for 40s, 48°C for 30s and 72°C for 2 min) and the final extension was performed at 72°C for 8 min. PCR products were analyzed on 1.0% agarose gel electrophoresis and the expected fragment was purified with the Agarose gel DNA Purification kit (TaKaRa, DV805A). The cDNA product was sequenced by an automatic sequencer ABI sequencer PRISM 3730 (Shanghai Sangon Biological Engineering Technology & Services Co, Shanghai, China).

### 1.3 T-A cloning and expression

The following primer set was used to introduce *Nde*I and *Bam*HI restriction endonuclease sites at the 5' and 3' ends, respectively, in the PCR-amplified cDNA: 5'-TTGCACATATGGAAGATGATCATAA-3' and 5'-AT-AAACGGATCCAATTACAGTTTGTATTTTTC-3' (underlined sequences represent the endonuclease site). The PCR product was ligated with pMD-18-T Simple Vector (TaKaRa, DV805A) and was transformed into the Top 10 competent cells. The PMD-18T-Pyluc vector was digested with *Nde*I and *Bam*HI sequentially and ligated into the corresponding clone sites on the pET-28a plasmid. PET28a-Luc recombinants were transformed into BL21 (DE3) competent cells by heat-shock treatment. Positive colonies were incubated in the fresh Luria-Bertain medium containing Kanamycin (50µg/mL) at 37°C until the optical density of liquid medium at 600 nm (OD600) reached 0.5-0.7 (mid-log phase). The culture was induced with IPTG (isopropyl β-D-thiogalactoside, 0.1mmol/L) and the incubation was continued for an additional 8-16 hours at 25°C. The bacterial cells were precipitated (5000× g, 20min) and the pellet was suspended in lyse buffer PBS (0.14 mol/L NaCl, 26 mmol/L KCl, 10mmol/L Na<sub>2</sub>HPO<sub>4</sub>, 2mmol/L KH<sub>2</sub>PO<sub>4</sub>, pH7.4) and frozen and thawed for 3-4 cycles to disrupt the bacterial cells. The homogenate was centrifuged at 12 000 g for 5 min and the supernatant was stored at -20°C. The cell lysate was centrifuged and the supernatant was analyzed by 12% SDS-PAGE.

### 1.4 Luciferase activity

To verify the luciferase expressed in *E. coli*, the cell extracts were assayed for luciferase activity by checking light generation in the dark. A volume of 20 µL of substrate mixture consisting of 1 mmol/L luciferin, 50 mmol/L Tris buffer, pH7.8, 2mmol/L ATP and 10mmol/L MgSO<sub>4</sub> was added to 50µL luciferase extraction solution

in a quartz cell. A photo of the luminescent wells was taken with a Sony DSC-707 digital camera with 30 s exposure.

### 1.5 Sequence and phylogenetic analysis

With 28 GenBank-registered amino acid luciferase sequences, initial alignment was conducted using Clustal X (Thompson et al, 1997), then confirmation of the alignment was done manually using BioEdit (Version 4.7.8). The phylogenetic tree was constructed using the neighbor-joining method (Saitoh & Nei, 1987) and the software PAUP\* 4.0 beta version (Swofford, 2002); the consistency of the branches was analyzed with a bootstrap value of 1000. The other sequences used in the phylogenetic analysis were obtained from GenBank and the accession numbers are as follows: *P. rufa* (AF328553, AY447202 and AY447203); *P. pectoralis* (EF155570); *P. miyako* (L39928); *Diaphanes pectinealis* (DQ408300); *L. noctiluca* (X89479, AY447204); *L. turkestanicus* (AY742225); *P. pyralis* (M15077); *Cratomorphus distinctus* (AY633557); *Nyctophila cf. caucasica* (DQ072141); *L. cruciata* (M26194); *Luciola lateralis* (X66919); *Luciola mingrelia* (S61961); *Luciola italica* (DQ138966); *Hotaria parvula* (L39929); *H. unimunsana* (AF420006, AF486800); *H. papariensis* (AF486802, AF486803); *H. tsushimana* (AF486801, AF486804); *Lampyroidea maculate* (DQ137139); *Photuris pennsylvanica* (U31240); *Phrixothrix hirtus* (AF139645); *Phrixothrix vivianii* (AF139644); *Rhagophthalmus ohbai* (AB255748).

### 1.6 Structure and homology modeling analysis of *P. pygidialis* luciferase

The structure of the *P. pygidialis* luciferase was modeled with the protein homology modeling SWISS-MODEL server using the crystal structure of *L. cruciata* luciferase (Protein Data Bank code: 2d1sA) as a template (swissmodel.expasy.org) (Guex & Peitsch, 1997; Schwede et al, 2003; Arnold et al, 2006). Analysis and comparison of the structures were carried out using Swiss-PdbViewer ver3.7. The domain structure map for the predicted amino acid sequence of *P. pygidialis* luciferase was performed using ProSite (Gattiker et al, 2002).

## 2 Results

### 2.1 RNA extraction, RT-PCR and sequence analysis

From the light organs of two larvae of the firefly *P. pygidialis*, total RNA was isolated, and its quality and quantity were verified by 1.0% agarose gel

electrophoresis. With the gene-specific primers, based on the luciferase gene of *L. noctiluca*, RT-PCR was performed to amplify *P. pygidialis* luciferase genes from the first strand of cDNA. The molecular sizes of the RT-PCR products (cDNA) were 1.7kb and identical to that expected (Fig. 1).

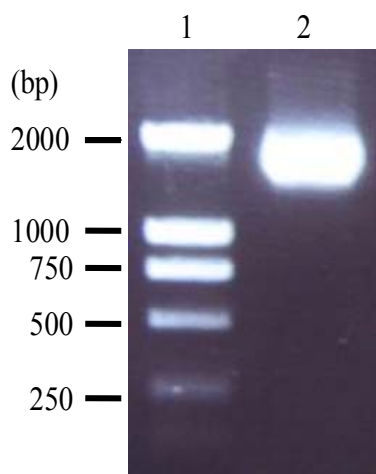


Fig. 1 RT-PCR products of luciferase gene on 1.0% agarose gel

Lane 1, nucleotide acid molecule size marker, DL2000; Lane 2, luciferase gene of *Pyrocoelia pygidialis*.

The nucleotide sequence of PCR products was sequenced and its amino acid sequence was deduced. The result of the complete nucleotide and amino acid sequences is shown in Fig. 2, the 1647bp luciferase gene has an open reading frame of 548 amino acid residues. The nucleotide sequence of *P. pygidialis* has been deposited in GenBank as entry EU826678.

## 2.2 Cloning and expression in *E. coli* BL21 (DE3)

*Nde*I and *Bam*HI restriction sites were appended to the *P. pygidialis* luciferase cDNA prior to being inserted into the pET-28a expression plasmid. Under the control of T7 promoter, luciferase was expressed in BL21(DE3) prokaryotic cells. The positive clones were verified by checking their sequences. The modification of a base pair, without amino acid change, was found. The protein synthesis in BL21 (DE3) cells were analyzed by 12% SDS-PAGE (Fig. 3). The luciferase expressed by the *P. pygidialis* luciferase gene was present as a band of about 70kDa in the cells transformed with recombinant vectors, but not in the cells infected with empty pET-28a plasmid (negative control). The Vector's initial codon and stop codon could have been under the T7 promoter or other reasons may have caused the expressed luciferase

molecular mass difference compared to the calculated value.

By checking luciferase activity of light generation in the dark, the bioluminescence emission of the luciferase from *P. pygidialis* was shown in the photo (Fig. 4), as a yellow-green emission.

## 2.3 Sequence analysis and phylogenetic implication

The deduced amino acid sequences of *P. pygidialis* luciferase gene were compared with those of known luciferase genes. The result of pairwise comparisons was not shown. The deduced amino acid sequence of the *P. pygidialis* luciferase was found to have a high identity value (97.8%) to those of *L. noctiluca*, *L. turkestanicus* and *Nyctophila* cf. *caucasica*. And 60%–70% sequence identity with those of Luciolinae. It had low similarity (48%–53%) with the Phengodidae and Rhagophthalmidae groups.

Phylogenetic analysis using deduced amino acid sequence showed that *P. pygidialis* did not form monophyly with its congeneric species *P. pectoralis*, *P. rufa* and *P. miyako*, but located at the base of *Lampyrus*+*Nyctophila* clade with a high bootstrap (97%) in NJ tree (Fig. 5).

## 2.4 Structure analysis and homology modeling

Using the ProSite (expasy.org/cgi-bin/prosite) online motif analysis, the structure of *P. pygidialis* luciferase was hypothesized. Tab. 1 defines part motifs, their sites and functions in *P. pygidialis* luciferase. *P. pygidialis* luciferase had two special motifs: Amidation site and Serpins signature compared to *L. maculata* luciferase (Emamzadeh et al, 2006). The c-terminal and microbody (peroxisome) targeting tripeptide is SKL, similar with other *Pyrocoelia* species. The putative AMP-binding sequence is 196-IMNSSGSTGLPK-206, corresponding to the *L. cruciata* luciferase P-loop (Nakatsu et al, 2006). The sequence is highly conserved among the various Lampyrinae firefly species (Alipour et al, 2004), small differences occurring in the Rhagophthalmidae and the Phengodidae luciferases. Three dimensional structure of *P. pygidialis* luciferase (K6-L540) and putative loops [L1 (N198-G208), L2 (T2

Tab. 1 Part motifs and sites in *Pyrocoelia pygidialis* luciferase

Amino acid position	Motif information
196 IMNSSGSTGLPK 207	Putative AMP-binding domain signature
458 ILLQHP 468	SERPIN Serpins signature
542 MGKK 545	Amidation site
546 SKL 548	Microbodies C-terminal targeting signal

1	ATG	GAA	GAT	GAT	GAT	AAA	AAT	ATT	ATG	CAC	GGT	CCA	GGG	CCA	TTC	45
1	Met	Glu	Asp	Asp	His	Lys	Asn	Ile	Met	His	Gly	Pro	Ala	Pro	Phe	15
46	TAT	GCT	TTC	GAG	GAT	GGA	AGT	GGT	GGA	GAA	GAA	TTC	GAC	AAA	GGG	90
16	Tyr	Pro	Leu	Glu	Asp	Gly	Thr	Ala	Gly	Glu	Gln	Leu	His	Lys	Ala	30
91	ATG	AAG	AGG	TAT	GGG	CAG	GTT	CCA	GGG	ACA	ATT	GCT	TTT	AGT	GAT	135
31	Met	Lys	Arg	Tyr	Ala	Gln	Val	Pro	Gly	Thr	Ile	Ala	Phe	Thr	Asp	45
136	GGG	CAC	GTA	GAG	GTA	AAT	ATT	ACA	TAT	TCC	GAA	TAT	TTT	GAA	ATG	180
46	Ala	His	Val	Glu	Val	Asn	Ile	Thr	Tyr	Ser	Glu	Tyr	Phe	Glu	Met	60
181	GGT	TGC	CGA	TTG	GGC	GAA	AGT	ATG	AAG	AGG	TAC	GGA	GTT	GGT	TTG	225
61	Ala	Cys	Arg	Leu	Ala	Glu	Thr	Met	Lys	Arg	Tyr	Gly	Leu	Gly	Leu	75
226	GAA	CAC	CAC	ATT	GGT	GTT	TGT	AGT	GAA	AAT	TCI	GTT	CAG	TTT	TTT	270
76	Gln	His	His	Ile	Ala	Val	Cys	Ser	Glu	Asn	Ser	Leu	Gln	Phe	Phe	90
271	ATG	CCT	GTA	TGC	GGT	GCT	CTA	TTT	ATT	GGG	GTT	GGA	GTT	GCA	CCA	315
91	Met	Pro	Val	Cys	Gly	Ala	Leu	Phe	Ile	Gly	Val	Gly	Val	Ala	Pro	105
316	ACA	AAT	GAT	ATT	TAC	AAT	GAA	GGT	GAA	TTG	TAC	AAG	AGT	TTG	TCC	360
106	Thr	Asn	Asp	Ile	Tyr	Asn	Glu	Arg	Glu	Leu	Tyr	Asn	Ser	Leu	Ser	120
361	ATA	TCA	CAA	GCT	ACA	ATA	GTA	TTC	TGT	TCC	AAA	AGA	GCG	CTG	CAA	405
121	Ile	Ser	Gln	Pro	Thr	Ile	Val	Phe	Cys	Ser	Lys	Arg	Ala	Leu	Gln	135
406	AAA	ATC	CTA	GGG	GTA	CAA	AAG	AAA	TTA	CCT	GTA	ATT	GAG	AAA	ATT	450
136	Lys	Ile	Leu	Gly	Val	Gln	Lys	Lys	Leu	Pro	Val	Ile	Glu	Lys	Ile	150
451	GTT	ATT	CTG	GAT	TCT	CGA	GAG	GAT	TAT	ATG	GGG	AAA	CAA	TCT	ATG	495
151	Val	Ile	Leu	Asp	Ser	Arg	Glu	Asp	Tyr	Met	Gly	Lys	Gln	Ser	Met	165
496	TAC	TGC	TTC	ATT	GAA	TCT	CAI	TIA	CCT	GCA	GGT	TTT	AAT	GAA	TAT	540
166	Tyr	Ser	Phe	Ile	Glu	Ser	His	Leu	Pro	Ala	Gly	Phe	Asn	Glu	Tyr	180
541	GAT	TAC	ATA	CCG	GAT	TCT	TTT	GAC	CGC	GAC	ACA	GCA	ACA	GCA	CTT	585
181	Asp	Tyr	Ile	Pro	Asp	Ser	Phe	Asp	Arg	Asp	Thr	Ala	Thr	Ala	Leu	195
586	ATA	ATG	AAT	TCA	TGG	GGA	TGT	AGT	GGA	TTA	CCG	AAG	GGA	GTT	GAG	630
196	Ile	Met	Asn	Ser	Ser	Gly	Ser	Thr	Gly	Leu	Pro	Lys	Gly	Val	Glu	210
631	CTT	ACT	CAC	AAA	AAT	GTT	TGT	GTT	AGA	TTT	TCT	CAC	TGC	AGA	GAT	675
211	Leu	Thr	His	Lys	Asn	Val	Cys	Val	Arg	Phe	Ser	His	Cys	Arg	Asp	225
676	GCT	GTG	TTT	GGT	AAT	GAA	ATT	ATT	GGG	GAT	AGT	GGG	ATT	TTA	ACA	720
226	Pro	Val	Phe	Gly	Asn	Gln	Ile	Ile	Pro	Asp	Thr	Ala	Ile	Leu	Thr	240
721	GTT	ATA	CCA	TTT	CAT	CAT	GGT	TTT	GGA	ATG	TTT	ACA	ACG	CTA	GGA	765
241	Val	Ile	Pro	Phe	His	His	Gly	Phe	Gly	Met	Phe	Thr	Thr	Leu	Gly	255
766	TAT	TTA	ACG	TGT	GGT	TTT	CGT	ATT	GTG	CTT	ATG	TAT	AGA	TTT	GAA	810
256	Tyr	Leu	Thr	Cys	Gly	Phe	Arg	Ile	Val	Leu	Met	Tyr	Arg	Phe	Glu	270
811	GAG	GAA	TTA	TTT	TTA	GGA	TCA	GTT	GAA	GAT	TAT	AAA	ATT	GAA	AGT	855
271	Glu	Glu	Leu	Phe	Leu	Arg	Ser	Leu	Gln	Asp	Tyr	Lys	Ile	Gln	Ser	285
856	GGG	TTG	CTG	GTA	CCA	ACC	CTA	TTT	TCA	TTC	TTT	GGC	AAA	AGC	ACC	900
286	Ala	Leu	Leu	Val	Pro	Thr	Leu	Phe	Ser	Phe	Phe	Ala	Lys	Ser	Thr	300
901	TTA	GTC	GAC	AAA	TAC	GAT	TTA	TCC	AAC	TTG	CAT	GAA	ATT	GCT	TCT	945
301	Leu	Val	Asp	Lys	Tyr	Asp	Leu	Ser	Asn	Leu	His	Glu	Ile	Ala	Ser	315
946	GGT	GGA	GCT	CCG	CTC	GGG	AAA	GAA	GTT	GGA	GAA	GCT	GTA	GCA	AAA	990
316	Gly	Gly	Ala	Pro	Leu	Ala	Lys	Glu	Val	Gly	Glu	Ala	Val	Ala	Lys	330
991	CGA	TTT	AAG	CTG	CCG	GGG	ATA	GGA	CAA	GGG	TAC	GGA	GTT	ACC	GAA	1035
331	Arg	Phe	Lys	Leu	Pro	Gly	Ile	Arg	Gln	Gly	Tyr	Gly	Leu	Thr	Glu	345
1036	ACT	ACG	TCA	GCT	ATT	ATA	ATT	ACA	CCA	GAA	GGG	GAT	GAT	AAA	CCA	1080
346	Thr	Thr	Ser	Ala	Ile	Ile	Ile	Thr	Pro	Glu	Gly	Asp	Asp	Lys	Pro	360
1081	GGA	GGA	TGT	GGT	AAA	GTT	GTT	GCA	TTC	TTT	AGT	GCC	AAA	ATT	GTT	1125
361	Gly	Ala	Cys	Gly	Lys	Val	Val	Pro	Phe	Phe	Ser	Ala	Lys	Ile	Val	375
1126	GAT	CTG	GAT	ACG	AGC	AAA	AGT	TTG	GGT	GTT	AAT	CAG	AGA	GGG	GAA	1170
376	Asp	Leu	Asp	Thr	Ser	Lys	Thr	Leu	Gly	Val	Asn	Gln	Arg	Gly	Glu	390
1171	TTA	TGT	GTG	AAA	GGT	CCA	ATG	ATA	ATG	AAG	GGT	TAC	GTA	AAC	AAC	1215
391	Leu	Cys	Val	Lys	Gly	Pro	Met	Ile	Met	Lys	Gly	Tyr	Val	Asn	Asn	405
1216	CGA	GAA	GCA	AGA	AGT	GGA	TTG	ATA	GAC	AAA	GAT	GGA	TGG	TTA	CAC	1260
406	Pro	Glu	Ala	Thr	Ser	Ala	Leu	Ile	Asp	Lys	Asp	Gly	Trp	Leu	His	420
1261	TGT	GGT	GAC	ATA	GCT	TAC	TAT	GAC	AAG	GAT	GGT	CAC	TTC	TTC	ATC	1305
421	Ser	Gly	Asp	Ile	Ala	Tyr	Tyr	Asp	Lys	Asp	Gly	His	Phe	Phe	Ile	435
1306	GTG	GAT	CGT	TTG	AAA	TGG	TTA	ATT	AAA	TAC	AAA	GGT	TAT	CAG	GTA	1350
436	Val	Asp	Arg	Leu	Lys	Ser	Leu	Ile	Lys	Tyr	Lys	Gly	Tyr	Gln	Val	450
1351	GGG	GGT	GGC	GAA	TTA	GAA	TGG	ATA	TTG	TTA	GAA	CAT	GGC	TTC	ATA	1395
451	Pro	Pro	Ala	Glu	Leu	Glu	Ser	Ile	Leu	Leu	Gln	His	Pro	Phe	Ile	465
1396	TTT	GAT	GCA	GGC	GTT	GCA	GGA	ATT	CCG	GAC	CCA	GAT	GCC	GGT	GAA	1440
466	Phe	Asp	Ala	Gly	Val	Ala	Gly	Ile	Pro	Asp	Pro	Asp	Ala	Gly	Glu	480
1441	CTA	GCT	GCA	GGC	GTT	GTT	GTC	TTA	GAG	GAA	GGG	AAA	ACG	ATG	ACT	1485
481	Leu	Pro	Ala	Ala	Val	Val	Val	Leu	Glu	Glu	Gly	Lys	Thr	Met	Thr	495
1486	GAA	CAA	GAA	GTG	ATG	GAT	TAT	GTT	GGG	GGA	CAA	GTA	ACT	GCT	TCT	1530
496	Glu	Gln	Glu	Val	Met	Asp	Tyr	Val	Ala	Gly	Gln	Val	Thr	Ala	Ser	510
1531	AAG	CGT	TTA	CGT	GGA	GGA	GTT	AAG	TTT	GTG	GAC	GAA	GTA	CCT	AAA	1575
511	Lys	Arg	Leu	Arg	Gly	Gly	Val	Lys	Phe	Val	Asp	Glu	Val	Pro	Lys	525
1576	GGT	CTA	ACT	GGA	AAG	ATT	GAT	TCA	AGG	AAA	ATC	AGG	GAG	ATC	CTT	1620
526	Gly	Leu	Thr	Gly	Lys	Ile	Asp	Ser	Arg	Lys	Ile	Arg	Glu	Ile	Leu	540
1621	ACG	ATG	GGA	AAA	AAA	TCA	AAA	CTG	TAA							1647
541	Thr	Met	Gly	Lys	Lys	Ser	Lys	Leu	End							

Fig. 2 The nucleotide and deduced amino acid sequences of *Pyrocoelia pygidialis* luciferase gene

40-G247), L3 (G317-K322), L4 (L343-I350) and L5 (G522-D532)] of the luciferase cleft area are shown in Fig. 6. It also shows one of the different amino acid residues between *P. pygidialis* luciferase and its

congeneric species.

### 3 Discussion

The cDNA encoding luciferase in *P. pygidialis* was

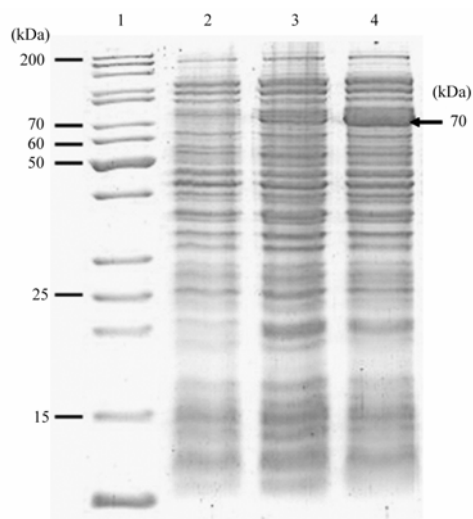


Fig. 3 SDS-PAGE analysis of *Pyrocoelia pygidialis* luciferase expressed in *E. coli* cells

Molecular weight standards were used as size marker (lane 1). The BL21 (DE3) *E. coli* cells were transformed by empty pET-28A (lane 2) and pET28a-Luc luciferase recombinant (lane 3, 4). The solid arrow on the right indicates the luciferase band of 70kDa.

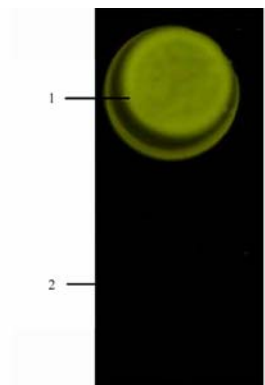


Fig. 4 Photo of a glowing clone, taken using a digital camera (Sony DSC-707) by long exposure and high light-sensitivity

- 1: *Pyrocoelia pygidialis* luciferase luminescence activity;  
2: the negative control (DE3 with pET-28a only).

cloned and expressed functionally in *E. coli* (Fig. 4). The ORF of this cDNA consisted of 1647 bp encoding a polypeptide of 548 amino acid residues. Phylogenetic analysis using deduced amino acid seq-

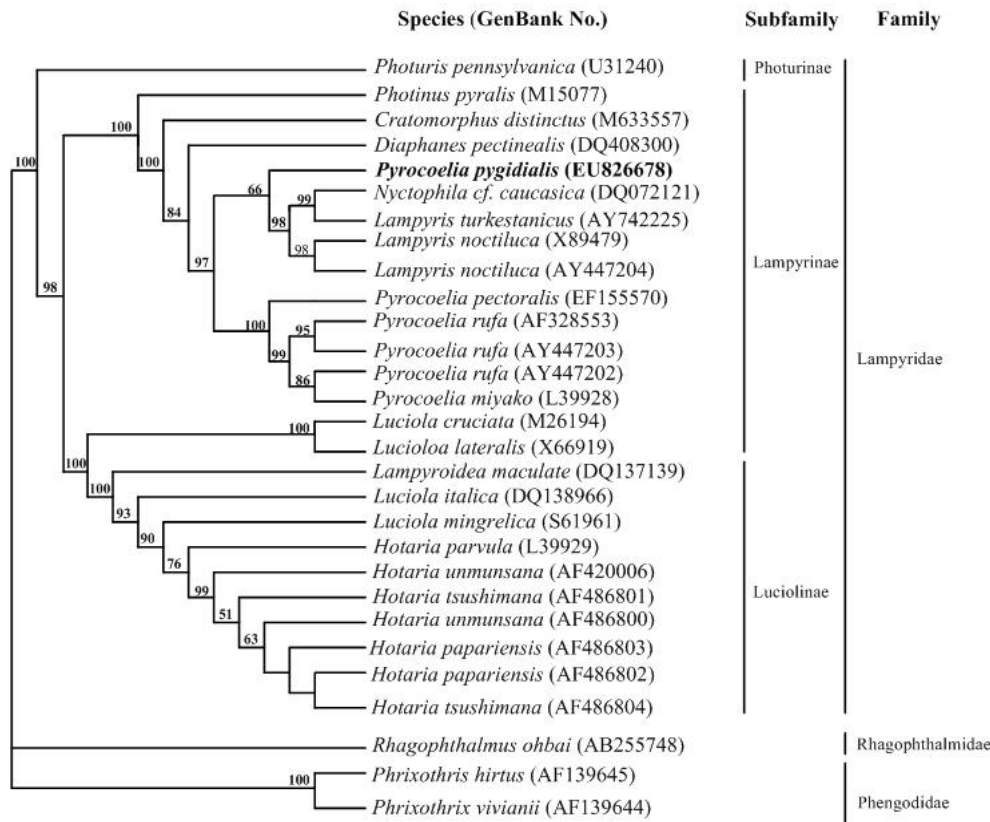


Fig. 5 A phylogenetic tree for aligned amino acid sequences of *Pyrocoelia pygidialis* luciferase and the known luciferase

The bootstrap values exceeding 50% were shown. *Rhagophthalmus ohbai* (Rhagophthalmidae), *Phrixothrix vivianii* and *Phrixothrix hirtus* (Phengodidae) were chosen as outgroups.



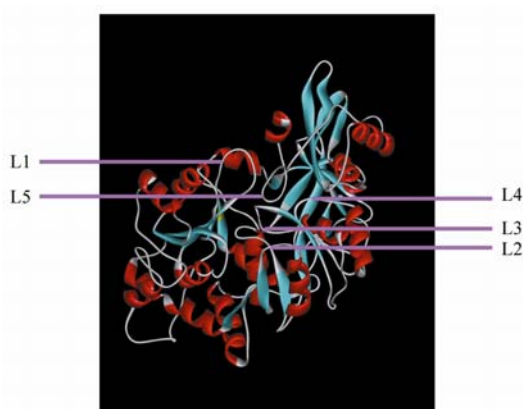


Fig. 6 Structural modeling of *Pyrocoelia pygidialis* Luciferase

The ribbon presentation of structural models were built using SWISS-MODEL server, based on crystal structure of *L. cruciata* (PDB code: 2d1sA). The  $\alpha$ -Helices are in red, the  $\beta$ -sheets are in blue, and the random coils are in grey. Five conserved loops (with purple line) located at the putative active area. The yellow position was the amino acid residue E210, while the correspondent position in other *Pyrocoelia* luciferases is aspartic acid. Note: E is glutamic acid.

uence showed that *P. pygidialis* did not form a monophyletic group with its congeneric species *P. rufa*, *P. pectoralis* and *P. miyako*, all three of which are strong luminous species. The analysis of 16S mitochondrial DNA sequences also supported *P. pygidialis*, together other weakly luminous congeneric species from Japan or China, constituted a separate group from those strong luminous taxa (Suzuki, 1997; Li et al, 2006b; Wang et al, 2007). *P. pygidialis*, with weakly luminous male adults,

is considered a diurnal firefly male, deduced from the morphology of its photic organs, i.e., degenerate photic organ of two-spot on the 7th ventrite (Li et al, 2008).

Five peptide loops (Fig. 6) located in the *P. pygidialis* luciferase cleft area between the C- and N-terminal domains, the active site for the substrate compound is entering and binding (Conti et al, 1996; Ohmiya et al, 1996; Nakatsu et al, 2006). Thirteen different residues of the luciferases of *P. pygidialis* with other known luciferases of *Pyrocoelia* species were located on the surface of the luciferase molecule. Future studies on these loops, residues or crystal structure analysis may be helpful in understanding the real difference between the luciferases from diurnal and nocturnal species.

Due to the lack of luciferase sequences from other weak luminous *Pyrocoelia* species (Suzuki, 1997), it is still too early to say that which is primitive for weak and strong luminous species in the Asian genus *Pyrocoelia* (Li et al, 2008). Nevertheless, it seems that weak luminous species, with degenerate photic organs in morphology, possibly evolved from a separated lineage to those possessing strong luminous species.

**Acknowledgements:** We are grateful to Dr WANG Wen for his special support to this project. Thanks to all other colleagues of Dr WANG's Max-Planck group for assistances in the laboratory work, too.

## Reference:

- Alipour BS, Hosseinkhani S, Nikkhah M, Naderi-Manesh H, Chaichi MJ, Osaloo SK. 2004. Molecular cloning, sequence analysis and expression of a cDNA encoding the luciferase from the glow-worm, *Lampyrus turkestanicus* [J]. *Biochem Biophys Research Comm*, **325**: 215-222.
- Arnold K, Bordoli L, Kopp J, Schwede T. 2006. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling [J]. *Bioinformatics*, **22**: 195-201.
- Branham MA, Wenzel JW. 2001. The evolution of bioluminescence in cantharoids (Coleoptera: Elateroidea) [J]. *Fla Entomol*, **84**: 565-586.
- Branham MA, Wenzel JW. 2003. The origin of photic behavior and the evolution of sexual communication in firefly (Coleoptera: Lampyridae) [J]. *Cladistics*, **19**: 1-22.
- Branchini BR, Southworth TL, DeAngelis JP, Roda A, Michelini E. 2006. Luciferase from the Italian firefly *Luciola italica*: Molecular cloning and expression [J]. *Comp Biochem Physiol B*, **145**: 159-167.
- Choi YS, Lee KS, Bae JS, Lee KM, Kim SR, Kim I, Lee SM, Sohn HD, Jin BR. 2002. Molecular cloning and expression of a cDNA encoding the luciferase from the firefly, *Hotaria unimunsana* [J]. *Comp Biochem Physiol B*, **132**: 661-670.
- Conti E, Franks NP, Brick P. 1996. Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes [J]. *Structure*, **4**: 287-298.
- DeLuca M. 1976. Firefly luciferase [J]. *Adv Enzymol*, **44**: 37-68.
- Devine JH, Kutuzova GD, Green VA, Ugarova NN, Baldwin TO. 1993. Luciferase from the East European firefly *Luciola mingrelica*: cloning and nucleotide sequence of the cDNA, overexpression in *Escherichia coli* and purification of the enzyme [J]. *Biochim Biophys Acta*, **1173**: 121-132.
- De Wet JR, Wood KV, Helinski DR, Helinski M, DeLuca. 1985. Cloning of firefly luciferase cDNA and the expression of active luciferase in *Escherichia coli* [J]. *Proc Natl Acad Sci USA*, **82**: 7870-7873.
- De Wet JR, Wood KV, DeLuca M, Subramani S. 1987. Firefly luciferase gene: Structure and expression in mammalian cells [J]. *Mol Cell Biol*, **7**: 725-737.
- Emamzadeh AR, Hosseinkhani S, Sadeghizadeh M, Nikkhah M. 2006. cDNA cloning, expression and homology modeling of a luciferase

- from the firefly *Lampyroidea maculata* [J]. *J Biochem Mol Biol*, **39**(5): 578-585.
- Gattiker A, Gasteige E, Bairoch A. 2002. ScanProsite: A reference implementation of a PROSITE scanning tool [J]. *Appl Bioinformatics*, **1**: 107-108.
- Gould SJ, Subramani S. 1988. Firefly luciferase as a tool in molecular and cell biology [J]. *Anal Biochem*, **175**: 5-13.
- Guex N, Peitsch MC. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling [J]. *Electrophoresis*, **18**: 2714-2723.
- Hirokawa K, Kajiyama N, Murakami S. 2002. Improved practical usefulness of firefly luciferase by gene chimerization and random mutagenesis [J]. *Biochim Biophys Acta*, **1597**: 271-279.
- Jeng ML, Yang PS, Sat M. 2002. Notes on the morphology and systematics of the genus *Pristolycus* Gorham (Coleoptera, Lampyridae) [J]. *Jpn J Syst Ent*, **8** (1): 87-108.
- Jeng ML, Yang PS, Sat M. 2006. Synopsis of *Cyphonocerus* (Coleoptera: Lampyridae) with the description of four new species and a key to the genus [J]. *Zoological Studies*, **45**(2): 157-167.
- Lee KS, Park HJ, Bae JS, Goo TW, Kim I, Sohn HD, Jin BR. 2001. Molecular cloning and expression of a cDNA encoding the luciferase from the firefly, *Pyrocoelia rufa* [J]. *J Biotechnol*, **92**: 9-19.
- Li XY, Yang S, Liang XC. 2006a. Phylogenetic relationship of the Firefly, *Diaphanes pectinealis* based on DNA sequence and gene structure of luciferase [J]. *Zool Res*, **27**: 367-374.
- Li XY, Yang S, Xie M, Liang XC. 2006b. Phylogeny of fireflies (Coleoptera: Lampyridae) inferred from mitochondrial 16S ribosomal DNA, with references to morphological and ethological traits [J]. *Prog Nat Sci*, **19**: 817-826.
- Li XY, Xie M, Dong PX, Liang XC. 2008. Morphology of *Pyrocoelia pygidialis* Pic (Coleoptera: Lampyridae) with notes on its biology [J]. *Entomotaxonomia*, Accepted
- Masuda T, Tatsumi H, Nakano E. 1989. Cloning and sequence analysis of a cDNA for luciferase of a Japanese firefly, *Luciola cruciata* [J]. *Gene*, **77**: 265-270.
- McElroy WD, Seliger HH, White EH. 1969. Spectral emission and quantum yield of firefly bioluminescence [J]. *Photochem Photobiol*, **10**: 153-170.
- Nakatsu T, Ichiyama S, Hiratake J, Saldanha A, Kobashi N, Sakata K, Kato H. 2006. Structural basis for the spectral difference in luciferase bioluminescence [J]. *Nature*, **440**: 372-376.
- Ohba, N. 2004. Mystery of Fireflies [M]. Kanagawa: Yokosuka City Museum, 199.
- Ohmiya Y, Ohba N, Toh H, Tsuji FI. 1995. Cloning, expression, and sequence analysis of cDNA for the luciferases from the Japanese fireflies, *Pyrocoelia miyako* and *Hotaria parvula* [J]. *Photochem. Photobiol*, **62**: 309-313.
- Ohmiya Y, Hirano T, Ohashi M. 1996. The structural origin of the color differences in the bioluminescence of firefly luciferase [J]. *FEBS Lett*, **384**: 83-86.
- Rong JJ, Chen ZY, Zhou GH. 2007. Molecular clone, expression and immobilization study on the Biotinylation luciferase [J]. *China Biotechnol*, **27**(9): 41-46.
- Saitoh N, Nei M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees [J]. *Mol Biol Evo*, **4**: 406-425.
- Sala-Newby GB, Thomson CM, Campbell AK. 1996. Sequence and biochemical similarities between the luciferases of the glow-worm *Lampyris noctiluca* and the firefly *Photinus pyralis* [J]. *Biochem J*, **313**: 761-767.
- Schwede T, Kopp J, Guex N, Peitsch MC. 2003. SWISS-MODEL: An automated protein homology-modeling server [J]. *Nucleic Acids Research*, **31**: 3381-3385.
- Suzuki H. 1997. Molecular phylogenetic studies of Japanese fireflies and their mating systems (Coleoptera: Cantharoidea) [J]. *TMU Bulletin of Natural History*, **3**: 1-53.
- Swofford DL. 2002. PAUP\*: Phylogenetic Analysis Using Parsimony (and other methods). Sinauer Associates, Sunderland, MA.
- Tatsumi H, Kajiyama N, Nakano E. 1992. Molecular cloning and expression in *Escherichia coli* of a cDNA clone encoding luciferase of a firefly, *Luciola lateralis* [J]. *Biochim Biophys Acta*, **1131**: 161-165.
- Thompson JD, Gibson TJ, Plewniak F. 1997. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools [J]. *Nucleic Acids Res*, **24**: 4876-4882.
- Ugarova NN, Brovko LY. 2002. Protein structure and bioluminescent spectra for firefly bioluminescence [J]. *Luminescence*, **17**: 321-330.
- Viviani VR, Bechara EJH, Ohmiya Y. 1999. Cloning, sequence analysis, and expression of active *Phrixothrix* railroad worm luciferases: relationship between bioluminescence spectra and primary structures [J]. *Biochemistry*, **38**: 8271-8279.
- Viviani VR, Ohmiya Y. 2000. Bioluminescence color determinants of *Phrixothrix* railroad-worm luciferases: Chimeric luciferases, site-directed mutagenesis of Arg 215 and guanidine effect [J]. *Photochem Photobiol*, **72**(2): 267-271.
- Viviani V, Uchida A, Suenaga N, Ryufuku M, Ohmiya Y. 2001. Thr 226 is a key residue for bioluminescence determination in beetle luciferases [J]. *Biochemical and Biophysical Research Communications*, **280**: 1286-1291.
- Viviani VR, Arnoldi FGC, Brochetto-Braga M, Ohmiya Y. 2004. Cloning and characterization of the cDNA for the Brazilian *Cratomorphus distinctus* larval firefly luciferase: similarities with European *Lampyris noctiluca* and Asiatic *Pyrocoelia luciferases* [J]. *Comp Biochem Physiol B*, **139**: 151-156.
- Viviani VR, Oehlmeyer TL, Arnoldi GC, Brochetto-Braga MR. 2005. A new firefly luciferase with bimodal spectrum: Identification of structural determinants of spectral pH-sensitivity in firefly luciferase [J]. *Photochem Photobiol*, **81**: 843-838.
- Wang YY, Zheng XL, Fu XH, Lei CL. 2007. Reproductive behavior of the terrestrial firefly, *Pyrocoelia pectoral* [J]. *Chinese Bulletin of Entomol*, **44**: 415-419.
- White EH, Rapaport E, Seliger HH, Hopkins TA. 1971. The chemi- and bioluminescence of firefly luciferin: an efficient chemical production of electronically excited states [J]. *Bioorg Chem*, **1**: 92-122.
- Wood KV. 1995. The chemical mechanism and evolutionary development of beetle bioluminescence [J]. *Photochem Photobiol*, **62**(4): 662-673.
- Ye L, Buck LM, Schaeffer HJ, Leach FR. 1997. Cloning and sequencing of a cDNA for firefly luciferase from *Photuris pennsylvanica* [J]. *Biochim Biophys Acta*, **1339**: 39-52.