Int. J. Indust. Entomol. Vol. 1, No. 2, 2000, pp. 165~169

[Communication]

International Journal of
Industrial Entomology

Analysis of Expressed Sequence Tags of the Firefly, *Pyrocoelia rufa*

Kwang Sik Lee, Jin Sik Bae, Tae Won Goo[1], Sam Eun Kim[1], Jong Gill Kim[1], Hung Dae Sohn and Byung Rae Jin*

College of Natural Resources and Life Science, Dong-A University, Pusan 604-714, Korea. [1]Department of Sericulture and Entomology, National Institute of Agricultural Science and Technology, RDA, Suwon 441-100, Korea.

(Received 15 October 2000; Accepted 15 November 2000)

We have constructed cDNA library from the larvae whole body of the firefly, *Pyrocoelia rufa*. Single direct partial sequencing of anonymous cDNA clones was performed to obtain genetic information on the firefly, *P. rufa*, of which genetic information is currently not available. This expressed sequence tags (EST) analysis of the 54 clones (54%) showed significant homology to the known genes registered in GenBank. Of these clones, twenty-four were related to the known insect genes, but these clones were not matched to previously identified firefly genes. Putative functional categories of these clones showed that the next abundant genes were associated with energy metabolism.

Key words : Firefly, *Pyrocoelia rufa*, Expressed sequence tags

Introduction

Fireflies are well known as luminous beetles, which emits flashes with species-specific duration and frequency as signals for mating and hunting (Llyod, 1983). Firefly luciferase catalyses the oxidative decarboxylation of D-luciferin in the presence of ATP and thereby light is emitted (Lembert, 1996). In firefly, there are well-known luciferase genes. The firefly luciferase genes also have been studied deeply in some species (de Wet *et al.*, 1987; Masuda *et al.*, 1989; Tatsumi *et al.*, 1992; Cho *et al.*, 1999; Devine *et al.*, 1993; Ohiyama *et al.*, 1995). The luciferase gene isolated from the firefly is increasingly used as a reporter gene in molecular biology. Furthermore, the gene encoding firefly luciferase has been shown to be a highly-effective reporter gene in many organisms including bacteria (Jacobs *et al.*, 1993), cellular slime moulds (Howard *et al.*, 1988), plants (Miller *et al.*, 1992), and mice (DiLella *et al.*, 1988; Kondo *et al.*, 1992).

The recent rapid progress in several genome projects has brought the advancement of a new research area, structural genomics. It is required to establish tools for the utilization of the information hidden in the genome. An efficient way for gathering information on the genome of an organism is to generate and analyze expressed sequence tags (ESTs). EST analysis is one of the powerful strategies for such approaches. Partial sequencing of cDNAs is a relatively rapid and economic way to obtain genetic information on structure and expression-profile of the genes expressed in different tissues and developmental stages of various organs (Adams *et al.*, 1992; Lee *et al.*, 2000; Lim *et al.*, 2000; Yun *et al.*, 2000).

Pyrocoelia rufa is an abundant in Korea and is only found at Tsushima in case of Japan (Suzuki, 1997). The body size of *P. rufa* is larger than that of the other species, and *P. rufa* has the largest luminescent organs among the fireflies occurring in Korea. *P. rufa* emits a strong light continuously (Suzuki, 1997). The nucleotide sequences of cDNA encoding luciferase of *P. miyako* in the *Pyrocoelia* group have been reported (Ohmiya *et al.*, 1995). However, the genetic information of the firefly, *P. rufa*, is not available currently.

In this report, we constructed and analyzed ESTs from *P. rufa* cDNA library. The information of the ESTs will be useful for the molecular genetic research of *P. rufa*.

Materials and Methods

Preparation of firefly

The larvae of the firefly *Pyrocoelia rufa* were collected at

*To whom correspondence should be addressed.
College of Natural Resources and Life Science, Dong-A University, Pusan 604-714, Korea. Tel & Fax: +82-51-200-7594;
E-mail: brjin@mail.donga.ac.kr

Muju, Chollabuk Province in Korea. The live larvae were directly used in this study.

Purification of poly(A)+ mRNA

Total RNAs were isolated from the whole body of *P. rufa* larvae following the procedure of Total RNA extraction kit (Promega). Poly(A)+ mRNA was purified using oligo(dT) columns of Quick mRNA isolation kit (Stratagene).

Construction of cDNA library

A cDNA library was constructed from poly(A)+ mRNA isolated from the whole body of *P. rufa* larvae by Uni-ZAP XR vector and Gigapack III Gold Packing Extract (Stratagene). The cDNA was ligated into *Eco*RI-*Xho*I sites of Uni-ZAP XR vector. Ligated library was transformed into *E. coli* XL1-Blue MRF strain.

Nucleotide sequencing and EST analysis

E. coli XL1-Blue MRF strain was infected by the Uni-ZAP XR library harboring *P. rufa* cDNA and cultured on the NZY agar medium. Each plaque was suspended in SM buffer [5.8 g/l NaCl, 2 g/l $MgSO_4 \cdot 7H_2O$, 0.05 M TrisCl (pH 7.5) and 0.01% gelatin solution] containing 0.02% (v/v) chloroform and stored at 4°C for 1 day. The pages were eluted into SM buffer. The pBluescript phagemids were *in vivo* excised from the Uni-ZAP XR vector using an ExAssist helper phage. *E. coli* strain, SOLR cell (Stratagene), was infected by the excised phagemids and plated on LB-Amp medium (50 μg/ml ampicillin). Plasmid DNA from the overnight culture was isolated. The size of inserted cDNA was estimated with a 1% agarose gel electrophoresis after treatment of restriction enzymes (*Eco*RI and *Xho*I). For DNA sequencing, plasmid DNA was extracted by Wizard mini-preparation kit (Promega). Sequence of the 5' end of each cDNA clone was determined using ABI PRISM BigDye Terminator Cycle Sequencing Kit (Perkin-elmer). The sequences were translated into 6 reading frames and compared using the DNASIS and BLAST programs provided by the NCBI. GenBank, EMBL and SwissProt databases were searched for sequence homology using a BLAST algorithm program.

Results and Discussion

Construction of cDNA library was prepared from the whole body of *P. rufa* larvae. The size of inserted cDNA was estimated on a 1% agarose gel (Fig. 1). As expected, the inserted cDNA was various in the size. The partial sequencing of randomly selected clones harboring cDNA inserts was performed to generate the *P. rufa* ESTs.

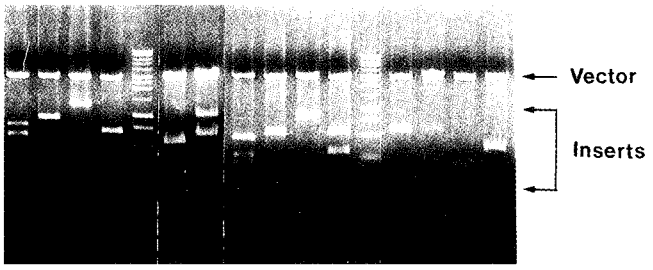

Fig. 1. Estimation of the size of inserted cDNA. The size of cDNA inserts randomly selected was estimated on a 1% agarose gel. One kb ladder is used as DNA size makers.

The putative identification of *P. rufa* ESTs is summarized in Table 1. Among 100 ESTs generated, 54 sequence tags showed significantly high amino acid sequence similarity to the known genes, but the other 46 clones were not homologous to the previously identified genes. Of these 54 clones, furthermore, twenty-four sequence tags are related to the known insect genes such as *Drosophila melanogaster*, but these clones were not matched to the previously identified firefly genes. Among the 24 ESTs related to the known insect genes, approximately 46% matched with *D. melanogaster* genes.

Putatively identified ESTs were classified into 9 functional groups, principally based on known functions (Table 2). The result showed that the most abundant genes were involved in gene expression and that the next abundant genes were associated with energy metabolism. Seven ESTs showed significantly high homology to the genes that are involved in cytoskeletal structure. Those genes are tropomyosin (PR65), myosin 2 essential light chain (PR242), myosin light chain 2 (PR1441), annexin (PRX-21), myophilin (PRB-6 and PRB-81), and unconventional myosin (PR2-19). There are high frequencies of translation-related genes encoding ribosomal proteins (PR768, PR824, PR1421, PR1505, PR1429, and PRB-74). Histone H3 (PRB-37), especially, showed on about 99% of identity to the matched region of *D. melanogaster* gene at the amino acid sequence level. Two ESTs including PR1439 and PRB-85 are related to pheromone binding protein.

The EST analyzed in this study was registered in GenBank. Results obtained from the *P. rufa* EST study can be applied for understanding of genetic information in the firefly.

Acknowledgments

This work was supported by a grant (Institute Research 2000) from the Dong-A University.

Table 1. Putative identification of Expressed Sequence Tags from *Pyrocoelia rufa* cDNA library

Clone	Putative identification	Species	Overlength (length)	Homology (%)	Acc. No.
PR1-10	CoA-dehydrogenase	*Homo sapiens*	572	63	AT003771
PR1-15	Troponin T	*Drosophila melanogaster*	476	81	AT003772
PR1-27	ATPase synthase 6	*Limulus polyphemus*	103	88	AT003773
PR1-36	CG3981	*Drosophila melanogaster*	443	51	AT003774
PR1-38	Integral membrane transporter protein	*Homo sapiens*	101	84	AT003775
PR1-43	ATPase subunit E	*Manduca sexta*	542	63	AT003776
PR1-82	RNase L inhibitor	*Homo sapiens*	278	88	AT003777
PR2-19	Unconventional myosin-15	*Mus musculus*	323	37	AT003778
PR2-21	CG3981	*Drosophila melanogaster*	257	53	AT003779
PR2-60	Sodium channel alpha subunit	*Homo sapiens*	308	41	AT003780
PR2-79	Chitinase	*Aeromonas hydrophila*	113	60	AT003781
PRX-1	Core 1 UDP-galactose	*Mus musculus*	326	51	AT003782
PRX-17	Helicase-like protein	*Arabidopsis thaliana*	131	54	AT003783
PRX-19	Non-funtional folate binding protein	*Homo sapiens*	389	42	AT003784
PRX-20	Integral membrane subunit	*Aedes aegypti*	125	80	AT003785
PRX-21	Annexin	*Bombyx mori*	374	64	AT003786
PRX-29	Carboxypeptide	*Astacus astacus*	221	60	AT003787
PRB-1	CG3884	*Drosophila melanogaster*	413	67	AT003788
PRB-6	Myophilin	*Echinococcus multilocularis*	293	74	AT003789
PRB-17	Male specific serum polypeptide beta	*Ceratitis capitata*	254	49	AT003790
PRB-26	Serine protease inhibitor	*Schistocerca gregaria*	89	56	AT003791
PRB-32	Alpha L1 nicotinic acetyl choline receptor	*Acheta domesticus*	74	72	AT003792
PRB-37	Histone H3	*Drosophila melanogaster*	320	99	AT003793
PRB-61	Phopholipase A2 inhibitor gamma subunit	*Oxyuranus microlepilepidotus*	227	44	AT003794
PRB-73	Sodium/Potassium exchanging ATPase alpha subunit	*Ctenocephalides felis*	524	92	AT003795
PRB-74	Ribosomal protein L27	*Mus musculus*	404	72	AT003796
PRB-76	Endothelin-converting enzyme	*Bos taurus*	380	68	AT003797
PRB-81	Myophilin	*Echinococcus multilocularis*	293	74	AT003798
PRB-85	Pheromone binding protein	*Choristoneura murinara*	173	55	AT003799
PRB-90	Isopenicillin N synthetase	*Streptomyces cattleya*	170	48	AT003800
PRB-92	Serpin	*Oryza sativa*	203	57	AT003801
PRB-94	Nucleolar protein p40	*Homo sapiens*	95	84	AT003802
PRB-99	Elongation factor 1-alpha	*Nanexila gracilis*	227	94	AT003803
PR13	CG3950	*Drosophila melanogaster*	137	65	AT003804
PR65	Tropomyosin	*Drosophila melanogaster*	164	54	AT003805
PR113	Reverse transcriptase	*Caenorhabditis elegans*	93	49	AT003806
PR126	Easily shocked protein	*Caenorhabditis elegans*	227	74	AT003807
PR241	Ubiquitin	*Sus scrofa*	203	85	AT003808
PR242	Myosin essential light chain	*Caenorhabditis elegans*	440	80	AT003809
PR768	40S ribosomal protein	*Xenopus laevis*	581	89	AT003810
PR770	Glycophorin-binding protein	*Caenorhabditis elegans*	239	48	AT003811
PR824	40S ribosomal protein	*Drosophila melanogaster*	425	85	AT003812
PR826	Mitochondrial outer membrane protein	*Saccharomyces cerevisiae*	99	62	AT003813

Underlines represent insect species.

Table 1. Continued

Clone	Putative identification	Species	Overlength (length)	Homology (%)	Acc. No.
PR884	Peritrophin-95 precursor	*Lucilia cuprina*	113	60	AT003814
PR951	Cytochrome p450-like protein	*Bacillus subtilis*	143	39	AT003815
PR972	Z83216	*Caenorhabditis elegans*	239	47	AT003816
PR1183	RNA helicase	*Drosophila melanogaster*	287	95	AT003817
PR1310	Endothelin converting enzyme	*Rattus norvegicus*	188	49	AT003818
PR1421	60S ribosomal protein	*Drosophila melanogaster*	323	46	AT003819
PR1429	60S ribosomal protein L15	*Chironomus tentans*	257	48	AT003820
PR1439	Pheromone binding protein	*Popillia japonica*	188	53	AT003821
PR1441	Myosin light chain	*Drosophila melanogaster*	407	74	AT003822
PR1505	40S ribosomal protein S19	*Gillichthys mirabilis*	143	64	AT003823
PR1512	Cytochrome b	*Rhytidoponera sp.*	140	89	AT003824

Underlines represent insect species.

Table 2. Funtional categories of *Pyrocoelia rufa* ESTs matching to known genes

Funtional catagories	
CYTOSKELETAL STRUCTURE	**TRANSCRIPTION/TRANSLATION**
Tropomyosin (PR65)	Ribosomal protein (PR768, PR824, PR1421,
Myosin 2 essential light chain (PR242)	PR1505, PR1429, PRB-74)
Myosin light chain 2 (PR1441)	RNA helicase (PR1183)
Annexin (PRX-21)	Helicase-like protein (PRX-17)
Myophilin (PRB-6, PRB-81)	Elongation factor 1-alpha (PRB-99)
Unconventional myosin (PR2-19)	Reverse transcriptase (PR113)
	RNase L inhibitor (PR1-82)
	Histone H3 (PRB-37)
ENERGY METABOLISM	
Cytochrome b (PR1512)	
ATPase subunit (PR1-27, PR1-43)	**MEMBRANE-ASSOCIATED**
Sodium/Potassium exchanging ATPase alpha	Integral membrane subunit (PRX-20)
Subunit (PRB-73)	Nucleolar protein p40 (PRB-94)
Sodium channel alpha subunit (PR2-60)	Mitochondrial outer membrane (PR826)
CoA-dehydrogenase (PR1-10)	Peritrphin-95 precursor (PR884)
Cytochrome p450-like protein (PR951)	Integral membrane transporter protein (PR1-38)
Core 1 UDP-galactose (PRX-1)	
Isopenicillin N synthetase (PRB-90)	
Carboxypeptide (PRX-29)	**HORMONE RESPONSE**
	Pheromone binding protein (PR1439, PRB-85)
SIGNAL TRANSDUCTION	
Glycophorin-binding protein (PR770)	**STRESS & DEFFENCE**
Alpha L1 nicotinic acetyl choline receptor	Easily shocked protein (PR126)
(PRB-32)	Chitinase (PR2-79)
	Troponin T (PR1-15)

Table 2. Continued

Funtional catagories	
PROTEIN MODIFICATION/DEGRADATION	**MICELLANNEOUS**
Ubiquitin (PR241)	Male specific serum polypeptide beta (PRB-17)
	Endothelin-converting enzyme (PRB-76, PR1310)
	Non-funtional folate binding protein (PRX-19)
	CG3950 (PR13)
	Z83216 (PR972)
	CG3981 (PR1-36, PR2-21)
	CG3884 (PRB-1)
	Protease inhibitor (PRB-26, PRB-61, PRB-92)

References

Adams, M. D., M. Dubnick, A. R. Kerlavage, R. Moreno, J. M. Kelly, T. R. Utterback, J. W. Nagle, C. Fields and C. Venter (1992) Sequence identification of 2375 human brain genes. *Nature* **355**, 632-634.

Cho, K. H., J. S. Lee, Y. D. Choi and K. S. Boo (1999) Structural polymorphism of the luciferase gene in the firefly, *Luciola lateralis*. *Insect Mol. Biol.* **8**, 193-200.

Devine, J. H., G. D. Kutuzova, V. A. Green, N. N. Ugarova and T. O. Baldwin (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. *Biochim. Biophys. Acta* **1173**, 121-132.

DiLella, A. G., D. A. Hope, H. Chen, M. Trumbauer, R. J. Schwartz and R. G. Smith (1988) Utility of firefly luciferase as a reporter gene for promoter activity in transgenic mice. *Nucleic Acids Res.* **16**, 4159.

Howard, P. K., K. G. Ahern and R. A. Firtel (1988) Establishment of a transient expression system for *Dictyostelium discoideum*. *Nucleic Acids Res.* **16**, 2613-2623.

Jacobs, W. R., R. G. Barletta, R. Udami, J. Chan, G. Kalkut, G. Sosne, T. Kieser, G. J. Sarkis, G. F. Hatfull and B. R. Bloom (1993) Rapid assessment of drug susceptibilities of *Mycobacterium tuberculosis* by means of luciferase reporter phages. *Science* **260**, 819-822.

Kondo, T., N. Takahashi and M. Muramatsu (1992) The regulation of the murine Hox-2.5 gene expression during cell differentiation. *Nucleic Acids Res.* **20**, 5729-5735.

Lee, E. K., S. B. Seo, T. H. Kim, S. K. Sung, G. An, C. H. Lee and Y. J. Kim (2000) Analysis of expressed sequence tags of *Porphyra yezoensis*. *Mol. Cells* **10**, 338-342.

Lembert, N. (1996) Firefly luciferase can use L-luciferin to produce light. *Biochem. J.* **317**, 273-277.

Lim, J. Y., C. S. Shin, E. J. Chung, J. S. Kim, H. U. Kim, S. J. Oh, W. B. Choi, C. S. Ryou, J. B. Kim, M. S. Kwon, T. Y. Chung, S. I. Song, J. K. Kim, B. H. Nahm, Y. S. Hwang, M. Y. Eun, J. S. Lee, J. J. Cheong and Y. D. Choi (2000) Analysis of expressed sequence tags from *Brassica rapa* L. ssp. *pekinensis*. *Mol. Cells* **10**, 399-404.

Lloyd, J. E. (1983) Bioluminescence and communication in insects. *Ann. Rev. Entomol.* **28**, 131-160.

Masuda, T., H. Tatsumi and E. Nakano (1989) Cloning and sequence analysis of cDNA for luciferase of a Japanese firefly, *Luciola cruciata*. *Gene* **77**, 265-270.

Millar, A. J., S. R. Short, N. H. Chua and S. A. Say (1992) A novel circadian phenotype based on firefly luciferase expression in transgenic plants. *Plant Cell* **4**, 1075-1087.

Ohiyama, Y., N. Ohba, H. Toh and F. I. Tsuji (1995) Cloning, expression and sequence analysis of cDNA for the luciferases from the Japanese fireflies, *Pyrocoelia miyako* and *Hotaria parvula*. *Photochem. Photobiol.* **62**, 309-313.

Suzuki, H. (1997) Molecular phylogenetic studies of Japanese fireflies and their mating systems (Coleoptera: Cantharoidea). *TMU Bull. Natural History* **3**, 1-53.

Tatsumi, H., N. Kajiyama and E. Nakano (1992) Molecular cloning and expression in *Escherichia coli* of a cDNA clone encoding luciferase of a firefly, *Luciola lateralis*. *Biochim. Biophys, Acta* **1131**, 161-165.

de Wet, J. R., K. W. Wood, M. DeLuca, D. R. Helinski and S. Subramani (1987) Firefly luciferase gene: structure and expression in mammalian cells. *Mol. Cell Biol.* **7**, 725-737.

Yun, E. Y., T. W. Goo, S. W. Kang, H. W. Lee, J. S. Hwang and H. R. Kim (2000) Analysis of expressed sequence tags generated from the posterior silkgland cDNA clones of *Antherea yamamai*. *Korean J. Life Sci.* **10**, 188-195.