

Characterisation of the luciferase gene and the 5' upstream region in the European glow-worm *Lampyris noctiluca* (Coleoptera: Lampyridae)

JOHN C. DAY

Centre for Ecology and Hydrology (CEH) – Oxford, Mansfield Road, Oxford, OX1 3SR, UK; e-mail: jcda@ceh.ac.uk

Keywords. *Lampyris noctiluca*, *Photinus pyralis*, luciferase, transposase, DDE domain, bioluminescence

Abstract. Beetle luciferase, a mono-oxygenase within the AMP-binding superfamily, is synthesized by bioluminescent beetles in concentrated levels within specialised cells clustering in the abdominal light organs. In vivo expression of luciferase has been rarely investigated and little is known about the role of enhancers and promoters in the expression of this gene. In order to investigate the gene structure and potential control of gene expression the luciferase gene along with 6 kb of upstream genomic sequence was characterised from the European glow-worm *Lampyris noctiluca*. Three TATA box motifs and a CAAT repeat were identified, two of these were found to be conserved in two other species of bioluminescent beetle. Although no enhancer regions were identified in the upstream sequence a region coding for a putative transposase DDE domain was identified 686 bp from the start codon of the luciferase gene. Although disrupted, the open reading frame also shared extensive identity to an mRNA transcript from the mosquito *Anopheles gambiae*. The remnants of an ancient transposase provide support for an ancestral luciferase transposition/insertion event that may have occurred within the genome of bioluminescent beetles.

INTRODUCTION

Along with fireflies the European glow-worm *Lampyris noctiluca* (Linnaeus) is one of over 2000 bioluminescent beetle species belonging to the family Lampyridae. The geographical distribution of *L. noctiluca* is extensive, ranging from Portugal in the West through Europe to China in the East. (Tyler, 1986). This range and also abundance probably makes the European glow-worm the most studied of all Palearctic lampyrids. *L. noctiluca*, along with other bioluminescent beetles, develop specialised cells localised in organs known as lanterns (Buck, 1948). In the glow-worm, these are located in the terminal abdominal segments where yellow-green light is produced functioning primarily for mate attraction in the adult (Newport, 1857; Tyler, 1994). The enzyme responsible for the light production, luciferase, is often concentrated within the lanterns of the beetle. Firefly luciferase (EC 1.13.12.7) from the American firefly species *Photinus pyralis* is a 62 kDa enzyme that catalyses the emission of yellow-green light (λ_{max} 560 nm) upon reaction of D-luciferin, ATP, and molecular oxygen (McElroy et al., 1969; White et al., 1971; Deluca & McElroy, 1978). Due to its high sensitivity and specificity for ATP, beetle luciferase has been used for ATP detection in a variety of biological samples including in vivo luminescence monitoring as well as medical and pharmaceutical protocols (Kricka, 2000). The utility of luciferase in a range of application has resulted in a large body of enzymatic studies (Viviani, 2000), but little research directed at the genetics and regulation of bioluminescence in beetles. The cloning and sequencing of luciferase cDNA from *Photinus pyralis* and many other bioluminescent beetle species has revealed that these enzymes are highly conserved proteins closely related to a large family of non-bioluminescent proteins that catalyse reactions of ATP with carboxylate substrates to form acyl-adenylates (Conti et al., 1996). The evolution of bioluminescence in beetles still remains a mystery but it is commonly thought that luciferase may have evolved from an ancestral AMP-binding protein with homologous presence in other insect genomes (Day et al., 2001). It has been shown that distantly related non-bioluminescent beetles are capable of emit-

ting extremely low levels of light although the enzymes responsible have yet to be fully characterised (Viviani & Bechara, 1996). Furthermore, it has recently been shown that luciferase has a dual function, in vitro and therefore possibly in situ, both as an oxygenase and as a long chain fatty acyl-CoA synthetase (Ohba et al., 2003).

Although numerous beetle luciferases have been characterised, little is known about the regulatory region upstream of the luciferase gene (*luc*). The regulation of luciferase in the firefly has been examined at a physiological level whereby studies in the concentrations and the localisation of luciferase and luciferin during development of the firefly *Photuris pennsylvanica* revealed that during pupation luciferase and luciferin levels remained constant in the posterior half of the pupa and in the anterior half showed an initial increase followed by a decrease (Strause et al., 1979). It was further reported that the whole body luminescence characteristic of the *P. pennsylvanica* pupa is correlated with widespread increasing levels of luciferin and luciferase that are not a result of release from larval lantern degeneration, but are a result of synthesis throughout the pupa that, upon maturation, become localised in the region of the adult lantern (Strause et al., 1979). Since the pupal glow persists after excision of the light organs it cannot be caused by dispersal of the contents of the larval light organs and the proposed widespread synthesis of luciferin and luciferase throughout the pupa (Strause et al., 1979). Although this physiological data suggest dramatic shifts in luciferase expression during pupation, to date no genetic information is available on the regulation of luciferase.

Genomic clones coding for *luc* have been published from a number of different genera of Lampyridae but little or no investigation has extended into these flanking regions (de Wet et al., 1987; Cho et al., 1999; Choi et al., 2003; Kim et al., 2005). One investigation into the nature of the *luc* upstream region in the Japanese firefly *Luciola lateralis* revealed allelic variation but described no significant open reading frames (Cho et al., 1999). Although the luciferase amino acid sequence has been reported for over 12 different species of beetle spanning three coleopteran families and extensive mutagenic studies have been con-

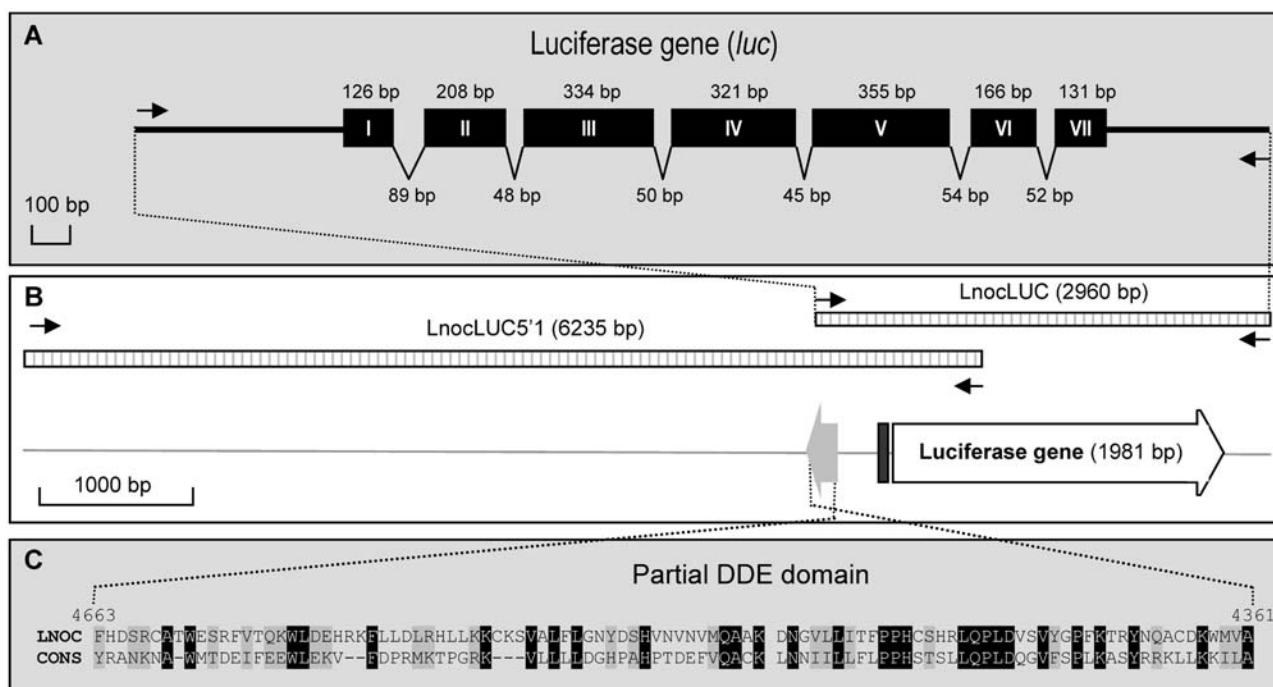


Fig. 1. Graphic representation of the PCR amplification and gene architecture of *luc* and flanking region genomic DNA from *Lampyrus noctiluca*. A. Gene structure of *luc* based upon 2960 bp PCR product, sizes of exons (black boxes) and introns (inverted triangles) are shown. B. Schematic representation of 8 kb sequenced genomic region containing *luc* based upon two PCR products (shaded boxes, arrows indicate primer positions). Positions of the core promoter region (black box) and the DDE transposase domain (gray box) are indicated. Open reading frame directions are indicated with open arrows. C. Amino acid translation of putative DDE domain from *L. noctiluca* (LNOC) aligned with the consensus sequence from pfam03184, DDE superfamily endonuclease (CONS).

ducted on the enzyme, little information has been provided on the *luc* upstream region and no identification of control elements, such as promoters and enhancer sequences has been attempted.

Although much has been speculated about the origins of the substrate luciferin in beetles (Day et al., 2004) nothing is known about the biosynthesis of luciferin. In order to investigate the possibility of luciferin biosynthesis genes being linked to the luciferase gene the *luc* flanking region was examined for putative open reading frames. This study presents the first investigation into the *luc* flanking region from the European glow-worm *L. noctiluca*. Furthermore, this study is the first to report the presence of a putative transposase domain in close proximity to the beetle luciferase gene region.

EXPERIMENTAL PROCEDURES

Specimen details and DNA preparation

A single adult female of the European glow-worm, *Lampyrus noctiluca*, was collected from an established colony in Southern England and stored at -70°C prior to use. Using the PCR template kit (Roche Diagnostics Ltd, Lewes, UK) total genomic DNA was extracted from the whole beetle and the majority used to construct an inverse PCR genome library using the Universal GenomeWalker Kit (Clontech, BD Biosciences Clontech, UK, Oxford, UK). In brief, approximately 2.5 mg of glow-worm DNA was digested separately with *Dra* I, *Eco* R V, *Pvu* II, *Sca* I and *Stu* I restriction enzymes to completion. To each digested DNA reaction GenomeWalker adaptors were added and ligations carried out overnight at 14°C . Unincorporated adaptors were removed by a phenol/chloroform extraction and after ethanol precipitation the DNA was resuspended in TE buffer and used as a template for PCR.

PCR, cloning and sequencing

Genome walking PCR primers were designed to sequentially walk out from the *L. noctiluca* luciferase gene in both a 5' and a 3' direction. Using the *L. noctiluca* luciferase cDNA sequence previously characterised (GenBank accession number: X89479) gene specific nested primer pairs were designed for upstream and downstream amplification (genome walking primer sequences are available from the author). Using each successive walker sequence the next primers were designed. The PCR was carried out in a 50 μl mixture containing 15 pmol of each primer (first gene specific primer and first adaptor primer), 75 mM Tris-HCl (pH 8.8), 2.5 mM magnesium chloride, 0.01% Tween 20, 1.3 M betaine, 400 mM of each dNTP. *Taq-Pfu* DNA polymerases mixture (15 : 1 units) was used. The cycling conditions were as follows: 7 cycles of 25 s at 94°C and 3 min at 72°C followed by 32 cycles of 25 s at 94°C and 3 min at 67°C . The last cycle was followed by the extension step for 7 min at 67°C . 1 μl of a 1 : 50 dilution of PCR mixture from the first round of PCR was used for second round PCR with the nested gene primer specific to the particular region and the second nested adaptor primer. PCR, the majority of the time, produced a single product that could be excised from the agarose gel and purified. In the result of multiple bands the whole PCR product was purified to remove small molecular weight products and the whole reaction cloned. The PCR products were cloned into pGEM Easy T (Promega UK Ltd, Southampton, UK), plasmid prepared using a plasmid mini prep kit (Qiagen Ltd, Crawley, UK) and the insert sequenced using a CEQ sequencing kit (Beckman Coulter Ltd, Uigh Wycombe, UK) with M13F and M13R primers and any internal oligonucleotides where necessary.

Based up the sixth sequential walk and the *luc* sequence, primers *LnocLUC5'F* 5' AGA GAT ACG AAG ATA GAT ATG

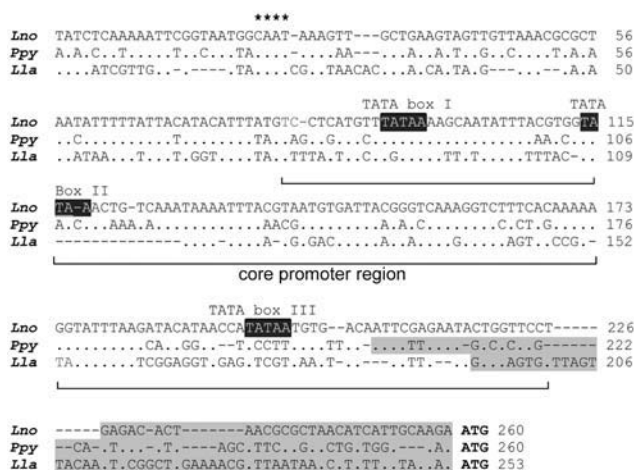


Fig. 2. Alignment of the *luc* upstream sequence from *Lampyris noctiluca* (Lno), *Photinus pyralis* (Ppy, GenBank accession number: M15077) and *Luciola lateralis* (Lla, GenBank accession number: U49182). The ATG *luc* start codon is shown at the terminus of the alignment. Gray shaded regions indicate 5' from cDNA sequences. TATA boxes are shown blocked in black and a CAAT box is highlighted with asterisks.

GAC ACG AC 3' and *LnocLUC5'R* 5' ATT TTT TTG CAG CGC TCT TTT GGA ACA GGA TAC 3' were designed to amplify a contiguous flanking region fragment extending over the length of six genomic walks through to the first 513 bp of *luc*. PCR amplification, cloning and sequencing was carried as described above. The fragment was primer sequenced in its entirety in both directions.

Sequence analysis

Putative promoter sites were determined using the Neural Network Promoter Prediction software via the Berkeley Drosophila Genome Project website. Translations of *LnocLUC5'I* in all six frames were used in an rpsBLAST search against a conserved domain database (CDD) at the NCBI website. tBLASTn searches were used to investigate the presence of open reading frames within flanking regions of *L. lateralis*.

RESULTS AND DISCUSSION

Luciferase gene of *Lampyris noctiluca*

Genome walking was carried out from the luciferase gene of *L. noctiluca* in both 5' and 3' directions. Only two walks out from the 3' end of the gene were successful before PCR failed to return contiguous product. However, 5' walking was successful for six overlapping walks. The initial genomic walks in both 5' and 3' directions gave sufficient sequence information to design primers (*LnocLUC F & R*) for the amplification of *LnocLUC* (GenBank accession number: AY748894), a PCR product 2960 bp in length composed of the entire *luc* gene along with upstream (540 bp) and downstream sequence (439 bp). The entire gene sequence was 1981 bp in length and showed *luc* to be composed of seven exons divided by six small introns (Fig. 1). When compared to the *luc* sequences from two other species of Lampyridae, *Photinus pyralis* (de Wet et al., 1987) and *Luciola lateralis* (Cho et al., 1999) exon/intron sites in *luc* from *L. noctiluca* were found to be completely conserved in both number and position. Comparisons of the predicted mRNA from *LnocLUC* with the *L. noctiluca* luciferase cDNA originally identified by Sala Newby et al. (1996) showed eleven substitutions within the coding region and one substitution in the untranscribed regions. Extensive population variation is evident at this locus as four out of the eleven coding positions were found to be nonsynonymous mutations. Recently the luciferase gene from *L. noctiluca* was reported from a Korean glow-worm specimen (Li et al., 2003). However, it was not possible to make comparisons with the GenBank deposited sequence (AAR20794) as the sequences contained a number of errors both at the gene level and at the protein level.

A search of conserved promoter motifs found in insect genomes revealed a core promoter region (CPR) in the *L. noctiluca* sequence *LnocLUC* 32 bp upstream of the *luc* start codon (Fig. 2). Three TATA boxes and one CAAT box were identified in this *L. noctiluca* CPR. Upstream sequence of *luc* has been characterised from two other species of firefly, *P. pyralis* and *L. lateralis* (deWet et al., 1987; Cho et al., 1999). 500 bp was identified from the former and up to 1980 bp from the latter. An alignment of the first 500 bases for *L. noctiluca*, *P. pyralis* and *L. lateralis* reveals a conserved CPR region present within all three species but with variable TATA box sites. There is a con-



Fig. 3. Amino acid sequence alignment of transposase conceptual translations. *Lampyris noctiluca* (Lnoc) is a composite of three conceptual translations in different open reading frames running from positions 4382 to 5283. *Agam* shows a conceptual translation of a partial mRNA sequence from *Anopheles gambiae* (GenBank accession number: XM554483).

servation of the CAAT box between all three species and one TATA box (TATA box I) between *P. pyralis* and *L. noctiluca* (Fig. 2). A third motif region (TATTTAA), 81 bp upstream of the *L. noctiluca* start codon was conserved in all three species and may function as a TATA box.

Motif search in upstream region of luciferase gene

Genome walking was continued upstream of *luc* until a problematic region was encountered generating multiple sequences with core sequence homology to the previous walk. In total six unambiguous genomic walks were carried out generating approximately 7kb of overlapping upstream sequence. Primers were designed to amplify 6173 bp of upstream region along with the start of *luc* (Fig. 1). A PCR product *LnocLUC5'1* (GenBank accession number: AY753186), 6173 bp in size was amplified generating 5661 bp of contiguous upstream sequence along with the first 513 bp of *luc*. As a result of allelic variation (the 1053 bp of overlapping sequence was found to contain four substitutions) the two genomic fragments *LnocLUC5'1* and *LnocLUC* where not combined and were deposited in GenBank separately.

All six open reading frames of the upstream region were investigated for conserved protein domains using rpsBLAST. Only one domain was identified: a 102 amino acid sequence sharing 29% identity with a consensus DDE superfamily domain (pfam03184) found in a number of endonucleases (Fig. 1). This domain appears related to integrases and transposases, both of which provide efficient DNA transposition. Conceptual translations extending beyond the partial domain, despite having disruptions to the open reading frame, were found to have further identity to a conceptual translation from a partial mRNA sequence in *Anopheles gambiae*, also classified as a member of the DDE superfamily pfam03184 (Fig. 3).

This partial domain in *L. noctiluca*, identified 686 bp upstream of the start codon of *luc*, presents the first evidence for a transposition event possibly occurring in bioluminescence evolution. Furthermore, tBLASTn searches of the 1980 bp upstream sequence of *L. lateralis* revealed regions with identity to transposases found in other insects (results not shown). A lack of strong identity with these transposase sequences combined with the interrupted open reading frame found in *L. noctiluca* suggests these elements are ancient and inactive but their close proximity to the luciferase gene in beetles indicates that they may have initially served to mobilise a luciferase precursor.

Based upon its catalytic properties firefly luciferase can be classified as a member of the adenylate-forming enzyme group which includes amino plant *p*-coumarate:CoA ligases, acyl-tRNA synthetases and long-chain acyl-CoA synthetases (McElroy et al., 1967; Conti et al., 1996). Recently, a novel catalytic function of firefly luciferase was identified - the ability to synthesise a long-chain fatty acyl-CoA from various long-chain fatty acids in the presence of ATP, coenzyme A (CoA) and Mg²⁺ (Oba et al., 2003). In vivo luciferase functions as a mono-oxygenase in the bioluminescent reaction but it is not yet known whether luciferase acts as an acyl-CoA synthetase in the beetle. It seems unlikely due to the concentration of the enzyme in the peroxisomes of the lantern, and its absence elsewhere in the adult (Strause et al., 1979). However, it may be possible that as a result of a gene duplication event from an ancestral AMP-binding enzyme in the ancestral beetle genome, luciferase has evolved a novel activity but retaining its original function in part. Some bioluminescent beetles produce different colours of light emitted from different lanterns on the body. This is particularly apparent in the railroad worm and the click beetles. The railroad-worms emit yellow-green light through eleven

pairs of lateral lanterns along the body and red light through two cephalic lanterns. The click beetles of Jamaica have ventral light organs producing light ranging from yellow-green to orange and their dorsal organs from green to yellow-green (Seliger et al., 1964; Biggley et al., 1967). Different light emission within the same individual is not a result of the substrate luciferin but amino acid differences in the luciferase sequence confirmed by cDNA characterisation (Viviani et al., 1999; Stolz et al., 2003). It is likely that at least two copies of the luciferase gene exist in the genomes of click beetles, railroad worms and possibly lampryids resulting from an ancestral duplication event enabling diverse colour production in localised lanterns. The findings in this paper suggest that this gene duplication may have been augmented by a transposition event facilitated by the putative transposase described upstream of the luciferase gene.

ACKNOWLEDGEMENTS. The author wishes to thank J. Tyler for the provision of specimens and M. Barclay for valuable comments on the manuscript. This work was supported by funding from DSTL, Porton Down, UK.

REFERENCES

- BIGGLEY W.H., LLOYD J.E. & SELIGER H.H. 1967: The spectral distribution of firefly light. II. *J. Gen. Physiol.* **50**: 1681–1692.
- BUCK J.B. 1948: The anatomy and physiology of the light organs in fireflies. *Ann. N.Y. Acad. Sci.* **49**: 397–482.
- CHO K.H., LEE J.S., CHOI Y.D. & BOO K.S. 1999: Structural polymorphism of the luciferase gene in the firefly, *Luciola lateralis*. *Insect Mol. Biol.* **8**: 193–200.
- CHOI Y.S., BAE J.S., LEE K.S., KIM S.R., KIM I., KIM J.G., KIM K.Y., KIM S.E., SUZUKI H., LEE S.M., SOHN H.D. & JIN B.R. 2003: Genomic structure of the luciferase gene and phylogenetic analysis in the Hotaria-group fireflies. *Comp. Biochem. Physiol. (B, Biochem. Mol. Biol.)* **134**: 199–214.
- CONTI E., FRANKS N.P. & BRICK P. 1996: Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes. *Structure* **4**: 287–298.
- DAY J.C., WHITE P.J., SQUIRRELL D.J. & BAILEY M.J. 2002: Origins of bioluminescence in beetles: A phylogenetic perspective. In Stanley P.E. & Kricka L.J. (eds): *Bioluminescence and Chemiluminescence: Progress and Current Applications*. World Scientific Publishing, Singapore, pp. 29–32.
- DELUCA M. & McELROY W.D. 1978: Purification and properties of firefly luciferase. *Methods Enzymol.* **57**: 3–15.
- DE WET J.R., WOOD K.V., HELINSKI D.R. & DELUCA M. 1985: Cloning of firefly luciferase cDNA and the expression of active luciferase in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **82**: 7870–7873.
- KIM J.G., CHOI Y.S., KIM K.Y., BAE J.S., KIM I., SOHN H.D. & JIN B.R. 2005: Genomic structure and phylogenetic analysis of the luciferase gene of the firefly, *Luciola lateralis* (Coleoptera: Lampyridae). *Eur. J. Entomol.* **101**: 1–11.
- KRICKA L.J. 2000: Application of bioluminescence and chemiluminescence in biomedical sciences. *Methods Enzymol.* **305**: 333–345.
- LI J., CHOI Y.S., LEE S.C., LEE S.M., KIM J.G., KIM I., KIM K.Y., SOHN H.D. & JIN B.R. 2003: Genomic structure of the luciferase gene of the firefly, *Lampyrus noctiluca*. *Int. J. Indust. Entomol.* **7**: 155–159.
- McELROY W.D., DELUCA M. & TRAVIS J. 1967: Molecular uniformity in biological catalyses. The enzymes concerned with firefly luciferin, amino acid, and fatty acid utilization are compared. *Science* **157**: 150–160.
- McELROY W.D., SELIGER H.H. & WHITE E.H. 1969: Mechanism of bioluminescence, chemiluminescence and enzyme function

- in the oxidation of firefly luciferin. *Photochem. Photobiol.* **10**: 153–170.
- NEWPORT G. 1857: On the natural history of the glowworm (*Lampyrus noctiluca*). *J. Linn. Soc. Zool.* **1**: 40–71.
- OBA Y., OJIKI M. & INOUE S. 2003: Firefly luciferase is a bifunctional enzyme: ATP-dependent monooxygenase and a long chain fatty acyl-CoA synthetase. *FEBS Letters* **540**: 251–254.
- SALA-NEWBY G.B., THOMSON C.M. & CAMPBELL A.K. 1996: Sequence and biochemical similarities between the luciferases of the glow-worm *Lampyrus noctiluca* and the firefly *Photinus pyralis*. *Biochem. J.* **313**: 761–767.
- SELIGER H.H., BUCK J.B., FASTIE W.G. & McELROY W.D. 1964: The spectral distribution of firefly light. *J. Gen. Physiol.* **48**: 95–104.
- STOLZ U., VELEZ S., WOOD K.V., WOOD M. & FEDER J.L. 2003: Darwinian natural selection for orange bioluminescent color in a Jamaican click beetle. *Proc. Natl Acad. Sci. USA* **100**: 14955–14959.
- STRAUSE L.G., DeLUCA M. & CASE J.F. 1979: Biochemical and morphological changes accompanying light organ development in the firefly, *Photuris pennsylvanica*. *J. Insect Physiol.* **25**: 339–347.
- TYLER J. 1986: The ecology and conservation of the glow worm, *Lampyrus noctiluca* (L.) in Britain. *Atala* **12**: 17–19.
- TYLER J. 1994: *Glow-worms*. Tyler-Scagell, Sevenoaks.
- VIVIANI V.R. 2002: The origin, diversity, and structure function relationships of insect luciferases. *Cell Mol. Life Sci.* **59**: 1833–1850.
- VIVIANI V.R. & BECHARA E.J.V. 1996: Larval *Tenebrio molitor* (Coleoptera: Tenebrionidae) fat body extracts catalyze firefly D-luciferin- and ATP-dependent chemiluminescence: a luciferase-like enzyme. *Photochem. Photobiol.* **63**: 713–718.
- VIVIANI V.R., BECHARA E.J. & OHMIYA Y. 1999: Cloning, sequence analysis, and expression of active *Phrixothrix* railroad-worms luciferases: relationship between bioluminescence spectra and primary structures. *Biochemistry* **38**: 8271–8279.
- WHITE E.H., RAPAPORT E., SELIGER H.H. & HOPKINS T.A. 1971: The chemi- and bioluminescence of firefly luciferin: an efficient chemical production of electronically excited states. *Bioorg. Chem.* **1**: 92–122.

Received July 13, 2005; revised and accepted August 15, 2005

Eur. J. Entomol. **102**: 791–792, 2005
ISSN 1210-5759

BOOK REVIEW

HECKMAN CH.W.: *ENCYCLOPEDIA OF SOUTH AMERICAN AQUATIC INSECTS: EPHEMEROPTERA*. Kluwer Academic Publishers, Dordrecht/Boston/London, 2002, viii + 419 pp. ISBN 1-4020-0775-2. Price USD 197.00.

This volume is the second part of this Encyclopedia dealing with the order mayflies, Ephemeroptera and is characterized by the subtitle “Illustrated Keys to Known Families, Genera and Species in South America”. The book consists of two main chapters. The introductory one deals with the history and main objectives of the project, and includes a five page general section called “An appeal for quality in taxonomic works”. The principal part of this book deals with the determination of both larvae and adults of genera and species of 11 South American mayfly families, of which there are 32 extant families currently recognised, but not all the genera of South America are keyed to species level and some extralimital taxa are included in the

keys. The basic general characteristics of the order are defined on p. 13–26 and in chapters on Morphology, Ecology, Preservation and Examination, Taxonomic problems and Suggestions for Improvement. The keys are followed by a list of about 250 references and an index of scientific names.

Anyone who purchases this book should ignore the keys. They are riddled with errors and contradictions, such as adult Siphonuridae (as *Metamonius*) and Ameletopsidae are grouped with families that do not have veinlets from CuA to the hind margin, an error that any reviewer would instantly find. The characters cited for adult Oligoneuriidae do not apply to South American or any Oligoneuriidae, and this part of the key must have been taken from a North American source when for a brief period the Oligoneuriidae included the Isonychiidae. Heckman keys (falsely) indicate that the following families have a series of veinlets attaching the CuA to the hind margin of the wing: Baetidae, Leptophebiidae, Ephemerellidae, Caenidae and of