## EVOLUTIONARY BIOLOGY

# Resurrecting the ancient glow of the fireflies

Y. Oba[1]*, K. Konishi[1,2], D. Yano[1], H. Shibata[3], D. Kato[3], T. Shirai[4]*

The color of firefly bioluminescence is determined by the structure of luciferase. Firefly luciferase genes have been isolated from more than 30 species, producing light ranging in color from green to orange-yellow. Here, we reconstructed seven ancestral firefly luciferase genes, characterized the enzymatic properties of the recombinant proteins, and determined the crystal structures of the gene from ancestral Lampyridae. Results showed that the synthetic luciferase for the last common firefly ancestor exhibited green light caused by a spatial constraint on the luciferin molecule in enzyme, while fatty acyl-CoA synthetic activity, an original function of firefly luciferase, was diminished in exchange. All known firefly species are bioluminescent in the larvae, with a common ancestor arising approximately 100 million years ago. Combined, our findings propose that, within the mid-Cretaceous forest, the common ancestor of fireflies evolved green light luciferase via trade-off of the original function, which was likely aposematic warning display against nocturnal predation.

## INTRODUCTION

Over the centuries, the bioluminescence of fireflies has attracted much attention as a charming seasonal sight, particularly in Asia, and, recently, as a useful diagnostic tool in the biomedical sciences (1). It is proposed that firefly bioluminescence originated as an aposematic warning display toward predators and later acquired a role in sexual communication for many firefly species (2).

Fireflies belong to the beetle family Lampyridae, which is composed of 10 subfamilies containing around 2200 recognized species across the world (3). Luminescent beetles are additionally found in the families Phengodidae, Rhagophthalmidae, and Elateridae. These four families are members of the superfamily Elateroidea, and it has been considered that Phengodidae and Rhagophthalmidae are sister groups of the Lampyridae and share a common origin of bioluminescence; conversely, bioluminescence in Elateridae evolved independently of the Lampyridae-Phengodidae-Rhagophthalmidae lineage (4).

The molecular system of bioluminescence is shared among these four families; the chemical structure of the luminescent substrate, D-luciferin, is considered identical for all luminescent beetles, and the luminescence reaction is catalyzed by homologous luciferases (>48% amino acid identity) in the presence of $O_2$, adenosine 5′-triphosphate (ATP), and $Mg^{2+}$ (Fig. 1A) (4). Despite the commonality in enzymatic reaction and components, luminescence color can vary widely between species. The European glowworm *Lampyris noctiluca*, for example, emits green light, the North American Big Dipper firefly *Photinus pyralis* yellow-green light, and the Japanese lesser firefly *Luciola parvula* orange-yellow light. The differences in luminescence color are considered to be the consequence of evolutionary strategies for warning predators and attracting mating partners more effectively (2, 5).

Fossil records of fireflies are limited, but a single adult male firefly was mined from Burmese amber dating back to 100 million years

(Ma) ago, which exhibited an obvious photophore structure on the abdominal segment (6). This specimen strongly suggested that the ancestral firefly of the Cretaceous period was already bioluminescent and used light emission for some purpose. Naturally, the color of ancestral bioluminescence is not evident from fossil records, however, making it difficult to predict the original function of the firefly light.

The colors of light in fireflies are regulated by the luciferase structure (7) and, probably, not by either the luciferin molecule or the effects of color filters (like the blue transmission filter of the hatchetfish photophore) (8, 9) and not by fluorescent substances (like *Aequorea* green fluorescent protein in jellyfish). For this reason, in vivo firefly luminescence colors match principally those of the in vitro luminescent reaction of the luciferase with D-luciferin (9–11). Thus, the evolutionary history of firefly bioluminescence is traceable by using the sequence data of extant luciferases.

Now, two types of firefly luciferase genes, Luc1 type and Luc2 type, have been isolated from more than 30 species in various subfamilies, most of which are Luc1 type genes; but recently, Luc2 type genes have also been isolated from several species of two major subfamilies, Lampyrinae and Luciolinae. The gene duplication occurred at the basal position of the lineage in Lampyridae, thus suggesting that probably all fireflies have both Luc1 and Luc2 (4, 10, 11). Gene expression profiles have suggested their subfunctionalization—that is, extant Luc1-type luciferase is responsible for luminescence in green-yellow of the lantern in larvae, pupae, and adults, while extant Luc2-type luciferase is responsible for the dim glow in green of oviposited eggs, the pupal body, and the ovaries (9, 11). On the basis of these premises, we recreated putative ancestral firefly luciferases by predicting their amino acid sequences with the maximum likelihood method of ancestral state reconstruction (12) and experimentally characterized the enzymatic properties including the luminescence colors.

## RESULTS

### Ancestral sequence reconstruction of firefly luciferases

To reconstruct the ancestral states, the molecular phylogenetic tree was constructed for a total of 53 amino acid sequences of beetle luciferases and close homologs (black line cladogram in Fig. 1B). The tree topology was mostly consistent with that of the species tree (red and blue line cladograms in Fig. 1B) based on the total evidence approach (2, 3). To trace back the evolution of firefly luciferase

[1]Department of Environmental Biology, Chubu University, Kasugai 487-8501, Japan. [2]Graduate School of Bioagricultural Sciences, Nagoya University, Nagoya 464-8601, Japan. [3]Graduate School of Science and Engineering, Kagoshima University, Kagoshima 890-0065, Japan. [4]Department of Bioscience, Nagahama Institute of Bio-Science and Technology, Nagahama 526-0829, Japan.
*Corresponding author. Email: yoba@isc.chubu.ac.jp (Y.O.); t_shirai@nagahama-i-bio.ac.jp (T.S.)
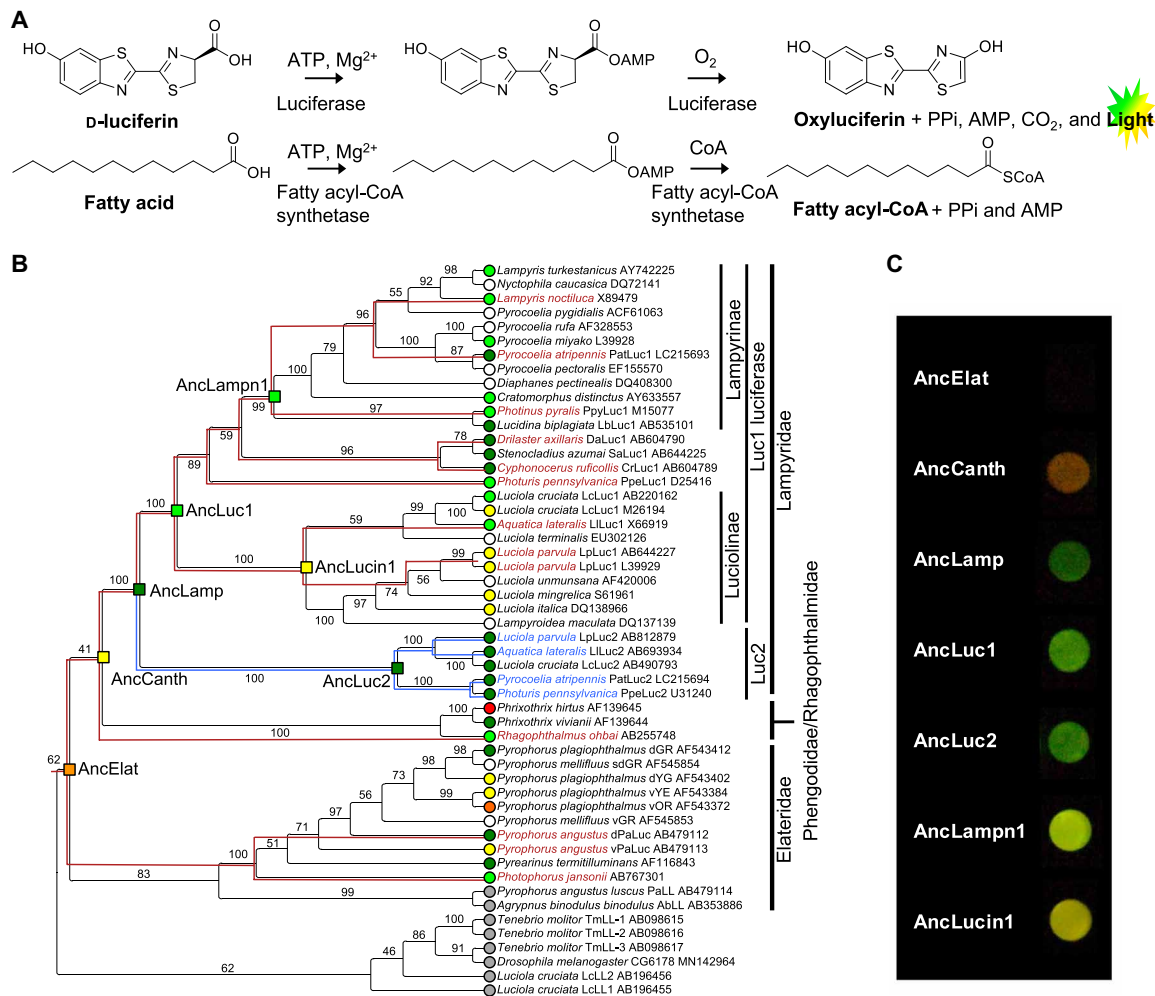
**Fig. 1. Firefly bioluminescence and color evolution.** (**A**) Coleopteran bioluminescence reaction. (**B**) Molecular phylogeny of luciferases and related enzymes. The leaf nodes are labeled with species name, protein name, and GenBank accession number. Branches are labeled with bootstrap probability (1000 reconstructions). The resurrected ancestral nodes are shown as a square. The leaf nodes are indicated with in vitro luminescent colors (green, yellow-green, yellow, orange, or red) judged by the luminescence maximum values in references (table S3). The resurrected ancestral nodes are indicated with in vitro luminescence colors judged by the combination of luminescence maximum (Table 1) and perceived coloration. Luciferases without spectral data and fatty acyl-CoA synthetase (nonluciferase) are denoted with white and gray circles, respectively. The 18*S* rRNA–based species trees (fig. S6) are superimposed onto the subtrees of Luc1 type (red lines) and Luc2 type (blue lines) to confirm consistency between the gene/protein phylogeny and the species phylogeny. (**C**) Photographs of the luminescence of seven ancestral luciferases in 96-well plate. The camera exposure time for each well is uneven to avoid the changes in coloration by overexposure. Luminescence of AncElat was not photographed even by longer exposure. Luminescence of AncCanth was categorized as yellow on the basis of λmax value but appears in orange on the photograph probably because of its broad spectrum.

properties, we targeted seven key ancestral genes encoding the ancestral luciferases of Elateroidea (AncElat); the Lampyridae-Phengodidae-Rhagophthalmidae lineage (AncCanth); Lampyridae (AncLamp), Luc1-type luciferase of Lampyridae (AncLuc1); Luc2-type luciferase of Lampyridae (AncLuc2); Luc1-type luciferase in Lampyrinae, a major subfamily of Lampyridae (AncLampn1); and Luc1-type luciferase of Luciolinae, another major Lampyridae subfamily (AncLucin1) (Fig. 1B). Monophyly of all ancestral nodes are supported with >99% bootstrap values, except for AncElat (62%) and AncCanth (41%).

The ancestral luciferase sequences were inferred on the basis of the molecular phylogeny of the extant amino acid sequences (fig. S1). The average posterior probability of the amino acid residues in the seven reconstructed ancestral luciferases were sufficiently high, ranging from 0.984 to 0.915 (Table 1 and fig. S2). Although 2 to 2 l sites

were estimated rather uncertainly, most of these sites distributed distant from the active sites (7) of luciferase with a few exceptions (fig. S3 and table S1).

## Luminescence properties of ancestral firefly luciferases
The seven codon-optimized ancestral gene sequences were synthesized and cloned into an expression vector. The recombinant proteins expressed in *Escherichia coli* were purified using a cobalt chelating column (fig. S4). The luminescence activities of the purified recombinant proteins were measured in the presence of D-luciferin, ATP, and Mg$^{2+}$, showing a similar intensity to those in the extant luciferases (*Luciola cruciata* LcLuc1, LcLuc2, and *P. pyralis* PpyLuc1) except for AncCanth and AncElat, for which the luminescence activities were lesser and at trace level, respectively (Fig. 2A).
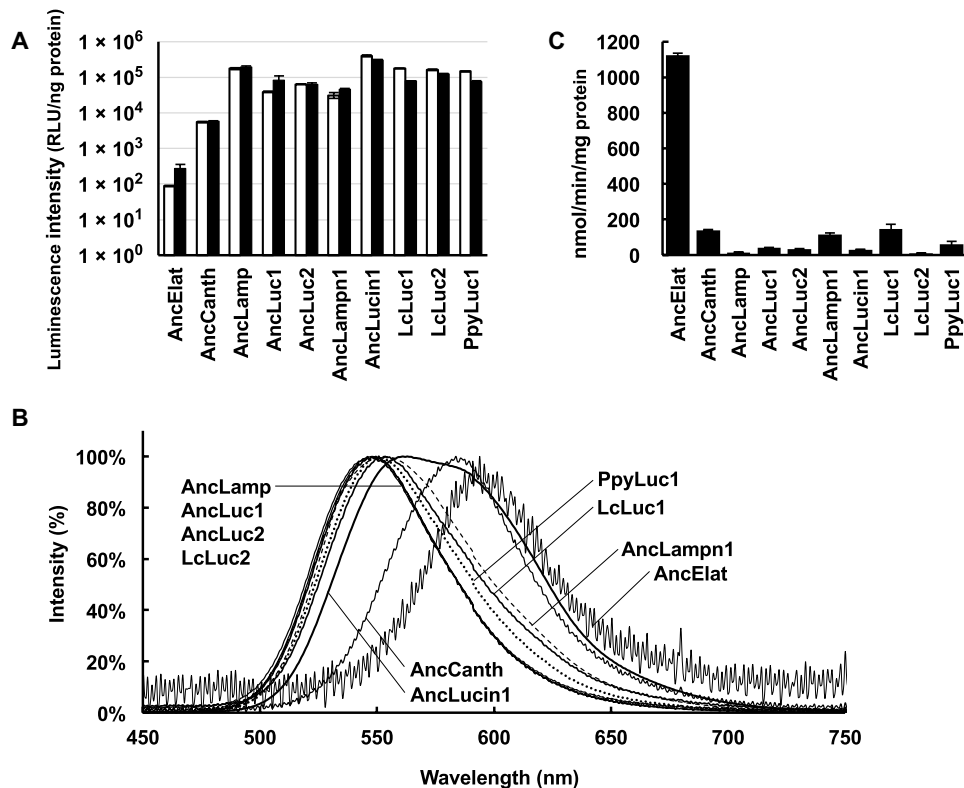
**Fig. 2. Biochemical properties of ancestral luciferases.** (**A**) Integration of luminescence intensities from 2 to 32 s. The values are shown by relative light unit per nanogram of protein (RLU/ng). White and black bars represent the mean with standard deviation values at 50 and 200 µM of D-luciferin, respectively ($n = 3$). (**B**) Luminescent spectra of seven recombinant ancestral luciferases and three extant firefly luciferases. (**C**) ACS activities of seven ancestral luciferases and three extant firefly luciferases to lauric acid are shown. Bars represent SE of the means ($n = 3$ to 6).

**Table 1. Reconstructed ancestral luciferases.**

| Ancestral luciferase | Position | Median divergence age (Ma ago) (95% HPD) | Av. P.P.* | λmax, FWHM (nm)† and perceived color |
|---|---|---|---|---|
| AncElat | Elateroidea ancestor | 115.85 (84.82–145.26) | 0.915 | 594, 61 (red) |
| AncCanth | Lampyridae-Phengodidae-Rhagophthalmidae ancestor | 102.55 (71.06–132.83) | 0.915 | 584, 67 (orange) |
| AncLamp | Lampyridae ancestor | not applicable | 0.935 | 548, 65 (green) |
| AncLuc1 | Luc1 ancestor | 69.59 (44.57–96.19) | 0.938 | 550, 64 (yellow-green) |
| AncLuc2 | Luc2 ancestor | 69.59 (44.57–96.19) | 0.949 | 548, 65 (green) |
| AncLampn1 | Lampyrinae Luc1 ancestor | 26.23 (22.91–66.83) | 0.984 | 553, 79 (yellow-green) |
| AncLucin1 | Luciolinae Luc1 ancestor | 12.18 (0.50–35.09) | 0.968 | 563, 92 (yellow) |

*Av. P.P., the average posterior probability of the amino acid residues protein at pH 8.0. FWHM, full width at half maximum (nanometers).  †λmax (nanometers), luminescence spectral maximum of the purified recombinant

The color of light produced by fireflies can be approximated from the λmax value of the spectrum, although not defined because of broad spectral bandwidth. In this study, we provisionally report the coloration of the light produced by luciferin-luciferase reaction as a visual description "green" for λmax 520 to 549 nm, "yellow-green" for λmax 550 to 559 nm, "yellow" for λmax 560 to 584 nm, and "orange" for λmax 585 to 619 nm, which are approximately aligned with our visual perception.

Measurement of the luminescence spectra for seven ancestors showed that AncElat and AncCanth emitted in the orange-red region (λmax 594 and 584 nm respectively); AncLuc1, AncLuc2, and AncLamp emitted in the green region (λmax ≤550 nm), which was most hypsochromic shifted (the peak shifted to shorter wavelength) in seven ancestors; and AncLampn1 and AncLucin1 showed a yellow-green to yellow (λmax 553 and 563 nm), respectively, (Table 1 and Figs. 1C and 2B).

### Acyl-CoA synthetic activities of ancestral firefly luciferases

Firefly luciferases are bifunctional enzymes; in addition to luciferase (luminescence) activity, they have a "promiscuous" acyl-CoA synthetic (ACS) activity to various fatty acids (Fig. 1A) (13–15). Consequently, it has been hypothesized that the firefly luciferases originated from a fatty acyl-CoA synthetase (15). We examined ACS activities of the seven ancestral luciferases using lauric acid as a substrate.

Results showed that ACS activity of AncElat was an order of magnitude higher than other ancestral and extant luciferases, demonstrating specific activity of protein (1126.1 ± 9.6 nmol/min per milligram) (Fig. 2C). The activities of other ancestral luciferases were rather low as ranging from 140.4 ± 3.2 (AncCanth) to 15.1 ± 2.7 (AncLamp), and these values were comparable with that of the extant luciferases ranging from 145.0 ± 28.0 (LcLuc1) to 9.7 ± 1.6 (LcLuc2).

### Crystal structure of ancestral firefly luciferases

The ancestral firefly luciferase AncLamp was cocrystallized with the reaction-intermediate analog $5'$-$O$-[$N$-(dehydroluciferyl)-sulfamoyl]adenosine (DLSA), and the crystal structure [Protein Data Bank (PDB) ID 6K4C] was determined to a 2.1-Å resolution (Fig. 3A and table S2). The overall structure of AncLamp was well conserved with LcLuc1 (PDB ID 2D1S), which shows 79% amino acid sequence identity and 1.64-Å root mean square deviation for 535 Cα atoms with AncLamp (7).

The DLSA molecule was in close contact (at least one atom of amino acid residue existed less than 3.5 Å from DLSA atoms) with His[244], Phe[246], Thr[250], Ser[313], Pro[317], Gly[338], Tyr[339], Gly[340], Leu[341], Thr[342], Thr[361], Asp[421], Ile[433], Arg[436], and Lys[528] in the AncLamp-DLSA complex structure. These substrate-binding residues are generally conserved with LcLuc1. The conformations of bound DLSA molecule were nearly identical between AncLamp and LcLuc1 as showing 0.38-Å root mean square deviation for 40 DLSA atoms in the superposed structures (Fig. 3B).

A noticeable difference in DLSA interaction between AncLamp and LcLuc1 was observed in proximity of the residues, namely, Ile[236] (Val[239] in LcLuc1), Ile[240] (Val[243]), and Leu[285] (Ile[288]) (Fig. 3C). Mainly because of the packing of Cδ1 atom of Ile[240] and Cδ2 atom of Leu[285], the phenyl ring of Phe[246] was slightly extruded toward the substrate-binding cavity. The distance from the Cε1 atom of Phe[246] (Phe[249] of LcLuc1) residue to the thiazole plane of DLSA (defined by atoms C8-S9-C4) was 3.18 and 3.43 Å in the AncLamp and LcLuc1 structures, respectively (Fig. 3D). Consequently, the space for luciferyl moiety appeared to be narrower in the AncLamp structure. The cavity volume for luciferyl moiety was 599 Å[3] in AncLamp, which was notably smaller than 719 Å[3] in LcLuc1.

The ancestral firefly luciferase AncLamp was also cocrystallized with genuine firefly luminescence substrates, namely, ATP and D-luciferin, and the crystal structure was determined to a 1.7-Å resolution (PDB
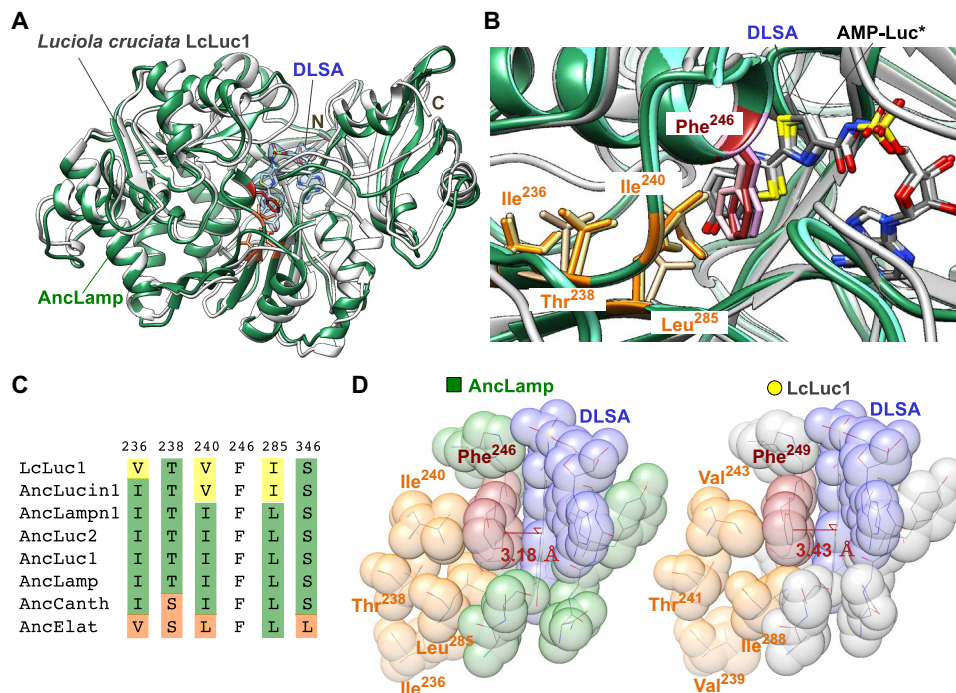


**Fig. 3. Structure of AncLamp.** (**A**) The structure of AncLamp-DLSA complex (green) is superposed on that of *Luciola cruciata* LcLuc1-DLSA complex (white). DLSA is represented as stick model. The all ligand omitting *Fo-Fc* map (contoured at 6 σ) is shown in blue. The positions of N and C termini are indicated. (**B**) Detail of the substrate-binding site of superposed AncLamp-DLSA (green), LcLuc1-DLSA (white), and AncLamp-substrates (D-luciferin/ATP) (light green) complex structures. The residues differing between AncLamp and LcLuc1, Ile[236] (Val[239] in LcLuc1), Ile[240] (Val[243]), and Leu[285] (Ile[288]), and those in close contact with these residues, Phe[246] (Phe[249]) and Thr[238] (Thr[241]), are shown in stick models. The molecules of the DLSA and probable intermediate, luciferyl-AMP (AMP-Luc*), are shown as stick models colored by element (carbon atoms are colored in white, gray, and light gray for AncLamp-DLSA, LcLuc1-DLSA, and AncLamp-substrates complex structures, respectively). (**C**) Amino acids of the ancestral luciferases and LcLuc1 at the sites discussed in the text. The residue numbers of AncLamp are indicated above the alignment. See fig. S1 for original alignment. (**D**) Close-up view of the substrate-binding site of AncLamp (left) and LcLuc1 (right) in van der Waals model. The different residues between two proteins, i.e., Ile[236] (Val[239] in LcLuc1), Ile[240] (Val[243]), and Leu[285] (Ile[288]), and those in close contact, Phe[246] (Phe[249]) and Thr[238] (Thr[241]), are shown. The distances from Cε1 atoms of Phe[246] (Phe[249]) to DLSA are indicated in red.

ID 6K4D). Although the protein structure was almost identical to that of the AncLamp-DLSA complex, the electron density map at the substrate-binding site demonstrated an intriguing feature, suggesting the presence of the reaction intermediate, most probably a luciferyl–adenosine monophosphate (luciferyl-AMP). The conformations of the substrate-interacting residues were similar between the complex structures DLSA and luciferyl-AMP complex. The conformation of luciferyl-AMP was also very close to that of DLSA as showing 0.42 Å root mean square deviation for 40 atoms (Fig. 3B). The cavity volume for luciferyl moiety was 558 Å$^3$. Thus, the space elimination for luciferyl moiety, which was observed for the intermediate analog (DLSA) complex, appeared to be consistent for a genuine intermediate molecule.

In the crystal structure of AncLamp with ATP and D-luciferin, substantial residual density was observed after building the luciferyl-AMP model, and it was interpreted to be unreacted D-luciferin. These residual densities fit well to sulfur or oxygen atoms when unreacted D-luciferin was built into the substrate-binding cavity (fig. S5B). Unexpectedly, the unreacted D-luciferin was trapped in the substrate-binding cavity upside-down for reaction (the reactive carboxylate moiety was pointing to the bottom of cavity), apparently showing unproductive configuration and suggesting an immature specificity for D-luciferin. No structure of luciferase complexed with D-luciferin or with D-luciferin and ATP has been reported so far, and this was the first observation of the substrate configuration for luciferase. Therefore, significance of the observed D-luciferin configuration for the evolution of acyl-CoA synthetase/luciferase activity was inconclusive at this point of time.

## Molecular clock analyses

For dating the ancestral luciferases, a phylogenetic tree of Elateriformia species was constructed on the basis of the 18S ribosomal RNA (rRNA) sequences. The geological ages of the ancestral luciferases (corresponding to the speciation nodes), except for AncLamp (common ancestor of duplicated genes of Luc1 and Luc2-type luciferases), were estimated with the Bayesian Markov chain Monte Carlo (MCMC) method by referring to the fossil date 152 ± 10 Ma ago of the most recent common ancestor of Elateroidea as the prior probability distribution (Table 1 and Fig. 4).

The oldest ancestral luciferase homolog of the present study, AncElat, was dated back to 115.85 Ma ago with 95% highest posterior density (HPD) of 84.82 to 145.26 Ma ago. As demonstrated in the luminescence activity assay, AncElat was nonluminescent. It suggested that the common ancestor of Elateroidea (indicated as ConElat, the common ancestor bearing AncElat gene, in Fig. 4) did not use bioluminescence for any purpose until the emergence of the Lampyridae-Phengodidae-Rhagophthalmidae lineage (ConCanth).

AncCanth, which demonstrated lesser luminescence activity to the extant luciferases, belonged to the last common ancestor of Lampyridae, Phengodidae, and Rhagophthalmidae. Thus, it might define the oldest date of bioluminescence activities in the Lampyridae-Phengodidae-Rhagophthalmidae lineage. Molecular clock analyses suggested that the Lampyridae originated in the mid-Cretaceous (16). This is consistent with the estimation of geological age of ConCanth of the present study, which suggested 102.55 (95% HPD of 71.06 to 132.83) Ma ago (Fig. 4). This dating was also consistent approximately to the oldest fossil record of ancient firefly having a photophore structure on the abdominal segment, which dated back to around 100 Ma ago (6).

AncLuc1 and AncLuc2 belonged to the last common ancestor of Lampyridae (ConLuc1/2) at 69.59 (95% HPD of 44.57 to 96.19) Ma ago. The duplication between Luc1-type and Luc2-type genes should occur between ConCanth and ConLuc1/2, where the last common ancestral gene encoded AncLamp. The results suggested it was from late-Cretaceous to early-Paleogene periods (Fig. 4).

## DISCUSSION
### Origin of beetle bioluminescence
Substantial evidence for luminescent color evolution in fireflies is provided in this study. The luminescence activity in five more-recent ancestral luciferases (AncLamp, AncLuc1, AncLuc2, AncLampn1, and AncLucin1) was comparable to that of extant luciferases, suggesting that the bioluminescence in the common ancestor of Lampyridae already has acquired some biological function(s).

Interpretation of the substantial but less active luminescence activity in AncCanth will be controversial. The most parsimonious
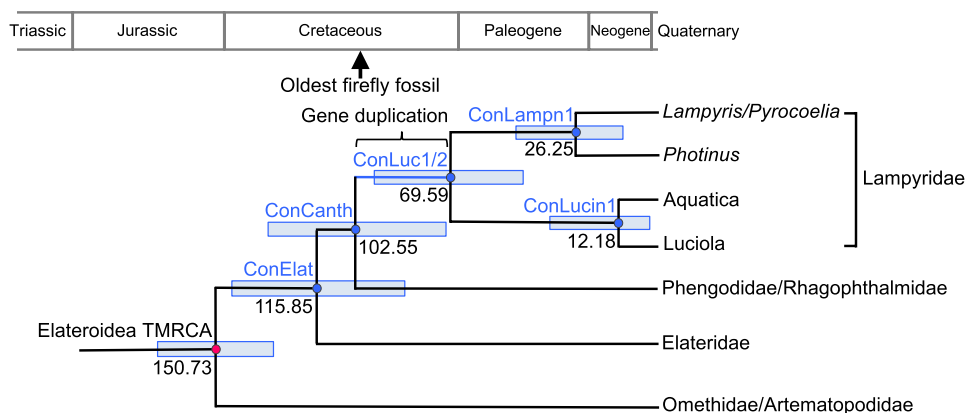


**Fig. 4. Dating of ancestral species.** Geological dating of the ancestral species (concestors) (50) bearing the ancestral luciferases based on the molecular clock analysis of 18S rRNA genes of Elateriformia species. The nodes corresponding to the hypothesized concestors, which genomes encoded AncElat, AncCanth, AncLampn1, and AncLucin1 are indicated as ConElat, ConCanth, ConLampn1, and ConLucin1, respectively. ConLuc1/2 is the concestor of Lampyridae, which had the duplicated genes of AncLuc1 and AncLuc2 as the first time and AncLamp as the last time (Fig. 1B). The median of estimated geological age and 95% HPD (light blue bar) from the MCMC estimations (fig. S6) are indicted on each node. The geohistorical positions of the oldest firefly fossil with abdominal photophore structure (6) and gene duplication between Luc1-type and Luc2-type genes are indicated on the phylogeny or geological time scale.

reconstruction of the ancestral state, based on molecular phylogenetic analyses, indicated that a common ancestor of the Lampyridae-Phengodidae-Rhagophthalmidae lineage was "bioluminescent" (4). Bioluminescent courtship is observed only in some species within the Lampyridae and Rhagophthalmidae, suggesting that the original function of luminescence in the common ancestor of the Lampyridae-Phengodidae-Rhagophthalmidae lineage was aposematic display from nocturnal predation, and that the mating function in bioluminescence evolved independently in both Lampyridae and Rhagophthalmidae (2). Thus, the substantial but less active bioluminescence of yellow-orange color in AncCanth may represent the ancestral weak luminescence. Aposematic luminescence in eggs and pupae in extant fireflies is very weak compared to the lantern luminescence in adult fireflies, which functions in mating (9, 17). On the other hand, inaccurate prediction of the ancestral sequence in AncCanth might be also the reason for its weaker luminescence activity, because of the relatively higher ambiguity of older sequence information (figs. S2 and S3). Although the main focus of this paper is the bioluminescent state of a common ancestor of Lampyridae (AncLamp), further research is necessary to understand the state of bioluminescence of the common ancestor in the Lampyridae-Phengodidae-Rhagophthalmidae lineage using the recent transcriptome data of Phengodidae (18).

A substantial lack of luminescence activity in the oldest AncElat confirms the hypothesis based on genome analysis (4), suggesting that a common ancestor of the Elateroidea was nonluminescent and that their bioluminescence abilities have evolved in parallel in the Lampyridae-Phengodidae-Rhagophthalmidae lineage and Elateridae (Fig. 1, B and C). The lack of luminescence activity in AncElat can be rationalized by the amino acid sequences. Amino acid sequence comparisons between ancestral/extant luciferases and a related fatty acyl-CoA synthetase highlighted two sites of interest, namely, Gly[315] and Ser[346]. They were the only residues in 12 substrate-binding sites (fig. S1, outlined in blue) consistent with luminous active/inactive features, i.e., conserved within luminescent active or inactive proteins but between them. On the basis of the crystal structure of AncLamp, Gly[315] and Ser[346] are responsible for the binding to AMP and luciferyl moieties, respectively. In the latter position, polar amino acids (serine or cysteine) are used in all extant luciferases and "luminescent active" ancestral luciferases, and the side chain can interact with the luciferin molecule via hydrogen bonding (fig. S5A). On the other hand, nonpolar leucine is used in luminescent inactive AncElat and AbLL, and the side chain probably cannot contribute to a polar interaction with luciferin (19). This is in congruent to that a single mutation of leucine at this position to serine recovered the substantial luminescence activity in both AbLL and CG6178, a fatty acyl-CoA synthetase in *Drosophila* (20). In relation to luminescence activity and this residue, it has been reported that the gene product of a luciferase-like gene from the tenebrionid beetle *Zophobas morio*, Zop, showed substantial luminescence activity with D-luciferin, and that position 329 in Zop (position 346 in AncLamp) was leucine (21). Luminescence is absent in the Leu[329]Ser mutant of Zop (21). Zop is also an acyl-CoA synthetase, but beetle luciferases are more closely related to AbLL and CG6178 than Zop. This implicates the multiplicity of possible pathways from acyl-CoA synthetase to luciferase.

## Trade-off evolution of acyl-CoA synthetase and luciferase
Our previous reports have suggested that firefly luciferases and their orthologs in beetles and *Drosophila* have ACS activity, and the most preferable substrate was lauric acid (14). On the basis of these results,

we have hypothesized that beetle luciferase originated from a fatty acyl-CoA synthetase (presumably, medium-chain fatty acyl-CoA synthetase) (4, 14, 15). More recently, enzymatic promiscuity of beetle luciferase to various carboxylate substrate has been demonstrated, suggesting the evolutionary origin of beetle luciferase from an acyl-CoA synthetase with broad substrate specificity including long-chain fatty acids (15). The results of the present study showed that oldest AncElat exhibited about 10 times higher activity to those of other ancestral luciferases and extant luciferases. This is, in notable contrast, to the luminescence activity profiles and agreement with our "ACS origin" hypothesis of beetle luciferase (Fig. 2, A and C). The presence of ACS activity in AncElat may also be rationalized by the amino acid substitution. It has been reported that fatty acids are also competitive inhibitor for firefly luciferase at higher concentration (22), indicating that both luciferin and fatty acid share the same substrate-binding site. It is expected that the nonpolar leucine at the position 346 in AncElat will favorably accept the aliphatic substrate into the binding site. Consequently, all AbLL mutants having leucine at this position exhibited higher ACS activities than those having serine (19).

This observation also fits the evolutionary "trade-off" model of enzyme neofunctionalization, in which the original catalytic function of an enzyme is decreased during the process of acquiring a new function (23). We previously suggested that the trade-off model fits the evolution of beetle luciferase by site-directed mutagenesis studies of beetle fatty acyl-CoA synthetase (19). We constructed all combinations of the mutants for three amino acids near the active site in AbLL, an ACS of a nonluminous click beetle, and measured both luciferase and ACS activities. The results showed negative trade-offs of these activities among mutants, that is, mutants having higher ACS activity exhibited lower luciferase activity, and vice versa (19). Since the activities of both the luciferase reaction and fatty ACS reaction are the sum of complex kinetics by multiple reaction steps, further studies on biochemical kinetics for each step are necessary to fully understand the evolution of firefly luciferase from acyl-CoA synthetase.

Recent firefly genome analysis showed that firefly luciferase evolved by tandem gene duplications of ACS and subsequent acquisition of the luciferase activity in a redundant gene (4). Probably, the multiple gene duplications of ACS occurred in an ancestral nonbioluminescent beetle leading to the evolution of a new function "luminescence activity" without suffering any adaptive disadvantages in organisms.

## Evolution of luminescence color
Luminescence spectra and photography of the ancestral luciferases suggested that the last common ancestor of the Lampyridae-Phengodidae-Rhagophthalmidae lineage emitted yellow (but rather orange by photograph, Fig. 1C) light (AncCanth); the last common ancestor of Lampyridae emitted green luminescence (AncLamp); and after gene duplication, bathochromic shifting (shifting the peak to longer wavelength) occurred in Luc1-type luciferase from green to yellow-green (AncLuc1 and AncLampn1), yellow (AncLucin1), and orange-yellow during the evolution of the lineages in Lampyrinae and Luciolinae, while Luc2-type luciferase conserved the ancestral green light (AncLuc2) in the extant species (Fig. 1B) (10, 11). The bathochromic luminescence of AncCanth fits the hypothesis that beetle "protoluciferases," having a less developed substrate-binding site environment, emitted red light (24, 25). The present study provides experimental evidence that the first luciferase "prototype" emitted bathochromic yellow-orange color.

Subsequently, the luminescent color of the first luciferase has been subjected to adaptive evolution. Green luminescence in AncLamp fits the current hypotheses of firefly biochemists and ecologists, which assumes the luminescence color of the common ancestor to be green (5, 26–28). Green light emission has been observed in the larval stages of the extant diurnal fireflies and also the egg and pupal stages of nocturnal fireflies. Green luminescence is actually common in a number of terrestrial luminous organisms, such as luminous mushrooms and millipedes (29, 30). As fireflies and millipedes are known to have distasteful toxins, they are considered to use luminescence as a warning signal (31, 32). The presence of nocturnal predators maximally sensitive to green light by may have given rise to aposematic displays conferring a selective advantage to fireflies emitting at night (27, 28). The presented results demonstrated that the green luminescent from AncLamp, of the last common ancestor of fireflies Con-Luc1/2, already hypsochromically shifted nearly to known limit for extant luciferases. Green luminescence may also have advantage for mate recognition at the early day of Lampyridae, because it can be postulated that "visual spectral sensitivity in nonbioluminescent ancestors of fireflies was conditioned by the predominance of green light reflected from foliage" by comparison of the extant firefly luminescence (5). A fossil male adult firefly in 100 Ma ago amber already had abdominal photophore in which "the location and size of the photic organ resemble those in many modern Lampyridae" (6), implicating the function of sexual communication.

To elucidate the molecular basis of hypsochromic shifted luminescence in AncLamp, the crystal structure was determined and compared with the structure of extant LcLuc1. The substrate-binding site of AncLamp is generally conserved with LcLuc1, but positions 240 and 285 have Ile and Leu substitutions, respectively (Fig. 3C). These substitutions appeared to affect the conformation of conserved Phe[246] and eliminate the space for accommodating the luciferyl moiety of D-luciferin from 719 to 558 and 599 Å$^3$. The distance from the Cε1 atom of Phe residue to the thiazole plane of DLSA is 3.43 and 3.18 Å in the LcLuc1 and AncLamp structures, respectively (Fig. 3D).

The previous structural analysis of LcLuc1 revealed that the steric constraint on luciferyl moiety was one of the most significant factors for luminescence wavelength; the Ser[286]Asn mutant of this protein caused a bathochromic shift in wavelength, and the crystal structure indicted that the mutation induced a conformation change in a substrate-binding residue, Ile[288], making a space for thermal relaxation of luciferyl moiety (7). An energy loss through thermal motion of luciferyl moiety was thought to responsible for a low quantum yield and bathochromic shift of luminescence wavelength. Thus, the hypsochromic shifted luminescence in AncLamp can also be explained by the spatial constraint on the luciferyl moiety via largely different atomic interactions as observed in the structure. It suggests that the spatial constraint is naturally occurring and the general strategy for luminescent hypsochromic shift in luciferases.

Hydrogen bonding patterns on DLSA in the crystal structures of AncLamp and LcLuc1 were also compared to assess the possible cause of the hypsochromic shifted luminescence (33). However, the luciferyl moiety of the DLSA formed only few hydrogen bonds through O10 (hydroxyl group of benzothiazole ring), N7 (amine group of thiazole ring; fig. S5A), and O39 (amide group between luciferyl moiety and AMP moiety) atoms, and they were mostly conserved between AncLamp and LcLuc1 (table S4). Only exception was that the last one at the peripheral of the luciferyl moiety was mediated by a water molecule in LcLuc1. This suggested that the hydrogen bond-

ing pattern has been conservative for the core luciferyl moiety, and it would not be a promising cause of the hypsochromic shift of luminescence in this case.

The relationship between luminescence color of firefly luciferase and the amino acid sequence has been extensively studied by site-directed and random mutagenesis, and most of these mutants exhibited bathochromic-shifted luminescence (21, 34). The reported mutations (fig. S1 and table S5) were comprehensively compared with the amino acid differences between ancestral luciferases, and it was found that mutation studies reported by Kajiyama and Nakano (35) and Branchini et al. (34) coincided with our current results in terms of the site position, substitution pattern, and the spectral shift phenotype (Fig. 3, C and D). The position 236 (in AncLamp) was Val in AncElat and Ile in all other ancestral luciferases; thus, the spectral hypsochromic shift in AncElat-AncLamp was consistent to the result of Val[239]Ile mutant in Luciola cruciata (LcLuc1), exhibiting hypsochromic shift from λmax 562 to 558 nm at pH 7.8 (35). The position 240 was Leu in AncElat, Val in AncLucin1, and Ile in all other ancestral luciferases; thus, the spectral bathochromic shift in AncLamp-AncLucin1 was consistent to the result of Val[241]Ile mutant in P. pyralis (PpyLuc1), exhibiting hypsochromic shift from λmax 557 to 555 nm at pH 7.8, from 562 to 557 nm at pH 7.0, and from 613 to 609 nm at pH 6.0 (34). Furthermore, this position in extant LcLuc1 was Val; thus, bathochromic shift in AncLamp-LcLuc1 was also consistent. These two positions were exactly located behind Phe[246], which faces to luciferin moiety in the crystal structure of AncLamp (Fig. 3D), demonstrating the involvement in evolution of luminescence color.

On the basis of the luminescence colors in AncLuc1, AncLuc2, AncLampn1, and AncLucin1, we proposed an evolutionary diversification scenario of bioluminescence in Lampyridae. After gene duplication event at about 70 Ma ago (Fig. 4), spatio-temporal subfunctionalization has happened; Luc1-type luciferase is used for the luminescence of lanterns in larval, pupal, and adult stages, while Luc2-type luciferase is used for the glow of eggs and pupal body (9, 11, 17). Green luminescence in AncLuc1 and AncLuc2 suggest that both Luc1 and Luc2 initially emitted in the same ancestral green spectrum, as also seen in the extant Pyrocoelia atripennis (9). In this regard, Luc2-type luciferases should be the authentic orthologs of AncLamp. After that, in both lineages of Lampyrinae and Luciolinae, some species independently began to adapt to a twilight environment for mating behavior. Now, the crepuscular fireflies needed to differentiate mate luminescence from the background noise of green foliage wherein lantern emissions became more bathochromic shifted toward yellow (AncLampn1 and AncLucin1) as seen in extant P. pyralis and L. parvula (5). The spectral sensitivities of crepuscular fireflies match their own yellowish light spectra to receive the conspecific luminescence effectively (5, 36). In contrast to the colorful evolution of Luc1-type luciferase, extant Luc2-type luciferase conserved the original green luminescence of AncLuc2 for warning displays in egg and pupal stages in both diurnal and nocturnal fireflies, because the gene duplication between Luc1 type and Luc2 type enabled a role differentiation between the paralogs.

In conclusion, we predict that the most ancestral fireflies emitted green light for aposematic display purposes, and some species evolved to use the light as a secondary role for complex sexual communications, shifting to a more yellowish light (Fig. 5). We believe that our results demonstrated the practical usefulness of the ancestral gene reconstruction methods (12), successfully reenacting visual scenarios of the lost world.
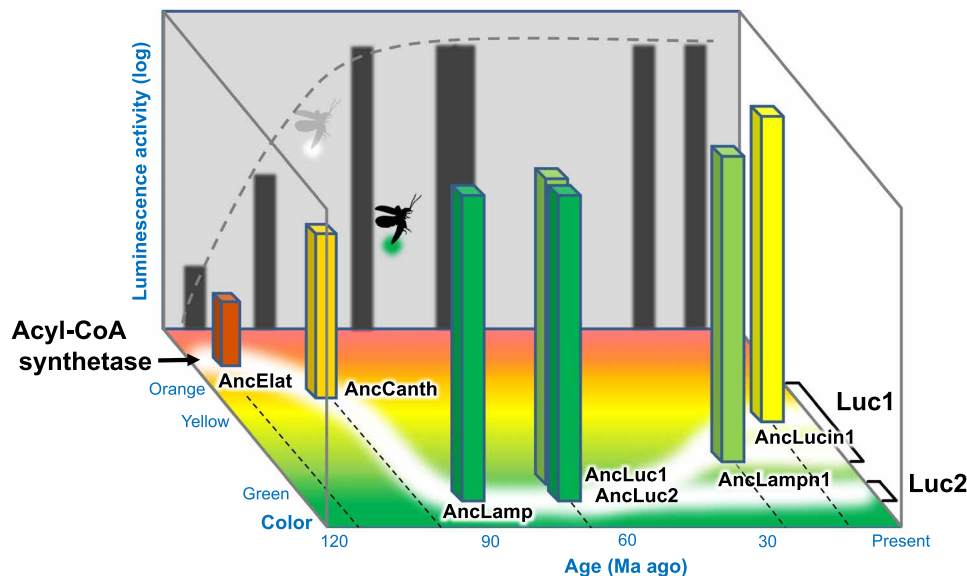
**Fig. 5. Scheme of the evolution of firefly's bioluminescence.** Horizontal and vertical axes represent the approximate age and luminescence activity, respectively. *Z* axis shows the color of ancestral luciferases.

## MATERIALS AND METHODS

### Ancestral luciferase sequence prediction

The initial set of amino acid sequences was retrieved from GenBank, Refseq, and UniProt databases. The sequences were aligned by using MAFFT v7.309 and manually refined with the XCED v8a.3.93 program (*37*). The final alignment contained a total of 52 sequences (Fig. 1B and table S3). The topology of phylogenetic tree was inferred on the basis of the amino acid sequences of the extant proteins with the neighbor-joining method on the Jones-Taylor-Thornton (JTT) matrix (*38, 39*). Then, the tree topology was manually refined so that it was coherent to the currently accepted species phylogeny by referring to the literatures (*2, 16*).

The phylogenetic tree and the alignment were applied to PAML v4.5 to infer the ancestral sequences (*40*). The empirical model with the JTT matrix was used for the substitution model. No site partition was defined, and substitution rate was uniform over the sites (*41*). The ancestral sequences were verified by reconstructing molecular phylogenies to determine whether the inferred sequences were connected to the corresponding nodes with zero evolutionary distance. The average posterior probabilities over the amino acid sites for each ancestor are shown in Table 1. The most probable amino acid sequences are aligned in fig. S1. The distributions of site posterior probabilities for each ancestral luciferase are shown in fig. S2. The numbers and distribution of less certainly identified residues, for which the posterior probability is less than 0.1 higher from that of the second most probable amino acid, are shown in fig. S3.

### Ancestral luciferase geological age estimation

The geological ages of the ancestral luciferases were estimated with the Bayesian MCMC method using BEAST v1.8.4. A total of 63 18*S* rRNA gene sequences of Elateriformia species were aligned by using MAFFT v7.309, and the alignment was manually corrected. A total of 665 aligned sites were selected by avoiding the sites on or near the undetermined or gapped sites. The alignment and the phylogeny based on that of luciferases were submitted for a MCMC calculation. The general time reversible (GTR) substitution model

with gamma categories of 4 was adapted. The fossil date 152 Ma ago (standard deviation, 10 Ma ago) of the most recent common ancestor (TMRCA) of Elateroidea (*16*) was used for the reference, and the dates of the nodes were estimated on the uncorrelated relaxed clock model. A total of $10^8$ MCMC iterations were performed, and the 95% HPD interval was obtained excluding the first $10^7$ iterations as a burn-in process. The trajectories were analyzed with TRACER v1.6.0 and FIGTREE v1.4.2. The results of ancestral node dating using the Bayesian MCMC method demonstrated that the date of the reference node (Elateroidea common ancestor) has adequately converged and distributed normally around the expected date with mean of 150.7 Ma ago and 95% HPD from 131.15 to 170.46 Ma ago based on 87,588 effective sample size. The estimated geological ages of ancestral luciferases are shown in Table 1.

### Expression and purification of recombinant proteins

Full ORF of the ancestral sequences were synthesized (Fasmac, Kanagawa, Japan) with *E. coli* optimized codon use (GenBank accession numbers, LC534642 – LC534648) and ligated in frame into the expression vector pCold-ZZ-P-X (*42*). The vectors were subcloned into *E. coli* BL21 (DE3) pLysS (Promega, Madison, WI, USA), and the recombinant proteins were expressed by cold shock at 15°C for 30 min with 0.2 mM of isopropyl-β-D-thiogalactopyranoside. The harvested cells were disrupted by sonication (4°C, 50 W, 5 s) using a Sonicator (Ohtake Works, Tokyo, Japan) three times, and the supernatant was adsorbed on a Talon cobalt-chelating column (Takara Bio, Shiga, Japan). To remove the fused ZZ domain and hexa histidine tag, the column was washed using a wash buffer [50 mM tris-HCl (pH 7.0), 1 M NaCl, and 15 mM imidazole], and then, on-column digestion reaction using PreScission protease (GE Healthcare, Piscataway, NJ, USA) was conducted overnight at 4°C for 12 hours, and the protein was eluted using an elution buffer [50 mM tris-HCl (pH 7.0), 150 mM NaCl, 1 mM EDTA, and 1 mM dithiothreitol]. Protein concentration was measured using Bio-Rad Protein Assay Dye Reagent (Bio-Rad, Hercules, CA, USA) with bovine serum albumin Fraction V (Sigma-Aldrich, St. Louis, MO, USA) as a standard. The homogeneity and

concentration estimated by protein assay were confirmed by SDS–polyacrylamide gel electrophoresis using 10% separation gel with Coomassie Brilliant Blue (CBB) staining (fig. S4). We tested the stability of each recombinant protein by leaving the solution on ice and confirmed that the luminescence activity had not decreased in the order of total light intensities for all proteins even after 12 hours (table S6). This suggests that enzymatic inactivation of the ancestral proteins during purification process was negligible. The purified protein was stored at −80°C until use.

## Measurement of the luminescence spectra
Luminescence spectra were measured using a spectrophotometer AB-1850 (ATTO, Tokyo, Japan). Spectral sensitivity was calibrated. A 2.5 or 5 μg of purified recombinant luciferase was mixed with 500 μM D-luciferin, 4 mM ATP, and 8 mM MgCl$_2$ in 0.1 M tris-HCl buffer (pH 8.0). Exposure time was 1 to 2 min. *P. pyralis* luciferase (PpyLuc1) was purchased from Sigma-Aldrich. Because of its very low luminescence activity, 10 μg of the protein was used for measuring the spectrum of AncElat.

## Measurement of the luminescence intensities
Luminescence intensity was measured using a luminometer CLX-101 (TOYOBO). Luminescence reaction was initiated by injecting the 50 μl of the purified recombinant luciferase (430 ng) into 50 μl of the mixture of D-luciferin (final concentration, 50 or 200 μM), ATP (final concentration, 100 μM), and MgCl$_2$ (final concentration, 5 mM) in tris-HCl buffer (pH 7.8) (final concentration, 50 mM). Light intensity was integrated from 2 to 32 s after mixing (total of 30 s).

## Measurement of the ACS activities
Thioesterification activity was determined using ultraviolet-visible spectrophotometry (Eppendorf, BioSpectrometer kinetic). In this assay, we measured the initial rate of AMP formation by coupling the thioesterification reaction with adenylate kinase, pyruvate kinase, and lactate dehydrogenase and monitoring the oxidation of reduced form of nicotinamide adenine dinucleotide (NADH) at 340 nm (6220 M$^{-1}$ cm$^{-1}$) (*43*). The standard reaction mixture for this assay contained 100 mM tris-HCl (pH 8.0), 10 mM ATP, 10 mM MgCl$_2$, 0.35 mM lauric acid (C12:0) substrate, 2 mM CoASH, 1 mM phosphoenolpyruvic acid, 0.4 mM NADH, adenylate kinase (40 μg/ml), pyruvate kinase (20 μg/ml), lactate dehydrogenase (20 μg/ml), and 10 μg of purified recombinant luciferase. The total volume was brought to 500 μl. The mixture containing all components except for the luciferase was incubated at room temperature (27 ± 2°C) for 10 min. The reaction was then initiated by addition of the enzyme and the data collected at 5 s intervals for 10 min.

## Crystal structure analyses
The structures of AncLamp were determined by x-ray crystallography. The AncLamp crystals were grown by the hanging drop vapor diffusion method, under conditions using 0.1 M trisodium citrate buffer (pH 5.5) containing 20% (w/v) PEG3000 (polyethylene glycol, molecular weight 3000) as a 0.5-ml reservoir, and a mixture of 2 μl of reservoir solution and 2 μl of protein solution in 50 mM tris-HCl (pH 8.0) buffer containing 1% (w/v) AncLamp in the hanging drop. To cocrystallize the protein with reaction-intermediate analog or substrates, 0.2 mM DLSA or 0.2 mM ATP and 0.2 mM D-luciferin were added to the protein solutions. All crystals were grown at 18°C for a few weeks.

X-ray diffraction data were collected from loop-mounted crystals under cryogenic conditions, with a charge-coupled device detector Quantum315 (ADSC) at BL38B1 or Eiger4M at BL26B2 in SPring-8 (Hyogo, Japan). The crystals were soaked for 10 to 30 s in the corresponding crystal growth buffer containing 15% (v/v) 2-methyl-2,4-pentanediol for cryoprotection. The diffraction images were processed with the MOSFLM program (*44*).

The crystal structures were solved by the molecular replacement method using the Phaser-MR application of PHENIX (*45*). The crystal structure of *L. cruciata* LcLuc1 (PDB code 2D1S) was used for a search model. The model refinements were conducted by using COOT and the phenix.refine application of PHENIX (*44*, *46*). The quality of the models was evaluated with the PROCHECK program (*47*). The crystallographic parameters, data collection and refinement statistics, and PDB codes are summarized in table S2. The atomic coordinates and structure factors of AncLamp-DLSA and AncLamp–D-luciferin–ATP complexes have been deposited in the PDB, with the accession codes 6K4C and 6K4D, respectively. The molecular graphics were prepared with CHIMERA (*48*).

The cavity volume of substrate-binding site was evaluated by using CASTp application (*49*). To close the hole for substrates, the AMP moiety of the DLSA or luciferyl-AMP was used for dummy atoms, and the volume for luciferyl moiety was evaluated.

## SUPPLEMENTARY MATERIALS

## REFERENCES AND NOTES
1. S. Lewis, *Silent Sparks: The Wondrous World of Fireflies* (Princeton Univ. Press, 2016).
2. G. J. Martin, M. A. Branham, M. F. Whiting, S. M. Bybee, Total evidence phylogeny and the evolution of adult bioluminescence in fireflies (Coleoptera: Lampyridae). *Mol. Phylogenet. Evol.* **107**, 564–575 (2017).
3. G. J. Martin, K. F. Stanger-Hall, M. A. Branham, L. F. L. Da Silveira, S. E. Lower, D. W. Hall, X.-Y. Li, A. R. Lemmon, E. M. Lemmon, S. M. Bybee, Higher-level phylogeny and reclassification of Lampyridae (Coleoptera: Elateroidea). *Insect Syst. Diver.* **3**, 11 (2019).
4. T. R. Fallon, S. E. Lower, C.-H. Chang, M. Bessho-Uehara, G. J. Martin, A. J. Bewick, M. Behringer, H. J. Debat, I. Wong, J. C. Day, A. Suvorov, C. J. Silva, K. F. Stanger-Hall, D. W. Hall, R. J. Schmitz, D. R. Nelson, S. M. Lewis, S. Shigenobu, S. M. Bybee, A. M. Larracuente, Y. Oba, J.-K. Weng, Firefly genomes illuminate parallel origins of bioluminescence in beetles. *eLife* **7**, e36495 (2018).
5. H. H. Seliger, A. B. Lall, J. E. Lloyd, W. H. Biggley, The colors of firefly bioluminescence—I. Optimization model. *Photochem. Photobiol.* **36**, 673–680 (1982).
6. S. V. Kazantsev, *Protoluciola albertalleni* gen.n., sp.n., a new Luciolinae firefly (Insecta: Coleoptera: Lampyridae) from Burmite amber. *Russian Entomol. J.* **24**, 281–283 (2015).
7. T. Nakatsu, S. Ichiyama, J. Hiratake, A. Saldanha, N. Kobashi, K. Sakata, H. Kato, Structural basis for the spectral difference in luciferase bioluminescence. *Nature* **440**, 372–376 (2006).
8. J. B. Buck, The anatomy and physiology of the light organ in fireflies. *Ann. N. Y. Acad. Sci.* **49**, 397–485 (1948).
9. M. Bessho-Uehara, K. Konishi, Y. Oba, Biochemical characteristics and gene expression profiles of two paralogous luciferases from the Japanese firefly *Pyrocoelia atripennis* (Coleoptera, Lampyridae, Lampyrinae): Insight into the evolution of firefly luciferase genes. *Photochem. Photobiol. Sci.* **16**, 1301–1310 (2017).
10. Y. Oba, N. Mori, M. Yoshida, S. Inouye, Identification and characterization of a luciferase isotype in the Japanese firefly, *Luciola cruciata*, involving in the dim glow of firefly eggs. *Biochemistry* **49**, 10788–10795 (2010).
11. M. Bessho-Uehara, Y. Oba, Identification and characterization of the Luc2-type luciferase in the Japanese firefly, *Luciola parvula*, involved in a dim luminescence in immobile stages. *Luminescence* **32**, 924–931 (2017).
12. J. W. Thornton, Resurrecting ancient genes: Experimental analysis of extinct molecules. *Nat. Rev. Genet.* **5**, 366–375 (2004).

13. Y. Oba, M. Ojika, S. Inouye, Firefly luciferase is a bifunctional enzyme: ATP-dependent monooxygenase and a long chain fatty acyl-CoA synthetase. *FEBS Lett.* **540**, 251–254 (2003).

14. Y. Oba, M. Sato, M. Ojika, S. Inouye, Enzymatic and genetic characterization of firefly luciferase and *Drosophila* CG6178 as a fatty acyl-CoA synthetase. *Biosci. Biotechnol. Biochem.* **69**, 819–828 (2005).

15. S. T. Adams Jr., S. C. Miller, Enzymatic promiscuity and the evolution of bioluminescence. *FEBS J.* **287**, 1369–1380 (2020).

16. D. D. McKenna, A. L. Wild, K. Kanda, C. L. Bellamy, R. G. Beutel, M. S. Caterino, C. W. Farnum, D. C. Hawks, M. A. Ivie, M. L. Jameson, R. A. B. Leschen, A. E. Marvaldi, J. V. McHugh, A. F. Newton, J. A. Robertson, M. K. Thayer, M. F. Whiting, J. F. Lawrence, A. Ślipiński, D. R. Maddison, B. D. Farrell, The beetle tree of life reveals that Coleoptera survived end-Permian mass extinction to diversify during the Cretaceous terrestrial revolution. *Syst. Entomol.* **40**, 835–880 (2015).

17. Y. Oba, M. Furuhashi, M. Bessho, S. Sagawa, H. Ikeya, S. Inouye, Bioluminescence of a firefly pupa: Involvement of a luciferase isotype in the dim glow of pupae and eggs in the Japanese firefly, *Luciola lateralis*. *Photochem. Photobiol. Sci.* **12**, 854–863 (2013).

18. D. T. Amaral, J. R. Silva, V. R. Viviani, Transcriptional comparison of the photogenic and non-photogenic tissues of *Phrixothrix hirtus* (Coleoptera: Phengodidae) and non-luminescent *Chauliognathus flavipes* (Coleoptera: Cantharidae) give insights on the origin of lanterns in railroad worms. *Gene Rep.* **7**, 78–86 (2017).

19. Y. Oba, K. Iida, S. Inouye, Functional conversion of fatty acyl-CoA synthetase to firefly luciferase by site-directed mutagenesis: A key substitution responsible for luminescence activity. *FEBS Lett.* **583**, 2004–2008 (2009).

20. Y. Oba, M. Ojika, S. Inouye, Characterization of *CG6178* gene product with high sequence similarity to firefly luciferase in *Drosophila melanogaster*. *Gene* **329**, 137–145 (2004).

21. R. A. Prado, J. A. Barbosa, Y. Ohmiya, V. R. Viviani, Structural evolution of luciferase activity in *Zophobas* mealworm AMP/CoA-ligase (protoluciferase) through site-directed mutagenesis of the luciferin binding site. *Photochem. Photobiol. Sci.* **10**, 1226–1232 (2011).

22. I. Ueda, A. Suzuki, Is there a specific receptor for anesthetics? Contrary effects of alcohols and fatty acids on phase transition and bioluminescence of firefly luciferase. *Biophys. J.* **75**, 1052–1057 (1998).

23. O. Khersonsky, C. Roodveldt, D. S. Tawfik, Enzyme promiscuity: Evolutionary and mechanistic aspects. *Curr. Opin. Chem. Biol.* **10**, 498–508 (2006).

24. V. R. Viviani, Y. Ohmiya, Bovine serum albumin displays luciferase-like activity in presence of luciferyl adenylate: Insights on the origin of protoluciferase activity and bioluminescence colours. *Luminescence* **21**, 262–267 (2006).

25. V. R. Viviani, V. Scorsato, R. A. Prado, J. G. C. Pereira, K. Niwa, Y. Ohmiya, J. A. R. G. Barbosa, The origin of luciferase activity in *Zophobas* mealworm AMP/CoA-ligase (protoluciferase): Luciferin stereoselectivity as a switch for the oxygenase activity. *Photochem. Photobiol.* **9**, 1111–1119 (2010).

26. J. E. Lloyd, Firefly mating ecology, selection and evolution, in *The Evolution of Mating Systems in Insects and Arachnids*, J. C. Choe, B. J. Crespi, Eds. (Cambridge Univ. Press, 1997), pp. 184–191.

27. R. De Cock, Larval and adult emission spectra of bioluminescence in three European firefly species. *Photochem. Photobiol.* **79**, 339–342 (2004).

28. D. W. Hall, S. E. Sander, J. C. Pallansch, K. F. Stanger-Hall, The evolution of adult light emission color in North American fireflies. *Evolution* **70**, 2033–2048 (2016).

29. Y. Oba, M. A. Branham, T. Fukatsu, The terrestrial bioluminescent animals of Japan. *Zoolog. Sci.* **28**, 771–789 (2011).

30. J. Sivinski, Arthropods attract to luminous fungi. *Psyche* **88**, 383–390 (1981).

31. P. Marek, D. Papaj, J. Yeager, S. Molina, W. Moore, Bioluminescent aposematism in millipedes. *Curr. Biol.* **21**, R680–R681 (2011).

32. B. C. Leavell, J. J. Rubin, C. J. W. McClure, K. A. Miner, M. A. Branham, J. R. Barber, Fireflies thwart bat attack with multisensory warnings. *Sci. Adv.* **4**, eaat6601 (2018).

33. V. R. Viviani, D. R. Neves, D. T. Amaral, R. A. Prado, T. Matsuhashi, T. Hirano, Bioluminescence of beetle luciferases with 6′-amino-D-luciferin analogues reveals excited keto-oxyluciferin as the emitter and phenolate/luciferin binding site interactions modulate bioluminescence colors. *Biochemistry* **53**, 5208–5220 (2014).

34. B. R. Branchini, D. M. Ablamsky, J. M. Rosenman, L. Uzasci, T. L. Southworth, M. Zimmer, Synergistic mutations produce blue-shifted bioluminescence in firefly luciferase. *Biochemistry* **46**, 13847–13855 (2007).

35. N. Kajiyama, E. Nakano, Isolation and characterization of mutants of firefly luciferase which produce different colors of light. *Protein Eng.* **4**, 691–693 (1991).

36. T. W. Cronin, M. Järvilehto, M. Weckström, A. B. Lall, Tuning of photoreceptor spectral sensitivity in fireflies (Coleoptera: Lampyridae). *J. Comp. Physiol. A* **186**, 1–12 (2000).

37. K. Katoh, K. Misawa, K.-i. Kuma, T. Miyata, MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).

38. N. Saitou, M. Nei, The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425 (1987).

39. D. T. Jones, W. R. Taylor, J. M. Thornton, The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**, 275–282 (1992).

40. Z. Yang, PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).

41. Z. Yang, B. Rannala, Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol. Biol. Evol.* **23**, 212–226 (2006).

42. S. Inouye, Y. Sahara, Expression and purification of the calcium binding photoprotein mitrocomin using ZZ-domain as a soluble partner in *E. coli* cells. *Protein Expres. Purif.* **66**, 52–57 (2009).

43. D.-i. Kato, Y. Hiraishi, M. Maenaka, K. Yokoyama, K. Niwa, Y. Ohmiya, M. Takeo, S. Negoro, Interconversion of ketoprofen recognition in firefly luciferase-catalyzed enantioselective thioesterification reaction using from *Pylocoeria miyako* (PmL) and *Hotaria parvura* (HpL) just by mutating two amino acid residues. *J. Biotechnol.* **168**, 277–283 (2013).

44. T. G. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell, A. G. W. Leslie, iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 271–281 (2011).

45. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *PHENIX*: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).

46. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2004).

47. R. A. Laskowski, D. S. Moss, J. M. Thornton, Main-chain bond lengths and bond angles in protein structures. *J. Mol. Biol.* **231**, 1049–1067 (1993).

48. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

49. W. Tian, C. Chen, X. Lei, J. Zhao, J. Liang, CASTp 3.0: Computed atlas of surface topography of proteins. *Nucleic Acids Res.* **46**, W363–W367 (2018).

50. R. Dawkins, Y. Wong, *The ancestor's tale: a pilgrimage to the dawn of life* (Weidenfeld & Nicolson, London, revised and expanded edition, 2016), pp. 1–11.

51. V. R. Viviani, E. J. H. Bechara, Y. Ohmiya, Cloning, sequence analysis, and expression of active *Phrixothrix* railroad-worms luciferases: Relationship between bioluminescence spectra and primary structures. *Biochemistry* **38**, 8271–8279 (1999).

52. V. R. Viviani, Y. Ohmiya, Bioluminescence color determinants of *Phrixothrix* railroad-worm luciferases: Chimeric luciferases, site-directed mutagenesis of Arg 215 and guanidine effect. *Photochem. Photobiol.* **72**, 267–271 (2000).

53. Y. Ohmiya, M. Sumiya, V. R. Viviani, N. Ohba, Comparative aspects of a luciferase molecule from the Japanese luminous beetle, *Rhagophthalmus ohbai*. *Sci. Rep. Yokosuka City Mus.* **47**, 31–38 (2000).

54. K. V. Wood, Y. A. Lam, H. H. Seliger, W. D. McElroy, Complementary DNA coding click beetle luciferases can elicit bioluminescence of different colors. *Science* **244**, 700–702 (1989).

55. Y. Oba, M. Kumazaki, S. Inouye, Characterization of luciferases and its paralogue in the Panamanian luminous click beetle *Pyrophorus angustus*: A click beetle luciferase lacks the fatty acyl-CoA synthetic activity. *Gene* **452**, 1–6 (2010).

56. V. R. Viviani, A. C. R. Silva, G. L. O. Perez, R. V. Santelli, E. J. H. Bechara, F. C. Reinach, Cloning and molecular characterization of the cDNA for the Brazilian larval click-beetle *Pyrearinus termitilluminans* luciferase. *Photochem. Photobiol.* **70**, 254–260 (1999).

57. A. J. Silva Neto, V. Scorsato, F. G. C. Arnoldi, V. R. Viviani, *Pyrearinus termitilluminans* larval click beetle luciferase: Active site properties, structure and function relationships and comparison with other beetle luciferases. *Photochem. Photobiol. Sci.* **8**, 1748–1754 (2009).

58. Y. Mitani, R. Futahashi, K. Niwa, N. Ohba, Y. Ohmiya, Cloning and characterization of luciferase from a Fijian luminous click beetle. *Photochem. Photobiol.* **89**, 1163–1169 (2013).

59. G. D. Kutuzova, R. R. Hannah, K. V. Wood, Bioluminescence color variation and kinetic behavior relationships among beetle luciferases, in *Bioluminescence and Chemiluminescence: Molecular Reporting with Photons*, J. W. Hastings, L. J. Kricka, P. E. Stanley, Eds. (John Wiley and Sons, 1997), pp. 248–252.

60. H. Tatsumi, N. Masuda, N. Kajiyama, E. Nakano, Luciferase cDNA from Japanese firefly, *Luciola cruciata*: Cloning, structure and expression in *Escherichia coli*. *J. Biolumin. Chemilumin.* **3**, 75–78 (1989).

61. Y. Oba, M. Yoshida, T. Shintani, M. Furuhashi, S. Inouye, Firefly luciferase genes from the subfamilies Psilocladinae and Ototretinae (Lampyridae, Coleoptera). *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **161**, 110–116 (2012).

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

63. M. I. Koksharov, N. N. Ugarova, Random mutagenesis of *Luciola mingrelica* firefly luciferase. Mutant enzymes with bioluminescence spectra showing low pH sensitivity. *Biochemistry (Mosc.)* **73**, 862–869 (2008).

64. B. R. Branchini, T. L. Southworth, J. P. DeAngelis, A. Roda, E. Michelini, Luciferase from the Italian firefly *Luciola italica*: Molecular cloning and expression. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **145**, 159–167 (2006).

65. B. R. Branchini, T. L. Southworth, L. J. Salituro, D. M. Fontaine, Y. Oba, Cloning of the Blue Ghost (*Phausis reticulata*) luciferase reveals a glowing source of green light. *Photochem. Photobiol.* **93**, 473–478 (2017).

66. M. Mortazavi, S. Hosseinkhani, K. Khajeh, B. Ranjbar, A. R. Emamzadeh, Spectroscopic and functional characterization of *Lampyris turkestanicus* luciferase: A comparative study. *Acta Biochim. Biophys. Sin.* **40**, 365–374 (2008).

67. Y. Oba, M. Furuhashi, S. Inouye, Identification of a functional luciferase gene in the non-luminous diurnal firefly, *Lucidina biplagiata*. *Insect Mol. Biol.* **19**, 737–743 (2010).

68. V. R. Viviani, F. G. C. Arnoldi, M. Brochetto-Braga, Y. Ohmiya, Cloning and characterization of the cDNA for the Brazilian *Cratomorphus distinctus* larval firefly luciferase: Similarities with European *Lampyris noctiluca* and Asiatic *Pyrocoelia* luciferases. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **139**, 151–156 (2004).

69. N. K. Tafreshi, M. Sadeghizadeh, R. Emamzadeh, B. Ranjbar, H. Naderi-Manesh, S. Hosseinkhani, Site-directed mutagenesis of firefly luciferase: Implication of conserved residue(s) in bioluminescence emission spectra among firefly luciferases. *Biochem. J.* **412**, 27–33 (2008).

70. G. B. Sala-Newby, C. M. Thomson, A. K. Campbell, Sequence and biochemical similarities between the luciferases of the glow-worm *Lampyris noctiluca* and the firefly *Photinus pyralis*. *Biochem. J.* **313**, 761–767 (1996).

71. M. I. Koksharov, N. N. Ugarova, Strategy of mutual compensation of green and red mutants of firefly luciferase identifies a mutation of the highly conservative residue E457 with a strong red shift of bioluminescence. *Photochem. Photobiol. Sci.* **12**, 2016–2027 (2013).

72. B. R. Branchini, R. A. Magyar, M. H. Murtiashaw, N. C. Portier, The role of active site residue arginine 218 in firefly luciferase bioluminescence. *Biochemistry* **40**, 2410–2418 (2001).

73. V. R. Viviani, F. G. C. Arnoldi, A. J. S. Neto, T. L. Oehlmeyer, E. J. H. Bechara, Y. Ohmiya, The structural origin and biological function of pH-sensitivity in firefly luciferases. *Photochem. Photobiol. Sci.* **7**, 159–169 (2008).

74. B. R. Branchini, T. L. Southworth, N. F. Khattak, E. Michelini, A. Roda, Red- and green-emitting firefly luciferase mutants for bioluminescent reporter applications. *Anal. Biochem.* **345**, 140–148 (2005).

75. B. R. Branchini, R. A. Magyar, M. H. Murtiashaw, S. M. Anderson, L. C. Helgerson, M. Zimmer, Site-directed mutagenesis of firefly luciferase active site amino acids: A proposed model for bioluminescence color. *Biochemistry* **38**, 13223–13230 (1999).

76. B. R. Branchini, D. M. Fontaine, T. L. Southworth, B. P. Huta, A. Racela, K. D. Patel, A. M. Gulick, Mutagenesis and structural studies reveal the basis for the activity and stability properties that distinguish the *Photinus* Luciferases *scintillans* and *pyralis*. *Biochemistry* **58**, 4293–4303 (2019).

77. Y. Wang, H. Akiyama, K. Terakado, T. Nakatsu, Impact of site-directed mutant luciferase on quantitative green and orange/red emission intensities in firefly bioluminescence. *Sci. Rep.* **3**, 2490 (2013).

78. D.-i. Kato, T. Kubo, M. Maenaka, K. Niwa, Y. Ohmiya, M. Takeo, S. Negoro, Confirmation of color determination factors for Ser286 derivatives of firefly luciferase from *Luciola cruciata* (LUC-G). *J. Mol. Catal. B: Enzym.* **87**, 18–23 (2013).

79. V. R. Viviani, T. L. Oehlmeyer, F. G. Arnoldi, M. R. Brochetto-Braga, A new firefly luciferase with bimodal spectrum: Identification of structural determinants of spectral pH-sensitivity in firefly luciferases. *Photochem. Photobiol.* **81**, 843–848 (2005).

80. H. Ueda, H. Yamanouchi, A. Kitayama, K. Inoue, T. Hirano, E. Suzuki, T. Nagamune, Y. Ohmiya, in *Bioluminescence and Chemiluminescence: Molecular Reporting with Photons*, J. W. Hastings, L. J. Kricka, P. E. Stanley, Eds. (John Wiley & Sons, 1996), pp. 216–219.

81. Y. Modestova, M. I. Koksharov, N. N. Ugarova, Point mutations in firefly luciferase C-domain demonstrate its significance in green color of bioluminescence. *Biochim. Biophys. Acta* **1844**, 1463–1471 (2014).

82. B. R. Branchini, M. H. Murtiashaw, R. A. Magyar, S. M. Anderson, The role of lysine 529, a conserved residue of the acyl-adenylate-forming enzyme superfamily, in firefly luciferase. *Biochemistry* **39**, 5433–5440 (2000).

**Citation:** Y. Oba, K. Konishi, D. Yano, H. Shibata, D. Kato, T. Shirai, Resurrecting the ancient glow of the fireflies. *Sci. Adv.* **6**, eabc5705 (2020).