Genetically encodable bioluminescent system from fungi

Alexey A. Kotlobay1,a, Karen S. Sarkisiyan1,a,b,c,d,1,a, Yuliya A. Mokrushina1,a, Marina Marct-Houben1,a,b, Ekaterina O. Serebrovskaya1,a, Nadezhda M. Markina1,a,b,g, Louisa Gonzalez Somermeyer2, André Y. Gorokhovatsky2, Andrey Vvedensky2, Konstantin V. Purto2, Valentin P. Petushkov2, Natalja S. Rodionova2, Tatiana V. Chepurnyh, Tatiana V. Chepurnyh, Alexey A. Kotlobay, and Sergey V. Dolgov2, Tatiana Y. Mitouchkina3,a,b, Eugene P. Kopantzev2, Hans E. Waldenmaier4, Anderson G. Oliveira5, Yuichi Obasa6, Ekaterina Barsova7,a,g, Ekaterina A. Bogdanova7,a, Toni Gabaldón8,a,b,p, Cassius V. Stevani9, Sergey Lukyanov10,a,j, Ivan V. Smirnov10, Josef I. Gitelson10, Fyodor A. Kondrashov10, and Iliya V. Yampolsky1,b,1,a,2

1Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, 117997 Moscow, Russia; 2Planta LLC, 121205 Moscow, Russia; 3Institute of Science and Technology Austria, 3400 Klosterneuburg, Austria; 4Medical Research Council London Institute of Medical Sciences, Imperial College London, W12 0NN London, United Kingdom; 5Centre for Genomic Regulation, The Barcelona Institute for Science and Technology, 08003 Barcelona, Spain; 6Universitat Pompeu Fabra, 08003 Barcelona, Spain; 7Evrogen JSC, 117997 Moscow, Russia; 8Institute of Biophysics, Federal Research Center Krasnoyarsk Science Center, Siberian Branch, Russian Academy of Sciences, 660036 Krasnoyarsk, Russia; 9Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences, 142290 Moscow, Russia; 10Progov Russian National Research Medical University, 117997 Moscow, Russia; 11Biomedical Nanomaterials, National Research Technological University (MISiS), 119049 Moscow, Russia; 12Skolkovo Institute of Science and Technology, 121205 Moscow, Russia; 13Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, 05508-000 São Paulo, Brazil; 14Departamento de Oceanografia Física, Quimica e Geológica, Instituto Oceanográfico, Universidade de São Paulo, 05508-120 São Paulo, Brazil; 15Department of Environmental Biology, Chubu University, 487-8510 Kasugai, Japan; 16Catalan Institute for Research and Advanced Studies (ICREA), 08010 Barcelona, Spain; and 17Departamento de Química Fundamental, Instituto de Química, Universidade de São Paulo, 05508-000 São Paulo, Brazil

Edited by E. Peter Greenberg, University of Washington, Seattle, WA, and approved October 25, 2018 (received for review February 28, 2018)

Bioluminescence is found across the entire tree of life, conferring a spectacular set of visually oriented functions from attracting mates to scaring off predators. Half a dozen different luciferins, molecules that emit light when enzymatically oxidized, are known. However, just one biochemical pathway for luciferin biosynthesis has been described in full, which is found only in bacteria. Here, we report identification of the fungal luciferase and three other key enzymes that together form the biosynthetic cycle of the fungal luciferin from caffeic acid, a simple and widespread metabolite. Introduction of the identified genes into the genome of the yeast *Pichia pastoris* along with caffeic acid biosynthesis genes resulted in a strain that is autoluminescent in standard media. We analyzed evolution of the enzymes of the luciferin biosynthesis cycle and found that fungal bioluminescence emerged through a series of events that included two independent gene duplications. The retention of the duplicated enzymes of the luciferin pathway in nonluminescent fungi shows that the gene duplication was followed by functional sequence divergence of enzymes at least one gene in the biosynthetic pathway and suggests that the evolution of fungal bioluminescence proceeded through several closely related stepping stone nonluminescent biochemical reactions with adaptive roles. The availability of a complete eukaryotic luciferin biosynthesis pathway provides several applications in biomedicine and bioengineering.

Bioluminescence | fungal luciferin biosynthesis | fungal luciferase

Bioluminescence is a natural phenomenon of light emission resulting from oxidation of a substrate, luciferin, catalyzed by an enzyme, luciferase. A variety of species are bioluminescent in nature (1); for many of them, the ability to emit light is a defining feature of their biology (2–4). Artificial integration of bioluminescent reactions into living systems has also become a reporting tool widely used in molecular and cell biology (5, 6). However, natural bioluminescent systems remain poorly characterized on a biochemical level, limiting more widespread application. Only 9 luciferins and 7 luciferase gene families have been described (7, 8) of at least 40 bioluminescent systems thought to exist in nature (9). Lamentably, only a single biochemical cascade starting from a widespread metabolite to a luciferin has been described in its entirety (10). The described pathway is bacterial and has limited application in eukaryotes (11). None of the eukaryotic bioluminescent systems have been described in sufficient detail to be expressed in another organism or to create artificial autonomously bioluminescent organisms. Here, we describe the function and evolution of the key genes responsible for the bioluminescence of the fungus *Neonothopanus nambi* and show that expression of fungal genes is sufficient to engineer autonomously bioluminescent eukaryotes.

Approximately 100 fungal species from the order Agaricales emit light utilizing the same biochemical reaction (12). Although the ecological role of their bioluminescence is not fully understood, there is evidence that it might be used by fungi to detect prey or to attract mates (13). Lamentably, only a single biocatalytic biochemical cascade is known. However, just one biochemical pathway for luciferin biosynthesis has been described in full, which is found only in bacteria. Here, we report identification of the fungal luciferase and three other key enzymes that together form the biosynthetic cycle of the fungal luciferin from caffeic acid, a simple and widespread metabolite. Introduction of the identified genes into the genome of the yeast *Pichia pastoris* along with caffeic acid biosynthesis genes resulted in a strain that is autoluminescent in standard media. We analyzed evolution of the enzymes of the luciferin biosynthesis cycle and found that fungal bioluminescence emerged through a series of events that included two independent gene duplications. The retention of the duplicated enzymes of the luciferin pathway in nonluminescent fungi shows that the gene duplication was followed by functional sequence divergence of enzymes of at least one gene in the biosynthetic pathway and suggests that the evolution of fungal bioluminescence proceeded through several closely related stepping stone nonluminescent biochemical reactions with adaptive roles. The availability of a complete eukaryotic luciferin biosynthesis pathway provides several applications in biomedicine and bioengineering.

Bioluminescence | fungal luciferin biosynthesis | fungal luciferase

Bioluminescence is a natural phenomenon of light emission resulting from oxidation of a substrate, luciferin, catalyzed by an enzyme, luciferase. A variety of species are bioluminescent in nature (1); for many of them, the ability to emit light is a defining feature of their biology (2–4). Artificial integration of bioluminescent reactions into living systems has also become a reporting tool widely used in molecular and cell biology (5, 6). However, natural bioluminescent systems remain poorly characterized on a biochemical level, limiting more widespread application. Only 9 luciferins and 7 luciferase gene families have been described (7, 8) of at least 40 bioluminescent systems thought to exist in nature (9). Lamentably, only a single biochemical cascade starting from a widespread metabolite to a luciferin has been described in its entirety (10). The described pathway is bacterial and has limited application in eukaryotes (11). None of the eukaryotic bioluminescent systems have been described in sufficient detail to be expressed in another organism or to create artificial autonomously bioluminescent organisms. Here, we describe the function and evolution of the key genes responsible for the bioluminescence of the fungus *Neonothopanus nambi* and show that expression of fungal genes is sufficient to engineer autonomously bioluminescent eukaryotes.

Edited by E. Peter Greenberg, University of Washington, Seattle, WA, and approved October 25, 2018 (received for review February 28, 2018)


Conflict of interest statement: K.S.S. and I.Y.V. are shareholders of Planta LLC. Planta LLC filed patent applications related to the use of the enzymes of fungal bioluminescent system.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: The data reported in this paper have been deposited in the NCBI Bio-project (accession no. PRJNA476325). Transcriptsomes of *N. nambi* and *M. cirticolor* are available at Fqshare (https://fqshare.com/articles/A_genetically_encodable_bioluminescent_system_from_fungi/6738953/3). Files used to reconstruct the Agaricales species tree, including raw and trimmed alignments and RAxML resulting files, are available at Fqshare (https://fqshare.com/articles/Species_tree_files/6572117/). Sequences of fungal luciferases are available at Figshare (https://figshare.com/articles/PRJNA476325).

1A.A.K., K.S.S., Y.A.M., M.M.H., and E.O.S. contributed equally to this work.

2To whom correspondence may be addressed. Email: karen@planta.bio or ivyamp@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1803615115/-/DCSupplemental.
Significance

We present identification of the luciferase and enzymes of the biosynthesis of a eukaryotic luciferin from fungi. Fungi possess a single biosynthetic system, with luciferins being only two enzymatic steps from well-known metabolic pathways. The expression of genes from the fungal biosynthetic pathway is not toxic to eukaryotic cells, and the luciferase can be easily co-opted to bioimaging applications. With the fungal system being a genetically encodable biosynthetic system from eukaryotes, it is now possible to create artificially bioluminescent eukaryotes by expression of three genes. The fungal biosynthetic system represents an example of molecular evolution of a complex ecological trait and with molecular details reported in the paper, will allow additional research into ecological significance of fungal bioluminescence.

attract spore-distributing insects (13). Fungal bioluminescence was known to utilize at least four components: molecular oxygen; the luciferin, which was recently identified as 3-hydroxyhispidin [a product of oxidation of the simple plant and fungal metabolite hispidin (14)]; and two previously undescribed key enzymes, an NAD(P)H-dependent hydroxylase and a luciferase (15, 16).

To identify enzymes of the fungal biosynthetic pathway, we first focused on isolation of the luciferase gene. By expressing the N. nambi cDNA library in Pichia pastoris and spraying agar plates with synthetic 3-hydroxyhispidin, we identified and sequenced a luminescent yeast colony expressing the luciferase gene (SI Appendix, Figs. S1 and S13). The N. nambi luciferase, nnLuz, is a 267-aa protein (SI Appendix, Fig. S2) and has no described homologs or pronounced sequence similarity to conserved protein domains, representing a novel protein family.

Genes coding for enzymes that synthesize secondary metabolites are often clustered in fungal genomes (17). We hypothesized that this may be the case for enzymes of the bioluminescent cascade, because it is thought that the cascade is conserved among the bioluminescent fungi (12). We thus looked for genes related to luciferin biosynthesis in the vicinity of the luciferase gene in the N. nambi genome. In addition to N. nambi, we also sequenced genomes and transcriptomes of bioluminescent fungi Neonothopanus gardneri, Mycena citricolor, and Panellus stipticus and compared them with publicly available genome sequences of bioluminescent and nonbioluminescent fungi (18, 19). We found that the luciferase is a member of a conserved gene cluster, which includes at least two other genes: h3h and hisps (Fig. 1 B and C and Datasets S1–S3).

The h3h gene showed sequence similarity with 3-hydroxybenzoate 6-monoxygenases, enzymes that catalyze oxidation of 3-hydroxybenzoate using NADH and molecular oxygen. This reaction is identical to that which converts hispidin into luciferin (Fig. L4); thus, we hypothesized that h3h gene codes for hispidin-3-hydroxylase (H3H), the enzyme corresponding to the predicted hydroxylase (15). We cloned the gene from N. nambi and found that P. pastoris colonies expressing both nnluz and nnh3h emit light when sprayed with luciferin precursor hispidin, unlike control colonies expressing nnluz alone (SI Appendix, Figs. S14 and S17)—confirming that nnH3H converts hispidin into luciferin.

The hisps gene (Fig. 1C) encodes a member of the polyketide synthase family, enzymes that produce secondary metabolites in a variety of organisms across the tree of life (20). Polyketide synthases typically add malonyl moieties to the growing carbon chain; thus, the α-pyrene nature of hispidin suggested that its biosynthesis may be performed by a polyketide synthase from caffeic acid by two cycles of addition of malonyl units followed by lactonization (Fig. L4 and SI Appendix, Fig. S3). Large modular polyketide synthases require posttranslational modifications for their activity (21), such as the transfer of a phosphopantetheiny group to a conserved serine residue of the acyl carrier protein.

Fig. 1. Luciferin biosynthesis pathway in fungal bioluminescence and gene cluster containing key enzymes in the clade of bioluminescent fungi. (A) Proposed pathway of fungal luciferin biosynthesis and recycling. Caffeic acid is converted to hispidin by hispidin synthase (HisPS) and hydroxylated by H3H, yielding 3-hydroxyhispidin (fungal luciferin). The luciferase (Luz) adds molecular oxygen, producing an endoperoxide as a high-energy intermediate with decomposition that yields oxyluciferin (caffeoylpyruvate) and light emission. Oxyluciferin can be recycled to caffeic acid by caffeoylpyruvate hydrolase (CPH). (B) Schematic depiction of the genomic cluster of N. nambi containing luciferase, H3H, hispidin synthase, and caffeoylpyruvate hydrolase (cph) genes. (C) Gene cluster in the clade of bioluminescent fungi. The species tree in Left is based on the comparison of protein-coding genes shared by most of the analyzed species. The red crosses mark the branches of the tree that eventually lost the ability to glow. Right shows the structure of the luciferase-containing gene cluster if such a cluster was found in the relevant genome. The genes coding for luciferase (Luz), h3h, hispidin synthase (hisps), and caffeoylpyruvate hydrolase (cph) are colored. The lighter blue and red colors of hisps and luz genes indicate that only a partial or truncated gene was found in Armillaria mellea and Guyanagaster necrophiza, respectively. Other genes that might belong to the cluster are named from O1 to O4 (colored in gray). Green ticks represent a cytochrome P450-like gene (different shades of green indicate different orthologous groups), and black ticks indicate other genes.
domain. To test whether hisps gene can produce luciferin precursor in a heterologous system, we integrated hisps, nnluz, and nnh3h genes together with the Aspergillus nidulans 4′-phosphopantetheinyl transferase gene ngpA into the genome of *P. pastoris*. When cultured in a medium containing caffeic acid, yeasts expressing all four genes emitted light seen by a naked eye (Fig. 3A), while no significant light production was observed in strains lacking ngpA or hisps genes (SI Appendix, Figs. S15 and S17). Therefore, hisps catalyzes synthesis of hispidan from caffeic acid, closing the chain of reactions (the “caffeic acid cycle”) from a common cellular metabolite with known biosynthesis to a eukaryotic luciferin.

In some bioluminescent species of fungi, the genomic cluster accommodates one or two additional genes (Fig. 1C): one belonging to the cytochrome P450 family, and the other belonging to the family of fumarylacetoacetate hydrolases. The latter (cph) likely encodes a caffeeylpyruvate hydrolase (Dataset S4) involved in oxyluciferin recycling, consistent with caffeeylpyruvate, the fungal oxyluciferin, being hydrolyzed to caffeate and pyruvate by a hydrolase present in fungal crude extracts (22).

Conservation of the gene cluster suggests that, in contrast to other groups of bioluminescent organisms (23), bioluminescence evolved in fungi only once, with *luz*, *h3h*, and *hisps* genes emerging through gene duplications. Reconstructed phylogenetic trees for *luz*, *h3h*, and *hisps* genes (SI Appendix, Figs. S4–S6) and the species tree of order Agaricales (Fig. 2) reveal the evolutionary history of the bioluminescence cascade in fungi. The primary *luz* enzyme of the fungal bioluminescence cascade emerged as a result of gene duplication at the base of Agaricales followed by the duplication of *h3h* and *hisps* a few million years later. Interestingly, many species in a large clade encoding *h3h* genes are nonbioluminescent, and the *hisps* homologs in bioluminescent species lack two domains, the ketoreductase and dehydratase domains (SI Appendix, Fig. S6). It is likely that the loss of these functional domains in the common ancestor of bioluminescent species favored the production of α-pyrones by ancestral *hisps*, possibly providing the final step for emergence of bioluminescence.

The gene cluster continued to evolve dynamically after the acquisition of bioluminescence. At least six independent complete or partial gene loss events of genes from the genomic cluster led to secondary loss of bioluminescence (Fig. 2). The *cph* gene was inserted into the cluster in the nonmycénoid clade, possibly twice (Fig. 1). This mosaic pattern resembles evolutionary history of fluorescent proteins (24), another visually relevant protein family with an unclear biological role, and may indicate that selective advantage provided by bioluminescence in fungi depends on a specific or transient ecological context.

Complex adaptations can be a source of biotechnologically relevant solutions. In addition to revealing the nature of basic photochemical processes and protein evolution, luciferases are among the primary types of reporter genes used in various research pipelines, methods of diagnostics, and environmental applications (5, 6, 25). To determine whether a fungal bioluminescent pathway can provide reporter genes, we characterized nnLuz in vitro and tested its ability to produce light in heterologous systems.

The nnLuz protein consists of 267 amino acids and has the molecular mass of about 28.5 kDa (SI Appendix, Fig. S7). Although its cellular localization in vivo remains unknown, the protein has a predicted N-terminal transmembrane helix, consistent with colocalization of luciferase activity with insoluble cell fractions in previous studies (22). When expressed in *P. pastoris*, nnLuz was associated with the microsomal fraction (SI Appendix, Fig. S8) and emitted green light with a maximum at 520 nm and spectrum identical to that of *N. nambi* mycelium (Fig. 3A and SI Appendix, Fig. S9). Recombinant nnLuz emits light optimally at around pH 8.0 and moderate temperatures, losing its activity at temperatures above 30 °C (SI Appendix, Fig. S10).

To test the potential of *nnluz* as a reporter gene in heterologous systems, we tested its expression in *Escherichia coli*, *P. pastoris*, early *Xenopus laevis* embryos, and human cells. Although luciferase accumulated mostly in inclusion bodies when expressed in bacteria (SI Appendix, Fig. S7), all tested cells and organisms expressing wild-type *nnluz* were clearly bioluminescent when 3-hydroxyhispidin was added to the medium (Fig. 3 and SI Appendix, Figs. S11 and S22). We also compared nnLuz qualitatively with the luciferase from the firefly *Photinus pyralis* in a whole-body imaging setup of tumor xenografts in mice. We s.c. implanted equal amounts of murine colon carcinoma cells expressing either *nnluz* or firefly luciferase under the same promoter, injected a
mixture of firefly and fungal luciferins i.p., and obtained almost identical signals from the implants (Fig. 3C).

Finally, we aimed to test whether luciferin biosynthesis can be achieved in organisms lacking caffeic acid biosynthesis. Introduction of three additional genes coding for enzymes of caffeic acid biosynthesis from tyrosine, *Rhodobacter capsulatus* tyrosine ammonia lyase and two *E. coli* 4-hydroxyphenylacetate 3-monooxygenase components (26), into the genome of *P. pastoris* strain carrying *npgA*, *hisps*, *h3h*, and *luz* genes resulted in a strain that was autonomously bioluminescent when grown in a standard yeast medium (*SI Appendix*, Figs. S12 and S16).

Thus, under all tested conditions, wild-type *N. nambi* luciferase is functional in heterologous systems, positioning itself as a promising reporter gene, and fungal luciferin can be synthesized from aromatic amino acids in other eukaryotes. In addition, fungal luciferin is a water-soluble and cell-permeable compound, and its light-emitting reaction is not dependent on the availability of ATP, making the fungal bioluminescent system attractive for numerous applications in biomedical imaging. Furthermore, various luciferin analogs can be used to enhance light emission and tune its spectra, improving light penetration in deep tissue imaging applications (22).

In conclusion, we present the enzymatic cascade that leads to light emission in fungi, which is a eukaryotic bioluminescence system with known biosynthesis of luciferin. We have shown that luciferin is synthesized from its precursor hispidin by *N. nambi* H3H and that hispidin can be directly synthesized by hispidin synthase from caffeic acid, a widespread cellular metabolite with efficient biosynthesis that was achieved in various organisms, including industrially relevant yeast strains (26). Just two enzymatic steps from the mainstream metabolic pathways, the fungal system has a high potential for synthetic biology to create autonomously glowing animals and plants: attempts to develop such organisms have so far been constrained by the poor performance in eukaryotes of the bacterial bioluminescent system, the only system for which luciferin biosynthesis was known (27, 28).

Reconstitution of fungal bioluminescent pathway in eukaryotic organisms might enable applications where tissues or organisms report changes in their physiological state with autonomous light emission. It might also push forward development of the next generation of organic architecture (29), where genetically modified glowing plants will be integrated into buildings and city infrastructure. Apart from that, with its intriguing evolutionary history, a family of luciferases, and overall simplicity, the fungal bioluminescent system presented here is a molecular playground holding numerous opportunities for basic and applied research.

**Materials and Methods**

*SI Appendix* includes the details of the materials and methods used in this study, including experiments with *Xenopus* embryos, mice, yeasts, bacteria, mammalian cells, and bioinformatic analyses. Animal experiments were approved by the local ethical committee of Pirogov Russian National Research Medical University and were carried out in accordance with European Union Directive 2016/63/EU.

Genomes of *P. stipticus*, *Lentinula edodes*, *N. gardneri*, *N. nambi*, and *M. citricolor* and transcriptomes of *P. stipticus*, *L. edodes*, and *N. gardneri* are available at the National Center for Biotechnology Information Bioproject PRJNA476325. Transcriptomes of *N. nambi* and *M. citricolor*, alignments of *P. pastoris* genome sequencing reads, and alignment of protein sequences of fungal luciferases are available at Figshare (https://figshare.com/articles/A_genetically_encodable_bioluminescent_system_from_fungi/67389532). Files used to reconstruct the Agaricaceae species tree, including raw and trimmed alignments and RAxML resulting files, are available at Figshare.