Predictable genetic recruitment for luciferin sulfation in the convergent evolution of bioluminescence

Emily S Lau[1*], Jessica A Goodheart[1,2], Nolan T Anderson[3], Vannie L Liu[3], Arnab Mukherjee[3,4,5,6*], Todd H Oakley[1*]

[1] Department of Ecology, Evolution, and Marine Biology, University of California Santa Barbara, Santa Barbara, CA 93106

[2] Division of Invertebrate Zoology, American Museum of Natural History, New York, NY 10025

[3] Department of Chemical Engineering, University of California Santa Barbara, Santa Barbara, CA 93106

[4] Department of Biological Engineering, University of California Santa Barbara, Santa Barbara, CA 93106

[5] Department of Chemistry, University of California Santa Barbara, Santa Barbara, CA 93106

[6] Neuroscience Research Institute, University of California Santa Barbara, Santa Barbara, CA 93106

[*]**Corresponding authors:** Emily Lau (emily.lau@lifesci.ucsb.edu); Arnab Mukherjee (arnabm@ucsb.edu); Todd Oakley (oakley@lifesci.ucsb.edu)

Abstract

Determinism and contingency shape evolutionary trajectories and the extent to which evolution is predictable depends on the relative contribution of each. However, the causes of patterns of predictable evolution are poorly understood. Here we propose evolution may be more deterministic — and therefore predictable — if functional evolution is constrained to fewer genetic solutions. We examine the bioluminescence system of ostracods and synthesize our findings with those reported previously for fireflies and sea pansies to show one aspect of bioluminescence, substrate metabolism, evolved convergently by recruiting homologous enzymes. We use gene expression, recombinant protein expression, and *in vitro* functional analyses to identify genes capable of substrate metabolism in ostracods. Our results, taken together with knowledge of firefly and sea pansy bioluminescence systems, provide empirical evidence supporting that these convergent mechanisms for substrate metabolism repeatedly evolved by co-opting sulfotransferases, a ubiquitous family of enzymes with broad substrate specificity. From this synthesis, we propose the lack of diverse genetic solutions for sulfation constrains evolution, resulting in more evolutionary determinism and increasing our ability to predict evolution.

Keywords: convergent evolution, complex traits, gene expression, bioluminescence, sulfotransferase

Introduction

Convergent evolution — the repeated evolution of similar phenotypes — provides replicates useful for investigating long-standing biological questions, such as to what extent evolution is predictable (Orgogozo 2015; Stayton 2015). Evolution is more predictable and deterministic when similar features are shared across convergently evolved phenotypes, but less predictable and more contingent when similar features are not shared (Blount et al. 2018). Importantly, since many convergent phenotypes are complex, the extent of convergence may differ across levels of biological organization, or “multi-level convergent evolution” (Lau and Oakley 2020). In some cases, phenotypic convergence can be similar at one level of biological organization but radically different at other levels (Rosenblum et al. 2014). As one example, many vertebrates convergently evolved viviparity — the ability to produce live young — using structurally and functionally similar placentas, yet also using non-homologous genes (Foster et al. 2022). For viviparity, structural and functional evolution are more predictable than genetic evolution. In some other cases, convergence extends across multiple levels of biological organization. For example, two squid species convergently evolved bioluminescence using structurally similar light-producing organs that express many genes homologous between the species. Therefore, squid light-producing organs show strong predictability of evolution across multiple levels of organization (Pankey et al. 2014). Fully understanding evolutionary predictability requires comparing convergent phenotypes across multiple biological levels of organization. Furthermore, understanding general rules of evolutionary predictability will require going beyond

individual case studies to determine causes of these patterns — when and why is convergent evolution predictable?

For several reasons, bioluminescence — the biological production of light by a living organism — provides an outstanding system for studying evolutionary predictability. First, bioluminescence convergently evolved at least 94 times across the tree of life (Lau and Oakley 2020). Second, genes used to produce bioluminescence can be reliably tested via *in vitro* functional experiments (Shimomura 2019). Third, across origins of bioluminescence the extent of evolutionary convergence at different levels of biological organization is variable, leading to great diversity in form and function (Haddock et al. 2010; Widder 2010). Bioluminescence is used for many biological functions: deterring predation, attracting prey, camouflaging, intraspecific communication, or courtship (Haddock et al. 2010). Bioluminescence is also produced under diverse conditions. In bioluminescent systems light may be produced intracellularly or extracellularly using a variety of complex structures (Harvey 1952), including glandular organs that may secrete (Dilly and Herring 1978; Takenaka et al. 2012) or house bioluminescence (Morin 1983), modified fin rays that house luminous bacteria (Hulet and Musil 1968), or luminous tubules and ducts (Thacker and Roje 2009). At a molecular level, bioluminescence is produced when enzymes called luciferases oxidize small molecule substrates called luciferins (Shimomura 2019). To harness bioluminescence, luminous organisms may modulate luciferase-luciferin reactions by housing luciferases and luciferins in different cellular or subcellular locations (Clarke et al. 1962; Huvard 1993), or inactivating luciferins by using enzymes

to bind to or biochemically modify them (Lee et al. 1993; Inouye 2007; Viviani et al. 2020). However, compared to knowledge of the function of bioluminescence and luciferase enzymes, little is known about how luciferins are metabolized for storage, transport, or modulation of light.

Fireflies, sea pansies, and cypridinid ostracods convergently evolved the same biochemical mechanism, sulfation — or the transfer of sulfate (SO_3^-) from the donor 3′-phosphoadenosine-5′-phosphosulfate (PAPS) to a substrate (Falany 1997) — for modulating the chemical state of their luciferins (Hori et al. 1972; Nakamura et al. 2014; Fallon et al. 2016; Tzertzinis et al. 2022) (Figure 1, Supplemental Figure S1). Once sulfated, these luciferins are unable to react with their respective luciferases (Fallon et al. 2016), and in the cases of sea pansies and ostracods, sulfation prevents non-specific oxidation of their luciferins (Cormier and Kazuo 1964; Nakamura et al. 2014). Sulfation is catalyzed by a family of enzymes called sulfotransferases that contain a sulfotransferase domain with two highly conserved motifs necessary for binding to PAPS and transferring sulfate to a substrate (Driscoll et al. 1995; Bidwell et al. 1999). Sulfotransferases may exhibit “substrate promiscuity”, which we define as an enzyme’s ability to catalyze the same biochemical reaction in more than one substrate. Sulfotransferase genes underlying luciferin metabolism have been identified and functionally tested in fireflies (Fallon et al. 2016) and recently in sea pansies (Tzertzinis et al. 2022), but not in ostracods. Previous work identified the presence of sulfated luciferin in the ostracod *Vargula hilgendorfii* (Nakamura et al. 2014; Nakamura et al. 2019). This work also found that crude protein extracts from the whole organism could

use PAPS to enzymatically sulfate luciferin — suggesting the presence of a luciferin sulfotransferase — to produce an inactive form of luciferin that could not react with luciferase and was less vulnerable to non-specific oxidation. Additionally, crude protein extracts of *V. hilgendorfii* can remove sulfate from sulfated luciferin to produce active luciferin (Nakamura et al. 2014; Nakamura et al. 2019) Since crude protein extracts from *V. hilgendorfii* can reversibly sulfate luciferin, this suggests sulfation may serve as a mechanism for luciferin metabolism in the organism (Nakamura et al. 2014; Nakamura et al. 2019). However, the genetic sequences of functional luciferin sulfotransferases in ostracods have yet to be identified.

Although fireflies, sea pansies, and cypridinid ostracods convergently evolved the same mechanism for luciferin sulfation, the extent of evolutionary convergence varies across levels of biological organization. These taxa convergently evolved bioluminescence as a defense mechanism — which was later co-opted for courtship in fireflies (Powell et al. 2022) and ostracods (Ellis et al. 2022) — but produce light by using structurally diverse organs or tissue (Peterson and Buck 1968; Spurlock and Cormier 1975; Huvard 1993), different luciferins (Tsarkova 2021), and non-homologous luciferases (Cormier and Kazuo 1964; Oba et al. 2003; Oakley 2005) (Figure 1). Understanding how these taxa genetically encode luciferin sulfation and synthesizing these findings with those from other levels of biological organization — namely, organismal, organ, and tissue — will reveal the extent to which convergently evolved phenotypes may vary across levels of biological organization.

Here, we propose how limited genetic solutions for certain biochemical functions may constrain convergent evolution, which may result in deterministic and predictable genetic evolution across complex traits that evolved convergently. Because bioluminescent ostracods sulfate their luciferin (Nakamura et al. 2014; Nakamura et al. 2019) and because sulfotransferases are the only known genetic mechanism for sulfation, we hypothesize luminous ostracods produce sulfotransferases capable of sulfating ostracod luciferin. To test our hypothesis, we use gene expression analyses, recombinant protein expression, and *in vitro* biochemical assays to identify and functionally test five candidate sulfotransferase genes highly expressed in the light-producing organ of the cypridinid ostracod *Vargula tsujii*. We perform two types of functional tests to show that all five candidates exhibit luciferin sulfation activity as evidenced by at least one assay. Our results, taken together with functional evidence of luciferin sulfotransferases previously published on sea pansies and fireflies (Fallon et al. 2016; Tzertzinis et al. 2022), provide empirical evidence supporting independent recruitment of paralogous genes from a single gene family — sulfotransferases — in the convergent evolution of luciferin metabolism from three phylogenetically distant organisms with distinct luciferin substrates. Because there are no other mechanisms known for sulfation, we assume there are limited genetic solutions, which leads us to propose predictable genetic evolution of luciferin sulfation.

Results

Four candidates highly expressed in the upper lip exhibit significant luciferin sulfotransferase activity in vitro

We identified 40 putative sulfotransferases expressed in the upper lip (the light-producing organ) of *Vargula tsujii*. To identify and functionally verify luciferin sulfotransferase genes in the ostracod *V. tsujii*, we hypothesized that, similar to fireflies (Fallon et al. 2016), sulfotransferases highly expressed in the light-producing organ are capable of sulfating luciferin in a PAPS-dependent manner. We selected the five most highly expressed sulfotransferase genes in the light organ (Figure 2, Supplemental Figure 3), hereafter referred to as ST1-5, to recombinantly express and test for luciferin sulfation function.

A universal sulfotransferase assay kit showed four candidate sulfotransferase genes — ST1, ST3, ST4, ST5 — exhibit significant sulfation activity on luciferin compared to the negative control, which lacked sulfotransferase (ST1: $p = 3.75e-3$, ST2: $p = 9.381e-2$, ST3: $p = 4.9e-7$, ST4: $p = 1.4e-5$, ST5: $p = 1.2e-4$, $N = 6$, Dunnett's test) (Figure 3A). This assay quantifies formation of phosphate from 3'-phosphoadenosine-5'-phosphate (PAP) — a metabolic by-product of sulfotransferases — as a proxy for sulfation activity. By avoiding a buildup of PAP in the reaction volume, this assay also mitigates potential product inhibition, which could limit sulfation activity *in vitro*.

Three candidates exhibit sulfation activity sufficient to modulate bioluminescence in vitro

Three of five candidate luciferin sulfotransferases — ST2, ST3, and ST5 — showed significant decreases in bioluminescence compared to PAPS-free controls (ST1: p =

2.91e-1, ST2: p = 3.79e-2, ST3: p = 2.05e-5, ST4: p = 9.6e-1, ST5: p = 2.16e-5, N = 3, unpaired 2-sample Student's t-test) (Figure 3B), suggesting the activities of ST2, ST3, and ST5 are sufficient to modulate bioluminescence because once sulfated, luciferins are unable to react with their corresponding luciferases to produce light.

Four candidates exhibit significant sulfation activity with p-nitrophenol

Four candidates — ST2, ST3, ST4, and ST5 — catalyzed sulfation of p-nitrophenol relative to PAPS-free controls (ST1: p = 1.45e-1, ST2: p = 1.17e-2, ST3: p = 4.22e-3, ST4: p = 6.91e-5, ST5: p = 2.92e-2, N = 3, unpaired 2-sample Student's t-test) (Figure 4). Since many sulfotransferases sulfate a range of small molecules, we hypothesized ST1-5 may sulfate small molecules besides luciferin. To test this for ST1-5, we examined their ability to catalyze sulfation of p-nitrophenol, a xenobiotic substrate commonly sulfated by cytosolic sulfotransferases. These results support substrate promiscuity in the sulfotransferase activities of ST2, ST3, ST4, and ST5.

Ostracods, fireflies, and sea pansies sulfate luciferin by using homologous genes

Ostracods, fireflies, and sea pansies convergently evolved sulfation as a mechanism for metabolizing luciferin by independently recruiting sulfotransferase genes (Figure 5). Phylogenetic analysis of functionally characterized sulfotransferases place the luciferin sulfotransferases of ostracods and fireflies in a clade of arthropod and mammalian sulfotransferases. This clade contains a mixture of sulfotransferases that may exhibit substrate promiscuity or specificity towards endogenous and xenobiotic small molecules, such as p-nitrophenol. Similarly, the luciferin sulfotransferase of sea pansies

falls in a separate clade, which also contains sulfotransferases with substrate promiscuity or specialized activity towards substrates such as hormones and xenobiotic substrates.

Discussion

Convergently evolved phenotypes often result from a mixture of similar and different features across and within levels of biological organization, providing excellent opportunities to explore general rules of evolutionary predictability. As one general rule, molecular evolution may be more predictable when biochemical functions have fewer possible genetic solutions, and less predictable when genetic redundancy permits many solutions. Here, we examine multi-level convergent evolution of bioluminescence in ostracods, fireflies, and sea pansies. These taxa similarly use bioluminescence to deter predation (Kastendiek 1976; Branham and Wenzel 2003; Ellis et al. 2022), but use structurally different organs and tissues (Peterson and Buck 1968; Satterlie and Case 1979; Huvard 1993) and non-homologous luciferase enzymes to oxidize luciferins for light production (Lorenz et al. 1991; Oba et al. 2003; Oakley 2005). At the same time, these taxa evolved a similar mechanism, sulfation, for luciferin metabolism (Hori et al. 1972; Nakamura et al. 2014; Fallon et al. 2016). We provide functional evidence for sulfotransferases capable of sulfating luciferin in the ostracod *Vargula tsujii*. We synthesize our results with those previously described in fireflies and sea pansies, leading us to propose luciferin sulfation convergently evolved by recruiting paralogous genes from the same gene family. These differences between sulfation recruiting

homologous genes and light production recruiting different genes raises the question: How might biochemical function shape the genetic evolution of complex biological systems like bioluminescence?

The evolutionary convergence of bioluminescence systems demonstrates how function may inform determinism and contingency in molecular evolution. Evolution may be deterministic and predictable when convergently evolved systems share functions constrained to homologous genetic solutions, and may be contingent and less predictable when functions can arise from many non-homologous genetic solutions. For example, since many marine luciferins are intrinsically sensitive to oxygen (Shimomura 2019), their luciferin-oxidizing functions may have convergently evolved using proteins, or “proto-luciferases”, that at first provided hydrophobic environments to facilitate the oxidation of luciferin to produce some light (Rees et al. 1998; Adams and Miller 2020). If increasingly higher amounts of light production continued to enhance survival and reproduction, selection could have led to the evolution of extant luciferases. This hypothesis proposes luciferin oxidation functions may be accomplished by many different genetic solutions, which may help to explain why many luciferases evolved from non-homologous ancestral proteins (Delroisse et al. 2021). In contrast, sulfation in biological systems — as far as we know — is catalyzed by sulfotransferases, which contain two highly conserved motifs used to bind to PAPS and to transfer sulfate to a substrate (Driscoll et al. 1995; Bidwell et al. 1999). Moreover, many sulfotransferases exhibit substrate promiscuity and can sulfate a range of substrates (Allali-Hassani et al. 2007), making these enzymes amenable to evolutionary recruitment. Selection may act

on this promiscuity, enabling the divergence and evolution of new functions (Tokuriki and Tawfik 2009), such as sulfation of luciferin. As such, the evolution of luciferin sulfation may have been constrained to a homologous genetic solution with a predisposition to functional evolution. Taken together, we propose molecular evolution of light production is less predictable due to genetic redundancy, and the evolution of sulfation-mediated luciferin metabolism is more deterministic and predictable because of genetic constraint.

We show the luminous ostracod *V. tsujii* expresses multiple sulfotransferases capable of sulfating luciferin *in vitro*, supporting the hypothesis that sulfotransferases sulfate luciferin in bioluminescent cypridinid ostracods. To test whether there are luciferin sulfotransferases in *V. tsujii*, we performed two types of functional tests for assessing luciferin sulfation activity, tested for significant activity as compared with control reactions, and then corrected for multiple comparisons. The sulfotransferase assay demonstrated significant luciferin sulfation activity for four of the five candidates — ST1, ST3, ST4, and ST5 — as compared with the negative control by using Dunnett's test. The luciferase assay identified three of the five candidates — ST2, ST3, and ST5 — to have significant luciferin sulfation activities capable of modulating bioluminescence *in vitro*. However, after accounting for multiple comparisons by using the Benjamini-Hochberg procedure, only ST3 and ST5 were statistically significant. This discrepancy between the results of the sulfotransferase and luciferase assays could be explained by experimental differences between the two assays. The sulfotransferase assay required a high concentration of luciferin substrate (100 μM) to generate enough product to

produce detectable absorbance measurements. In the luciferase assay, the concentration of luciferin (10 μM) was ten times less than that of the sulfotransferase assay. We found ostracod luciferase exhibits substrate inhibition (Supplemental Figure S6), which limited the concentration of luciferin that could be used in this assay. Since the rate of sulfation reactions depends on substrate concentrations, (Atkinson 1966; James 2021), the differences in substrate concentrations between the two assays may have led to the observed differences in product formation of sulfated luciferin. Inhibition of sulfotransferases by PAP (Chapman et al. 2004), a by-product of sulfation produced in both assays but removed only in the sulfotransferase assay, may have also contributed to the observed differences in sulfation. Nevertheless, results from both assays support the presence of multiple sulfotransferases capable of sulfating luciferin and modulating bioluminescence in the light-producing organ of *V. tsujii*. Future research on the *in vivo* conditions of bioluminescence — such as concentrations of luciferin, luciferase, PAPS, and PAP — may further refine these *in vitro* assays by confirming physiologically relevant testing conditions. These results — taken together with those from firefly and sea pansy bioluminescence systems — support the repeated co-option of sulfotransferases in the convergent evolution of luciferin sulfation, leading us to propose how genetic constraints on functional evolution may result in the predictable and deterministic evolution of this mechanism.

Functional evolution may be predictable if constrained by genetic solutions, but functional specialization of the recruited genes may be unpredictable. Functional evolution does not necessarily require functional specialization of recruited genes

(O'Brien and Herschlag 1999; Khersonsky et al. 2006). New functions may be catalyzed by multi-functional and promiscuous enzymes, or by functionally specialized enzymes. For example, the luciferases of fireflies and click beetles convergently evolved by recruiting homologous genes via gene duplication and neofunctionalization. In fireflies, both copies of the gene retained the ancestral function and one copy gained luciferase function (Oba et al. 2003; Oba et al. 2006). In contrast, click beetles only have one copy of the gene that retained the ancestral function, while the other copy gained and specialized in luciferase function while losing the ancestral function (Oba et al. 2010). In line with this, our findings suggest that, while homologous sulfotransferase genes are recruited in the evolution of luciferin sulfation, functional specialization in these genes may differ across convergent systems. Out of the luciferin sulfotransferases identified by at least one assay, four — ST2, ST3, ST4, and ST5 — can also sulfate p-nitrophenol. Conversely, in fireflies, Fallon et al. (2016) demonstrated only one of the three highly expressed sulfotransferase genes from the light-producing organ was capable of sulfating firefly luciferin and could not sulfate p-nitrophenol. Two other sulfotransferases readily sulfated p-nitrophenol but not firefly luciferin (Fallon et al. 2016). In sea pansies, Tzertzinis et al. showed a luciferin sulfotransferase gene encodes multiple isoforms, two of which can function as luciferin sulfotransferases (Tzertzinis et al. 2022). Therefore, although fireflies might use a functionally specialized luciferin sulfotransferase to sulfate luciferin, ostracods may use many promiscuous luciferin sulfotransferases, and sea pansies may use one or both luciferin sulfotransferase isoforms. To test these hypotheses, further work, such as gene knockout experiments and gene co-expression

analyses, will be necessary for confirming which sulfotransferases are associated with bioluminescence function *in vivo*.

Understanding how complex systems convergently evolved across levels of biological organization may inform general rules of predictability in evolution. The convergent bioluminescence systems of fireflies, sea pansies, and ostracods illustrate how a shared organismal function — using bioluminescence to deter predation — uses structurally different organs and tissues, and a mixture of non-homologous and homologous genes to produce and modulate bioluminescence. Therefore, molecular evolution of bioluminescence may sometimes be more predictable than structural evolution, depending on the range of genetic solutions to a function. These ideas may be broadly extended to the general rule that evolution may be predictable if functions are limited to one or a few genetic solutions and may be far less predictable if redundancy permits many non-homologous genetic solutions.

Materials and Methods

We purchased TRIzol™ reagent, Terrific Broth, and the pCMV-Cypridina Luc vector (encoding ostracod luciferase) from Thermo Fisher Scientific (Waltham, MA). We purchased the pQE80L *E. coli* expression vector and Ni-NTA agarose from Qiagen (Hilden, Germany). The pET-6xHis-SUMO-MarathonRT plasmid (source of SUMO tag) was a gift from Dr. Anna Pyle (Addgene plasmid # 109029). We purchased *Cypridina* (ostracod) luciferin from Prolume LTD (Pinetop, AZ) and prepared single-use aliquots by

dissolving 0.5 mg luciferin in 100 % methanol to a stock concentration of 6.15 mM, then stored the aliquots at -80 °C. We obtained 3′-phosphoadenosine 5′-phosphosulfate (PAPS) lithium salt from Santa Cruz Biotechnology (Dallas, TX), dissolved it in phosphate buffered saline (pH 7.4) to prepare 2 mM aliquots, and stored aliquots at -80 °C. We purchased the sulfotransferase assay kit and lithium-free PAPS substrate from R&D Systems (Minneapolis, MN). We obtained p-nitrophenol from MP Biomedicals (Santa Ana, CA), dissolved it in 20 mM tris hydrochloride (pH 7.4) to prepare a stock solution of 10 mM, and stored the stock at 4 °C. We purchased Chinese hamster ovary (CHO) cell lines from Takara Bio (San Jose, CA), oligonucleotide primers and gBlocks™ from Integrated DNA Technologies (Coralville, IA), and Q5 DNA polymerase, agarose gel extraction kit, reagents for Gibson assembly, competent *E. coli* 10-beta cells and NEB Stable cells (for plasmid propagation), and competent *E. coli* BL21 cells (for protein expression) from New England Biolabs (Ipswich, MA). We purchased reagents for Quick Start™ Bradford protein assay and denaturing polyacrylamide gels (Mini-PROTEAN TGX stain-free protein gels) from BioRad (Hercules, CA), Amicon® ultracentrifugal filter units (30 kDa and 10 kDa molecular weight cutoff) from Sigma Aldrich (St. Louis, MO), and reagents for purifying plasmid DNA from Promega (Madison, WI). We purchased linear polyethyleneimine (25 kDa) transfection reagent from Polysciences (Warrington, PA), dissolved it in deionized water to prepare stock solutions at a concentration of 0.258 mg/mL, and stored stock solutions at -20 °C.

Construction of a high-quality reference transcriptome for Vargula tsujii

We used PacBio full-length isoform sequencing (Iso-Seq) technology to generate a high-quality reference transcriptome for the California luminous ostracod, *Vargula tsujii*. We collected total RNA from 88 *V. tsujii* specimens across various instar stages (carapace lengths, measured from the rostrum to the keel, ranging from 0.68 to 2.02 mm) (Supplemental Table S1), including adult males and brooding females (Supplemental Figure S2). We extracted RNA using TRIzol™ reagent, following the manufacturer′s protocol with two minor modifications. First, we used a higher centrifugation speed (16,100 x g) to pellet cell debris and separate the organic and aqueous layers. Second, we increased precipitation time to ~ 1 hour to increase RNA yield. We shipped purified RNA to the DNA Sequencing Center at Brigham Young University for library preparation and Single Molecule Real-Time (SMRT) sequencing (PacBio Sequel II). To maximize coverage of different sized transcripts, we used two fractions for sequencing: one that was resolved by pulsed field electrophoresis (BluePippin, Sage Science, Beverly, MA) to isolate 4-10 kb transcripts; and another that was not size-selected. We clustered and polished circular consensus sequencing (CCS) reads using the IsoSeq v3 pipeline (https://github.com/PacificBiosciences/IsoSeq). We combined the ensuing long reads with Illumina short reads from NCBI (SRR21201581) (Ellis et al. 2022) using the transcriptome assembler, rnaSPAdes, with the following flags: --rna, --only-assembler, and --trusted-contigs (Bushmanova et al. 2019). To reduce duplicated transcripts, we used cd-hit-est (Li and Godzik 2006; Fu et al. 2012) to cluster sequences with ≥ 95 % identity. We assessed the completeness of the final assembly by using Benchmarking Universal Single-Copy Orthologs (BUSCO) v.3.0.2

(Simão et al. 2015; Waterhouse et al. 2018). BUSCO analysis identified 91.1 % of the total orthologs from the arthropod gene set, 80% of which was classified as complete and 11.1 % fragmented.

Gene expression profiling and identification of sulfotransferase candidates in Vargula tsujii

To identify the most highly expressed sulfotransferase sequences in *Vargula tsujii*, we collected RNA from the light producing organ (known as the upper lip) of ostracod specimens maintained in the Oakley lab culture (Goodheart et al. 2020). We extracted upper lips from nine ostracods (five males and four females) by dissection under a stereomicroscope, immediately mixed them with TRIzol™, and stored the samples at -80 °C until further use. We extracted RNA — using the same extraction protocol as described above, for constructing the Iso-Seq transcriptome — and shipped purified RNA samples to the Genomic Sequencing and Analysis Facility at The University of Texas at Austin for library preparation and gene expression profiling using Tag-Seq (Illumina HiSeq 2500, SR100), resulting in an average of 1.1 million reads (range: 0.61-2.1 million) per sample after filtering (Supplemental Table 2) (Meyer et al. 2011). We processed Tag-Seq reads using a modified version of the Tag-Seq pipeline freely available on GitHub (https://github.com/z0on/tag-based_RNAseq) (Meyer et al. 2011), which mapped the reads to a reference transcriptome — the Iso-Seq transcriptome of *V. tsujii* — using Bowtie2 v.2.3.4.3 (Langmead and Salzberg 2012). On average, 0.87 million reads per sample (range: 0.45-1.5 million) produced unambiguous alignments to the reference transcriptome (Supplemental Table 2). We translated the reference

transcriptome using TransDecoder v.5.5.0 (Haas, BJ. https://github.com/TransDecoder/TransDecoder) and used HMMER v.3.2.1 (hmmer.org) to identify protein sequences containing a sulfotransferase domain (Pfam ID: PF00685) with an e-value score less than or equal to 1e-3. We used MAFFT (v.7.453) (Katoh and Standley 2013) to align protein sequences and used IQ-TREE v.1.6.12 (Nguyen et al. 2015) to infer a maximum likelihood phylogeny, assuming the best-fit substitution model (LG+I+G4) — according to Bayesian Information Criterion — as determined with ModelFinder (Kalyaanamoorthy et al. 2017). To midpoint root the phylogeny and visualize normalized gene expression counts, we used the R package phytools v.1.0.3 (Revell 2012).

Cloning, expression, and purification of ostracod (Cypridina noctiluca) luciferase

To facilitate protein purification, we used Gibson assembly (Supplemental Table S3) to incorporate 3 x FLAG and hexahistidine tags at the C-terminus of an ostracod (*Cypridina noctiluca*) luciferase encoding gene in the pCMV-Cypridina Luc vector (Supplemental Figure S4). We propagated this modified vector by transforming chemically competent NEB Stable *E. coli* cells. After DNA sequencing (Genewiz, South Plainfield, NJ) to confirm successful cloning, we co-transfected CHO cells with this vector and a marker plasmid encoding green fluorescent protein (GFP) by transient transfection of ~ 70% confluent cultures using polyethylenimine. We used fluorescence microscopy (Olympus CKX-41) to confirm transfection success by monitoring the production of GFP, then we harvested the cell culture media — containing secreted ostracod luciferase — 48 hours post transfection, and incubated the media with Ni-NTA

agarose beads in 10 % (v/v) sample buffer (500 mM NaH_2PO_4, 1.5 M NaCl, 100 mM imidazole, pH 8) on a rotary shaker overnight at 4 °C. Subsequently, we loaded the Ni-NTA agarose slurry in a gravity-flow column, collected the flowthrough, and washed the column twice with 3 mL wash buffer (20 mM NaH_2PO_4, 500 mM NaCl, 10 mM imidazole, pH 7.4) to remove nonspecifically bound proteins. Then, we eluted the column twice with 2 mL elution buffer (20 mM NaH_2PO_4, 500 mM NaCl, 500 mM imidazole, pH 7.4). To confirm the presence of ostracod luciferase throughout purification, we added 20 µL of each purification step — namely, the flowthrough, wash, and each elution step — to 50 µL of 20 mM tris hydrochloride pH 8.0. We then tested for luciferase activity by adding 5 µL of 50 µM luciferin to each sample and recorded luminescence using a microplate reader (Tecan Spark) with an integration time of 2 seconds. After confirming the presence of luciferase in our eluate samples (Supplemental Figure S5), we concentrated the eluate and exchanged it into a storage buffer (20 mM tris hydrochloride, 20 % glycerol, pH 8.0) by spin filtration using a 30 kDa molecular weight cutoff filter. We assessed the purity of recombinant luciferase by performing denaturing polyacrylamide gel electrophoresis, visualized the bands by activating the stain-free gels with UV (BioRad Gel Doc), (Supplemental Table S4, Supplemental Figure S5) and quantified protein yield using the Bradford assay. We flash-froze aliquots of recombinant luciferase in dry ice and stored them at -80 °C until use.

Characterizing ostracod (Cypridina noctiluca) luciferase kinetics

To characterize the enzyme kinetics of ostracod luciferase, we added 65 µL of purified ostracod luciferase to 10 µL of luciferin, resulting in a total volume of 75 µL and final

concentrations of 250 nM luciferase and varying final concentrations — 0 μM, 0.39 μM, 0.78125 μM, 1.5625 μM, 3.125 μM, 6.25 μM, 18.75 μM, 37.5 μM, 75 μM, and 150 μM — of luciferin. We incubated the reactions at room temperature for 12 minutes to stabilize the reading and quantified luminescence using a microplate reader (Tecan Spark) with an integration time of 2 seconds (N = 3). Since we observed decreasing light production at final substrate concentrations equal to or greater than 3.125 μM, we plotted and fitted these data using a Michaelis-Menten equation modified for modeling substrate inhibition (Copeland 2000), where v = velocity, V_{max} = maximum velocity, S = substrate concentration, K_m = Michaelis-Menten constant, and K_i = the dissociation constant for the inactive ternary complex formed when another substrate molecule binds to an enzyme-substrate complex (Supplemental Figure S6):

$$v = \frac{V_{max}\,[S]}{K_m + [S]\left(1 + \frac{[S]}{K_i}\right)}$$

Cloning, expression, and purification of Vargula tsujii sulfotransferase candidates

To express and purify candidate luciferin sulfotransferase proteins, we codon-optimized and synthesized DNA sequences corresponding to ST1-5 as gBlocks™ (Integrated DNA Technologies). We used Gibson assembly (Supplemental Table S3) to clone these fragments into a bacterial expression vector, pQE80L, allowing expression of the resulting proteins with an N-terminal hexahistidine tag for purification (Supplemental Figure S4). All plasmid constructs were propagated by electroporation of competent *E. coli* NEB® 10-beta cells, verified by DNA sequencing (Genewiz, South Plainfield, NJ), and subsequently electroporated in competent *E. coli* BL21 cells for protein expression.

Initial efforts to express these proteins in *E. coli* BL21 resulted in soluble expression only for ST3 and ST4 (Supplemental Table S4, Supplemental Figure S7). For the remaining proteins — namely ST1, ST2, and ST5 — we attempted to improve expression by cloning the genes into a different expression vector, pET SUMO, which appended a small ubiquitin-like modifier (SUMO) tag at the N-terminus to improve protein solubility (Butt et al. 2005) (Supplemental Figure S4). Upon re-expression in *E. coli* BL21, the SUMO-tagged constructs showed clearly detectable bands at the expected size in a denaturing gel (Supplemental Table S4, Supplemental Figure S8).

We transformed *E. coli* BL21 cells with the sulfotransferase expressing plasmids and grew cells in 50 mL Terrific Broth at 37 ºC to mid-log phage (optical density at 600 nm ~ 0.4). Then, we induced protein expression by adding 50 µL isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM, moved induced cultures to an orbital shaker at room temperature, and continued protein expression for 18 hours. To harvest bacterial cells, we centrifuged the cell cultures (5000 x g, 10 minutes) and froze cell pellets at -20 °C for 3 hours. We resuspended frozen cell pellets in lysis buffer (20 mM tris hydrochloride, 500 mM NaCl, 30 mM imidazole, 0.5 % tween 20, 5 mM β-mercaptoethanol, pH 7.4), lysed cells by sonication (5 cycles of 30 s on, 10 s off at 30 % amplitude), pelleted cellular debris by centrifuging the crude lysate (12,000 x g, 30 minutes) at 4 °C, and incubated the clarified supernatant with Ni-NTA agarose beads at 4 °C overnight on a rotary shaker. Subsequently, we loaded the Ni-NTA slurry in a gravity-flow chromatography column and purified recombinant protein by washing twice with 2 mL wash buffer (20 mM tris hydrochloride, 500 mM NaCl, 30 mM imidazole pH

7.4), and eluting with 4 mL elution buffer (20 mM tris hydrochloride, 500 mM NaCl, 500 mM imidazole, pH 7.4). The eluates were concentrated and exchanged into a storage buffer (20 mM tris hydrochloride, 300 mM NaCl, 5 mM DTT, 20 % glycerol, pH 8) by spin filtration using 10 kDa molecular weight cutoff filters. We assessed protein purity by running a denaturing polyacrylamide gel, visualizing the bands by activating the stain-free gels with UV (BioRad Gel Doc), and performing densitometry analysis of protein bands on a denaturing polyacrylamide gel (Supplemental Figure S8), then quantified the yield via Bradford assay. Using this approach, we successfully purified all five candidate proteins, achieving > 90 % homogeneity for the purified fractions, as determined by densitometry analysis of protein bands resolved by denaturing gel electrophoresis (Supplemental Table S5).

Assaying for sulfotransferase activity based on PAP formation

We freshly purified and tested sulfotransferase candidates by using a sulfotransferase assay kit (R&D Systems) that examines the formation of 3′-phosphoadenosine-5′-phosphate (PAP), which is released as a byproduct during biochemical sulfation. In this assay, a coupling phosphatase is used to hydrolyze PAP, releasing phosphate, which is quantified by adding malachite green and monitoring the increase in absorbance at 620 nm due to the formation of a complex between malachite green and phosphate. As recommended by the kit, we prepared all reagents in 50 mM Tris, 15 mM magnesium chloride, pH 7.5, hereafter referred to as “assay buffer”. We set up the assay by incubating 5 μL coupling phosphatase with 10 μL ostracod luciferin, 10 μL PAPS, and 25 μL sulfotransferase, resulting in a final volume of 50 μL and working concentrations

of 10 ng/μL coupling phosphatase, 100 μM ostracod luciferin, 200 μM PAPS, and 10 μM sulfotransferase. We set up control reactions — as described above — but omitted PAPS to prevent sulfation activity (Supplemental Figure S9). We included a positive control to confirm the enzyme activity of the coupling phosphatase by incubating 20 μL assay buffer with 5 μL coupling phosphatase (100 ng/μL) and 25 μL PAP (100 μM) in a total volume of 50 μL with working concentrations of 10 ng/μL coupling phosphatase and 50 μM PAP. Importantly, since coupling phosphatase exhibits activity with PAPS and PAP at about a ratio of 1:350, we ran a negative control to measure this background activity by incubating 5 μL coupling phosphatase with 10 μL ostracod luciferin, 10 μL PAPS, and 25 μL of assay buffer, resulting in a final volume of 50 μL and working concentrations of 10 ng/μL coupling phosphatase, 100 μM ostracod luciferin, and 200 μM PAPS. Lastly, we included an assay blank, which contained 50 μL of the assay buffer. We incubated the reactions at room temperature for 3 hours, then measured the formation of PAP — catalyzed by the coupling phosphatase — by adding malachite green and measuring the increase in absorbance at 620 nm with a microplate reader (Tecan Spark). To test whether the sulfotransferase candidates exhibited significant sulfation activity using the luciferin substrate, we used Dunnett's test to compare the absorbance measurements from each of the five experimental groups with the absorbance measurements from the negative control.

In vitro modulation of bioluminescence by luciferin sulfation

We incubated freshly purified sulfotransferase candidates with ostracod luciferin and PAPS, then measured their ability to enzymatically modulate bioluminescence by

incubating the reaction product with ostracod luciferase and recording light production. To set up this experiment, we mixed 10 μL of 10 μM ostracod luciferin and 100 uM PAPS prepared in 20 mM tris hydrochloride pH 7.4, with 1 μL of each purified protein (final concentrations ST1: 15.16 μM, ST2: 21.79 μM, ST3: 51.22 μM, ST4: 31.59 μM, ST5: 24.63 μM). We set up control reactions — as described above — but omitted PAPS to prevent sulfation activity. We incubated the reactions at 25 °C for 3 hours and terminated the reaction by heating at 65 °C for 5 minutes. To measure the decrease in bioluminescence due to the conversion of luciferin to a sulfated form, we cooled and added 90 μL of 20 mM tris hydrochloride pH 7.4 to dilute the reaction products. We then added 10 μL of the diluted reaction product to 55 μL purified ostracod luciferase (final concentration ~ 136 nM), then measured luminescence using a microplate reader (Tecan Spark) with an integration time of 2 seconds. To test whether each candidate sulfotransferase exhibited significant sulfation activity with the luciferin substrate and PAPS, as compared with the control reaction, we performed Student's t-tests. To conclude whether *V. tsujii* contains a sulfotransferase gene with significant sulfation activity towards luciferin, we corrected for multiple comparisons by performing the Benjamini-Hochberg procedure to control the false discovery rate.

In vitro p-nitrophenol sulfation assay

We tested sulfotransferase candidates for their ability to sulfate p-nitrophenol, a common substrate for cytosolic sulfotransferases. To set up this assay, we mixed 25 μL purified luciferin sulfotransferase, 50 μL p-nitrophenol, and 25 μL PAPS resulting in a total reaction volume of 100 μL and working concentrations of 4 μM purified

sulfotransferase, 100 μM p-nitrophenol, and 500 μM PAPS. We set up control reactions — as described above — but omitted PAPS to prevent sulfation. We incubated the reactions at room temperature for 15 hours and monitored p-nitrophenol sulfation by measuring the decrease in absorbance at 405 nm with a microplate reader (Tecan Spark). We used Student's t-test to test whether each candidate sulfotransferase exhibited significant sulfation activity with the p-nitrophenol substrate as compared with the control reaction.

Phylogenetic analysis of functionally tested sulfotransferases

Based on literature searches, we compiled sequence data from functionally tested sulfotransferases (Supplemental Table S6, Supplemental References), and used MAFFT (v.7.453) (Katoh and Standley 2013) to align protein sequences. We inferred a maximum likelihood phylogeny using IQ-TREE (v. 1.6.12) (Nguyen et al. 2015), assuming the best-fit substitution model (LG+R5) — according to Bayesian Information Criterion — as determined with ModelFinder (Kalyaanamoorthy et al. 2017). The phylogeny was midpoint rooted and visualized via the Interactive Tree of Life v6 (Letunic and Bork 2021).

Acknowledgements

We thank members of the Oakley Lab, especially RM Varney and NM Hensley, and CM Pomory for discussion and feedback. This work was supported by National Science

Foundation (NSF) awards (IOS-1754770 and DEB-1457754 to THO), by 2019 (THO) and 2021 (AM) UCSB Academic Senate Faculty Research Grants, a R35 Maximizing Investigators' Research Award from the National Institutes of Health (5R35GM133530 to AM), a Discovery Award from the Department of Defense Congressionally Directed Medical Research Program (PR191154 to AM), and by the US Army Research Office under Cooperative Agreement Numbers W911NF-19-D-0001-0009 and ICB-2019-SSB-12 for the Institute for Collaborative Biotechnologies. ESL was supported by the NSF Graduate Research Fellowship Program (NSF GRFP 1650114) and JAG by the NSF Postdoctoral Fellowship in Biology (BIO PRFB-1711201). The authors acknowledge the use of the Biological Nanostructures Laboratory within the California NanoSystems Institute, supported by the University of California, Santa Barbara and the University of California, Office of the President. Use was made of computational facilities purchased with funds from the National Science Foundation (CNS-1725797) and administered by the Center for Scientific Computing (CSC). The CSC is supported by the California NanoSystems Institute and the Materials Research Science and Engineering Center (MRSEC; NSF DMR 1720256) at UC Santa Barbara.

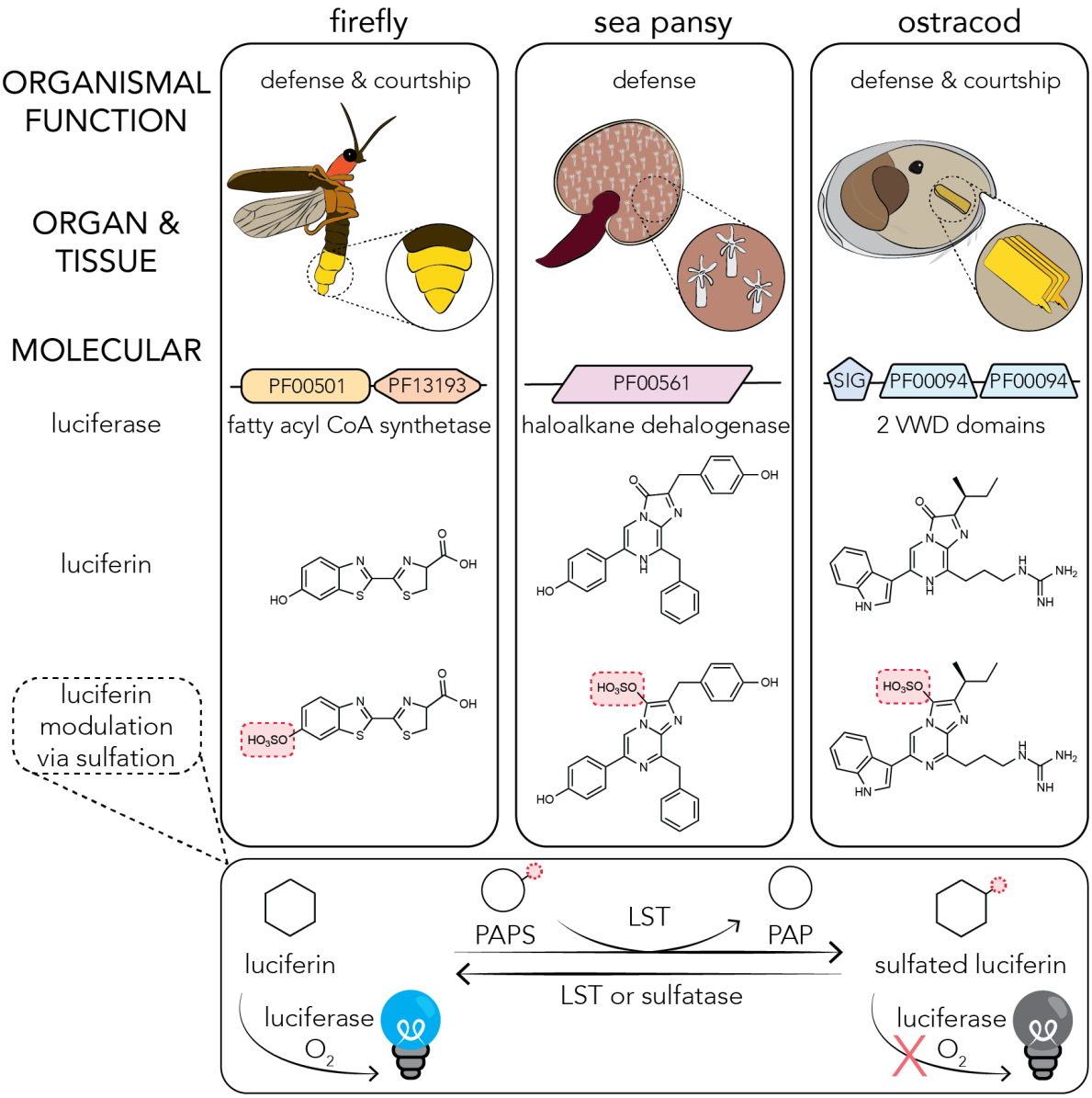
Author Contributions

ESL: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data Curation, Writing - Original Draft and Review & Editing, Visualization. THO: Conceptualization, Resources, Writing - Review & Editing, Supervision, Project administration, Funding acquisition. AM: Methodology, Resources, Writing - Review &

Editing, Supervision, Project administration, Funding acquisition. JAG: Software, Formal analysis, Investigation, Writing - Review & Editing, Resources, Data Curation. NTA: Methodology, Investigation. VLL: Investigation.

Data Availability

All raw sequencing data are available on NCBI (BioProject: PRJNA935772). Sulfotransferase sequences are available on GenBank (Accessions: OQ626349, OQ626350, OQ626351, OQ626352, OQ626353).



1

2 **Figure 1. Convergently evolved bioluminescence systems of fireflies, sea**

3 **pansies, and ostracods may vary in the extent of convergence across and within**

4 **levels of biological organization.** These taxa use bioluminescence as a defense

5 mechanism but produce light by using structurally different organs and tissues, non-

6 homologous luciferases (PF00501 = AMP-dependent synthetase/ligase, PF13193 =

7 AMP-binding enzyme C-terminal domain, PF00561 = alpha/beta hydrolase fold, SIG =

signal peptide, PF00094 = von Willebrand factor type D domain), and different luciferins. However, they use the same mechanism — sulfation — for modulating their luciferins. Luciferin sulfation is catalyzed by a luciferin sulfotransferase (LST), which transfers a sulfate from the donor 3′-phosphoadenosine-5′-phosphosulfate (PAPS) to luciferin and produces 3’-phosphoadenosine-5’-phosphate (PAP) and sulfated luciferin, which is unable to react with their respective luciferase. The sulfate may be removed, either by a LST or sulfatase, to produce the active form of luciferin, which reacts with luciferase to produce light.

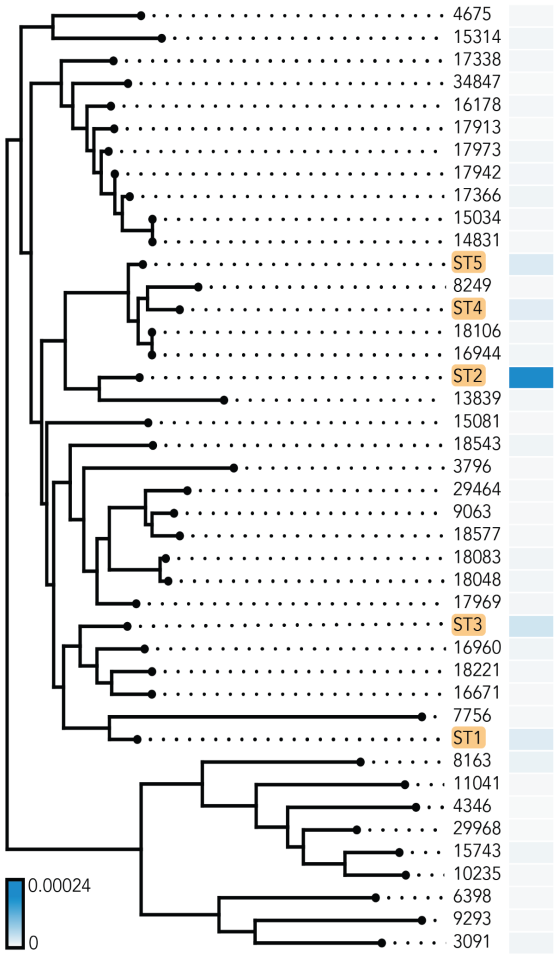


Figure 2. Maximum-likelihood inferred phylogeny and normalized gene expression of sulfotransferase-domain containing proteins in *Vargula tsujii*. We inferred a maximum-likelihood phylogeny using candidate sulfotransferase sequences identified from the Iso-Seq reference transcriptome. The five most highly expressed candidates (labeled ST1-5) were selected for further biochemical testing. The color bar represents the average (N = 9 ostracods, Supplemental Figure S3) of raw transcript counts normalized to the total number of mapped reads in each ostracod sample.

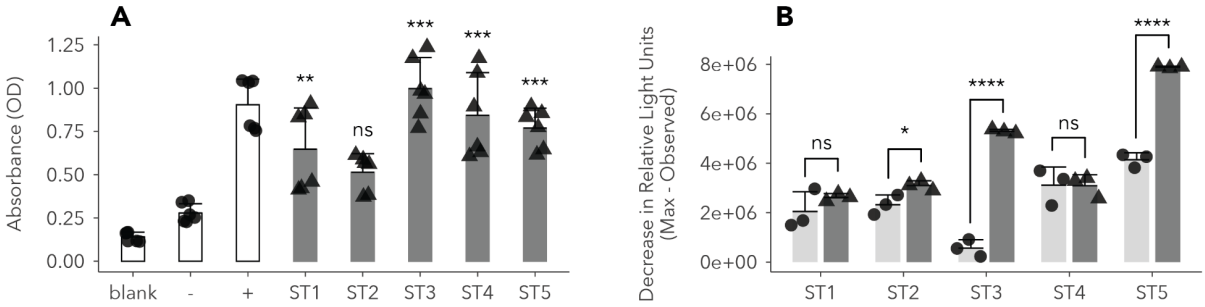


Figure 3. Candidate sulfotransferases from *Vargula tsujii* exhibit luciferin sulfation activity. Sulfotransferase activity was assayed by incubating ostracod luciferin and PAPS with each of the 5 candidate enzymes for 3 hours (circles = control reactions, triangles = experimental reactions). (A) The sulfotransferase assay assesses sulfotransferase activity by using a coupling phosphatase to hydrolyze PAP — a metabolic by-product of sulfation — to form phosphate. At the end of the reaction, malachite green is added to bind free phosphate, forming a complex that can be quantified via absorbance. To quantify background absorbance due to the assay buffer, we included a “blank” measurement. We included a positive (+) control — which consisted of coupling phosphatase and PAP — to confirm the enzyme activity of the coupling phosphatase. Compared to the negative (–) control reactions lacking sulfotransferase, four sulfotransferases — ST1, ST3, ST4, ST5 — showed significant luciferin sulfation activity (N = 6). (B) The luciferase assay uses luminescence to assess luciferin sulfation activity. If a candidate’s luciferin sulfation activity is sufficient to modulate bioluminescence, the product of their sulfation reaction should produce less bioluminescence upon addition of ostracod luciferase. Compared to the control reactions lacking PAPS, three candidates — ST2, ST3, and ST5 — exhibit decreased

light production indicative of sulfation activity sufficient to modulate bioluminescence *in vitro* (N = 3). Each datum is an average of three technical replicates.

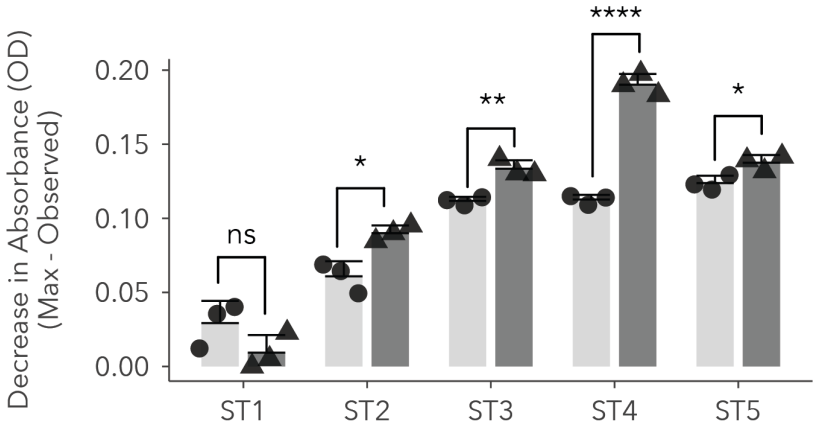


Figure 4. Four candidates sulfate p-nitrophenol. Sulfation of p-nitrophenol, a common substrate for many mammalian sulfotransferases, was assayed by measuring change in absorbance at 405 nm following a 15 hour incubation with purified sulfotransferase proteins in the absence (circles) and presence (triangles) of PAPS. Four candidates — ST2, ST3, ST4, and ST5 — catalyze significant p-nitrophenol sulfation under the conditions tested (N = 3).

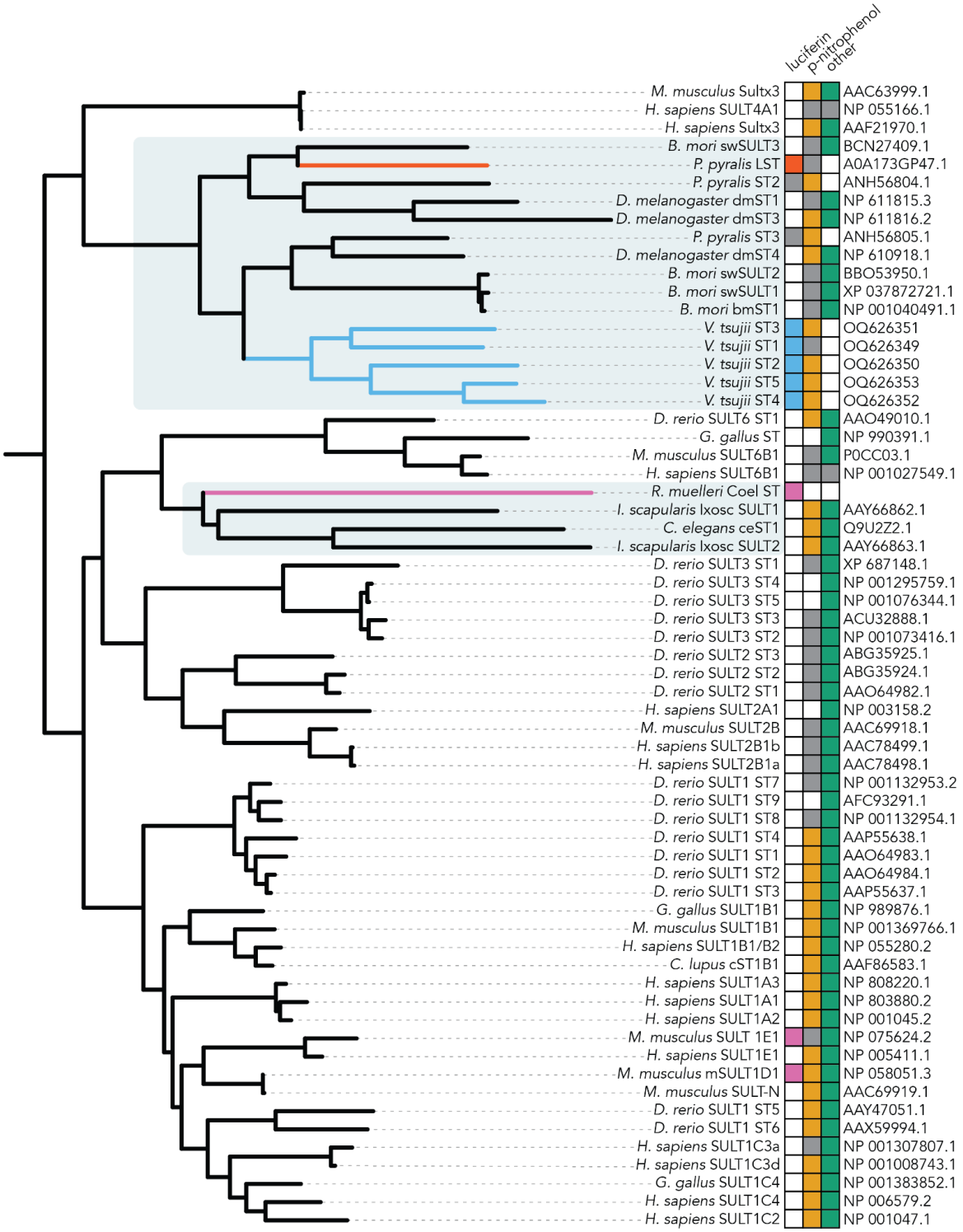


Figure 5. Maximum-likelihood inferred phylogeny of functionally tested sulfotransferases. Luciferin sulfotransferases from luminous organisms are denoted by

colored branches (firefly = red, ostracods = blue, sea pansies = purple). Invertebrate sulfotransferases are highlighted in gray boxes. Colored squares (red = firefly luciferin, blue = ostracod luciferin, purple = sea pansy luciferin, orange = p-nitrophenol, green = other) indicate functionally demonstrated substrates for each sulfotransferase; gray squares indicate molecules that cannot be sulfated by the respective sulfotransferase; white squares indicate molecules that have not yet been tested. Accession numbers for each sulfotransferase are listed on the right.

References

Adams ST Jr, Miller SC. 2020. Enzymatic promiscuity and the evolution of bioluminescence. *FEBS J.* 287:1369–1380.

Allali-Hassani A, Pan PW, Dombrovski L, Najmanovich R, Tempel W, Dong A, Loppnau P, Martin F, Thornton J, Edwards AM, et al. 2007. Structural and chemical profiling of the human cytosolic sulfotransferases. *PLoS Biol.* 5:e97.

Atkinson DE. 1966. Regulation of Enzyme Activity. *Annu. Rev. Biochem.* 35:85–124.

Bidwell LM, McManus ME, Gaedigk A, Kakuta Y, Negishi M, Pedersen L, Martin JL. 1999. Crystal structure of human catecholamine sulfotransferase. *J. Mol. Biol.* 293:521–530.

Blount ZD, Lenski RE, Losos JB. 2018. Contingency and determinism in evolution: Replaying life's tape. *Science* [Internet] 362. Available from: http://dx.doi.org/10.1126/science.aam5979

Branham MA, Wenzel JW. 2003. The origin of photic behavior and the evolution of sexual communication in fireflies (Coleoptera: Lampyridae). *Cladistics* 19:1–22.

Bushmanova E, Antipov D, Lapidus A, Prjibelski AD. 2019. rnaSPAdes: a de novo transcriptome assembler and its application to RNA-Seq data. *GigaScience* [Internet] 8. Available from: http://dx.doi.org/10.1093/gigascience/giz100

Butt TR, Edavettal SC, Hall JP, Mattern MR. 2005. SUMO fusion technology for difficult-to-express proteins. *Protein Expr. Purif.* 43:1–9.

Chapman E, Best MD, Hanson SR, Wong C-H. 2004. Sulfotransferases: structure, mechanism, biological activity, inhibition, and synthetic utility. *Angew. Chem. Int. Ed Engl.* 43:3526–3548.

Clarke GL, Conover RJ, David CN, Nicol JAC. 1962. Comparative studies of luminescence in copepods and other pelagic marine animals*. *J. Mar. Biol. Assoc. U. K.* 42:541–564.

Copeland RA. 2000. Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis. John Wiley & Sons

Cormier MJ, Kazuo H. 1964. Studies on the bioluminescence of Renilla reniformis. *Biochimica et Biophysica Acta (BBA) - Specialized Section on Biophysical Subjects* 88:99–104.

Delroisse J, Duchatelet L, Flammang P, Mallefet J. 2021. Leaving the Dark Side? Insights Into the Evolution of Luciferases. *Frontiers in Marine Science* [Internet] 8. Available from: https://www.frontiersin.org/articles/10.3389/fmars.2021.673620

Dilly PN, Herring PJ. 1978. The light organ and ink sac of Heteroteuthis dispar (Mollusca: Cephalopoda). *J. Zool.* 186:47–59.

Driscoll WJ, Komatsu K, Strott CA. 1995. Proposed active site domain in estrogen sulfotransferase as determined by mutational analysis. *Proc. Natl. Acad. Sci. U. S. A.* 92:12328–12332.

Ellis EA, Goodheart JA, Hensley NM, González VL, Reda NJ, Rivers TJ, Morin JG, Torres E, Gerrish GA, Oakley TH. 2022. Sexual signals persist over deep time: ancient co-option of bioluminescence for courtship displays in cypridinid ostracods. *Syst. Biol.* [Internet]. Available from: http://dx.doi.org/10.1093/sysbio/syac057

Falany CN. 1997. Enzymology of human cytosolic sulfotransferases. *FASEB J.* 11:206–216.

Fallon TR, Li F-S, Vicent MA, Weng J-K. 2016. Sulfoluciferin is Biosynthesized by a Specialized Luciferin Sulfotransferase in Fireflies. *Biochemistry* 55:3341–3344.

Foster CSP, Van Dyke JU, Thompson MB, Smith NMA, Simpfendorfer CA, Murphy CR, Whittington CM. 2022. Different Genes are Recruited During Convergent Evolution of Pregnancy and the Placenta. *Mol. Biol. Evol.* 39:msac077.

Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28:3150–3152.

Goodheart JA, Minsky G, Brynjegard-Bialik MN, Drummond MS, David Munoz J, Fallon TR, Schultz DT, Weng J-K, Torres E, Oakley TH. 2020. Laboratory culture of the California Sea Firefly Vargula tsujii (Ostracoda: Cypridinidae): Developing a model system for the evolution of marine bioluminescence. *Sci. Rep.* 10:1–15.

Haddock SHD, Moline MA, Case JF. 2010. Bioluminescence in the Sea. *Ann. Rev. Mar. Sci.* 2:443–493.

Harvey EN. 1952. Bioluminescence Academic Press. *New York, NY USA* 649.

Hori K, Nakano Y, Cormier MJ. 1972. Studies on the bioluminescence of Renilla reniformis. XI. Location of the sulfate group in luciferyl sulfate. *Biochim. Biophys. Acta* 256:638–644.

Hulet WH, Musil G. 1968. Intracellular Bacteria in the Light Organ of the Deep Sea Angler Fish, Melanocetus murrayi. *Copeia* 1968:506–512.

Huvard AL. 1993. Ultrastructure of the light organ and immunocytochemical localization of luciferase in luminescent marine ostracods (Crustacea: Ostracoda: Cypridinidae). *J. Morphol.* 218:181–193.

Inouye S. 2007. Expression, purification and characterization of calcium-triggered luciferin-binding protein of Renilla reniformis. *Protein Expr. Purif.* 52:66–73.

James MO. 2021. Enzyme Kinetics of PAPS-Sulfotransferase. *Methods Mol. Biol.* 2342:285–300.

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14:587–589.

Kastendiek J. 1976. Behavior of the sea pansy Renilla kollikeri Pfeffer (Coelenterata: Pennatulacea) and its influence on the distribution and biological interactions of the species. *Biol. Bull.* 151:518–537.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30:772–780.

Khersonsky O, Roodveldt C, Tawfik DS. 2006. Enzyme promiscuity: evolutionary and mechanistic aspects. *Curr. Opin. Chem. Biol.* 10:498–508.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat.*

Methods 9:357–359.

Lau ES, Oakley TH. 2020. Multi-level convergence of complex traits and the evolution of bioluminescence. *Biol. Rev. Camb. Philos. Soc.* [Internet]. Available from: https://onlinelibrary.wiley.com/doi/10.1111/brv.12672

Lee DH, Mittag M, Sczekan S, Morse D, Hastings JW. 1993. Molecular cloning and genomic organization of a gene for luciferin-binding protein from the dinoflagellate Gonyaulax polyedra. *J. Biol. Chem.* 268:8842–8850.

Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49:W293–W296.

Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* [Internet] 22:1658–1659. Available from: http://dx.doi.org/10.1093/bioinformatics/btl158

Lorenz WW, McCann RO, Longiaru M, Cormier MJ. 1991. Isolation and expression of a cDNA encoding Renilla reniformis luciferase. *Proc. Natl. Acad. Sci. U. S. A.* 88:4438–4442.

Meyer E, Aglyamova GV, Matz MV. 2011. Profiling gene expression responses of coral larvae (Acropora millepora) to elevated temperature and settlement inducers using a novel RNA-Seq procedure. *Mol. Ecol.* 20:3599–3616.

Morin JG. 1983. Coastal Bioluminescence: Patterns and Functions. *Bull. Mar. Sci.* 33:787–817.

Nakamura M, Matsuda K, Nakamura M, Yamashita K, Suzuki T, Inouye S. 2019. Enzymatic Conversion of Cypridina Luciferyl Sulfate to Cypridina Luciferin with Coenzyme A as a Sulfate Acceptor in Cypridina (Vargula) hilgendorfii. *Photochem.*

Photobiol. 95:1376–1386.

Nakamura M, Suzuki T, Ishizaka N, Sato J-I. 2014. Identification of 3-enol sulfate of Cypridina luciferin, Cypridina luciferyl sulfate, in the sea-firefly Cypridina (Vargula) hilgendorfii. *Tetrahedron* 70:2161–2168.

Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32:268–274.

Oakley TH. 2005. Myodocopa (Crustacea: Ostracoda) as models for evolutionary studies of light and vision: multiple origins of bioluminescence and extreme sexual dimorphism. *Hydrobiologia* 538:179–192.

Oba Y, Kumazaki M, Inouye S. 2010. 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.

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

Oba Y, Sato M, Ohta Y, Inouye S. 2006. Identification of paralogous genes of firefly luciferase in the Japanese firefly, Luciola cruciata. *Gene* 368:53–60.

O'Brien PJ, Herschlag D. 1999. Catalytic promiscuity and the evolution of new enzymatic activities. *Chem. Biol.* 6:R91–R105.

Orgogozo V. 2015. Replaying the tape of life in the twenty-first century. *Interface Focus* 5:20150057.

Pankey MS, Minin VN, Imholte GC, Suchard MA, Oakley TH. 2014. Predictable

transcriptome evolution in the convergent and complex bioluminescent organs of squid. *Proc. Natl. Acad. Sci. U. S. A.* 111:E4736–E4742.

Peterson MK, Buck J. 1968. Light organ fine structure in certain Asiatic fireflies. *Biol. Bull.* 135:335–348.

Powell GS, Saxton NA, Pacheco YM, Stanger-Hall KF, Martin GJ, Kusy D, Felipe Lima Da Silveira L, Bocak L, Branham MA, Bybee SM. 2022. Beetle bioluminescence outshines extant aerial predators. *Proceedings of the Royal Society B: Biological Sciences* 289:20220821.

Rees JF, De Wergifosse B, Noiset O, Dubuisson M, Janssens B, Thompson EM. 1998. The origins of marine bioluminescence: Turning oxygen defence mechanisms into deep-sea communication tools. *J. Exp. Biol.* 201:1211–1221.

Revell LJ. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3:217–223.

Rosenblum EB, Parent CE, Brandt EE. 2014. The Molecular Basis of Phenotypic Convergence. *Annu. Rev. Ecol. Evol. Syst.* 45:203–226.

Satterlie RA, Case JF. 1979. Development of bioluminescence and other effector responses in the pennatulid coelenterate Renilla köllikeri. *Biol. Bull.* 157:506–523.

Shimomura O. 2019. Bioluminescence: Chemical Principles and Methods. World Scientific

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212.

Spurlock BO, Cormier MJ. 1975. A fine structure study of the anthocodium in Renilla

mülleri: Evidence for the existence of a bioluminescent organelle, the luminelle. *J. Cell Biol.* 64:15–28.

Stayton CT. 2015. What does convergent evolution mean? The interpretation of convergence and its implications in the search for limits to evolution. *Interface Focus* 5:20150039.

Takenaka Y, Yamaguchi A, Tsuruoka N, Torimura M, Gojobori T, Shigeri Y. 2012. Evolution of bioluminescence in marine planktonic copepods. *Mol. Biol. Evol.* 29:1669–1681.

Thacker CE, Roje DM. 2009. Phylogeny of cardinalfishes (Teleostei: Gobiiformes: Apogonidae) and the evolution of visceral bioluminescence. *Mol. Phylogenet. Evol.* 52:735–745.

Tokuriki N, Tawfik DS. 2009. Protein dynamism and evolvability. *Science* 324:203–207.

Tsarkova AS. 2021. Luciferins Under Construction: A Review of Known Biosynthetic Pathways. *Frontiers in Ecology and Evolution* [Internet] 9. Available from: https://www.frontiersin.org/articles/10.3389/fevo.2021.667829

Tzertzinis G, Baker B, Benner J, Brown E, Corrêa IR Jr, Ettwiller L, McClung C, Schildkraut I. 2022. Coelenterazine sulfotransferase from Renilla muelleri. *PLoS One* 17:e0276315.

Viviani VR, Silva JR, Amaral DT, Bevilaqua VR, Abdalla FC, Branchini BR, Johnson CH. 2020. A new brilliantly blue-emitting luciferin-luciferase system from Orfelia fultoni and Keroplatinae (Diptera). *Sci. Rep.* 10:9608.

Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, Kriventseva EV, Zdobnov EM. 2018. BUSCO Applications from Quality

1 Assessments to Gene Prediction and Phylogenomics. *Mol. Biol. Evol.* 35:543–548.

2 Widder EA. 2010. Bioluminescence in the Ocean: Origins of Biological, Chemical, and

3 Ecological Diversity. *Science* [Internet] 328:704–708. Available from:

4 http://dx.doi.org/10.1126/science.1174269

5