Multi-level convergence of complex traits and the evolution of bioluminescence

Emily S. Lau[1]* and Todd H. Oakley[1]*

[1]Ecology, Evolution, and Marine Biology, University of California Santa Barbara, Santa Barbara CA 93106
Corresponding authors information:
Emily S. Lau: emily.lau@lifesci.ucsb.edu; Todd H. Oakley: oakley@lifesci.ucsb.edu

Abstract

Evolutionary convergence provides natural opportunities to investigate how, when, and why complex traits evolve. The complexity of convergent traits highlights the importance of explicitly considering convergence on different levels of biological organization, or "multi-level convergent evolution". In order to extend this approach from studying convergence to an integrative workflow for studying complex trait evolution, we propose a holistic and hierarchical framework, which emphasizes breaking down the traits to several functional modules to investigate multi-level convergent evolution. We begin this review by identifying long-standing questions on the origins of complexity and the mechanisms underlying phenotypic convergence to discuss how their research can be enhanced by examining convergent systems. We argue that bioluminescence, a complex trait that evolved dozens of times through either novel mechanisms or conserved toolkits, is particularly well-suited for these studies. We present an updated estimate of at least 84 independent origins of bioluminescence across the tree of life, which was calculated by improving the taxon sampling in cnidarians, fishes, and cephalopods. Then, we use the presented framework to review the biology, chemistry, and evolution of bioluminescence, and for each biological level, identify questions that arise after our systematic review. We focus on luminous organisms that use the shared luciferin substrates coelenterazine or vargulin to produce light because these organisms convergently evolved bioluminescent proteins that use the same luciferins to produce bioluminescence. Lastly, we highlight promising avenues for developing bioluminescence as a model system for studying multi-level convergent evolution.

Key words

multi-level convergence, evolution, bioluminescence, biological organization, complex trait

Contents

I. Introduction

Evolutionary convergence, the independent evolution of a similar phenotype in response to comparable ecological problems, produces biological replicates useful for investigating fundamental questions about the interplay of conservation and divergence during the evolution of complex traits. The classic approach to identifying cases of phenotypic convergence, called ancestral state reconstruction, uses a well-represented, robust phylogeny to map and trace the origin of a particular trait of interest (Joy *et al.*, 2016). In the genomics era, we can extend this approach to examine phenotypic convergence on multiple biological levels (i.e. phenotype, organ, cellular and subcellular, molecular), or "multi-level convergent evolution", especially as research moves towards characterizing complex trait systems at the cellular, subcellular, and molecular levels. Investigating convergent evolution with a hierarchical perspective is necessary because convergent phenotypes may or may not have conserved morphological or molecular mechanisms (Losos, 2011; Stern, 2013; Rosenblum, Parent, & Brandt, 2014). What factors determine whether disparate or conserved mechanisms underlie phenotypic convergence, and how are these mechanisms influenced by genetic constraints (gene pleiotropy or genomic content), or functional/structural constraints (the need to retain certain biochemical activities) (Stern & Orgogozo, 2009; Gompel & Prud'homme, 2009; Christin, Weinreich, & Besnard, 2010; Frankel, Wang, & Stern, 2012)? Recognizing which biological levels of convergence are conserved and what affects their conservation will identify multi-level drivers of predictability in phenotypically convergent traits. Conversely, we can study multiple levels of convergent traits to understand how constraints led to different evolutionary pathways, and how these pathways ultimately converged on similar solutions.

Phenotypic convergence results from underlying genetic mechanisms that are either diverged or conserved. If conserved genetic mechanisms underlie phenotypic convergence, this is considered parallel evolution. It is important to recognize that there are many historical usages of

the term "parallel evolution", the two main ones being phylogenetic and molecular-based definitions. The former describes the independent evolution of a similar phenotype in closely related organisms, while the latter is used to describe the independent origin of a similar phenotype resulting from homologous molecular mechanisms. Although the meaning and usage of the term "parallel evolution" is often debated, in this review we follow the definition presented in Rosenblum et al. 2014 and define parallel evolution as a special case of convergent evolution produced by homologous genetic mechanisms (Scotland, 2011; Rosenblum *et al.*, 2014). Using this definition to distinguish between convergent and parallel evolution will clearly identify cases of phenotypic convergence that rely on the same genetic mechanisms and studying them will provide insight into the drivers and mechanisms of genetic conservation in convergent evolution. Therefore, we consider parallel evolution a subset of convergent evolution, and like convergence, it must be understood in a hierarchical fashion (i.e. regulatory mechanism, gene network, single gene). This hierarchical way of thinking is especially important when studying the convergent evolution of complex traits.

Unlike simple traits, which are controlled by a simple genetic architecture, complex traits (e.g. eyes, venom, and bioluminescence) are multifaceted and functionally integrate many genes and their products (Ogura, Ikeo, & Gojobori, 2004; Fernald, 2006; Haddock, Moline, & Case, 2010; Pankey *et al.*, 2014; Schendel *et al.*, 2019; Arbuckle, 2020). Using a holistic approach to unravel, identify, and evolutionarily trace individual functionally important modules contributing to a complex trait is essential to understanding the evolution of the complex trait as a whole, especially since each module may have different evolutionary histories (Nilsson, 2009; Oakley & Speiser, 2015; Suzuki, 2017). Knowledge of each module's evolutionary history generates hypotheses for when and how separate modules convened to produce extant complexity. However, complex phenotypes may consist of variable modules, and because of their inherent genetic complexity, the phenotypic convergence of complex traits may result from a combination of convergent and parallel genetic mechanisms. Therefore, in order to systematically study the convergent evolution of complex traits, we need to break down complex traits to identify their modules, then examine the multi-level convergent evolution of each module. This strategy merges the holistic approach used to understand complex trait evolution with the hierarchical lens used to study convergent evolution.

In this review, we introduce bioluminescence as an excellent trait for studying the multi-level convergent evolution of complex traits. Then, we use our proposed framework to broadly summarize available information at each biological level for luminous organisms that use convergent molecular mechanisms to produce light. Lastly, we identify gaps in knowledge that should be investigated to develop bioluminescence as a model system for studying convergence. We propose that bioluminescent systems provide exceptionally good opportunities for studying predictability in evolutionary trajectories and the roles of evolutionary and ecological constraints in complex trait convergence.

II. Bioluminescence: an excellent system for studying convergence in complex traits

(1) Harnessing biochemical reactions to produce light

Bioluminescence, the biochemical production of light by a living organism, is generated by proteins called luciferases that facilitate the oxidation of substrates called luciferins. Photoproteins, a distinctive class of luciferase, are protein complexes that bind to luciferin and oxygen prior to initializing the biochemical reaction and require the addition of a cofactor to produce light (McCapra, 1976). Upon biochemical oxidation, oxidized luciferins enter an excited state and emit light as they return to their ground state (Hastings & Wilson, 1976). These biochemical reactions can be produced by the organism itself, called autogenic bioluminescence, or organisms can house and control luminous bacterial colonies in specialized light organs, called bacteriogenic bioluminescence. These biochemical reactions, harnessed by many organisms via autogenic or bacteriogenic bioluminescence, are used for diverse biological functions such as prey luring or mating, predator defence through a startling or sacrificial lure mechanism, or counterillumination, a form of marine camouflage that uses ventrally produced light to match the intensity of downwelling light (Case *et al.*, 1977; Morin, 1983; Young, 1983; Deheyn, Mallefet, & Jangoux, 1998; Harper & Case, 1999; Marek *et al.*, 2011; Gerrish & Morin, 2016; Hellinger *et al.*, 2017). To harness this biochemical reaction in a bioluminescent system, luminous organisms use modules to perform several physiological functions: (i) production/dietary acquisition of a photoprotein/luciferase, (ii) production/dietary acquisition of a luciferin, (iii) maintenance of an organ used to produce bioluminescence or house luminous bacterial symbionts, (iv) behavioural control of light production, (v) catabolism and/or recycling of bioluminescent molecules, and in some cases, (vi) storage of bioluminescent substrates and (vii) modification of light emission using biochemical or physical mechanisms (e.g. green fluorescent proteins, filters, chromatophores) (Denton *et al.*, 1985; Thomson, Herring, & Campbell, 1997; Wilson & Hastings, 1998; Gomi, Hirokawa, & Kajiyama, 2002; Tong *et al.*, 2009; Bessho-Uehara *et al.*, 2020).

However, not all bioluminescence systems have all these functions, and the modules used to achieve these functions can vary (Figure 1). For example, organisms produce autogenic bioluminescence by using proteins that evolved from the parallel evolution of conserved genes (e.g. calcium binding photoproteins in ctenophores and cnidarians), or non-homologous genes (e.g. luciferases in copepods and sea pansies) (Prasher, McCann, & Cormier, 1985; Lorenz *et al.*, 1991; Tsuji *et al.*, 1995; Markova *et al.*, 2004, 2015). What ecological and evolutionary factors determine whether disparate or parallel genetic mechanisms will underlie convergence in bioluminescent systems' modules? Does the conservation or divergence in modules used for the same functional categories extend across multiple biological levels? Bioluminescence is a truly exceptional convergent system that is well-suited for these questions because it is a complex trait that was re-invented dozens of times by using a mix of conserved, similar, and radically diverse ancestral genes.

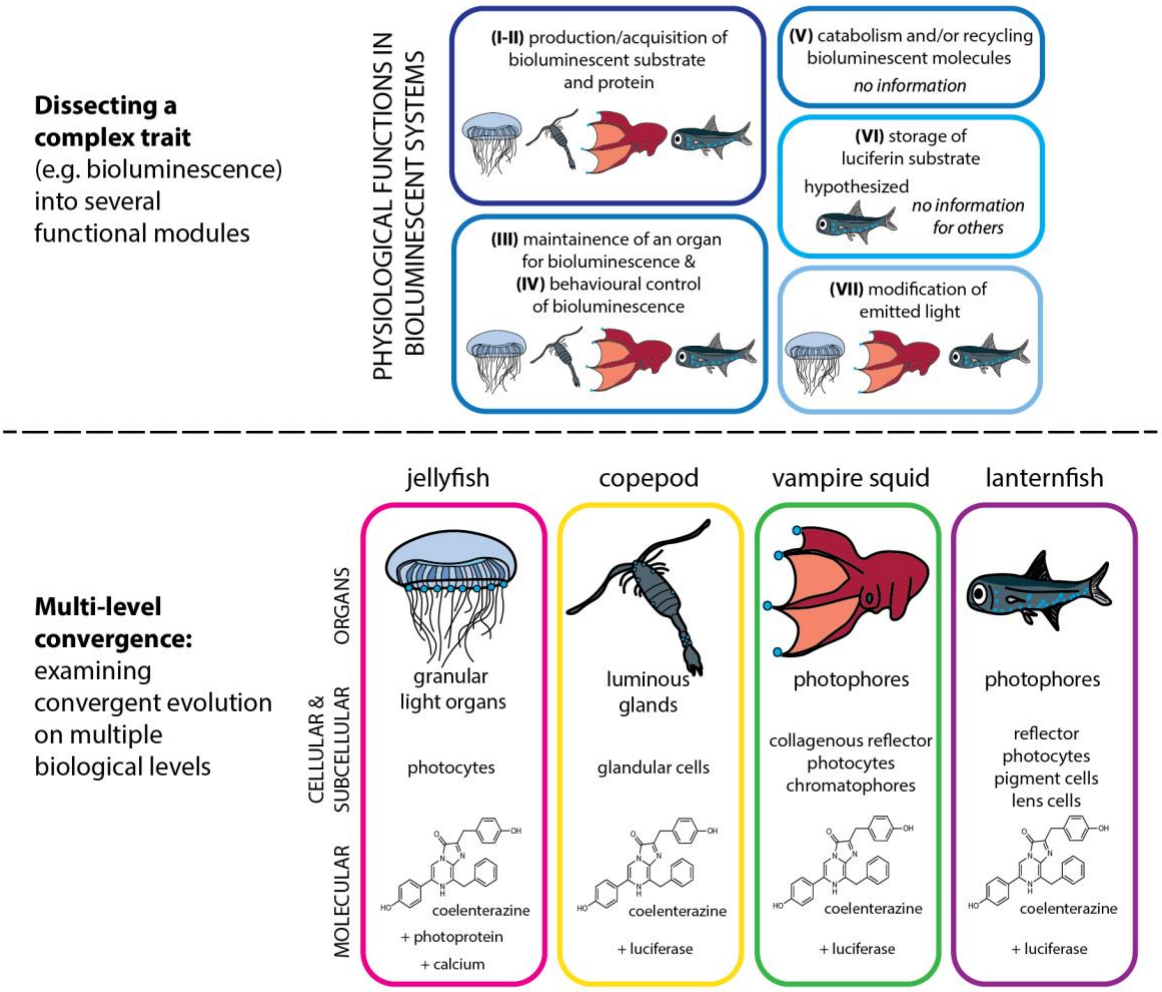

Figure 1. A systematic approach to investigating complex trait convergence involves dissecting a complex trait into several functional modules (top) and examining convergence on all levels of life (bottom). Luminous organisms may or may not share several functional modules to harness the ability to produce light. Although jellyfish, copepods, vampire squids, and lanternfishes independently evolved molecularly convergent proteins that produce light using the same luciferin substrate, this convergence is not necessarily found on other biological levels.

(2) Widespread convergent evolution of bioluminescence

Bioluminescence is found in freshwater, terrestrial, and marine environments, and independently evolved dozens of times across the tree of life. Haddock et al.'s review on convergence in bioluminescence, published in 2010, estimated that bioluminescence evolved at least 40 times, and likely more than 50 times across lifeforms (Haddock *et al.*, 2010). In their review, Haddock et al. also suggested that the number of independent origins may increase upon increasing the resolution of bioluminescence character mapping in ray-finned fishes and

cephalopods; the former was done later in 2016 by Davis et al. (Davis, Sparks, & Smith, 2016). We estimated that bioluminescence originated at least 10 times in cephalopods by using a genus-level cephalopod phylogeny produced by Sanchez et al. (E.S. Lau, B.A. Vincent, T.H. Oakley, unpublished data) (Sanchez *et al.*, 2018). Recently, Bessho-Uehara et al. examined the bioluminescent systems of deep sea Anthozoans and estimated at least six origins of bioluminescence in Cnidaria; three in Anthozoa, two in Scyphozoa, and one in Hydrozoa (Bessho-Uehara, M., Francis, & W.R., Haddock, S.H.D., 2020).

Calculating independent origins can be difficult because ancestral state reconstruction can overestimate or underestimate independent origins, depending on the assumptions used to describe the probability of character gain and loss. Assuming equal probabilities of gain and loss, ray-finned fishes evolved bioluminescence 29 times. However, if the cost to evolve bioluminescence is weighted two or three times, the origins are estimated to be 27 and 22, respectively. A bare minimum for the number of origins in fishes is two (bacteriogenic and autogenic), but that would imply an unreasonable number of over 120 losses. Aside from difficulties in calculation, defining an independent origin may not be straightforward because many luminous organisms do not produce their own luciferin and/or bioluminescent protein (i.e. organisms that rely on bioluminescent bacterial symbionts or organisms that acquire luciferins and/or bioluminescent proteins through their diet). Despite these caveats, Haddock et al. and Davis et al. considered the aforementioned cases in their estimates because these organisms evolved modules used for regulating bioluminescence (i.e. light organ, behavioural control of bioluminescence, modification of light emission). By following this line of reasoning and expanding on the taxon sampling in cnidarian, cephalopods and ray-finned fishes tips of Haddock et al.'s previous supertree, we estimate at least 84 independent origins of bioluminescence across the tree of life (Figure 2).

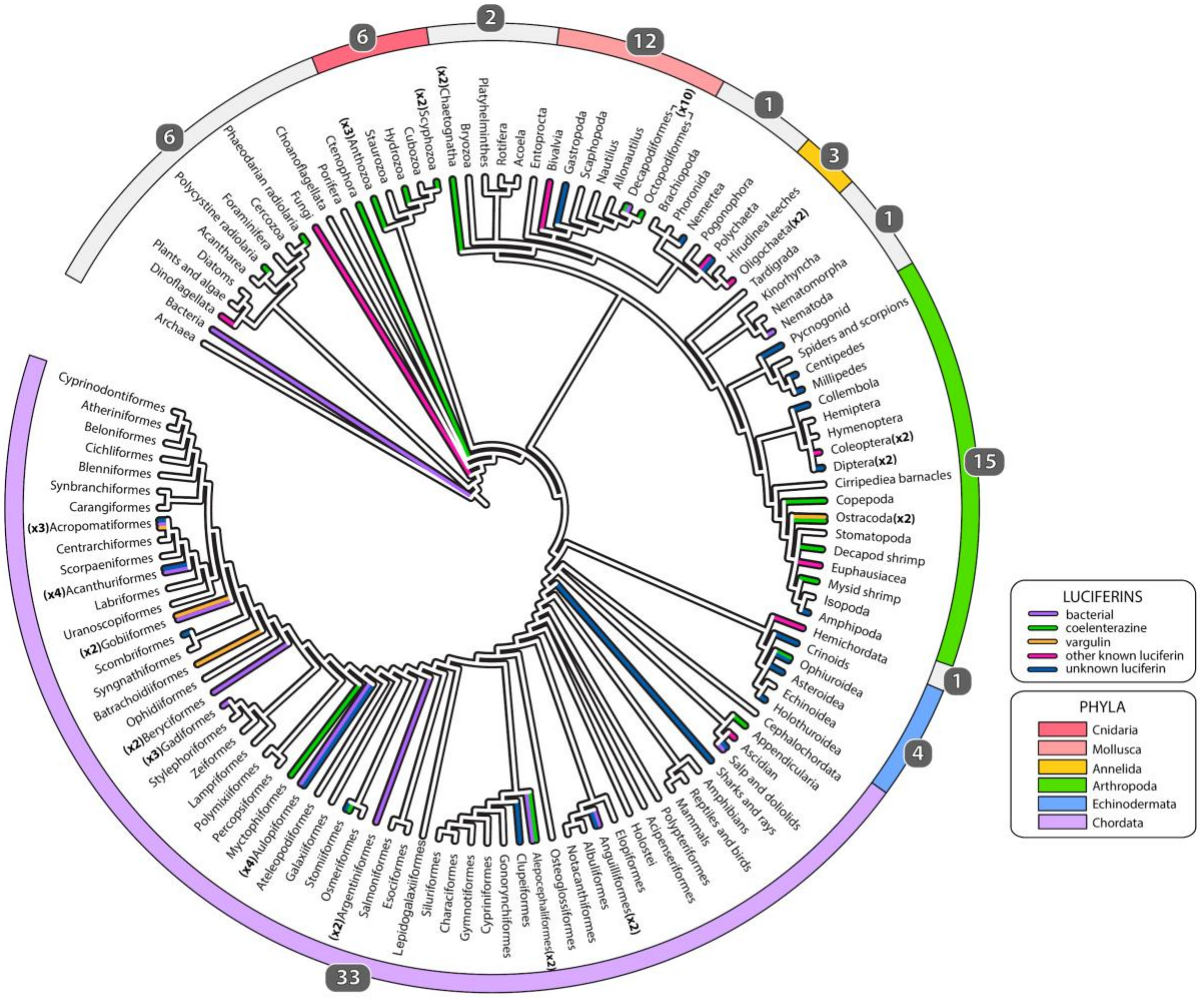

Figure 2. Tree showing at least 84 independent origins of bioluminescence
Numbers in the grey boxes represent the number of independent origins within each arc. Coloured arcs group the tips falling under different phyla. Grey branches represent non-luminous taxa and coloured branches indicate luminous taxa, with different branch colours corresponding to different luciferins. Many taxa use more than one type of luciferin, as shown by the multi-coloured branches. Numbers next to taxon names indicate the number of origins within that taxon. The supertree topology is based on previously published phylogenies (Dunn *et al.*, 2008; Haddock *et al.*, 2010; Davis *et al.*, 2016; Sanchez *et al.*, 2018)

Bioluminescence systems use many different types and combinations of organs, cells, and biochemistries (Sweeney, 1980; Herring, 1987, 2000; Shimomura, 2006). The diversity of biological levels found within this complex trait is precisely what makes bioluminescence an excellent system for examining how conservation and divergence shape the evolution of complex traits. While bioluminescence can be convergent on one biological level, this convergence may not be found in other levels (Figure 3). For example, bioluminescence is phenotypically

convergent, but its usage as a behavioural mechanism (e.g. mating, defence, prey attraction) may not be convergent. Secreted bioluminescence, which is used for mating and defence, is discharged from light organs such as modified mouths, photophores, and appendages (Angel, 1968; Nicol, 1969; Barnes & Case, 1972; Abe *et al.*, 2000; Robison *et al.*, 2003; Wong *et al.*, 2015). Alternatively, internal extracellular and intracellular bioluminescence, which is used for counterillumination, prey capture, mating, and communication, is produced in light spots associated with the epithelium, light organs connected to the digestive tract, luminous lanterns, or structurally complex photophores (Peterson & Buck, 1968; Arnold & Young, 1974; Baguet, 1975; Haddock & Case, 1999; Thacker & Roje, 2009; Galeazzo *et al.*, 2019).

We see divergence and conservation on an organ level, because the biological function of bioluminescence may be constrained by the types of organs associated with a particular behaviour. On a cellular level, we see divergence and conservation of the cell types found in light organs. Furthermore, convergent organ types do not necessarily contain morphologically similar cell types. For example, while photophores in the viperfish *Chauliodus* contain secretory photocytes that are disparate from photocytes found in the midshipman fish *Porichthys*, the photocytes found in the squids *Watasenia scintillans* (Berry, 1911) and *Bathothauma lyromma* (Chun, 1906) contain crystal rods that are structurally similar to the ones found in krill photophores. (Sweeney, 1980; Herring, Dilly, & Cope, 1994). It is important to recognize that past bioluminescence research historically grouped these “cell types” based on morphological features or functions (i.e. subcellular content, production of light). With advances in genetic sequencing on a cellular level, which we highlight later in this review, we can genetically define a cell type and produce cell type phylogenies to understand how cells used in bioluminescent systems evolved.

At the molecular level (i.e. genes, proteins, substrates), we may see divergence, convergence, and in some cases, parallelism of the biochemical components used to produce light (e.g. luciferin, luciferase/photoprotein). Organisms that use different luciferins to produce light have non-convergent bioluminescent proteins. On the other hand, organisms that use identical luciferins for bioluminescence have molecularly convergent bioluminescent proteins. The molecular convergence in bioluminescent proteins can result from parallel evolution or convergent evolution (the evolution of homologous proteins in the former case, or non-homologous proteins in the latter case). Despite many cases of molecular convergence, most known bioluminescent proteins exhibit wide molecular diversity and are non-homologous across distantly related taxa, which suggests that most origins of bioluminescent proteins are the result of convergent, but not parallel, evolution (Hastings, 1983; Rees *et al.*, 1998; Oba, Branham, & Fukatsu, 2011).

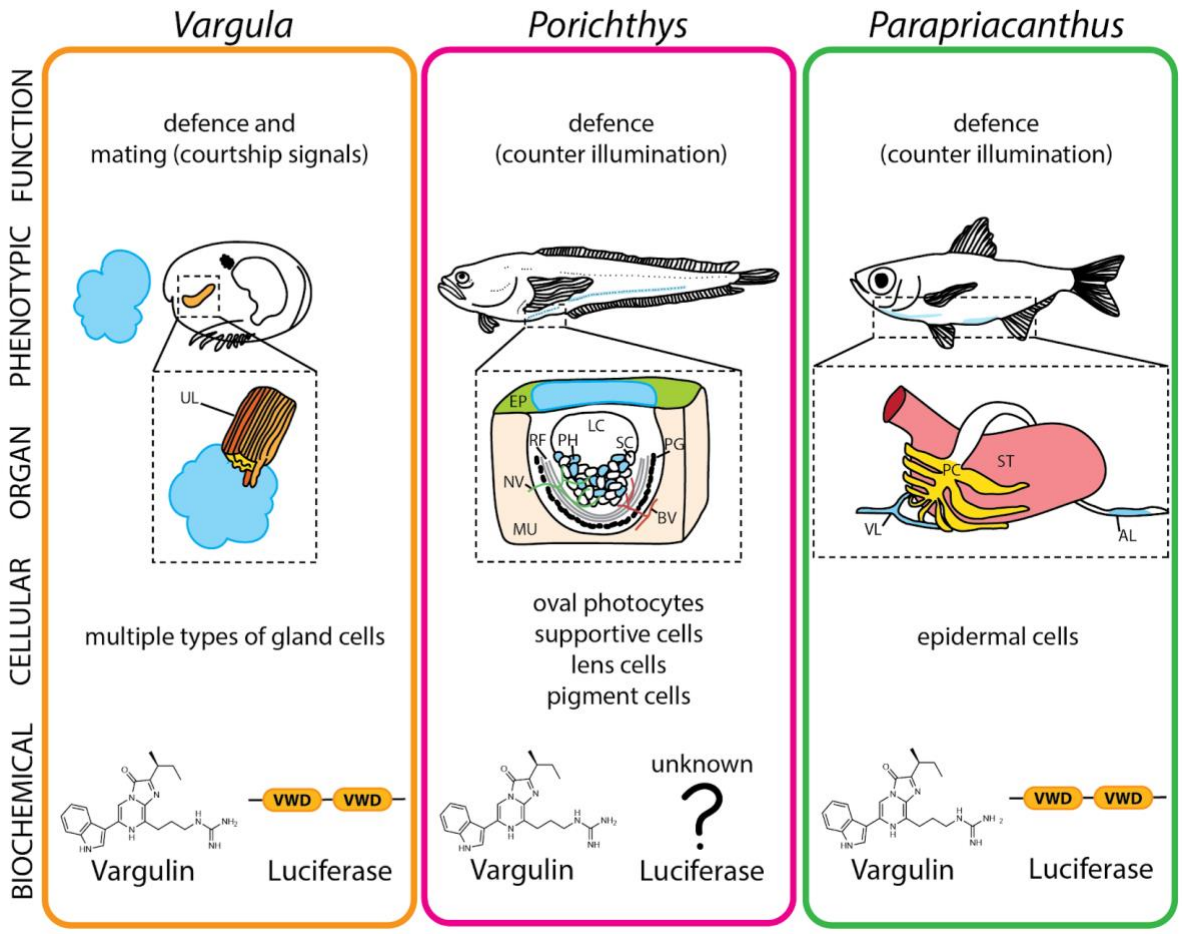

Figure 3. Cypridinid ostracods (*Vargula*), midshipman fishes (*Porichthys*), and golden sweepers (*Parapriacanthus*) all use vargulin to produce light, but these organisms rely on different biological components (upper lip, photophore, and visceral light organs, respectively) to produce bioluminescence (exemplified in blue). The methods used to produce bioluminescence (secreted or intracellular) and its function (defence and/or mating) varies between ostracods and the fishes. Labels: *Vargula* UL=upper lip; *Porichthys* EP=epidermis, MU=muscle, NV=nerves, BV=blood vessel, PG=pigment cell, SC=supportive cell, RF=reflector, PH=photocyte, LC=lens cells; *Parapriacanthus* ST=stomach, PC=pyloric caeca, VL=ventral light organ, AL=anal light organ

Unlike the diversity of luciferases and photoproteins, there are only 11 known luciferins and one untested luciferin candidate for the New Zealand glowworm (Figure 4) (Shimomura & Johnson, 1968; Shimomura, 2006; Petushkov *et al.*, 2014; Purtov *et al.*, 2015; Watkins *et al.*, 2018; Kotlobay *et al.*, 2019). The genetic bases responsible for luciferin biosynthetic pathways are only known for bacterial and fungal luciferins. Bacteria produce their luciferin using the lux operon, which contains genes used to produce both subunits of bacterial luciferase in addition to bacterial luciferin (Meighen, 1991). Similarly, the genes involved in the biosynthesis of fungal luciferin are found in a genomic cluster, along with fungal luciferase (Kotlobay *et al.*, 2018). Although the

biosynthetic pathways responsible for coelenterazine and vargulin are unknown, the amino acids responsible for their formation are phenylalanine/tyrosine/tyrosine and arginine/isoleucine/tryptophan, respectively (Kato *et al.*, 2004; Oba *et al.*, 2009).

The chemical structures of most luciferins are distinct from each other. The only luciferins that share a part of their chemical structure are coelenterazine and vargulin, and dinoflagellate luciferin and krill luciferin (Nakamura *et al.*, 1989; Campbell & Herring, 1990; Shimomura, 2006). Coelenterazine and vargulin share an imidazopyrazinone core, which is composed of a fused bicyclic ring containing four nitrogens and a ketone; however, these luciferins cannot be used interchangeably. Dinoflagellate and krill luciferins, molecules that are classified as tetrapyrroles, are almost structurally identical. The structural similarities observed in krill and dinoflagellate luciferins led to the formation of a hypothesis (not yet experimentally supported) suggesting that krill obtain their luciferin through their diet. Additionally, krill and dinoflagellate luciferins are structurally similar to chlorophyll, a molecule that is hypothesized to be the precursor for dinoflagellate luciferin (Dunlap, Hastings, & Shimomura, 1981; Nakamura *et al.*, 1988, 1989; Shimomura, 1995). Except for these two examples, luciferins have diverse chemical structures, thus showing that the independent evolution of luciferins was not constrained to one type of small molecule (Hastings, 1983).

Although most luciferins are used by the taxa in which they are found, the luciferins coelenterazine and vargulin are used as bioluminescent substrates in phylogenetically distant organisms. Coelenterazine is the most common marine luciferin, shared across eight phyla, while vargulin is shared across two phyla (Tsuji *et al.*, 1971; Rees *et al.*, 1990; Thomson *et al.*, 1997; Haddock *et al.*, 2010; Markova & Vysotski, 2015). Despite the prevalence of coelenterazine-based bioluminescence, most organisms that use it do not produce it themselves. Because the only organisms known to have coelenterazine biosynthesis capabilities are deep sea copepods, ctenophores, and decapod shrimps (the former determined through tracing stable isotopes and latter two determined by observing either an increase in luciferin or bioluminescence in embryos), many other organisms that use coelenterazine for bioluminescence probably obtain it through their diet (Thomson, Herring, & Campbell, 1995; Oba *et al.*, 2009; Francis *et al.*, 2015). This is supported by experimental evidence, which show that many bioluminescent organisms reared in the laboratory on a coelenterazine-free diet gradually lose the ability to luminesce, but can regain bioluminescence activity upon consuming food items containing coelenterazine (Frank, Widder, & Case, 1984; Haddock, Rivers, & Robison, 2001; Mallefet, Duchatelet, & Coubris, 2020). Furthermore, coelenterazine can be detected in digestive tracts, stomachs, and livers of both luminous and non-luminous organisms, suggesting that coelenterazine is readily available in the marine food web (Shimomura, 1987; Rees *et al.*, 1990; Mallefet & Shimomura, 1995). Similarly, only ostracod crustaceans are known to produce vargulin. Luminous fishes rely on the dietary acquisition of this substrate, and some fishes can give it to their young through maternal provisioning (Haneda & Johnson, 1958; Tsuji *et al.*, 1971; Tsuji, Barnes, & Case, 1972; Duchatelet *et al.*, 2019). In addition to dietary acquisition of luciferin, a recent discovery found that the luminous fish *Parapriacanthus ransonnetti* (Steindachner, 1870) acquires both vargulin and

luciferase from bioluminescent ostracods. This phenomenon, known as "kleptoprotein bioluminescence", is a novel mechanism used in autogenic bioluminescence, which produces light by using a bioluminescent protein that is sequestered from the fish's diet (Bessho-Uehara *et al.*, 2020).

The rampant dietary acquisition of the luciferins coelenterazine and vargulin suggests that many luminous organisms do not produce their own luciferins. Of the Metazoa hypothesized to produce their own luciferins, all marine groups (i.e. cypridinid ostracods, ctenophores, decapod shrimps, deep sea copepods, syllid polychaetes) and two terrestrial groups (i.e. *Diplocardia* earthworms, *Latia* freshwater snails) contain taxa (some if not all) that are capable of secreting bioluminescence into the environment (Herring, 1985; Widder, Greene, & Youngbluth, 1992; Verdes & Gruber, 2017). Is the evolution of these two characters, luciferin production and bioluminescence secretion, correlated? This hypothesis could be tested by running a phylogenetic correlation test for the two binary characters (luciferin production and secretory bioluminescence) (Pagel & Meade, 2006). As exemplified here, we can use a holistic approach to examine bioluminescence systems, hypothesize how they evolved, and better understand how luminous systems' functional modules evolved and integrated.

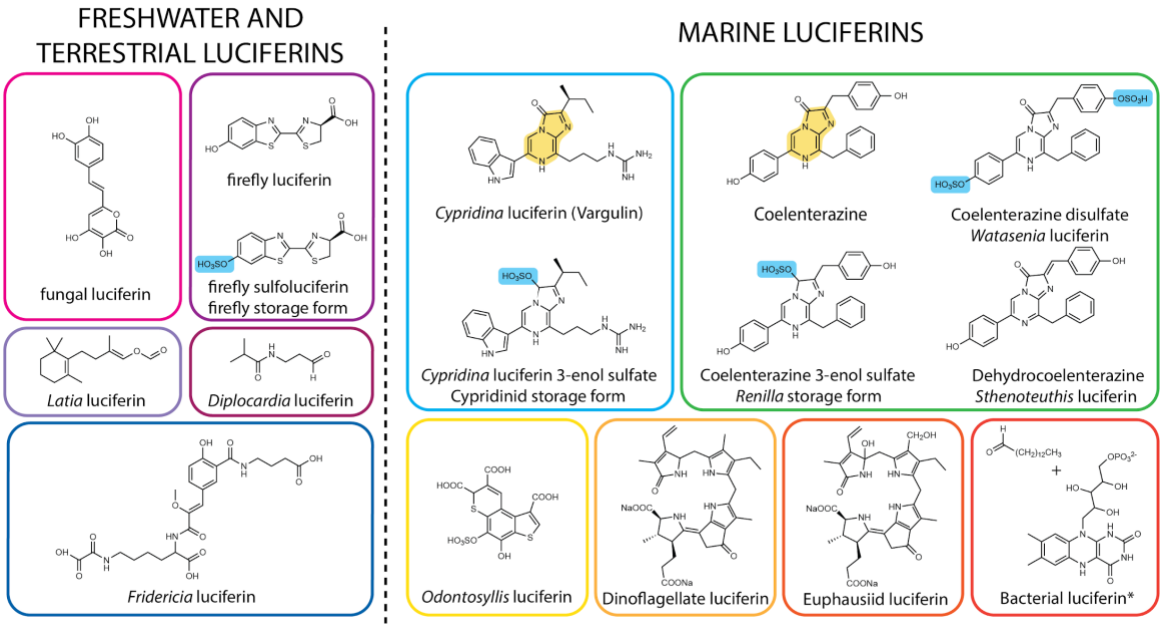

Figure 4. Chemical structures of known luciferins, their derivatives, and their storage forms Vargulin and coelenterazine share imidazopyridine cores (highlighted in yellow). Firefly sulfoluciferin and the 3-enol sulfated vargulin and coelenterazine are proposed to be luciferin storage forms that can be enzymatically converted to active luciferins using a sulfotransferase (sulfates highlighted in blue). *Sthenoteuthis oualaniensis* (Lesson & Duperrey, 1830) and *Watasenia scintillans* use modified coelenterazine as their luciferins.

*In addition to marine systems, luminous bacteria are responsible for nematode bioluminescence and have a symbiotic relationship with these luminous nematodes (Thanwisai *et al.*, 2012)

(3) Evolutionary origins of bioluminescence

The ubiquitous nature of bioluminescence, coupled with the biochemical diversity found across bioluminescent systems, suggests that the evolutionary origins of bioluminescence are disparate. Like other complex traits, understanding the origin and evolution of bioluminescence requires dissecting and investigating each functional module used in this trait. In the case of the module responsible for the biochemical production of light, studying the evolutionary histories of luciferins and photoproteins/luciferases will provide testable hypotheses for how these substrates and proteins evolved across bioluminescent systems. For example, in luminous fungi and bacteria, genes for luciferin-synthesis and luciferase-production are found in a genomic cluster (Meighen, 1991; Kotlobay *et al.*, 2018). Does this pattern of genetic organization extend to luminous metazoans? How did these genes and their corresponding products originate and evolve?

Although the functional modules in bioluminescent systems include more than just the ones used to biochemically produce light (luciferin, luciferases/photoproteins), an approach that emphasizes studying a complex trait's functional modules is historically constrained by earlier bioluminescence research that focused on characterizing the biochemistry of light production in luminous systems. Now, owing to advancements in specimen collection, imaging, and sequencing, we can begin to understand the diversity of modules used to perform the physiological functions needed in this complex trait (i.e. maintaining light organ(s), controlling light production, light modification, substrate storage).

> "Apparently there is no rhyme or reason in the distribution of luminescence throughout the plant or animal kingdom. It is as if the various groups had been written on a blackboard and a handful of sand cast over the names. Where each grain of sand strikes, a luminous species appears" - E. Newton Harvey

E. Newton Harvey, a pioneer in bioluminescence research, observed that bioluminescence is widespread but questioned why it is seemingly scattered randomly across lifeforms (Harvey, 1920). Is bioluminescence truly distributed haphazardly across the tree of life? Why do some groups have more luminous species than others? One hundred years later, we can propose hypotheses to these questions by identifying and comparing selective pressures and possible evolutionary routes leading to bioluminescence. By exploring these routes in luminous and non-luminous organisms, we can test whether certain organisms are predisposed to bioluminescence and hypothesize why we see this scattered phylogenetic distribution of bioluminescence.

(a) Origin of luciferases and photoproteins

Most luciferases and photoproteins are molecularly disparate but produce bioluminescence by playing analogous roles as substrate binders and oxidation facilitators. How did

luciferases/photoproteins evolve and in most cases, how can non-homologous proteins produce light using the same substrate? Is the oxygenase function of a photoprotein or luciferase ancestral or derived? One hypothesis proposes that luciferins were non-specifically oxidized by pre-existing proteins. This promiscuous protein hypothesis postulates that proto luciferases/photoproteins may either function as an oxygenase and/or provide a hydrophobic environment for the oxidation of luciferin (Seliger, 1975; Hastings, 1983; Rees *et al.*, 1998; Vassel *et al.*, 2012; Adams & Miller, 2020). The latter can be accomplished by intrinsic protein pockets, or a protein cage effect, which is hypothesized to provide an isolating environment to prevent the quenching of the luciferin's excited state (Rees *et al.*, 1998). The hydrophobic environment hypothesis is supported by firefly luciferase, which evolved from an ancestral fatty acyl-CoA synthetase, thus showing that its oxygenase abilities are not ancestral but derived (Oba, Ojika, & Inouye, 2003; Oba *et al.*, 2006).

Those ancestral proteins that facilitated luminescence did not necessarily have to produce the full bioluminescence that we see today, but if they offered an evolutionary advantage, selection may have acted to improve substrate binding efficiency and light production. These hypotheses are supported by the diversity of extant, non-homologous photoproteins and luciferases, which demonstrates that their convergent evolution is not constrained to one protein family (Figure 5) (Gould, Keller, & Subramani, 1987; Abe, Nagata, & Hashizume, 1996; Oba *et al.*, 2003, 2006; Oakley, 2005; Loening, Fenn, & Gambhir, 2007; Markova *et al.*, 2012; Schnitzler *et al.*, 2012; Delroisse *et al.*, 2017; Brugler *et al.*, 2018; Fallon *et al.*, 2018; Kotlobay *et al.*, 2018; Schultz *et al.*, 2018). The origins of luciferases/photoproteins and their ancestral functions are mostly unknown. Understanding the origin and evolution of these proteins involves tracing bioluminescent proteins' evolutionary histories, performing ancestral protein sequence reconstruction, then expressing and characterizing the function of ancestral proteins (Takenaka *et al.*, 2013). Ultimately, this workflow will reveal how diverse luciferases and photoproteins functionally converged to produce light.

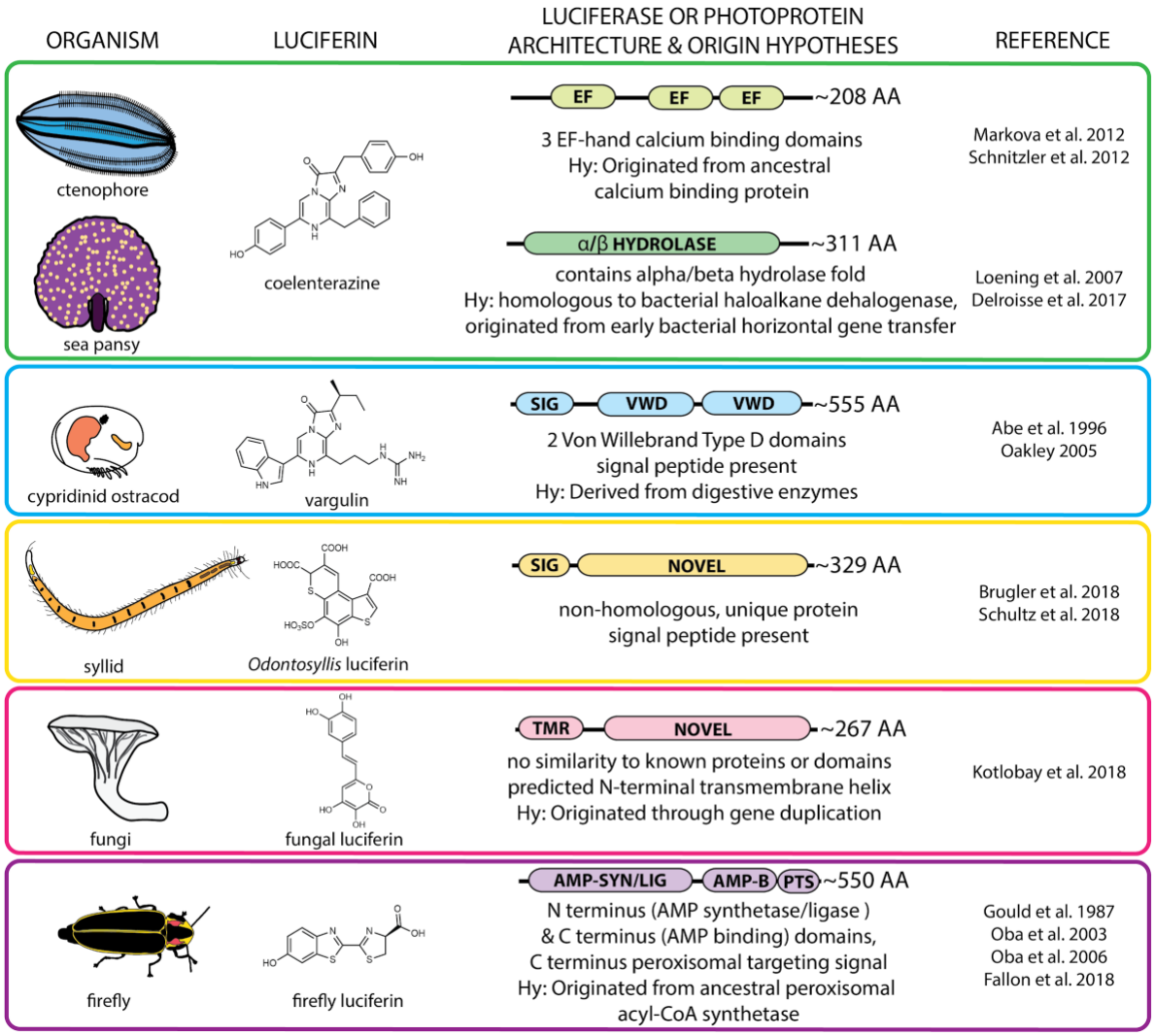

Figure 5. Subset of photoproteins and luciferases exemplifying the heterogeneity found in bioluminescent proteins

Photoproteins and luciferases vary in amino acid length, protein domain architecture, and evolutionary history. Most luciferases and photoproteins are non-homologous, even in organisms that use the same luciferin substrate (such as ctenophores and sea pansies). As a result of this disparity, bioluminescence research proposed different hypotheses (Hy) for the origins of these proteins.

(b) Origin of luciferins

Although luciferins are chemically distinct, their roles in the light producing biochemical reactions are the same—luciferins are oxidized and produce chemiluminescence. Despite their disparate chemical structures, how can luciferins play convergent biochemical roles? Hypotheses for the origin and evolution of luciferins are strongly tied to luciferins' sensitivity to oxygen and

their intrinsic chemiluminescence. The oxidative stress hypothesis posits that luciferins or their precursors, called "proto luciferins", played a role in scavenging detrimental molecules, called reactive oxygen species, which damage biologically important materials (Hastings, 1983; Rees *et al.*, 1998). Combined with another hypothesis proposing that (proto)luciferins had innate chemiluminescence capabilities, selection could act on (proto)luciferins that produce light upon oxidation.

Recent discoveries in fungal bioluminescence systems, specifically the identification of the biosynthetic pathway for fungal luciferin, support the oxidative stress hypothesis. Fungi produce luciferin by modifying a luciferin precursor called hispidin (Oliveira *et al.*, 2012; Purtov *et al.*, 2015; Oba *et al.*, 2017b; Kotlobay *et al.*, 2018). Hispidin is also found in non-luminous fungi, where it is used as pigment precursor and has antioxidant biochemical properties (Bu'Lock, Leeming, & Smith, 1962; Khushbaktova *et al.*, 1996). With that information in hand, the chemiluminescence hypothesis can be tested by investigating whether the proto luciferin, hispidin, has innate chemiluminescent properties. In the future, identifying luciferins' biosynthetic pathways will ultimately reveal how different luciferins originated and integrated into bioluminescence systems, and whether there are patterns in their evolutionary histories.

III. Reviewing the multi-level convergent evolution of bioluminescence: aims and scope

We focus our review on organisms that use the luciferins coelenterazine and vargulin because in addition to phenotypic convergence, they represent cases of interphylum convergence (on a molecular level) in luminous systems that use the same luciferins to produce light. Although krill and dinoflagellate luciferins are almost structurally identical, we do not include their bioluminescent systems in this review because there is no experimental evidence supporting the hypothesis that krill obtain their luciferin from dinoflagellates. Additionally, although fireflies and click beetles independently evolved their bioluminescence systems and biochemistries through the parallel evolution of their luciferases, they are members of the same phylum, therefore we do not review their bioluminescent systems here (Fallon *et al.*, 2018).

There is an uneven distribution of functional information in bioluminescent systems because a large portion of past bioluminescence research focused on identifying the molecular bases of bioluminescence biochemistries and the larger structural components of light organs. We anticipate that, with advancements in research technologies, future bioluminescence research will characterize more modules used to perform the functions involved in luminous systems. Specifically, we hope our review will drive researchers to investigate: (i) morphology and/or genetics of cell and subcellular components of luminous systems, (ii) biochemistry of luciferin storage mechanisms, (iii) genetic and physiological mechanisms to behaviourally control light production, (iv) mechanisms for catabolism and recycling of bioluminescent molecules.

There are many prior reviews that are focused on the biodiversity, biochemistry, biological function, sexual dimorphism, and vision of bioluminescent organisms, but none that address the biological level diversity in organisms with molecularly convergent bioluminescent systems

(Harvey, 1956; Locket, 1977; Herring & Clarke, 1983; Herring, 1987, 1990, 2000, 2007; Herring & Widder, 2001; Viviani, 2002; Warrant & Locket, 2004; Haddock *et al.*, 2010; Widder, 2010; Oba *et al.*, 2011, 2017a). Therefore, in this review we will broadly highlight the diverse biological levels of organization used for light production in luminous systems that rely on the shared luciferins (coelenterazine and vargulin) and identify questions that arise after using a holistic and hierarchical approach to systematically examine these systems.

(1) Coelenterazine systems

A number of cephalopods, chaetognaths, cnidarians, crustaceans, ctenophores, fishes, larvaceans, ophiuroids, and radiolarians independently evolved coelenterazine-based bioluminescence, and this taxonomic diversity is reflected in the diverse structures and mechanisms underlying light production. How did bioluminescence evolve in this diverse array of organisms?

Coelenterazine-based bioluminescence ranges from violet to green, and the colour of emitted light can be altered by using fluorescent proteins in the bioluminescent system (e.g. green fluorescent protein in *Aequorea*) (Shimomura, Johnson, & Saiga, 1962). Bioluminescence can be secreted into the environment by using specialized glands and organs to eject the bioluminescence, and/or be produced intracellularly and/or housed within the organism in structured photophores or light organs. Interestingly, *Vampyroteuthis infernalis* (Chun, 1903) can use its photophores to produce both secreted and intracellular bioluminescence (Robison *et al.*, 2003). Decapod and mysid shrimp secrete bioluminescence from their mouths and through a pore on the maxilla, respectively, while some copepods and halocyprid ostracods use glands to secrete bioluminescence through pores, chaetognaths secrete luminous particles from hexagonal-chambered light organs, and searsid fishes use a black lined, post-cleithral organ to secrete bioluminescence (Angel, 1968; Barnes & Case, 1972; Herring, 1972, 1976; Thuesen, Goetz, & Haddock, 2010; Wong *et al.*, 2015). A number of decapod shrimps, many fishes, and most cephalopods produce internal and/or intracellular bioluminescence in photophores ranging in structural complexity (composed of structures such as: lenses, reflectors, pigment, or photogenic tissue), while ctenophores, cnidarians, radiolarians, ophiuroids, and some copepods house bioluminescence in luminous areas of their bodies such as: the granular light organs along the rim of *Aequorea*, the gelatinous outer layer of radiolarians, or the spines, radial shields, ventral/lateral plates in some ophiuroids. As shown here, the organs used to produce coelenterazine-based bioluminescence are diverse despite having molecularly convergent bioluminescent biochemistries. On an organ level, how did each type of light organ (i.e. photophores, glands, and mouths) convergently evolve?

Like the diversity found in bioluminescent organs, the cells used to produce coelenterazine-based bioluminescence are morphologically variable and contain different subcellular contents both within and across taxa. Although photogenic organs are structurally variable, most contain photocytes (cells grouped together by the common function of light production). Within and across groups, did photocytes evolve from the same ancestral cell type? By distinguishing and

morphologically characterizing the cell types associated with bioluminescence systems, especially photocytes, we can begin to form hypotheses on how these cells originated and evolved. In ophiuroids, *Amphipholis squamata* (Chiaie, 1823) luminous organs contain small, spherical photocytes filled with granules that change as the bioluminescence biochemical reaction progresses (Deheyn, Mallefet, & Jangoux, 2000). In fishes, hatchetfish photophores have small photocytes which are densely packed and polyhedral shaped, while lanternfish photophores have long, spindle-shaped photocytes (Nicol, 1969). Within photocytes, subcellular organization can vary even within a family, like in members of the squid family Enoploteuthidae (Young & Bennett, 2013). Although these photocytes are morphologically dissimilar, did they evolve from similar ancestral cell types? If so, did parallel genetic mechanisms govern photocyte evolution, size, and organization? If not, how did photocytes convergently evolve from disparate cell types?

On a molecular level, both photoproteins and luciferases are capable of producing light with coelenterazine, but these proteins have low percentages of sequence similarity, suggesting that there are many routes that proteins can take to converge to produce bioluminescence (Markova & Vysotski, 2015). Although some closely related taxa use homologous photoproteins/luciferases, that is not always the case. In Cnidaria, the luminous scyphozoan *Periphylla periphylla* (Péron & Lesueur, 1810) uses a luciferase rather than a photoprotein used by other scyphozoans, and the anthozoan sea pansy *Renilla* uses a luciferase (Ward & Cormier, 1978; Shimomura *et al.*, 2001). Similarly, both photoprotein and luciferase systems are found in the ophiuroids *Amphiura filiformis* (Müller, 1776) and *Ophiopsila californica* (Clark, 1921), which use a luciferase and calcium dependent photoprotein, respectively (Shimomura, 1986; Delroisse *et al.*, 2017). Photoproteins using coelenterazine as a substrate can also vary in terms of the cofactors required to produce bioluminescence. For example, cnidarians, ctenophores, and radiolarians use the divalent cation calcium as a photoprotein cofactor, but the squids *Dosidicus gigas* (Orbigny, 1847) and *Sthenoteuthis oualaniensis* use a monovalent cation cofactor such as sodium or potassium (Shimomura, 1985; Takahashi & Isobe, 1994; Tsuji *et al.*, 1995; Galeazzo *et al.*, 2019). The variability in bioluminescence biochemistry across independent origins reveal that there are numerous biological approaches to harnessing coelenterazine-based bioluminescence.

(2) Vargulin systems

Although not as widely used as coelenterazine, vargulin-based bioluminescence systems still exhibit some of the variability found in coelenterazine-based systems. Bioluminescence can be intracellular (in *Porichthys*) or secreted into the surround environment (in cypridinid ostracods), the colours of emitted light can range from blue to green, and organs housing internal and/or intracellular bioluminescence vary in morphology and origin (Tsuji, Lynch, & Haneda, 1970; Anctil, 1977; Sweeney, 1980; Herring, 1985). The luminous fishes that produce autogenic bioluminescence in the Pempheridae, Apogonidae, and Batrachoididae families rely on the dietary acquisition of vargulin from their bioluminescent ostracod prey (Haneda & Johnson, 1958; Tsuji *et al.*, 1971; Thompson, Nafpakntis, & Tsuji, 1987; Thacker & Roje, 2009). If coelenterazine is so

ubiquitous in marine systems, why do these fishes use vargulin? Do their diets lack coelenterazine, or was the evolution of their bioluminescence systems constrained by their digestive systems' ability to absorb the substrate?

Even though vargulin-based bioluminescence independently originated fewer times than coelenterazine-based systems, there are still similar cases of morphological diversity in the light organs used for vargulin-based bioluminescence in ostracods and fishes. Despite being non-convergent on a biochemical level, how did luminous organisms convergently evolve structurally similar light organs? Similar to the oral light organs in the decapod/mysid shrimps, cypridinid ostracods secrete bioluminescent mucus from a specialized light organ called the upper lip, which contains tusks ending in numerous nozzles (Huvard, 1993; Abe *et al.*, 2000; Cohen & Morin, 2003). The photophores used for vargulin and coelenterazine-based bioluminescence also share similar structural components. The toadfishes *Porichthys* sp., commonly known as midshipman fishes, have hundreds of ventral dermal photophores composed of a pear shaped lens, photogenic tissue, a reflector, and pigment (Nicol, 1957; Anctil, 1977). However, instead of photophores, pempherid and apogonid fishes independently evolved visceral light organs that are morphologically unique to vargulin-based bioluminescence. The pempherid *Parapriacanthus ransonnetti* produces bioluminescence in ventral thoracic and anal light organs, the former extending from the pyloric caeca (a structure where luciferin is stored), and the latter connecting to the rectum and anus (Haneda & Johnson, 1958). Although pempherid and apogonid vargulin-based bioluminescence is produced in light organs that are connected to the gut, the light organs in pempherids are distinct structures containing numerous tubules, while the ones in apogonids are less structurally complex (Thacker & Roje, 2009). Why is this type of visceral light organ unique to vargulin-based bioluminescence? In addition to using vargulin stored in the pyloric caeca, *P. ransonnetti* also sequesters and uses ostracod luciferase (the aforementioned "kleptoprotein bioluminescence") (Bessho-Uehara *et al.*, 2020). Did this dual functional requirement of sequestering both ostracod luciferin and luciferase constrain the evolution of pempherid light organs?

Like coelenterazine systems, cellular morphology and subcellular composition is variable across independent origins of vargulin-based bioluminescence. In the ostracods *Vargula hilgendorfii* (Mueller, 1890), *V. tsujii* (Kornicker & Jh, 1977), and *Photeros graminicola* (Cohen & Morin, 1986), the light organ called the upper lip houses different types of gland cells in the light organ (estimated to have at least two and up to six different cell types), with some types of gland cells localized in specific areas and other cell types appearing in multiple regions of the upper lip. One particular cell type contains yellow vesicles hypothesized to hold vargulin, which is yellow/orange coloured when pure (Huvard, 1993; Abe *et al.*, 2000). Midshipman fish photophores contain many types of cells: lens cells, photocytes, pigment cells, and supportive cells (Strum, 1969; Baguet & Zietz-Nicolas, 1979). Their photophore reflector is made of cells containing guanine crystals, similar to that of many other fishes. Although other fishes with photophores use coelenterazine, do the cellular components of these photophores originate from similar progenitor cells or similar cell types? Comparing the cell types in pempherid and apogonid

light organs, which independently evolved from the extension of the gut, will be especially interesting, because of their differences in organ level structural complexity. In addition to differences in structural complexity, do the cell-types in pempherid and apogonid light organs differ?

Unlike any known coelenterazine-based luciferase or photoprotein, the luciferase in cypridinid ostracods contains two von Willebrand factor type D domains, which are domains shared in mucins and glycoproteins (Oakley, 2005). Although the enzymatic function of the ancestral protein is unknown, the dual role of ostracods' upper lip (secreting bioluminescence and consuming food), coupled with the von Willebrand factor type D domains present, suggest the hypothesis that ostracod luciferase originated from an ancestral digestive enzyme (Abe *et al.*, 1996; Oakley, 2005). As previously mentioned, *Parapriacanthus* sequesters both vargulin and ostracod luciferase from its diet and uses them to produce light. The recent discovery of kleptoprotein bioluminescence in *Parapriacanthus* may extend to *Pempheris*, the sister genus with similar light organs that also use vargulin for bioluminescence (Haneda, Johnson, & Shimomura, 1966). Because the apogonid *Jaydia* has a similar light organ, future work testing kleptoprotein bioluminescence in apogonids will reveal whether this recently discovered mechanism for producing autogenic bioluminescence convergently evolved in other taxa. Although the molecular basis of cypridinid ostracod bioluminescence is well known, it is unknown in midshipman fishes. Midshipman luciferase is difficult to extract; perhaps it is membrane bound, or their photophores may not contain a lot of luciferase (Cormier, Crane, & Nakano, 1967; Tsuji *et al.*, 1971). *In vivo* observations reveal that midshipman fish bioluminescence is much dimmer relative to that of ostracods. Differences in their bioluminescence biochemistry, coupled with differences in their bioluminescence emission spectra, suggest that these luciferases are unrelated (Cormier *et al.*, 1967; Tsuji *et al.*, 1975). It is unlikely that midshipman fishes exhibit kleptoprotein bioluminescence, because non-luminous specimens can bioluminescence after injecting only vargulin into their stomachs (Barnes, Case, & Tsuji, 1973). Thus, identifying the luciferase gene in midshipman fishes, the only other record of a molecularly convergent vargulin-based bioluminescence system, is crucial to understanding how this interphylum molecular convergence evolved.

IV. Luciferin storage and regulation: an unexplored area of bioluminescence research

In addition to investigating the traditional luciferin and luciferase/photoprotein bioluminescence system, we highlight that an untapped area in bioluminescence research is the presence of stored luciferins and the biochemical mechanisms used to store them. Filling that gap in bioluminescence research will not only reveal how various bioluminescent systems tackle the issue presented by unstable luciferins, which will improve our understanding of how functional convergence (i.e. preservation of a substrate) evolved through conserved or disparate genetic mechanisms, but will also improve our understanding of bioluminescence as a whole by characterizing the substrate storage functional module used for this complex trait. Furthermore, by

genetically characterizing the proteins used for storing luciferins, the results of this research can be applied to regulate the availability of active luciferins in bioluminescence-based biotechnological imaging assays, which currently rely on luciferins that are vulnerable to non-specific oxidation.

Luciferins are double edged swords because they are typically unstable molecules that must be oxidized in order to perform their function and yet this very attribute makes them prone to non-specific oxidation by the surrounding environment. To prevent non-specific oxidation and/or to regulate usage, some animals such as the fireflies, ostracods, and sea pansies biochemically modify their luciferins to store them (Cormier, Hori, & Karkhanis, 1970; Nakamura *et al.*, 2014; Fallon *et al.*, 2016). Additionally, lanternfish tissues contain low amounts of coelenterazine derivatives (relative to coelenterazine), but it is not known whether lanternfishes have a mechanism to convert luciferin into a storage form (Duchatelet *et al.*, 2019). Alternatively, fungi and dinoflagellates have luciferin precursors that can be readily converted to active luciferins upon the addition of NADPH (Fresneau *et al.*, 1986; Oba *et al.*, 2017b). In addition to biochemical modification following luciferin biosynthesis or acquisition, organisms can possibly store luciferins by biochemically regulating the conversion of luciferin precursors to active luciferins. We hypothesize that evolution can find a way to regulate available luciferin by simply controlling the production of light from luciferin precursors.

Interestingly, fireflies, ostracods, and sea pansies use different luciferins but are shown or proposed to use sulfotransferases to store the substrates in a sulfated form. However, the genetic basis for luciferin sulfotransferase is only known in fireflies (Fallon *et al.*, 2016). The activity of this luciferin sulfotransferase is specific to fireflies' luciferin. Similar to the photoprotein/luciferase origin hypothesis, fireflies' luciferin sulfotransferase may have originated from an ancestral, promiscuous sulfotransferase. Luciferins can also be stored by luciferin-binding-proteins, as in the case of dinoflagellates, or "substrate binding fractions" in the luminous fly *Orfelia fultoni* (Fisher, 1940) (Lee *et al.*, 1993; Viviani *et al.*, 2020). Interestingly, sea pansies have both a storage form of luciferin and a luciferin binding protein, the latter which releases coelenterazine upon the addition of calcium (Anderson, Charbonneau, & Cormier, 1974; Inouye, 2007). As of this moment, we only know of a few organisms that can actively store their luciferins (dinoflagellates, fireflies, sea pansies, and ostracods), and three types of hypothesized luciferin storage mechanisms: (i) biochemical modification into a sulfated form, (ii) luciferin binding proteins, (iii) biochemical regulation of luciferin production. Do organisms that share luciferins store luciferins in the same way? If so, how and why? Fireflies, ostracods, and dinoflagellates produce their own luciferins, yet they still store it. Can organisms with a dietary dependence on luciferin store luciferin as well? Testing for luciferin storage capabilities in bioluminescent systems and chemically identifying storage forms in shared luciferins such as coelenterazine and vargulin will be especially valuable, because it will reveal other cases of molecular convergence or perhaps parallelism, within this complex trait.

V. Promising new tools for bioluminescence research

Characterizing bioluminescence systems incorporates many fields such as biology, ecology, chemistry, physics, and computational biology. The most comprehensive chemical and biochemical methods are outlined in a book by Osamu Shimomura, who shared the 2008 Nobel Prize in Chemistry with Martin Chalfie and Roger Y. Tsien for discovering the green fluorescent protein in *Aequorea* and developing it for its use as a biotechnological tool (Shimomura, 2006). In the following sections, we highlight newly developed biological and chemical tools that will be useful for characterizing bioluminescent systems.

(1) Advances in genetic sequencing

Most past bioluminescence research examined bioluminescence on the organ or molecular levels. Sequencing on a cellular resolution will fill in current gaps in our knowledge of bioluminescence on a cellular level. The development of single cell sequencing techniques such as scRNA-seq (single cell RNA sequencing) or single cell ATAC-seq (assay for transposase-accessible chromatin using sequencing), which generates RNA data and identifies accessible chromatin on a cellular level, respectively, allows us to move towards understanding regulatory elements used to control bioluminescence, characterizing cell types associated with bioluminescence systems, and tracing the developmental history and evolutionary relationships of cells used in luminous systems (Buenrostro *et al.*, 2015; Kolodziejczyk *et al.*, 2015; Shafer, 2019). Using these sequencing methods on a cellular level rather than an organ level is advantageous because they offer a better genetic resolution of the organ as a whole. Furthermore, by using spatial transcriptomics technology, which sequences RNA given a section of tissue, we can spatially resolve transcriptomic data and connect morphological data with genetic data (Ståhl *et al.*, 2016). For example, we can apply these technologies to identify genes unique to luminous photocytes or other cells involved in light production, followed by identifying the molecular basis of bioluminescence by functionally testing candidate genes (i.e. CRISPR genome editing to test regulatory elements, using protein expression systems to produce and functionally test proteins) (Cregg *et al.*, 2000; Schnitzler *et al.*, 2012; Ran *et al.*, 2013; Rosano & Ceccarelli, 2014).

(2) Mass spectrometry: improving resolution and advancements in mass spectrometry imaging

Mass spectrometry can be extremely sensitive and useful in characterizing bioluminescence systems. For example, researchers used mass spectrometry to quantify coelenterazine using just one halocyprid ostracod specimen, an animal that is about 1 millimetre long (Oba *et al.*, 2004). Rather than mixing methanol extracts from luminous tissue with a known luciferase, a technique called “cross reactivity” traditionally used to test for the presence of a specific luciferin, mass spectrometry can analyse these extracts and chemically identify luciferin,

if it is known. Given a luciferin standard, analysis tools (e.g. spectral matching tandem mass spectra, molecular networking) can identify components in the extract that are similar but not an exact match, which will be useful for identifying derivatives of luciferins such as the modified coelenterazine used in *Sthenoteuthis* and *Watasenia* bioluminescence (Guthals *et al.*, 2012; Yang *et al.*, 2013; Wang *et al.*, 2016).

In addition to small molecule analyses, mass spectrometry is also used in proteomic workflows and top-down/bottom-up proteomics, which quantifies the mass and abundance of proteins, and identifies the amino acid sequence and its precursor gene given an extracted and purified protein, respectively (Aebersold & Mann, 2003; Bantscheff *et al.*, 2007). Like single cell genetic sequencing, mass spectrometry can also be used to identify proteins on a cellular level (Su, Shi, & Wei, 2017; Specht & Slaevov, 2018). Furthermore, a technique called mass spectrometry imaging offers the spatial distribution of molecules given a tissue sample (Stoeckli *et al.*, 2001; McDonnell & Heeren, 2007; Stauber *et al.*, 2010; Hamilton *et al.*, 2020). This technique, used to track proteins and cell metabolism, will have promising applications in bioluminescence research. We suggest that this approach can be applied to tracing and identifying molecular and genetic pathways to luciferin biosynthesis. Specifically, we envision that, as protein identification improves in mass spectrometry imaging techniques, this approach can identify proteins in the regions of the tissue that contain luciferins/luciferin-like molecules (Piehowski *et al.*, 2020). After obtaining this information, we can propose and functionally test potential biosynthetic pathways using protein expression or genome editing techniques. Using these chemical tools will bolster data generated from new sequencing techniques and will allow us to understand how bioluminescence systems function on a cellular, biochemical, and genetic level.

VI. Conclusions

(1) Central goals in evolutionary biology aim to understand how complexity originates, what shapes its evolution, and how complex traits with similar functions repeatedly evolve. We can systematically study the convergent evolution of a complex trait by using a hierarchical and holistic approach that investigates the multi-level convergent evolution of a complex trait and its functional modules.

(2) Bioluminescence, an ecologically important trait, is an excellent system for studying the roles of divergence and conservation in convergent evolution because of its biological complexity, biodiversity, and ubiquity across lifeforms. We present an updated estimate of at least 84 independent origins of bioluminescence across the tree of life.

(3) Developing bioluminescence as a system for studying convergence requires studying all levels of life and an interdisciplinary approach that incorporates computational biology, molecular biology, ecology, evolution, chemistry, biochemistry, and physics.

(4) Organisms that use the same luciferins have bioluminescent proteins that are molecularly convergent. Despite this molecular convergence, we see both conservation and divergence in the biological functions, organs, and cells associated with the production of

bioluminescence. We anticipate that, with advances in mass spectrometry and genetic sequencing, future bioluminescence research will further characterize the functional modules involved in a bioluminescent system (i.e. luciferin/luciferase production/acquisition, bioluminescent molecule recycling or catabolism, maintenance of light organs, behavioural control of light production, modification of light emission, storage of luciferin).

(5) Ultimately, using convergent systems such as bioluminescence to study broad evolutionary questions will reveal how divergent evolutionary pathways can converge on similar traits and will further our knowledge on the origins and evolution of biological complexity.

. Acknowledgments

We extend our many thanks to S.H.D Haddock, M. Bessho-Uehara, T.R. Fallon, J.A. Goodheart, N.M. Hensley, N. Picciani, L.Y. Mesrop, and B.A. Vincent for their insightful comments and many fruitful discussions on bioluminescence biology, biochemistry, and phenotypic convergence. We thank T.R. Fallon for suggesting regulating luciferin production as a mechanism of luciferin storage. Funding from the National Science Foundation supported this work, specifically the NSF Graduate Research Fellowship Program (NSF GRFP 1650114) to E.S. Lau and NSF 1754770 to T.H. Oakley.

VII. References

Abe, K., Nagata, T. & Hashizume, H. (1996) Digestive Enzymes from the Upper Lip of a Bioluminescent Ostracod, Vargula hilgendorfii. *Reports of Faculty of Science, Shizuoka University* **30**, 35–40. Shizuoka University.

Abe, K., Ono, T., Yamada, K., Yamamura, N. & Ikuta, K. (2000) Multifunctions of the upper lip and a ventral reflecting organ in a bioluminescent ostracod Vargula hilgendorfii (Müller, 1890). *Hydrobiologia* **419**, 73–82. Springer.

Adams, S.T., Jr & Miller, S.C. (2020) Enzymatic promiscuity and the evolution of bioluminescence. *The FEBS journal* **287**, 1369–1380. Wiley Online Library.

Aebersold, R. & Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422**, 198–207. nature.com.

Anctil, M. (1977) Development of bioluminescence and photophores in the midshipman fish, Porichthys notatus. *Journal of morphology* **151**, 363–395. Wiley Online Library.

Anderson, J.M., Charbonneau, H. & Cormier, M.J. (1974) Mechanism of calcium induction of Renilla bioluminescence. Involvement of a calcium-triggered luciferin binding protein.

Biochemistry **13**, 1195–1200. ACS Publications.

Angel, M.V. (1968) Bioluminescence in planktonic halocyprid ostracods. *Journal of the Marine Biological Association of the United Kingdom. Marine Biological Association of the United Kingdom* **48**, 255–257. Cambridge University Press.

Arbuckle, K. (2020) From molecules to macroevolution: Venom as a model system for evolutionary biology across levels of life. *Toxicon: X* **6**. Elsevier Ltd.

Arnold, J.M. & Young, R.E. (1974) Ultrastructure of a cephalopod photophore I. Structure of the photogenic tissue. *The Biological bulletin* **147**, 507–521.

Baguet, F. (1975) Excitation and control of isolated photophores of luminous fishes. *Progress in neurobiology* **5**, 97–125.

Baguet, F. & Zietz-Nicolas, A.M. (1979) Fluorescence and luminescence of isolated photophores of Porichthys. *The Journal of experimental biology* **78**, 47–57. Citeseer.

Bantscheff, M., Schirle, M., Sweetman, G., Rick, J. & Kuster, B. (2007) Quantitative mass spectrometry in proteomics: a critical review. *Analytical and bioanalytical chemistry* **389**, 1017–1031. Springer.

Barnes, A.T. & Case, J.F. (1972) Bioluminescence in the mesopelagic copepod, Gaussia princeps (T. Scott). *Journal of experimental marine biology and ecology* **8**, 53–71.

Barnes, A.T., Case, J.F. & Tsuji, F.I. (1973) Induction of bioluminescence in a luciferin deficient form of the marine teleost, Porichthys, in response to exogenous luciferin. *Comparative biochemistry and physiology. A, Comparative physiology* **46**, 709–723. Elsevier.

Berry, S.S. (1911) Note on a new Abraliopsis from Japan (A. scintillans n. sp.). *The Nautilus* **25**, 93–94. ci.nii.ac.jp.

Bessho-Uehara, M., Francis, & W.R., Haddock, S.H.D. (2020) Biochemical characterization of diverse deep-sea anthozoan bioluminescence systems. *Marine biology*.

Bessho-Uehara, M., Yamamoto, N., Shigenobu, S., Mori, H., Kuwata, K. & Oba, Y. (2020) Kleptoprotein bioluminescence: Parapriacanthus fish obtain luciferase from ostracod prey. *Science advances* **6**, eaax4942. advances.sciencemag.org.

Brugler, M.R., Aguado, M.T., Tessler, M. & Siddall, M.E. (2018) The transcriptome of the bermuda fireworm odontosyllis enopla (annelida: Syllidae): A unique luciferase gene family and putative epitoky-related genes. *PloS one* **13**, 1–14.

Buenrostro, J.D., Wu, B., Chang, H.Y. & Greenleaf, W.J. (2015) ATAC-seq: A Method for

Assaying Chromatin Accessibility Genome-Wide. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]* **109**. Wiley Online Library.

Bu'Lock, J.D., Leeming, P.R. & Smith, H.G. (1962) Pyrones. Part II. Hispidin, a new pigment and precursor of a fungus 'lignin'. *Journal of the Chemical Society*, 2085–2089.

Campbell, A.K. & Herring, P.J. (1990) Imidazolopyrazine bioluminescence in copepods and other marine organisms. *Marine biology* **104**, 219–225.

Case, J.F., Warner, J., Barnes, A.T. & Lowenstine, M. (1977) Bioluminescence of lantern fish (Myctophidae) in response to changes in light intensity. *Nature* **265**, 179–181. nature.com.

Chiaie, S.D. (1823) *Memorie sulla storia e notomia degli animali senza vertebre del regno di Napoli*. Dalla stamperia de' Fratelli Fernandes.

Christin, P.A., Weinreich, D.M. & Besnard, G. (2010) Causes and evolutionary significance of genetic convergence. *Trends in genetics: TIG* **26**, 400–405. Elsevier Ltd.

Chun, C. (1903) Aus den Tiefen des Weltmeeres: Schilderungen von der Deutschen Tiefsee. *Expedition* **549**.

Chun, C. (1906) System der Cranchien. *Zoologischer Anzeiger* **31**, 82–86.

Clark, A.H. (1921) A new Ophiuran of the genus Ophiopsila from Southern California. *Proceedings. Biological Society of Washington* **34**, 109–110.

Cohen, A.C. & Morin, J.G. (1986) *Three new luminescent ostracodes of the genus Vargula (Myodocopida, Cypridinidae) from the San Blas region of Panama*. Natural History Museum of Los Angeles County.

Cohen, A.C. & Morin, J.G. (2003) Sexual Morphology, Reproduction and the Evolution of Bioluminescence in Ostracoda. *The Paleontological Society Papers* **9**, 37–70. Cambridge University Press.

Cormier, M.J., Crane, J.M., Jr & Nakano, Y. (1967) Evidence for the identity of the luminescent systems of Porichthysporosissimus (fish) and Cypridinahilgendorfii (crustacean). *Biochemical and biophysical research communications*. Elsevier.

Cormier, M.J., Hori, K. & Karkhanis, Y.D. (1970) Studies on the bioluminescence of Renilla reniformis. VII. Conversion of luciferin into luciferyl sulfate by luciferin sulfokinase. *Biochemistry* **9**, 1184–1189. ACS Publications.

Cregg, J.M., Cereghino, J.L., Shi, J. & Higgins, D.R. (2000) Recombinant protein expression in Pichia pastoris. *Molecular biotechnology* **16**, 23–52. Springer.

Davis, M.P., Sparks, J.S. & Smith, W.L. (2016) Repeated and Widespread Evolution of Bioluminescence in Marine Fishes. *PloS one* **11**, e0155154.

Deheyn, D., Mallefet, J. & Jangoux, M. (1998) Bioluminescence in Amphipholis squamata: an anti-predator aposematic deterrent function. *Echinoderms: San Francisco*, 351. vliz.be.

Deheyn, D., Mallefet, J. & Jangoux, M. (2000) Cytological changes during bioluminescence production in dissociated photocytes from the ophiuroid Amphipholis squamata (Echinodermata). *Cell and tissue research* **299**, 115–128. Springer.

Delroisse, J., Ullrich-Lüter, E., Blaue, S., Ortega-Martinez, O., Eeckhaut, I., Flammang, P. & Mallefet, J. (2017) A puzzling homology: a brittle star using a putative cnidarian-type luciferase for bioluminescence. *Open biology* **7**. royalsocietypublishing.org.

Denton, E.J., Herring, P.J., Widder, E.A., Latz, M.F. & Case, J.F. (1985) The roles of filters in the photophores of oceanic animals and their relation to vision in the oceanic environment. *Proceedings of the Royal Society of London. Series B. Biological Sciences* **225**, 63–97. Royal Society.

Duchatelet, L., Hermans, C., Duhamel, G., Cherel, Y., Guinet, C. & Mallefet, J. (2019) Détection de la coelentérazine chez cinq espèces de Myctophidae du plateau des Iles Kerguelen, Coelenterazine detection in five myctophid species from the Kerguelen Plateau. In *Second Kerguelen Plateau Symposium on Marine Ecosystems and Fisheries. November 2017, Hobart, Tasmania*. archimer.ifremer.fr.

Dunlap, J.C., Hastings, J.W. & Shimomura, O. (1981) Dinoflagellate luciferin is structurally related to chlorophyll. *FEBS letters* **135**, 273–276.

Dunn, C.W., Hejnol, A., Matus, D.Q., Pang, K., Browne, W.E., Smith, S.A., Seaver, E., Rouse, G.W., Obst, M., Edgecombe, G.D., Sørensen, M.V., Haddock, S.H.D., Schmidt-Rhaesa, A., Okusu, A., Kristensen, R.M., et al. (2008) Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* **452**, 745–749.

Fallon, T.R., Li, F.-S., Vicent, M.A. & Weng, J.-K. (2016) Sulfoluciferin is Biosynthesized by a Specialized Luciferin Sulfotransferase in Fireflies. *Biochemistry* **55**, 3341–3344. ACS Publications.

Fallon, T.R., Lower, S.E., Chang, C.H., Bessho-Uehara, M., Martin, G.J., Bewick, A.J., Behringer, M., Debat, H.J., Wong, I., Day, J.C., Suvorov, A., Silva, C.J., Stanger-Hall, K.F., Hall, D.W., Schmitz, R.J., et al. (2018) Firefly genomes illuminate parallel origins of bioluminescence in beetles. *eLife* **7**, 1–146.

Fernald, R.D. (2006) Casting a genetic light on the evolution of eyes. *Science* **313**, 1914–1918.

science.sciencemag.org.

Fisher, E.G. (1940) New Mycetophilidae from North Carolina (Diptera). *Entomological news* **51**, 243–247.

Francis, W.R., Shaner, N.C., Christianson, L.M., Powers, M.L. & Haddock, S.H.D. (2015) Occurrence of isopenicillin-N-synthase homologs in bioluminescent ctenophores and implications for coelenterazine biosynthesis. *PloS one* **10**, 1–20.

Frankel, N., Wang, S. & Stern, D.L. (2012) Conserved regulatory architecture underlies parallel genetic changes and convergent phenotypic evolution. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 20975–20979.

Frank, T.M., Widder, E.A. & Case, J.F. (1984) Dietary maintenance of bioluminescence in a deep-sea mysid. *The Journal of experimental biology* **109**, 385–389.

Fresneau, C., Hill, M., Lescure, N., Arrio, B., Dupaix, A. & Volfin, P. (1986) Dinoflagellate luminescence: Purification of a NAD(P)H-dependent reductase and of its substrate. *Archives of biochemistry and biophysics* **251**, 495–503. Elsevier.

Galeazzo, G.A., Mirza, J.D., Dorr, F.A., Pinto, E., Stevani, C.V., Lohrmann, K.B. & Oliveira, A.G. (2019) Characterizing the Bioluminescence of the Humboldt Squid, Dosidicus gigas (d'Orbigny, 1835): One of the Largest Luminescent Animals in the World. *Photochemistry and photobiology* **95**, 1179–1185.

Gerrish, G.A. & Morin, J.G. (2016) Living in sympatry via differentiation in time, space and display characters of courtship behaviors of bioluminescent marine ostracods. *Marine biology* **163**. Springer Berlin Heidelberg.

Gomi, K., Hirokawa, K. & Kajiyama, N. (2002) Molecular cloning and expression of the cDNAs encoding luciferin-regenerating enzyme from Luciola cruciata and Luciola lateralis. *Gene* **294**, 157–166. Elsevier.

Gompel, N. & Prud'homme, B. (2009) The causes of repeated genetic evolution. *Developmental biology* **332**, 36–47. Elsevier.

Gould, S.G., Keller, G.A. & Subramani, S. (1987) Identification of a peroxisomal targeting signal at the carboxy terminus of firefly luciferase. *The Journal of cell biology* **105**, 2923–2931. rupress.org.

Guthals, A., Watrous, J.D., Dorrestein, P.C. & Bandeira, N. (2012) The spectral networks paradigm in high throughput mass spectrometry. *Molecular bioSystems* **8**, 2535–2544. pubs.rsc.org.

Haddock, S.H.D. & Case, J.F. (1999) Bioluminescence spectra of shallow and deep-sea gelatinous zooplankton: ctenophores, medusae and siphonophores. *Marine biology* **133**, 571–582. Springer.

Haddock, S.H.D., Moline, M.A. & Case, J.F. (2010) Bioluminescence in the sea. *Annual review of marine science* **2**, 443–493.

Haddock, S.H., Rivers, T.J. & Robison, B.H. (2001) Can coelenterates make coelenterazine? Dietary requirement for luciferin in cnidarian bioluminescence. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 11148–11151.

Hamilton, B.R., Marshall, D.L., Casewell, N.R., Harrison, R.A., Blanksby, S.J. & Undheim, E.A.B. (2020) Mapping Enzyme Activity on Tissue by Functional Mass Spectrometry Imaging. *Angewandte Chemie* **59**, 3855–3858. Wiley Online Library.

Haneda, Y. & Johnson, F.H. (1958) The luciferin-luciferase reaction in a fish, Parapricanthus beryciformis, of newly discovered luminescence. *Proceedings of the National Academy of Sciences of the United States of America* **44**, 127–129. ncbi.nlm.nih.gov.

Haneda, Y., Johnson, F.H. & Shimomura, O. (1966) The origin of luciferin in the luminous ducts of Parapriacanthus ransonneti, Pempheris klunzingeri, and Apogon ellioti. *Bioluminescence in progress*, 533–545. Princeton University Press Princeton, NJ, USA.

Harper, R.D. & Case, J.F. (1999) Disruptive counterillumination and its anti-predatory value in the plain ® sh midshipman Porichthys notatus **001**, 529–540.

Harvey, E.N. (1920) *The Nature of Animal Light*. Lippincott.

Harvey, E.N. (1956) Evolution and bioluminescence. *The Quarterly review of biology* **31**, 270–287.

Hastings, J.W. (1983) Biological diversity, chemical mechanisms, and the evolutionary origins of bioluminescent systems. *Journal of molecular evolution* **19**, 309–321.

Hastings, J.W. & Wilson, T. (1976) Bioluminescence and chemiluminescence. *Photochemistry and photobiology* **23**, 461–473. Wiley Online Library.

Hellinger, J., Jägers, P., Donner, M., Sutt, F., Mark, M.D., Senen, B., Tollrian, R. & Herlitze, S. (2017) The Flashlight Fish Anomalops katoptron Uses Bioluminescent Light to Detect Prey in the Dark. *PloS one* **12**, e0170489. journals.plos.org.

Herring, P.J. (1972) Bioluminescence of searsid fishes. *Journal of the Marine Biological Association of the United Kingdom. Marine Biological Association of the United Kingdom*

52, 879–887.

Herring, P.J. (1976) Bioluminescence in decapod Crustacea. *Journal of the Marine Biological Association of the United Kingdom. Marine Biological Association of the United Kingdom* **56**, 1029–1047. Cambridge University Press.

Herring, P.J. (1985) Bioluminescence in the crustacea. *Journal of crustacean biology: a quarterly of the Crustacean Society for the publication of research on any aspect of the biology of crustacea* **5**, 557–573.

Herring, P.J. (1987) Systematic distribution of bioluminescence in living organisms. *Journal of bioluminescence and chemiluminescence* **1**, 147–163.

Herring, P.J. (1990) Bioluminescent communication in the sea. *Light and life in the sea*, 245–264. Cambridge University Press Cambridge, UK.

Herring, P.J. (2000) Bioluminescent signals and the role of reflectors. *Journal of Optics A: Pure and Applied Optics* **2**.

Herring, P.J. (2007) Sex with the lights on? A review of bioluminescent sexual dimorphism in the sea. *Journal of the Marine Biological Association of the United Kingdom. Marine Biological Association of the United Kingdom* **87**, 829–842. Cambridge University Press.

Herring, P.J. & Clarke, M.R. (1983) The spectral characteristics of luminous marine organisms. *Proceedings of the Royal Society of London. Series B. Biological Sciences* **220**, 183–217. Royal Society.

Herring, P.J., Dilly, P.N. & Cope, C. (1994) The bioluminescent organs of the deep-sea cephalopod Vampyroteuthis infernalis (Cephalopoda: Vampyromorpha). *Journal of zoology* **233**, 45–55. Wiley Online Library.

Herring, P.J. & Widder, E.A. (2001) Bioluminescence, 376–384.

Huvard, A.L. (1993) Ultrastructure of the light organ and immunocytochemical localization of luciferase in luminescent marine ostracods (Crustacea: Ostracoda: Cypridinidae). *Journal of morphology* **218**, 181–193. Wiley Online Library.

Inouye, S. (2007) Expression, purification and characterization of calcium-triggered luciferin-binding protein of Renilla reniformis. *Protein expression and purification* **52**, 66–73. Elsevier.

Joy, J.B., Liang, R.H., McCloskey, R.M., Nguyen, T. & Poon, A.F.Y. (2016) Ancestral Reconstruction. *PLoS computational biology* **12**, e1004763. journals.plos.org.

Kato, S.-I., Oba, Y., Ojika, M. & Inouye, S. (2004) Identification of the biosynthetic units of Cypridina luciferin in Cypridina (Vargula) hilgendorfii by LC/ESI-TOF-MS. *Tetrahedron* **60**, 11427–11434. Elsevier.

Khushbaktova, Z.A., Yusupova, S.M., Badal'Yants, K.L. & Syrov, V.N. (1996) Isolation of hispidin from a walnut-tree fungus and its antioxidant activity. *Chemistry of Natural Compounds* **32**, 27–29.

Kolodziejczyk, A.A., Kim, J.K., Svensson, V., Marioni, J.C. & Teichmann, S.A. (2015) The technology and biology of single-cell RNA sequencing. *Molecular cell* **58**, 610–620. Elsevier.

Kornicker, L.S. & Jh, B. (1977) Vargula tsujii, a new species of luminescent Ostracoda from Lower and Southern California (Myodocopa: Cypridininae). pascal-francis.inist.fr.

Kotlobay, A.A., Dubinnyi, M.A., Purtov, K.V., Guglya, E.B., Rodionova, N.S., Petushkov, V.N., Bolt, Y.V., Kublitski, V.S., Kaskova, Z.M., Ziganshin, R.H., Nelyubina, Y.V., Dorovatovskii, P.V., Eliseev, I.E., Branchini, B.R., Bourenkov, G., et al. (2019) Bioluminescence chemistry of fireworm Odontosyllis. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 18911–18916. National Acad Sciences.

Kotlobay, A.A., Sarkisyan, K.S., Mokrushina, Y.A., Marcet-Houben, M., Serebrovskaya, E.O., Markina, N.M., Somermeyer, L.G., Gorokhovatsky, A.Y., Vvedensky, A., Purtov, K.V., Petushkov, V.N., Rodionova, N.S., Chepurnyh, T.V., Fakhranurova, L.I., Guglya, E.B., et al. (2018) Genetically encodable bioluminescent system from fungi. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 12728–12732.

Lee, D.H., Mittag, M., Sczekan, S., Morse, D. & Hastings, J.W. (1993) Molecular cloning and genomic organization of a gene for luciferin-binding protein from the dinoflagellate Gonyaulax polyedra. *The Journal of biological chemistry* **268**, 8842–8850. ASBMB.

Lesson, R.P. & Duperrey, L.-I. (1830) *Voyage autour du monde: Exécuté par Ordre du Roi, Sur la Corvette de Sa Majesté, La Coquille, pendant les années 1822, 1823, 1824 et 1825, sous le ministère et conformént aux instructions de S. E. M. le Marquis de Clermont-Tonnerre, ministre de la marine; et publié sous les auspices de son excellence Mgr le Cte de Chabrol, ministre de la marine et des colonies. Zoologie / par M. Lesson. Tome 2e, 2e Partie*. Bertrand.

Locket, N.A. (1977) Adaptations to the Deep-Sea Environment. In *The Visual System in Vertebrates* (eds F. Crescitelli, C.A. Dvorak, D.J. Eder, A.M. Granda, D. Hamasaki, K. Holmberg, A. Hughes, N.A. Locket, W.N. McFarland, D.B. Meyer, W.R.A. Muntz, F.W.

Munz, E.C. Olson, R.W. Reyer & F. Crescitelli), pp. 67–192. Springer Berlin Heidelberg, Berlin, Heidelberg.

Loening, A.M., Fenn, T.D. & Gambhir, S.S. (2007) Crystal Structures of the Luciferase and Green Fluorescent Protein from Renilla reniformis. *Journal of molecular biology* **374**, 1017–1028. Elsevier.

Lorenz, W.W., McCann, R.O., Longiaru, M. & Cormier, M.J. (1991) Isolation and expression of a cDNA encoding Renilla reniformis luciferase. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 4438–4442. National Acad Sciences.

Losos, J.B. (2011) Convergence, adaptation, and constraint. *Evolution; international journal of organic evolution* **65**, 1827–1840.

Mallefet, J., Duchatelet, L. & Coubris, C. (2020) Bioluminescence induction in the ophiuroid Amphiura filiformis (Echinodermata). *The Journal of experimental biology* **223**. jeb.biologists.org.

Mallefet, J. & Shimomura, O. (1995) Presence of coelenterazine in mesopelagic fishes from the Strait of Messina. *Marine biology* **124**, 381–385. Springer.

Marek, P., Papaj, D., Yeager, J., Molina, S. & Moore, W. (2011) Bioluminescent aposematism in millipedes. *Current biology: CB* **21**, R680–R681. Elsevier.

Markova, S.V., Burakova, L.P., Golz, S., Malikova, N.P., Frank, L.A. & Vysotski, E.S. (2012) The light-sensitive photoprotein berovin from the bioluminescent ctenophore Beroe abyssicola: a novel type of Ca2+-regulated photoprotein. *The FEBS journal* **279**, 856–870. Wiley Online Library.

Markova, S.V., Golz, S., Frank, L.A., Kalthof, B. & Vysotski, E.S. (2004) Cloning and expression of cDNA for a luciferase from the marine copepod Metridia longa. A novel secreted bioluminescent reporter enzyme. *The Journal of biological chemistry* **279**, 3212–3217. ASBMB.

Markova, S.V., Larionova, M.D., Burakova, L.P. & Vysotski, E.S. (2015) The smallest natural high-active luciferase: cloning and characterization of novel 16.5-kDa luciferase from copepod Metridia longa. *Biochemical and biophysical research communications* **457**, 77–82. Elsevier.

Markova, S.V. & Vysotski, E.S. (2015) Coelenterazine Dependent Luciferases. *Biochemistry* **80**, 714–732.

McCapra, F. (1976) Chemical mechanisms in bioluminescence. *Accounts of chemical research*

9, 201–208. American Chemical Society.

McDonnell, L.A. & Heeren, R.M.A. (2007) Imaging mass spectrometry. *Mass spectrometry reviews* **26**, 606–643. Wiley Online Library.

Meighen, E.A. (1991) Molecular biology of bacterial bioluminescence. *Microbiological reviews* **55**, 123–142. Am Soc Microbiol.

Morin, J.G. (1983) Coastal Bioluminescence: Patterns and Functions. *Bulletin of marine science* **33**, 787–817. ingentaconnect.com.

Mueller, G.W. (1890) Neue Cypridiniden. *Zoologische Jahrbuecher Abteilung fuer Systematik Geographie und Biologie der Tiere*, 211–252.

Müller, O.F. (1776) *Zoologiae Danicae Prodromus, Seu Animalium Daniae Et Norvegiae Indigenarum Characteres, Nomina Et Synonyma Imprimis Popularium*. Impensis auctoris.

Nakamura, H., Kishi, Y., Shimomura, O., Morse, D. & Hastings, J.W. (1989) Structure of dinoflagellate luciferin and its enzymic and nonenzymic air-oxidation products. *Journal of the American Chemical Society* **111**, 7607–7611. American Chemical Society.

Nakamura, H., Musicki, B., Kishi, Y. & Shimomura, O. (1988) Structure of the light emitter in krill (Euphausia pacifica) bioluminescence. *Journal of the American Chemical Society* **110**, 2683–2685. American Chemical Society.

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.

Nicol, B.J.A.C. (1957) Observations on Photophores and Luminescence in the Teleost Porichthys **98**, 179–188.

Nicol, J.A.C. (1969) 7 Bioluminescence. *Fish Physiology*. Elsevier.

Nilsson, D.-E. (2009) The evolution of eyes and visually guided behaviour. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **364**, 2833–2847. royalsocietypublishing.org.

Oakley, T.H. (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. Springer.

Oakley, T.H. & Speiser, D.I. (2015) How Complexity Originates: The Evolution of Animal Eyes. *Annual review of ecology, evolution, and systematics* **46**, 237–260. Annual Reviews.

Oba, Y., Branham, M.A. & Fukatsu, T. (2011) The terrestrial bioluminescent animals of Japan. *Zoological science* **28**, 771–789. BioOne.

Oba, Y., Kato, S.I., Ojika, M. & Inouye, S. (2009) Biosynthesis of coelenterazine in the deep-sea copepod, Metridia pacifica. *Biochemical and biophysical research communications* **390**, 684–688. Elsevier Inc.

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 letters* **540**, 251–254. Elsevier.

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. Elsevier.

Oba, Y., Stevani, C.V., Oliveira, A.G., Tsarkova, A.S., Chepurnykh, T.V. & Yampolsky, I.V. (2017a) Selected Least Studied but not Forgotten Bioluminescent Systems. *Photochemistry and photobiology* **93**, 405–415. Wiley Online Library.

Oba, Y., Suzuki, Y., Martins, G.N.R., Carvalho, R.P., Pereira, T.A., Waldenmaier, H.E., Kanie, S., Naito, M., Oliveira, A.G., Dörr, F.A., Pinto, E., Yampolsky, I.V. & Stevani, C.V. (2017b) Identification of hispidin as a bioluminescent active compound and its recycling biosynthesis in the luminous fungal fruiting body. *Photochemical & photobiological sciences: Official journal of the European Photochemistry Association and the European Society for Photobiology* **16**, 1435–1440. Royal Society of Chemistry.

Oba, Y., Tsuduki, H., Kato, S.-I., Ojika, M. & Inouye, S. (2004) Identification of the Luciferin--Luciferase System and Quantification of Coelenterazine by Mass Spectrometry in the Deep-Sea Luminous Ostracod Conchoecia pseudodiscophora. *Chembiochem: a European journal of chemical biology* **5**, 1495–1499. Wiley Online Library.

Ogura, A., Ikeo, K. & Gojobori, T. (2004) Comparative analysis of gene expression for convergent evolution of camera eye between octopus and human. *Genome research* **14**, 1555–1561. genome.cshlp.org.

Oliveira, A.G., Desjardin, D.E., Perry, B.A. & Stevani, C.V. (2012) Evidence that a single bioluminescent system is shared by all known bioluminescent fungal lineages. *Photochemical & photobiological sciences: Official journal of the European Photochemistry Association and the European Society for Photobiology* **11**, 848–852.

Orbigny, A. d. (1847) *Voyage dans l'Amérique Méridionale ... exécuté pendant les années 1826-1833*. books.google.com.

Pagel, M. & Meade, A. (2006) Bayesian analysis of correlated evolution of discrete characters by

reversible-jump Markov chain Monte Carlo. *The American naturalist* **167**, 808–825. journals.uchicago.edu.

Pankey, M.S., Minin, V.N., Imholte, G.C., Suchard, M.A. & Oakley, T.H. (2014) Predictable transcriptome evolution in the convergent and complex bioluminescent organs of squid. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E4736–E4742. National Acad Sciences.

Péron, F. & Lesueur, C.-A. (1810) Tableau des caractères génériques et spécifiques de toutes les espèces de méduses connues jusqu'à ce jour. In *Annales du Muséum d'Histoire Naturelle* pp. 325–366.

Peterson, M.K. & Buck, J. (1968) Light organ fine structure in certain Asiatic fireflies. *The Biological bulletin* **135**, 335–348. journals.uchicago.edu.

Petushkov, V.N., Dubinnyi, M.A., Tsarkova, A.S., Rodionova, N.S., Baranov, M.S., Kublitski, V.S., Shimomura, O. & Yampolsky, I.V. (2014) A novel type of luciferin from the Siberian luminous earthworm Fridericia heliota: structure elucidation by spectral studies and total synthesis. *Angewandte Chemie* **53**, 5566–5568. Wiley Online Library.

Piehowski, P.D., Zhu, Y., Bramer, L.M., Stratton, K.G., Zhao, R., Orton, D.J., Moore, R.J., Yuan, J., Mitchell, H.D., Gao, Y., Webb-Robertson, B.-J.M., Dey, S.K., Kelly, R.T. & Burnum-Johnson, K.E. (2020) Automated mass spectrometry imaging of over 2000 proteins from tissue sections at 100-µm spatial resolution. *Nature communications* **11**, 8.

Prasher, D., McCann, R.O. & Cormier, M.J. (1985) Cloning and expression of the cDNA coding for aequorin, a bioluminescent calcium-binding protein. *Biochemical and biophysical research communications* **126**, 1259–1268. Elsevier.

Purtov, K.V., Petushkov, V.N., Baranov, M.S., Mineev, K.S., Rodionova, N.S., Kaskova, Z.M., Tsarkova, A.S., Petunin, A.I., Bondar, V.S., Rodicheva, E.K., Medvedeva, S.E., Oba, Y., Oba, Y., Arseniev, A.S., Lukyanov, S., et al. (2015) The Chemical Basis of Fungal Bioluminescence. *Angewandte Chemie* **54**, 8124–8128. Wiley Online Library.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A. & Zhang, F. (2013) Genome engineering using the CRISPR-Cas9 system. *Nature protocols* **8**, 2281–2308. nature.com.

Rees, J.F., De Wergifosse, B., Noiset, O., Dubuisson, M., Janssens, B. & Thompson, E.M. (1998) The origins of marine bioluminescence: Turning oxygen defence mechanisms into deep-sea communication tools. *The Journal of experimental biology* **201**, 1211–1221.

Rees, J.F., Thompson, E.M., Baguet, F. & Tsuji, F.I. (1990) Detection of Coelenterazine and Related Luciferase Activity in the Tissues of the. *Comparative biochemistry and physiology*

96A, 425–430.

Robison, B.H., Reisenbichler, K.R., Hunt, J.C. & Haddock, S.H.D. (2003) Light production by the arm tips of the deep-sea cephalopod Vampyroteuthis infernalis. *The Biological bulletin* **205**, 102–109. journals.uchicago.edu.

Rosano, G.L. & Ceccarelli, E.A. (2014) Recombinant protein expression in Escherichia coli: advances and challenges. *Frontiers in microbiology* **5**, 172. frontiersin.org.

Rosenblum, E.B., Parent, C.E. & Brandt, E.E. (2014) The Molecular Basis of Phenotypic Convergence. *Annual review of ecology, evolution, and systematics* **45**, 203–226. Annual Reviews.

Sanchez, G., Setiamarga, D.H.E., Tuanapaya, S., Tongtherm, K., Winkelmann, I.E., Schmidbaur, H., Umino, T., Albertin, C., Allcock, L., Perales-Raya, C., Gleadall, I., Strugnell, J.M., Simakov, O. & Nabhitabhata, J. (2018) Genus-level phylogeny of cephalopods using molecular markers: current status and problematic areas. *PeerJ* **6**, e4331.

Schendel, V., Rash, L.D., Jenner, R.A. & Undheim, E.A.B. (2019) The Diversity of Venom: The Importance of Behavior and Venom System Morphology in Understanding Its Ecology and Evolution. *Toxins* **11**. mdpi.com.

Schnitzler, C.E., Pang, K., Powers, M.L., Reitzel, A.M., Ryan, J.F., Simmons, D., Tada, T., Park, M., Gupta, J., Brooks, S.Y., Blakesley, R.W., Yokoyama, S., Haddock, S.H., Martindale, M.Q. & Baxevanis, A.D. (2012) Genomic organization, evolution, and expression of photoprotein and opsin genes in Mnemiopsis leidyi: a new view of ctenophore photocytes. *BMC biology* **10**, 107. Springer.

Schultz, D.T., Kotlobay, A.A., Ziganshin, R., Bannikov, A., Markina, N.M., Chepurnyh, T.V., Shakhova, E.S., Palkina, K., Haddock, S.H.D., Yampolsky, I.V. & Oba, Y. (2018) Luciferase of the Japanese syllid polychaete Odontosyllis umdecimdonta. *Biochemical and biophysical research communications* **502**, 318–323. Elsevier Ltd.

Scotland, R.W. (2011) What is parallelism? *Evolution & development* **13**, 214–227. Wiley Online Library.

Seliger, H.H. (1975) The origin of bioluminescence. *Photochemistry and photobiology* **21**, 355–361. Wiley Online Library.

Shafer, M.E.R. (2019) Cross-Species Analysis of Single-Cell Transcriptomic Data. *Frontiers in cell and developmental biology* **7**, 175. frontiersin.org.

Shimomura, O. (1985) Bioluminescence in the sea: photoprotein systems. *Symposia of the*

Society for Experimental Biology **39**, 351–372. ncbi.nlm.nih.gov.

Shimomura, O. (1986) BIOLUMINESCENCE OF THE BRITTLE STAR OPHIOPSILA CALIFORNICA. *Photochemistry and photobiology* **44**, 671–674. Wiley Online Library.

Shimomura, O. (1987) Presence of coelenterazine in non-bioluminescent marine organisms. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry* **86**, 361–363. Elsevier.

Shimomura, O. (1995) The roles of the two highly unstable components F and P involved in the bioluminescence of euphausiid shrimps. *Journal of bioluminescence and chemiluminescence* **10**, 91–101.

Shimomura, O. (2006) *Bioluminescence: Chemical Principles And Methods*. World Scientific.

Shimomura, O., Flood, P.R., Inouye, S., Bryan, B. & Shimomura, A. (2001) Isolation and properties of the luciferase stored in the ovary of the scyphozoan medusa Periphylla periphylla. *The Biological bulletin* **201**, 339–347. journals.uchicago.edu.

Shimomura, O. & Johnson, F.H. (1968) The structure of Latia luciferin. *Biochemistry* **7**, 1734–1738. ACS Publications.

Shimomura, O., Johnson, F.H. & Saiga, Y. (1962) Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *Journal of cellular and comparative physiology* **59**, 223–239. Wiley Online Library.

Specht, H. & Slavov, N. (2018) Transformative Opportunities for Single-Cell Proteomics. *Journal of proteome research* **17**, 2565–2571. ACS Publications.

Ståhl, P.L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J.F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J.O., Huss, M., Mollbrink, A., Linnarsson, S., Codeluppi, S., Borg, Å., Pontén, F., et al. (2016) Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78–82. science.sciencemag.org.

Stauber, J., MacAleese, L., Franck, J., Claude, E., Snel, M., Kaletas, B.K., Wiel, I.M.V.D., Wisztorski, M., Fournier, I. & Heeren, R.M.A. (2010) On-tissue protein identification and imaging by MALDI-ion mobility mass spectrometry. *Journal of the American Society for Mass Spectrometry* **21**, 338–347. ACS Publications.

Steindachner, F. (1870) Ichthyologische Notizen, X: Barbus albanicus n. sp.--Sitzungsber. k. *Akad. Wiss. Math. -Naturwiss. Cl. Abt* **1**, 630–631.

Stern, D.L. (2013) The genetic causes of convergent evolution. *Nature reviews. Genetics* **14**,

751–764.

Stern, D.L. & Orgogozo, V. (2009) Is genetic evolution predictable? *Science* **323**, 746–751. science.sciencemag.org.

Stoeckli, M., Chaurand, P., Hallahan, D.E. & Caprioli, R.M. (2001) Imaging mass spectrometry: a new technology for the analysis of protein expression in mammalian tissues. *Nature medicine* **7**, 493–496. researchgate.net.

Strum, J.M. (1969) Fine structure of the dermal luminescent organs, photophores, in the fish, Porichthys notatus. *The Anatomical record* **164**, 433–461. Wiley Online Library.

Su, Y., Shi, Q. & Wei, W. (2017) Single cell proteomics in biomedicine: High-dimensional data acquisition, visualization, and analysis. *Proteomics* **17**, 1600267. Wiley Online Library.

Suzuki, T.K. (2017) On the Origin of Complex Adaptive Traits: Progress Since the Darwin Versus Mivart Debate. *Journal of experimental zoology. Part B, Molecular and developmental evolution* **328**, 304–320. Wiley Online Library.

Sweeney, B.M. (1980) Intracellular Source of Bioluminescence. *International review of cytology* **68**, 173–195.

Takahashi, H. & Isobe, M. (1994) Photoprotein of Luminous Squid, Symplectoteuthis oualaniensis and Reconstruction of the Luminous System. *Chemistry letters* **23**, 843–846. The Chemical Society of Japan.

Takenaka, Y., Noda-Ogura, A., Imanishi, T., Yamaguchi, A., Gojobori, T. & Shigeri, Y. (2013) Computational analysis and functional expression of ancestral copepod luciferase. *Gene* **528**, 201–205. Elsevier B.V.

Thacker, C.E. & Roje, D.M. (2009) Phylogeny of cardinalfishes (Teleostei: Gobiiformes: Apogonidae) and the evolution of visceral bioluminescence. *Molecular phylogenetics and evolution* **52**, 735–745. Elsevier Inc.

Thanwisai, A., Tandhavanant, S., Saiprom, N., Waterfield, N.R., Ke Long, P., Bode, H.B., Peacock, S.J. & Chantratita, N. (2012) Diversity of Xenorhabdus and Photorhabdus spp. and their symbiotic entomopathogenic nematodes from Thailand. *PloS one* **7**, e43835.

Thompson, E.M., Nafpakntis, B.G. & Tsuji, F.I. (1987) Induction of bioluminescence in the marine fish, Porichthys, by Vargula (Crustacean) luciferin. Evidence for de novo synthesis or recycling of luciferin. *Photochemistry and photobiology* **45**, 529–533.

Thomson, C.M., Herring, P.J. & Campbell, A.K. (1995) Evidence for De Novo Biosynthesis of

Coelenterazine in the Bioluminescent Midwater Shrimp, Systellaspis Debilis. *Journal of the Marine Biological Association of the United Kingdom. Marine Biological Association of the United Kingdom* **75**, 165–171. Access paid by the UCSB Libraries.

Thomson, C.M., Herring, P.J. & Campbell, A.K. (1997) The widespread occurrence and tissue distribution of the imidazolopyrazine luciferins. *Journal of bioluminescence and chemiluminescence* **12**, 87–91. Wiley Online Library.

Thuesen, E.V., Goetz, F.E. & Haddock, S.H.D. (2010) Bioluminescent organs of two deep-sea arrow worms, Eukrohnia fowleri and Caecosagitta macrocephala, with further observations on bioluminescence in chaetognaths. *The Biological bulletin* **219**, 100–111.

Tong, D., Rozas, N.S., Oakley, T.H., Mitchell, J., Colley, N.J. & McFall-Ngai, M.J. (2009) Evidence for light perception in a bioluminescent organ. *Proceedings of the National Academy of Sciences* **106**, 9836–9841.

Tsuji, F.I., Barnes, A.T. & Case, J.F. (1972) Bioluminescence in the marine teleost, Porichthys notatus, and its induction in a non-luminous form by Cypridina (ostracod) luciferin. *Nature* **237**, 515–516. nature.com.

Tsuji, F.I., Haneda, Y., Lynch, R.V. & Sugiyama, N. (1971) Luminescence cross-reactions of Porichthys luciferin and theories on the origin of luciferin in some shallow-water fishes. *Comparative biochemistry and physiology* **4**, 163–179.

Tsuji, F.I., Lynch, R.V., 3rd & Haneda, Y. (1970) Studies on the bioluminescence of the marine ostracod crustacean Cypridina serrata. *The Biological bulletin* **139**, 386–401. journals.uchicago.edu.

Tsuji, F.I., Nafpaktitist, B.G., Goto, T., Cormier, M.J., Wampler, J.E. & Anderson, J.M. (1975) Spectral characteristics of the bioluminescence induced in the marine fish, Porichthys notatus, by Cypridina (Ostracod) luciferin. *Molecular and cellular biochemistry* **9**.

Tsuji, F.I., Ohmiya, Y., Fagan, T.F., Toh, H. & Inouye, S. (1995) Molecular evolution of the Ca2+-binding photoproteins of the hydrozoa. *Photochemistry and photobiology* **62**, 657–661. Wiley Online Library.

Vassel, N., Cox, C.D., Naseem, R., Morse, V., Evans, R.T., Power, R.L., Brancale, A., Wann, K.T. & Campbell, A.K. (2012) Enzymatic activity of albumin shown by coelenterazine chemiluminescence. *Luminescence: the journal of biological and chemical luminescence* **27**, 234–241.

Verdes, A. & Gruber, D.F. (2017) Glowing Worms: Biological, Chemical, and Functional Diversity of Bioluminescent Annelids. *Integrative and comparative biology* **57**, 18–32.

academic.oup.com.

Viviani, V.R. (2002) The origin, diversity, and structure function relationships of insect luciferases. *Cellular and molecular life sciences: CMLS* **59**, 1833–1850. Springer.

Viviani, V.R., Silva, J.R., Amaral, D.T., Bevilaqua, V.R., Abdalla, F.C., Branchini, B.R. & Johnson, C.H. (2020) A new brilliantly blue-emitting luciferin-luciferase system from Orfelia fultoni and Keroplatinae (Diptera). *Scientific reports* **10**, 9608. nature.com.

Wang, M., Carver, J.J., Phelan, V.V., Sanchez, L.M., Garg, N., Peng, Y., Nguyen, D.D., Watrous, J., Kapono, C.A., Luzzatto-Knaan, T., Porto, C., Bouslimani, A., Melnik, A.V., Meehan, M.J., Liu, W.-T., et al. (2016) Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature biotechnology* **34**, 828–837. nature.com.

Ward, W.W. & Cormier, M.J. (1978) Energy transfer via protein-protein interaction in Renilla bioluminescence. *Photochemistry and photobiology* **27**, 389–396. Wiley Online Library.

Warrant, E.J. & Locket, N.A. (2004) Vision in the deep sea. *Biological reviews of the Cambridge Philosophical Society* **79**, 671–712. Wiley Online Library.

Watkins, O.C., Sharpe, M.L., Perry, N.B. & Krause, K.L. (2018) New Zealand glowworm (Arachnocampa luminosa) bioluminescence is produced by a firefly-like luciferase but an entirely new luciferin. *Scientific reports* **8**, 3278. nature.com.

Widder, E.A. (2010) Bioluminescence in the Ocean: Origins of Biological, Chemical, and Ecological Diversity. *Science*. Http://dx.doi.org/10.1126/science.1174269.

Widder, E.A., Greene, C.H. & Youngbluth, M.J. (1992) Bioluminescence of sound-scattering layers in the Gulf of Maine. *Journal of plankton research* **14**, 1607–1624. Oxford University Press.

Wilson, T. & Hastings, J.W. (1998) Bioluminescence. *Annual review of cell and developmental biology*, 197–230.

Wong, J.M., Pérez-Moreno, J.L., Chan, T.Y. & Frank, T.M. (2015) Phylogenetic and transcriptomic analyses reveal the evolution of bioluminescence and light detection in marine deep-sea shrimps of the family Oplophoridae. *Molecular phylogenetics and evolution*. Elsevier.

Yang, J.Y., Sanchez, L.M., Rath, C.M., Liu, X., Boudreau, P.D., Bruns, N., Glukhov, E., Wodtke, A., de Felicio, R., Fenner, A., Wong, W.R., Linington, R.G., Zhang, L., Debonsi, H.M., Gerwick, W.H., et al. (2013) Molecular networking as a dereplication strategy.

Journal of natural products **76**, 1686–1699. ACS Publications.

Young, R.E. (1983) Oceanic Bioluminescence: an Overview of General Functions. *Bulletin of marine science* **33**, 829–845. ingentaconnect.com.

Young, R.E. & Bennett, T.M. (2013) Enoploteuthinae (Cephalopoda). *Paleontology and Neontology of Cephalopods*, 241. Elsevier.