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

Emily S. Lau1* and Todd H. Oakley1*

1Ecology, 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
II. Bioluminescence: an excellent system for studying convergence in complex traits
   (1) Harnessing biochemical reactions to produce light
   (2) Widespread convergent evolution of bioluminescence
Evolutionary origins of bioluminescence
(a) Origin of luciferases and photoproteins
(b) Origins of luciferins

III. Reviewing the multi-level convergent evolution of bioluminescence: aims and scope
(1) Coelenterazine systems
(2) Vargulin systems

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

V. Promising new tools for bioluminescence research
(1) Genetic sequencing
(2) Mass spectrometry

VI. Conclusions
. Acknowledgements

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 colored 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](image)

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)*
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; Obi 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, *Amphipolis 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 *Amphipora filiformis* (Müller, 1776) and *Ophiopsila Californica* (Clark, 1921), which use a luciferase and calcium dependent photoprotein, respectively (Shimomura, 1986; Delrosie 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 varguin 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 apognids 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


gelatinous zooplankton: ctenophores, medusae and siphonophores. *Marine biology* 133,
571–582. Springer.

of marine science* 2, 443–493.

Dietary requirement for luciferin in cnidarian bioluminescence. *Proceedings of the National
Academy of Sciences of the United States of America* 98, 11148–11151.


Haneda, Y. & Johnson, F.H. (1958) The luciferin-luciferase reaction in a fish, Parapriacanthus
beryctiformis, of newly discovered luminescence. *Proceedings of the National Academy of

of Parapriacanthus ransonneti, Pempheris klunzingeri, and Apogon ellioti. *Bioluminescence

the plain ® sh midshipman Porichthys notatus 001, 529–540.


287.

Hastings, J.W. (1983) Biological diversity, chemical mechanisms, and the evolutionary origins

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

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*


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


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


751–764.


academic.oup.com.


