Reference: *Biol. Bull.* **213:** 95–100. (October 2007)

Endogenous Green Fluorescent Protein (GFP) in Amphioxus

DIMITRI D. DEHEYN[1,*], KAORU KUBOKAWA[2], JAMES K. MCCARTHY[1], AKIO MURAKAMI[3], MAGALI PORRACHIA[1], GREG W. ROUSE[1], AND NICHOLAS D. HOLLAND[1]

[1]*Marine Biology Research Division, Scripps Institution of Oceanography (UCSD), La Jolla, California 92093-0202;* [2]*Research Institute, University of Tokyo, Nakano, Tokyo, 164-8639, Japan;* [3]*Research Center for Inland Seas, Kobe University, Awaji, Hyogo, 656-2401, Japan*

Green fluorescent proteins (GFPs) are well known for their intensive use in cellular and molecular biology in applications that take advantage of the GFPs self-folding and built-in fluorophore characteristics as biomarker. Occurrence and function of GFPs in nature is less known. For a long time GFPs were described only from some cnidarians, and it is only recently that they were also found in copepod crustaceans. Here we describe the occurrence of a GFP from three species of amphioxus, namely Branchiostoma floridae, B. lanceolatum, *and* B. belcheri *(Chordata: Cephalochordata). This is the first time an endogenous GFP has been found in any representative of the deuterostome branch of the Animal Kingdom. We have isolated and characterized a gene (*AmphiGFP*) from* B. floridae *that encodes a GFP protein related to those of cnidarians and copepods in both its amino acid sequence and its predicted higher order structure (an 11-stranded β-barrel enclosing a fluorophore). Bayesian and maximum parsimony phylogenetic analyses demonstrate that the AmphiGFP protein is markedly more closely related to copepod than to cnidarian GFPs. In adults of all three amphioxus species, the green fluorescence is strikingly concentrated anteriorly. The anterior end is the only body part exposed to light in these shallow-water dwellers, suggesting possible photoreceptive or photoprotective functions for the endogenous GFP.*

Green fluorescent proteins (GFPs) are familiar to most biologists as invaluable tools for cellular and molecular biology (1). However, in spite of the considerable effort spent on developing GFPs for laboratory reagents, much remains to be learned about the taxonomic distribution and biological function of these proteins in nature. To date, GFPs have been found in only two major groups in the metazoan tree: specifically, in a number of cnidarians, relatively near the base of the tree, and in a few copepod crustaceans, relatively derived within the protostome branch (2, 3). The cnidarian GFPs are often associated with bioluminescence, but those found so far in copepods are not. We now report that the limited taxonomic distribution of animals with endogenous GFPs may be partially due to inadequate sampling efforts, because we have found such molecules in the cephalochordate amphioxus. About 10 years ago, we began to suspect that endogenous GFPs are present in amphioxus, because the eggs and embryos emit a uniform green fluorescence when illuminated with UV light (this phenomenon is illustrated in reference 4). The present note reports on the isolation and molecular characterization of indubitable GFPs from three amphioxus species (none of which are bioluminescent). This is the first demonstration of the presence of these distinctive molecules in any deuterostome. In addition, the tissue distribution of amphioxus fluorescence, interestingly localized at the anterior end of the adult body, gives insights into possible functions of the endogenous GFPs (discussed below).

The present note concerns three amphioxus species: *Branchiostoma floridae* Hubbs, 1922 (the Florida amphioxus, collected in Tampa, Florida), *Branchiostoma lanceolatum* (Pallas, 1774) (the European amphioxus, collected in Banyuls-sur-Mer, France), and *Branchiostoma belcheri* Gray, 1847 (the Asian amphioxus, collected in Enshu-nada Sea, Japan). For adults of these three species, the fluorescence spectra, stimulated by incident UV (380 nm), had

Received 23 January 2007; accepted 8 May 2007.

* To whom correspondence should be addressed. E-mail: ddeheyn@ucsd.edu

peaks at 524 nm, 526 nm, and 527 nm, respectively (Fig. 1). To establish a link between this green fluorescence and a possible endogenous GFP in the tissues of amphioxus, a cnidarian protein (GenBank AY157666) was blasted (tblastn) using the EST_OTHERS database in GenBank. The search identified numerous clones of GFP from unfertilized egg, gastrula, neurula, larval, and adult libraries for *B. floridae*, all sharing the same sequence that was yet distinct from the cnidarian sequence. Using primers based on the EST nucleotide sequence, we obtained cDNA of the expressed gene by reverse transcriptase RT-PCR with mRNA extracted from the adult *B. floridae* as the template. The cDNA sequence is identical to the EST sequence. The sequence, which we named *AmphiGFP* and deposited in GenBank (EF157660), encodes a protein of 218 amino acids. Conversion of the amino acid sequence of AmphiGFP to a three-dimensional structure by Swiss-Model (ExPASy server) and modeling from the known crystal structure of a fluorescent protein from a copepod (5) showed that the predicted higher order structure of the amphioxus protein closely resembles that of endogenous fluorescent proteins in cnidarians and copepods. All these molecules compose an 11-stranded β-barrel enclosing a central strand that includes a cyclized tripeptide fluorophore—based on glycine-tyrosine-glycine in both amphioxus and copepods, but on serine-tyrosine-glycine in most fluorescent cnidarians.

We used UV irradiation (380 nm) to study the tissue distribution of the green fluorescence in living developmental stages and adults of the Florida amphioxus, in living adults of the European amphioxus, and in RNA*later*-preserved adults of the Asian amphioxus (Fig. 2). In the Florida species, the fluorescence, which is ubiquitous in the eggs and larvae (4), first becomes patchily distributed in the larvae (Fig. 2A, B), and finally becomes localized at the anterior end of the juveniles and adults, exclusively in the support cells of the oral cirri (although not in the skeletal arch from which they spring) (Fig. 2C, D). In ripening adults of the Florida amphioxus, fluorescence is also detected in the ovaries, specifically in the growing oocytes, but not in the testes (data not shown). In adults of the European amphioxus, UV also induces green fluorescence intensely in the oral cirri and more diffusely in the epidermis—chiefly at the anterior end of the body, but also very inconspicuously near the posterior end (Fig. 2E, F); the anterior fluorescence is in the support cells of the cirri, but not in the skeletal arch (Fig. 2G). Adults of the Asian amphioxus also emit green fluorescence—strongly from the supporting cells of the oral cirri and their skeletal arch support (Fig. 2H), as well as weakly from the epidermis at the anterior end of the animal.

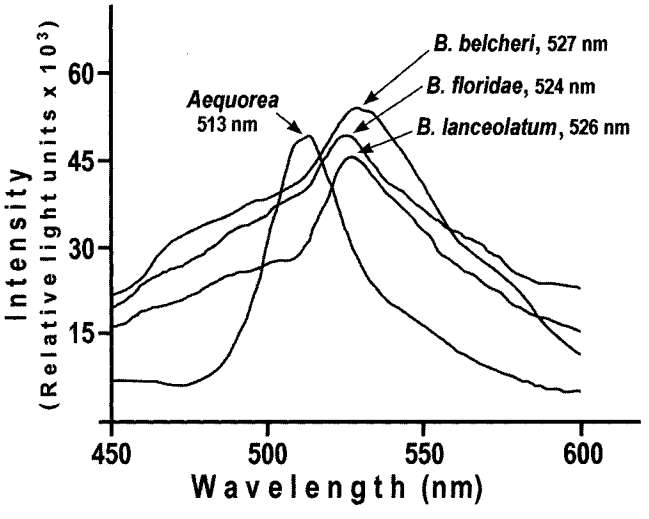

Figure 1. Fluorescence spectra (with emission maxima) from intact adult amphioxus illuminated with UV (380 nm) and measured by SE200 low-light digital spectrograph (Catalina Scientific Instruments, Tucson, AZ). Spectra from living specimens of *Branchiostoma floridae* and *B. lanceolatum*, and from a specimen of *B. belcheri* preserved in RNA*later* (Ambion, Austin, TX) (exposure to RNA*later* does not alter GFP emission spectra; our unpublished observations). For comparison, emission spectrum for EGFP from hydrozoan cnidarian *Aequorea* (Biovision Research Products, Mountain View, CA) measured in same apparatus is included.

AmphiGFP belongs to the 11-stranded β-barrel superfamily of proteins. This superfamily is considered to include some proteins that fluoresce in colors other than green, some chromoproteins that do not fluoresce at all, and G2FP motif proteins, which are components of the extracellular matrix of most metazoans (6–9). We assessed the relationships between AmphiGFP and other known proteins in the superfamily by phylogenetic analyses with Bayesian and maximum parsimony methods (Fig. 3). In our analysis, amphioxus GFP is more closely related to fluorescent proteins of copepods than to those of cnidarians, an arrangement in accord with current and previous hypotheses of metazoan phylogeny (10). A fair indication of the degree of similarity among fluorescent proteins of these three animal groups is the only 19% amino acid identities between AmphiGFP and a cnidarian (*Aequorea victoria*) GFP, as contrasted to 35% amino acid identities between AmphiGFP and a copepod (*Pontellina plumata* ppluGFP2) GFP. Importantly, our phylogenetic analysis is consistent with earlier work (2) suggesting that fluorescent proteins of bilaterian animals originated from a single ancestral fluorescent protein in a basal metazoan. Thus, AmphiGFP is evidently not independently derived from a deuterostome G2FP motif protein (several of which are included in our analysis).

The ecological significance of endogenous fluorescent proteins is poorly understood (2, 11, 12). Even so, the tissue distribution of AmphiGFP suggests a couple of possible functions for the endogenous molecule in amphioxus. One is photoreception, as indicated by the localization of fluorescence in support cells of the oral cirri, although these structures are not one of the four currently recognized photoreceptive tissues in amphioxus (13). There is, however, ultrastructural evidence consistent with the possible

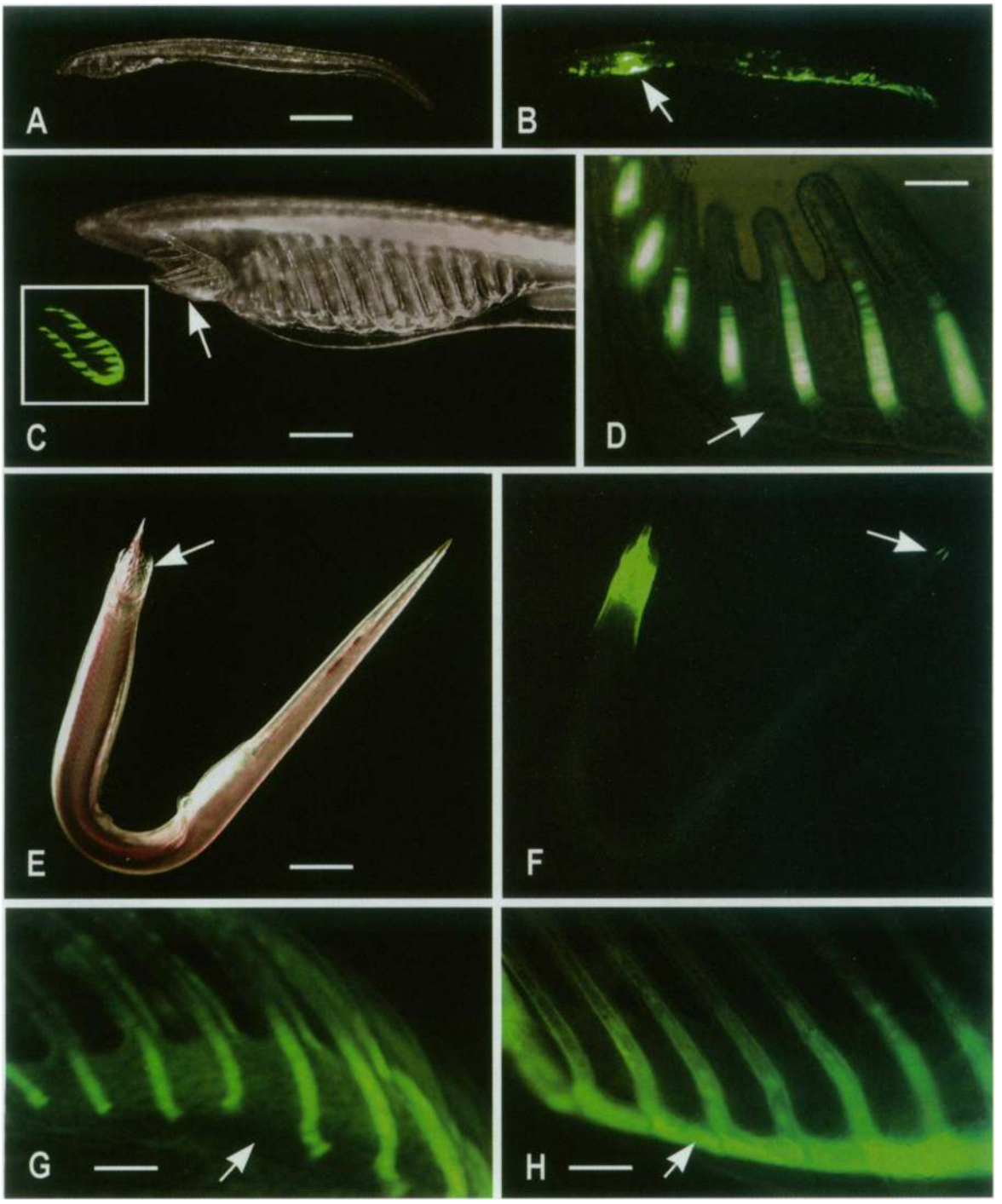

Figure 2. Fluorescence of intact specimens of three amphioxus species under UV illumination (380 nm). (A) Early larva of *Branchiostoma floridae* under visible light; (B) same under UV with patchy green fluorescence, most intense just posterior to the mouth (arrow). (C) Postmetamorphic juvenile of *B. floridae* under visible light, with oral cirri indicated by arrow; inset, oral cirri of same animal fluorescing under UV. (D) Detail of fluorescing support cells of *B. floridae* in oral cirri of adult under UV (arrow indicates that no cells of skeletal arch fluoresce). (E) Adult specimen of *B. lanceolatum* under visible light, with oral cirri indicated by arrow. (F) Same animal under UV with fluorescence conspicuous anteriorly and weak posteriorly (arrow). (G) Detail of oral cirri of adult *B. lanceolatum* (arrow indicates no cells in skeletal arch fluoresce). (H) Detail of oral cirri of adult *B. belcheri* (arrow indicates fluorescence in skeletal arch). Scale line is 100 μm for A and B; 250 μm for D; 300 μm for C, G, and H; and 3 mm for E and F. Scale line for B is the same as A; scale line for F is the same as E.

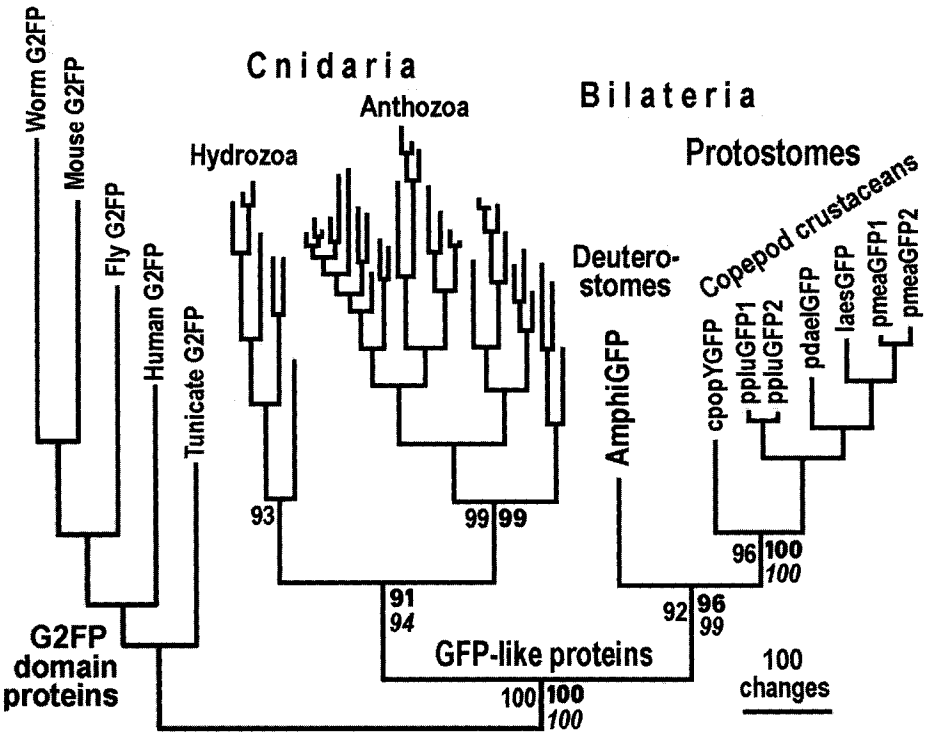

Figure 3. Phylogenetic analysis of the 11-stranded β-barrel protein superfamily in animals. Sequences for AmphGFP and a recently discovered copepod fluorescent protein from *Chiridius poppei* (9) were aligned, using ClustalW (29) and manual adjustment, against the cDNA dataset and alignment from a previous analysis (2); this database includes a range of fluorescent proteins from Cnidaria and copepod arthropods as well as G2FP motif proteins from three deuterostomes (tunicate *Halocynthia roretzi*, mouse *Mus musculus*, and human *Homo sapiens*) and two protostomes (worm *Coenorhabditis elegans*, and fly *Drosophila melanogaster*). The alignment used and results are available at Treebase (www.treebase.org). Maximum parsimony analyses of the 816 characters (696 parsimony informative) were performed using PAUP* 40b10 (30), with default settings, except for stepwise addition using 100 random-addition sequences. This resulted in a single shortest tree of length 5335. Bootstrap and jackknife (with 37% deletion of characters) values were assessed in PAUP* using 1000 replicates, and the majority-rule consensus tree matched the most parsimonious tree. Bayesian analyses with MrBayes 3.1 (31) were run with default priors (rate matrix: 0–100, branch lengths: 0–10, gamma shape: 0–1), a random starting tree, and six Markov chains. MrModelTest 2.2 (32) indicated use of the GTR+I+G model under both the hLRT and AIC criteria. Of the 1,100,000 generations run, with a tree saved every 1000 generations, the first 100,000 generations (100 trees) were discarded as burn-in. The majority-rule consensus tree of the remaining 1000 trees for each analysis gave the posterior probabilities for each clade. Both maximum parsimony and Bayesian analyses gave congruent results and show that AmphGFP is most closely related to its homologs in copepod arthropods (protostomes) and is not derived independently from non-fluorescent G2FP proteins. Bayesian posterior probability values to the left of supported nodes; bootstrap (bold) and parsimony jackknife (italics) values to the right (values less than 90% not shown).

role of oral cirri in photoreception: cirral support cells stack on top of each other, delimiting between-cell extracellular pockets into which project microvilli and cilia (one per cell) (14, see fig. 89), the whole stack of cells being associated with axonal processes of presumed sensory neurons (15). These characteristics are those of rhabdomeric photoreceptor organs (16, 17), whose activity could be coupled with AmphiGFP given that, like other fluorescent proteins, AmphiGFP can probably undergo reversible photochemical transformations (18) and thus has the potential to transduce light energy into chemical energy. The photoreceptive function is also suggested when we consider the ecological behavior of the animal, which lives burrowed in the sand except for its head, from which oral cirri face the water column and thus the downwelling sunlight. The animal in this almost-completely-burrowed position has long been shown to be sensitive to change in light exposure (19).

A second possible function of AmphiGFP may be photoprotection against intense visible light, lower wavelength (UVA, blue) light, or both. This possibility is suggested by the presence of fluorescence throughout the tissues of the pelagic amphioxus embryos as well by as the concentration of fluorescence at the anterior end of the body of the adult—the end that usually projects slightly from the bur-

row of these relatively shallow-living marine animals. At a first level of defense, the aromatic fluorophore core of AmphiGFP could absorb high-energy light and scatter it or dissipate it as less damaging, lower energy fluorescence (12, 20, 21), possibly by a mechanism involving Förster resonance energy transfer (22). Moreover, at a second level of defense, the molecule could act as an antioxidant to detoxify reactive oxygen/free radicals (23), because imidazolinone fluorophores are known to have a high affinity for molecular oxygen (24–26).

In conclusion, with our demonstration of AmphiGFP, amphioxus becomes the only deuterostome known to contain an endogenous fluorescent protein. Even with this discovery, the distribution of fluorescent proteins among animals remains sparse and widely scattered—with known representatives in only one isolated group of deuterostomes (amphioxus), in one isolated group of protostomes (a few copepods), and in one group of relatively basal metazoans (namely, some hydrozoan and anthozoan cnidarians). This sparse distribution could be indicative of horizontal gene transfer, although there are not many well-accepted examples of this phenomenon in metazoans (27, 28); of secondary loss from most taxa; or of inadequate taxonomic sampling. It is possible that more members of the highly distinctive 11-stranded β-barrel protein superfamily (other than the ubiquitous G2FP proteins) remain to be discovered, and some of these, not necessarily fluorescent, might be relatively common and involved in functions more general than the production of conspicuous fluorescence.

Acknowledgments

We are indebted to L. Z. Holland, J. M. Lawrence, M. Izeki, M. Paris, S. Podell, and K. Thamatrakoln for assistance and advice. This study was partly under the auspices of the AFOSR Biomimetics, Biomaterials, and Biointerfacial Sciences program (funding to D. D. D.) and KAKENHI (Grant-in-Aid for Scientific Research) from the Ministry of Education, Culture, Sports, Science and Technology of Japan (to K. K.).

Literature Cited

1. **Stewart, C. N. 2006.** Go with the glow: fluorescent proteins to light transgenic organisms. *Trends Biotechnol.* **24:** 155–162.
2. **Shagin, D. A., E. V. Yanushevich, A. F. Fradkov, K. A. Lukyanov, Y. A. Labas, T. N. Semenova, J. A. Ugalde, A. Meyers, J. N. Nunez, E. A. Widder, S. A. Lukyanov, and M. V. Matz. 2004.** GFP-like proteins as ubiquitous metazoan superfamily: evolution of functional features and structural complexity. *Mol. Biol. Evol.* **21:** 841–850.
3. **Shimomura, O. 2005.** The discovery of aequorin and green fluorescent protein. *J. Microsc.* **217:** 3–15.
4. **Yu, J. K., N. D. Holland, and L. Z. Holland. 2004.** Tissue-specific expression of *FoxD* reporter constructs in amphioxus embryos. *Dev. Biol.* **274:** 452–461.
5. **Wilmann, P. G., J. Battad, J. Petersen, M. C. J. Wilce, S. Dove, R. J. Devenish, M. Prescott, and J. Rossjohn. 2006.** The 2.1 Å crystal structure of copGPF, a representative member of the copepod clade within the green fluorescent protein superfamily. *J. Mol. Biol.* **359:** 890–900.
6. **Hopf, M., W. Göhring, A. Ries, R. Timpl, and E. Hohenester. 2001.** Crystal structure and mutational analysis of a perlecan-binding fragment of nidogen-1. *Nat. Struct. Biol.* **8:** 634–640.
7. **Labas, Y. A., N. G. Gurskaya, Y. G. Yanushevich, A. F. Fradkov, K. A. Lukyanov, S. A. Lukyanov, and M. V. Matz. 2002.** Diversity and evolution of the green fluorescent protein family. *Proc. Natl. Acad. Sci. USA* **99:** 4256–4261.
8. **Prescott, M., M. Ling, T. Beddoe, A. J. Oakley, S. Dove, O. Hoegh-Guldberg, R. J. Devenish, and J. Rossjohn. 2003.** The 2.2 Å crystal structure of pocilloporin pigment reveals a non planar chromophore conformation. *Structure* **11:** 275–284.
9. **Masuda, H., T. Takenaka, A. Yamaguchi, S. Nishikawa, and H. Mizuno. 2006.** A novel yellowish-green fluorescent protein from the marine copepod *Chiridius poppei*, and its use as a reporter protein in HeLa cells. *Gene* **372:** 18–25.
10. **Rokas, A., D. Krüger, and S. B. Carroll. 2005.** Animal evolution and the molecular signature of radiations compressed in time. *Science* **310:** 1933–1938.
11. **Herring, P. J. 1987.** Systematic distribution of bioluminescence in living organisms. *J. Biolumin. Chemilumin.* **1:** 147–163.
12. **Dove, S. G., O. Hoegh-Guldberg, and S. Ranganathan. 2001.** Major colour patterns of reef-building corals are due to a family of GFP-like proteins. *Coral Reefs* **19:** 197–204.
13. **Glardon, S., L. Z. Holland, W. J. Gehring, and N. D. Holland. 1998.** Isolation and developmental expression of the amphioxus Pax-6 gene (*AmphiPax6*): insights into eye and photoreceptor evolution. *Development* **125:** 2701–2710.
14. **Ruppert, E. E. 1997.** Cephalochordata (Acrania). Pp. 349–504 in *Microscopic Anatomy of Invertebrates,* Vol. 1, *Hemichordata, Chaetognatha, and the Invertebrate Chordates,* G. W. Harrison and E. E. Ruppert, eds. Wiley-Liss, New York.
15. **Demski, L. S., J. A. Beaver, and J. B. Morrill. 1996.** The cutaneous innervation of amphioxus: a review incorporating new observations with DiI tracing and scanning electron microscopy. *Isr. J. Zool.* **42 Suppl.:** 117–120.
16. **Ruiz, M. S., and R. Anadon. 1991.** Some considerations on the fine-structure of rhabdomeric photoreceptors in the amphioxus, *Branchiostoma lanceolatum* (Cephalochordata). *J. Hirnforschung* **32:** 159–164.
17. **Arendt, D. 2003.** Evolution of eyes and photoreceptor cell types. *Int. J. Dev. Biol.* **47:** 563–571.
18. **Andresen, M., M. C. Wahl, A. C. Stiel, F. Gräter, L. V. Schäfer, S. Trowitzsch, G. Weber, C. Eggeling, H. Grubmüller, S. W. Hell, and S. Jakobs. 2005.** Structure and mechanism of the reversible photoswitch of a fluorescent protein. *Proc. Natl. Acad. Sci. USA* **102:** 13070–13074.
19. **Parker, G. H. 1908.** The sensory reactions of amphioxus. *Proc. Am. Acad. Arts Sci.* **43:** 415–455.
20. **Salih, A., A. Larkum, G. Cox, M. Kuhl, and O. Hoegh-Guldberg. 2000.** Fluorescent pigments in corals are photoprotective. *Nature* **408:** 850–853.
21. **Oswald, F., F. Schmitt, A. Leutenegger, S. Ivanchenko, C. D'Angelo, A. Salih, S. Maslakova, M. Bulina, R. Schirmbeck, G. U. Nienhaus, M. Matz, and J. Wiedenmann. 2007.** Contributions of host and symbiont pigments to the coloration of reef corals. *FEBS J.* **274:** 1102–1109.
22. **Gilmore, A. M., A. W. D. Larkum, A. Salih, S. Itoh, Y. Shibata, C. Bena, H. Yamasaki, M. Papina, and R. Van Woesik. 2003.** Simultaneous time resolution of the emission spectra of fluorescent proteins and zooxanthellar chlorophyll in reef-building corals. *Photochem. Photobiol.* **77:** 515–523.

23. **Bou-Abdallah, F., N. D. Chasteen, and M. P. Lesser. 2006.** Quenching of superoxide radicals by green fluorescent protein. *Biochim. Biophys. Acta* **1760:** 1690–1695.
24. **Heim, R., D. C. Prasher, and R. Y. Tsien. 1994.** Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc. Natl. Acad. Sci. USA* **91:** 12501–12504.
25. **Inouye, S., and F. I. Tsuji. 1994.** Evidence for redox forms of the *Aequorea* green fluorescent protein. *FEBS Lett.* **351:** 211–214.
26. **Rosenow, M. A., H. N. Patel, and R. M. Wachter. 2005.** Oxidative chemistry in the GFP active site leads to covalent cross-linking of a modified leucine side chain with a histidine imidazole: implications for the mechanism of chromophore formation. *Biochemistry* **44:** 8303–8311.
27. **Steele, R. E., S. E. Hampson, N. A. Stover, D. F. Kiber, and H. R. Bode. 2004.** Probable transfer of a gene between a protist and a cnidarian. *Curr. Biol.* **14:** R298–R299.
28. **Bapteste, E., E. Susko, J. Leigh, D. MacLeod, R. L. Charlebois, and W. F. Doolittle. 2005.** Do orthologous gene phylogenies really support tree-thinking? *BMC Evol. Biol.* **5:** 33.
29. **Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994.** CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22:** 4673–4680.
30. **Swofford, D. L. 2002.** *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods).* Sinauer, Sunderland, MA.
31. **Ronquist, F., and J. P. Huelsenbeck. 2003.** 1003 MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19:** 1572–1574.
32. **Nylander, J. A. A. 2004.** *MrModeltest v2.2.* Program distributed by the author. Evolutionary Biology Centre, Uppsala University, Sweden.