Evolution of Coral Pigments Recreated

Juan A. Ugalde,[1,3] Belinda S. W. Chang,[4] Mikhail V. Matz[1,2]*

In proteins homologous to the green fluorescent protein from *Aequorea victoria* (GFP), green and cyan emission colors require two consecutive autocatalytic reactions to complete chromophore synthesis. Red fluorescent proteins and purple-blue chromoproteins require a third reaction, thereby manifesting a higher level of functional complexity (*1*). Multiple events of red/green color diversification within the GFP superfamily (*2*) may therefore reflect convergent evolution of molecular complexity.

To examine this issue, we studied one of these events, that which gave rise to the color diversity exhibited by the great star coral *Montastraea cavernosa.* This coral possesses several genes coding for fluorescent GFP-like proteins of cyan, shortwave green, longwave green, and red emission colors (*3*) (Fig. 1C). We statistically inferred and synthesized the ancestral genes corresponding to the common ancestor of all *M. cavernosa* colors ("ALL ancestor"), the common ancestor of red proteins ("Red ancestor"), and two intermediate nodes corresponding to the possible common ancestors of red and longwave green proteins ("Red/Green ancestor" and "pre-Red ancestor") (Fig. 1B). Bacteria were transformed with plasmids carrying the ancestral genes, and the color of the expressed proteins was evaluated spectroscopically (Fig. 1A). Green color at any of the ancestral nodes would indicate that red proteins from other parts of the phylogeny [such as dendRFP (Fig. 1B)] were the result of convergent evolution, whereas if red only arose once, then all ancestors would be red.

Statistical inference of the ancestral sequences was performed with the data set described elsewhere (*2*, *3*) under three models of evolution based on different types of sequence information: amino acids, codons, and nucleotides (*4*). The reconstructions of all four ancestral sequences were robust under these models, with average posterior probabilities at a site ranging from 0.96 to 0.99. Still, the models were in disagreement at several sites (between 4 and 8 out of a total of 217). When planning ancestral gene synthesis, we designed the codons corresponding to these sites to be degenerate in order to incorporate alternative predictions.

For each type of the ancestral gene, the protein products displayed identical fluorescent phenotypes, even though the gene sequences were different at the degenerate sites. The ALL ancestor turned out to be shortwave green. The two possible common ancestors of red and green proteins (Red/Green and pre-Red) showed an intermediate longwave green/red phenotype: Although the main bulk of the protein remained longwave green, a small fraction was able to complete the third chromophore synthesis reaction, resulting in a minor peak of red emission. Clones of the Red ancestor showed an "imperfect red" phenotype: Although the red emission peak dominated, the rate of green to red conversion was still less efficient than in extant reds, resulting in a prominent peak of green fluorescence (Fig. 1, A and C).

Our results indicate that because the ALL ancestor was green and not red, red color within the superfamily of GFP-like proteins has more than one origin, demonstrating the convergent evolution of a complex molecular system. The more complex red color evolved from green through small incremental transitions (a stepwise accumulation of improvements), each identified in our experiments by ancestral gene reconstruction (Fig 1D). This mode of evolution has been anticipated since Darwin, but has only recently been demonstrated in computer simulation experiments (*5*).

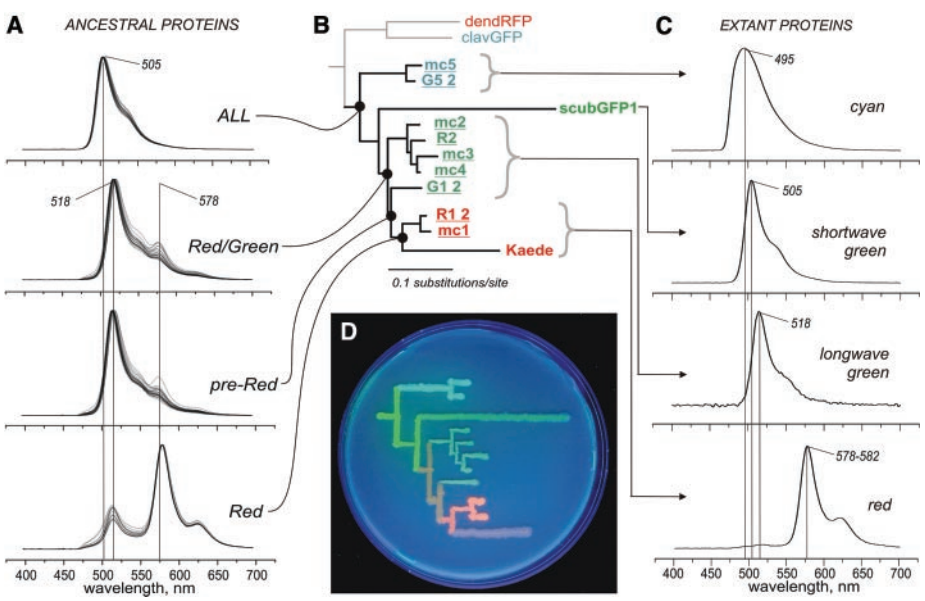

Fig. 1. (**A**) Fluorescence spectra of the reconstructed ancestral proteins. Multiple curves correspond to clones bearing variations at degenerate sites. (**B**) Phylogeny of GFP-like proteins from the great star coral *M. cavernosa* (sequence names are underlined) and closely related coral species. The red and cyan proteins from soft corals (dendRFP and clavGFP) represent an outgroup. (**C**) Fluorescence spectra of extant proteins. (**D**) Phylogenetic tree of colors from the great star coral, drawn on a petri dish with bacteria expressing extant and ancestral proteins, under ultraviolet light.

References and Notes

1. A. Miyawaki, T. Nagai, H. Mizuno, *Curr. Opin. Chem. Biol.* **7**, 557 (2003).
2. D. A. Shagin *et al.*, *Mol. Biol. Evol.* **21**, 841 (2004).
3. I. Kelmanson, M. Matz, *Mol. Biol. Evol.* **20**, 1125 (2003).
4. Z. Yang, *Comput. Appl. Biosci.* **13**, 555 (1997).
5. R. E. Lenski, C. Ofria, R. T. Pennock, C. Adami, *Nature* **423**, 139 (2003).
6. We thank N. V. Grishin for providing access to computer resources and D. Zacharias and K. Lukyanov for critical reading of the manuscript. Supported by grants from the Grass Foundation (J.A.U.), the U.S. Department of Defense and NIH (M.V.M.), and NSF and the Natural Sciences and Engineering Research Council of Canada (B.S.W.C.).

Supporting Online Material
www.sciencemag.org/cgi/content/full/305/5689/1433/DC1
Materials and Methods
Fig. S1
References and Notes

26 April 2004; accepted 7 July 2004

[1]Whitney Laboratory, [2]Department of Molecular Genetics and Microbiology, University of Florida, Gainesville, FL, USA. [3]Laboratory of Bioinformatics and Gene Expression, University of Chile, Santiago, Chile. [4]Department of Zoology, University of Toronto, Toronto, Canada.

*To whom correspondence should be addressed. E-mail: matz@whitney.ufl.edu