

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
4 November 2004 (04.11.2004)

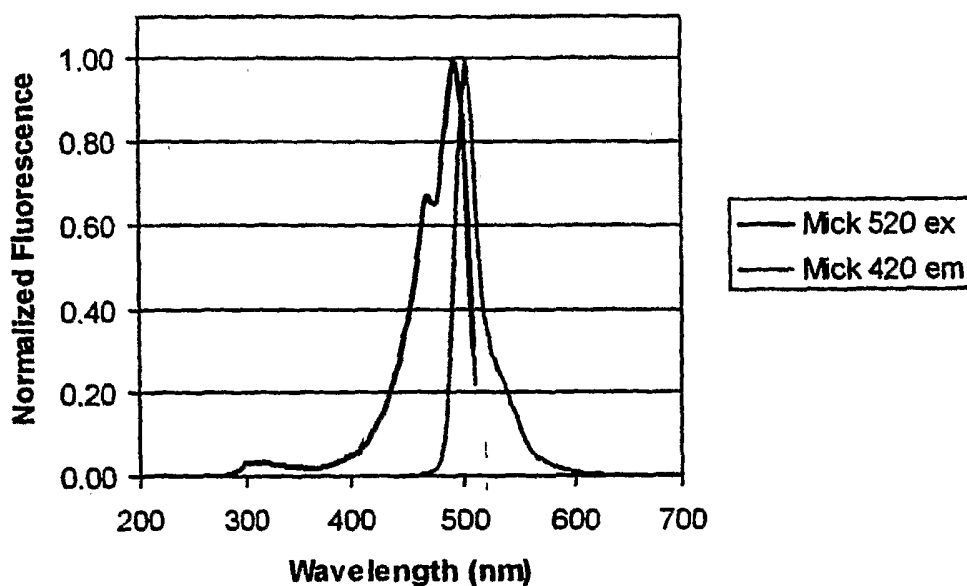
PCT

(10) International Publication Number
WO 2004/094597 A2

- (51) International Patent Classification⁷: **C12N** 33144 (US). SCHMALE, Michael, C. [US/US]; Um-Rs-mas-Mbf, 4600 Rickenbacker Cswy, Miami, FL 33144 (US).
- (21) International Application Number: PCT/US2004/011902
- (22) International Filing Date: 16 April 2004 (16.04.2004)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/463,038 16 April 2003 (16.04.2003) US
- (71) Applicant (for all designated States except US): **UNIVERSITY OF MIAMI** [US/US]; Office of Technology Transfer, 1475 12th Avenue, Sewell Bldg., Suite 2012, Miami, FL 33136 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **CARTER, Robert, W.** [US/US]; Um-Rsmas-Mbf, 4600 Rickenbacker Cswy, Miami, FL 33144 (US). **GIBBS, Patrick, D.** [US/US]; Um-Rsmas-Mbf, 4600 Rickenbacker Cswy, Miami, FL
- (74) Agent: **HOBBS, Ann, S.**; Venable LLP, P.O. Box 34385, Washington, DC 20043-9998 (US).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: ANTHOZOAN FLUORESCENT PROTEIN GENES



(57) Abstract: The present invention relates to fluorescent proteins, genes encoding the proteins, and methods for their use. In particular, the invention relates to a fluorescent protein identified as "Mick", having an excitation maximum of about 492 nm and an emission maximum of about 502 nm.



Declaration under Rule 4.17:

- *of inventorship (Rule 4.17(iv)) for US only*

Published:

- *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

ANTHOZOAN FLUORESCENT PROTEIN GENES

This application claims priority to U.S. provisional application no. 60/463,038, filed April 16, 2003, which is incorporated herein by reference in its entirety.

5 The work resulting in this invention was funded in part by NIEHS-MFBSC grant no. ES05705 and NIH-NINDS grant no. NS36998. The U.S. government has certain rights in the invention.

BACKGROUND OF THE INVENTION

10 1. Field of the Invention

The present invention relates to fluorescent proteins, genes encoding the proteins, and methods for their use.

2. Background Information

15 Genes encoding fluorescent proteins (FPs) have been isolated from a wide variety of cnidarian species and have been useful in a variety of contexts, including the tracking of proteins in heterologous cells, production of transgenic animals, and other purposes. The products of these genes fluoresce under certain wavelengths of light, allowing efficient measurement under a wide variety of *in vivo* and *in vitro* conditions. For example, such genes have been operably linked to genes encoding
20 proteins of interest and used to study gene expression and protein localization, as fluorescent tags to aid in visualizing protein trafficking, and as components of indicator systems that allow fluorescent sensing of small molecules and pH. Transgenic animals containing such genes are useful, for example, in environmental applications, and are becoming popular as pets.

25 In spite of the many fluorescent proteins and genes isolated thus far, a need remains for new fluorescent proteins with several favorable properties, including better stability of fluorescence, lower toxicity, and better resistance to photobleaching.

SUMMARY OF THE INVENTION

30 The invention provides an isolated green fluorescent protein identified as "Mick", having an excitation maximum of about 492 nm and an emission maximum of about 502 nm. The invention also provides a nucleic acid encoding Mick. Mick can be isolated from the scleractinian coral *Montastraea cavernosa*, and is highly resistant to photobleaching.

Mick has exceptional long-term stability and resistance to photobleaching compared to EGFP (the most widely used FP on the market, available through ClonTech) and other available FPs. Bacterial cells expressing Mick will remain fluorescent for long periods of time while the fluorescence of EGFP-expressing cells will diminish rapidly after c.a. 48 hours.

The excitation and emission spectra of Mick were obtained as described below and are shown in Figure 3. Amino acid and nucleic acid sequences of Mick are shown in Figure 4.

The invention further provides an isolated cyan fluorescent protein identified as "Pd1", having an excitation maximum of about 482 nm and an emission maximum of about 494 nm. The invention also provides a nucleic acid encoding Pd1. Pd1 can be isolated from the scleractinian coral *Pocillopora damicornis*. The excitation and emission spectra of Pd1 were obtained as described below and are shown in Figure 5. Amino acid and nucleic acid sequences are shown in Figure 6.

The invention further provides an isolated blue-cyan fluorescent protein identified as "Dst1", having an excitation maximum of about 438 nm and an emission maximum of about 482 nm. The invention also provides a nucleic acid encoding Dst1. Dst1 can be isolated from the non-scleractinian anthozoan *Discosoma striata*. The excitation and emission spectra of Dst1 were obtained as described below and are shown in Figure 7. Amino acid and nucleic acid sequences are shown in Figure 8.

In addition, included in the invention are fragments and variants of Mick, Pd1 and Dst1 that exhibit similar fluorescent properties, and nucleic acids encoding such fragments and variants. Suitable fragments and variants can be obtained through routine experimentation by methods known to persons of skill in the art. Variants include 1-20 (preferably no more than 10) amino acid deletions, substitutions or insertions, preferably conservative substitutions, as will be known to those of skill in the art. It is expected that fragments of close to full length (i.e. containing at least about 90%, preferably 95%, and most preferably 98 or 99% of the original amino acid sequence) will retain the properties of the protein. Such functional fragments can be discovered by routine experimentation using methods described herein and those known in the art. For example, Dopf J, Horiagon TM (1996) Deletion mapping of the *Aequorea victoria* green fluorescent protein. Gene 173:39-44, describes suitable techniques.

The invention also provides an isolated nucleic acid sequence that encodes a fluorescent indicator or chimeric construct, the indicator having a sensor polypeptide or protein that is responsive to a chemical, biological, electrical or physiological parameter and a fluorescent protein moiety, wherein the sensor polypeptide is
5 operatively linked to the fluorescent protein moiety, and wherein the fluorescent protein moiety is affected by the responsiveness of the sensor polypeptide.

Techniques for obtaining such indicators and constructs are known to those of skill in the art, as described, for example in Halloran MC, *et al.* (2000) Laser-induced gene expression in specific cells of transgenic zebrafish. *Development* 127(9):1953-1960.

10 The invention also includes transgenic nonhuman animals having a nucleic acid sequence that encodes a fluorescent protein of the invention. Techniques for making such transgenic animals are also known in the art. Methods for making transgenic zebrafish can be found, for example, in Gibbs PDL and Schmale MC (2000) GFP as a genetic marker scorable throughout the life cycle of transgenic
15 zebrafish. *Marine Biotechnology* 2(2): 107-125 and Gibbs PDL, Peek A, and Thorgaard G (1994). An *in vivo* screen for the luciferase transgene in zebrafish. *Molecular Marine Biology and Biotechnology* 3:307-316.

The fluorescent peptides of the invention can be used as markers to detect expression of a gene of interest, for example by inserting the gene of interest along
20 with DNA encoding the fluorescent protein into a vector. Such vectors can be transfected into eukaryotic or prokaryotic host cells, including, for example, bacterial cells, insect cells, yeast cells, and mammalian cells. Techniques useful for carrying out this and other aspects of the invention can be found in Sambrook *et al.* (*Molecular Cloning. A Laboratory Manual*, 3rd ed. (2001), Cold Spring Harbor Laboratory
25 Press).

Fluorescent proteins of the invention can also be used in biochemical assays and as reagents, for example to monitor fermentation processes or quantify gene expression and to label cells for tracking purposes.

The fluorescent proteins of the invention can also be used in a range of
30 applications including use as easily scorable transgenic markers and as taxonomic markers for genetic studies of corals and other cnidarians. They can also be incorporated into diagnostic kits as a color change indicators or used in FRET systems. Other uses include food colorants, cosmetics, paints and dyes.

Other aspects of the invention include luminescent ornamental fish (e.g. for hobbyists) and a wide variety of other living organisms, including, for example, bacteria, fungi, viruses, animals and plants, into which the nucleic acids and proteins can be incorporated for recreational, medical, research and other purposes.

5 The invention provides an isolated protein comprising the amino acid sequence identified herein as *Mick* (SEQ ID NO:2), or a fragment or variant thereof having the excitation and emission properties shown in Figure 1, and an isolated nucleic acid encoding the protein.

10 The invention also provides an isolated nucleic acid comprising the nucleic acid sequence identified herein as *Mick* (SEQ ID NO:1).

The invention provides an isolated protein comprising the amino acid sequence identified herein as *Pd1* (SEQ ID NO:4) or a fragment or variant thereof having the excitation and emission properties shown in Figure 4, and an isolated nucleic acid encoding *Pd1*.

15 The invention further provides an isolated nucleic acid comprising the nucleic acid sequence identified herein as *Pd1* (SEQ ID NO:3).

20 The invention also provides an isolated protein comprising the amino acid sequence identified herein as *Dst1* (SEQ ID NO:6), or a fragment or variant thereof having the excitation and emission properties shown in Figure 5, and an isolated nucleic acid encoding *Dst1*.

The invention further provides an isolated nucleic acid comprising the nucleic acid sequence identified herein as *Dst1* (SEQ ID NO:5).

25 It will be appreciated by the skilled practitioner that the invention also includes nucleic acids and proteins that consist essentially of, or consist of, SEQ ID NOS:1, 2, 3, 4, 5, and 6.

30 The invention also provides a vector comprising one of any of the nucleic acids of the invention, a host cell comprising one of the nucleic acids of the invention linked to an operable promoter, a transgenic animal comprising one of the nucleic acids of the invention. Preferably the transgenic animal is a fish, more preferably, a zebrafish.

The invention further provides a nucleic acid construct comprising

- i) a coding sequence encoding a polypeptide or protein of interest, and

- ii) a nucleic acid of one of claims 3, 4, 7, 8, 11 or 12 encoding a fluorescent protein,

wherein the coding sequence and the nucleic acid are linked in such a manner that expression of the fused sequence yields a fluorescent hybrid protein in which the

5 protein of interest is fused to the fluorescent protein. Also included in the invention is a cell comprising this nucleic acid construct. For example, Mick, Pd1 or Dst1 can be linked to myc or other genes of interest and analyzed as described by Langenau DM, *et al.* (2003) Myc-induced T cell leukemia in transgenic zebrafish. Science 299(5608):887-890.

10 The invention also provides a method of detecting expression of a nucleic acid comprising introducing a nucleic acid construct, as described above, into a cell or organism, and detecting expression of the nucleic acid by emission of fluorescent light.

The term "isolated" is intended to refer to any compound or agent that is in a
15 form not found in its original environment or in nature, *e.g.*, more concentrated, more purified, separated from at least one other component with which it is naturally associated.

The terms "peptide", "polypeptide", and "protein" as used herein, are intended to refer to amino acid sequences of various sizes, and are not considered to imply any
20 size limitation on the compositions of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1a: Photobleaching of *Montastraea* green FPs. Two FPs are shown, Mick and
25 a second FP, MfaG1, cloned independently by Gibbs and Carter. Note the lack of photobleaching in Mick compared to the strong reduction in MfaG1 fluorescence under the same conditions. MfaG1 is more representative of the currently known *Montastraea* green FPs. Measurements were made in a SPEX spectrofluorometer by exposing diluted aqueous stock solutions to low intensity excitation light. Bleaching
30 was induced by continual exposure of each sample to the maximum available excitation light. The rebound in MfaG1 fluorescence occurred when the continual excitation light was shut off.

Figure 1b: Photobleaching of Mick and EGFP under identical conditions. Dilute samples of Mick and EGFP were mixed with 0.5% low melt agarose and poured into polycarbonate cuvettes. Each cuvette was placed into the excitation beam of a SPEX spectrofluorometer and fluorescence was measured each second for 120 seconds.

- 5 Shown is a representative bleaching curve of three samples of each FP under low light exposure. Time-based photobleaching in EGFP is more than twice as fast as Mick. This relationship held true for the entire range of light intensities that could be produced by the spectrofluorometer.

10 Figure 2:

Bacterial streaks under blue light excitation:

- a) EGFP (left) and Mick (right) after 24 hours at 37 °C
- b) Same two streaks (slightly different location) after an additional four days at room temperature

- 15 Bacterial plates after one day at 37 °C and four additional days at room temperature (photos taken with identical exposure settings):

- c) EGFP colonies
- d) Mick colonies

Transgenic and wild type fish:

- 20 e) Three days old, bright field
- f) Three days old, fluorescence
- g) Five days old, bright field
- h) Five days old, fluorescence
- i) Three months old (photo taken under blue light with no additional color
- 25 modification or filtering)

Figure 3: Excitation and Emission Spectra of Mick. The excitation spectrum was obtained using a SPEX spectrofluorometer by setting the emission wavelength to 520 nm and scanning all excitation wavelengths at 2 nm increments between 200 and 510 nm. The emission spectrum was obtained by setting the excitation light to 420 nm and scanning all emission wavelengths at 2 nm intervals between 440 and 700 nm.

- 35 Figure 4: Nucleotide (SEQ ID NO:1) and Amino Acid (SEQ ID NO:2) Alignment of Mick

Figure 5: Excitation and Emission Spectra of Pd1

Figure 6: Nucleotide (SEQ ID NO:3) and Amino Acid (SEQ ID NO:4) Alignment of

5 Pd1

Figure 7: Excitation and Emission Spectra of Dst1

Figure 8: Nucleotide (SEQ ID NO:5) and Amino Acid (SEQ ID NO:6) Alignment of

10 Dst1

DETAILED DESCRIPTION OF THE INVENTION

We have cloned several new fluorescent proteins from three anthozoan species. Each has its own set of useful and unique attributes. They are as follows:

15

1. Mick: a new green FP from the Caribbean coral, *Montastraea cavernosa*. This FP expresses at a high level in both *E. coli* and zebrafish. It is also highly resistant to photobleaching. *In vivo* fluorescence does not decrease even after prolonged exposure to intense UV light. In aqueous solution, other FPs (including other green FPs from

20

Montastraea spp.) photobleach rapidly. Compared to EGFP, the most popular FP available on the market today, Mick produces more stable fluorescence in *E. coli* and can be visualized in zebrafish under room lighting.

2. Pd1: a new cyan fluorescent protein from the Pacific coral, *Pocillopora damicornis*. This cyan FP expresses well in *E. coli* and the Pd1 sequence is dissimilar from most other anthozoans.

25

3. Dst1: a new blue-cyan fluorescent protein from the Pacific soft coral, *Discosoma striata*.

30

Table 1: Comparison of Characteristics of Mick, Pd1 and Dst1

	<u>Mick</u>	<u>Pd1</u>	<u>Dst1</u>
35 Excitation Maximum:	492 nm	482 nm	438 nm
Emission Maximum:	502 nm	494 nm	482 nm
Nucleotides:	681	666	681
Amino Acids:	227	222	227

5 Materials and Methods

Species collection and animal husbandry

Single, bright green colonies of *Montastraea cavernosa* and *Montastraea faveolata* colonies were collected from reefs in the South Florida area. A single branch from a large colony of *Pocillopora damicornis* was collected from the Pacific
10 coast of Panama. Colony of *Discosoma striata* was purchased at a local aquarium store. All corals were maintained in flow through aquariums using filtered sea water at the Rosenstiel School of Marine and Atmospheric Sciences (University of Miami) until tissue processing could occur.

15 *Isolation of total RNA*

RNA was isolated using the Totally RNA kit (Ambion, Inc.) from a cellular mass lightly airbrushed from the underlying skeleton of each colony. Briefly, the tissue was homogenized and mixed by tube inversion with approximately 10 volumes of denaturation solution and extracted with 1 volume of phenol/chloroform/isoamyl
20 alcohol. 1/10 volume of 3M sodium acetate was added to the supernatant which was extracted an additional time with 1 volume of acid-phenol/chloroform. One volume of isopropanol was added to the supernatant and total RNA was precipitated by centrifugation, the pellet washed with 75% ethanol and resuspended in 50 uL RNase-free water. Yield was measured by spectrophotometry and quality was assessed by gel
25 electrophoresis.

Reverse transcription, amplification and recovery of cDNAs

The RLM-RACE kit (Ambion, Inc.) was used to create cDNAs using the "small reaction" protocol with 1 ug of total RNA. Calf intestinal phosphatase was
30 used to remove 5' phosphates from degraded and non-capped RNA. Tobacco acid pyrophosphatase was then used to remove the cap from mRNAs.

A 5' RNA adaptor:

5'-

GCUGAUGGCGAUGAAUGAACACUGCGUUUGCUGGCUUUGAUGAAA-3'
(SEQ ID NO:7)

5 was ligated to total decapped mRNAs with T4 RNA ligase.

Reverse transcription was accomplished using MMLV reverse transcriptase and a polyT primer:

10 5'-CTCGAGAAGCTTGAATTCGGATCCTTTTTTTTTTTTTTTTTT-3' (SEQ
ID NO:8)

Initial cDNA amplification was performed using the Outer RNA Adapter Primer
supplied in the kit:

15

5'-GCTGATGGCGATGAATGAACACTG-3' (SEQ ID NO:9)

and the polyT primer described above. The PCR conditions were: 94°C/5 min, 80°C
hold, add polymerase, 34 cycles of 94°C/30 sec, 60°C/1 min, 72°C/1 min and a final
20 10 min of 72°C. Amplified cDNAs were gel purified. The area between 500 - 1200 bp
was removed from the gel with a clean razor blade and recovered by slow speed
centrifugation through siliconized glass wool followed by isopropanol precipitation.

Cloning of PdaC1 and DstC1

25 One microliter of the recovered PCR amplified cDNA was ligated into the
PCR 2.1 TOPO vector (Invitrogen). The resulting plasmids were used to
electrotransform Top-10 *E. coli* cells (Invitrogen), plated out on 100 µg/ml LB-Amp
medium and incubated overnight.

30 *Cloning of Mick*

Secondary amplification of the *Montastraea cavernosa* cDNA was done using
SuperTaq polymerase (Ambion, Inc.) and two gene-specific primers we had
previously designed for another fluorescent protein gene, McaG1. The primers were
initially designed for cloning McaG1 into the plasmid vector pBS II KS+ [It had
35 originally been cloned into the pCR 2.1 TA vector (Invitrogen)]. The upper primer
included a Kpn I restriction site, both Shine-Dalgarno and Marilyn Kozak

upregulation sequences (SD-MK) and 16 bases of the coding region, starting with the initial start codon, ATG. The lower primer included 13 bases of the 3' untranslated region (UTR) from McaG1, a stop codon, and a Xba I restriction site.

5 Upper Primer (McaG1 U1-37)

5'-CCGGTACCTAA GGAGGCCACC ATGAGTGTGATAAAAC-3' (SEQ ID NO:10)

Kpn-SD-MK Coding Region

10 Lower Primer (McaG1 L1-24)

5'-GGTCTAGA TTA CTTGGCCTGCCTC-3' (SEQ ID NO:11)

Xba Stop 3'UTR

Because we did not expect 100% similarity between the McaG1 primers and the new
15 gene sequences, we used the PCR reaction conditions for what we call "dirty PCR". This includes a low annealing temperature and long extension times for 10 cycles plus an additional 30 cycles under more normal conditions. PCR conditions were: 94°C/5 min, 80°C hold, add polymerase, 10 cycles of 94°C/1 min, 42°C/2 min, 72°C/3 min, followed by 30 cycles of 94°C/30 sec, 55°C/1 min, 72°C/1 min and a final 10 min of
20 72°C.

Cloning, Plasmid preparation and DNA Sequencing

The PCR products were gel purified, digested with Kpn I/Xba I, and ligated into a Kpn I/Xba I digested pBS II KS+ plasmid. Resultant plasmids were
25 electrotransformed into Top 10 *E. coli* (Invitrogen). Transformed bacteria were grown on LB-ampicillin (100 ug/ml) plates and colonies were screened for fluorescence using a Leica MZFLIII fluorescence stereo dissection microscope. Fluorescent colonies were screened for brightness after 24 hours at 37°C. Single colonies were picked, restreaked for several rounds to resolve mosaicism and grown in liquid culture
30 for plasmid preparation by the Qiagen Midi kit. Sequencing was accomplished using the "Big Dye Terminator" reaction on an Applied Biosystems Model 377 DNA Sequencer at the DNA Core Facility.

Characteristics of Mick

Figure 1a shows a comparison of photobleaching between Mick and MfaG1 (MfaG1 is a green FP with typical photobleaching characteristics and was cloned independently from the scleractinian coral *Montastraea faveolata* by Gibbs and Carter. Under bright excitation light, MfaG1 rapidly loses fluorescence (decending
5 part of curve). When the excitation light is turned off, fluorescence rebounds (ascending part of curve) and remains stable as long as the excitation light remains off. This figure was generated using aqueous solutions of purified cloned FPs. Measurements were made using a SPEX spectrofluorometer.

Figure 1b shows a comparison of photobleaching resistance between Mick and
10 EGFP. Measurements were made as described above but at several different light intensities. Fluorescence decays more than twice as fast in EGFP than in Mick. This relationship was true at all light intensities tested.

Figure 2: a) 24-hour old Mick- and EGFP-expressing bacterial streaks under blue light excitation. b) the same two streaks after 96 additional hours at room
15 temperature. Note that the EGFP cells have lost a considerable amount of fluorescence while the relative intensity of the Mick-expressing cells has incereased. c-d) single colonies of EGFP (c) and Mick (d) after 24 hours at 37 degrees C followed by 96 hours at room temperature. e-f) Mick-expressing and wild-type three-day-old zebrafish embyos under white light (e) and blue light (f). g-h) Five-day-old zebrafish
20 embryos as in e and f. i) Mick-expressing and wild-type zebrafish under blue light. Photos a-h were taken with a digital camera mounted on a Leica MZFLIII fluorescence stereo dissecting microscope. Photo i was taken with a digital camera in a dark room with a blue light source for illumination and no additional camera filters or post-processing of color

25 Excitation and emission spectra of Mick are given in Figure 3. The excitation spectrum was obtained using a SPEX spectrofluorometer by setting the emission wavelength to 520 nm and scanning all excitation wavelengths at 2 nm increments between 200 and 500 nm. The emission spectrum was obtained by setting the excitation light to 480 nm and scanning all emission wavelengths at 2 nm intervals
30 between 490 and 700 nm.

Transgenic Animals

Transgenic zebrafish were made according to the methods described by Gibbs and Schmale (2000), and are shown in Figure 2 (e-i). Similar techniques can be used

to obtain other types of ornamental fish, for example, goldfish, koi, angelfish, tiger barbs, etc.

The embodiments illustrated and discussed in the present specification are intended only to teach those skilled in the art the best way known to the inventors to make and use the invention, and should not be considered as limiting the scope of the present invention. The exemplified embodiments of the invention may be modified or varied, and elements added or omitted, without departing from the invention, as appreciated by those skilled in the art in light of the above teachings. It is therefore to be understood that, within the scope of the claims and their equivalents, the invention may be practiced otherwise than as specifically described.

References cited herein are hereby incorporated by reference.

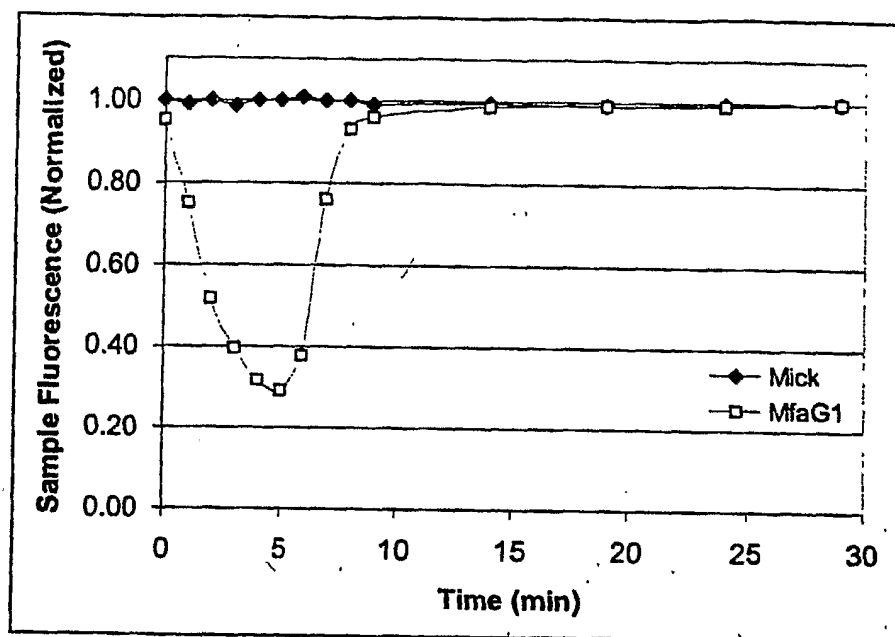
WE CLAIM:

1. An isolated protein comprising the amino acid sequence identified herein as Mick (SEQ ID NO:2), or a fragment or variant thereof having the excitation and emission properties shown in Figure 3.
2. The isolated protein of claim 1 comprising Mick (SEQ ID NO:2).
3. An isolated nucleic acid encoding the protein of claim 1.
4. An isolated nucleic acid comprising the nucleic acid sequence identified herein as *Mick* (SEQ ID NO:1).
5. An isolated protein comprising the amino acid sequence identified herein as Pd1 (SEQ ID NO:4) or a fragment or variant thereof having the excitation and emission properties shown in Figure 5.
6. The isolated protein of claim 5 comprising Pd1 (SEQ ID NO:4).
7. An isolated nucleic acid encoding the protein of claim 5.
8. An isolated nucleic acid comprising the nucleic acid sequence identified herein as *Pd1* (SEQ ID NO:3).
9. An isolated protein comprising the amino acid sequence identified herein as Dst1 (SEQ ID NO:6), or a fragment or variant thereof having the excitation and emission properties shown in Figure 6.
10. The isolated protein of claim 9 comprising Dst1 (SEQ ID NO:6).
11. An isolated nucleic acid encoding the protein of claim 9.
12. An isolated nucleic acid comprising the nucleic acid sequence identified herein as *Dst1* (SEQ ID NO:5).
13. A vector comprising the nucleic acid of one of claims 3, 4, 7, 8, 11 or 12.
14. A host cell comprising the nucleic acid of one of claims 3, 4, 7, 8, 11 or 12 linked to an operable promoter.
15. A transgenic organism comprising the nucleic acid of one of claims 3, 4, 7, 8, 11 or 12 .
16. The organism of claim 15 that is a transgenic animal.
17. The animal of claim 16 that is a fish.
18. The fish of claim 17 that is a zebrafish.
19. A nucleic acid construct comprising
 - iii) a coding sequence encoding a polypeptide or protein of interest,
and

- iv) a nucleic acid of one of claims 3, 4, 7, 8, 11 or 12 encoding a fluorescent protein,
wherein the coding sequence and the nucleic acid are linked in such a manner that expression of the fused sequence yields a fluorescent hybrid protein in which the protein of interest is fused to the fluorescent protein.
- 20. A method of detecting expression of a nucleic acid comprising introducing the nucleic acid construct of claim 19 into a cell or organism, and detecting expression of the nucleic acid by emission of fluorescent light.
- 21. A cell comprising the nucleic acid construct of claim 19.

1/9

Figure 1a



2/9

Photobleaching of Mick and EGFP Under Identical Conditions

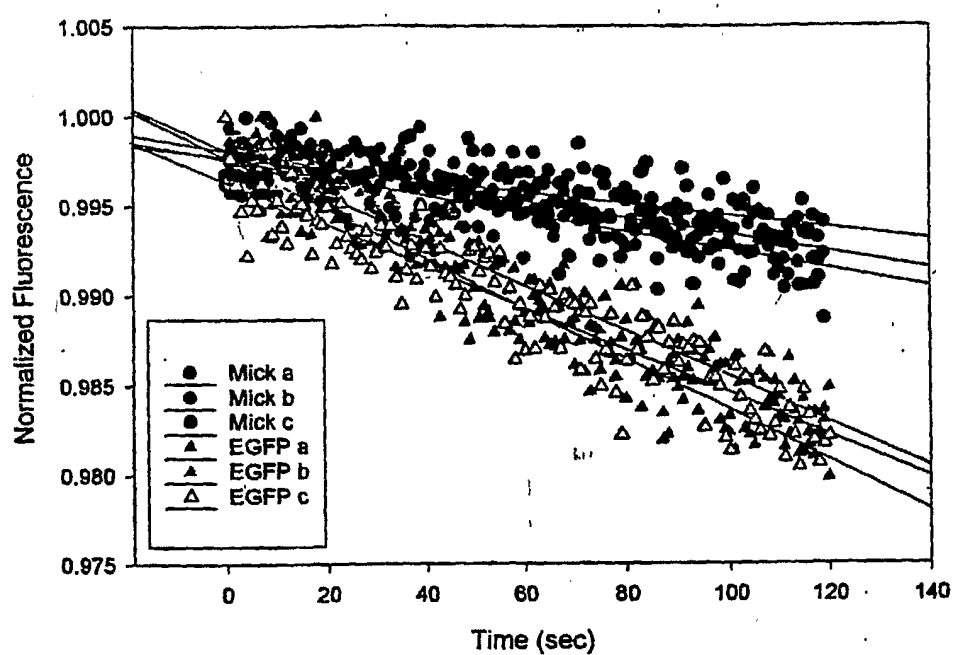
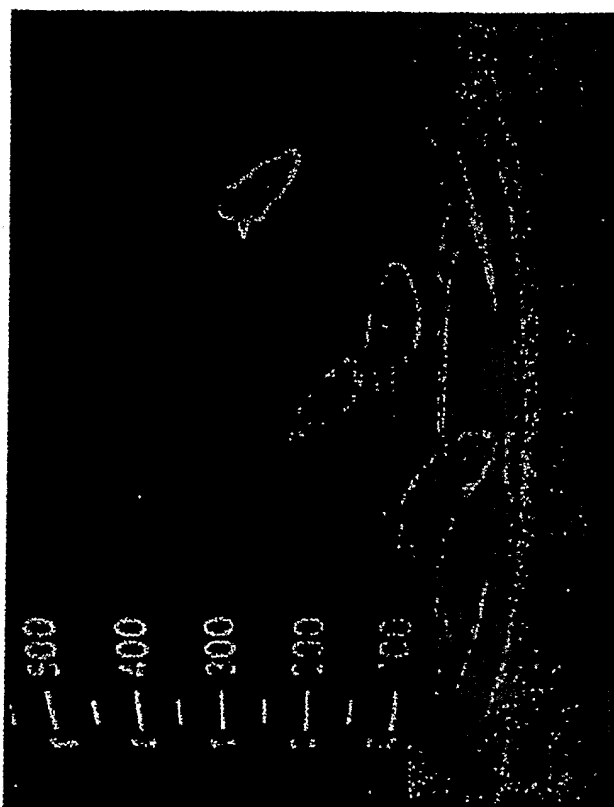
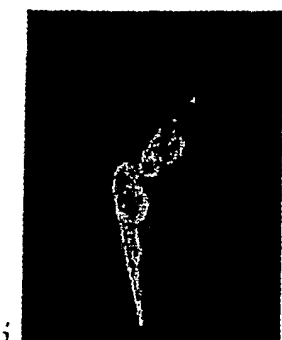
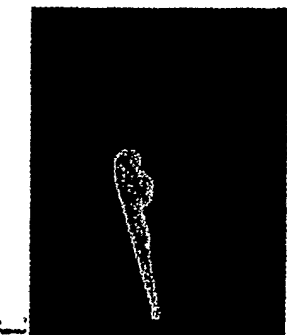
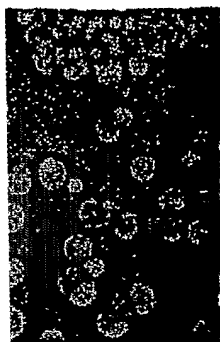


Figure 1b

Figures 2a-2i

3/9



4/9

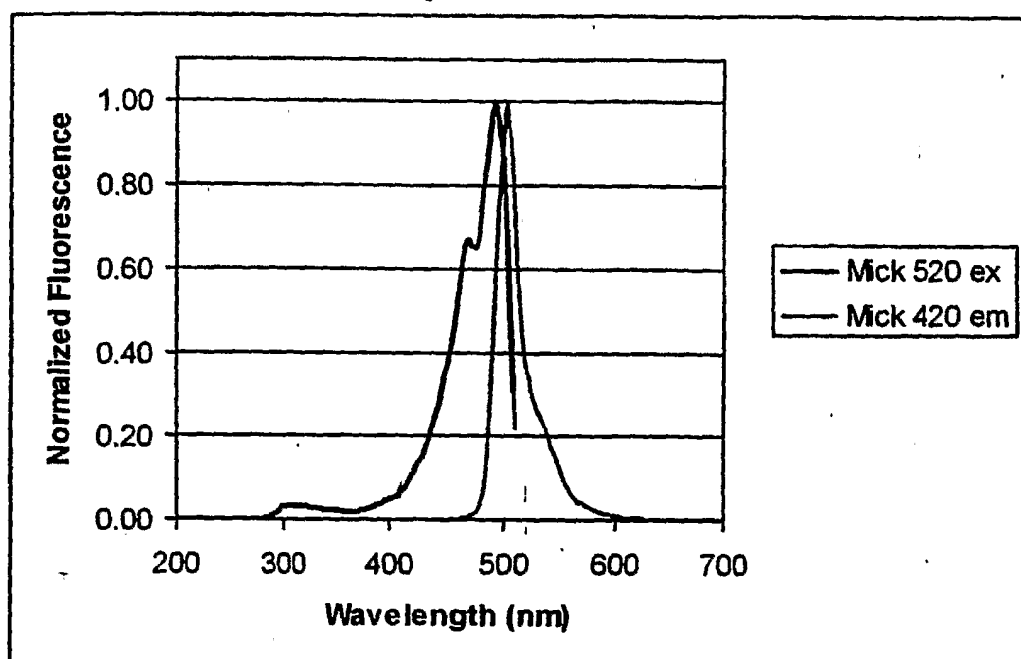


Figure 3

5/9

Figure 4: Nucleotide and Amino Acid Alignment of Mick

ATGAGTGTGATAAAACCAGACATGAAGATCAAGCTGCGTATGGAAGGCGC - 50
 M S V I K P D M K I K L R M E G A

GTAAACGGGCACAACCTTCGTGATTGAAGGAGAAGGAAAAGGCAAGCCTT - 100
 V N G H N F V I E G E G K G K P

CGAGGGAACACAGACTATAAACCTTACAGTCAAAGAAGGCGGACCTCTG - 150
 P E G T Q T I N L T V K E G G P L

CTTTTGCTTACGATATCTTGACAGCAGCATTCCAGTACGGCAACAGGGC - 200
 P F A Y D I L T A A F Q Y G N R A

TTCACCAAATACCCAAGAGACATAGCAGACTATTTCAAGCAGTCTTTTC - 250
 F T K Y P R D I A D Y F K Q S F

TGAGGGGTATTCTCTGGGAACGAAGCATGACTTATGAAGACCAGGGCATT - 300
 P E G Y S W E R S M T Y E D Q G I

GCATCATCAAGAGCGACATAAGAATGGAAGGCGACTGCTTTATCTATGA - 350
 C I I K S D I R M E G D C F I Y E

ATTTCGATATGATGGTGTGAACTTTCCCCAAGTGGTCCAGTTATGCAAA - 400
 I R Y D G V N F P P S G P V M Q

AGAAGACGCTGAAATGGGAGCCATCCACTGAGAAAATGTATGTGCGTGAT - 450
 C K T L K W E P S T E K M Y V R D

GAGTGCTGAAGGGTGAAGTTAACATGGCTCTGTTGCTTGAAGGAGGTGG - 500
 G V L K G E V N M A L L L E G G G

CATTACCGATGTGACTTTTGAAGTACTTACAAAGCGAAGAAACGTGTTC - 550
 H Y R C D F R S T Y K A K K R V

GTTGCCAGACTATCACTTTGTGGACCACCGCATTGAGATTTTGAGCCAT - 600
 P L P D Y H F V D H R I E I L S H

ACAATGACTACAACACCGTAAAGCTGTCCGAGAATGCCCGAGGCTCGCT - 650
 D N D Y N T V K L S E N A R G S L

TTCTATGCTGCCGAGGCAGGCCAAGTAATC - 681
 F Y A A E A G Q V I

6/9

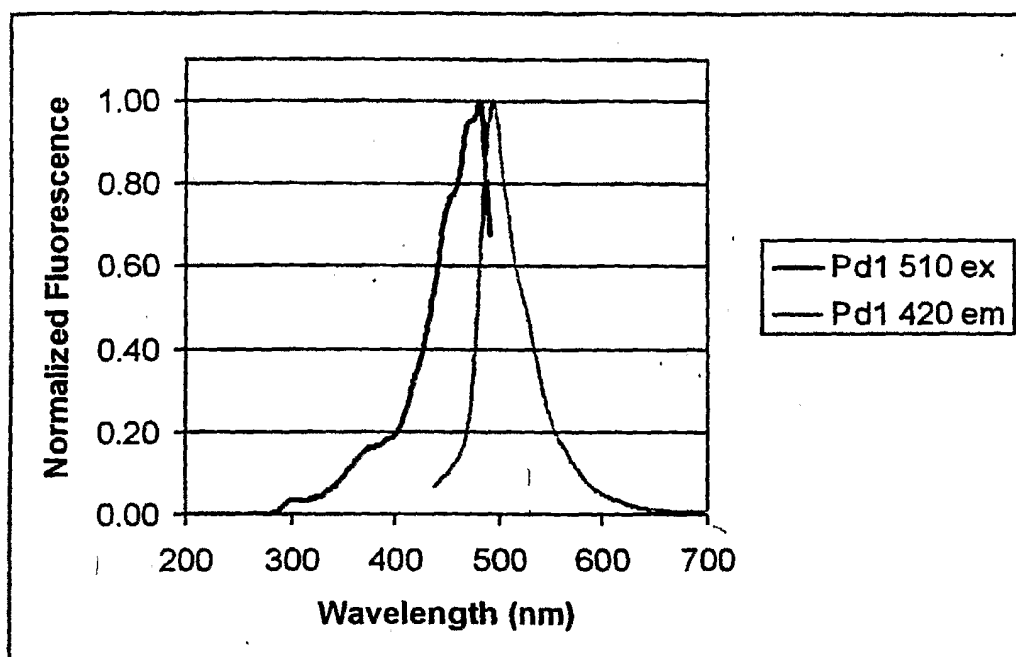


Figure 5

7/9

Figure 6: Nucleotide and Amino Acid Alignment of Pd1

ATGTCTCTTTCAAAGCAGGTCATTCTAAAAGACATGAATTTGAAATTTGA - 50
 M S L S K Q V I L K D M N L K F E

AATGAAAGGTAGTGTCAATGGACATTACTTTGAGATCGAAGGTGAAGGAA - 100
 M K G S V N G H Y F E I E G E G

GAGGAAAGCCTTACGAAGGAGTACAGAAGTCCACTTTTTGGGTAACCAAG - 150
 R G K P Y E G V Q K S T F W V T K

GGAGGACCCCTTCCATTTTCCTTTGACATACTTTTCGTCAGCCTTCAAATA - 200
 G G P L P F S F D I L S Z A F K Y

TGGAAACAGATGCTTTACTAAGTACCCTGCCGACATGCCTGACTATTTCA - 250
 G N R C F T K Y P A D M P D Y F

AAGAAGCATTTTCCTGCTGGAATGTCATTTGAAAGGACATTTACATTTGAG - 300
 K E A F P A G M S F E R T F T F E

SATGGAGGAGTTGCTACAGCCAGTGGACACATTTGTCTTGAGGGGAATTG - 350
 D G G V A T A S G H I C L E G N W

GTTTAAGCACACATCCATGTTTCATGGCGTTAACTTTCCCGCTAACGGAC - 400
 F K H T S M F H G V N F P A N G

CCATAATGCAAAAGAGGACAATTGGCTGGGACCCTTCCTTTGAGAAAATG - 450
 P I M Q K R T I G W D P S F E K M

ACCGTCTCCAACAACATATTGAGAGGTGATGTAACATATGTTCCCTTCAACT - 500
 T V S N N I L R G D V T M F L Q L

CAAAGGGGGAGGTTATCACAGTTGCCAGTTTCACACTTCTTACAAAACAA - 550
 K G G G Y H S C Q F H T S Y K T

AGGAGCCGGTCACGCTGCCTCAGAACCACGTCGTAGAACATCGCATTACG - 600
 K E P V T L P Q N H V V E H R I T

AGGACCGACATTGAAGATAAAAAGGTCCTGCTGGAAGAACTGCTGTGGC - 650
 R T D I E D K K V L L E E T A V A

CATGTTAACCCTTTG - 666
 H V N P L

8/9

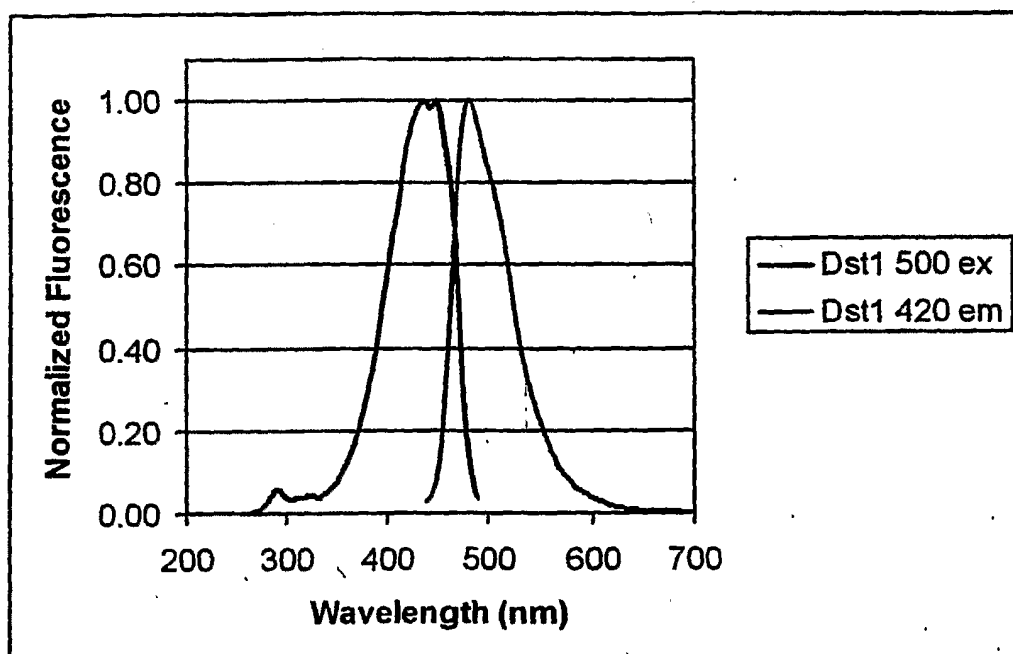


Figure 7

9/9

Figure 8: Nucleotide and Amino Acid Alignment of Dst1

ATGAGTTGGTCCAAGAGTGTGATCAAGGAAGAAATGTTGATCAATCTTCA - 50
 M S W S K S V I K E E M L I N L H

TCTGGAAGGAACGTTCAATGGGCACTACTTTGAAATAAAAGGCAAAGGAA - 100
 L E G T F N G H Y F E I K G K G

AAGGGAACCCTAATGAAGGCACCAATACCGTCACGCTCGAGGTTACCAAG - 150
 K G N P N E G T N T V T L E V T K

GGTGGACCTCTGCCATTTGGTTGGCATATTTTGTGCCCAATTTTCAGTA - 200
 G G P L P F G W H I L C P Q F Q Y

TGGAAACAAGGCATTTGTCCACCACCCTGACGACATACCCGATTATCTAA - 250
 G N K A F V H H P D D I P D Y L

AGCTGTCAATTTCCGGAGGGATATACATGGGAACGGTCCATGCACTTTGAA - 300
 K L S F P E G Y T W E R S M H F E

GACGGTGGCTTGTGTTGTATACCAATGATATCAGTTTGACAGGCAACTG - 350
 D G G L C C I T N D I S L T G N C

TTTCAACTACGACATCAAGTTCACTGGCTTGAACCTTCCTCCAAATGGAC - 400
 F N Y D I K F T G L N F P P N G

CCGTTGTGCAGAAGAAGACAACCTGGCTGGGAACCGAGCACTGAGCGTTTG - 450
 P V V Q K K T T G W E P S T E R L

TATCCTCGTGATGGCGTGTGATAGGAGACATCCATCATGCTCTCACAGT - 500
 Y P R D G V L I G D I H H A L T V

GGCAGGAGGTGGTCATTACGTATGTGACATTAAAACTGTTTACAGGGCCA - 550
 A G G G H Y V C D I K T V Y R A

AAAASCCCGTAAAGATGCCAGGGTATCACTATGTTGACCCCAAACCTGGTT - 600
 K Z P V K M P G Y H Y V D S K L D

ATAAGGAGCAACGACAAAGAATTCATGAAAGTTGAGGAGCATGAAATCGC - 650
 I R S N D K E F M K V E E H E I A

CGTTGCACGCCNCCATCCGCTCCAAAGCCAA - 681
 V A R Z H P L Q S Q

1/8

SEQUENCE LISTING

<110> UNIVERSITY OF MIAMI

<120> ANTHOZOAN FLUORESCENT PROTEIN GENES

<130> 39532-202060

<140> 60/463,038

<141> 2003-04-16

<160> 11

<170> PatentIn Ver. 3.2

<210> 1

<211> 681

<212> DNA

<213> Montastraea cavernosa

<220>

<221> CDS

<222> (1)..(681)

<400> 1

atg agt gtg ata aaa cca gac atg aag atc aag ctg cgt atg gaa ggc	48
Met Ser Val Ile Lys Pro Asp Met Lys Ile Lys Leu Arg Met Glu Gly	
1 5 10 15	
gct gta aac ggg cac aac ttc gtg att gaa gga gaa gga aaa ggc aag	96
Ala Val Asn Gly His Asn Phe Val Ile Glu Gly Glu Gly Lys Gly Lys	
20 25 30	
cct ttc gag gga aca cag act ata aac ctt aca gtc aaa gaa ggc gga	144
Pro Phe Glu Gly Thr Gln Thr Ile Asn Leu Thr Val Lys Glu Gly Gly	
35 40 45	
cct ctg cct ttt gct tac gat atc ttg aca gca gca ttc cag tac ggc	192
Pro Leu Pro Phe Ala Tyr Asp Ile Leu Thr Ala Ala Phe Gln Tyr Gly	
50 55 60	
aac agg gca ttc acc aaa tac cca aga gac ata gca gac tat ttc aag	240
Asn Arg Ala Phe Thr Lys Tyr Pro Arg Asp Ile Ala Asp Tyr Phe Lys	
65 70 75 80	
cag tct ttt cct gag ggg tat tcc tgg gaa cga agc atg act tat gaa	288
Gln Ser Phe Pro Glu Gly Tyr Ser Trp Glu Arg Ser Met Thr Tyr Glu	
85 90 95	
gac cag ggc att tgc atc atc aag agc gac ata aga atg gaa ggc gac	336
Asp Gln Gly Ile Cys Ile Ile Lys Ser Asp Ile Arg Met Glu Gly Asp	
100 105 110	
tgc ttt atc tat gaa att cga tat gat ggt gtg aac ttt ccc cca agt	384
Cys Phe Ile Tyr Glu Ile Arg Tyr Asp Gly Val Asn Phe Pro Pro Ser	
115 120 125	

2/8

ggt cca gtt atg caa aag aag acg ctg aaa tgg gag cca tcc act gag 432
 Gly Pro Val Met Gln Lys Lys Thr Leu Lys Trp Glu Pro Ser Thr Glu
 130 135 140

aaa atg tat gtg cgt gat gga gtg ctg aag ggt gaa gtt aac atg gct 480
 Lys Met Tyr Val Arg Asp Gly Val Leu Lys Gly Glu Val Asn Met Ala
 145 150 155 160

ctg ttg ctt gaa gga ggt ggc cat tac cga tgt gac ttt cga agt act 528
 Leu Leu Leu Glu Gly Gly Gly His Tyr Arg Cys Asp Phe Arg Ser Thr
 165 170 175

tac aaa gcg aag aaa cgt gtt cag ttg cca gac tat cac ttt gtg gac 576
 Tyr Lys Ala Lys Lys Arg Val Gln Leu Pro Asp Tyr His Phe Val Asp
 180 185 190

cac cgc att gag att ttg agc cat gac aat gac tac aac acc gta aag 624
 His Arg Ile Glu Ile Leu Ser His Asp Asn Asp Tyr Asn Thr Val Lys
 195 200 205

ctg tcc gag aat gcc cga ggc tgc cta ttc tat gct gcc gag gca ggc 672
 Leu Ser Glu Asn Ala Arg Gly Ser Leu Phe Tyr Ala Ala Glu Ala Gly
 210 215 220

caa gta atc 681
 Gln Val Ile
 225

<210> 2

<211> 227

<212> PRT

<213> Montastraea cavernosa

<400> 2

Met Ser Val Ile Lys Pro Asp Met Lys Ile Lys Leu Arg Met Glu Gly
 1 5 10 15

Ala Val Asn Gly His Asn Phe Val Ile Glu Gly Glu Gly Lys Gly Lys
 20 25 30

Pro Phe Glu Gly Thr Gln Thr Ile Asn Leu Thr Val Lys Glu Gly Gly
 35 40 45

Pro Leu Pro Phe Ala Tyr Asp Ile Leu Thr Ala Ala Phe Gln Tyr Gly
 50 55 60

Asn Arg Ala Phe Thr Lys Tyr Pro Arg Asp Ile Ala Asp Tyr Phe Lys
 65 70 75 80

Gln Ser Phe Pro Glu Gly Tyr Ser Trp Glu Arg Ser Met Thr Tyr Glu
 85 90 95

Asp Gln Gly Ile Cys Ile Ile Lys Ser Asp Ile Arg Met Glu Gly Asp
 100 105 110

Cys Phe Ile Tyr Glu Ile Arg Tyr Asp Gly Val Asn Phe Pro Pro Ser
 115 120 125

3/8

Gly Pro Val Met Gln Lys Lys Thr Leu Lys Trp Glu Pro Ser Thr Glu
 130 135 140
 Lys Met Tyr Val Arg Asp Gly Val Leu Lys Gly Glu Val Asn Met Ala
 145 150 155 160
 Leu Leu Leu Glu Gly Gly Gly His Tyr Arg Cys Asp Phe Arg Ser Thr
 165 170 175
 Tyr Lys Ala Lys Lys Arg Val Gln Leu Pro Asp Tyr His Phe Val Asp
 180 185 190
 His Arg Ile Glu Ile Leu Ser His Asp Asn Asp Tyr Asn Thr Val Lys
 195 200 205
 Leu Ser Glu Asn Ala Arg Gly Ser Leu Phe Tyr Ala Ala Glu Ala Gly
 210 215 220
 Gln Val Ile
 225

<210> 3
 <211> 666
 <212> DNA
 <213> Pocillopora damicornis

<220>
 <221> CDS
 <222> (1)..(666)

<400> 3
 atg tct ctt tca aag cag gtc att cta aaa gac atg aat ttg aaa ttt 48
 Met Ser Leu Ser Lys Gln Val Ile Leu Lys Asp Met Asn Leu Lys Phe
 1 5 10 15
 gaa atg aaa ggt agt gtc aat gga cat tac ttt gag atc gaa ggt gaa 96
 Glu Met Lys Gly Ser Val Asn Gly His Tyr Phe Glu Ile Glu Gly Glu
 20 25 30
 gga aga gga aag cct tac gaa gga gta cag aag tcc act ttt tgg gta 144
 Gly Arg Gly Lys Pro Tyr Glu Gly Val Gln Lys Ser Thr Phe Trp Val
 35 40 45
 acc aag gga gga ccc ctt cca ttt tcc ttt gac ata ctt tcg tca gcc 192
 Thr Lys Gly Gly Pro Leu Pro Phe Ser Phe Asp Ile Leu Ser Ser Ala
 50 55 60
 ttc aaa tat gga aac aga tgc ttt act aag tac cct gcc gac atg cct 240
 Phe Lys Tyr Gly Asn Arg Cys Phe Thr Lys Tyr Pro Ala Asp Met Pro
 65 70 75 80
 gac tat ttc aaa gaa gca ttt cct gct gga atg tca ttt gaa agg aca 288
 Asp Tyr Phe Lys Glu Ala Phe Pro Ala Gly Met Ser Phe Glu Arg Thr
 85 90 95

4/8

```

ttt aca ttt gag gat gga gga gtt gct aca gcc agt gga cac att tgt 336
Phe Thr Phe Glu Asp Gly Gly Val Ala Thr Ala Ser Gly His Ile Cys
          100                      105                      110

ctt gag ggg aat tgg ttt aag cac aca tcc atg ttt cat ggc gtt aac 384
Leu Glu Gly Asn Trp Phe Lys His Thr Ser Met Phe His Gly Val Asn
          115                      120                      125

ttt ccc gct aac gga ccc ata atg caa aag agg aca att ggc tgg gac 432
Phe Pro Ala Asn Gly Pro Ile Met Gln Lys Arg Thr Ile Gly Trp Asp
          130                      135                      140

cct tcc ttt gag aaa atg acc gtc tcc aac aac ata ttg aga ggt gat 480
Pro Ser Phe Glu Lys Met Thr Val Ser Asn Asn Ile Leu Arg Gly Asp
          145                      150                      155                      160

gta act atg ttc ctt caa ctc aaa ggg gga ggt tat cac agt tgc cag 528
Val Thr Met Phe Leu Gln Leu Lys Gly Gly Gly Tyr His Ser Cys Gln
          165                      170                      175

ttt cac act tct tac aaa aca aag gag ccg gtc acg ctg cct cag aac 576
Phe His Thr Ser Tyr Lys Thr Lys Glu Pro Val Thr Leu Pro Gln Asn
          180                      185                      190

cac gtc gta gaa cat cgc att acg agg acc gac att gaa gat aaa aag 624
His Val Val Glu His Arg Ile Thr Arg Thr Asp Ile Glu Asp Lys Lys
          195                      200                      205

gtc ctg ctg gaa gaa act gct gtg gct cat gtt aac cct ttg 666
Val Leu Leu Glu Glu Thr Ala Val Ala His Val Asn Pro Leu
          210                      215                      220

```

<210> 4

<211> 222

<212> PRT

<213> Pocillopora damicornis

<400> 4

```

Met Ser Leu Ser Lys Gln Val Ile Leu Lys Asp Met Asn Leu Lys Phe
  1                      5                      10                      15

Glu Met Lys Gly Ser Val Asn Gly His Tyr Phe Glu Ile Glu Gly Glu
          20                      25                      30

Gly Arg Gly Lys Pro Tyr Glu Gly Val Gln Lys Ser Thr Phe Trp Val
          35                      40                      45

Thr Lys Gly Gly Pro Leu Pro Phe Ser Phe Asp Ile Leu Ser Ser Ala
          50                      55                      60

Phe Lys Tyr Gly Asn Arg Cys Phe Thr Lys Tyr Pro Ala Asp Met Pro
          65                      70                      75                      80

Asp Tyr Phe Lys Glu Ala Phe Pro Ala Gly Met Ser Phe Glu Arg Thr
          85                      90                      95

```

5/8

Phe Thr Phe Glu Asp Gly Gly Val Ala Thr Ala Ser Gly His Ile Cys
 100 105 110
 Leu Glu Gly Asn Trp Phe Lys His Thr Ser Met Phe His Gly Val Asn
 115 120 125
 Phe Pro Ala Asn Gly Pro Ile Met Gln Lys Arg Thr Ile Gly Trp Asp
 130 135 140
 Pro Ser Phe Glu Lys Met Thr Val Ser Asn Asn Ile Leu Arg Gly Asp
 145 150 155 160
 Val Thr Met Phe Leu Gln Leu Lys Gly Gly Gly Tyr His Ser Cys Gln
 165 170 175
 Phe His Thr Ser Tyr Lys Thr Lys Glu Pro Val Thr Leu Pro Gln Asn
 180 185 190
 His Val Val Glu His Arg Ile Thr Arg Thr Asp Ile Glu Asp Lys Lys
 195 200 205
 Val Leu Leu Glu Glu Thr Ala Val Ala His Val Asn Pro Leu
 210 215 220

<210> 5
 <211> 681
 <212> DNA
 <213> Discosoma striata

<220>
 <221> CDS
 <222> (1)..(681)

<220>
 <221> modified_base
 <222> (662)
 <223> a, t, c or g

<400> 5
 atg agt tgg tcc aag agt gtg atc aag gaa gaa atg ttg atc aat ctt 48
 Met Ser Trp Ser Lys Ser Val Ile Lys Glu Glu Met Leu Ile Asn Leu
 1 5 10 15

 cat ctg gaa gga acg ttc aat ggg cac tac ttt gaa ata aaa ggc aaa 96
 His Leu Glu Gly Thr Phe Asn Gly His Tyr Phe Glu Ile Lys Gly Lys
 20 25 30

 gga aaa ggg aac cct aat gaa ggc acc aat acc gtc acg ctc gag gtt 144
 Gly Lys Gly Asn Pro Asn Glu Gly Thr Asn Thr Val Thr Leu Glu Val
 35 40 45

 acc aag ggt gga cct ctg cca ttt ggt tgg cat att ttg tgc cca caa 192
 Thr Lys Gly Gly Pro Leu Pro Phe Gly Trp His Ile Leu Cys Pro Gln
 50 55 60

6/8

ttt cag tat gga aac aag gca ttt gtc cac cac cct gac gac ata ccc	240
Phe Gln Tyr Gly Asn Lys Ala Phe Val His His Pro Asp Asp Ile Pro	
65 70 75 80	
gat tat cta aag ctg tca ttt ccg gag gga tat aca tgg gaa cgg tcc	288
Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ser	
85 90 95	
atg cac ttt gaa gac ggt ggc ttg tgt tgt atc acc aat gat atc agt	336
Met His Phe Glu Asp Gly Gly Leu Cys Cys Ile Thr Asn Asp Ile Ser	
100 105 110	
ttg aca ggc aac tgt ttc aac tac gac atc aag ttc act ggc ttg aac	384
Leu Thr Gly Asn Cys Phe Asn Tyr Asp Ile Lys Phe Thr Gly Leu Asn	
115 120 125	
ttt cct cca aat gga ccc gtt gtg cag aag aag aca act ggc tgg gaa	432
Phe Pro Pro Asn Gly Pro Val Val Gln Lys Lys Thr Thr Gly Trp Glu	
130 135 140	
ccg agc act gag cgt ttg tat cct cgt gat ggc gtg ttg ata gga gac	480
Pro Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Ile Gly Asp	
145 150 155 160	
atc cat cat gct ctc aca gtg gca gga ggt ggt cat tac gta tgt gac	528
Ile His His Ala Leu Thr Val Ala Gly Gly Gly His Tyr Val Cys Asp	
165 170 175	
att aaa act gtt tac agg gcc aaa aas ccc gta aag atg cca ggg tat	576
Ile Lys Thr Val Tyr Arg Ala Lys Xaa Pro Val Lys Met Pro Gly Tyr	
180 185 190	
cac tat gtt gac ccc aaa ctg gtt ata agg agc aac gac aaa gaa ttc	624
His Tyr Val Asp Pro Lys Leu Val Ile Arg Ser Asn Asp Lys Glu Phe	
195 200 205	
atg aaa gtt gag gag cat gaa atc gcc gtt gca cgc cnc cat ccg ctc	672
Met Lys Val Glu Glu His Glu Ile Ala Val Ala Arg Xaa His Pro Leu	
210 215 220	
caa agc caa	681
Gln Ser Gln	
225	

<210> 6

<211> 227

<212> PRT

<213> Discosoma striata

<220>

<221> MOD_RES

<222> (185)

<223> Asn or Lys

7/8

<220>

<221> MOD_RES

<222> (221)

<223> Leu, Pro, His or Arg

<400> 6

Met Ser Trp Ser Lys Ser Val Ile Lys Glu Glu Met Leu Ile Asn Leu
 1 5 10 15

His Leu Glu Gly Thr Phe Asn Gly His Tyr Phe Glu Ile Lys Gly Lys
 20 25 30

Gly Lys Gly Asn Pro Asn Glu Gly Thr Asn Thr Val Thr Leu Glu Val
 35 40 45

Thr Lys Gly Gly Pro Leu Pro Phe Gly Trp His Ile Leu Cys Pro Gln
 50 55 60

Phe Gln Tyr Gly Asn Lys Ala Phe Val His His Pro Asp Asp Ile Pro
 65 70 75 80

Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ser
 85 90 95

Met His Phe Glu Asp Gly Gly Leu Cys Cys Ile Thr Asn Asp Ile Ser
 100 105 110

Leu Thr Gly Asn Cys Phe Asn Tyr Asp Ile Lys Phe Thr Gly Leu Asn
 115 120 125

Phe Pro Pro Asn Gly Pro Val Val Gln Lys Lys Thr Thr Gly Trp Glu
 130 135 140

Pro Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Ile Gly Asp
 145 150 155 160

Ile His His Ala Leu Thr Val Ala Gly Gly Gly His Tyr Val Cys Asp
 165 170 175

Ile Lys Thr Val Tyr Arg Ala Lys Xaa Pro Val Lys Met Pro Gly Tyr
 180 185 190

His Tyr Val Asp Pro Lys Leu Val Ile Arg Ser Asn Asp Lys Glu Phe
 195 200 205

Met Lys Val Glu Glu His Glu Ile Ala Val Ala Arg Xaa His Pro Leu
 210 215 220

Gln Ser Gln
 225

<210> 7

<211> 45

<212> RNA

<213> Artificial Sequence

8/8

<220>

<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 7

gcugauggcg augaaugaac acugcguug cuggcuuuga ugaaa

45

<210> 8

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
primer

<400> 8

ctcgagaagc ttgaattcg atcctttttt ttttttttt t

41

<210> 9

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
primer

<400> 9

gctgatggcg atgaatgaac actg

24

<210> 10

<211> 37

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
primer

<400> 10

ccggtaccta aggaggccac catgagtgtg ataaaac

37

<210> 11

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
primer

<400> 11

ggtctagatt acttggcctg cctc

24