Journal of the Marine Biological Association of the United Kingdom

cambridge.org/mbi

Original Article

Cite this article: Dizon EGS, Da-Anoy JP, Roth MS, Conaco C (2021). Fluorescent protein expression in temperature tolerant and susceptible reef-building corals. *Journal of the Marine Biological Association of the United Kingdom* **101**, 71–80. https://doi.org/10.1017/ S0025315421000059

Received: 28 March 2020 Revised: 11 January 2021 Accepted: 19 January 2021 First published online: 2 March 2021

Key words:

Bleaching tolerance; fluorescent proteins; thermal stress; transcriptome

Author for correspondence: Cecilia Conaco, E-mail: cconaco@msi.upd.edu.ph

© The Author(s), 2021. Published by Cambridge University Press on behalf of The Marine Biological Association of the United Kingdom



Fluorescent protein expression in temperature tolerant and susceptible reef-building corals

Exequiel Gabriel S. Dizon¹ ⁽¹), Jeric P. Da-Anoy¹ ⁽¹), Melissa S. Roth² ⁽¹) and Cecilia Conaco¹ ⁽¹)

¹Marine Science Institute, College of Science, University of the Philippines, Diliman, Quezon City, 1101, Philippines and ²Department of Plant and Microbial Biology, University of California, Berkeley CA, 94720-3102, USA

Abstract

Fluorescent proteins (FPs) are reported to play an important role as photoprotectants and antioxidants in corals subjected to stressful conditions. Identifying the various FP genes expressed and FP gene expression patterns under stress in diverse coral species can provide insight into FP function. In this study, we identified 16 putative FP homologues from the transcriptomes of corals with varying susceptibility to elevated temperature, including *Acropora digitifera*, *Favites colemani*, *Montipora digitata* and *Seriatopora caliendrum*. Each coral expressed a different complement of FP transcripts, which were predicted to have distinct spectral properties. The most diverse and abundant repertoire of FP transcripts, including at least 6 green FPs, were expressed fewer FP types. Specific FP transcripts exhibited variable expression profiles in coral fragments subjected to 32 ± 1 °C (treatment) or 28 ± 1 °C (control) for up to 72 h, suggesting that distinct FPs may have different roles. Further studies on the expression of the proteins encoded by these FP transcripts, their fluorescence activity, tissue localization, and possible antioxidant properties, are needed to reveal their contribution to thermal stress tolerance in certain species of corals.

Introduction

Scleractinian corals are responsible for building and maintaining the three-dimensional calcium carbonate matrix of the coral reef (Erez et al., 2011). The interactions between the coral host and its microalgal symbionts from family Symbiodiniaceae are crucial for the stability of the coral holobiont. However, the symbiotic relationship between the coral host and its microalgal symbionts is vulnerable to disturbances such as rising sea surface temperatures that accompany climate change (Hughes et al., 2003; Baird et al., 2009; Hoegh-Guldberg & Bruno, 2010). Various factors contribute to the susceptibility of corals to stressors and the overall stress response. In the case of thermal stress, pre-existing variations due to acclimatization (Bellantuono et al., 2011; Hume et al., 2013), symbiont composition (Cziesielski et al., 2018), and genetic variability (Bay & Palumbi, 2014) all contribute to differences in susceptibility. The upregulation of stress-related genes prior to exposure to stress, or frontloading (Barshis et al., 2013), may increase the resilience of corals that have been exposed to elevated temperatures for either short-term (Bellantuono et al., 2011) or long-term (Hume et al., 2013) periods prior to exposure to bleaching temperatures. The array of genes expressed by the host and symbionts in response to thermal stress, including genes involved in antioxidant systems, heat-shock proteins, calcium homeostasis, cytoskeletal reorganization, and genes with photoprotective properties such as fluorescent proteins (Salih et al., 2000; Desalvo et al., 2008; Baird et al., 2009; Mayfield et al., 2012; Deschaseaux et al., 2014; Pinzon et al., 2015) support further differences in the observed response across coral species and even among individuals of the same species (Parkinson et al., 2015). Of particular interest in this study are the fluorescent proteins, which are reported to confer various protective functions to the coral holobiont against environmental stressors (Schlichter et al., 1986; Salih et al., 1998, 2000, 2006; Mazel et al., 2003; Bou-Abdallah et al., 2006; Matz et al., 2006; Palmer et al., 2009; Roth et al., 2010; Smith et al., 2013; Gittins et al., 2015; Aihara et al., 2019).

The first green fluorescent protein (GFP) was discovered in the jellyfish, *Aequorea victoria*, as a component of a bioluminescent system with aequorin and luciferin (Shimomura, 2005). Structural characterization of GFP revealed an 11-stranded β -barrel flanked by lids on both sides, enclosing a central α -helix (Ormö *et al.*, 1996). The chromophore structure consists of a few well-conserved amino acid residues within the β -barrel. The glycine at position 67 and the tyrosine residue at position 66 are found in all naturally forming fluorescent proteins regardless of emitted colour (Barondeau *et al.*, 2006; Chudakov *et al.*, 2010; Stepanenko *et al.*, 2011, 2013). The residue at position 65 varies greatly among GFP-like proteins and affects the resulting chemical structure of the chromophore, as well as its spectral properties (Stepanenko *et al.*, 2011). The chemical environment surrounding the chromophore also affects the spectral properties of the protein (Heim *et al.*, 1994; Follenius-Wund *et al.*, 2003). Since the discovery of GFP, a variety of proteins with similar structures have been discovered in many other

species (Shagin *et al.*, 2004; Stepanenko *et al.*, 2013). The most abundant and colour diverse set of FPs, with a spectral range spanning cyan, green and red, as well as non-fluorescent chromo-proteins, are found in class Anthozoa of Phylum Cnidaria (Verkhusha & Lukyanov, 2004; Alieva *et al.*, 2008).

Fluorescent proteins contribute to the diversity of colours observed in coral reef organisms (Dove et al., 2001) and have been ascribed various roles. For example, green fluorescence emitted by GFPs promote symbiont phototaxis towards corals (Aihara et al., 2019) while cyan and green fluorescence are thought to counterbalance the brown colouration of symbiont pigments to make corals appear less palatable to herbivorous fishes (Matz et al., 2006). FPs have also been linked to protective functions, which include photoprotection (Salih et al., 1998, 2000, 2006; Roth et al., 2010; Smith et al., 2013) and photoacclimation (Roth et al., 2010; Smith et al., 2013; Gittins et al., 2015). Autofluorescent coral species, such as Faviids, Agariciids and some Poritids, usually possess dense layers of fluorescent protein granules concentrated in large clusters above the symbionts to form a pigment screen that may protect against intense illumination (Salih et al., 1998), while these granules were found among or below the symbionts in dim light environments suggesting a light-enhancing role for photosynthesis through wavelength transformation and back scattering (Schlichter et al., 1986). The abundance of fluorescent proteins in corals living in both high light and dim light environments support a conserved, integral role of these proteins in managing the internal light environment of the coral holobiont through photoprotection and photoacclimation (Roth et al., 2010). Corals exhibit intraspecific variation in FP expression among individuals (Gittins et al., 2015) and subpopulations (Takahashi-Kariyazono et al., 2018), and across developmental stages (Kenkel et al., 2011), even under identical light environments. Variation among similar individuals may be due to differences in the number of FP gene copies or polymorphisms in the genome (Takahashi-Kariyazono et al., 2018), as well as to differences in the regulation of FP expression (Gittins et al., 2015).

Fluorescent proteins possess antioxidant properties and can eliminate reactive oxygen species and peroxides (Mazel et al., 2003; Bou-Abdallah et al., 2006; Palmer et al., 2009) that are produced as a consequence of elevated temperatures and are the primary cause of coral bleaching (Lesser, 2006; Gardner et al., 2017). Purified recombinant GFP protein was demonstrated to have superoxide radical quenching activity that increased linearly with protein concentration (Bou-Abdallah et al., 2006). In addition, a significant positive correlation between in vivo peroxide scavenging rates and fluorescent protein concentration has been observed in Caribbean corals (Palmer et al., 2009). Different fluorescent types of FPs exhibit varying levels of peroxide scavenging activity (Palmer et al., 2009), suggesting that the FP complement in corals may be an important determinant of coral stress tolerance. However, the relationship between fluorescence, fluorescent protein concentrations, and stress has been studied in only a few coral species (Roth & Deheyn, 2013). Further studies comparing the FP gene complement and FP expression in corals that are known to be susceptible or tolerant to stressors, such as elevated temperature, remain to be conducted.

This study thus aimed to identify FP transcripts in the corals *Acropora digitifera* (Dana, 1846), *Favites colemani* (Veron, 2000), *Montipora digitata* (Dana, 1846) and *Seriatopora calien-drum* (Ehrenberg, 1834) and to determine the response of selected transcripts to elevated temperature. A total of 16 transcripts putatively encoding FPs of different spectral types were identified by homology search from the transcriptomes of these corals. Real-time PCR (qPCR) revealed that the expression of selected FP transcripts varied among coral species subjected to elevated

temperature conditions but did not show common trends correlating with duration of exposure.

Materials and methods

Coral collection and acclimation

Three colonies each of Acropora digitifera (16'17.287N 120'00.448E), Favites colemani (16'24.708N 119'54.270E), Montipora digitata (16'26.513N 119'56.494E) and Seriatopora caliendrum (16'22.293N 120'00.228E) were collected in November 2016 from various sites within the Bolinao-Anda Reef Complex, north-western Philippines, where each coral genus is abundant. These coral species are ranked based on increasing thermal bleaching susceptibility as follows: F. colemani < A. digitifera < M. digitata < S. caliendrum (Da-Anoy et al., 2019). Based on regular monitoring by the Bolinao Marine Laboratory, sea surface temperatures within the reef complex range from 25-32 °C with annual mean temperature of 28.89 ± 0.90 °C and a maximum of 31-33 °C during the summer season. Shallow collection sites (2-9 m) had similar temperature profiles but some variation in light (~400-1250 lux). Sample collection was conducted with the permission of the Philippines Department of Agriculture Bureau of Fisheries and Aquatic Resources (DA-BFAR Gratuitous Permit no. 0102-15). Colonies at least 10-15 m apart horizontally were collected to minimize genotypic similarity, although validation of genotypes was not conducted. Corals were fragmented into 2.5-5.0 cm long nubbins and allowed to heal for 2 weeks in outdoor tanks with running seawater maintained at 28 ± 1 °C and illumination under low photosynthetic photon flux density of $\sim 80-90 \,\mu\text{mol}\,\text{m}^{-2}\,\text{s}^{-1}$ on a 12:12 h light-dark cycle. Fragments were tagged to keep track of their colony of origin. Healed fragments were acclimated for 2 weeks in indoor experimental tanks with running seawater maintained at 28 ± 1 °C and illumination of ~80 µmol m⁻² s⁻¹ on a 12:12 h light-dark cycle.

Thermal stress experiment

Thermal stress experiments were conducted in 401 tanks with constantly aerated, 10 µm-filtered flow-through seawater, as described in Da-Anoy et al. (2019). Two independent replicate tanks were used for each temperature treatment. Each tank received seawater from chilled reservoirs maintained at 27 °C and the temperature in individual experimental tanks was adjusted using submersible thermostat heaters. Flow rate was maintained at $\sim 5-81h^{-1}$ with additional mixing provided by 6001 h⁻¹ pumps. All setups received illumination under low photosynthetic photon flux density of ${\sim}80\,\mu\text{mol}\,m^{-2}\,s^{-1}$ on a 12:12 light-dark cycle to avoid light stress. Light and temperature in each experimental tank were monitored using submersible loggers (Onset HOBO). 10-12 coral fragments of each species were quickly transferred into each experimental tank set at either elevated temperature (32 ± 1 °C, treatment) or ambient temperature (28 ± 1 °C, control). Although the thermal treatment does not approximate what happens in nature, it allows testing of the robustness of the response of each species to acute temperature shock, as has been done in other studies (Barshis et al., 2013). Coral fragments were collected after 4, 24, 48 and 72 h and flash frozen in liquid nitrogen for storage and transport.

RNA extraction and sequencing

Coral fragments frozen in liquid nitrogen were ground into a fine powder using a mortar and pestle. Total RNA was extracted using TRIzol Reagent (Invitrogen) following the manufacturer's instructions. Contaminating DNA was removed using the TURBO DNA-free Kit (Invitrogen). RNA concentration was measured using a NanoDrop 2000c spectrophotometer (ThermoScientific) and quality was checked by gel electrophoresis. Libraries were prepared using the Illumina TruSeq RNA Sample Prep Kit. A total of 12 libraries were constructed for each species with three replicates each of the control and heated samples from the 4 and 24 h timepoints. Acropora digitifera and S. caliendrum libraries were sequenced on the HiSeq 2500 platform at BGI Genomics, Hong Kong, while M. digitata and F. colemani were sequenced on the HiSeq 4000 platform at Macrogen, South Korea. Sequencing generated 100 bp paired-end reads. Transcriptomes were assembled de novo using Trinity (Grabherr et al., 2011). Transcript abundance estimation was performed by mapping individual paired-end reads back to the reference transcriptome assembly using RNASeq by Expectation Maximization (RSEM) (Li & Dewey, 2011) with the Bowtie alignment method (Langmead et al., 2009). Analysis of differentially expressed genes was conducted using the edgeR (Robinson et al., 2010) package in R. Transcript abundance expressed as transcripts per million reads (TPM) was generated in RSEM and used for visualization of FP gene expression. De novo transcriptome assemblies have been deposited at DDBJ/EMBL/GenBank under the following accessions: GIVI00000000 (A. digitifera), GIVN00000000 (F. colemani), GIVM00000000 (M. digitata) and GIVG00000000 (S. caliendrum). Sequence reads are available from NCBI under the following BioProject accessions: PRJNA 421253 (A. digitifera), PRJNA422022 (F. colemani), PRJNA 422015 (M. digitata) and PRJNA422012 (S. caliendrum).

Identification of putative FPs in corals

Known coral FP sequences from Alieva et al. (2008) were used as query to identify similar sequences in the translated transcriptomes of A. digitifera, F. colemani, M. digitata and S. caliendrum using the blastp algorithm of Basic Local Alignment Search Tool (BLAST) at an *E*-value cutoff of 1×10^{-5} . Peptides were aligned in BioEdit (Hall, 1999) using ClustalW (Thompson et al., 1994) (Supplementary File S1). Alignments were visualized using CLC Sequence Viewer 8 (QIAGEN Bioinformatics) and used to identify the presence of a chromophore region, along with diagnostic residues for chromophore maturation and structure stability. Domains were identified by mapping predicted peptides against the Pfam 32.0 seed database (El-Gebali et al., 2019) using HMMER v3.2.1 (Eddy, 1998) accessed with Cygwin v2.11.2 (Lazenby, 2000). To estimate the average expression of FPs and allow comparisons across species, we obtained the log2--transformed expression value for each FP (TPM_{FP}) normalized against the average expression value of selected reference genes $(\text{mean TPM}_{\text{REF}})$ in each sample using the equation below.

Average FP expression =
$$\left(\log_2\left(\left(\frac{\text{TPM}_{\text{FP}}}{\text{mean TPM}_{\text{REF}}} \times 10,000\right) + 1\right)\right)$$

The list of reference genes and expression normalization method were adapted from Takahashi-Kariyazono *et al.* (2018). Transcripts corresponding to each reference gene were identified using corresponding accessions from EggNOG (Jensen *et al.*, 2008) or COG (Tatusov *et al.*, 2000) (Table S1). Only reference genes common across all four coral species were included in the analysis.

Phylogenetic analysis

Nucleotide sequences of FP genes were aligned with other FP sequences retrieved from Alieva et al. (2008), the NCBI non-

redundant nucleotide database, and from the ReefGenomics database (Liew *et al.*, 2016) (Table S2). Nucleotide sequences were aligned using ClustalW (Thompson *et al.*, 1994) and ambiguous positions were trimmed using Gblocks (Castresana, 2000). A total of 130 sequences were included in the alignment (Supplementary File S2). The appropriate model of evolution was identified as GTR + G + I (Tavare, 1986) as indicated by MEGA7 (Kumar *et al.*, 2016). Phylogenetic analysis was performed using MrBayes (Huelsenbeck & Ronquist, 2001) with arthropod FP sequences as outgroup. The MCMC chain was run for 1,000,000 iterations with a sample frequency of 100 resulting in 10,000 trees, of which the first 2500 were discarded while summarizing the data.

Quantitative real-time PCR assay (qPCR)

cDNA was produced by reverse transcription using the GoScriptTM (Promega) kit. Briefly, 3 µl of total RNA was mixed with 1 µl each of $500 \,\mu g \,ml^{-1}$ oligo(dT) and random hexamer primers, incubated at 70 °C for 5 min then on ice for 5 min. 15 µl of reverse transcription reaction mix was added (4.5 µl nuclease-free water, 4 µl 10 × GoScriptTM reaction buffer, 4 µl 25 mM MgCl₂, $1.0\,\mu l~0.5\,m M$ nucleotide mix, $0.5\,\mu l~40\,U\,\mu l^{-1}$ Recombinant RNasin[®] Ribonuclease Inhibitor, and 1µl GoScript[™] Reverse Transcriptase) and samples were incubated at 25 °C for 5 min, then at 42 °C for 1 h on a PCR heat block. cDNAs were stored at -20 °C. Primers for qPCR were designed using Primer-BLAST (Ye et al., 2012) against selected FP transcripts exhibiting the highest mean expression in each coral species (Table S3). Note that FP transcripts in A. digitifera and S. caliendrum have overlapping sequences and can be amplified by the same primer set for each species. qPCR reactions were performed on a CFX96 Touch[™] Real-Time PCR Detection System with activation at 95 °C for 2 min, denaturation at 95 °C for 30 s, and annealing/elongation at 60 °C for 1 min. Each reaction contained $5\,\mu l$ of $2\times GoTaq^{*}$ (Promega) qPCR Master Mix, 0.5 μl each of 10 µM forward and reverse primers, 1 µl template cDNA, and 3 µl nuclease-free H₂O. Three biological replicates and three technical replicates were used in the quantitation of each gene alongside negative controls. Primer efficiency and primer specificity were assessed using the dilution curves and melt curves, respectively. The abundance of target transcripts was determined from primerspecific relative standard curves generated using 10-fold serial dilutions of coral cDNA (Figure S1). Target transcript abundances were normalized to actin as a reference gene. Actin has previously been used as a reference gene in other cnidarian gene expression studies (Rodriguez-Lanetty et al., 2006; Gajigan & Conaco, 2017) and its expression was unchanged in the temperature treatments. Mann-Whitney U tests (Mann & Whitney, 1947) were used to compare the relative fold change in expression of treated vs control samples. Statistical analyses were conducted in R v. 3.5.1 (R Core Team, 2013) using the base statistical package.

Results

Identification of putative FPs in corals

Eight FP transcripts were identified in *F. colemani*, two in *A. digitifera*, three in *M. digitata* and three in *S. caliendrum*. Some transcripts had similar sequences and common best BLAST hits (Table 1) indicating that they are transcript fragments from the same gene. Most of the transcripts encoded partial FPs with predicted peptide lengths ranging from 119–280 amino acids, which may be attributed to incomplete assembly of sequences in the transcriptome. Nevertheless, each of the peptides had a domain

Transcript	Top BLAST match	Species	NCBI accession	E value	% Identity	PFAM domain
AdigFP1	S/Me fluorescent protein FP2_S1603_2	Acropora digitifera	LC177541.1	0.0	100	PF01353.22
AdigFP2	PREDICTED: fluorescent chromoprotein amFP486 (LOC114976946)	Acropora millepora	XM_029357520.1	0.0	100	PF01353.22
FcolFP1	Green fluorescent protein	<i>Scleractinia</i> sp. Lizard Island 35	GQ385210.1	0.0	99	PF01353.22
FcolFP2	Green fluorescent GFP-like protein	Platygyra lamellina	EU498724.1	0.0	92.4	PF01353.22
FcolFP3	Red fluorescent GFP-like protein	Mycedium elephantotus	DQ206386.1	0.0	95	PF01353.22
FcolFP4	Red fluorescent GFP-like protein	Mycedium elephantotus	DQ206386.1	0.0	94.4	PF01353.22
FcolFP5	GFP-like fluorescent chromoprotein cFP484	Orbicella faveolata	XM_020749854.1	6×10^{-144}	88.6	PF01353.22
FcolFP6	Green fluorescent GFP-like protein	Favites abdita	EU498723.1	0.0	95.0	PF01353.22
FcolFP7	Green fluorescent protein	<i>Scleractinia</i> sp. Lizard Island 36	GQ385223.1	0.0	85.1	PF01353.22
FcolFP8	Green fluorescent protein	Montastraea cavernosa	EU035529.1	0.0	89	PF01353.22
MdigFP1	Cyan fluorescent GFP-like protein	Montipora millepora	DQ206392.1	0.0	97.7	PF01353.22
MdigFP2	GFP-like chromoprotein	Montipora efflorescens	DQ206377.1	0.0	98.8	PF01353.22
MdigFP3	Green fluorescent protein	<i>Montipora</i> sp. M5	LC029025.1	0.0	99.6	PF01353.22
ScalFP1	PREDICTED: fluorescent chromoprotein amFP486 (LOC111319708)	Stylophora pistillata	XM_022922450.1	9×10 ⁻¹²⁹	92.8	PF01353.22
ScalFP2	GFP-like chromoprotein	Stylophora pistillata	DQ206398.1	2×10^{-130}	91.2	PF01353.22
ScalFP3	PREDICTED: fluorescent chromoprotein amFP486 (LOC111319708)	Stylophora pistillata	XM_022922450.1	9×10 ⁻¹²⁹	92.8	PF01353.22

Table 1. Top nucleotide BLAST alignments of the putative FP coding sequences from *A. digitifera*, *F. colemani*, *M. digitata* and *S. caliendrum* against the NCBI non-redundant database and HMMER domain matches of the translated sequences to the Pfam 32.0 seed database

with significant similarity (*E*-value $\leq 6.70 \times 10^{-20}$) to GFP (PF01353.22) in the PFAM database.

Alignment of the 16 FP peptides revealed extensive sequence similarity to *Aequorea victoria* GFP (P42212) (Figure 1). Glycine residues necessary for the stability of the β -barrel structure (Fu *et al.*, 2015) were observed in most of the FP sequences, although some positions were not represented in the partial sequences of FcoIFP7, ScaIFP1 and ScaIFP2. Residues necessary for chromophore maturation, specifically Thr62 (Yang *et al.*, 1996), Arg96 (Sniegowski *et al.*, 2005), Ser205 (Ormö *et al.*, 1996) and Glu222 (Sniegowski *et al.*, 2005) were also found in the majority of the FP sequences. Chromophore residues aligning with positions 65–67 of GFP, which are essential for fluorescence (Barondeau *et al.*, 2006; Stepanenko *et al.*, 2011, 2013), were present in all sequences except in the incomplete sequences of AdigFP1, FcoIFP5 and FcoIFP6.

Although the presence of specific chromophore residues can be informative in predicting the spectral properties of an FP, interactions within the β -barrel (Stepanenko *et al.*, 2013) and external factors such as pH (Chudakov *et al.*, 2010) also affect the resulting excitation and emission spectra. Thus, in this study we looked at both chromophore sequences and phylogenetic evidence to infer the potential spectral properties of candidate coral FPs. Phylogenetic analysis of the 16 coral FPs identified in this study revealed their affiliation with different FP clades (Figure 2) that represent groupings based on both taxonomy and predicted fluorescence types (Alieva *et al.*, 2008). A total of eight FP transcripts were found in the transcriptome of *F. colemani*, with 6 complete and 2 incomplete sequences. All the

transcripts clustered in clade D with other faviid FPs. FcolFP1 was 99% identical to a green FP from an unknown scleractinian coral (GQ385210.1: Gruber et al., 2009) and a green FP from Echinophyllia echinata (DQ206383: Alieva et al., 2008). FcolFP2 closely clustered with another faviid sequence from ReefGenomics (Favia sp. 44353: Bhattacharya et al., 2016) and its sequence was 92.4% identical to a green fluorescent FP from Platygyra lamellina (EU498724.1: Alieva et al., 2008). FcolFP5 and FcolFP6 grouped with green FPs from E. echinata (DQ206395: Alieva et al., 2008) and Favites abdita (EU498723: Alieva et al., 2008). FcolFP7, the longest translated sequence described in this study at 280 aa, had a nucleotide sequence with 85.1% identity to green FP from an unknown scleractinian coral (GQ385223.1: Gruber et al., 2009) and grouped with cyan FPs from Montastraea cavernosa (AY181556) and Mycedium elephantotus (DQ206382). FcolFP8 clustered with green FPs KikG from Dipsastraea favus (AB193294) and Montastraea annularis (AY037766). FcolFP3 and FcolFP4 both clustered closely with a red FP from M. elephantotus (DQ206386: Alieva et al., 2008). Both FcolFP1 and FcolFP2 have the EYG chromophore, while FcolFP3, FcolFP4 and FcolFP8 have HYG, and FcolFP7 has QYG (Figure 1). These data suggest that FcolFP1, FcolFP2, FcolFP5, FcolFP6, FcolFP7 and FcolFP8 are transcripts for green fluorescent proteins while FcolFP3 and FcolFP4 encode red fluorescent proteins.

The two FP transcripts from *A. digitifera*, AdigFP1 and AdigFP2, clustered among other *Acropora* cyan FP sequences in clade C2 (Figure 2). These FPs displayed significant sequence similarity to a predicted FP from the genome of *A. digitifera* (XM_015895650), a short/middle wavelength-emitting fluorescent

		20		40		60		80
GFP						MGKGEELFTG	VVPILVELDG	VNGHKFSVS 30 AVNGHIFEID 44 AVNGHKFVII 34 TVNGHYFVIE 30 TVNGHYFVIE 49
FcolFP1				NVKH	KCCIRRISTF	TSTIMSVIKP	DMKIKLHMKG	AVNGHIFEID 44
FcolFP2					AGEE	YKKSENVIKP	DMKIKLRMEG	AVNGHKFVII 34
FCOIFP3				I KERVGOOO	OFIDETDUIR	MSTIMSVIKP	DMKLKLRMEG	TVNGHYFVIE 30
FcolFP5				- LKENVGQQQ	QELUKIDILF	13111131114	DWKTKLRWEG	1 VINGH 1 F VIE 49
FcolFP6								
FcolFP7	MASFFFVLFV	LLCCTSANPV	PDNHEQTKKS	YFSPEKVEDF	TSDVRDAGEE	YMETKSVVRP	IMNIMLHMTG	NVNGHAFEVK 80
FcolFP8						MSVIKK	DMKIELRMEG	AVNGHKFVVT 26
AdigFP1								
AdigFP2 MdigEP1						MAI PK	EMKITYHMEG	NVNGHEELIK 25
MdigFP2						MSVIAK	QMTYKVYMSG	TVNGHYFEVQ 26
MdigFP3						MALSKQGVEG	KMDLKFHMEG	SVNGHEFTIK 30
ScalFP1					MLLA	VRVVTKVLAD	TMKMTWLMEG	NINGHAFIIE 34
ScalFP2							THE	SINGHAFTIE 10
ScalFP3					MLLA	VHVVIKVLAD	IMKMIWLMEG	NVNGHAFEVK 80 AVNGHKFVVT 26 NVNGHFFIIK 25 TVNGHYFEVQ 26 SVNGHEFTIK 30 NINGHAFIIE 34 SINGHAFTIE 10 NINGHAFIIE 34
Conservation						000000000000000000000000000000000000000		
0%		100		120		140		160
CEP	REGERDATING	KITIKEICTT	GKI PV PWPT	INTERVOVO	CESEVEDHMK	BHDEEKSAMP	E. OVVOERTI	FFKDDGNYKT 108
EcolEP1	GEGNGKPFEG	KOTIELKVVD	GGPLPFAFDI	LTTVFEYGNB	VFAKYPPEI-	- VDYFKOSEP	E- GYSWERSM	MYEDGGICIA 121
EcolEP2	GKGEGKPYEG	TQTIDLEVIA	GAPLPFAFDI	LTTVFEYGNR	VFAKYPQKI-	- EDYFKQSFP	G-GYSWERSM	TYEDGGICIA 111
FcolFP3	GDGKGKPFEG	KQSMDLDVKE	GGPLPFAYDI	LTTAFHYGNR	VFAEYPDHI-	- PDYFKQSFP	G-GYSWERSL	TFEDGGICIA 107 TFEDGGICIA 126
FcoIFP4	GDGKGKPFDG	KQSMDLDVKE	GGPLPFAYDI	LTTAFHYGNR	VFAEYPDHI-	- QDYFKQSFP	G-GYSWERSL	TFEDGGICIA 126
FcolFP5				· · · · · · · · · · · · · · · · · · ·			YLWERSM	TYEDGGICIA 17 TYEDGGICIA 17
FCOIFP6	GEGEGKAVEG	TOTIVIEVTE	GARIPEAEDI	ITTAFOYOND	AETNYPPDI	PDEEKESEP	E. CHSW/PTM	I FDDGGVCDV 157
FcolFP8	GEGRGOPFEG	IQNMNLTVID	GGPLPFAFDI	LTTVFHYGNB	AFVKYPKDI-	- PDYFKOSFP	E- GFSWERSM	TYEDGGICIA 103
AdiaFP1					M-	- HDYFKQAFP	D- GMSYERSF	LYEDGGVATA 29
AdiaFP2		LVITKPA	GKPLPFSFDI	LSTVFQYGNR	CFTKYPEGM-	- TDYFKQAFP	D- GMSYERSF	LYEDGGVATA 64
MdigFP1	GEGVGEPHEG	THTIKLQVVE	GSPLPFSADI	LSTVFQYGNR	CFTKYPPNI-	- VDYFKNSCS	GGGYTFGRSF	LYEDGAVCTA 103
MdigFP2	GDGKGKPYEG	EQTVKLTVTK	GGPLPFAWDI	LSPLSQYGSI	PFTKYPEDI-	- PDYVKQSFP	E- GYTWERIM	NFEDGAVCTV 103
MdigFP3	GEGTGQPYEG	TOCTOLRVEK	GGPLPFSVDI	LSAVFLYGNR	CITKYPRGI-	- VDYFKNSCP	D- GYKWERSF	LFEDGAVCTA 107 IFEDGGFATA 111
		KQTRTFRVTK		VAPTIKYGEK	CEMKYPADI-	- PDYEKLAES	E-GLTYIRSI	TFEDGGSATA 87
ScalFP3	GEGTGKPYEG	KQTGTFRVTK	GGPLPFAFDI	VAPTLKYGFK	CFMKYPADI-	- PDYFKLAFP	E- GLTYDTSI	IFEDGGFATA 111
100%		-						
Conservation								
0.0		180 		200		220		240
GFP	RAEVKFE	180 I GDT L VNR I	ELKGIDFKED	GNILGHKL-E	YNYNSHNVYI	MADKQKNGIK	VNFKIRHNIE	DGSVQLADHY 182
GFP FcolFP1	TNNITLLKDA	180 I GDTLVNRI HGVDYFYYNI	ELKGIDFKED RFDGVNFPAN	GNILGHKL-E GPVMQKKTVK	YN YN SHNV Y I WE P ST EKMYV	I MADKQKNGIK RDGVLKGDVN	VNFKIRHNIE MALLIE	DGSVQLADHY 182 GGGHNRCDFK 197
GFP FcoIFP1 FcoIFP2	TNN I T L L K D A TNN I T L SK DN	180 I GDTLVNRI HGVDYFYYNI DNCFDYDI	ELKGIDFKED RFDGVNFPAN RFDGVNFPPK	I GNILGHKL-E GPVMQKKTVK SPVLQKTTIK	YN YN SHNV Y I WE P ST EKMYV WE P SS ENMYV	I MADKQKNGIK RDGVLKGDVN RNGVLKGDIN	VNFKIRHNIE MALLIE MALLLE	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185
GFP FcolFP1 FcolFP2 FcolFP3	TNN ITLLKDA TNN ITLSKDN RND I KMV	180 GDT L VNR I HGVD YF Y YN I DNC FD YD I GDT F YNT V	ELKGIDFKED RFDGVNFPAN RFDGVNFPPK RFDGVNFPPN	GNILGHKL-E GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK	YN YN SHNV Y I WEPSTEKMYV WEPSSENMYV WEPSTEK I YV	I MADKQKNGIK RDGVLKGDVN RNGVLKGDIN RDGVLTGDIT	VN FK I RHN I E MA LL I E MA LL LE MA LL LE	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178
GFP FcolFP1 FcolFP2 FcolFP3 FcolFP4	TNN IT LLKDA TNN IT LSKDN RND I KMV RND I TMV	180 GDTLVNRI HGVDYFYYNI DNCFDYDI GDTFYNTV GDTFYNTV	ELKGIDFKED RFDGVNFPAN RFDGVNFPPK RFDGVNFPPN RFDGVNFPPN	GNILGHKL-E GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK	YN YN SHNVYI WEPSTEKMYV WEPSSENMYV WEPSTEKIYV WEPSTEKIYV	I MADKQKNGIK RDGVLKGDVN RNGVLKGDIN RDGVLTGDIT RDGVLTGDIT	VNFKIRHNIE MALLIE MALLLE MALLLE MALLLK	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197
GFP FcoIFP1 FcoIFP2 FcoIFP3 FcoIFP4 FcoIFP5 FcoIFP6	TNN ITLLKDA TNN ITLSKDN RND I KMV RND I TMV TND I TLD TND I TLE	180 GDTLVNRI HGVDYFYYNI DNCFDYDI GDTFYNTV GDTFYNTV GDCFHYKI GDCFVYKI	ELKGIDFKED RFDGVNFPAN RFDGVNFPPN RFDGVNFPPN RFDGVNFPAN RFDGVNFPAN	GNILGHKL-E GPVMQKRTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK SPVMKKTTQK SPVLQKKTQK	YN YN SHNV YI WEPSTEKMYV WEPSTEKIYV WEPSTEKIYV WEPSTEKIYV WEPSTEKMYV WEPSTEKLYV	I MADKQKNGIK RDGVLKGDVN RNGVLKGDIN RDGVLTGDIT RDGVLTGDIT RDGVLKGDVN RDGVLKGDVN	VNFKIRHNIE MALLIE MALLLE MALLLE MALLLQ MALLLQ	DG SVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFK 88 GGGHYRCDFR 88
GFP FcoIFP1 FcoIFP2 FcoIFP3 FcoIFP4 FcoIFP5 FcoIFP6 FcoIFP7	TNN IT LLKDA TNN IT LSKDN RND I KMV RND ITMV TND IT LD TND IT LE TND IT MH	180 GDTLVNRI HGVDYFYNI DNCFDYDI GDTFYNTV GDTFYNTV GDCFHYKI GDCFYYKI GDVFYYRI	ELKGIDFKED RFDGVNFPAN RFDGVNFPPK RFDGVNFPPN RFDGVNFPAN RFDGVNFPAN RFDGVNFPAK TFNCHNFDPS	I GNILGHKL-E GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK SPVMQKRTLK SPVMKKTTQK SPVLQKKTQK GPVMMKKTLK	YN YN SHNV YI WEPSTEKMYV WEPSSENMYV WEPSTEKIYV WEPSTEKIYV WEPSTEKMYV WEPSCESMYA	I MADKQKNGIK RDGVLKGDVN RNGVLKGDIN RDGVLTGDIT RDGVLKGDVN RDGVLKGDVN VNGDLQGDDF	VNFKIRHNIE MALLIE MALLLE MALLLE MALLLK MALLLE RHLLLE	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFR 88 GGGHYRCDFR 88 GGGHLRCDFK 228
GFP FcoIFP2 FcoIFP3 FcoIFP4 FcoIFP5 FcoIFP6 FcoIFP7 FcoIFP8	TNN ITLLKDA TNN ITLSKDN RND I KMV RND ITMV TND ITLD TND ITLE TND ITMH TND IKLKEGA	180 GDTLVNRI HGVDYFYYNI -DNCFDYDI -GDTFYNTV -GDTFYNTV -GDCFYYKI -GDCFYYKI -GDVFYYRT -DNCFYYEI	ELKGIDFKED RFDGVNFPAN RFDGVNFPPN RFDGVNFPPN RFDGVNFPAN RFDGVNFPAK TFNCHNFDPS RFDGVNFPPN	GNILGHKL-E GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK SPVMKKTTQK SPVLQKKTQK GPVMMKKTLK GPVMQRKTVK	YNYNSHNVYI WEPSTEKMYV WEPSTEKIYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSTEEMFV	I MADKQKNG I K RDGV LKGD VN RDGV LKGD I N RDGV LTGD I T RDGV LKGD VN RDGV LKGD VN VNGD LQGDD F RDGV LKGD LN	VNFKIRHNIE MALLIE MALLLE MALLLK MALLLQ MALLLE RHLLLE MALLLE	GGHYRCDFK 182 GGGHYRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGGHYRCDFR 187 GGGHYRCDFR 88 GGGHYRCDFR 88 GGGHYRCDFK 228 GGGHYRCNFK 177
GFP FcoIFP1 FcoIFP2 FcoIFP3 FcoIFP5 FcoIFP5 FcoIFP6 FcoIFP7 FcoIFP8 AdigFP1	TNN ITLLKDA TNN ITLSKDN RND IKMV TND ITMV TND ITLD TND ITLE TND ITKH TND ITKKEGA SWN IRLE-	1380 I HGVDYFYYNI - DNCFDYDI - GDTFYNTV - GDTFYNTV - GDCFHYKI - GDCFVYKI - GDVFYYRI - DNCFVYEI - RDCFIHKS	ELKG ID FKED RFDG VN FPAN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAK T FNCHN FDPS RFDG VN FPAN I YHG VN FPAD	GNILGHKI-E GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK SPVLQKKTQK GPVMMKKTLK GPVMQRKTVK GPVMQRKTVK GPVMKKTIG	YNYNSHNVYI WEPSTEKMYV WEPSTEKIYV WEPSTEKIYV WEPSTEKMYV WEPSTEKMYV WEPSTEKMYV WEPSTEKMYV WEPSTESMYA WEPSTEMFV WDKAFEKMTV	I MADKQKNGIK RDGVLKGDVN RNGVLKGDIN RDGVLTGDIT RDGVLKGDVN RDGVLKGDVN VNGDLQGDDF RDGVLKGDLN SKDVLRGDVT	VNFKIRHNIE MALLIE MALLLE MALLLE MALLLK MALLLE RHLLLE MALLLE EFLMLE	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 185 GGVHYRCDFR 197 GGGHYRCDFK 188 GGGHYRCDFK 88 GGGHLRCDFK 228 GGGHLRCDFK 228 GGGYHRCNFK 177 GGGYHSCOFH 100
GFP FcoIFP1 FcoIFP2 FcoIFP3 FcoIFP5 FcoIFP5 FcoIFP6 FcoIFP7 FcoIFP7 AdigFP1 AdigFP2	TNN ITLLKDA TNN ITLSKDN RND IKMV TND ITLD TND ITLD TND ITLE TND ITKH TND IKLKEGA SWN IRLE-	1380 I HGVDYFYNI - DNCFDYDI - GDTFYNTV - GDCFYNTV - GDCFYYKI - GDCFYYKI - GDCFYYKI - DNCFYYEI - RDCFIHKS - RDCFIHKS	ELKG IDFKED RFDG VN FPAN RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN TFNC HN FDPS RFDG VN FPAN I YHG VN FPAD	GNILGHKLI GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK SPVLQKKTQK GPVMKKTQK GPVMKKTUK GPVMKKTIG GPVMKKKTIG	YNYNSHNVYI WEPSTEKMYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSTEKLYV WEPSTEMFV WDKAFEKMTV WDKAFEKMTV	MADKQKNGIK RDGVLKGDVN RNGVLKGDIT RDGVLTGDIT RDGVLKGDVN RDGVLKGDVN VNGDLQGDDF RDGVLKGDLN SKDVLRGDVT	VNFKIRHNIE MALLLE MALLLE MALLLK MALLLK MALLLE RHLLLE MALLLE EFLMLE EFLMLE	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFK 88 GGGHLRCDFK 286 GGGHLRCDFK 287 GGGYHSCQFH 100 GGGYHSCQFH 100
GFP FcolFP1 FcolFP2 FcolFP3 FcolFP4 FcolFP5 FcolFP6 FcolFP7 FcolFP7 AdigFP1 AdigFP1	TNN ITLLKDA TNN ITLSKDN RND IKMV TND ITLD TND ITLE TND ITLE TND IKLKEGA SWN IRL - E- SWN IRL - E- SGD ITLSAD-	1380 I HGVDYFYYNI - DNCFDYDI - GDTFYNTV - GDCFHYKI - GDCFYYKI - GDCFYYKI - DNCFVYEI - RDCFIHKS - RDCFIHKS - KKSFEHKS	ELKG IDFKED RFDG VN FPAN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN FFDG VN FPAK TFNCHN FDPS RFDG VN FPAD I YHG VN FPAD I YHG VN FPAD	GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK SPVMQKRTLK SPVMKKTTQK GPVMMKKTLK GPVMMKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG	YNYNSHNVYI WEPSTEKMYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSCESMYA WEPSCESMYA WCKAFEKMTV WDKAFEKMTV	I MADKQKNGIK RDGVLKGDIN RDGVLTGDIT RDGVLTGDIT RDGVLKGDVN RDGVLKGDVN VNGDLQGDDF RDGVLKGDLN SKDVLRGDVT SKDVLRGDVT	VNFKIRHNIE MALLE MALLE MALLE MALLQ MALLQ RHLLQ EFLLQ EFLLQ EFLMLE EFLMLE EFLKK	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 188 GGGHYRCDFR 197 GGGHYRCDFR 88 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 228 GGGHYRCDFK 177 GGGYHSCQFH 100 GGGYHSCQFH 135 DGKRYKCQFH 135
GFP FoolFP1 FoolFP2 FoolFP3 FoolFP4 FoolFP6 FoolFP6 FoolFP7 FoolFP8 AdigFP2 MdigFP2 MdigFP2 MdigFP3	TNN IT LLKDA TNN IT LSKDN RND IKMV RND IT LD TND IT LE TND IT LE TND IT KLKEGA SWN IRLE- SWN IRLE- SGD IT LSAD- SND SS IQ CAD IR VSVE-	1380 I HGVDYFYVNI - DNCFDYDI - GDTFYNTV - GDCFHYKI - GDCFHYKI - GDCFYYKI - DNCFVYEI - RDCFIHKS - RDCFIHKS - KKSFEHKS - GKCFIYNV - DNCFYHES	ELKG IDFKED RFDG VN FPPK RFDG VN FPPK RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAD I YHG VN FPAD V FLG VN FPAD K FLG VN FPAD K FLG UN FPAD	GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK GPVMQKRTLK GPVMMKKTQK GPVMMKKTQ GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKATIG GPVMKKATIG GPVMKKATIG	YNYNSHNVYI WEPSTEKMYV WEPSTEKIYV WEPSTEKIYV WEPSTEKKYV WEPSTEKMYV WEPSTEKMYV WEPSTEKMYV WDKAFEKMTV WDKAFEKMTV WEPSCEKMTP WEPSCEKMTP	I MADKQKNG I K RDGVLKGDVN RNGVLKGDIN RDGVLTGDIT RDGVLKGDVN RDGVLKGDVN VNGDLQGDF RDGVLKGDLN SKDVLRGDVT NGMTLIGDVT RDGMLIGNNF CVGILNGDVT	VNFKIRHNIE MALLIE MALLLE MALLLE MALLLQ MALLLQ MALLLQ EFLLLQ EFLMLE GFLMLE GFLKKE MALKKE MFLLKK	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFK 88 GGGHLRCDFK 88 GGGHLRCDFK 228 GGGHYRCDFK 107 GGGYHSCOFH 105 DGKRYKCQFH 176 GGGYLCEFK 174 DGTRYRCQFH 187
GFP FcoIFP1 FcoIFP2 FcoIFP3 FcoIFP6 FcoIFP6 FcoIFP6 AdigFP1 MdigFP2 MdigFP2 MdigFP2 ScaIFP1	TNN ITLLKDA TNN ITLLKDN RND IKMV RND ITMV TND ITLE TND ITLE TND ITMH TND IKKEGA SWN IRLE- SWN IRLE- SWN IRLE- SWN IRLE- SWD SIQ CAD IRVSVE- TVEMSIN	1380 I - GDTLVNRI HGVDYFYNI - DNCFDYDI - GDTFYNTV - GDCFHYKI - GDCFHYKI - GDCFYYKI - DNCFYYEI - RDCFIHKS - RDCFIHKS - RDCFIHKS - KKSFEHKS - GNCFIYNV - DNCFYHES - SDNNSILE	ELKG IDFKED RFDG VN FPAN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN FDG VN FPAN I YHG VN FPAD I YHG VN FPAD I YHG VN FPAD K I SG LN FPSN K FCG VN FPAD RT I EQV IH-	GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK SPVMQKRTLK GPVMQKKTQK GPVMMKKTLK GPVMQKKTQG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKTTM	YNYNSHNVYI WEPSTEKMYV WEPSTEKIYV WEPSTEKIYV WEPSTEKKYV WEPSTEKMYV WEPSTEKMYV WEPSTEKMYV WDKAFEKMTV WDKAFEKMTV WEPSCEKMTP WEPSCEKMTP	I MADKQKNG I K RDGVLKGDVN RNGVLKGDIN RDGVLTGDIT RDGVLKGDVN RDGVLKGDVN VNGDLQGDF RDGVLKGDLN SKDVLRGDVT NGMTLIGDVT RDGMLIGNNF CVGILNGDVT	VNFKIRHNIE MALLIE MALLLE MALLLE MALLLQ MALLLQ MALLLQ EFLLLQ EFLMLE GFLMLE GFLKKE MALKKE MFLLKK	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFR 88 GGGHVRCDFK 88 GGGHVRCDFK 88 GGGHVRCDFK 100 GGGYHSCQFH 100 GGGYHSCQFH 100 GGGYHSCQFH 135 DGKRYKCQFH 176 GGGHYLCEFK 174 DGTRYRCQFH 180
GFP FcoIFP1 FcoIFP2 FcoIFP5 FcoIFP6 FcoIFP6 FcoIFP7 FcoIFP8 AdigFP2 MdigFP2 MdigFP3 ScaIFP1 ScaIFP2	TNN IT LLKDA TNN IT LLKDA RND I KMV RND IT MV TND IT LE TND IT LE TND IK LK EGA SWN I RL - E- SGD IT LSAD- SND SSIQ CAD I RV SVE- TVEMS LK	1380 I HGVDYFVNI - DNCFDYDI - GDTFYNTV - GDCFYKI - GDCFYKI - GDCFYKI - GDCFYKI - RDCFIHKS - RDCFIHKS - KKSFEHKS - GNCFINV - DNCFYHES - SDNNSILE - GNTLAHKT	ELKG IDFKED RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAD YHG VN FPAD KFLG VN FPAD KFLG VN FPAD KFCG VN FPAD RT I EQV IH NFQG NFPAD	GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK SPVMQKRTLK GPVMQKKTQK GPVMMKKTLK GPVMQKKTQG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKTTM	YNYNSHNVYI WEPSSEKMYV WEPSSEKNYV WEPSTEKIYV WEPSTEKHYV WEPSTEKMYV WEPSCESMYA WEPSTEEMFV WDKAFEKMTV WDKAFEKMTV WEPSCEKMTP WEPNTERLFA WEPSSEKMVP	MADKQKNG IK RDGVLKGDVN RDGVLTGDIT RDGVLTGDIT RDGVLKGDVN RDGVLKGDVN VNGDLQGDDF RDGVLKGDLN SKDVLRGDVT NGMTLIGDVT RDGMLIGNNF CVGILNGDVT	VNFKIRHNIE MALLIE MALLLE MALLLE MALLLQ MALLLQ RHLLLQ EFLMLE GFLMLE GFLMLE GFLKKE MALKLE	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGGHYRCDFK 187 GGGHYRCDFK 197 GGGHYRCDFK 88 GGGHYRCDFK 177 GGGYHSCOFH 178 GGGYHSCOFH 176 GGGYHSCOFH 176 GGGYHSCOFH 176 GGGHYRCOFH 176 GGGHYCOFH 180 DGTRYRCQFH 180
GFP FcoIFP1 FcoIFP2 FcoIFP5 FcoIFP6 FcoIFP6 FcoIFP7 FcoIFP8 AdigFP2 MdigFP2 MdigFP3 ScaIFP1 ScaIFP2	TNN IT LLKDA TNN IT LLKDA RND I KMV RND IT MV TND IT LE TND IT LE TND IK LK EGA SWN I RL - E- SGD IT LSAD- SND SSIQ CAD I RV SVE- TVEMS LK	1380 I HGVDYFYVNI - DNCFDYDI - GDTFYNTV - GDCFYYKI - GDCFYYKI - GDCFYYKI - GDCFYYKI - RDCFIHKS - RDCFIHKS - KKSFEHKS - GNCFYNV - DNCFYHES - SDNNSILE - GNTLAHKT - S	ELKG IDFKED RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAD YHG VN FPAD KFLG VN FPAD KFLG VN FPAD KFLG VN FPAD KFCG VN FPAD RFI EQV IH- NFQG RFP ID	GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK GPVMQKRTLK GPVMQKKTQK GPVMMKKTLK GPVMMKKTLK GPVMKKKTIG GPVMKKTIG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG G	YNYNSHNVYI WEPSSENMYV WEPSSENMYV WEPSTEKIYV WEPSTEKHYV WEPSTEKMYV WEPSCESMYA WEPSTEEMFV WDKAFEKMTV WEPSCEKMTP WEPNTERLFA WEPSSEKMVP	MADKQKNG IK RDGVLKGDVN RNGVLKGDIN RDGVLTGDIT RDGVLKGDVN RDGVLKGDVN VNGDLQGDDF RDGVLKGDLN SKDVLRGDVT NGMTLIGDVT RDGMLIGNNF CVGILNGDVT	VNFKIRHNIE MALLIE MALLLE MALLLE MALLLQ MALLLQ RHLLLQ EFLMLE GFLMLE GFLMLE GFLKKE MALKLE MFLLKK	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFR 88 GGGHVRCDFK 88 GGGHVRCDFK 88 GGGHVRCDFK 100 GGGYHSCQFH 100 GGGYHSCQFH 100 GGGYHSCQFH 135 DGKRYKCQFH 176 GGGHYLCEFK 174 DGTRYRCQFH 180
GFP FcoIFP1 FcoIFP2 FcoIFP3 FcoIFP6 FcoIFP6 FcoIFP6 AdigFP1 MdigFP1 MdigFP2 MdigFP3 ScaIFP2 ScaIFP3 ScaIFP3 ScaIFP3	TNN IT L LKDA TNN IT L SKDN RND I KMV RND I TMV TND IT LE TND IT LE TND I KLKEGA SWN I RL - E SGD I T LSAD SND SSIQ CAD I RV SVE- TVEMS LK TVEMS LK TVEMS LK	1380 I HGVDYFYVNI - DNCFDYDI - GDTFYNTV - GDCFYYKI - GDCFYYKI - GDCFYYKI - GDCFYYKI - RDCFIHKS - RDCFIHKS - KKSFEHKS - GNCFYNV - DNCFYHES - SDNNSILE - GNTLAHKT - S	ELKG IDFKED RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAD YHG VN FPAD KFLG VN FPAD KFLG VN FPAD KFLG VN FPAD KFCG VN FPAD RFI EQV IH- NFQG RFP ID	GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK GPVMQKRTLK GPVMQKKTQK GPVMMKKTLK GPVMMKKTLK GPVMKKKTIG GPVMKKTIG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG G	YNYNSHNVYI WEPSSENMYV WEPSSENMYV WEPSTEKIYV WEPSTEKHYV WEPSTEKMYV WEPSCESMYA WEPSTEEMFV WDKAFEKMTV WEPSCEKMTP WEPNTERLFA WEPSSEKMVP	MADKQKNG IK RDGVLKGDVN RNGVLKGDIN RDGVLTGDIT RDGVLKGDVN RDGVLKGDVN VNGDLQGDDF RDGVLKGDLN SKDVLRGDVT NGMTLIGDVT RDGMLIGNNF CVGILNGDVT	VNFKIRHNIE MALLIE MALLLE MALLLE MALLLQ MALLLQ RHLLLQ EFLMLE GFLMLE GFLMLE GFLKKE MALKLE MFLLKK	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGGHYRCDFK 187 GGGHYRCDFK 197 GGGHYRCDFK 88 GGGHYRCDFK 177 GGGYHSCOFH 178 GGGYHSCOFH 176 GGGYHSCOFH 176 GGGYHSCOFH 176 GGGHYRCOFH 176 GGGHYCOFH 180 DGTRYRCQFH 180
GFP FcoIFP1 FcoIFP2 FcoIFP5 FcoIFP6 FcoIFP5 FcoIFP6 FcoIFP7 FcoIFP8 AdigFP1 AdigFP2 MdigFP1 MdigFP2 MdigFP2 ScaIFP1 ScaIFP1 ScaIFP1	TNN IT L LKDA TNN IT L SKDN RND I KMV RND I TMV TND IT LE TND IT LE TND I KLKEGA SWN I RL - E SGD I T LSAD SND SSIQ CAD I RV SVE- TVEMS LK TVEMS LK TVEMS LK	1380 I HGVDYFYVNI - DNCFDYDI - GDTFYNTV - GDCFYYKI - GDCFYYKI - GDCFYYKI - GDCFYYKI - RDCFIHKS - RDCFIHKS - KKSFEHKS - GNCFYNV - DNCFYHES - SDNNSILE - GNTLAHKT - S	ELKG IDFKED RFDG VN FPAN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAD YHG VN FPAD KFLG VN FPAD KFLG VN FPAD KFLG VN FPAD RFDG VN FPAD RFDG VN FPAD RFDG VN FPAD RFDG VN FPAD	GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK GPVMQKRTLK GPVMQKKTQK GPVMMKKTLK GPVMMKKTLK GPVMKKKTIG GPVMKKTIG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG G	YNYNSHNVYI WEPSTEKMYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSTEKLYV WEPSCESMYA WEPSCESMYA WEPSCEKMTP WEPNTERLFA WEPSSEKMVP	MADKQKNG IK RDGVLKGDVN RNGVLKGDIN RDGVLTGDIT RDGVLKGDVN RDGVLKGDVN VNGDLQGDDF RDGVLKGDLN SKDVLRGDVT NGMTLIGDVT RDGMLIGNNF CVGILNGDVT	VNFKIRHNIE MALLIE MALLLE MALLLE MALLLQ MALLLQ RHLLLQ EFLMLE GFLMLE GFLMLE GFLKKE MALKLE MFLLKK	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGGHYRCDFK 187 GGGHYRCDFK 197 GGGHYRCDFK 88 GGGHYRCDFK 177 GGGYHSCOFH 178 GGGYHSCOFH 176 GGGYHSCOFH 176 GGGYHSCOFH 176 GGGHYRCOFH 176 GGGHYCOFH 180 DGTRYRCQFH 180
GFP FcoIFP1 FcoIFP2 FcoIFP5 FcoIFP6 FcoIFP5 FcoIFP6 FcoIFP7 FcoIFP8 AdigFP1 AdigFP2 MdigFP3 ScaIFP1 ScaIFP2 ScaIFP2 ScaIFP2 ScaIFP3	TNN IT LLKDA TNN IT LSKDN RND IKMV RND IT MV TND IT LE TND IT LE TND IT LKEGA SWN I RLE- SGD IT LSAD SND SS IQ CAD I RV SVE- TVEMSLK TVEMSLK TVEMSLK	1380 I HGVDYFYNI - DNCFDYDI - DNCFDYDI - GDTFYNTV - GDCFYYKI - GDCFYYKI - GDCFYYKI - RDCFIHKS - RDCFIHKS - KKSFEHKS - SNNSILE - GNCFYHES - SNNSILE - GNTLAHKT - S	ELKG IDFKED RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAD YHG VN FPAD KFLG VN FPAD KFLG VN FPAD KFLG VN FPAD KFCG VN FPAD RT I EQV IH- NFQG RFP ID	GPVMQKKTVK SPVLQKKTVK GPVMQKRTLK GPVMQKRTLK GPVMQKRTLK GPVMQKKTQK GPVMKKTQG GPVMKKKTIG GPVMKKTIG GPVMKKTIG GPVMKKTQK GPVMKKTQ GPVMKKTQK GPVMKKTQK GPVMKKTK GPVMKKTQK GPVMKKTK GPVMKKTQK GPVMKKTQK GPVMKKTQK GPVMKKTQK GPVMKKTQK GPVMKKTQK GPVMK GPVMKTQK GPVMK GPV GPV GPV GPV GPV GPV GPV GPV GPV GPV	YNYNSHNVYI WEPSTEKMYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSTEKMYV WEPSTEKMYV WEPSCESMYA WEPSCESMYA WEPSCEKMTP WEPNTERLFA WEPSCEKMYP	ADGVLKGDVN RNGVLKGDIN RNGVLKGDIN RDGVLTGDIT RDGVLKGDVN RDGVLKGDVN VNGDLQGDDF RDGVLKGDLN SKDVLRGDVT RDGVLKGDVT RDGMLIGNNF CVGILNGDVT	VNFKIRHNIE MALLIE MALLLE MALLLE MALLLQ MALLLQ MALLLQ EFLLLQ EFLMLE GFLMLE GFLLKE MALLKE MALLKE	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGGHYRCDFK 187 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 177 GGGYHSCOFH 178 GGGYHSCOFH 176 GGGYHSCOFH 176 GGGYHSCOFH 176 GGGHYRCOFH 176 GGGHYCOFH 180 DGTRYRCOFH 180
GFP FcoIFP1 FcoIFP2 FcoIFP2 FcoIFP6 FcoIFP6 FcoIFP6 AdigFP1 AdigFP2 MdigFP2 MdigFP3 ScaIFP2 ScaIFP3 ScaIF93 Sc	TNN ITLLKDA TNN ITLLKDN RND IKMV RND ITMV TND ITLE TND ITLE TND ITKKEGA SWN IRLE- SWN IRLE- SGD ITLSAD- SND SSIQ CAD IRVSVE- TVEMSLK TVEMSLK TVEMSLK	1380 - GDTLVNR1 HCVDYFYN1 - DNCFDYD1 - GDTFYNTV - GDCFHYK1 - GDCFHYK1 - GDCFHYK1 - GDCFYYK1 - DNCFVYE1 - RDCF1HKS - RDCF1HKS - RDCF1HKS - SDNS1LE - GNTLAHKT - SONS1LE - SONS1LE - SDNS1LE - SD	ELKG IDFKED RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAK TFNC HN FDPS RFDG VN FPAD I YHG VN FPAD KFLG VN FPAD KFLG VN FPAD RT I EQV IH NFQG NFPID NFQG NFPID NFQG NFPID	GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK GPVMQKRTLK SPVMKKTTQK GPVMMKKTTQK GPVMMKKTTG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKETTN GPVMKKETTN GPVMKKETTN GPVMKKETTN GPVMKKETTN GPVMKKETTN GPVMKKETTN GPVMKKETN GPVMKKTUK	YNYNSHNVYI WEPSTEKHYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSTEKLYV WEPSTEKLYV WEPSTEKKYV WEPSCESMYA WEPSCESMYA WEPSCEKMTP WEPSCEKMTP WEPSCEKMTP	MADKQKNG IK RDGVLKGDIN RNGVLKGDIN RDGVLTGDIT RDGVLTGDIT RDGVLKGDVN VNGDLQGDD SKDVLRGDVT SKDVLRGDVT SKDVLRGDVT SKDVLRGDVT NGMTLIGNVF CVGILNGDVT 	VNFKIRHNIE MALLIE MALLIE MALLIE MALLIE RHLLIE RHLLIE RHLLIE EFLILE EFLIK MALLIC EFLKK MALLK MFLLK MFLLK MFLLK MFLLK MFLLK MFLLK	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGGHYRCDFK 187 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 177 GGGYHSCOFH 178 GGGYHSCOFH 176 GGGYHSCOFH 176 GGGYHSCOFH 176 GGGHYRCOFH 176 GGGHYCOFH 180 DGTRYRCOFH 180
GFP FcoIFP1 FcoIFP2 FcoIFP3 FcoIFP6 FcoIFP6 FcoIFP6 AdigFP1 AdigFP2 MdigFP2 MdigFP2 ScaIFP2 ScaIFP3 ScaIF93 Sc	TNN ITLLKDA TNN ITLLKDN RND IKMV RND ITMV TND ITLD TND ITLE TND ITKKEGA SWN IRLE- SGD ITLSAD- SWN IRLE- SGD ITLSAD- SWN IRLE- SGD ITLSAD- TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK	1380 I HGVDYFYVNI - DNCFDYDI - GDTFYNTV - GDCFHYKI - GDCFHYKI - GDCFYYKI - GDCFYYKI - GDCFYYKI - RDCFIHKS - RDCFIHKS - KKSFEHKS - GNCFIHKS - GNCFYHES - SNNNSILE - GNTLAHKT - S	ELKG IDFKED RFDG VN FPPK RFDG VN FPPK RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAD KFCG VN FPAD KFLG VN FPAD KFLG VN FPAD KFLG VN FPAD RFLG VN FPAD STQSALSKDP DHR IE IVSHD DHR II VSHD	GPVMQKKTUK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK GPVMQKRTLK GPVMMKKTQK GPVMKKKTQ GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKATIG GPVMKKATG GPVMKKATG GPVMKKTQ KTQ KQ KQ KQ KQ KQ KQ KQ KQ KQ KQ KQ KQ KQ	YNYNSHNVYI WEPSTEKMYV WEPSTEKIYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKMYV WEPSCESMYA WEPSCESMYA WEPSCEKMTP WEPSCEKMTP WEPSCEKMTP WEPSCEKMTP	MADKQKNG IK RDGVLKGDUN RNGVLKGDIN RDGVLTGDIT RDGVLKGDUN RDGVLKGDVN NGDLQGDDF RDGVLKGDUN SKDVLRGDVT NGMTLIGDVT RDGMLIGNNF CVGILKGDVT 	VNFKIRHNIE MALLIE MALLE MALLE MALLE RHLLE RHLLE RHLLE GFLMLE GFLMLE GFLKKE MALLK MALLK MALLK MALLK MFLLK MALLK MALLK MALLK MALLK MALLK MALLK MALLK MALLK MALLK MALLK MALLK MA	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGGHYRCDFK 187 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 177 GGGYHSCOFH 178 GGGYHSCOFH 176 GGGYHSCOFH 176 GGGYHSCOFH 176 GGGHYRCOFH 176 GGGHYCOFH 180 DGTRYRCOFH 180
GFP FcoIFP1 FcoIFP2 FcoIFP2 FcoIFP6 FcoIFP6 FcoIFP6 FcoIFP6 AdigFP1 AdigFP2 MdigFP3 ScaIFP1 ScaIFP3 ScaIFP3 ScaIFP1 ScaIFP2 FcoIFP3	TNN ITLLKDA TNN ITLLKDN RND IKMV RND ITMV TND ITLE TND ITLE TND ITKKEGA SWN IRL-E- SGD ITLSAD- SND SSIQ CAD IRV SVE- TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TYKAR-K TTYKAK-K	1380 - GDTLVNR1 HGVDYFYN1 - DNCFDYD1 - GDTFYNTV - GDCFHYK1 - GDCFHYK1 - GDCFYYK1 - DNCFYYE1 - NDCFIHKS - RDCFIHKS - RDCFIHKS - RDCFIHKS - RDCFIHKS - SDNSILE - GNTLAHKT - S	ELKG IDFKED RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAN I YHG VN FPAD I YHG VN FPAD KFLG VN FPAD KFLG VN FPAD KFLG VN FPAD STOSALSKDP DHR I E I VSHD DHS IE ILSHD	GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK GPVMQKRTLK GPVMQKKTQK GPVMMKKTTQK GPVMMKKTTQ GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKETTN GPVMKKETTN GPVMKKETTN GPVMKKETTN GPVMKKETTN GPVMKKETTN GPVMKKETTN GPVMKKETTN GPVMKKETTN GPVMKKETTN GPVMKKETTN GPVMKKETK GPVMKKETK GPVMKKETK GPVMKKETK GPVMKKETK GPVMKKETK GPVMKKETK GPVMKKETK GPVMKKETK GPVMKKE KKKD KKKD KKKD KKKK KKKK KKKK KKKKKKKK	YNYNSHNVYI WEPSTEKHYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSTEKLYV WEPSTEKLYV WEPSTEKKYV WDKAFEKMTV WEPSCESMYA WEPSSEKMVP	MADKQKNGIK RDGVLKGDUN RNGVLKGDIT RDGVLTGDIT RDGVLKGDVN RDGVLKGDVN VNGDLQGDDF RDGVLKGDVN SKDVLRGDVT SKDVLRGDVT SKDVLRGDVT NGMTLIGNVF CVGILNGDVT 1 HGMDELYK- SGLPSSRQAK YGLQKEAK- SGLPGNKRQ-	VNFKIRHNIE MALLIE MALLIE MALLIE MALLIE RHLLIE RHLLIE RHLLIE EFLILE EFLILE GFLIKK MALKIE MFLKIE MFLKIE MFLKIE MFLIKK MALKIE MFLIKK MALKIE MFLIKK MALKIE MFLIKK MALKIE MF	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 228 GGGYHSCQFH 100 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGHYLCFK 174 GGGHYLCFK 174 JOTRYRCQFH 180 113
GFP FoolFP1 FoolFP2 FoolFP3 FoolFP5 FoolFP6 FoolFP6 AdigFP2 MdigFP1 ScalFP1 ScalFP1 ScalFP2 ScalFP3 ScalFP1 ScalFP2 ScalFP3 ScalFP2 ScalFP3 ScalFP2 ScalFP3 ScalFP3 ScalFP3 ScalFP4 ScalFP5 Sc	TNN IT LLKDA TNN IT LLKDA TNN IT LKDN RND IKMV RND IT MU TND IT LE TND IT LE TND IT LKEGA SWN IR L SGD IT LSAD CAD I RV SVE TVEMS IK TVEMS IK TYKAKEK TTYKAKEK	1380 I HGVDYFYVNI - DNCFDYDI - GDTFYNTV - GDCFYYKI - GDCFYYKI - GDCFYYKI - GDCFYYKI - RDCFIHKS - RDCFIHKS - RDCFIHKS - KKSFEHKS - GNCFYHES - GNCFYHES - SDNNSILE - GNTLAHKT - S	ELKG IDFKED RFDG VN FPPK RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAD KFCG VN FPAD KFLG VN FPAD KFLG VN FPAD KFLG VN FPAD KFCG VN FPAD STOSALSKOP DHR I E IVSHD DHR I E IVSHD DHS I E ILSHD	GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK GPVMQKRTLK GPVMMKKTQK GPVMKKTQK GPVMKKTTG GPVMKKTIG GPVMKKTIG GPVMKKTIG GPVMKKTTG GPVMKKTTG GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ KQPV-KVKL KQPV-KVKL KQPV-KVKL	YNYNSHNVYI WEPSTEKMYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSTEKLYV WEPSTEKMYV WEPSCESMYA WEPSCESMYA WEPSCEKMTP WEPSCEKMTP WEPSCEKMTP	MADKQKNG IK RDGVLKGDUN RNGVLKGDIT RDGVLTGDIT RDGVLKGDUN RDGVLKGDVN VNGDLQGDF RDGVLKGDUN SKDVLRGDVT NGMTLIGDVT 	VNFKIRHNIE MALLIE MALLE MALLE MALLE RHLLE RHLLE GFLLE GFLMLE GFLKE MALKE MFLLK MF	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 177 GGGYHSCQFH 100 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGYHSCQFH 176 JGGRYLCDFK 174 JI11 JI19 JI119
GFP FcoIFP1 FcoIFP2 FcoIFP2 FcoIFP6 FcoIFP6 FcoIFP6 AdigFP1 AdigFP2 MdigFP2 MdigFP2 ScaIFP2 ScaIFP3 ScaIFP3 ScaIFP3 ScaIFP3 ScaIFP3 FcoIFP5 FcoIFP5 FcoIFP5 FcoIFP3 FcoIFP3 FcoIFP3 FcoIFP5	TNN ITLLKDA TNN ITLLKDN RND IKMV RND ITMV TND ITLE TND ITLE TND ITKKEGA SWN IRLE- SWN IRLE- TVEMSK TVEMSK 	1380 - GDTLVNR1 HGVDYFYN1 - DNCFDYD1 - GDTFYNTV - GDCFHYK1 - GDCFHYK1 - GDCFYYK1 - DNCFVYE1 - NDCFIHKS - RDCFIHKS - RDCFIHKS - SDNSILE - GNTLAHKT - S 200 1 PVLLPDNHFU AVVLPDNHFV GIQLPGYHFV GIQLPGYHFV GIQLPGYHFV - SNLSA - SUNSILE - SDNSILE - SDNS	ELKG IDFKED RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAD KISG VN FPAD KFLG VN FPAD KFLG VN FPAD RFLG VN FPAD RT L GV IH	GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK SPVLQKKTQK SPVUKKTQK GPVMMKKTLK GPVMMKKTTG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG JVKKTIG JVKKKTIG JVKKKTIG JVKKKTIG JV	YNYNSHNVYI WEPSTEKHYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSTEKLYV WEPSTEKLYV WDKAFEKMTV WDKAFEKMTV WEPSCESKMYP WEPSCEKMTP WEPSCEKMTP WEPSCEKMTP WEPSCEKMTP WEPSCEKMTP HEFAKAR- YEHAEAH YEHAEAH YEHAEAH	MADKQKNGIK RDGVLKGDIN RDGVLTGDIT RDGVLTGDIT RDGVLTGDIT RDGVLKGDVN RDGVLKGDVN NGDVLKGDVN SKDVLRGDVT SKDVLRGDVT SKDVLRGDVT NGMTLIGDVT RDGMLIGNNF CVGILGND 1 HGMDELYK- SGLPSSRQAK SGLPGNKRQ- SGLPGNKRQ- SGLPGNKRQ- SGLPGNKRQ-	VNFKIRHNIE MALLIE MALLE MALLE MALLE RHLLE RHLLE RHLLE EFLLE EFLLE RHLLE EFLLK MALLK MFLLK MALLK MALLK MALLK MALLK MALLK MFLLK MFLLK MFLLK MFLLK MFLLK MFLLK MFLLK MFLLK MALLK MALLK MALLK MALLK MALLK MF	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 228 GGGYHSCQFH 100 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGHYLCFK 174 GGGHYLCFK 174 JOTRYRCQFH 180 113
GFP FoolFP1 FoolFP2 FoolFP3 FoolFP5 FoolFP6 FoolFP6 FoolFP6 AdigFP1 AdigFP2 MdigFP2 MdigFP2 ScalFP2 ScalFP3 ScalFP1 ScalFP2 ScalFP3 ScalFP1 ScalFP2 FoolFP1 FoolFP1 FoolFP1 FoolFP3 FoolFP3 FoolFP3 FoolFP3 FoolFP4 FoolFP3 FoolFP3	TNN IT LLKDA TNN IT LLKDA RND IKMV RND ITMU TND ITLE TND ITLE TND IKLKEGA SWN IRLE- SGD ITLSAD CAD IRVSVE- TVEMSIK TVEMSIK TVEMSIK TVEMSKK TVEMSKK TYKAK-K TTYKAKEK TYKAK-K TYKAK-K	1380 - GDTLVNR1 HGVDYYVN1 - DNCFDYD1 - GDTFYNTV - GDCFYNTV - GDCFYYK1 - GDCFYYK1 - GDCFYYK1 - NDCFYYE1 - NDCFIHKS - RDCF1HKS - RDCF1HKS - RDCF1HKS - RDCF1HKS - SDNN51LE - GNTLAHKT - SONN51LE - GNTLAHKT - SONN51LE - QNLPDHFY 200 0 0 0 0 0 0 0 0 0 0 0 0	ELKG IDFKED RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN I YHG VN FPAD I YHG VN FPAD KFLG VN FPAD KFLG VN FPAD KFLG VN FPAD STOSALSKDP DHR IE I VSHD DHS IE ILSHD DHS IE ILSHD DHR IE ILSHD DHR IE ILSHD DHR IE ILSHD	GPVMQKKTVK SPVLQKKTVK SPVLQKKTVK GPVMQKRTLK GPVMQKRTLK GPVMQKKTQK SPVLQKKTQK GPVMKKTLG GPVMKKKTIG GPVMKKKTIG GPVMKKTTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG SPVN-KVKL KDYN-KVKL KDYN-KVKL RDYN-VVKL	YNYNSHNVYI WEPSTEKMYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSTEKLYV WEPSCEKMTV WDKAFEKMTV WDKAFEKMTV WEPSCEKMTP WEPSCEKMTP 	MADKQKNG IK RDGVLKGDIN RNGVLKGDIN RDGVLTGDIT RDGVLKGDVN RDGVLKGDVN RDGVLKGDVN SKDVLRGDVT SKDVLRGDVT SKDVLRGDVT SKDVLRGDVT 	VNFKIRHNIE MALLIE MALLLE MALLLE MALLLE MALLLQ MALLLQ EFLLLQ EFLLLQ EFLLLQ EFLLLQ EFLLLQ EFLLQ MALLQ ALLQ ALLQ ALLQ ALLQ A	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 228 GGGYHSCQFH 100 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGHYLCFK 174 GGGHYLCFK 174 JOTRYRCQFH 180 113
GFP FcoIFP1 FcoIFP3 FcoIFP3 FcoIFP5 FcoIFP6 FcoIFP6 AdigFP1 AdigFP2 MdigFP2 MdigFP2 MdigFP3 ScaIFP2 ScaIFP3 ScaIF9 ScaIF9 ScaIF9 ScaIF9 ScaIF9 ScaIF9 ScaIF9 ScaIF9 ScaIF9 ScaIF9 ScaIF9 ScaIF9 ScaIF3	TNN IT LLKDA TNN IT LLKDA TNN IT LKHD RND IKMV RND IT MU TND IT LE TND IT LE TND IT LKEGA SWN IR L SGD IT LSAD CAD IR VY VE- TVEMS IK TVEMS IK TYKAK K TT YKAK K TT YKAK K TT YKAK K TT YKAK K TT YKAK K	1380 - GDTLVNR1 HGVDYFYN1 - DNCFDYD1 - GDTFYNTV - GDCFHYK1 - GDCFHYK1 - GDCFYYK1 - DNCFVYE1 - NDCFIHKS - RDCFIHKS - RDCFIHKS - SDNSILE - GNTLAHKT - S 200 1 PVLLPDNHFU AVVLPDNHFV GIQLPGYHFV GIQLPGYHFV GIQLPGYHFV - SNLSA - SUNSILE - SDNSILE - SDNS	ELKG IDFKED RFDG VN FPPK RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAD KFCG VN FPAD KFCG VN FPAD KFCG VN FPAD KFCG VN FPAD KFCG VN FPAD STQ SALSKDP DHR IE IVSHD DHR IE ILSHD DHR IE ILSHD DHR IE ILSHD DHR IE ILSHD DHR IE ILSHD	GNILGHKLI- GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK GPVMQKKTQK GPVMKKTQ GPVMKKTG GPVMKKTIG CON CON CON CON CON CON CON CON CON CON	YNYNSHNVYI WEPSTEKHYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSTEKLYV WEPSTEKKYV WEPSTEKKYV WEPSCESMYA WEPSCEKMTP WEPSCEKMTP WEPSCEKMTP WEPSCEKMTP WEPSCEKMTP HEPSCEKMYP	MADKQKNG IK RDGVLKGDUN RNGVLKGDIN RDGVLTGDIT RDGVLKGDUN RDGVLKGDVN VNGDLQGDDF RDGVLKGDUN SKDVLRGDVT NGMTLIGDVT RDGMLIGNNF CVGILKGDVT 1 HGMDELYK- SGLPSSRQAK SGLPSSRQAK SGLPGNKRQ- SGLPGNKRQ- SGLPRQAK- SGLSKKAK-	VNFKIRHNIE MALLIE MALLE MALLE MALLE RHLLE RHLLE RHLLE GFLMLE GFLKE GFLKKE MALLK MFLK MFLLK MF	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 228 GGGYHSCQFH 100 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGHYLCFK 174 GGGHYLCFK 174 JOTRYRCQFH 180 113
GFP FoolFP1 FoolFP2 FoolFP3 FoolFP5 FoolFP6 FoolFP6 FoolFP7 AdigFP1 AdigFP2 MdigFP2 MdigFP2 MdigFP2 ScalFP1 ScalFP1 ScalFP1 ScalFP1 ScalFP2 ScalFP2 ScalFP2 FoolFP2 FoolFP2 FoolFP2 FoolFP2 FoolFP2 FoolFP3 FoolFP4 FoolFP2 FoolFP4 FoolFP5 FoolFP4 FoolFP5 FoolFP4 FoolFP5 FoolFP4 FoolFP5 FoolFP4 FoolFP5 FoolFP5 FoolFP5 FoolFP5 FoolFP5 FoolF95 Fo	TNN IT LLKDA TNN IT LLKDA TNN IT LKDN RND IKMV RND IT MU TND IT LE TND IT LE TND IT LKEGA SWN IR L SGD IT LSAD SND SS IQ CAD I RV SVE- TVEMSLK TVEMSLK TVEMSLK TVEMSLK TYEMSLK TYKAK TTYKAK TTYKAK TTYKAK TTYKAK TTYKAK TTYKAK TTYKAK TTYKAK TTYKAK TTYKAK STYKPE K	1380 I HGVDYFYNI - GDTLVNRI - DNCFDYDI - GDTFYNTV - GDCFYYKI - GDCFYYKI - GDCFYYKI - NDCFYYEI - NDCFIHKS - KKSFEHKS - SNCFYHES - SDNNSILE - GNCFYHES - SDNNSILE - GNLAHKT - S	ELKG IDFKED RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAD KFCG VN FPAD HT I EV IH- STOSALSKOP DHR I EI VSHD DHR I EI LSHD DHR I EI LSHD	GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK GPVMQKRTLK GPVMQKKTQK GPVMKKTQ GPVMKKTTG GPVMKKTIG GPVMKKTTG GPVMKKTTG GPVMKKTTG GPVMKKTTG GPVMKKTTG GPVMKKTG GPVMKKTQ GPVK GPVMKKTQ GPV GPV GPV GPV GPV GPV GPV GPV GPV GPV	YNYNSHNVYI WEPSTEKHYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSTEKLYV WEPSTEKLYV WEPSTEKLYV WEPSCESMYA WEPSCESMYA WEPSCEKMTP WEPSCEKMTP WEPSCEKMTP WEPSCEKMTP WEPSCEKMTP WEPSCEKMYP LEFVTAAGIT SEHAEAH	MADKQKNG IK RDGVLKGDUN RNGVLKGDIN RDGVLTGDIT RDGVLKGDUN RDGVLKGDVN NGDLQGDF RDGVLKGDUN SKDVLRGDVT NGMLIGDVT CVGILNGDVT HGMDELYK- SGLPSRQAK YGLQKEAK- SGLPGNKRQ- SGLPRQAK- SGLPRQAK- SGLPRQAK- YSPIPEIEPM SGLPRQAK-	VNFKIRHNIE MALLIE MALLE MALLE MALLE RHLLE RHLLE GFLLE GFLLE GFLLE GFLLE MALLE GFLLE MALLE CHLLE CHLLE CH	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 228 GGGYHSCQFH 100 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGHYLCFK 174 GGGHYLCFK 174 JOTRYRCQFH 180 113
GFP FcoIFP1 FcoIFP2 FcoIFP2 FcoIFP6 FcoIFP6 FcoIFP6 AdigFP1 MdigFP2 MdigFP2 MdigFP2 ScaIFP2 ScaIFP3 ScaIFP3 ScaIFP3 ScaIFP3 FcoIFP6 FcoIFP7 FcoIFP5 FcoIFP6 FcoIFP7 FcoIFP5 FcoIFP6 FcoIFP7 FcoIFP5 FcoIFP6 FcoIFP7 FcoIFP8 FcoIFP6 FcoIFP7 FcoIFP8 FcoIFP8 AdigFP1 AdigFP2	TNN ITLLKDA TNN ITLLKDN RND IKMV RND ITMV TND ITLD TND ITLE TND ITKEKGA SWN IRLE- SGD ITLSAD- SND SSIQ CAD IRVSVE- TVEMSLK TVEMSKS TVEMSKS TVEMSKS TYKAK-K TTYKAK-K TTYKAK-K TTYKAK-K TTYKAK-K STYRAK-K STYRAK-K STYKAK-K STYKAK-K	1380 - GDTLVNRI HGVDYFYNI - DNCFDYDI - GDTFYNTV - GDCFHYKI - GDCFHYKI - GDCFYYKI - NDCFVYEI - NDCFVYEI - NDCFIHKS - RDCFIHKS - RDCFIHKS - GNCFIYNV - DNCFYHES - SDNNSILE - GNTLAHKT - S	ELKG IDFKED RFDG VN FPPK RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN KFCG VN FPAD KISG UN FPAD KIS	GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK SPVLQKKTQK SPVLQKKTQK GPVMMKKTQ GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKATIG GPVMKKATG GPVMKKATG GPVMKKTQ MKKD MKKD MKKTQ CPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ GPVMKKTQ CPV SQ SQ SQ SQ SQ SQ SQ SQ SQ SQ SQ SQ SQ	YNYNSHNVYI WEPSTEKHYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSTEKLYV WEPSTEKMYV WEPSCESMYA WEPSCESMYA WEPSCEKMTP	MADKQKNG IK RDGVLKGDUN RNGVLKGDIN RDGVLTGDIT RDGVLTGDIT RDGVLKGDVN RDGVLKGDVN SKDVLRGDVT SKDVLRGDVT SKDVLRGDVT SKDVLRGDVT CVGILKGDVN 1000	VNFKIRHNIE MALLIE MALLE MALLE MALLE RH	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 228 GGGYHSCQFH 100 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGHYLCFK 174 GGGHYLCFK 174 JOTRYRCQFH 180 113
GFP FoolFP1 FoolFP2 FoolFP3 FoolFP5 FoolFP6 FoolFP6 FoolFP6 AdigFP1 MdigFP2 MdigFP2 ScalFP2 ScalFP3 ScalFP1 ScalFP2 ScalFP3 ScalFP1 ScalFP2 FoolFP1 FoolFP1 FoolFP3 FoolFP4 FoolFP3 FoolFP4 FoolFP4 FoolFP5 FoolFP4 FoolFP5 FoolFP4 FoolFP5 FoolFP4 FoolFP5 FoolFP4 FoolFP5 FoolFP4 FoolFP5 FoolFP4 FoolFP5 FoolFP4 FoolFP5 FoolFP4 FoolFP5 FoolFP4 FoolFP5 FoolFP4 FoolFP5 FoolFP4 FoolFP5 Fo	TNN ITLLKDA TNN ITLLKDA TNN ITLSKDN RND IKMV RND ITMU TND ITLE TND ITLE TND ITLE SND SIQ CAD IRVSVE- TVEMSIK TVEMSIK TVEMSIK TVEMSIK TVEMSIK TVEMSIK TTYKAK TTYKAK TTYKAK TTYKAKEK TTYKAKEK TTYKAKEK TTYKAKEK TTYKAKEK TTYKAKEK TTYKAKEK TTYKAKEK TTYKAK STYKPEK STYKPEK	1380 - GDTLVNR1 HGVDYYN1 - DNCFDYD1 - GDTFYNTV - GDCFYNTV - GDCFYYK1 - GDCFYYK1 - GDCFYYK1 - NCCFYK5 - RDCF1HKS - RDCF1HKS - RDCF1HKS - RDCF1HKS - GNCFYHES - GNCFYHES - GNTLAHKT - GNTLAHKT - GNTLAHKT - S	ELKG IDFKED RFDG VN FPAN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAD KFLG VN FPAD KFL	GPVMQKKTVK SPVLQKKTVK GPVMQKRTLK GPVMQKRTLK GPVMQKRTLK GPVMQKKTQK GPVMKKTQ GPVMKKTTQ GPVMKKTTG GPVMKKTTG GPVMKKTTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVMKKTG GPVK KTG DVN-KVKL KDYN-KVKL KDYN-KVKL KDYN-KVKL CDYN-KVKL CDYN-KVKL CSAGFTVKL GSMQTWQLT-	YNYNSHNVYI WEPSTEKMYV WEPSTEKIYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSCEKMTV WEPSCEKMTP WEPSCEKMTP WEPSCEKMTP WEPSCEKMTP WEPSSEKMVP WEPSSEKMVP WEPSSEKMVP WEPSSEKMVP WEPSSEKMYP WEPNTERLFA WEPSSEKMYP CONSTRUCTION WEPSSEKMVP WEPSSEKMVP CONSTRUCTION WEPSSEKMVP WEPSSEKMVP WEPSSEKMVP CONSTRUCTION WEPSSEKMVP WEPSSEKMVP CONSTRUCTION WEPSSEKMVP CONSTRUCTION WEPSSEKMVP WEPSSEKMVP CONSTRUCTION WEPSSEKMVP WEPSSEKMVP WEPSSEKMVP WEPSSEKMVP CONSTRUCTION WEPSSEKMVP WEPSSEKMVP CONSTRUCTION WEPSSEKMVP WEPSSEKMVP WEPSSEKMVP WEPSSEKMVP CONSTRUCTION WEPSSEKMVP	MADKQKNG IK RDGVLKGDUN RNGVLKGDIN RDGVLTGDIT RDGVLKGDUN RDGVLKGDUN RDGVLKGDVN SKDVLRGDVT SKDVLRGDVT SKDVLRGDVT NGMLIGNNF CVGILNGDVT I HGMDELYK- SGLPSRQAK YGLQKEAK- YSPIPEIEPM SGLPRQAK- YSPIPEIEN	VNFKIRHNIE MALLIE MALLLE MALLLE MALLLQ MALLLQ MALLLQ EFLLLQ EFLLLQ EFLLKE MALLC Chrono 138 280 227 151 139 227	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 228 GGGYHSCQFH 100 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGHYLCFK 174 GGGHYLCFK 174 JOTRYRCQFH 180 113
GFP FcolFP1 FcolFP2 FcolFP3 FcolFP5 FcolFP6 FcolFP6 FcolFP7 FcolFP8 AdigFP1 MdigFP2 MdigFP2 ScalFP2 ScalFP2 ScalFP3 ScalFP2 ScalFP3 ScalFP1 ScalFP2 ScalFP3 FcolFP4 FcolFP5 FcolFP5 FcolFP5 FcolFP5 FcolFP5 FcolFP5 FcolFP5 FcolFP5 FcolFP5 FcolFP6 FcolFP5 Fc	TNN IT LLKDA TNN IT LLKDA TNN IT LKDN RND IKMV RND IT MV TND IT LE TND IT LE TND IT LKEGA SWN IR L SGD IT LSAD CAD IR VY VE- TVEMS IN TVEMS IK TVEMS IK TVEMS IK TVEMS IK TVEMS IK TYKAK TTYKAK TTYKAK TTYKAK STYRAE STYKAE	1380 - GDTLVNRI HGVDYFYNI - GDTFYNTV - GDTFYNTV - GDCFHYKI - GDCFHYKI - GDCFYYKI - DNCFVYEI - NDCFIHKS - RDCFIHKS - RDCFIHKS - RDCFIHKS - RDCFIHKS - SDNSILE - GNTLAHKT - S 200 - 200 - 200	ELKG IDFKED RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAN KFCG VN FPAD KFCG VN FPAD KFCG VN FPAD KFCG VN FPAD KFCG VN FPAD KFCG VN FPAD KFCG VN FPAD HT IE VIH- NFQG RFPID STOSALSKOP DHR IE IVSHD DHR IE IVSHD DHR IE ILSHD DHR IE ILSHD CHR IE IL	GNILGHKIL GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK GPVMQKKTQK GPVMKKTQG GPVMKKTG GPVMKKTIG GPVMKKTIG GPVMKKTTG GPVMKKTTG GPVMKKTTG GPVMKKTQ GPVMKKTQ COMPACT	YNYNSHNVYI WEPSTEKHYV WEPSTEKIYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSTEKMYV WEPSTEKMTV WEPSCESMYA WEPSCESMYA WEPSCEKMTP	MADKQKNG IK RDGVLKGDUN RNGVLKGDIN RDGVLTGDIT RDGVLKGDUN RDGVLKGDVN VNGDLQGDDF RDGVLKGDUN SKDVLRGDVT NGMTLIGDVT NGMTLIGNF CVGILKGDVN 1000	VNFKIRHNIE MALLIE MALLE MALLE MALLE RHLLE RHLLE GFLLE GFLLE GFLLE GFLLE GFLLE MF	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 228 GGGYHSCQFH 100 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGHYLCFK 174 GGGHYLCFK 174 JOTRYRCQFH 180 113
GFP FcolFP1 FcolFP2 FcolFP3 FcolFP5 FcolFP6 FcolFP6 FcolFP7 FcolFP6 MdigFP1 MdigFP2 MdigFP2 MdigFP2 ScalFP3 ScalFP1 ScalFP3 ScalFP1 ScalFP2 FcolFP6 FcolFP7 FcolFP6 FcolFP7 FcolFP8 FcolFP5 FcolFP7 FcolFP8 AdjgFP2 MdigFP1 MdigFP2 MdigFP1 MdigFP2 FcolFP8 FcolFP6 FcolFP6 FcolFP7 FcolFP6 FcolFP7 FcolFP8 FcolFP6 FcolFP7 FcolFP8 FcolFP6 FcolFP7 FcolFP8 FcolFP6 FcolFP7 FcolFP8 FcolFP6 FcolFP7 FcolFP8 FcolFP6 FcolFP7 FcolFP8 FcolFP6 FcolFP7 FcolFP8 FcolFP6 FcolFP7 FcolFP8 FcolFP6 FcolFP7 FcolFP8 FcolFP6 FcolFP7 FcolFP8 FcolFP6 Fc	TNN IT LLKDA TNN IT LLKDA TNN IT LSKDN RND I KMV RND I TMU TND I TLE TND I TLE TND I TLE SGD I TLSAD CAD I RV SVE- TVEMS IN CAD I RV SVE- TVEMS IK TVEMS IK TYKAK TTYKAK TTYKAK TTYKAK TYKAK TYKAK TYKAK	1380 - GDTLVNR1 HGVDYFYN1 - DNCFDYD1 - GDTFYNTV - GDCFHYK1 - GDCFHYK1 - GDCFYYR1 - DNCFYYE1 - NDCFIHKS - RDCFIHKS - RDCFIHKS - RDCFIHKS - RDCFIHKS - SDNSILE - GNTLAHKT - S	ELKG IDFKED RFDG VN FPAN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAN I YHG VN FPAD I YHG VN FPAD K FLG VN FPAD K FLG VN FPAD K FLG VN FPAD K FLG VN FPAD T I EQ VI FPAD STQSALSKDP DHR I E I VSHD DHR I E I VSHD DHS I E I LSHD DHS I E I LSHD EHS I VST SHN SHS SHS SHS SHS SHS SHS SHS SHS SHS	GPVMQKKTUK GPVMQKRTLK GPVMQKRTLK GPVMQKRTLK GPVMQKRTLK GPVMQKKTQK GPVMKKTTQC GPVMKKTIG GPVMKKTIG GPVMKKTIG GPVMKKTIG GPVMKKTG GPVMKKTQC GPVMKKC GPVMKKTQC GPVMKKTQC GPVMKKTQC GPVMKKTQC GPVMKKTQC GPVMKKTQC GPVMKKTQC GPVMKKTQC GPVMKKTQC GPVMKKTQC GPVMKKTQC GPVMKKTQC GPVMKKQC GPVMKKQC GPVMKKQC GPVMKKQC GPVMKKQC GPVMKKQC GPVMKKQC GPVMKKQC GPVMKKQC GPVMKC GPVMC GPVMKC GPVMC GPVMKC GPVMC GPVMC GPVMC GPVMC GPVMC GPVMC GPVMC GPVMC GPVM	YNYNSHNVYI WEPSTEKMYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSTEKLYV WDKAFEKMTV WDKAFEKMTV WDKAFEKMTV WEPSCEKMTP WEPSCEKMTP WEPSCEKMTP WEPSSEKMVP 	MADKQKNG IK RDGVLKGDVN RDGVLKGDIN RDGVLKGDIN RDGVLKGDVN RDGVLKGDVN RDGVLKGDVN SKDVLRGDVT SKDVLRGDVT SKDVLRGDVT SKDVLRGDVT SKDVLRGDVT 	VNFKIRHNIE MALLIE MALLLE MALLLE MALLLE MALLLE MALLLE MALLE MALLE MALLE EFLLKE MA	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 228 GGGYHSCQFH 100 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGHYLCFK 174 GGGHYLCFK 174 JOTRYRCQFH 180 113
GFP FcolFP1 FcolFP2 FcolFP3 FcolFP5 FcolFP6 FcolFP6 FcolFP7 FcolFP6 MdigFP1 MdigFP2 MdigFP2 MdigFP2 ScalFP3 ScalFP1 ScalFP3 ScalFP1 ScalFP2 FcolFP6 FcolFP7 FcolFP6 FcolFP7 FcolFP8 FcolFP5 FcolFP7 FcolFP8 AdjgFP2 MdigFP1 MdigFP2 MdigFP1 MdigFP2 FcolFP8 FcolFP6 FcolFP6 FcolFP7 FcolFP6 FcolFP7 FcolFP8 FcolFP6 FcolFP7 FcolFP8 FcolFP6 FcolFP7 FcolFP8 FcolFP6 FcolFP7 FcolFP8 FcolFP6 FcolFP7 FcolFP8 FcolFP6 FcolFP7 FcolFP8 FcolFP6 FcolFP7 FcolFP8 FcolFP6 FcolFP7 FcolFP8 FcolFP6 FcolFP7 FcolFP8 FcolFP6 FcolFP7 FcolFP8 FcolFP6 Fc	TNN ITLLKDA TNN ITLLKDN RND IKMV RND ITMV TND ITLD TND ITLE TND ITLE SND ISI SND ISI SNN IRL SGD ITLSAD- SNN IRL SSWN IRL SSWN IRL SSWN IRL TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TYKAK-K TTYKAK-K TTYKAK-K TTYKAK-K TTYKAK-K STYRAK	1380 - GDTLVNRI HGVDYFYNI - DNCFDYDI - GDTFYNTV - GDCFHYKI - GDCFHYKI - GDCFYYKI - RDCFIHKS - RDCFIHKS - RDCFIHKS - RDCFIHKS - RDCFIHKS - SDNSILE - GNCFYHES - SDNSILE - GNTLAHKT - DNCFYHES - SDNSILE - GNTLAHKT - DNCFYHES - SDNSILE - GNTLAHKT - S	ELKG IDFKED RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAN KISG VN FPAD KISG UN FPAD KIS	GNILGHKI GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK GPVMQKKTQK SPVLQKKTQK GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKATIG GPVMKKATIG GPVMKKATIG GPVMKKTUK MDYN-HVKL KDYN-KVKL KDYN-KVKL KDYN-KVKL KDYN-KVKL GSMQTWQLT- KDYL-SVEQ GRDQKWQLV-	YNYNSHNVYI WEPSTEKHYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSTEKMYV WEPSTEKMYV WEPSCESMYA WEPSCESMYA WEPSCEKMTP	MADKQKNG IK RDGVLKGDUN RNGVLKGDIN RDGVLTGDIT RDGVLKGDUN RDGVLKGDVN RDGVLKGDVN SKDVLRGDVT NGMTLIGDVT RDGMLIGNVT CVGILKGDVT 1000000000000000000000000000000000000	VNFKIRHNIE MALLIE MALLE MALLE MALLE RH	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 228 GGGYHSCQFH 100 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGHYLCFK 174 GGGHYLCFK 174 JOTRYRCQFH 180 113
GFP FcoIFP1 FcoIFP2 FcoIFP3 FcoIFP5 FcoIFP6 FcoIFP6 AdigFP1 AdigFP2 MdigFP2 ScaIFP2 ScaIFP2 ScaIFP3 ScaIFP2 ScaIFP3 ScaIFP1 ScaIFP2 ScaIFP1 ScaIFP2 ScaIFP1 ScaIFP2 ScaIFP3 ScaIFP1 ScaIFP1 FcoIFP5 FcoIFP6 FcoIFP7 FcoIFP8 FcoIFP5 FcoIFP6 FcoIFP7 FcoIFP8 ScaIFP1 ScaIF1	TNN ITLLKDA TNN ITLLKDN RND IKMV RND ITMV TND ITLD TND ITLE TND ITLE SND ISI SND ISI SNN IRL SGD ITLSAD- SNN IRL SGD ITLSAD- SNN IRL TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TYKAK-K TTYKAK-K TTYKAK-K TTYKAK-K TTYKAK-K STYRAK	1380 - GDTLVNR1 HGVDYFYN1 - DNCFDYD1 - GDTFYNTV - GDCFHYK1 - GDCFHYK1 - GDCFYYR1 - DNCFYYE1 - NDCFIHKS - RDCF1HKS - RDCF1HKS - RDCF1HKS - RDCF1HKS - SDNS1LE - GNTLAHKT - S	ELKG IDFKED RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAN KISG VN FPAD KISG UN FPAD KIS	GNILGHKI GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK GPVMQKKTQK SPVLQKKTQK GPVMKKKTIG GPVMKKKTIG GPVMKKKTIG GPVMKKATIG GPVMKKATIG GPVMKKATIG GPVMKKTUK MDYN-HVKL KDYN-KVKL KDYN-KVKL KDYN-KVKL KDYN-KVKL GSMQTWQLT- KDYL-SVEQ GRDQKWQLV-	YNYNSHNVYI WEPSTEKHYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSTEKMYV WEPSTEKMYV WEPSCESMYA WEPSCESMYA WEPSCEKMTP	MADKQKNG IK RDGVLKGDUN RNGVLKGDIN RDGVLTGDIT RDGVLKGDUN RDGVLKGDVN RDGVLKGDVN SKDVLRGDVT NGMTLIGDVT RDGMLIGNVT CVGILKGDVT 300 1 HGMDELYK- SGLPSSRQAK- SGLPGNKRQ	VNFKIRHNIE MALLIE MALLE MALLE MALLE RH	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 228 GGGYHSCQFH 100 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGHYLCFK 174 GGGHYLCFK 174 JOTRYRCQFH 180 113
GFP FcoIFP1 FcoIFP2 FcoIFP3 FcoIFP5 FcoIFP6 FcoIFP6 FcoIFP7 FcoIFP8 AdigFP1 MdigFP2 MdigFP2 ScaIFP3 ScaIFP3 FcoIFP5 FcoIFF5 FcoIFP5 FcoIFP5 FcoIFF5 FcoIF55 FcoIFF5 FcoIF55 FcoIFF5 FcoIF55 Fc	TNN ITLLKDA TNN ITLLKDN RND IKMV RND ITMV TND ITLD TND ITLE TND ITLE SND ISI SND ISI SNN IRL SGD ITLSAD- SNN IRL SGD ITLSAD- SNN IRL TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TVEMSLK TYKAK-K TTYKAK-K TTYKAK-K TTYKAK-K TTYKAK-K STYRAK	1380 - GDTLVNRI HGVDYFYNI - DNCFDYDI - GDTFYNTV - GDCFHYKI - GDCFHYKI - GDCFYYKI - NDCFVYEI - NDCFVYEI - NDCFYHES - KKSFEHKS - SDNSILE - GNTLAHKT - S	ELKG IDFKED RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPPN RFDG VN FPAN RFDG VN FPAN RFDG VN FPAN KISG VN FPAD KISG UN FPAD KIS	GNILGHKLI GPVMQKKTVK SPVLQKTTIK GPVMQKRTLK GPVMQKRTLK GPVMQKKTQK GPVMKKTTQ GPVMKKKTIG GPVMKKKTIG GPVMKKTTM GPVMKKTTM GPVMKKTTM GPVMKKTTM GPVMKKTTM GPVMKKTTM GPVMKKTTM GPVMKKTTM GPVMKKTTM COMMAN C	YNYNSHNVYI WEPSTEKHYV WEPSTEKIYV WEPSTEKIYV WEPSTEKLYV WEPSTEKLYV WEPSTEKMYV WEPSTEKMYV WEPSCESMYA WEPSCESMYA WEPSCEKMTP	MADKQKNG IK RDGVLKGDUN RNGVLKGDIN RDGVLTGDIT RDGVLKGDUN RDGVLKGDVN RDGVLKGDVN SKDVLRGDVT NGMTLIGDVT RDGMLIGNVT CVGILKGDVT 300 1 HGMDELYK- SGLPSSRQAK- SGLPGNKRQ	VNFKIRHNIE MALLIE MALLE MALLE MALLE RH	DGSVQLADHY 182 GGGHNRCDFK 197 GGGHYRCDFK 185 GGVHYRCDFR 178 GGVHYRCDFR 197 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 88 GGGHYRCDFK 228 GGGYHSCQFH 100 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGYHSCQFH 176 GGGHYLCFK 174 GGGHYLCFK 174 JOTRYRCQFH 180 113

Fig. 1. Alignment of the fluorescent protein (FP) peptide sequences from *Acropora digitifera, Favites colemani, Montipora digitata* and *Seriatopora caliendrum* with *Aequorea victoria* GFP. Positions necessary for FP activity, including the putative chromophore triad (green box), glycine residues for structural integrity (grey boxes), and residues needed for chromophore maturation (blue boxes), are shown. A histogram indicating the conservation at each position is shown below the sequence alignment. See Supplementary Material File S1 for the peptide alignment.

protein FP2 S1603 2 (LC177541.1) identified from the same species (Takahashi-Kariyazono *et al.*, 2016), and to cyan FPs from *A. nobilis* (AY646071, AY646072: Alieva *et al.*, 2008). The sequences of AdigFP1 and AdigFP2 were also similar to amFP486 (XM_029357520.1), a cyan fluorescent protein from *A. millepora* (Table 1). Similar to its closest neighbouring cyan FPs in clade C2, AdigFP2 has a QYG chromophore (Figure 1). Altogether, these data suggest that AdigFP1 and AdigFP2 encode cyan fluorescent proteins.

Three transcripts potentially encoding full-length FPs were found in the transcriptome of *M. digitata*. MdigFP1 and MdigFP2 grouped into clade C3 (Figure 2). MdigFP1 was most closely affiliated with cyan FPs from *M. millepora* (DQ206392) and *M. efflorescens* (DQ206381: Alieva *et al.*, 2008). The MdigFP2 sequence was 98.8% identical to a chromoprotein from *M. efflorescens* (DQ206377.1: Alieva *et al.*, 2008; Table 1) and it grouped with other chromoproteins in clade B (Figure 2). MdigFP3, on the other hand, clustered with a green FP from *Montipora* sp. (LC029025: Takahashi-Kariyazono *et al.*, 2015). MdigFP1 and MdigFP2 have the QYG chromophore whereas MdigFP3 has LYG (Figure 1). These data suggest that MdigFP1, MdigFP2 and MdigFP3 encode a cyan FP, a chromoprotein, and a green FP, respectively.

The three partial transcripts from *S. caliendrum* encoded incomplete FPs, all with significant BLAST matches to a chromoprotein from *Stylophora pistillata* (DQ206398.1: Alieva *et al.*,

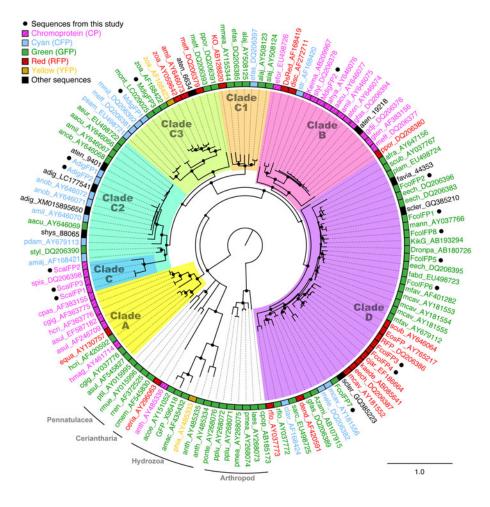


Fig. 2. Bayesian phylogenetic tree showing the relationship amongst fluorescent proteins. FPs with known spectral type are shown in colour (green, GFP; red, RFP; cyan, CFP; yellow, YFP; pink, chromoprotein). Sequences described in this study are marked by black dots and coloured based on their predicted spectral type. Other FP sequences retrieved from the Reefgenomics 20-species comparative database are shown in black text. Clades are based on designations by Alieva et al. (2008). Circles at the nodes indicate posterior probability >0.75. Arthropod FPs are shown as the outgroup. See Supplementary Material Table S2 for the list of accession numbers and species and Supplementary Material File S2 for the trimmed nucleotide alignment.

2008) and a fluorescent chromoprotein homologue of amFP486 also from *S. pistillata* (XM_022922450.1; Table 1). Their respective nucleotide sequences clustered on the branch marked clade C within clade C2. Like the sequence from *S. pistillata*, all three *S. caliendrum* sequences possessed the KYG chromophore (Figure 1). As such, these data suggest that these transcripts encode the same chromoprotein.

Quantitation of FP expression and changes during thermal stress

At ambient temperature conditions, the F. colemani green FP transcripts exhibited the greatest abundance compared with FP transcripts in other coral species (Figure 3A). In coral fragments that had been subjected to 32 ± 1 °C or 28 ± 1 °C for 4, 24, 48 and 72 h, the selected FPs exhibited different expression patterns (Figure 3B). The green FP transcript FcolFP7 was upregulated by 134-fold at 24 h, remained highly expressed by 20-fold at 48 h, but was downregulated by 0.004-fold at 72 h of exposure. Expression of A. digitifera cyan FP transcripts were significantly upregulated by 11.33-fold in heated samples relative to controls at 4 h of exposure then were downregulated at 24 h, with a 0.33-fold decrease in expression at 48 h, before returning to basal levels. The M. digitata green FP transcript MdigFP3 was already downregulated at 24 h, with a 0.06-fold reduction in expression by 48 h, before returning to near basal levels at 72 h. The S. caliendrum transcripts remained close to basal levels throughout the temperature exposure but showed a significant 8.10-fold upregulation in heated samples relative to controls at the 48 h time point. Significant changes in expression did not correspond to changes

in photochemical efficiency reported in the study of Da-Anoy et al. (2019) (Table S4).

Discussion

In this study, we identified a total of 16 transcripts coding for FPs in the adult transcriptomes of four scleractinian corals. Based on their chromophore sequences and phylogenetic clustering, these FP transcripts are predicted to encode FPs with different spectral properties. Although many of the sequences obtained from *de novo* transcriptome assembly were partial or incomplete, structural glycine residues (Fu *et al.*, 2015) were highly conserved, indicating that the protein products would be able to form stable barrel structures. In addition, most of the sequences possessed the conserved tyrosine and glycine amino acids within the chromophore triad (Chudakov *et al.*, 2010; Stepanenko *et al.*, 2011, 2013).

Fluorescent proteins in *A. digitifera* have been thoroughly examined (Takahashi-Kariyazono *et al.*, 2018), while fluorescent colourmorphs and three different FPs have been observed in *M. digitata* (Klueter *et al.*, 2006). However, little is known about the FPs of *F. colemani* and *S. caliendrum.* The number of FP sequences per species that we identified is relatively fewer than the 25 fluorescent protein genes, including short, middle and long wavelength-emitting proteins, and non-fluorescent GFP-like proteins, that were identified in the genome of *A. digitifera* (Takahashi-Kariyazono *et al.*, 2016, 2018). Identification of fewer FPs transcripts from the *de novo* transcriptome assemblies may be due to varied expression of FPs depending on life stage (Roth *et al.*, 2013; Takahashi-Kariyazono *et al.*, 2018), tissue

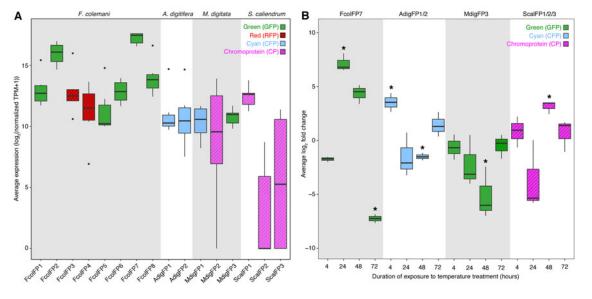


Fig. 3. Expression of FP transcripts in coral species. Species have been ordered by thermal bleaching susceptibility (leftmost, least susceptible; rightmost, most susceptible). (A) Average expression of FP transcripts based on transcriptome data from corals maintained at the control temperature of 28 °C. Values (N = 6) are presented as log₂ of transcripts per million (TPM) + 1. TPM values for each transcript were normalized to the mean expression of reference genes. The species from where the transcripts were identified is indicated above the graph. (B) Average expression of selected transcripts in corals subjected to elevated temperature. Transcripts amplified by the qPCR primers are indicated above the graph. Some primer pairs target similar transcripts, and were thus grouped together. Values are presented as log₂ fold change in expression for samples subjected to 32 °C treatments relative to 28 °C controls for 4, 24, 48 and 72 h (N = 3). Expression values were normalized to the actin transcript as reference. Asterisks indicate statistically significant comparisons between treatment and control groups (Mann–Whitney *U* test, $\alpha = 0.10$). Bars are coloured according to predicted fluorescence types. Error bars indicate standard deviation.

type (Salih *et al.*, 1998) or environmental conditions (Roth *et al.*, 2010; Hume *et al.*, 2013). For example, *Seriatopora hystrix* exhibits a shift from green to cyan FPs as it develops from larvae to adult (Roth *et al.*, 2013). Corals also exhibit fluorescent colour polymorphisms among individuals, subpopulations and across developmental stages (Kenkel *et al.*, 2011; Strader *et al.*, 2016; Takahashi-Kariyazono *et al.*, 2018). Such variation among similar individuals may be attributed to presence-absence polymorphisms or gene dosage effects, and are thought to contribute to the ability of corals to acclimatize to variable light environments (Gittins *et al.*, 2015; Takahashi-Kariyazono *et al.*, 2018). As such, it is possible that the corals in this study possess more FP variants that are not constitutively expressed, warranting further analysis of their genomes, as well as expression in other tissues and growth conditions.

The diversity and level of expression of FP transcripts in the corals that were examined in this study coincided with their reported susceptibility to bleaching (Da-Anoy et al., 2019; Marshall & Baird, 2000). Among these corals, F. colemani and M. digitata expressed the most diverse set of FPs, with representatives of different spectral types. The temperature-tolerant species, F. colemani, also had a larger and more abundant complement of green FP transcripts, whereas only one green FP transcript was found in M. digitata and none in A. digitifera and S. caliendrum. Although it is well-known that many host and symbiont characteristics contribute to the overall resilience of corals, it is possible that FPs also contribute to this characteristic through their ROS scavenging (Bou-Abdallah et al., 2006; Palmer et al., 2009) and light scattering properties (Lyndby et al., 2016). In addition, green fluorescence has been shown to enhance symbiont phototaxis (Aihara et al., 2019). Thus, the presence of multiple green FP transcripts in F. colemani may reduce thermal stress susceptibility by providing the coral with an enhanced mechanism for recovery of symbionts in the event of bleaching.

The expression of transcripts encoding FP homologues varied in response to elevated temperatures and duration of exposure. However, their expression patterns did not necessarily reflect observed changes in symbiont photosynthetic efficiency in the corals (Da-Anoy et al., 2019). Upregulation of fluorescent proteins after the onset of the temperature treatment, as seen in A. digitifera and F. colemani, may be due to the increased rate of cellular processes at warmer temperatures and may contribute to the mitigation of reactive oxygen species that are generated under these conditions. In fact, these corals were able to withstand 3 days of constant exposure to 32 °C with no apparent detrimental effect (Da-Anoy et al., 2019). In contrast, the green fluorescent protein transcript MdigFP3 in M. digitata was significantly downregulated at 48 h of exposure, coinciding with a reduction in symbiont density (Da-Anoy et al., 2019). This expression profile mirrors reports on GFP downregulation in other coral species subjected to temperature stress (Smith-Keune & Dove, 2007; Roth & Deheyn, 2013). On the other hand, S. caliendrum FP transcripts, which belong to the chromoprotein family, were significantly upregulated at 48 h of exposure to higher temperature, coinciding with the decrease in photosynthetic efficiency and symbiont density prior to bleaching (Da-Anoy et al., 2019). The upregulation of transcripts encoding chromoproteins, which are predicted to possess efficient ROS scavenging properties (Palmer et al., 2009), may be an additional protective response in the coral. Seriatopora caliendrum has been noted to be more susceptible to thermal stress (Da-Anoy et al., 2019; Marshall & Baird, 2000), partly due to their thinner tissue layers and the light scattering properties of their skeleton.

Other studies have shown variable responses of fluorescent proteins to temperature exposure. Cyan fluorescence in *Porites lobata* diminished after 17 days of exposure to 31.5 °C (Hume *et al.*, 2013). A green fluorescent protein in *A. millepora*, AmA1a, was significantly downregulated after 6 h at 32 °C and 33 °C (Smith-Keune & Dove, 2007). Green fluorescence and GFP protein concentration both decreased with declining coral health before onset of bleaching due to temperature stress in *A. yongei* (Roth & Deheyn, 2013). In contrast, a chromoprotein gene in *A. millepora*, AmCh, was upregulated in samples undergoing natural bleaching compared with healthy corals (Seneca *et al.*, 2009). These findings suggest that patterns of expression of FPs may be specific to fluorescent type and that FPs may play different roles in corals. Quantification of the expression of different types of FPs, as well as changes in fluorescence or colour intensity, in corals subjected to different environmental conditions are needed to gain further insights into the role of these proteins in the coral stress response.

In this study, we identified a diverse array of 16 fluorescent protein transcripts representing 12 unique genes in four adult scleractinian corals. Among the four coral species, the bleaching resistant species F. colemani had the greatest number of identified FP genes. Different types of FPs exhibited distinct expression trends in corals subjected to elevated temperature. The constitutive expression of fluorescent protein transcripts in some coral species may provide additional protection against the detrimental effects of thermal stress through their reported ROS scavenging properties, as well as their ability to attract symbionts. Together with their reported photoprotective functions, these findings suggest that FPs may contribute to the tolerance of corals to thermal stress. However, it should be noted that, because we are working with de novo assembled transcriptomes that may contain fragmented sequences, the putative fluorescent and non-fluorescent types of FPs that were identified in this study may be incomplete. Further work would be required to retrieve the complete sequences of these genes and to characterize their spectral properties. Analyses of fluorescent protein abundance, localization and antioxidant activity, and their role in symbiont phototaxis in the corals, would also be important future steps to determine their range of functions within the coral holobiont.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0025315421000059

Acknowledgements. The authors extend their gratitude to the staff of the Bolinao Marine Laboratory and the Marine Molecular Biology Laboratory of the University of the Philippines Diliman for assistance with coral collection and experiments.

Financial support. This study was funded in part by a research grant from the Department of Science and Technology Philippine Council for Agriculture, Aquatic, and Natural Resources Research and Development (C.C.) and a thesis writing grant from the University of the Philippines Marine Science Institute (E.G.D.).

References

- Aihara Y, Maruyama S, Baird AH, Iguchi A, Takahashi S and Minagawa J (2019) Green fluorescence from cnidarian hosts attracts symbiotic algae. *Proceedings of the National Academy of Sciences USA* **116**, 2118–2123.
- Alieva NO, Konzen KA, Field SF, Meleshkevitch EA, Hunt ME, Beltran-Ramirez V, Miller DJ, Wiedenmann J, Salih A and Matz MV (2008) Diversity and evolution of coral fluorescent proteins. *PLoS ONE* **3**, e2680.
- Baird AH, Bhagooli R, Ralph PJ and Takahashi S (2009) Coral bleaching: the role of the host. *Trends in Ecology and Evolution* 24, 16–20.
- Barondeau DP, Kassmann CJ, Tainer JA and Getzoff ED (2006) Understanding GFP posttranslational chemistry: structures of designed variants that achieve backbone fragmentation, hydrolysis, and decarboxylation. *Journal of the American Chemical Society* **128**, 4685–4693.
- Barshis DJ, Ladner JT, Oliver TA, Seneca FO, Traylor-Knowles N and Palumbi SR (2013) Genomic basis for coral resilience to climate change. Proceedings of the National Academy of Sciences USA 110, 1387–1392.
- Bay RA and Palumbi SR (2014) Multilocus adaptation associated with heat resistance in reef-building corals. *Current Biology* 24, 2952–2956.
- Bellantuono AJ, Hoegh-Guldberg O and Rodriguez-Lanetty M (2011)
 Resistance to thermal stress in corals without changes in symbiont composition. Proceedings of the Royal Society B: Biological Sciences 279, 1100–1107.
 Bhattacharya D, Agrawal S, Aranda M, Baumgarten S, Belcaid M, Drake JL,
- Erwin D, Foret S, Gates RD, Gruber DF, Kamel B, Lesser MP, Levy O,

Liew YJ, MacManes M, Mass T, Medina M, Mehr S, Meyer E, Price DC, Putnam HM, Qiu H, Shinzato C, Shoguchi E, Stokes AJ, Tambutté S, Tchernov D, Voolstra CR, Wagner N, Walker CW, Weber AP, Weis V, Zelzion E, Zoccola D and Falkowski PG (2016) Comparative genomics explains the evolutionary success of reef-forming corals. *eLife* 5, e13288.

- Bou-Abdallah F, Chasteen ND and Lesser MP (2006) Quenching of superoxide radicals by green fluorescent protein. *Biochimica et Biophysica Acta* 1760, 1690–1695.
- Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* 17, 540–552.
- Chudakov DM, Matz MV, Lukyanov S and Lukyanov KA (2010) Fluorescent proteins and their applications in imaging living cells and tissues. *Physiological Reviews* **90**, 1103–1163.
- Cziesielski MJ, Liew YJ, Cui G, Schmidt-Roach S, Campana S, Marondedze C and Aranda M (2018) Multi-omics analysis of thermal stress response in a zooxanthellate cnidarian reveals the importance of associating with thermotolerant symbionts. *Proceedings of the Royal Society B: Biological Sciences* 285, 20172654.
- Da-Anoy JP, Cabaitan PC and Conaco CG (2019) Species variability in the response to elevated temperature of select corals in northwestern Philippines. Journal of the Marine Biological Association of the United Kingdom 99, 1273–1279.
- Desalvo MK, Voolstra CR, Sunagawa S, Schwarz JA, Stillman JH, Coffroth MA, Szmant AM and Medina MM (2008) Differential gene expression during thermal stress and bleaching in the Caribbean coral *Montastraea faveolata*. *Molecular Ecology* 17, 3952–3971.
- Deschaseaux ESM, Beltran VH, Jones GB, Deseo MA, Swan HB, Harrison PL and Eyre BD (2014) Comparative response of DMS and DMSP concentrations in *Symbiodinium* Clades C1 and D1 under thermal stress. *Journal of Experimental Marine Biology and Ecology* **459**, 181–189.
- **Dove SG, Hoegh-Guldberg O and Ranganathan S** (2001) Major colour patterns of reef-building corals are due to a family of GFP-like proteins. *Coral Reefs* **19**, 197–204.
- Eddy SR (1998) Profile hidden Markov models. Bioinformatics 14, 755-763.
- El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, Sonnhammer EL, Hirsh L, Paladin L, Piovesan D, Tosatto SCE and Finn RD (2019) The Pfam protein families database in 2019. Nucleic Acids Research 47, D427–D432.
- Erez J, Reynaud S, Silverman J, Schneider K and Allemand D (2011) Coral calcification under ocean acidification and global change. In Dubinsky Z and Stambler N (eds), *Coral Reefs: An Ecosystem in Transition*. Berlin: Springer, pp. 151–173.
- Follenius-Wund A, Bourotte M, Schmitt M, Iyice F, Lami H, Bourguignon JJ, Haiech J and Pigault C (2003) Fluorescent derivatives of the GFP chromophore give a new insight into the GFP fluorescence process. *Biophysical Journal* 85, 1839–1850.
- Fu JL, Kanno T, Liang S-C, Matzke AJM and Matzke M (2015) GFP loss-of-function mutations in *Arabidopsis thaliana*. G3: Genes, Genomics, Genetics 5, 1849–1855.
- Gajigan AP and Conaco C (2017) A microRNA regulates the response of corals to thermal stress. *Molecular Ecology* 26, 3472–3483.
- Gardner SG, Raina J, Ralph PJ and Petrou K (2017) Reactive oxygen species (ROS) and dimethylated sulphur compounds in coral explants under acute thermal stress. *Journal of Experimental Biology* **220**, 1787–1791.
- Gittins JR, D'Angelo C, Oswald F, Edwards RJ and Wiedenmann J (2015) Fluorescent protein-mediated colour polymorphism in reef corals: multicopy genes extend the adaptation/ acclimatization potential to variable light environments. *Molecular Ecology* 24, 453–465.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N and Regev A (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29, 644–652.
- Gruber DF, DeSalle R, Lienau EK, Tchernov D, Pieribone VA and Kao HT (2009) Novel internal regions of fluorescent proteins undergo divergent evolutionary patterns. *Molecular Biology and Evolution* **26**, 2841–2848.
- Hall TA (1999) Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41, 95–98.

- Heim R, Prasher DC and Tsien RY (1994). Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proceedings of the National Academy of Sciences USA* **91**: 12501–12504.
- Hoegh-Guldberg O and Bruno JF (2010) The impact of climate change on the world's marine ecosystems. *Science (New York, N.Y.)* **328**, 1523–1528.
- Huelsenbeck JP and Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754–755.
- Hughes TP, Baird AH, Bellwood DR, Card M, Connolly SR, Folke C, Grosberg R, Hoegh-Guldberg O, Jackson JBC, Kleypas J, Lough JM, Marshall P, Nyström M, Palumbi SR, Pandolfi JM, Rosen B and Roughgarden J (2003) Climate change, human impacts, and the resilience of coral reefs. Science (New York, N.Y.) 301, 929–933.
- Hume B, D'Angelo C, Burt J, Baker AC, Riegl B and Wiedenmann J (2013) Corals from the Persian/Arabian Gulf as models for thermotolerant reefbuilders: prevalence of clade C3 *Symbiodinium*, host fluorescence and ex situ temperature tolerance. *Marine Pollution Bulletin* **72**, 313–322.
- Jensen LJ, Julien P, Kuhn M, von Mering C, Muller J, Doerks T and Bork P (2008). eggNOG: automated construction and annotation of orthologous groups of genes. *Nucleic Acids Research* **36**(Database issue), D250–D254.
- Kenkel CD, Traylor MR, Wiedenmann J, Salih A and Matz MV (2011) Fluorescence of coral larvae predicts their settlement response to crustose coralline algae and reflects stress. *Proceedings of the Royal Society B: Biological Sciences* 278, 2691–2697.
- Klueter A, Loh W, Hoegh-Guldberg O and Dove S (2006) Physiological and genetic properties of two fluorescent colour morphs of the coral *Montipora digitata*. Symbiosis 42, 123–134.
- Kumar S, Stecher G and Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33, 1870–1874.
- Langmead B, Trapnell C, Pop M and Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* 10, R25.
- Lazenby D (2000) Cygwin: for Windows NT. Linux Journal 75es, art. 14.
- Lesser MP (2006) Oxidative stress in marine environments: biochemistry and physiological ecology. *Annual Review of Physiology* **68**, 253–257.
- Li B and Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12, 323.
- Liew YJ, Aranda M and Voolstra CR (2016) Reefgenomics.Org a repository for marine genomics data. *Database* (*Oxford*) 2016, baw152.
- Lyndby NH, Kühl M and Wangpraseurt D (2016) Heat generation and light scattering of green fluorescent protein-like pigments in coral tissue. *Scientific Reports* 6, 26599.
- Mann HB and Whitney DR (1947) On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics* 1, 50–60.
- Marshall PA and Baird AH (2000) Bleaching of corals on the Great Barrier Reef: differential susceptibilities among taxa. *Coral Reefs* **19**, 155–163.
- Matz MV, Labas YA and Ugalde J (2006) Evolution of function and color in GFP-like proteins. *Methods of Biochemical Analysis* 47, 139–161.
- Mayfield AB, Chan PH, Putnam HM, Chen CS and Fan TY (2012) The effects of a variable temperature regime on the physiology of the reefbuilding coral *Seriatopora hystrix*: results from a laboratory-based reciprocal transplant. *Journal of Experimental Biology* **215**, 4183–4195.
- Mazel CH, Lesser MP, Gorbunov MY, Barry TM, Farrell JH, Wyman KD and Falkowski PG (2003) Green-fluorescent proteins in Caribbean corals. *Limnology and Oceanography* 48, 402–411.
- Ormö M, Cubitt AB, Kallio K, Gross LA, Tsien RY and Remington SJ (1996) Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science (New York, N.Y.)* **273**, 1392–1395.
- Palmer CV, Modi CK and Mydlarz LD (2009) Coral fluorescent proteins as antioxidants. *PLoS ONE* 4, e7298.
- Parkinson JE, Banaszak AT, LaJeunesse TC and Baums IB (2015) Intraspecific diversity among partners drives functional variation in coral symbioses. *Scientific Reports* 5, 15667.
- Pinzon JH, Kamel B, Burge CA, Harvell CD, Medina M, Weil E and Mydlarz LD (2015) Whole transcriptome analysis reveals changes in expression of immune-related genes during and after bleaching in a reefbuilding coral. *Royal Society Open Science* 2, 140214.
- **R Core Team** (2013) *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing. Available at http://www.R-project.org/.

- Robinson MD, McCarthy DJ and Smyth GK (2010) Edger: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)* 26, 139–140.
- Rodriguez-Lanetty M, Phillips WS and Weis VM (2006) Transcriptome analysis of a cnidarian – dinoflagellate mutualism reveals complex modulation of host gene expression. *BMC Genomics* 7, 23.
- Roth MS and Deheyn DD (2013) Effects of cold stress and heat stress on coral fluorescence in reef-building corals. *Scientific Reports* **3**, 1421.
- Roth MS, Latz MI, Goericke R and Deheyn DD (2010) Green fluorescent protein regulation in the coral *Acropora yongei* during photoacclimation. *Journal of Experimental Biology* **213**, 3644–3655.
- Roth MS, Fan TY and Deheyn DD (2013) Life history changes in coral fluorescence and the effects of light intensity on larval physiology and settlement in *Seriatopora hystrix*. *PLoS ONE* **8**, e59476.
- Salih A, Larkum A, Cox G, Kuhl M and Hoegh-Guldberg O (1998) Photoprotection of symbiotic dinoflagellates by fluorescent pigments in reef corals. In Greenwood JG and Hall NJ (eds), Proceedings of the Australian Coral Reef Society 75th Anniversary Conference, Heron Island, October 1997. Brisbane: School of Marine Science, The University of Queensland, pp. 217–230.
- Salih A, Larkum A, Cox G, Kühl M and Hoegh-Guldberg O (2000) Fluorescent pigments in corals are photoprotective. *Nature* **408**, 850–853.
- Salih A, Cox G, Szymczak R, Coles SL, Baird A, Dunstan A, Mills J and Larkum AW (2006) The role of host-based color and fluorescent pigments in photoprotection and in reducing bleaching stress in corals. In *Proceedings* of the 10th International Coral Reef Symposium, Okinawa, Japan, 28 June – 2 July 2004, pp. 746–756.
- Schlichter D, Fricke HW and Weber W (1986) Light harvesting by wavelength transformation in a symbiotic coral of the Red Sea twilight zone. *Marine Biology* 91, 403–407.
- Seneca FO, Foret S, Ball EE, Smith-Keune C, Miller DJ and van Oppen MJH (2009) Patterns of gene expression in a scleractinian coral undergoing natural bleaching. *Marine Biotechnology* 12, 594–604.
- Shagin DA, Barsova EV, Yanushevich YG, Fradkov AF, Lukyanov KA, Labas YA, Semenova TN, Ugalde JA, Meyers A, Nunez JM, Widder EA, Lukyanov SA and Matz MV (2004) GFP-like proteins as ubiquitous metazoan superfamily: evolution of functional features and structural complexity. *Molecular Biology and Evolution* 21, 841–850.
- Shimomura O (2005) The discovery of aequorin and green fluorescent protein. Journal of Microscopy 217, 3–15.
- Smith EG, D'Angelo C, Salih A and Wiedenmann J (2013) Screening by coral green fluorescent protein (GFP)-like chromoproteins supports a role in photoprotection of zooxanthellae. *Coral Reefs* 32, 463–474.
- Smith-Keune C and Dove S (2007). Gene expression of a green fluorescent protein homolog as a host-specific biomarker of heat stress within a reefbuilding coral. *Marine Biotechnology* 10, 166–180.
- Sniegowski JA, Lappe JW, Patel HN, Huffman HA and Wachter RM (2005). Base catalysis of chromophore formation in Arg96 and Glu222 variants of green fluorescent protein. *Journal of Biological Chemistry* 280, 26248– 26255.
- Stepanenko OV, Stepanenko OV, Shcherbakova DM, Kuznetsova IM, Turoverov KK and Verkhusha VV (2011) Modern fluorescent proteins: from chromophore formation to novel intracellular applications. *Biotechniques* 51, 313–327.
- Stepanenko OV, Stepanenko OV, Kuznetsova IM, Verkhusha VV and Turoverov KK (2013) Beta-barrel scaffold of fluorescent proteins: folding, stability and role in chromophore formation. *International Review of Cellular and Molecular Biology* 302, 221–278.
- Strader ME, Aglyamova GV and Matz MV (2016) Red fluorescence in coral larvae is associated with a diapause-like state. *Molecular Ecology* 25, 559–569.
- Takahashi-Kariyazono S, Satta Y and Terai Y (2015) Genetic diversity of fluorescent protein genes generated by gene duplication and alternative splicing in reef-building corals. *Zoological Letters* **1**, 23.
- Takahashi-Kariyazono S, Gojobori J, Satta Y, Sakai K and Terai Y (2016) *Acropora digitifera* encodes the largest known family of fluorescent proteins that has persisted during the evolution of Acropora Species. *Genome Biology and Evolution* **8**, 3271–3283.
- Takahashi-Kariyazono S, Sakai K and Terai Y (2018) Presence-absence polymorphisms of highly expressed FP sequences contribute to fluorescent

polymorphisms in Acropora digitifera. Genome Biology and Evolution 10, 1715–1729.

- Tatusov RL, Galperin MY, Natale DA and Koonin EV (2000) The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Research* 28, 33–36.
- Tavare L (1986) Some probabilistic and statistical problems of the analysis of DNA sequences. Lecture Notes on Mathematical Modelling in the Life Sciences 17, 57–86.
- Thompson JD, Higgins DG and Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence

weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673–4680.

- Verkhusha VV and Lukyanov KA (2004) The molecular properties and applications of Anthozoa fluorescent proteins and chromoproteins. *Nature Biotechnology* 22, 289–296.
- Yang F, Moss LG and Phillips GN Jr. (1996) The molecular structure of green fluorescent protein. *Nature Biotechnology* 14, 1246–1251.
- Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S and Madden TL (2012) Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* 13, 134.