

# A *Pyropia haitanensis* putative ABA-GPCR homologous protein as an early evolution transmembrane receptor protein?

**Zhujun Zhu**

Ningbo Institute of Oceanography <https://orcid.org/0000-0002-9391-9838>

**Zhang-he Zhen**

Ningbo University

**Haimin Chen** (✉ [chenhaimin@nbu.edu.cn](mailto:chenhaimin@nbu.edu.cn))

Ningbo University

**Zhen Jin**

Ningbo University

**Juanjuan Chen**

Ningbo University

**Denghui Gu**

Ningbo University

**Qijun Luo**

Ningbo University

**Xiaojun Yan**

Ningbo University

**Rui Yang**

Ningbo University

---

## Research article

**Keywords:** ABA-GPCR, G protein-coupled receptor, oligoagar, *Pyropia haitanensis*, salicylic acid

**Posted Date:** November 2nd, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-74933/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

Algae respond to chemical and environmental stimuli in a way similar to higher plants. However, their receptors sensing stimuli remain a mystery.

## Results

We identified and characterized one putative abscisic acid-G-protein coupled receptor homologous protein from red alga *Pyropia haitanensis*, Ph1205, and analyzed its evolution relationship, subcellular localization, and ligand-binding properties. Ph1205 contains 9 predicted transmembrane domains. Phylogenetic analysis revealed its early evolution. Recombinant expression of Ph1205-Egfp in HEK293T cells and immunofluorescence detection against N-terminal extracellular 14 amino acids (N14) of Ph1205 in Ph1205-transfected HEK293T cells and *P. haitanensis* protoplast confirmed its membrane location. Both N14 and Ph1205-transfected cells displayed high affinity with salicylic acid (SA) and oligoagar, which activated cAMP response element in HEK293T cells and upregulated luciferase activity. SA and oligoagar elicited similar but different transcriptional profiles of *P. haitanensis*, mainly by being involved in anabolic metabolism and cell growth. However, related intracellular signaling system in *P. haitanensis* has not yet fully evolved and evidently responded.

## Conclusions

Our results demonstrate that Ph1205 displays some structural and functional characteristics as an early evolutionary membrane receptor, has potential extracellular-binding affinity to SA and oligoagar, and can activate heterotrimeric G-protein and CRE in animal cells.

## Background

Algae are a variety of independent lineages positioned in the early stage of the evolution tree of eukaryotes [1]. Unlike higher plants and mammals, they have no complex tissues or differentiated organs. However, they are exposed to various environmental stresses. For example, algae living in the intertidal zone are usually exposed to such stresses as desiccation, high temperature, and UV damage. In particular, marine algae resemble terrestrial plants in their basic mechanisms for pathogen recognition and signaling transduction, suggesting possible existence of essential cellular functions of terrestrial plants in marine algae [2]. Algae can respond to various stimuli, including oligosaccharides, salicylic acid (SA), methyl jasmonic acid, and Flg22, which can upregulate defense-related genes, e.g. respiratory burst oxidase homolog (rboh), and heat shock protein 70, activating signal pathways, such as oxylipins pathway, shikimate pathway, and halogen metabolism; and stimulating the physiological defense response, such as active oxygen burst, volatile compound emission and halogen production [3–7].

Responses of organisms to chemical or environmental stimuli usually initiate with recognition of extracellular signals by the corresponding membrane receptors [8–10]. Although there are a few studies on the defense-related genes in algae, it is not clear how algae transduce external signals from extracellular space into cells; whether algae have already evolved the related receptor genes and expressed receptor proteins and whether such proteins may have receptor functions still remain a mystery. These capabilities are crucial to the integrity of algal defense system.

Currently, a large number of G protein-coupled receptors (GPCRs), various heterotrimeric G-proteins and their corresponding effectors with different functions have been identified in mammal and fungi. Approximately 800 GPCRs are identified in the human genome, making GPCRs the largest family of cell-surface receptors [11]. These receptors are responsible for sensing a wide variety of structurally diverse ligands. With their activation, numerous cellular signaling cascades are triggered. GPCRs can couple with intracellular heterotrimeric G-protein, leading to changes in cellular functions, including regulation of growth, differentiation and defense [12]. However, to date, there have been fewer reports on plant-derived GPCRs. For instance, one *Arabidopsis* gene (*GCR1*), most similar to GPCR E family cAMP receptors of *Dictyostelium*, encodes a protein with significant sequence similarity to non-plant GPCRs and a predicted 7-transmembrane (TM) domain structure characteristics of GPCRs. It physically interacts with the single prototypical G $\alpha$  (GPA1) in *Arabidopsis*. A null *GCR1* mutant, *gcr1*, was less sensitive toward plant hormones [13]. Two novel *Arabidopsis* proteins, named GPCR-type G proteins 1 and 2 (GTG1/2), were identified and characterized. They contain 9 predicted TM domains, which are not only similar to an orphan vertebrate GPCR (GPCR89) and interact with GPA1 but also have intrinsic GTP-binding and GTPase activity [14]. *Arabidopsis* mutants lacking both GTG1 and GTG2 exhibited hyposensitivity toward abscisic acid (ABA). However, no algal GPCR orthologues have been identified yet. We found a gene functionally annotated as GPCR-like protein, named Ph1205, in *Pyropia haitanensis* genome. It triggered our curiosity and prompted us to investigate Ph1205.

*P. haitanensis*, a species of red alga of Bangiales, Rhodophyta, is an early multicellular alga having evolved after the first endosymbiont. It possesses typical intertidal characteristics not found in single-celled red algae [15, 16]. Therefore, studying its receptors will lead us to gain significant insights into how multicellular red algae have been evolved and how they adapt to complex environment of intertidal zone. In this study, we firstly reported a putative membrane receptor in algae, and conducted a relatively comprehensive characterization of functional properties and evolution relationships of Ph1205. We found that Ph1205 was different from the classic 7-TM GPCRs. It is likely the ancestral form of GPCR exhibiting membrane receptor characteristics, and is possibly involved in activating G-protein signaling pathways. Furthermore, we identified several candidate compounds displaying binding affinity to Ph1205.

## Results

### *Ph1205* sequence analysis and protein structure prediction

The *Ph1205* gene derived from *P. haitanensis* genomic DNA is composed of 3 exons (5'UTR-CDS1, CDS2, and CDS3-3'UTR) (Fig. 1a), with a complete ORF of 1965 bp. It encodes a protein with predicted molecular weight of about 67.18 kDa. Conserved domain search identified Ph1205 belonging to the Golgi pH regulator family N-terminal domain- (GPHR-N, 184 aa-227 aa) and abscisic acid-G-protein coupled receptor (ABA-GPCR, 483 aa-577 aa) domain-containing protein family (G-A family).

Ph1205 contains 9 predicted transmembrane domains (TM1-9) (Fig. 1b). Its N-terminal 14 amino acids, N14, is predicted to be extracellular. Its C-terminus, as long as 74 amino acids, is predicted to be intracellular and contains an extended  $\alpha$  helix by secondary structure prediction (Fig. 1b, red underline).

### Multiple sequence alignment and evolution analysis of Ph1205 homologous proteins

BLAST analysis showed that Ph1205 shared > 83% amino acid sequence identity with homologous sequences from two red algae of *Bangiophyceae*, *Pyropia yezoensis* and *Porphyra umbilicalis* (Additional file 1: Table S1). But the sequence identity between Ph1205 and those from other selected eukaryotes, including unicellular red algae (*Galdieria sulphuraria* and *Porphyridium purpureum*), multicellular red algae *Floridophyceae* (*Gracilariopsis chorda* and *Chondrus crispus*), Chromalveolata, green algae, fungi, plants and animals, are < 37%. Conserved domain search showed that they all belonged to G-A family (Additional file 1: Table S1). However, no orthologs were found in the genomes of unicellular red algae *Cyanidioschyzon morulae*, some selected prokaryotic cyanobacteria and proteobacteria.

We analyzed TM structures of the eukaryotic G-A family members. The results showed that while their TM distribution patterns were similar, their TM numbers were quite different (Additional file 2: Table S2). Among which, TMs were distributed on both terminals of the sequence, and connected with a large cytoplasmic loop (LCL). In fungi, higher plants and mammals, the TM numbers are relatively consistent, showing the distribution of 5TM-LCL-4TM. However, in some homologous proteins of algae, the predicted TM numbers at both terminals of LCL changed. For example, in red algae, Ph1205 is 5TM-LCL-4TM; while in *P. umbilicalis*, OSX76732 is 5TM-LCL-1TM; and in *G. chorda*, PXF46398 is 4TM-LCL-4TM.

We performed multiple sequence alignments among eukaryotic G-A family members. The conserved regions were mainly located in TM domains except TM1 (Fig. 2). GPHR-N domain is located in TM5-CL3. ABA-GPCR domain is located between TM7-TM9 region, and the highly conserved amino acid sites are mainly distributed in TM8-TM9. Additionally, three short conserved motifs were found in the junction region of TM and extracellular loop (EL) (Fig. 2, # mark). For example, [EQ]-[YF]-x-F-[FY] was found in TM9-EL4 junction region. Relatively high frequency of [RK] and [ML] appeared in several TM and cytoplasmic loop (CL) junction regions (Fig. 2, \* mark).

In other regions, especially in the two terminal fragments, the homology among G-A family is very low. The homologous proteins of *Bangiophyceae*, including Ph1205, have the characteristic sequences at both terminals not found in other species. They are longer in length than those derived from animals or plants, and are of less conservatism (Fig. 2). For example, the extracellular N-terminal of Ph1205 (N14 peptide) contains 7 and 4 more amino acids than human GPCR89 and *Arabidopsis* GTG do, with low

identity. Compared to the C-terminus of *Arabidopsis* GTG, Ph1205 held a longer C-terminal extension fragment of 61 amino acids that may form the tenth predicted helix structure. In addition, some unique long fragments are inserted in loop regions of *Bangiophyceae* homologous proteins. For example, a long amino acid sequence (23 amino acids) not found in other eukaryotic G-A orthologues was inserted at EL2. Ph1205 also contains unique fragments longer than other eukaryotic members in EL3 and CL3. Furthermore, from several CL or EL regions, e.g. CL3, more amino acid mutations were observed in homologous proteins of red algae or multicellular red algae, but more conservative in those of vesicular algae or green algae in evolution (Fig. 2, orange background).

PROSITE motif analysis identified a conserved ATP-/GTP-binding domain in *Arabidopsis* GTG (Fig. 2, region 382-411 for GTG1), which is not found in Ph1205 and human GPCR89. Compared with the consensus pattern of GTP-binding motif ([LIV]-G-G-{P}-[FYWMGSTNH]-[SGA]-{PW}-[LIVCAT]-{PD})-x-[GSTACLIVMFY]-x(5,18)-[LIVMFYWCSTAR]-[AIVP]-[LIVMFAGCKR]-K), Ph1205 maintained only seven conserved amino acids (Fig. 2, blue background and @ mark). Other eukaryotic G-A family members except high plant orthologs are also less conserved, especially at the N-terminal two amino acids, which do not conform to [LIV]-G (Fig. 2). In addition, the LCL of *Arabidopsis* GTG shares a 62%-68% similarity with the degenerate Ras GTPase-activating protein domain ([GSNA]-x-[LIVMF]-[FYCI]-[LIVMFY]-R-[LIVMFY](2)-[GACNS]-[PAV]-[AV]-[LIV]-[LIVM]-[SGANT]-P) (Fig. 2, amino acids 230–243 in GTG1) [14]. Ph1205 has only 28.57% homology in this region (Fig. 2, cyan background and @ mark).

To assess the evolutionary relationship among these proteins, we constructed phylogenetic tree (Fig. 3). The protein from unicellular red alga *G. sulphuraria* appears to be the earliest origin of G-A family, thereafter they form two clades: one is for red algae, including multicellular red algae and unicellular *P. purpureum*. Their branching situations are consistent with the evolutionary relationship of their genomes [17-19]. Another gradual evolutionary branch has formed two large clusters, one large group includes mainly homologous proteins of fungi, amoeba and metazoans, which share an ancestral homologous sequence with that of cryptoalgae *Guillardia theta* and that of a collar flagellate *Monosiga brevicollis*. Another large group includes mainly the homologous sequences of the green plant kingdom, including green algae, and terrestrial plants. The evolution of the homologous sequences of the brown algae *Ectocarpus siliculosus* and *Oomycetes*, lies between the two clades of red algae and other eukaryotes. However, several members of diatoms are in parallel with homologous sequences of *G. sulphuraria* on the evolutionary tree, being an independent branch.

### **Analysis of subcellular localization of Ph1205 protein**

In order to analyze subcellular localization of Ph1205 protein, we performed three experiments. First, we transfected the HEK293T cells with pCI recombinant plasmids carrying *Egfp* or the fused *Ph1205-Egfp*. When only *Egfp* was expressed in HEK293T cells, *Egfp* green fluorescence was distributed throughout the cytoplasm (Fig. 4a). However, when the fused *Ph1205-Egfp* complex was expressed, *Egfp* green fluorescence was mainly concentrated on the cellular membrane. The green fluorescence was overlapped very well with the red fluorescence derived from the membrane probe Dil with a Pearson correlation

coefficient ( $r$ ) of 0.98. These results indicate that Ph1205 can be selectively localized to the cellular membrane (Fig. 4a).

Then we transfected HEK293T cells with pCI-*Ph1205* and immunochemically hybridized cells with N14 polyclonal antibody, produced by a N-terminus peptide of Ph1205, followed by detection with FITC-linked secondary antibody. Fig. 4b showed that the green fluorescence was also concentrated in the cell membrane, and overlapped mainly with the red fluorescence derived from Dil-red with the Pearson correlation coefficient ( $r$ ) of 0.98, further confirming that the recombinant Ph1205 is localized to the cellular membrane of HEK293T cells (Fig. 4b).

We then prepared protoplast of *P. haitanensis* and similarly probed the protoplast with the primary antibody of N14. The results showed that the green fluorescence only existed at the edge of *P. haitanensis* cells and the spontaneous red fluorescence was seen inside *P. haitanensis* cells, clearly confirming that Ph1205 is indeed localized on the membrane of *P. haitanensis* cells (Fig. 4c).

### The interactions between Ph1205 and candidate compounds

We determined whether the membrane-localized Ph1205 can act as a receptor for binding with candidate compounds. Oligoagar, JA, SA, 1-octen-3-ol, and BL, that have been studied in algae and can cause the resistant responses of algae [3, 4, 6, 20-22], were selected. Phytohormones, including ABA, gibberellic acid (GA), IAA and Flg22, were also selected. Their binding kinetics to N14 polypeptide of Ph1205 were analyzed by FortéBio Octet system. Fig. 5a showed that the N14 polypeptide could interact with SA, oligoagar, JA and Flg22. Except for Flg22, the binding and dissociation of SA, oligoagar and JA showed dose-dependent changes. Their binding constants ( $K_{on}$ ), dissociation constants ( $K_{off}$ ) and affinity constants ( $K_d$ ) were listed in Table 1. Binding rates of both SA and Flg22 to N14 were fast, but there was no significant difference in the dissociation rate among three compounds. The  $K_d$  values reflected that the binding affinity of SA to N14 peptide was the highest, reaching  $5.13E-05$ , followed by that of Flg22. JA displayed only a weaker binding affinity for N14. Because oligoagar is a mixture with no precise molecular weight, its binding affinity constant could not be calculated. However, oligoagar exhibited a quite high binding affinity to N14. At high concentrations, the binding thickness with the probe reached 0.6 nm.

**Table 1.** The affinity constants between N14 peptide of Ph1205 and candidate compounds.

Compounds	$K_{on}$ ( $ms^{-1}$ )	$K_{off}$ ( $s^{-1}$ )	$K_d$ (M)	Full $R^2$
Salicylic acid	4.66E+03	2.39E-01	5.13E-05	0.9889
Jasmonic acid	4.88E+01	5.99E-01	1.23E-02	0.9788
Flg22	3.82E+03	6.43E-01	1.68E-04	0.6553

We also tested their binding to HEK293T cells transfected with pCI-*Ph1205*. These compounds bound to the probed HEK293T cells (Fig. 5b), but were not completely consistent with their binding to N14 described above. Oligoagar, Flg22 and JA were bound with the cells in a concentration-dependent manner while JA binding was still weaker. The dissociation of oligoagar at moderate and high concentrations were better. Flg22 showed higher binding to probed HEK293T cells. At high concentrations, the probe thickness could reach 0.88 nm. However, SA no longer showed a concentration-dependent binding trend.

### **CRE-mediated signal transduction in Ph1205-expressing HEK293T cells induced by candidate Compounds**

After binding to its ligand, GPCR activates intracellular G-protein and triggers the downstream gene expression via the regulation of CRE present in the cells [23, 24]. Thus, by taking the advantage of the endogenous CRE in cells, we co-transfected the HEK293T cells with PCI-*Ph 1205* and Dual-Luciferase reporter systems and further screened the compounds binding to Ph1205 (Fig. 5c). Both SA and oligoagar aroused significant signal responses ( $P < 0.05$ ), with a concentration-dependent trend. At high concentration, SA stimulated the highest signal response, consistent with the SA-N14 interaction ( $P < 0.01$ ), while treatments of HEK293T cells with Flg22 and JA also caused the increasing fluorescence signals though not significant.

### **Transcriptional changes of *Ph1205* regulated by candidate compounds**

Both JA and oligoagar significantly promoted the mRNA level of *Ph1205* gene (Fig. 6). When thalli were stimulated by JA for 0.5 h at 50  $\mu\text{M}$ , *Ph1205* transcription was increased by 10.85-fold ( $P < 0.01$ ). Oligoagar at 50  $\mu\text{g mL}^{-1}$  for 3 h significantly up-regulated overall Ph1205 expression by 13.45-fold ( $P < 0.01$ ). However, both SA and Flg22 hardly changed but even down-regulated its transcription.

### **Transcriptome patterns of *P. haitanensis* stimulated by SA and oligoagar**

The transcriptome sequencing of *P. haitanensis* after being stimulated with SA and oligoagar was conducted. Comparing two sets of transcriptome data revealed that their effects on types of GO biological processes and KEGG pathways were quite similar (Fig. 7a). For instance, in the GO classification, the three mostly enriched biological processes were all Metabolic processes, Cellular process, and Signal-organism process. The mostly enriched in Cellular components were Membrane, and Membrane part, etc. In KEGG Pathway Kierarchy1, the mostly significant differential pathways in Cellular processes were Transport and catabolism. In Metabolism, the mostly significant differential pathways were Carbohydrate metabolism and Energy metabolism. However, comparing the changes in gene transcription profiles between two groups revealed that the up- and down-regulation trends induced by SA and oligoagar were quite different, and many genes have shown even the opposite changes. For instance, in the SA treatment group, the up-regulation pathways were enriched in Carbon fixation photosynthetic organisms, Carbon metabolism, and Pentose phosphate pathway, etc. In the oligoagar treatment group, these pathways were just down-regulated. Thus, we chose Carbon fixation photosynthetic organisms to

conduct a detailed comparison. Fig. 7b showed that most of the genes were reversely regulated when *P. haitanensis* was treated with SA and oligoagar. The heat-map showed this opposition.

There were also significant differences in the expression profiles of the transcriptional factors after being stimulated by these two compounds (Additional file 3: Table S3). For example, after stimulation with oligoagar at 200  $\mu\text{g mL}^{-1}$  for 1 h, the transcription factors that were significantly up-regulated were mainly those in the C2H2 family and the TIG family. Pha005926 of C2H2 family gene was up-regulated by 247.28-fold ( $\text{Log}_2\text{FC} = 7.95$ ,  $q < 0.01$ ); and TIG family gene, Pha005519, was upregulated by 26.72-fold ( $\text{Log}_2\text{FC} = 4.74$ ,  $q < 0.01$ ); Under SA stimulation, Pha005926 was even significantly down-regulated by 4.23-fold ( $\text{Log}_2\text{FC} = -2.08$ ,  $q < 0.01$ ).

Referring to the SA signaling pathway in plants (<http://www.yeastrc.org/pdr/viewGONode.do?acc=GO:0009862>), we retrieved genes in this pathway from *P. haitanensis* genome (Additional file 4: Table S4). However, after SA stimulation, almost no significant transcriptional differences were seen, indicating that the SA signaling pathway in *P. haitanensis* has not been evolved as that in plants. We also performed homologous search of the G-protein of *P. haitanensis* and related genes involved in cAMP signaling pathway. No information about heterotrimeric Ga and  $\beta/\gamma$ -subunit homology in *P. haitanensis* genome was obtained, but many members of other Ras-like GTPase family were found (Additional file 5: Table S5). After stimulation by SA and oligoagar at 100  $\mu\text{M}$  for 1 h, these genes did not appear to be differentially expressed, except for one ribosome-associated GTPase EngA gene, Pha007322, whose transcript was significantly up-regulated by 1.74-fold ( $\text{Log}_2\text{FC} = 0.80$ ,  $q < 0.05$ ). Furthermore, the cAMP signaling pathway of *P. haitanensis* was also found imperfect (Additional file 5: Table S5). For example, no adenylyate cyclase gene and cAMP-dependent protein kinase gene were detected, but two genes encoding adenylyate cyclase-associated protein (CAP) genes, Pha000384 and Pha006544, and one gene encoding the conserved domain of catalytic domain of Ser/Thr kinases, Pha009109, were found while they were not differentially regulated by either SA or oligoagar.

## Discussion

GPCRs represent the largest family of transmembrane receptors. However, far fewer GPCRs are present in plants [14, 25]. Currently, no detailed studies on algal GPCRs have been reported. We conducted a search for genes encoding GPCR-like protein from *P. haitanensis* genome and retrieved only one related gene annotated as GPCR, namely Ph1205. Interestingly, we observed that Ph1205 was similar to *Arabidopsis* GTG. They belonged to the G-A family and have 9 predicted TM domains, indicating that Ph1205 and *Arabidopsis* GTG may have originated from the same ancestor. The evolutionary origin of this family can be traced back to the single celled red algae, since the homologous sequence can be found in the genomes of *G. sulphuraria* and *P. purpureum*. In red algae, homologous proteins of this family are clustered on the earliest branch of evolution, and the sequence of *G. sulphuraria* is the ancient origin of other eukaryotes. However, several members of the diatom are placed in a syngeneic position with *G. sulphuraria* on the phylogenetic tree, but are scattered in different branches with other algae of

Chromalveolata, including brown algae, *E. siliculosus* and Cryptophyta, *Guillardia theta*, perhaps associated with the secondary endosymbiont of the red algae cells engulfed by different host cells independently [26].

In prokaryotic genomes, such as those of cyanobacteria and proteobacteria, no homologous sequences of G-A family have been found. The bacterial FeoB is a seven-transmembrane transporter that mediates GTPase-dependent iron uptake in bacteria [27]. Pandey *et al.* suggested that FeoB and the GTG homologs present in unicellular organisms were possibly primordial GPCRs that have proliferated as receptor-effector systems in higher organisms [14]. However, in terms of sequence homology, the TM number and distribution pattern of conserved domains, FeoB is positioned far from G-A family. By performing motif analysis, we found no conserved motifs in the homologous proteins of red algae, with neither the conserved ATP-/GTP-binding region nor the degenerate Ras GTPase-activating protein domain. Thus, we speculated that the ancestors of G-A family proteins might not yet have acquired GTPase activity. Multiple sequence alignment also revealed the consensus pattern of the junction region between cytoplasmic loop 4 (CL4) and TM8 in this family. In this junction region, except for those of higher plants, all eukaryotic G-A members were of low conservation to the ATP-/GTP-binding motif [14]. In particular, only the sequences of higher plants strictly conform to the conservative first two amino acid sites as V-G. While in Ph1205, the two amino acids are replaced by T-N, suggesting that these two sites may have a key influence on the spatial location of ATP-/GTP-binding.

The eukaryotic G-A family proteins showed more highly conserved sites appeared on TM regions. For instance, a conserved motif [EQ]-[YF]-x-F-[FY] exists at the extracellular terminus of TM9 among all eukaryotic homologous proteins, including Ph1205, *Arabidopsis* GTG and human GPCR89. In some classic 7TM-GPCR family, a short-conserved motif is also present. For example, in GPCR A family, both the TM3 intracellular terminus of human  $\beta 2$  adrenoceptor ( $\beta 2$ AR) and bovine rhodopsin have the conserved [DE]-R-[YW] motifs. They are involved in maintaining the basic conformation of the receptor [28, 29]. Therefore, it can be speculated that among the G-A family, conservatism is mainly in the TM region, which may be important for maintaining the conformation of the protein.

However, the variable sequences at two terminals and loop regions (CL and EL) are present in G-A family. For example, the obvious difference among Ph1205, *Arabidopsis* GTG and human GPCR89 is that Ph1205 has a potential helix extension in the predicted C-terminus. Moreover, this fragment only exists in homologous proteins of *Bangiophyceae*. In some members of the classic 7TM-GPCR family, such as  $\beta 2$ AR, there also exists the 8<sup>th</sup> helix structure at the intracellular C-terminus, which was reported to be involved in the interactions with GPCR kinases, arrestins and other signaling molecules [29]. Ph1205 contains a longer C-terminal intracellular extension than *Arabidopsis* GTG does, indicating that the intracellular region of Ph1205 may have evolved into a specific conformation to participate in certain signaling functions. Additionally, Ph1205 also contains unique fragments longer than GTG in three loop regions, i.e. EL2, EL3 and CL3. Thus, the flexibility and variability of loop region and terminal regions may be critical for functionality, diversity and specificity in G-protein activation for the extracellular binding ligands and intracellular signal transduction [28].

Expression of the recombinant Ph1205 fused with Egfp in HEK293T cells displayed membrane localization of Ph1205. Furthermore, the immunofluorescence detection in Ph1205-expressing HEK293T cells and isolated protoplast of *P. haitanensis* with N14 antibody all clearly and strongly demonstrate that Ph1205 is a membrane-localized protein.

We next addressed whether it can actually perform the membrane receptor functions. The extracellular N-terminal of GPCRs is crucial for ligand binding. For example, secretin-like receptors (GPCR B1) has a conserved cysteine ligand binding region at its extracellular N-terminus [30]. Arg46 of two GPCR A proteins, vasotocin receptor (V1aR) and oxytocin receptor (OTR), is the extracellular amino acid necessary for agonist binding [31-33]. Herein, we first used N14 of Ph1205 to test its candidate ligands. Among the 9 tested compounds, the binding of N14 to SA and oligoagar were quite good. However, the N14 binding alone might not fully account for the entire protein binding. For instance, due to the short N-terminal fragment of rhodopsin receptor, most rhodopsin-like ligands will also interact with the extracellular loop to activate the receptor [28, 34, 35]. The EL2 and EL3 of Ph1205 are 35-46 amino acids longer than those of *Arabidopsis* GTG and human GPCR89. In addition to N-terminus, the extracellular loop region of Ph1205 is likely to participate in or affect the ligand recognition and binding. Thus, we expressed the entire Ph1205 protein on the membrane surface of HEK293T cells and observed the cell binding to compounds mentioned above. We observed that oligoagar was consistently and dose-dependently interacted with both N14 and Ph1205-expressing cells. Its binding affinity to them was all high; while SA exhibited the highest affinity binding just with N14, indicating that the extracellular loop region of Ph1205 is likely to affect the overall binding. *Arabidopsis* GTG was reported to function as the classic membrane-localized ABA receptors [14]. However, Ph1205 displayed no binding affinity to ABA in our study, further indicating that the low conserved extracellular loop may affect its ligand binding selectivity.

By taking advantage of the presence of endogenous CRE in HEK293T cells and application of Dual-Luciferase reporter assay, we observed that several candidate compounds could activate the CRE of HEK293T cells by binding to recombinant Ph1205, indicating that the intracellular region of Ph1205 has already acquired at least partial function as a membrane receptor to activate G-protein. Among the compounds tested, both SA and oligoagar exhibited the most significant signal response, consistent with the binding results. Resistance experiments already found that both SA and oligoagar could inhibit the infection of pathogenic bacteria and partially restore the health of algae [7, 36].

Several 7TM-GPCRs crystal structure characterization suggests that at the terminals of the TM helices, certain amino acids will participate in the G-protein coupling. For example, in the GPCR-B family, some highly conserved amino acids (R, K, N, M, L) are the key sites for receptor binding to intracellular G protein [37, 38]. Among Ph1205 and other eukaryotic homologous proteins, the CL regions are poorly conserved. In spite of this, several conserved amino acids (R, K, M, L) are present in the position where the CL connects with TM1-TM7. It may be involved in the interaction between Ph1205 and G protein in the cell, so that recombinant Ph1205 can activate human G protein in HEK293T cells and initiate the downstream CRE, i.e. the sequence and conformation of ph1205 have already evolved the partial functionality of membrane receptors.

However, the homologous sequence of G $\alpha$  protein was not found in *P. haitanensis* genome. In other red algae, such as *C. merolae* and *C. crispus*, heterotrimeric G proteins, at least the alpha and beta subunits, are missing in their genomes [39]. However, *C. crispus* harbors a series of conserved Ras superfamily GTPases, including Ran, MIRO, SpG1/TEM1, LIP1/RabL3, and signal recognition particle receptor (SR) subunit [39]. We also retrieved several similar genes related to their members, such as Ran, SpG, Rab, Arf and EngA, which might be the transitional state of G $\alpha$  protein, but also indicates that in the early stage of evolution, the intracellular signal system of G-A family proteins is not perfect yet. However, binding and activation of Ph1205 are not directly related to Ph1205 transcription regulation. SA and flg22 bound to Ph1205 but did not stimulate Ph1205 transcription. On the other hand, although its binding affinity was poor, JA could significantly stimulate Ph1205 transcription. Oligoagar not only interacted with Ph1205 but also promoted its transcription. Both JA and oligoagar were reported to enhance resistance of algae to environmental stresses [3, 4, 20, 21]. Ph1205 may act as a member responsible for enhancing the resistance of *P. haitanensis* to environmental stresses.

Comparing the transcriptomes induced by SA and oligoagar revealed that both GO and KEGG signaling pathways affected by SA and oligoagar were quite consistent, indicating that *P. haitanensis*, as a lower organism after endosymbiosis, may be stimulated to respond resistance by a relatively simple mechanism, mainly via regulating carbon assimilation, anabolic metabolisms, and energy metabolism etc. i.e., to resist stresses by controlling cell growth. Unlike higher plants and animals, *P. haitanensis* has not yet evolved into too complex, multi-cellular and cooperative adaptation mechanisms. The SA regulatory system in *P. haitanensis* has not evolved as that of higher plants, and its cAMP signaling pathway is also incomplete. SA and oligoagar utilized quite different ways to regulate the expression of genes in GO and KEGG pathways, and even reversed their regulatory effects, indicating a large difference in the signal transduction or regulation patterns. For example, we observed that the expression levels of transcription factors regulated by SA and oligoagar were significantly different.

## Conclusions

Overall, we identified and characterized a new and putative ABA-GPCR homologous protein, Ph1205, from *P. haitanensis* for the first time. Ph1205 was different from the classic 7 TM-GPCRs, but could be an earlier and primitive form of GPCR. This protein has exhibited some characteristics of membrane GPCRs, including membrane localization, affinity binding with certain phytohormones or elicitors, activation of G-protein and downstream CRE-regulating genes in human cells. Ph1205 may act as the membrane receptor for the signal transduction of both SA and oligoagar, activating and enhancing the defensive resistance of *P. haitanensis* by stimulating carbon assimilation, energy metabolism and cell growth through quite different regulation mechanisms. However, the intracellular signaling system of GPCR in *P. haitanensis* is not yet fully evolved.

## Methods

### Materials

A variety of cultivars of *P. haitanensis* ((T.J.Chang & B.F.Zheng) N.Kikuchi & M.Miyata 2011) [40], “Zhedong-1” (artificially bred by Q. Luo and his team) were used and the gametophytes were collected from Hepu, Xiangshan Harbor, Ningbo, Zhejiang Province, China (29°09'18"N, 121°54'05"W), and cultivated in sterile seawater at 20°C for 24 h, under illumination of 20  $\mu\text{mol m}^{-2} \text{s}^{-1}$  of the photon flux density.

HEK293T cells from China Center for Type Culture Collection (CCTCC) (Wuhan, Hubei, China) were cultured with MEM medium (Gibco, Gaithersburg, MD, USA) containing 10% fetal bovine serum at 37°C under 5% CO<sub>2</sub>.

Flg22 (QRLSTGSRINSAKDDAAGLQIA) was synthesized by Ningbo KangBei Biochemical Co. Ltd (Ningbo, Zhejiang, China). Oligoagar was prepared in the laboratory with degree of polymerization of 2-10 as described previously [41].

### **Cloning and sequencing of *Ph1205***

The genomic DNA and total RNA were isolated from *P. haitanensis* gametophytes using Plant DNA Extraction kit (TianGen, Beijing, China) and Takara RNAiso Plus kit (Takara, Japan), respectively. Total RNA was reversely transcribed into cDNA with the Primescript RT Reagent Kit with gDNA Eraser kit (Takara, Japan).

A sequence, functionally associated to GPCR-like protein, was obtained from the genome DNA of *P. haitanensis*, and named as *Ph1205* (GeneBank ID, KU600033). Gene-specific primers (GsP), P1205T-S/A, were designed to clone the open reading frame (ORF) of *Ph1205* at 64°C (Additional file 6: Table S6). The PCR products were purified and sequenced.

### **Structural analysis of Ph1205 protein**

The secondary structures of Ph1205 protein were analyzed by its deduced primary amino acid sequence with HNN secondary structure prediction method ([https://npsa-prabi.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=/NPSA/npsa\\_hnn.html](https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_hnn.html)). Transmembrane regions were predicted at TMHMM v2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>). Conserved domains were analyzed on NCBI Conserved domain Research (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>). The protein functional motifs were identified on ExpASY PROSITE (<https://prosite.expasy.org/>).

### **Multiple sequence alignment and phylogenetic analyses**

Sequences homologous to Ph1205 protein were retrieved from NCBI GeneBank database. Multiple sequence alignment was performed with the Muscle method in MEGA 5 and edited by GeneDoc software. Phylogenetic tree was constructed with MEGA 5 software using neighbor-joining (N-J) method with bootstrap of 1000 replications and edited by Coreldraw v8 software.

### **Construction of recombinant expressing vectors**

A green fluorescent protein (Egfp) was used as an indicator. Three pairs of GsP (Additional file 6: Table S6) were designed to construct different recombinant pCI-Egfp, pCI-Ph1205 (with stop codon) and pCI-Ph1205-Egfp vectors. All PCR products were purified and then subjected to the corresponding double digestion together with pCI-based vectors at 37°C for 3 h. Digested products were purified, ligated with T4 ligase (Takara, Japan) at 16°C overnight, and transformed into *E. coli* DH-5 $\alpha$ . Single colony was selected and cultured. Endotoxin-free plasmids were extracted by using the QIAprep Spin Miniprep kit (QIAGEN, Germany).

### **Cell transfection and fluorescence signal observation**

HEK293T cells were cultured in a 20 mm diameter confocal culture dish until 50-60% confluence, mixed with 7.5  $\mu$ L of Lipofectamine 3000 Reagent, 10  $\mu$ L of P3000 Reagent (Invitrogen, USA) and 5  $\mu$ g endotoxin-free recombinant plasmids (pCI-Egfp, pCI-Ph1205 or pCI-Ph1205-Egfp) in serum-free MEM medium, and cultured for 24 h. Then the cells were stained with 200  $\mu$ L of 10  $\mu$ M Dil red solution (Beyotime, Shanghai, China) and incubated at 37°C for 20 min, fixed with 4% paraformaldehyde. Finally, the cells were incubated in 4,6-diamidino-2-phenylindole (DAPI) reagent (Beyotime, Shanghai, China) at room temperature for 10 min, washed and sealed. High resolution laser confocal microscope (Zeiss LSM 880, Germany) was used to visualize the subcellular localization of Egfp protein. The  $\lambda_{Ex}/\lambda_{Em}$  for DAPI, Egfp and Dil red were 359/461 nm, 490/520 nm and 549/565 nm, respectively.

### **Antibody preparation and immunofluorescence detection**

The N-terminal 14 amino acids of Ph1205 protein (MVLSGPLAMAASDA, named N14) was synthesized and used as antigen to prepare mouse polyclonal antibody. Mice were injected with 0.5 mL of 1 mg mL<sup>-1</sup> N14 solution emulsified by 1:1 (v/v) DMSO (Sigma, USA) and Freund's complete adjuvant FCA (F-5881, Sigma, USA) once a week for three weeks. On the fourth day after the final injection, mouse blood was collected. Serum was separated as antibody solution.

After transfection with pCI-Ph1205 and fixation as described above, HEK293T cells were incubated with N14 antibody solution in 1:50 dilution at 4°C overnight. After washing, they were incubated in 1:500 FITC-labeled goat anti-mouse IgG (H+L) (Beyotime, Shanghai, China) at room temperature for 1 h, washed with PBS and incubated consecutively with 200  $\mu$ L of 10  $\mu$ M Dil-red and 200  $\mu$ L of DAPI. After sealing, the fluorescent signal was visualized as described above. The  $\lambda_{Ex}$  and  $\lambda_{Em}$  for FITC-linked secondary antibody were 494 nm and 518 nm, respectively.

### **Preparation of *P. haitanensis* protoplasm and immunofluorescence detection**

*Littorina brevicula* gland was grinded on ice in 1:5 (w/v) PBS (0.1 M, pH 7.2). The mixture was centrifuged at 2,000 g for 10 min. The supernatant was collected and added into sterile seawater with a ratio of 1:5 (v/v) as the digestive solution.

The healthy *P. haitanensis* were cut into 2-3 mm pieces, which were digested in 1 mL of digestive solution at 20°C for 2.5 h, filtered by 80 mesh filter, and rinsed with sterile seawater 3 times. *P. haitanensis* cells were suspended in sterile seawater containing 6 mM glucose.

The N14 antibody serum was diluted with sterile seawater (1:100, v/v), added into the cell suspension and incubated at 20°C overnight. After being rinsed with sterile seawater three times, the cells were incubated in FITC-labeled goat anti-mouse IgG (H+L) seawater dilution (1:200, v/v) at room temperature for 1 h and then washed the FITC fluorescence was visualized as above.

### **FortéBio Octet system assay for binding affinity of N14 peptide and Ph1205-transfected HEK293T cells with candidate compounds**

Binding affinities between Ph1205 protein and candidate compounds were evaluated by biolayer interferometry on an Octet K2 system (FortéBio, Fremont, CA, USA). The candidate compounds included SA, ABA, jasmonic acid (JA), brassinolide (BL), gibberellic acid (GA), indoleacetic acid (IAA) (Yuanye, Shanghai, China), 1-octen-3-ol (Sigma, USA), Flg22 and oligoagar. Oligoagar concentrations used were 0.25, 0.5 and 1 mg mL<sup>-1</sup>, respectively. The concentrations of other compounds were 0.25 mM, 0.5 mM and 1 mM, respectively.

N14 peptide was biotinylated by incubating with equivalent amounts of biotin for 30 min. The sensors (Super Streptavidin, and SSA) were loaded with biotinylated N14 for 15 min, quenched in 10 μM biotin for 1 min, and then dipped into the test compound solutions for association for 120 s and then dissociation. After referencing corrections, based on the subtracted binding interference data, the kinetics parameters (K<sub>on</sub> and K<sub>off</sub>) and affinity constants (K<sub>d</sub>) were calculated using the FortéBio analysis software v9.0.0.10.

Binding studies of Ph1205-transfected HEK293T cells with compounds above were also performed on the Octet K2 system. After transfection with pCI-*Ph1205*, HEK293T cells were immobilized on FortéBio's aminopropyl silane (APS) biosensors at 30°C and then dipped into the test compound solutions for association and dissociated.

### **Transient co-transfection and luciferase reporter assays**

HEK293T cells were transiently co-transfected with 20 ng pCRE-Luc (PT3286-1, CloneTech, San Francisco, CA, USA), 20 ng pSV40-Renilla (E2231, Promega, Madison, WI, USA), and 50 ng pCI-*Ph1205* by using X-tremeGENE HP DNA Transfection Reagent (Roche, Switzerland). After 24 h of culture, cells were exposed to 25 μL of candidate compounds at the indicated concentrations for 4 h. Oligoagar concentrations were 10, 50 and 100 μg mL<sup>-1</sup>, respectively. Other compounds' concentrations were 10, 50 and 100 μM, respectively. Then firefly and Renilla luciferase activities in cell lysates were determined using Dual-Glo<sup>®</sup> Luciferase Assay System (Promega, Madison, WI, USA). The relative luciferase activity (F value) was normalized using Renilla luciferase activity. Activity differences between groups were evaluated using *t* tests. *P*<0.05 was considered to be statistically significant and experiments were performed in triplicate.

## Real-time quantitative PCR detection for *Ph1205*

The *P. haitanensis* thalli were grown at a density of 1.5 mg mL<sup>-1</sup> and exposed to SA, JA, Flg22 and oligoagar at different concentrations above at 20°C for 0.5, 1 and 3 h. Thalli samples were collected at each time point. Total RNA was isolated and reversely transcribed into cDNA with PrimeScript RT reagent Kit Perfect Real Time kit (Takara, Japan).

Real-time quantitative PCR (qPCR) assay was performed to analyze the mRNA level of *Ph1205* with P1205-QF/QR (Additional file 6: Table S6). *P. haitanensis18s* was used as internal control. The relative transcriptional changes (C value) were calculated with the 2<sup>-ΔΔCt</sup> method and differences between groups were evaluated by *t* tests.

## Transcriptome analysis

The thalli (150 mg) were cultured in 100 mL of sterile seawater containing 20 and 100 μM SA, and 100 and 200 μg mL<sup>-1</sup> oligoagar at 20°C for 1 h in dark and quickly collected. Total RNA was extracted and the cDNA libraries were constructed. Library preparations were sequenced on an Illumina HiSeq (PE150) platform by Frasergen Company (Wuhan, Hubei, China). Paired-end reads were generated and aligned to the reference genome using TopHat v2.1.1. Gene expression levels were quantified in terms of fragments per kilobase of exon per million mapped fragments (FPKM) using DEseq2 with default parameters. Differentially expressed genes (DEGs) between groups were identified based on following thresholds: absolute of log<sub>2</sub> (fold-change) > 1 (<-1) and *q*-value (false discovery rate (FDR)) < 0.05. All DEGs were mapped to terms in the GO, KEGG and Swiss-prot databases and were ready for the comparison. The raw reads are deposited in the NCBI SRA repository (PRJNA642213).

## Abbreviations

ABA-GPCR: abscisic acid-G-protein coupled receptor; GPHR-N: the Golgi pH regulator family N-terminal domain; G-A family: GPHR-N and ABA-GPCR domain-containing protein family; TM: transmembrane domain; LCL: large cytoplasmic loop; EL: extracellular loop; CL: cytoplasmic loop; N14 peptide: N-terminal extracellular 14 amino of Ph1205; DAPI: 4,6-diamidino-2-phenylindole; SA: salicylic acid; JA: jasmonic acid; BL: brassinolide; GA: gibberellic acid; IAA: indoleacetic acid; OA: oligoagar; GsP: Gene-specific primers; UTR: untranslated region; CDS: coding sequence.

## Declarations

### Ethics approval and consent to participate

Materials were collected from our experimental base at the coast of Xiangshan harbor in Ningbo, Zhejiang province, China. Sampling was permitted by the local government (Xiangshan County Government) and the local department of fisheries (Ningbo Ocean & Fishery Bureau).

## Consent for publication

Not applicable.

## Availability of data and materials

The raw reads for transcriptome were uploaded to the NCBI SRA repository under accession number PRJNA642213. Other data generated during this study are included in this published article and its supplementary information files, and all materials generated during this study are available upon request.

## Competing interests

The authors declare that they have no competing interests.

## Funding

This work was supported by the National Key R&D Program of China (2018YFD0900305) for collecting samples, National Natural Science Foundation of China (31872540, and 31772871) for analyzing data, Natural Science Foundation of Zhejiang Province (LY18C190004) for analyzing data, State Key Laboratory Project (ZS 20190110), Major Scientific and Technological Project of Zhejiang Province (2016C02055-6B) for culturing algae, Earmarked Fund for China Agriculture Research System (CARS-50) for culturing algae, K. C. Wong Magna Fund in Ningbo University for providing instrument and Ningbo Public Service Platform for High-Value Utilization of Marine Biological Resources (NBHY-2017-P2) for writing the manuscript.

## Authors' contributions

HC, XY and RY planned and designed the research. ZZ, Z-HZ and ZJ performed experiments. QL and DG collected samples. ZZ, JC and ZJ analyzed data. ZZ and HC wrote the manuscript. All authors read and approved the final manuscript.

## Acknowledgements

We thank FraserGen Company for the transcriptome analysis.

## References

1. Baldauf SL, Roger AJ, Wenk-Siefert I, Doolittle WF. A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science*. 2000;290(5493):972-7. doi: 10.1126/science.290.5493.972.
2. Potin P, Bouarab K, Salaün J-P, Pohnert G, Kloareg B. Biotic interactions of marine algae. *Current Opinion in Plant Biology*. 2002;5(4):308-17. doi: 10.1016/s1369-5266(02)00273-x.
3. Küpper FC, Müller DG, Peters AF, Kloareg B, Potin P. Oligoalgininate recognition and oxidative burst play a key role in natural and induced resistance of sporophytes of Laminariales. *Journal of Chemical*

- Ecology. 2002;28:2057-81.
4. Weinberger F, Guillemain M, Destombe C, Valero M, Faugeron S, Correa JA, et al. Defense Evolution in the Gracilariaceae (Rhodophyta): Substrate-Regulated Oxidation of Agar Oligosaccharides Is More Ancient Than the Oligoagar-Activated Oxidative Burst. *Journal of Phycology*. 2010;46(5):958-68. doi: 10.1111/j.1529-8817.2010.00887.x.
  5. Tarakhovskaya ER, Maslov YI, Shishova MF. Phytohormones in algae. *Russian Journal of Plant Physiology*. 2007;54(2):163-70. doi: 10.1134/s1021443707020021.
  6. Tu S, Yang R, Xu X, Chen J, Luo Q, Zhu Z, et al. Flg22-triggered oxylipin production in *Pyropia haitanensis*. *Phycological Research*. 2017;65(1):86-93. doi: 10.1111/pre.12163.
  7. Weinberger F. Pathogen-Induced Defense and Innate Immunity in Macroalgae. *Biology Bulletin*. 2007;213:290-302.
  8. Shan X, Yan J, Xie D. Comparison of phytohormone signaling mechanisms. *Current opinion in plant biology*. 2012;15(1):84-91. doi: 10.1016/j.pbi.2011.09.006.
  9. Boller T, He SY. Innate immunity in plants: an arms race between pattern recognition receptors in plants and effectors in microbial pathogens. *Science*. 2009;324(5928):742-4. doi: 10.1126/science.1171647.
  10. Luttrell LM. Reviews in molecular biology and biotechnology: transmembrane signaling by G protein-coupled receptors. *Molecular biotechnology*. 2008;39(3):239-64. doi: 10.1007/s12033-008-9031-1.
  11. Lagerstrom MC, Schiöth HB. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nature reviews Drug discovery*. 2008;7(4):339-57. doi: 10.1038/nrd2518.
  12. Cabrera-Vera TM, Vanhauwe J, Thomas TO, Medkova M, Preininger A, Mazzoni MR, et al. Insights into G protein structure, function, and regulation. *Endocrine reviews*. 2003;24(6):765-81. doi: 10.1210/er.2000-0026.
  13. Pandey S, Assmann SM. The Arabidopsis putative G protein-coupled receptor GCR1 interacts with the G protein alpha subunit GPA1 and regulates abscisic acid signaling. *The Plant cell*. 2004;16(6):1616-32. doi: 10.1105/tpc.020321.
  14. Pandey S, Nelson DC, Assmann SM. Two novel GPCR-type G proteins are abscisic acid receptors in Arabidopsis. *Cell*. 2009;136(1):136-48. doi: 10.1016/j.cell.2008.12.026.
  15. Sutherland JE, Lindstrom SC, Nelson WA, Brodie J, Müller KM. A new look at an ancient order: Generic revision of the bangiales (Rhodophyta). *Journal of Phycology*. 2011;47(5):1131-51.
  16. Diehl N, Michalik D, Zuccarello GC, Karsten U. Stress metabolite pattern in the eulittoral red alga *Pyropia plicata* (Bangiales) in New Zealand - mycosporine-like amino acids and heterosides. *Journal of Experimental Marine Biology and Ecology*. 2019;510:23-30.
  17. Cao M, Xu K, Yu X, Bi G, Liu Y, Kong F, et al. A chromosome-level genome assembly of *Pyropia haitanensis* (Bangiales, Rhodophyta). *Molecular ecology resources*. 2020;20(1):216-27. doi: 10.1111/1755-0998.13102.

18. Brawley SH, Blouin NA, Ficko-Blean E, Wheeler GL, Lohr M, Goodson HV, et al. Insights into the red algae and eukaryotic evolution from the genome of *Porphyra umbilicalis* (Bangiophyceae, Rhodophyta). *Proceedings of the National Academy of Sciences of the United States of America*. 2017;114(31):E6361-E70. doi: 10.1073/pnas.1703088114.
19. Lee J, Yang EC, Graf L, Yang JH, Qiu H, Zelzion U, et al. Analysis of the Draft Genome of the Red Seaweed *Gracilariopsis chorda* Provides Insights into Genome Size Evolution in Rhodophyta. *Molecular biology and evolution*. 2018;35(8):1869-86. doi: 10.1093/molbev/msy081.
20. Hervé C, Tonon T, Collé J, Corre E, Boyen C. NADPH oxidases in Eukaryotes: red algae provide new hints! *Current Genetics*. 2006;49:190-204. doi: 10.1007/s00294-0050044-z.
21. Wang F, Wang C, Zou T, Xu N, Sun X. Comparative transcriptional profiling of *Gracilariopsis lemaneiformis* in response to salicylic acid- and methyl jasmonate-mediated heat resistance. *PLoS one*. 2017;12(5):e0176531. doi: 10.1371/journal.pone.0176531.
22. Chen H, Yang R, Chen J, Luo Q, Cui X, Yan X, et al. 1-Octen-3-ol, a self-stimulating oxylipin messenger, can prime and induce defense of marine alga. *BMC plant biology*. 2019;19(1):37. doi: 10.1186/s12870-019-1642-0.
23. Fitzgerald LR, Mannan IJ, Dytko GM, Wu HL, Nambi P. Measurement of responses from Gi-, Gs-, or Gq-coupled receptors by a multiple response element/cAMP response element-directed reporter assay. *Analytical Biochemistry* 1999;275(1):54-61.
24. Dressler H, Economides K, Favara S, Wu NN, Pang Z, Polites HG. The CRE luc bioluminescence transgenic mouse model for detecting ligand activation of GPCRs. *Journal of biomolecular screening*. 2014;19(2):232-41. doi: 10.1177/1087057113496465.
25. Jones AM, Assmann SM. Plants: the latest model system for G-protein research. *EMBO reports*. 2004;5(6):572-8. doi: 10.1038/sj.embor.7400174.
26. Zimorski V, Ku C, Martin WF, Gould SB. Endosymbiotic theory for organelle origins. *Current opinion in microbiology*. 2014;22:38-48. doi: 10.1016/j.mib.2014.09.008.
27. Marlovits TC, Haase W, Herrmann C, Aller SG, Unger VM. The membrane protein FeoB contains an intramolecular G protein essential for Fe(II) uptake in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*. 2002;99(25):16243-8. doi: 10.1073/pnas.242338299.
28. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, et al. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science*. 2000;289:739-45.
29. Rasmussen SG, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, et al. Crystal structure of the human  $\beta$ 2 adrenergic G-protein-coupled receptor. *Nature*. 2007;450(7168):383-7. doi: 10.1038/nature06325.
30. Couvineau A, Laburthe M. The family B1 GPCR: structural aspects and interaction with accessory proteins. *Current Drug Targets* 2012;13:103-15.
31. Hawtin SR, Wesley VJ, Parslow RA, Patel S, Wheatley M. Critical role of a subdomain of the N-terminus of the V1a vasopressin receptor for binding agonists but not antagonists; functional rescue by the oxytocin Receptor N-Terminus. *Biochemistry*. 2000;39:13524-33.

32. Hawtin SR, Simms J, Conner M, Lawson Z, Parslow RA, Trim J, et al. Charged extracellular residues, conserved throughout a G-protein-coupled receptor family, are required for ligand binding, receptor activation, and cell-surface expression. *The Journal of biological chemistry*. 2006;281(50):38478-88. doi: 10.1074/jbc.M607639200.
33. Wesley VJ, Hawtin SR, Howard HC, Wheatley M. Agonist-Specific, High-Affinity Binding Epitopes Are Contributed by an Arginine in the N-Terminus of the Human Oxytocin Receptor. *Biochemistry*. 2002;41:5086-92.
34. Bockaert J, Pin JP. Molecular tinkering of G protein-coupled receptors: an evolutionary success. *Embo Journal*. 1999;18(7):1723-9.
35. Sandoval A, Eichler S, Madathil S, Reeves PJ, Fahmy K, Bockmann RA. The Molecular Switching Mechanism at the Conserved D(E)RY Motif in Class-A GPCRs. *Biophysical journal*. 2016;111(1):79-89. doi: 10.1016/j.bpj.2016.06.004.
36. Weinberger F, Friedlander M. Response of *Gracilaria conferta* (Rhodophyta) to oligoagars results in defense against agar-degrading epiphytes. *Journal of Phycology*. 2000;36:1079-86.
37. Song G, Yang D, Wang Y, Graaf Cd, Zhou Q, Jiang S, et al. Human GLP-1 receptor transmembrane domain structure in complex with allosteric modulators. *Nature*. 2017;546(7657):312-5. doi: 10.1038/nature22378.
38. Zhang H, Qiao A, Yang D, Yang L, Dai A, Graaf Cd, et al. Structure of the full-length glucagon class B G-protein-coupled receptor. *Nature*. 2017;546(7657):259-64. doi: 10.1038/nature22363.
39. Collén J, Porcel B, Carré W, Ball SG, Chaparro C, Tonon T, et al. Genome structure and metabolic features in the red seaweed *Chondrus crispus* shed light on evolution of the Archaeplastida. *Proceedings of the National Academy of Sciences of the United States of America*. 2013;110(13):5247-52. doi: 10.1073/pnas.1221259110.
40. Sutherland JE, Lindstrom SC, Nelson WA, Brodie J, Lynch MD, Hwang MS, et al. A new look at an ancient order: generic revision of the Bangiales (Rhodophyta). *Journal of Phycology*. 2011;47(5):1131-51. doi: 10.1111/j.1529-8817.2011.01052.
41. Chen HM, Zheng L, Lin W, Yan XJ. Product monitoring and quantitation of oligosaccharides composition in agar hydrolysates by precolumn labeling HPLC. *Talanta*. 2004;64(3):773-7. doi: 10.1016/j.talanta.2004.04.002.

## Additional Files

**Additional files were not provided with this version of the manuscript.**

**Additional file 1:Table S1.** A comparative analysis of GPHR-N and ABA-GPCR domain containing proteins in *Pyropia haitanensis* and other selected species. (PDF 121 kb)

**Additional file 2:Table S2.** Prediction of transmembrane domains in GPHR-N and ABA-GPCR domain containing proteins. (PDF 87 kb)

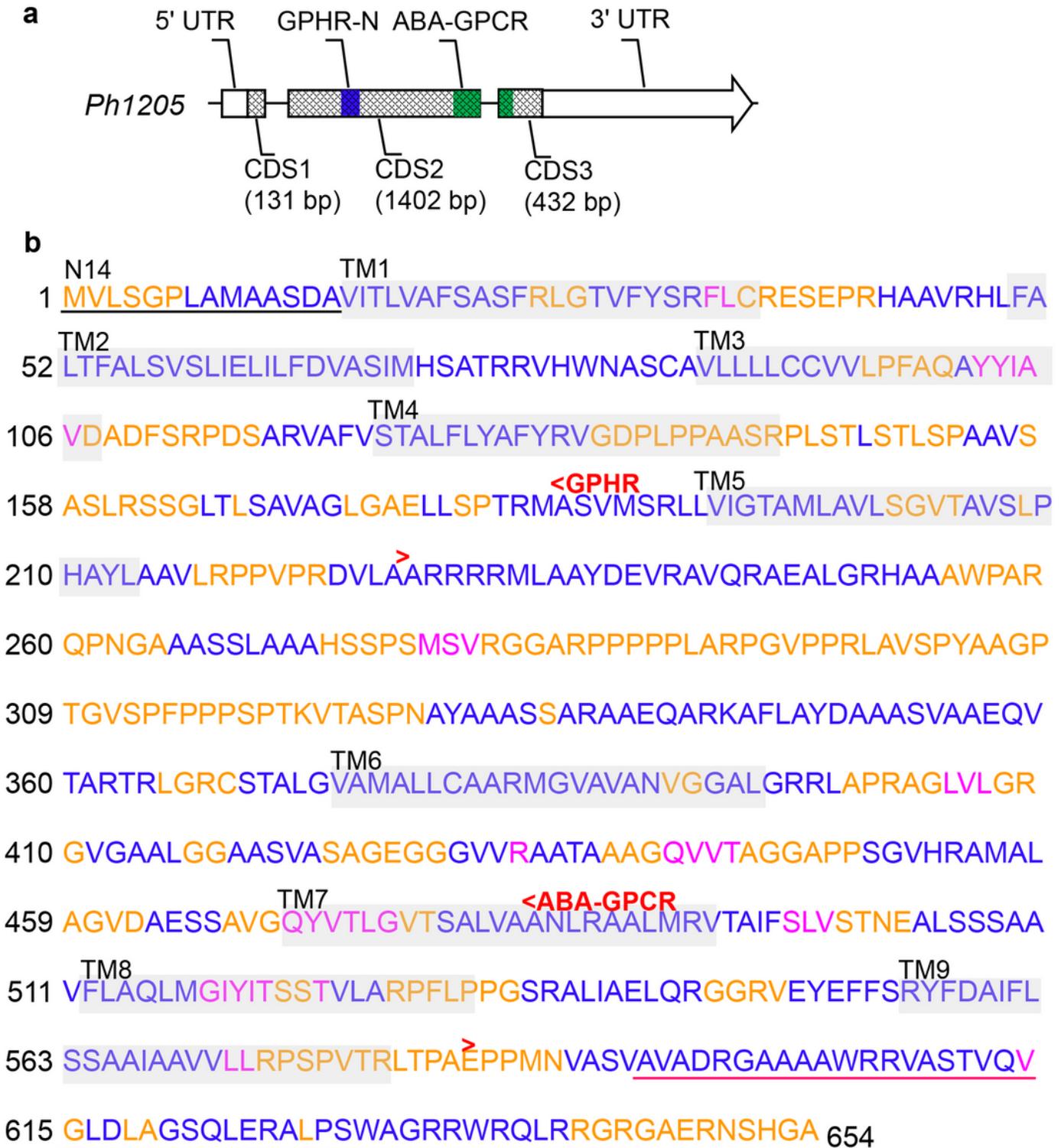
**Additional file 3:Table S3.** Differential expression of transcription factors in *Pyropia haitanensis* transcriptome after stimulation by salicylic acid and oligoagar. (PDF 94 kb)

**Additional file 4:Table S4.** Homologous information of salicylic acid signaling pathway-related genes in *Pyropia haitanensis* genome and differential expression after salicylic acid stimulation. (PDF 115 kb)

**Additional file 5:Table S5.** Homologous information of the Ras-like GTPase family and cAMP-signaling components in *Pyropia haitanensis* genome and differential expression under stimulation. (PDF 115 kb)

**Additional file 6:Table S6.** Various gene special primers used in the experiment. (PDF 76 kb)

## Figures

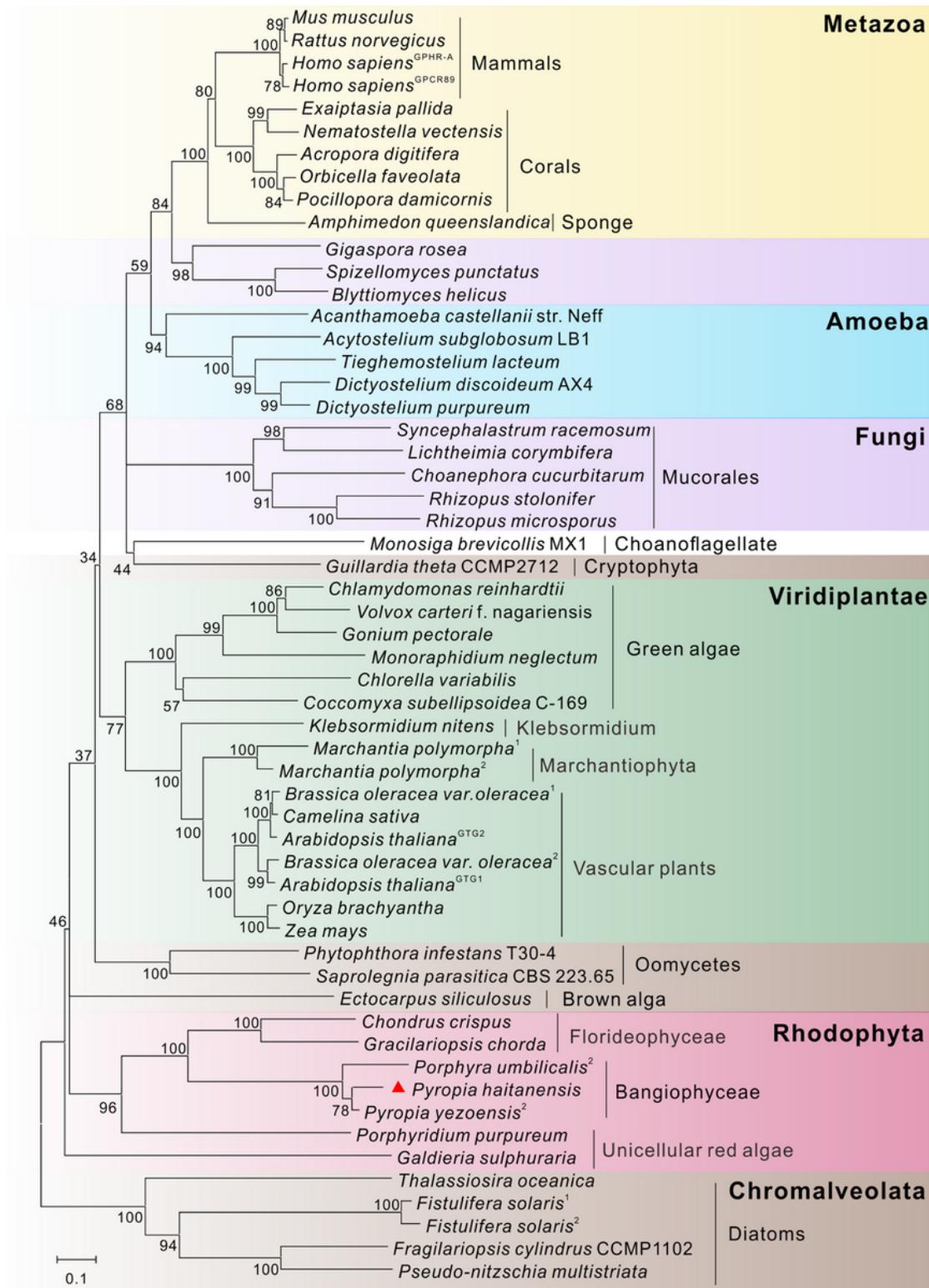


**Figure 1**

The sequence analysis of Ph1205. a Ph1205 gene structure. The black frame area is the full-length sequence of Ph1205 mRNA composed of three exons on the genome. The grid fill indicates the open read frame (ORF) of Ph1205 consisted of three coding sequences (CDS). Unfilled regions are untranslated regions (UTR) at both terminals of the gene. The blue filling indicates GPHR-N domain and the green filling indicates ABA-GPCR domain. b The predicted secondary structures and transmembrane (TM)

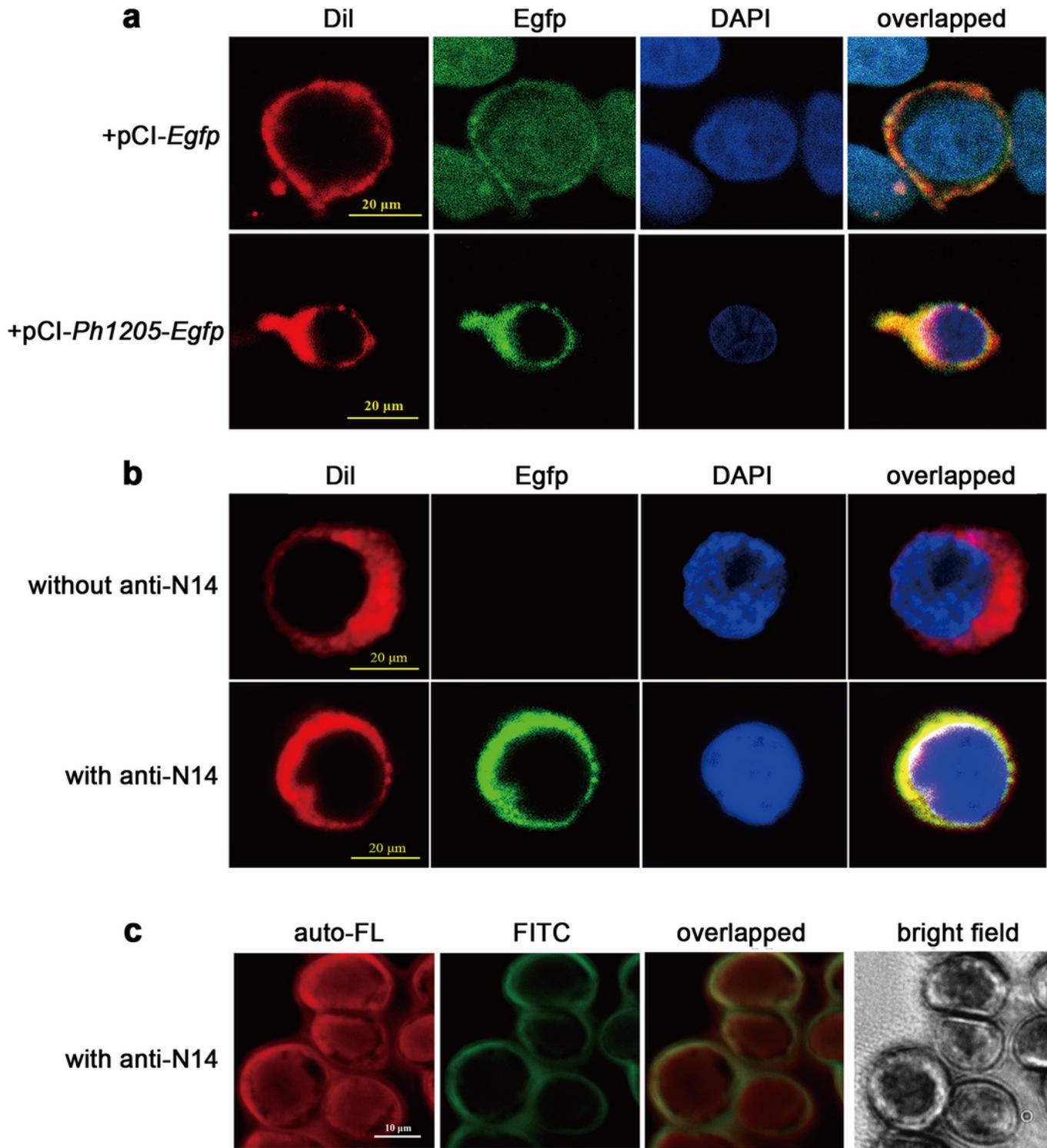


Multiple sequence alignment of GPHR-N and ABA-GPCR domain containing proteins from selected species. TM, transmembrane domains, CL, cytoplasmic loops, EL, extracellular loops. Conserved domains such as GPHR-N and ABA-GPCR were showed by red and green boxes. Residues related to the Ras GTPase-activating protein motif in CL3 or the ATP-/GTP-binding motif in CL4-TM8 were marked in blue or cyan background. Conserved motifs near the extracellular part of TM5, TM7 and TM9 were marked by the symbol # and red residues. Conserved residues related to [RK] and [ML] in the cytoplasmic terminus of TM1, TM4, TM5 and TM7 were marked by the symbol \* and showed in green. The characteristic segments in Ph1205 or homologous proteins from Bangiophyceae were marked by pink background. The conserved sites appear in loops of most homologous proteins except that of red algae or multicellular red algae were marked by orange background.



**Figure 3**

The unrooted phylogenetic tree of GPHR-N and ABA-GPCR domain containing proteins. Sequence alignment and construction of evolutionary tree in MEAG5, using the neighbor-joining (N-J) method, with 1000 bootstraps and other default parameters. The red triangle indicates Ph1205. The sequence ID from selected species refer to Additional file 1: Table S1.



**Figure 4**

Subcellular localization of Ph1205 protein. a Subcellular localization of recombinant Ph1205-Egfp in HEK293T cells. HEK293T cells were transfected with pCI-Egfp or pCI-Ph1205-Egfp plasmids, respectively. The cells were then stained with Dil red and DAPI. b Subcellular localization of recombinant Ph1205 in HEK293T cells immunochemically detected with antibody against N14. HEK293T cells were transfected with pCI-Ph1205 plasmids. N14 primary antibody and FITC-labeled secondary antibody were treated

consecutively, followed by staining with Dil red and DAPI dye. Both A and B were visualized under laser confocal microscope. Bar: 10  $\mu\text{m}$ . c Subcellular localization of Ph1205 in thalli immunochemically detected with antibody against N14. The protoplasm of *P. haitanensis* was incubated with N14 primary antibody, followed by being coupled with FITC-labeled secondary antibody. The antibody-labeled protoplasm was visualized under fluorescence microscope. Bar: 20  $\mu\text{m}$ .

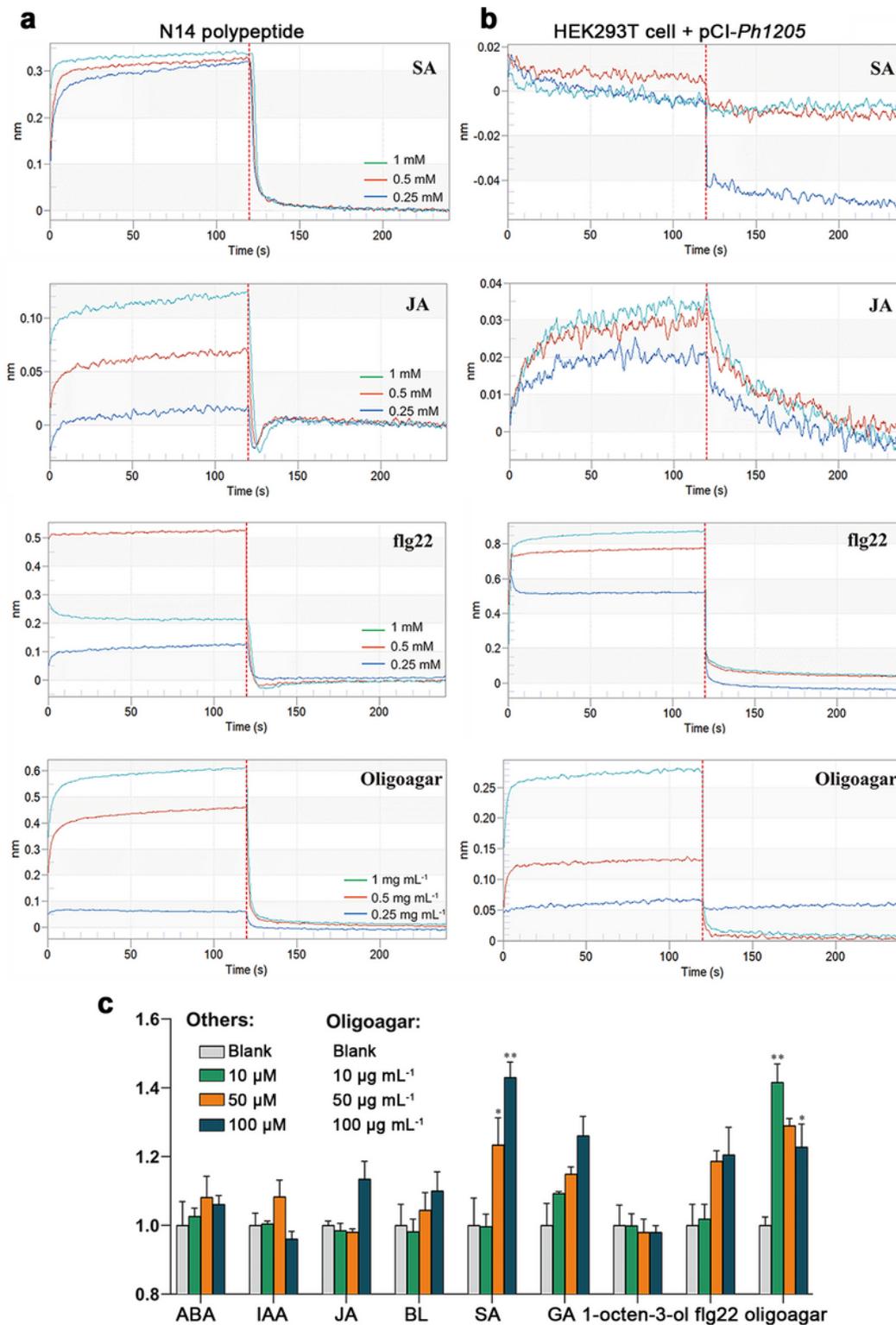
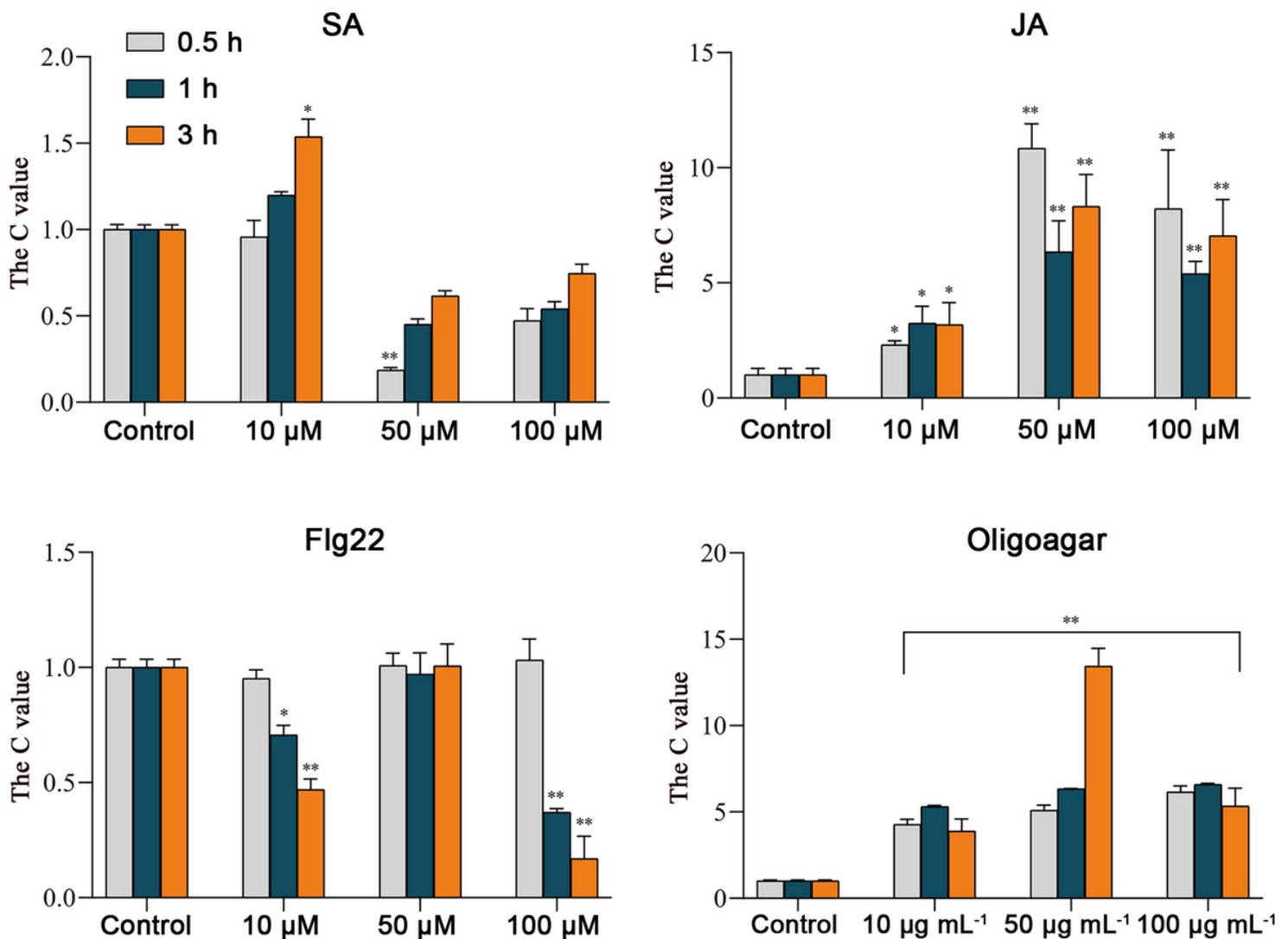


Figure 5

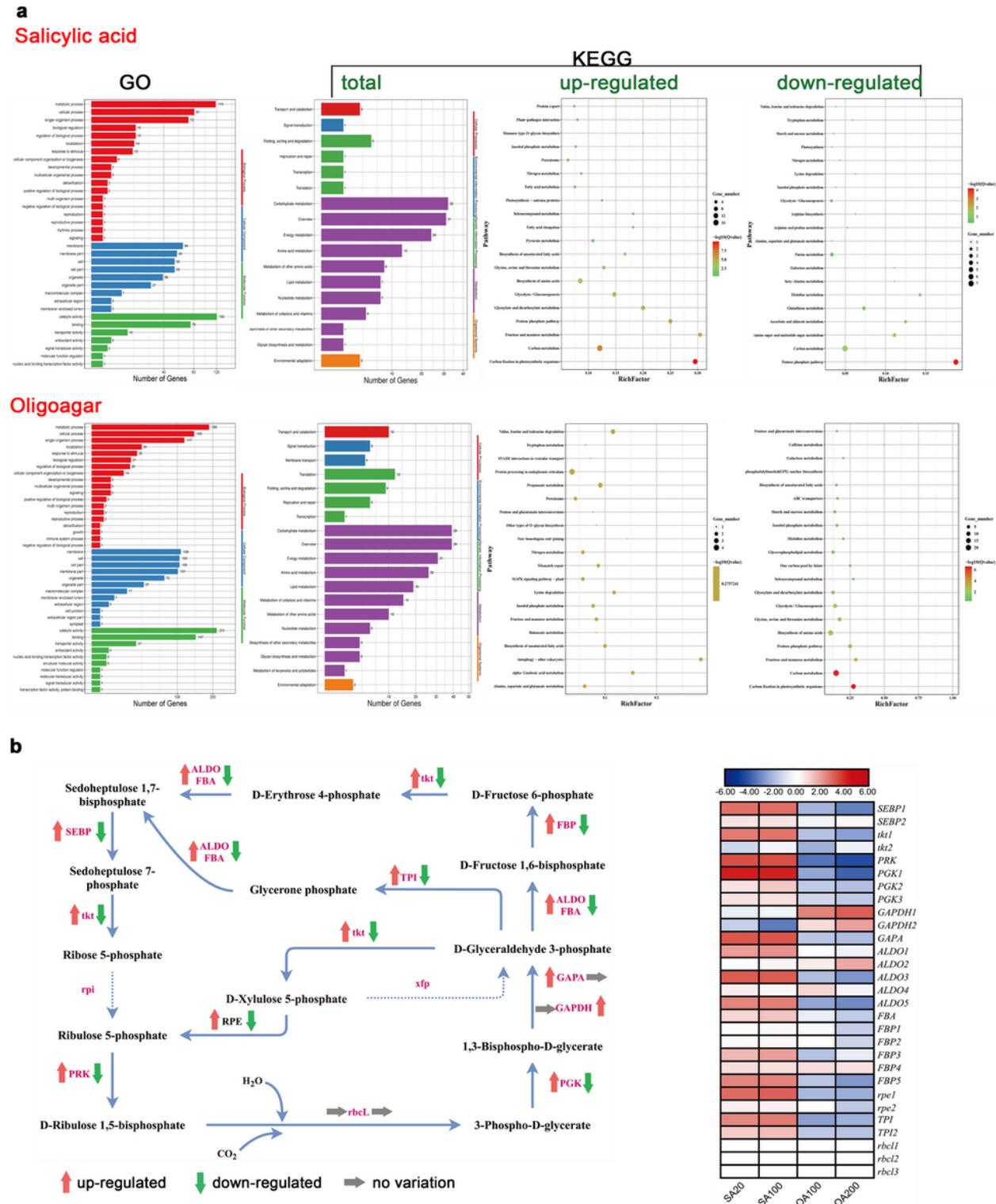
Interactions of N14 peptide, Ph1205-transfected HEK293T cells with candidate compounds and intracellular signal activation. a Interactions between N14 peptide with the indicated candidate compounds. b Interactions between Ph1205-transfected HEK293T cells and the indicated candidate compounds. The binding affinities of N14 peptide and HEK293T cells with these compounds were detected on an Octet K2 system. N14 peptide and HEK293T cells were immobilized on biosensors, which were then dipped into the compound solutions for association and returned to the blocking solution for dissociation. c Measurement of luciferase activity after Ph1205-transfected HEK293T cells were stimulated with different compounds. HEK293T cells were transiently co-transfected with pCRE-Luc, pSV40-Renilla, and pCI-Ph1205 and incubated for 24 h. The cells were exposed to 25  $\mu$ L of the indicated compounds solution for 4 h. The relative luciferase activity was normalized using Renilla luciferase activity and showed as F value. \* $P$ <0.05, \*\* $P$ <0.01.



**Figure 6**

Transcriptional differences of Ph1205 induced by salicylic acid, oligoagar, jasmonic acid and Flg22. Thalli samples were exposed to SA, JA, Flg22 and oligoagar at different concentrations at 20°C for 0.5 h,

1 h and 3 h. The Ph1205 transcriptional levels were assayed with RT-qPCR analysis and showed as the C value. *P. haitanensis* 18s was used as internal control. \*P <0.05, \*\*P <0.01.



**Figure 7**

Comparison in transcriptome of *Pyropia haitanensis* after being stimulated by salicylic acid and oligoagar. a Histogram and enrichment scatter plot of GO and KEGG classification of differential expression genes after being stimulated by stimuli. b Differential expression patterns of genes involved in

pathways of carbon fixation photosynthetic organisms and Heatmap analysis of related genes. SA20, SA100, OA100 and OA200 represented the experimental groups using 20  $\mu\text{M}$ , 100  $\mu\text{M}$  SA; and 100  $\mu\text{g mL}^{-1}$ , 200  $\mu\text{g mL}^{-1}$  oligoagar (OA) as elicitors.