Heterologous production of novel and rare C\textsubscript{30}-carotenoids using \textit{Planococcus} carotenoid biosynthesis genes

Miho Takemura\textsuperscript{1*†, Chiharu Takagi\textsuperscript{2†, Mayuri Aikawa\textsuperscript{2, Kanaho Araki\textsuperscript{2, Seon-Kang Choi\textsuperscript{3, Mitsuhiro Itaya\textsuperscript{4, Kazutoshi Shindo\textsuperscript{2 and Norihiko Misawa\textsuperscript{1}}}}}}

Abstract

Background: Members of the genus \textit{Planococcus} have been revealed to utilize and degrade solvents such as aromatic hydrocarbons and alkanes, and likely to acquire tolerance to solvents. A yellow marine bacterium \textit{Planococcus maritimus} strain iso-3 was isolated from an intertidal sediment that looked industrially polluted, from the Clyde estuary in the UK. This bacterium was found to produce a yellow acyclic carotenoid with a basic carbon 30 (C\textsubscript{30}) structure, which was determined to be methyl 5-glucosyl-5,6-dihydro-4,4′-diapolyneurosporne. In the present study, we tried to isolate and identify genes involved in carotenoid biosynthesis from this marine bacterium, and to produce novel or rare C\textsubscript{30}-carotenoids with anti-oxidative activity in \textit{Escherichia coli} by combinations of the isolated genes.

Results: A carotenoid biosynthesis gene cluster was found out through sequence analysis of the \textit{P. maritimus} genomic DNA. This cluster consisted of seven carotenoid biosynthesis candidate genes (\textit{orf1}–\textit{7}). Then, we isolated the individual genes and analyzed the functions of these genes by expressing them in \textit{E. coli}. The results indicated that \textit{orf2} and \textit{orf1} encoded 4,4′-diapophytoene synthase (CrtM) and 4,4′-diapophytoene desaturase (CrtNa), respectively. Furthermore, \textit{orf4} and \textit{orf5} were revealed to code for hydroxydiaponeurosporene desaturase (CrtNb) and glucosyltransferase (GT), respectively. By utilizing these carotenoid biosynthesis genes, we produced five intermediate C\textsubscript{30}-carotenoids. Their structural determination showed that two of them were novel compounds, 5-hydroxy-5,6-dihydro-4,4′-diapolyneurosporene and 5-glucosyl-5,6-dihydro-4,4′-diapolycopene, and that one rare carotenoid 5-hydroxy-5,6-dihydro-4,4′-diapolyneurosporene is included there. Moderate singlet oxygen-quenching activities were observed in the five C\textsubscript{30}-carotenoids including the two novel and one rare compounds.

Conclusions: The carotenoid biosynthesis genes from \textit{P. maritimus} strain iso-3, were isolated and functionally identified. Furthermore, we were able to produce two novel and one rare C\textsubscript{30}-carotenoids in \textit{E. coli}, followed by positive evaluations of their singlet oxygen-quenching activities.

Introduction

Carotenoids, essential pigments for photosynthesis, are known to protect cells from oxidative stress [1, 2]. In the domain bacteria, all of the photosynthetic bacteria, including cyanobacteria, as well as some of the non-photosynthetic bacteria, usually chemoheterotrophs, can produce carotenoids [3]. Among the non-photosynthetic bacteria, several species have been shown to produce acyclic carotenoids with a...
basic carbon 30 (C₃₀) structure instead of the typical C₄₀-carotenoids. Specifically, the carotenogenic-bacterial species of the phylum Firmicutes of low GC Gram-positive bacteria that have been reported to synthesize only C₃₀-carotenoids include Streptococcus faecium [3], Staphylococcus aureus [4], Bacillus firmus [5], Halobacillus halophilus [6, 7], Planococcus maritimus [8, 9], Sporosarcina aquimarinia [10], and Lactobacillus plantarum (recently reclassified to Lactiplantibius plantarum) [11]. C₃₀-Carotenoids have also been found in bacterial species that belong to other phyla, such as Rubritalea squalenificiens of the phylum Verrucomicrobia [12] and Methylomonas sp. strain 16a of the phylum Proteobacteria [13].

The biosynthesis of C₃₀-carotenoids has mostly been studied in Staphylococcus aureus [4, 14–16]. Farnesyldiphosphate (FPP) is converted to 4,4'-diapophytoene (dehydroquinsalene) by 4,4'-diapophytoene synthase (CrtM) and to 4,4'-diaponeurosporene by 4,4'-diapophytoene desaturase (CrtN) [16], which is likely to be the common biosynthetic route in all C₃₀-carotenogenic bacteria. Lastly, staphyloxanthin [β-d-glucosyl-1-O-(4,4'-diaponeurosporene-4-oate)-6-O-(12-methyltetradecanoate)], which has shown to confer oxidative stress tolerance on S. aureus [17], is produced from 4,4'-diaponeurosporene as the final product with the mediation of four genes that encode oxidase (CrtP), dehydrogenase (AldH), glycosyltransferase (CrtQ), and acyltransferase (CrtO) [14, 15]. As for other C₃₀-carotenogenic bacteria, a few studies have been published on their genes for metabolizing 4,4'-diaponeurosporene [13, 18–20].

Members of the genus Planococcus have been revealed to utilize and degrade solvents such as aromatic hydrocarbons and alkanes, and likely to acquire tolerance to solvents [21, 22]. A yellow marine bacterium Planococcus maritimus strain iso-3 was isolated from an intertidal sediment that looked industrially polluted, from the Clyde estuary in the UK [8]. Its yellow pigment was identified as methyl 5-glucosyl-5,6-dihydro-4,4'-diapocopenoate (methyl 5-glucosyl-5, 6-dihydro-apo-4,4'-lycopenoate), which showed potent antioxidant activity [8, 9]. Kim et al. [23] also reported that Planococcus faecalis AJ003 produced a similar C₃₀-carotenoid glycosyl-4,4'-diaponeurosporene-4'-ol-4-oate.

In this study, we tried to isolate genes involved in the production of the glucosyl C₃₀-carotenoid from P. maritimus strain iso-3, to identify the functions of the isolated genes using E. coli, and to produce novel and rare C₃₀-carotenoids with anti-oxidative activity in E. coli by combinations of the obtained genes.

Materials and methods

Bacterial strains and growth conditions

Escherichia coli K12 DH5α was used for DNA manipulation. E. coli K12 JM101(DE3) was constructed from strain JM101 using the λDE3 lysogenization kit (Merk, Darmstadt, Germany) and used for expression of the carotenoid biosynthesis genes. These E. coli strains and their transformants were maintained as the frozen stocks including 20% glycerol at −80 °C, and were grown in 2×YT medium (16 g/L of tryptone, 10 g/L of yeast extract, 5 g/L of NaCl) at 37 °C as needed. Bacillus subtilis strain MI112 was kindly provided from Dr. Atsuhiko Shimmyo (NAIST, Japan). This bacterium was similarly maintained as the frozen stock and routinely grown in LB medium (10 g/L of tryptone, 5 g/L of yeast extract, 10 g/L of NaCl) at 37 °C.

DNA isolation of Planococcus maritimus strain iso-3

Genome DNA was prepared from Planococcus maritimus strain iso-3 according to a method described by Nishida et al. [24].

Functional cloning experiments

The genome DNA was digested with BamHI, ligated with vector pGETS103, and used to transform Bacillus subtilis strain MI112 as described previously [25]. In generated genome library, one yellow colony appeared, which was found to contain a 7344-bp insert.

Inverse PCR

To isolate the flanking region of the carotenoid biosynthesis gene cluster, inverse PCR was performed. One μg of genomic DNA from the Planococcus was digested with EcoRI. Digested DNA (0.1 μg/μL) was self-ligated with ligation solution (Takara Bio, Ohtsu, Japan). PCR was performed using 50 pg of these circular DNAs as the template in 25 μL of the reaction solution with the gene-specific primers which we designed. The amplified fragments were cloned into the pBC phagemid vector (Agilent, CA, USA), transformed into E. coli (DH5α) and sequenced. The primers used are summarized in Additional file 1: Table S1.

Sequence analysis

Homology search was performed by BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Amino acid alignments and phylogenetic trees were constructed using MAFFT (http://www.mafft.cbrc.jp/).

Cloning of the carotenoid biosynthesis genes from P. maritimus strain iso-3

Based on the sequences obtained, the primers containing the restriction sites were designed as shown in Additional
Expression of the Planococcus carotenoid biosynthesis genes in Escherichia coli

Plasmids used in this study are summarized in Additional file 1: Fig. S1. Firstly, we constructed the plasmid pAC-HI which contained the Haematococcus pluvialis IDI (isopentenyl diphosphate isomerase) gene between the tac promoter (Ptac) and the rrnB terminator (TrrnB) in the pACYC184 vector. The exogenous expression of the IDI gene in E. coli results in the increase of carotenoid production [26]. Then, the coding region of the Planococcus orf3 was inserted into the plasmid pAC-HI. The resultant plasmid was named as pAC-HIO3. The coding regions of the orf1, orf2, orf4 were also inserted into the pAC-HIMF (same as pAC-HIMNO7) and the obtained plasmids were pAC-HIMNO2, pAC-HIMNO4, and pAC-HIMNO7, respectively. The orf2 and orf4 were also inserted into the pAC-HIMNF (same as pAC-HIMNO7) and the obtained plasmids were pAC-HIMNFO2 and pAC-HIMNFO4, respectively. The orf5 was further inserted into the pAC-HIMNFnB (same as pAC-HIMNFnB04) and this plasmid was named as pAC-HIMNFnB05 (renamed pAC-HIMNFnBG). All plasmids were independently introduced into the wild type E. coli (JM101 (DE3)). Each transformed E. coli was grown in 2YT medium at 37 °C. Next day, this culture was inoculated with JM101 (DE3). Each transformed E. coli cultures were centrifuged and cell pellets were extracted with methanol (MeOH) using mixer for 5 min. Tris–HCl (50 mM, pH 7.5) and 1 M NaCl were added and mixed. Then chloroform was added to the mixture and mixed for 5 min. After centrifugation, the chloroform phase was collected and dried by centrifugal evaporation. Dried residues were re-suspended with ethyl acetate (EtOAc), and applied to HPLC with a Waters Alliance 2695–2996 (PDA) system (Waters, Milford, MA, USA). HPLC was carried out according to the method described [28] using TSKgel ODS-80Ts (4.6 × 150 mm, 5 μm; Tosoh, Tokyo, Japan). Briefly, the extract was eluted at a flow rate of 1.0 mL/min at 25 °C with solvent A [water (H2O)-MeOH, 5:95] for 5 min, followed by a linear gradient from solvent A to solvent B (tetrahydrofuran-MeOH, 3:7) for 5 min, and solvent B alone for 8 min. The produced carotenoids were identified by comparing both retention times and absorption spectra with those of our authentic standards. When the produced carotenoids are not compounds in our authentic standards as described in the following sections, we isolated the produced carotenoids and determined their structures using HRESI-MS (high resolution electrospray ionization-mass spectrometry) and NMR (nuclear magnetic resonance) analyses.

Isolation of 5-hydroxy-5,6-dihydro-4,4'-diaponeurosporone (3)

The transformed E. coli cells carrying crtM, crtNa, and cruF were collected by centrifugation from 4 L culture, and extracted with 50 mL acetone-MeOH (7:2) and dichloromethane (CH2Cl2) (×2) with sonication in a stepwise manner. The combined extract (150 mL) was concentrated to small volume in vacuo, and partitioned between EtOAc/H2O (each 100 mL). The n-hexane layer evaporated to dryness (54.2 mg) and subjected to silica gel chromatography (20 mm × 200 mm) using n-hexane–acetone (8:1), and fractionated by 10 mL. The fractions 15–30 containing a yellow compound [Rf 0.4 by silica gel TLC using developing solvent n-hexane–acetone (4:1)] were collected and concentrated to afford a yellow oil (27.1 mg). The yellow oil was applied on preparative HPLC (column: Develosil C-30-UG-5 (20 × 250 mm, Nomura Chemical, Co. Ltd, Aichi, Japan), solvent: CH2Cl2-CH3CN (4:6), flow rate: 8.0 mL/min, detect: PDA 250–700 nm), and a yellow peak eluted at 15.0 min was collected and concentrated to afford pure 5-hydroxy-5,6-dihydro-4,4'-diaponeurosporone (3, 13.9 mg).

Isolation of 5-glucosyl-5,6-dihydro-4,4'-diapolyepocene (5)

The transformed E. coli cells carrying crtM, crtNa, cruF, crtnB, and GT were collected by centrifugation from 4 L culture, and extracted with acetone-MeOH (7:2) and CH2Cl2-MeOH (1:1) (×2) with sonication in a stepwise manner. The extract (150 mL) was concentrated to small volume in vacuo, and partitioned between EtOAc/H2O (each 100 mL). The EtOAc layer evaporated to dryness (243.9 mg) and subjected to silica gel chromatography (25 mm × 190 mm) using CH2Cl2-MeOH (10:1) and fractionated by 10 mL. The fractions 13–15 containing
an UV compound [Rf 0.36 by silica gel TLC using developing solvent CH2Cl2-MeOH (10:1)] to afford a yellow oil (11.2 mg). The yellow oil was applied on preparative HPLC [column: Develosil C-30-UG-5 (20 × 250 mm), solvent: CH2Cl2-MeOH (1:4), flow rate: 8.0 mL/min, detect: PDA 250–700 nm], and a yellow peak eluted at 17.0 min was collected and concentrated to afford pure 5-glucosyl-5,6-dihydro-4,4′-diapolyycopene (5, 1.4 mg).

Singlet oxygen-quenching activity
For the measurement of singlet oxygen-quenching activity, 80 μL of 25 μM methylene blue and 100 μL of 0.24 M linoleic acid, with or without 40 μL of carotenoid (final concentration, 1–100 μM; each dissolved in ethanol), were added to 5 mL glass test tubes. The tubes were mixed well and were illuminated at 7000 lx at 22 °C for 3 h in a styrofoam box. Then, 50 μL of the reaction mixture was removed and diluted to 1.5 mL with ethanol, and OD235 was measured to estimate the formation of conjugated dienes [29]. The OD235 in the absence of carotenoids relative to this reference value. The activity was indicated as the IC50 value, which represents the concentration at which 50% inhibition was observed.

Results
Isolation of the carotenoid biosynthesis gene cluster from P. maritimus strain iso-3
A yellow colony including a 7.3-kb genomic insert from P. maritimus strain iso-3 was obtained by functional cloning experiments using Bacillus subtilis as the host. Sequence analysis of this insert showed the presence of seven open reading frames (orf1–7), named orf2–8 (Fig. 1). Six orfs, orf2–7, were predicted to be involved in carotenoid biosynthesis, and the remaining orf8 was unlikely to be involved in it. When the 7.3-kb (7344-bp) DNA fragment was expressed in E. coli, the recombinant cells did not become yellow. This DNA fragment contains the 4,4′-diapophytoene synthase (crtM) gene, but it does not include the 4,4′-diapophytoene desaturase (crtNa) gene, as shown later. It is thus likely that the recombinant E. coli produced only the colorless 4,4′-diapophytoene. On the other hand, the B. subtilis used as the host retained 4,4′-diapophytoene desaturase activity; therefore, the recombinant Bacillus cells turned yellow.

The complete genome sequences of several Planococcus, such as P. plakortidis (NZ_CP016539.2), P. faecalis (NZ_CP019401.1), and P. halocrytophilus (NZ_CP016537.2), have been determined. Each of the sequences contains a similar carotenoid biosynthesis gene cluster that includes one more carotenoid biosynthesis gene other than the orfs2–8. Then, we performed inverse PCR to isolate the flanking region of orfs2–8. As a result, the 2.1-kb fragment adjacent to orf2 was isolated and the additional orf, named orf1, was found. In total, a 9.4-kb genomic DNA was analyzed to find out that eight orfs existed (Fig. 1). The accession no. of the sequences of this genomic DNA is LC620265.

Sequence analysis of the carotenoid biosynthesis gene candidates
Next, we performed a sequence analysis of orf1–8. Homology searches suggested that seven orfs, orf1–7, were involved in carotenoid biosynthesis, but not orf8, which was homologous to the genes encoding the MurR/RpiR family of transcriptional regulators. The predicted amino acid sequences of Orf1, Orf2, and Orf4 were homologous to that of diapophytoene desaturase. Orf1 was most homologous, displaying approximately 61% identity to the H. halophilus CrtNa. Whereas Orf2 and Orf4 were homologous to the H. halophilus CrtNc and CrtNb, displaying about 50 and 52% identity, respectively (Fig. 2A). On the other hand, Orf3 showed homology to diapophytoene synthase (CrtM) (Fig. 2B). Meanwhile, Orf5 and Orf6 were similar to glycosyltransferase and acyltransferase, respectively. Lastly, Orf7 was homologous to the genes that were annotated as carotenoid biosynthesis genes but with unknown functions. However, we found that the Orf7 was slightly similar, at 22% identity, to CruF (1′-hydroxylase).

Orf3 acts as a diapophytoene synthase
To investigate the function of orf3 which was homologous to diapophytoene synthase, we made the plasmid pAC-HIO3 which included H. pluvialis IDI and orf3, and introduced it into the E. coli (JM101(DE3)). Wild type E. coli cannot produce diapophytoene from farnesyl pyrophosphate (FPP). On the other hand, a new peak (named 1) was observed in the E. coli expressing pAC-HIO3 with IDI and orf3 (Fig. 3A). Then, the produced compound 1 was purified from the cells via acetone-MeOH extraction, n-hexane/90% MeOH partition, and silica gel column chromatography and analyzed by ESI–MS (+), 1H,
and $^{13}$C NMR spectra. From these spectra, the produced compound was identified as 15-cis-4,4'$\text{′}$-diapophytoene (carotenoid 1) (Fig. 4) [30]. Thus, orf3 was confirmed to encode a 15-cis-4,4'$\text{′}$-diapophytoene synthase (CrtM).

**Orf1 acts as a diapophytoene desaturase**

Because orf1, orf2, and orf4 were found to be homologous to phytoene desaturase, the catalytic activities of these Orfs were examined by constructing the plasmids pAC-HIMO1, pAC-HIMO2, and pAC-HIMO4 containing *H. pluvialis* IDI, *Leuconostoc mesenteroides* crtM, and each orf. The recombinant *E. coli* expressing pAC-HIMO1 generated a new carotenoid product, as indicated by the new peak (2) (Fig. 3B). On the other hand, the introductions of pAC-HIMO2 and pAC-HIMO4 did not affect the carotenoid profiles (Fig. 3B). We
Takemura et al. Microb Cell Fact (2021) 20:194

Fig. 3  Functional analysis of orfs. A HPLC chromatogram of cell extracts of the E. coli expressing pAC-HIMO3. B HPLC chromatograms of the extracts of the E. coli expressing the plasmids pAC-HIMO1 (upper), pAC-HIMO2 (middle), and pAC-HIMO4 (lower). C HPLC chromatograms of the extracts of the E. coli expressing the plasmids pAC-HIMNO2 (upper), pAC-HIMNO4 (middle), and pAC-HIMNO7 (lower). D HPLC chromatograms of the extracts of the E. coli expressing the plasmids pAC-HIMNO2 (upper) and pAC-HIMNO4 (lower). E HPLC chromatogram of the extracts of the E. coli expressing the plasmid pAC-HIMNFbO5. F The spectra correspond to the peak 1: 4,4′-diapophytoene, peak 2: 4,4′-diaponeurosporene, peak 3: 5-hydroxy-5,6-dihydro-4,4′-diaponeurosporene, peak 4: 5-hydroxy-5,6-dihydro-4,4′-diapolyopene, peak 5: 5-glucosyl-5,6-dihydro-4,4′-diapolyopene. Inset show the bacterial cell pelette of each recombinant E. coli.

cultured the E. coli expressing pAC-HIMO1 and purified the carotenoid 2 using the same process described in the previous section. The ESI–MS (+), 1H, and 13C NMR spectral analyses indicated carotenoid 2 as 4,4′-diaponeurosporene (Fig. 4) [31]. These results suggest that the orf1, but not orf2 or orf4, acts as a diapophytoene desaturase (CrtNa). However, the activity of Orf1 was low in E. coli; therefore, we use L. mesenteroides crtN for further experiments.

Orf7 acts as a diaponeurosporene hydratase

The synthetic pathway from FPP to diaponeurosporene is thought to be general, but the reaction after diaponeurosporene varies. So, we investigated the activities of Orf2, Orf4, and Orf7 by constructing the plasmids, pAC-HIMNO2, pAC-HIMNO4, and pAC-HIMNO7, which contained H. pluvialis IDI, L. mesenteroides crtM, L. mesenteroides crtN and each orf. When pAC-HIMNO2 or pAC-HIMNO4 was introduced into the E. coli, it only
produced 4,4′-diaponeurosporene (Fig. 3C). In the E. coli expressing pAC-HIMNO7 (orf7), a new peak (3) was detected, which could not be identified using our standards, in addition to 4,4′-diaponeurosporene (Fig. 3C). Thus, we cultured this recombinant E. coli and purified the new carotenoid 3 (13.9 mg) as orange powder as described in the Material and Methods section. HRESI-MS (+) analysis ([C₃₀H₄₄ONa] m/z 443.33015 (M+Na)+, calcd. for 443.32896) determined the molecular formula of peak 3 as C₃₀H₄₄O (4,4′-diaponeurosporene (2 + H₂O). Detailed 1D NMR (¹H and ¹³C) and 2D NMR [¹H–¹H COSY (correlation spectroscopy), HMQC (heteronuclear multiple quantum correlation), and HMBC (heteronuclear multiple bond correlation)] spectral analyses of 3 showed the ¹H and ¹³C signals of carotenoids 2 and 3 to be quite similar except the 5,6Δ signals in carotenoid 2 and one methylene [H-6 (δH 1.42–1.49) and C-6 (δC 43.4)] and one sp³ quaternary carbon [C-5 (δC 80.0) signals in carotenoid 3 (Additional file 1: Fig. S2)]. Finally, the ¹H–¹H vicinal spin network of H-6–H-8 observed in ¹H–¹H COSY, and ¹H–¹³C long-range couplings from H-4 and H-18 to C-5 and C-6 observed in HMBC indicated the structure of carotenoid 3 as 5-hydroxy-5,6-dihydro-4,4′-diaponeurosporene (Fig. 4). Carotenoid 3 was a new one according to the CAS database. The assigned ¹H and ¹³C signals in 3 were listed in Table 1. These results indicate that orf7 encodes a diaponeurosporene hydratase; thus, we named orf7 as cruF, encoding a γ-carotene 1′ hydratase involved in myxoxanthophyll biosynthesis in Synechococcus. We also renamed the plasmid pAC-HIMNO7 to pAC-HIMNF.

**Orf4 acts as a 5-hydroxy-5,6-dihydro-4,4′ diaponeurosporene desaturase**

Since the subsequent reaction of carotenoid synthesis was expected to be desaturation, we constructed the plasmids pAC-HIMNO2 and pAC-HIMNO4. When pAC-HIMNO2 was introduced into E. coli, only the production of 5-hydroxy-5,6-dihydro-4,4′-diaponeurosporene was detected (Fig. 3D). But in the E. coli expressing pAC-HIMNO4, a new peak (4) was observed near the peak of 5-hydroxy-5,6-dihydro-4,4′-diaponeurosporene (Fig. 3D). The new
compound 4 was identified as 5-hydroxy-5,6-dihydro-4,4′-diaponeurosporene, which was a rare compound in nature, using HPLC analysis with the standards (Fig. 4) [6, 7]. This data suggested that orf4 encodes a carotenoid desaturase; thus, we named orf4 as crtNb and changed the name of the plasmid pAC-HIMNFO4 to pAC-HIMNFNb.

Table 1. 1H and 13C NMR data for 5-hydroxy-5,6-dihydro-4,4′-diaponeurosporene (3) in CDCl3 and 5-glucosyl-5,6-dihydro-4,4′-diaplycopene (5) in DMSO-d6

| Position | 3 | | 5 | |
|----------|---|---|---|---|
|          | δH | δC | δH | δC |
| 4        | 1.22 (3H, s) | 29.3 (q) | 1.14 (3H, s) | 26.4 (q) |
| 5        | 1.42–1.49 (2H) | 43.4 (t) | 2.32 (2H) | 45.1 (t) |
| 6        | 1.46–1.52 (2H) | 22.6 (q) | 5.88 (1H, m) | 126.9 (d) |
| 7        | 2.11 (1H, dd, 7.0, 7.3) | 40.5 (t) | 6.14 (1H, d, 16.4) | 137.0 (d) |
| 8        | 139.5 (s) | 135.6 (s) |
| 9        | 5.96 (1H, d, 8.9) | 126.0 (d) | 6.13 (1H, d, 10.0) | 130.3 (d) |
| 10       | 6.49 (1H, dd, 8.9, 14.9) | 125.0 (d) | 6.62 (1H, dd, 10.0, 15.9) | 125.5 (d) |
| 11       | 2.11 (1H, dd, 7.0, 7.3) | 40.5 (t) | 6.14 (1H, d, 16.4) | 137.0 (d) |
| 12       | 136.0 (s) | 136.1 (s) |
| 13       | 6.19 (1H, m) | 131.5 (d) | 6.32 (1H, m) | 132.8 (d) |
| 14       | 6.60 (1H, m) | 130.1 (d) | 6.69 (1H, m) | 130.5 (d) |
| 15       | 1.22 (3H, s) | 29.3 (q) | 1.14 (3H, s) | 26.8 (q) |
| 16       | 1.82 (3H, s) | 16.8 (q) | 1.87 (3H, s) | 13.0 (q) |
| 17       | 1.96 (3H, s) | 12.8 (q) | 1.92 (3H, s) | 12.7 (q) |
| 18       | 1.82 (3H, s) | 18.6 (q) | 1.78 (3H, s) | 26.2 (q) |
| 19       | 136.0 (s) | 135.6 (s) |
| 19′      | 5.93 (1H, d, 9.6) | 126.1 (d) | 5.91 (1H, d, 10.0) | 126.9 (d) |
| 20       | 6.47 (1H, dd, 9.6, 15.2) | 124.8 (d) | 6.46 (1H, dd, 10.0, 14.9) | 125.1 (d) |
| 21       | 6.22 (1H, d, 15.2) | 135.0 (d) | 6.22 (1H, d, 14.9) | 169.1 (s) |
| 22       | 136.4 (s) | 136.4 (s) |
| 23       | 6.25 (1H, d, 11.8) | 132.6 (d) | 6.21 (1H, d, 10.0) | 131.6 (d) |
| 24       | 6.61 (1H, dd, 11.8, 14.9) | 124.9 (d) | 6.64 (1H, dd, 10.0, 15.2) | 125.4 (d) |
| 25       | 6.35 (1H, d, 14.9) | 137.3 (d) | 6.37 (1H, d, 15.2) | 137.2 (d) |
| 26       | 136.2 (s) | 136.4 (s) |
| 27       | 6.19 (1H, m) | 131.5 (d) | 6.32 (1H, m) | 132.6 (d) |
| 29       | 6.60 (1H, m) | 129.5 (d) | 6.69 (1H, m) | 130.5 (d) |
| 30       | 1.82 (3H, s) | 26.3 (q) | 1.78 (3H, s) | 18.6 (q) |
| 31       | 1.95 (3H, s) | 12.8 (q) | 1.92 (3H, s) | 12.7 (q) |
| 32       | 1.96 (3H, s) | 12.9 (q) | 1.92 (3H, s) | 13.0 (q) |
| 33       | 4.32 (1H, d, 7.9) | 97.4 (d) |
| 34       | 2.89 (1H, dd, 7.9, 8.0) | 73.7 (d) |
| 35       | 3.15 (1H, dd, 8.0, 8.8) | 76.8 (d) |
| 36       | 3.05 (1H, m) | 70.4 (d) |
| 37       | 3.05 (1H, m) | 71.6 (d) |
| 38       | 3.40 (1H, m) | 3.61 (1H, m) | 61.4 (t) |

a, b, c, d Interchangeable

Orf5 acts as a glycosyltransferase
A novel carotenoid previously isolated from Planococcus iso-3, methyl 5-glucosyl-5,6-dihydro-4,4′-diaplycopene, was glycosylated. Because orf5 showed homology to glycosyltransferase, we investigated the activity of Orf5. The introduction of the
plasmid pAC-HIMNFNbO5 into E. coli resulted in the detection of a new peak (5), which could not be identified using the standards (Fig. 3E). Thus, we cultured this recombinant E. coli and purified the new carotenoid (5, 13.9 mg) as orange powder. HRESI-MS (+) analysis [C36H52O6Na m/z 603.36794 (M+Na)+, calcd. for 603.36616] determined the molecular formula of carotenoid 5 as C36H52O6 (5-hydroxy-5,6-dihydro-4,4′-diapolyycopene 4 + C6H13O5–H2O). Detailed 1D NMR (1H and 13C) and 2D NMR (1H–1H COSY, HSQC, and HMBC) spectral analyses of 5 demonstrated that the structure of 5-hydroxy-5,6-dihydro-4,4′-diapolyycopene (carotenoid 4) was present also in carotenoid (5) and the attachment of a hexose at C-5 of carotenoid 4. The hexose in carotenoid 5 was shown to be β-glucose because all vicinal coupling constants among H-1′ to H-5′ were large (7.9–9.1 Hz) in the 1H NMR spectrum of acetylated carotenoid 5 in CDCl3 (Additional file 1: Fig. S3). From these observations, carotenoid 5 was determined to be 5-glucosyl-5,6-dihydro-4,4′-diapolyycopene (Fig. 4). Carotenoid 5 was a new compound according to the CAS database. The assigned 1H and 13C signals in 5 were listed in Table 1. These results indicate that the orf5 encodes a glycosyltransferase and we renamed pAC-HIMNFNbO5 as pAC-HIMNFNbG.

**Singlet oxygen-quenching activity of carotenoids 1 to 5**

In the previous study, we reported 1O2-quenching activity of methyl 5-glucosyl-5,6-dihydro-4,4′-diapolyycopene [8, 9]. Thus, we evaluated 1O2-quenching activities of the intermediates, carotenoids 1–5 isolated in this study. These five carotenoids showed milder 1O2-quenching activities than methyl 5-glucosyl-5,6-dihydro-4,4′-diapolyycopene (Table 2).

### Table 2 Singlet oxygen quenching activity of the intermediate carotenoids

| Carotenoid | Singlet oxygen quenching IC50 (μM) |
|------------|-----------------------------------|
| 15-cis-4,4′-Diapophytoene (1) | > 100 |
| 4,4′-Diaponeurosporene (2) | 45 |
| S-Hydroxy-5,6-dihydro-4,4′-diaponeurosporene (3) | 56 |
| S-Hydroxy-5,6-dihydro-4,4′-diapolyycopene (4) | 30a |
| S-Glucosyl-5,6-dihydro-4,4′-diapolyycopene (5) | 30 |
| Methyl 5-glucosyl-5,6-dihydro-4,4′-diapolyycopene | 5.1b |
| Astaxanthin | 3.7b |

a Cited from the data of Shindo et al. [9]

**Discussion**

Until now, there are a few studies on the genes involved in the carotenoid biosynthesis in *Planococcus*. Recently, Lee et al. [19] reported the C30-carotenoid biosynthesis genes of *Planococcus faecalis* AJ003T, which mediated glycosyl-4,4′-diaponeurosporene-4′-ol-4-oic acid. Herein, we report the carotenoid biosynthesis genes in *Planococcus maritimus* strain iso-3, which produces the rare C30 carotenoid, methyl 5-glucosyl-5,6-dihydro-4,4′-diapolyycopenoate.

The whole-genome sequences of several kinds of *Planococcus* were available in the genome database (https://www.ncbi.nlm.nih.gov/genome/). All of them, including *P. faecalis* AJ003T, had identical carotenoid biosynthesis gene clusters to that found in this study [19, 23]. Similar gene clusters have also been reported in *Halobacillus halophilus*, *Staphylococcus aureus*, *Bacillus indicus*, and *Bacillus firmus*, which all belong to the order of Bacillales (Fig. 5) [14, 20]. In particular, *H. halophilus* has nearly the same gene organization as *Planococcus* regarding gene order and transcriptional direction, whereas they fall into distinct families, the Bacillaceae and Planococcaceae families, respectively [18]. On the other hand, the gene organization is much different between *B. indicus* and *B. firmus* [20]. These results suggested that the ancestor of Bacillales retained the same carotenoid biosynthesis gene cluster, and during evolution, the gene organizations have been rearranged independently.

Functional analysis of the genes isolated from *P. maritimus* indicated that the *crtM* and *crtNa* genes encoded proteins with enzymatic activities as predicted from the sequence homology analysis. On the other hand, the enzyme encoded by *crtNb* was found to generate a product different from those by other *crtNb* gene products. For example, Steiger et al. [20] demonstrated that the proteins encoded by *B. indicus* and *B. firmus* *crtNb* functioned as aldehyde synthases for utilizing 4,4′-diapolyycopene or 4,4′-diaponeurosporene. In contrast, the functions of the *crtNb* genes in *H. halophilus* and *P. fecalis* sp. Nov have not been revealed [18, 19, 32]. To discuss the evolution of *crtNb* gene, it will be required to characterize the function of various *crtNb* genes.

We also found a carotenoid 1,2-hydratase gene that showed a slight similarity to *cruF*, initially found in *Deinococcus* [33]. *CruF* catalyzes the reaction from γ-carotene to 1′-OH-γ-carotene and is required to produce deinoxanthin in *Deinococcus*. The homologous genes with this were present not only in the *Planococcus* genomes but also in the *Halobacillus* genomes at the same position, whereas their functions have not been elucidated. As for other-type hydratase genes, *crtC* has been reported in *Rubrivivax gelatinosus* [34, 35]. However, the Orf7 did not show any similarity to the CrtC protein.
Here, the entire carotenoid biosynthetic pathway was almost elucidated using complementation analysis with *E. coli* (Fig. 4). After 15-cis-4,4′-diapophytoene (1) is synthesized, it is desaturated to 4,4′-diaponeurosporene (2), hydrated to produce 5-hydroxy-5,6-dihydro-4,4′-diaponeurosporene (3), and then desaturated to produce 5-hydroxy-5,6-dihydro-4,4′-diapolycopene (4). Finally, 5-glucosyl-5,6-dihydro-4,4′-diapolycopene (5) is produced by the glycosylation of a terminal group of 5-hydroxy-5,6-dihydro-4,4′-diapolycopene (4). However, according to previous reports, the final product is methyl 5-glucosyl-5,6-dihydro-apo-4,4′-lycopenoate, indicating that more reactions should occur. We found *orf6*, homologous to acyltransferase in the carotenoid gene cluster and speculated that it catalyzed the esterification. On the other hand, the genes involved in generating carotenoid carboxylic acid were not identified in the gene cluster. Indeed, the introduction of *orf2* into the *E. coli* carrying the plasmid pAC-HIMNFNb or pAC-HIMNFNbO5 did not affect the carotenoid profile (Additional file 1: Fig. S4). Lee et al. [19] reported that *crtP* and *aldH* genes coded for 4,4-diaponeurosporene oxidase and aldehyde dehydrogenase, respectively; that these genes were located away from the carotenoid gene cluster in *P. faecalis*. Thus, it is necessary to analyze the genes outside of the carotenoid synthesis gene cluster in the *Planococcus* iso-3 genome.

We evaluated the 1O2-quenching activities of carotenoids 1–5, and compared their activities with 5-glucosyl-5,6-dihydro-4,4′-diapolycopene (4). It is very interesting that the compound produced through more advanced biosynthetic pathway shows more potent 1O2-quenching activity. *P. maritimus* strain iso-3 may have completed this biosynthetic pathway to obtain a strong singlet oxygen scavenger.

**Conclusions**

We isolated seven carotenoid biosynthesis gene candidates from *P. maritimus* strain iso-3, and functionally identified five of the seven candidates as the *crtM*, *crtNa*, *crtNb*, *cruF* and *GT* genes. By utilizing these genes, we produced two novel C30-carotenoids, 5-hydroxy-5,6-dihydro-4,4′-diaponeurosporene (3) and 5-glucosyl-5,6-dihydro-4,4′-diapolycopene (5), in addition to one rare C30-carotenoid 5-hydroxy-5,6-dihydro-4,4′-diapolycopene (4), in *E. coli*. The yield of them were
estimated to be 10.8 mg/L, 2.2 mg/L and 2.6 mg/L for 3, 4 and 5, respectively. These compounds were shown to exhibit moderate singlet oxygen-quenching activities. In the near future, by using new carotenogenic genes, other novel and rare C₃₀-carotenoids will be produced in E. coli.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12934-021-01683-3.

Additional file 1: Figure S1. The scheme of the constructed plasmids. Figure S2. NMR spectra of 5-hydroxy-5,6-dihydro-4,4′-diaponeurosporene (3) in CDCl₃. Figure S3. NMR spectra of 5-glucoyl-5,6-dihydro-4,4′-diapolycolpenone (5) in DMSO-d₆. Figure S4. Functional analysis of orf2. Table S1. Primers used in this study. Table S2. Accession numbers of the genes used in Fig. 2A. Table S3. Accession numbers of the genes used in Fig. 2B.

Acknowledgements
We thank Chisako Fuchimoto and Miyuki Murakami for their technical assistances.

Authors’ contributions
MT, CT, KS and NM designed this study and interpreted the data. SC, MI and MT cloned the genes and analyzed the carotenoids from E. coli. All authors read and approved the final manuscript.

Availability of data and materials
The Planococcus genomic sequences decoded during this study are available in GenBank with the accession no. LC620265. The sequences of the plasmids, pAC-Hl, pAC-HIM, pAC-HIMN, pAC-HIMNfB, and pAC-HIMN-FNBG are available in GenBank with the accession no. LC635748, LC647532, LC647533, LC647534, LC647535 and LC647536, respectively. Data and materials used during this study are available from the corresponding author on reasonable request.

Declarations
Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
1Research Institute for Bioresources and Biotechnology, Ishikawa Prefectural University, 1-308 Suematsu, Nojihiga, Ishikawa 921-8836, Japan. 2Department of Food and Nutrition, Japan Women’s University, Bunkyo-ku, Tokyo 112-8681, Japan. 3Department of Agriculture and Life Industry, Kangwon National University, Cheoncheon, Republic of Korea. 4Department of Biomedical Engineering, Graduate School of Science and Technology, Shinshu University, Wakasa 4-17-1, Nagano 380-8553, Japan.

Received: 6 July 2021 Accepted: 22 September 2021 Published online: 09 October 2021

References
1. Krinsky NI. Antioxidant functions of carotenoids. Free Radic Biol Med. 1989;7:617–35.
2. Stahl W, Sies H. Bioactivity and protective effects of natural carotenoids. Biochim Biophys Acta. 2004;1700:101–7.
3. Britton G, Liaoen-Jensen S, Pander H. Carotenoid handbook. Basel: Birkhäuser Verlag, 2004.
4. Marshall JH, Wilmoth G. Proposed pathway of triterpenoid carotenoid biosynthesis in Staphylococcus aureus: evidence from a study of mutants. J Bacteriol. 1981;147:914–9.
5. Steiger S, Perez-Fons L, Fraser PD, Sandmann G. Biosynthesis of a novel C30 carotenoid in Bacillus firmus isolates. J Appl Microbiol. 2012;113:888–95.
6. Osaka A, Ishii Y, Sasamura N, Morita M, Kocher S, Müller V, Sandmann G, Shindo K. Hydroxy-3,4-dehydro-apo-8′-lycopene and methyl hydroxy-3,4-dehydro-apo-8′-lycopene, novel C₃₀-carotenoids produced by a mutant of marine bacterium Halobacillus halophilus. J Antibiot. 2010;63:291–5.
7. Osaka A, Ishii Y, Sasamura N, Morita M, Kocher S, Müller V, Sandmann G, Shindo K. Hydroxy-5,6-dihydro-apo-4,4′-lycopene and methyl 5-hydroxy-5,6-dihydro-apo-4,4′-lycopene, novel C₃₀-carotenoids produced by a mutant of marine bacterium Halobacillus halophilus. J Antibiot. 2014;67:733–5.
8. Shindo K, Endo M, Miyake Y, Wakisugi K, Morritt D, Bramley PM, Fraser PD, Kasai H, Misawa N. Methyl hydroxy-3,4-dehydro-apo-8′-lycopene, a novel antioxidant glyco-C₃₀-carotenoid acid produced by a marine bacterium Planococcus marinus. J Antibiot. 2008;61:729–35.
9. Shindo K, Endo M, Miyake Y, Wakisugi K, Morritt D, Bramley PM, Fraser PD, Kasai H, Misawa N. Methyl 5-hydroxy-5,6-dihydro-apo-4,4′-lycopene, a novel antioxidant glyco-C₃₀-carotenoid acid produced by a marine bacterium Planococcus marinus. J Antibiot. 2014;67:731–2.
10. Steiger S, Perez-Fons L, Fraser PD, Sandmann G. The biosynthetic pathway to a novel derivative of 4,4′-diapolycolpenone-4,4′-oate in a red strain of Sprosarcina aquimarinus. Arch Microbiol. 2012;194:779–84.
11. Garrido-Fernández J, Malonda-Barragán A, Caballero-Guerrero B, Homero-Méndez D, Ruiz-Barba JL. Carotenoid production in Lactobacillus plantarum. Int J Food Microbiol. 2010;140:34–9.
12. Shindo K, Asagi E, Sano A, Hotta E, Mimura N, Mikami K, Tamesada E, Misawa N, Maoka T. Diapolycolpenedioic acid xylosyl esters A, B, and C, novel antioxidant glyco-C₃₀-carotenoid acids produced by a new marine bacterium Rubribacteria squamifaciens. J Antibiot. 2008;61:185–91.
13. Tao L, Schenze A, Odom JM, Cheng Q. Novel carotenoid oxalide involved in biosynthesis of 4,4′-diapoytolene dialdehyde. Appl Environ Microbiol. 2005;71:3294–301.
14. Kim SH, Lee PC. Functional expression and extension of Staphylogoccal staphyloxanthin biosynthesis pathway in Escherichia coli. J Biol Chem. 2012;287:21575–83.
15. Pelz A, Wieland K-P, Pitizbach K, Hentschel KP, Alberl K, Götz F. Structure and biosynthesis of staphyloxanthin from Staphylococcus aureus: evidence from a study of mutants. J Bacteriol. 1981;147:914–9.
16. Clauditz A, Resch A, Wieland KP, Poralla K, Gotz F. Genetic and biochemical analyses of the biosynthesis of the yellow carotenoid 4,4′-diapeneurosporene of Staphylococcus aureus. J Bacteriol. 1994;176:7719–26.
17. Clauditz A, Resch A, Wieland KP, Poralla K, Gotz F. Staphyloxanthin plays a role in the fitness of Staphylococcus aureus and its ability to cope with oxidative stress. Infect Immun. 2006;74:4950–3.
18. Kocher S, Breitenbach J, Müller V, Sandmann G. Structure, function and biosynthesis of carotenoids in the moderately halophilic bacterium Halobacillus halophilus. Arch Microbiol. 2009;191:95–104.
19. Lee JH, Kim JW, Lee PC. Genome mining reveals two missing CrtP and CrtE, Misawa N, Maoka T. Diaplycolpenedioic acid xylosyl esters A, B, and C, novel antioxidant glyco-C₃₀-carotenoid acids produced by a new marine bacterium Rubribacteria squamifaciens. J Antibiot. 2008;61:185–91.
20. Steiger S, Perez-Fons L, Cutting SM, Fraser PD, Sandmann G. Annotation and functional assignment of the genes for the C30 carotenoid pathways from the genomes of two bacteria: Bacillus indicus and Bacillus firmus. Microbiology. 2015;161:194–202.
21. Engelhardt MA, Daly K, Swanell RJL, Head IM. Isolation and characterization of a novel hydrocarbon-degrading, Gram-positive bacterium,
isolated from intertidal beach sediment, and description of Planococcus alkanoclasticus sp. Nov. J Appl Microbiol. 2001;90:237–47.

22. Li H, Liu YH, Luo N, Zhang XY, Luan TG, Hu JM, Wang ZY, Wu PC, Chen MJ, Lu JQ. Biodegradation of benzene and its derivatives by a psychrotolerant and moderately haloalkaliphilic Planococcus sp. strain ZD22. Res Microb. 2006;157:629–36.

23. Kim JW, Choi BH, Kim JH, Kang H-J, Ryu H, Lee PC. Complete genome sequence of Planococcus faecalis AJ0033, the type species of the genus Planococcus and a microbial C30 carotenoid producer. J Biotechnol. 2018;266:72–6.

24. Nishida Y, Adachi K, Kasai H, Shizuri Y, Shindo K, Sawabe A, Komemushi S, Miki W, Misawa N. Elucidation of a carotenoid biosynthesis gene cluster encoding a novel enzyme, 2,2′-β-hydroxylase, from Brevundimonas sp. strain SD212 and combinatorial biosynthesis of new or rare xanthophylls. Appl Environ Microbiol. 2005;71:4286–96.

25. Tsuge K, Ito M. Recombinational transfer of 100-kilobase genomic DNA to plasmid in Bacillus subtilis 168. J Bacteriol. 2001;183:5453–8.

26. Albrecht M, Misawa N, Sandmann G. Metabolic engineering of the terpenoid biosynthetic pathway of Escherichia coli for production of the carotenoids β-carotene and zeaxanthin. Appl Environ Microbiol. 1999;65:791–5.

27. Fraser PD, Elisabete M, Pinto S, Holloway DE, Bramley PM. Application of high-performance liquid chromatography with photodiode array detection to the metabolic profiling of plant isoprenoids. Plant J. 2000;24:139–44.

28. Yokoyama A, Miki W. Composition and presumed biosynthetic pathway of carotenoids in the astaxanthin-producing bacterium Agrobacterium aurantiacum. FEMS Microbiol Lett. 1995;128:139–44.

29. Hirayama O, Nakamura K, Hamada S, Kobayashi Y. Singlet oxygen quenching ability of naturally occurring carotenoids. Lipids. 1994;29:149–50.

30. Tiziani S, Schwartz SJ, Vodovotz Y. Profiling of carotenoids in tomato juice by one- and two-dimensional NMR. J Agric Food Chem. 2006;54:6094–100.

31. Inomata M, Hira I, Yoshida R, Ohigashi H. Biosynthesis of abscisic acid by the direct pathway via ionylideneethane in a fungus, Cerocapsor cruenta. Biosci Biotechnol Biochem. 2004;68:2571–80.

32. Kim JH, Kang HJ, Yu BJ, Kim SC, Lee PC. Planococcus faecalis sp. nov., a carotenoid-producing species isolated from stools of Antarctic penguins. Int J Syst Evol Microbiol. 2015;65:3373–8.

33. Sun Z, Shen S, Wang C, Wang H, Hu Y, Jiao J, Ma T, Tian B, Hua Y. A novel carotenoid 1,2-hydrolase (CrtF) from two species of the non-photosynthetic bacterium Deinococcus. Microbiology. 2009;155:2775–83.

34. Hiessen A, Arends IWCE, Otten LG. Biochemical characterization of the carotenoid 1,2-hydrolases (CrtF) from Rubrivivax gelatinosus and Thiocapsa roseopersicina. Appl Microbiol Biotechnol. 2011;91:1029–36.

35. Steiger S, Mazet A, Sandmann G. Heterologous expression, purification, and enzymatic characterization of the acyclic carotenoid 1,2-hydrolase from Rubrivivax gelatinosus. Arch Biochem Biophys. 2003;414:51–8.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.