



HHS Public Access

Author manuscript

Environ Microbiol. Author manuscript; available in PMC 2016 June 01.

Published in final edited form as:

Environ Microbiol. 2015 June ; 17(6): 2158–2171. doi:10.1111/1462-2920.12669.

Sioxanthin, a novel glycosylated carotenoid reveals an unusual subclustered biosynthetic pathway

Taylor K.S. Richter¹, Chambers C. Hughes¹, and Bradley S. Moore^{1,2,*}

¹Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093

²Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA 92093

Summary

Members of the marine actinomycete genus *Salinispora* constitutively produce a characteristic orange pigment during vegetative growth. Contrary to the understanding of widespread carotenoid biosynthesis pathways in bacteria, *Salinispora* carotenoid biosynthesis genes are not confined to a single cluster. Instead, bioinformatic and genetic investigations confirm that four regions of the *S. tropica* CNB-440 genome, consisting of two gene clusters and two independent genes, contribute to the *in vivo* production of a single carotenoid. This compound, namely (2'S)-1'-(β-D-glucopyranosyloxy)-3',4'-didehydro-1',2'-dihydro-φ,ψ-caroten-2'-ol, is novel and has been given the trivial name “sioxanthin”. Sioxanthin is a C₄₀-carotenoid, glycosylated on one end of the molecule and containing an aryl moiety on the opposite end. Glycosylation is unusual amongst actinomycete carotenoids, and sioxanthin joins a rare group of carotenoids with polar and non-polar head groups. Gene sequence homology predicts that the sioxanthin biosynthetic pathway is present in all of the *Salinispora* as well as other members of the family Micromonosporaceae. Additionally, this study's investigations of clustering of carotenoid biosynthetic genes in heterotrophic bacteria show that a non-clustered genome arrangement is more common than previously suggested, with nearly half of the investigated genomes showing a non-clustered architecture.

Introduction

Salinispora is a bacterial genus of obligate marine actinomycetes isolated from tropical and temperate sediments from around the globe (Jensen and Mafnas, 2006; Goo *et al.*, 2014). Like their terrestrial actinomycete cousins, the *Salinispora* are notable for their high secondary metabolic capacity and are a known source of novel bioactive chemistry, including salinosporamide A, a potent proteasome inhibitor with promising anticancer properties (Feling *et al.*, 2003; Gulder and Moore, 2010). The three identified species, *Salinispora tropica*, *Salinispora arenicola*, and *Salinispora pacifica*, are very closely related, sharing greater than 98% identity in their 16S rRNA sequence, but are differentiated by species-specific chemotype, as determined by their secondary metabolism (Jensen *et al.*, 2007). Despite differences in their specialized metabolism, the three species of *Salinispora*

*For correspondence bsmoore@ucsd.edu; Tel (1)8588226650 Fax (1)8588226652.

are visually and morphologically indistinguishable in part due to their pigmentation (Maldonado *et al.*, 2005; Ahmed *et al.*, 2013).

Pigmentation is a distinctive feature of the *Salinispora* phenotype – orange during vegetative growth, followed by the production of black spores as the culture ages. This chemical pattern is nearly ubiquitous among the *Salinispora*, but has been overlooked in studies of their biosynthetic capabilities. Presumably, the distinctive orange pigmentation in *Salinispora* is due to the accumulation of carotenoids, because in addition to the production of carotenoids in related actinomycetes, carotenoid biosynthesis gene homologs had been previously identified in the *Salinispora* genome (Udwary *et al.*, 2007; Penn *et al.*, 2009).

Carotenoids are highly conjugated, linear tetraterpenoids responsible for the majority of yellow, orange, and red pigmentation seen throughout nature (Armstrong, 1997). Though over 750 diverse carotenoid structures have been identified in nature, they share a linear, conjugated backbone and a common biosynthesis (Britton, 2004) (Figure 1). Carotenoids are built from the linear condensation of isoprene units, derived from primary metabolism (Armstrong, 1994; Kuzuyama *et al.*, 2002). Chain desaturation produces the chromophore, characteristic of carotenoids, which results in a region of delocalized electrons that are easily excited (resulting in pigmentation) or donated to quench oxidized molecules (making carotenoids effective antioxidants) (Britton, 1995). Further modifications on either end of the molecule give rise to the structural diversity seen in nature (Armstrong, 1997).

Carotenoid biosynthesis occurs in every branch of the tree of life, with the exception of animals where carotenoids are introduced through diet (Britton, 1995). Perhaps the most well-known role for carotenoids is in photosynthetic organisms, where they serve as accessory pigments in their light harvesting centers (Vershinin, 1999). The biological function of carotenoids in non-phototrophic microorganisms is primarily in the roles of oxidative stress relief and membrane stability (Britton, 1995). As powerful antioxidants, carotenoid biosynthesis has been shown to respond to activated oxygen species under conditions such as high copper concentration, light irradiation, and biofilm formation (Moraleda-Munoz *et al.*, 2005; Takano *et al.*, 2005; Takano *et al.*, 2006; Pezzoni *et al.*, 2011; Galbis-Martinez *et al.*, 2012; Zheng *et al.*, 2013). Carotenoids are often localized to cell membranes, where their influence on membrane fluidity is a function of the molecule's exact structure (Lutnaes *et al.*, 2004; Gruszecki, 2009; Diesler *et al.*, 2010).

Carotenoids are also important contributors to industry, human health and nutrition. As a natural product, carotenoids represented a \$1.2 USD billion market in 2010, with isolated compounds being used in everything from food colorants to supplements in the prevention of cancer and macular degeneration (Marz, 2011). Carotenoids serve as virulence factors in pathogenic microbes such as *Staphylococcus aureus*, where the production of carotenoids makes them less susceptible to neutrophil killing, and have been implicated in suppressing the human immune system (Liu and Nizet; 2009). Better understanding of the biosynthesis and biological functions of carotenoids may impact both the food and health industries.

The availability of microbial genome sequences provides necessary insight into the genome arrangement and evolution of secondary metabolism such as carotenoid biosynthetic

pathways. Bacterial carotenoid biosynthetic genes were first identified in *Rhodobacter capsulatus* where they form a gene cluster, in which all genes in the biosynthetic pathway occupy neighboring loci on the genome (Armstrong and Hearst, 1996). Similar arrangements have been found throughout bacteria, where clustering is the common arrangement except in the cyanobacteria (Armstrong and Hearst, 1996; Martin *et al.*, 2003; Fischbach and Voigt, 2010). Microbial secondary metabolite biosynthetic genes are generally organized into gene clusters (Fischbach *et al.*, 2008; Osbourn, 2010; O'Brien and Wright, 2011; Ziemert *et al.*, 2014). The clustering of genes in a single locus is proposed to improve coordination during horizontal gene transfer, increase regulation efficiency, shorten diffusion times for proteins finding their targets and forming complexes, and limit the probability of loss of function due to mutation (Pal and Hurst, 2004; Ballouz *et al.*; 2010).

Here, we report the isolation and characterization of a novel glycosylated carotenoid from the marine bacterial genus *Salinispora*. The chemical structure aided in uncovering the compound's biosynthetic pathway, which has an unusual, sub-clustered genomic architecture.

Results and Discussion

Identification of carotenoid biosynthetic genes

To investigate the biosynthetic origin of the orange pigment, we performed gene homology searches of known carotenoid biosynthetic genes in the *S. tropica* CNB-440 genome. Using sequences of biochemically characterized carotenoid biosynthetic enzymes as a search query, putative carotenogenesis genes were identified in four distinct regions of the *S. tropica* genome (Figure 2). Two regions had previously been identified as gene clusters, *terp1* and *terp2*, predicted to function in terpene biosynthesis (Udwary *et al.*, 2007). The *terp2* cluster, consisting of five genes (*strop4437-strop4441*), was considered to be the most likely candidate for carotenoid biosynthesis, in that it contained all the genes necessary for lycopene biosynthesis in addition to a *merR* regulator. The *terp1* cluster (*strop3244-strop3251*) is comprised of eight genes, only two of which had significant homology to known carotenoid biosynthetic genes, a phytoene dehydrogenase and polyprenyl synthetase.

Two additional genes with homology to carotenoid biosynthesis, predicted to be a *crtY* (*strop2408*; lycopene cyclase) and a *crtU* (*strop0241*; β -carotene desaturase), were found in the *S. tropica* genome. These genes were found in isolated regions of the genome without any proximity to other carotenoid biosynthetic genes (Figure 2).

To explore the involvement of the predicted genes in the biosynthesis of the *Salinispora* pigment, we performed gene inactivation experiments by PCR targeting (Eustaquio *et al.*; 2008) and analyzed the resulting mutants by pigment extraction and HPLC analysis. Both the *terp1* and *terp2* clusters contain genes with predicted roles in lycopene (5) biosynthesis. Lycopene, which is derived from the head-to-head condensation and dehydrogenation of two geranylgeranyl pyrophosphate (GGPP) (3) molecules, serves as the central precursor to the majority of carotenoids (Figure 1) (Armstrong, 1994; Britton, 1995). Thus, phytoene synthase (*strop4441*), from the *terp2* locus, and polyprenyl synthetases (*strop4440* and *strop3251*), in both *terp1* and *terp2* loci, were genetically inactivated. Chemical analysis of

the *strop4441* deletion mutant revealed a complete loss of orange pigmentation as demonstrated by both the visible phenotype, as well as the absence of metabolites detectable at 450 nm (Figure 2). The genetic interrogation of the two polyprenyl synthetases that share 67% sequence identity revealed that just the *strop4440* homolog in *terp2* was associated with visible carotenoid pigmentation, as the *strop3251* deletion mutant showed no alteration from the wild-type. These observations suggest that the *terp2* locus is largely responsible for the biosynthesis of carotenoid precursors in *Salinispora*. Further, these experiments demonstrate for the first time that carotenoids alone are responsible for vegetative pigmentation in these bacteria.

Inactivation of the non-clustered genes *strop2408* and *strop0241* resulted in the continued biosynthesis of carotenoids, but with altered structures based on HPLC analysis and, in the case of the *strop2408* mutant, a visibly altered phenotype (Figure 2). These results show that the genes found in *terp2* in addition to *strop0241* and *strop2408* are required for the biosynthesis of the *Salinispora* carotenoid.

The suite of genes identified in *Salinispora* is consistent with those found in members of the genus *Streptomyces*, where they have been shown to produce isorenieratene (**6**) (Figure 3) (Krugel *et al.*, 1999.; Takano *et al.*, 2005). The *Streptomyces* carotenoid biosynthetic genes, like most other described bacterial carotenoid pathways, comprise a single gene cluster (Krugel *et al.*, 1999; Takano *et al.*, 2005). However, an HPLC trace showing the major *Salinispora* carotenoid differed from that of an isorenieratene standard isolated from *Streptomyces coelicolor* A3(2) (Supporting information Figure S1). The shorter HPLC retention time of the *Salinispora* carotenoid indicated a more hydrophilic compound than the hydrocarbon isorenieratene. Bioinformatic predictions of the *Salinispora* carotenoid failed to predict the chemical compound, suggesting that there were additional biosynthetic genes that had been overlooked due to low sequence homology with known biosynthetic genes. To aid in the identification of these missing genes, we next set out to solve the structure of the *Salinispora* carotenoid.

Isolation and characterization of sioxanthin

Mass spectral and comprehensive NMR analyses (Supporting Information Table S1, Figures S2–S15) of the HPLC purified compound resulted in a novel glycosylated carotenoid (2'*S*)-1'-(β-D-glucopyranosyloxy)-3',4'-didehydro-1',2'-dihydro-φ,ψ-caroten-2'-ol (**7**), which we named “sioxanthin”. Structure elucidation was conducted on an acetylated derivative, which was necessary to aid in its isolation and purification (Supporting Information Figure S2). APCI-MS returned an exact mass of 937.5008 m/z (M+H)⁺ corresponding to a molecular formula of C₅₆ H₇₃ O₁₂ (δ = 1.0 ppm).

The UV/Vis spectrum for sioxanthin has peaks at 446, 472, and 504 nm (with a fine structure of %III/II 84), showing a chromophore similar to that of lycopene (Supporting Information Figure S9). Stereochemical analyses confirmed the *S* conformation of the 2'-hydroxyl and identified the glycosyl group as glucose. Though sioxanthin has never before been described, the modifications on both ends of the molecule are known in other carotenoids. The aromatic ring portion is known, for example, in isorenieratene, the primary carotenoid in many *Streptomyces* (Krugel *et al.*, 1999). The other end of the molecule

contains several modifications, namely a 1'-glycosylation, a 2'-hydroxyl, and a 3',4'-desaturation, all of which are known structural elements in pheixanthophyll, from a single strain of *Mycobacterium phlei* (Hertzber and Jensen, 1967).

Identification of additional sioxanthin biosynthesis genes

Using the structure as a guide, we identified four additional enzymatic steps required for the biosynthesis of sioxanthin – a 3',4'-desaturation, a 2'-hydroxyl group addition, and a 1'-hydroxyl group addition followed by the transfer of a glucose residue. Glycosylated carotenoids, though well studied among the cyanobacteria, are rare among heterotrophic actinobacteria. Prior to this study, glycosylated carotenoids had been identified in only seven species within the actinomycetes and, oftentimes, only within a single species in a genus (Britton, 2004; Takaichi *et al.*, 2008). These strains lack genome sequences and/or carotenoid biosynthesis characterization making them a poor resource for gene homology searches in *Salinispora*. Query sequences for these biosynthetic genes came from the phylogenetically distant phylum Deinococcus-Thermus (Tian and Hua, 2010) which have carotenoids with similar functional groups whose biosynthesis has been characterized.

Homology searches in the *S. tropica* CNB-440 genome identified four genes within the *terp1* cluster that were the most likely candidates for the missing biosynthetic steps (Supporting Information Figure S16). These *terp1* genes include a putative *crtA*-like monooxygenase (*strop3244*), a *cruF*-like 1',2'-hydratase (*strop3246*), a *cruC*-like glycosyl transferase (*strop3247*), and a *crtD*-like 3',4'-desaturase (*strop3248*), as well as several other genes with no predicted function in sioxanthin production (Figure 2).

The remaining genes in the *terp1* cluster were similarly interrogated through gene inactivation and chemical extract analyses to determine their roles in the biosynthetic pathway. Of the eight genes investigated in the *strop3244*—*strop3251* locus, carotenoid extracts of three gene knockout mutants (*strop3246*, *strop3247* and *strop3248*) showed compounds at different HPLC retention times and, in some cases, with different UV spectral features than the wild-type (Figure 2). These observations indicate that although the mutants continue to synthesize carotenoids, the molecules produced are different from the wild-type carotenoid, thereby confirming a role for these genes in sioxanthin biosynthesis. The remaining genes in the cluster (*strop3244*, *strop3245*, *strop3248*, *strop3249*, and *strop3251*) did not alter *Salinispora* pigmentation upon inactivation. These were deemed unnecessary to the sioxanthin biosynthetic pathway, despite the *strop3244* and *strop3251* homology to carotenoid biosynthetic genes.

Structure elucidation of mutant bacteria pigments and confirmation of sioxanthin biosynthetic pathway

To further confirm the biosynthetic function of the sioxanthin gene products, we analyzed the mutant carotenoid structures. MS and UV/Vis data (Supporting Information Table S2) confirmed the structures of the carotenoids produced by the mutant bacteria (Figure 3). Therefore, *strop2408* encodes a lycopene cyclase (*crtY*) that forms a cyclized end, *strop3246* encodes a 1',2'-hydratase (*cruF*) that hydroxylates C 1', *strop3247* encodes a glycosyl transferase (*cruC*) which adds a glucose moiety to the 1'-hydroxyl group, and *strop3248*

encodes a 3',4'-desaturase which extends the chromophore by adding a double bond. The *strop3247* deficient mutant produced a novel structure, a sioxanthin aglycone, which was further confirmed by ¹H NMR spectroscopy (Supporting Information Figure S17). Though not confirmed by MS data, *strop0241* is predicted to be the desaturase/isomerase responsible for the conversion of the β-ring in to an aryl functional group. The gene responsible for encoding a 2'-hydroxylase, however, was not identified. A biosynthetic gene responsible for the addition of this functional group is presently unknown in the carotenoid literature. A 2,2'-hydroxylase is also missing from the biosynthetic pathways of (2*S*,2'*S*)-oscillole 2,2'-di-(α-L-rhamnoside), from *Gemmatimonas aurantiaca*, and deinoxanthin, from *Deinococcus radiodurans* (Takaichi *et al.*, 2010; Tian and Hua, 2010). Both of these pathways also have non-clustered arrangements (Takaichi *et al.*, 2010; Tian and Hua, 2010), as found in *Salinispora*. Cyanobacteria are also known to produce carotenoids with 2,2'-hydroxy group additions, but no gene has been confirmed to be responsible for this enzymatic reaction (Britton, 2004; Graham and Bryant, 2009).

Identification of these structural features is suggestive of the order in which enzymes act during the biosynthesis of sioxanthin. The intermediates isolated from the biosynthetic mutants show that both ends of the compound are biosynthesized independently (Figure 1), though in each of these ends, biosynthesis occurs in a set order. Unsurprisingly, on the aromatic side of the molecule, cyclization takes place prior to desaturation/isomerization. On the other end, 1',2'-hydration occurs first, resulting in a 1'-hydroxy group as well as a saturation of the 1',2' bond. Next is the desaturation of the 3',4'-bond, followed by the addition of the 2'-hydroxy, and finally, the glycosylation of the 1'-hydroxyl group.

Sioxanthin biosynthetic pathway in other bacteria

Since a sub-clustered and dispersed genomic arrangement of secondary metabolites is unusual in bacteria, the conservation and arrangement of these genes in other members of the genus was investigated. Two closed and 86 draft *Salinispora* genomes are currently available covering all three species and dozens of strains. Analyses confirmed that the carotenoid gene subclustered arrangement was identical in every strain of *Salinispora* (Supporting Information Figures S18–S21). The identification of these genes in other species of *Salinispora* demonstrates that the carotenoid biosynthesis genomic pattern, in addition to the compound produced, is conserved throughout the genus.

The sioxanthin biosynthetic pathway and genome arrangement are not localized to the genus *Salinispora*. Other bacteria in the bacterial family Micromonosporaceae produce orange pigments, presumably carotenoids. The sioxanthin pathway genes, as they are found in *Salinispora*, are also present in the genera *Micromonospora*, *Verrucosipora* and some species in the genus *Actinoplanes* (Supporting Information Figure S22). It is difficult to predict the biosynthesis of sioxanthin in these bacteria with certainty, as there is still an unidentified gene in the pathway. However, the high degree of pathway synteny between these genera (Supporting Information Figures S23 and S24) provides some confidence that sioxanthin contributes to their orange pigmentation. The *Hamadea* and *Longispora* contain genes for both terpene clusters, though the genes are sometimes found in a different arrangement (Supporting Information Figure S22). As neither of these genera harbors *crtY*

or *crtU* homologs, they are not predicted to produce sioxanthin. Sioxanthin biosynthesis appears to be a feature of *Salinispora* and closely related genera, though not of Micromonosporaceae as a whole. This suggests that the sioxanthin pathway developed in a common ancestor of a few genera, but after the divergence of the Micromonosporaceae family.

While the sioxanthin biosynthetic pathway genes are too ancient to be considered the result of horizontal gene transfer, evaluation of the ancestral traits does suggest acquisition of different parts of the pathway at different times in *Salinispora* evolution. Character trees showing the presence and absence of sioxanthin gene homologs were analyzed to show the gene traits of ancestral species in the actinomycetes (Supporting Information Figures S25–S30). The evolution of the sioxanthin biosynthetic pathway can be summarized as shown in Figure 4. Three of the genes investigated (*crtE* and *crtB* of *terp2* and the acetyltransferase of *terp1*) have a long history within this group of bacteria and were vertically inherited from a common ancestor of the Order Actinomycetales. The other three genes (*cruC* of *terp1* and the unclustered *crtU* and *crtY*) are much more recent additions, having been acquired after the split of the Micromonosporaceae from other bacterial families, but before the divergence of the *Salinispora*, *Micromonospora*, *Verrucosispora*, and *Actinoplanes* genera. This acquisition was then likely followed by gene loss of *crtY* homologs in several of the *Actinoplanes* species. No actinomycetes outside of the Micromonosporaceae contained the full suite of carotenogenesis genes as found in *Salinispora*, thus making it unlikely that sioxanthin is found outside of this bacterial family.

Gene clustering and modularity in carotenoid biosynthesis

Gene clustering is considered to be the standard configuration of bacterial secondary metabolic genes, including carotenoid biosynthesis (Armstrong and Hearst, 1996; Armstrong, 1997; Ballouz *et al.*, 2010). Since the first identification of a bacterial carotenoid gene cluster in *Rhodobacter capsulatus* in 1976 (Armstrong and Hearst, 1996; Yen and Marrs, 1976), carotenoid gene clusters have been characterized from at least 35 bacteria of diverse phylogenies, contributing to the notion that gene clustering is the standard arrangement. However, three characterized examples are now known in bacteria which diverge from this model. These bacteria with dispersed pathway subclusters are from three distantly related phyla, *Gemmatimonas aurantiaca* (Gemmatimonadetes) (Takaichi *et al.*, 2010), *Deinococcus radiodurans* (*Deinococcus-Thermus*, other species within this phylum have bioinformatically predicted pathways with similar architecture) (Tian and Hua, 2010), and now *Salinispora tropica* (Actinobacteria). This observation shows that the non-clustered phenomenon is phylogenetically widespread, though the prevalence is unknown.

As published studies of carotenoid biosynthesis have overwhelmingly uncovered clustered gene sets, attention was focused on bacterial genomes with unstudied carotenoid biosynthetic pathways. This was accomplished by searching for relative genome locations of phytoene synthase and lycopene cyclase genes in 86 pathways from 75 genomes in 5 bacterial classes whose carotenoid pathways had not been reported. The results showed that these two genes were clustered in only about half of these genomes, much greater than the representation in the published literature. This observation showed no pattern related to

phylogeny (with the exception of the gamma-proteobacteria, where only five genomes were investigated), number of carotenoid pathways within a genome, or environment (marine, freshwater, soil, and host-association environments as well as in both high and low temperature extremes were investigated (Figure 5)). Though it cannot yet be shown that all of these genes are active in their predicted biosynthetic pathways, these results do suggest that clustering of carotenoid biosynthetic genes is not as standard as the literature suggests. The mechanisms for development and maintenance of this genome architecture are not yet known, but may provide insights towards bacterial genome organization as well as the evolution and regulation of secondary metabolites.

The observation that single gene clusters are not required in carotenoid biosynthesis led to questions regarding the patterns of sub-clustering in these pathways. In sioxanthin biosynthesis, the pathway has a pseudomodular arrangement in which individual genomic regions are responsible for different carotenoid structural features. Modularity has also been implicated in carotenoid biosynthesis, as a means of explaining the large chemical diversity and the relatively low number of known biosynthetic genes (Garcia-Asua *et al.*, 1998). It has been shown that exchange and addition of genes in bacterial carotenoid pathways result in the biosynthesis of new compounds (Garcia-Asua *et al.*, 1998; Kim *et al.*, 2010). This observation suggests that carotenoid biosynthetic pathways have the flexibility to alter their final product in response to changes in their gene composition, and that carotenoid structural diversity may arise from the exchange or addition of biosynthetic modules.

In *S. tropica*, the functional role of carotenoid biosynthetic genes is correlated with its genomic arrangement. As shown in Figure 1, the combined efforts of genes in *terp2* produce lycopene, which provides the backbone for further modification. The *terp1* genes encode enzymes that modify only one side of the molecule, contributing the glycosylated functional group as well as the additional desaturation. The opposite side of the molecule, the result of cyclization and aromatization, is constructed by the non-clustered genes. However, this pseudomodular pattern is not found in the other two known non-clustered pathways from *Deinococcus* and *Gemmatomonas*. More work is needed to understand the mechanisms driving the genomic architecture in these cases and the impacts upon regulation.

Carotenoid glycosylation and relationship of structure and function

The unique structure of the sioxanthin molecule begs questions regarding its biological function. Like many carotenoids, it likely has a role in oxidative stress prevention due to its central conjugated chain. Carotenoids also serve as structural molecules, attaching themselves to proteins or, more commonly, integrating within lipid membranes (Britton, 1995; Gruszecki, 2009). Their position and influence on membrane fluidity is determined by the presence of polar functional groups (Gruszecki, 2009). No studies have been reported that have investigated the membrane interactions of carotenoids with polar and non-polar ends, as is the case in sioxanthin.

In actinomycetes, glycosylated carotenoids are rare. Only nine glycosylated carotenoids have been identified from eight genera, including sioxanthin biosynthesis in *Salinispora* (Britton, 2004; Takaichi *et al.*, 2008). Glycosyl groups can assist in the structural role of carotenoids as has been demonstrated in cyanobacteria where glycosylated carotenoids are

common (Britton, 2004). Studies have shown that glycosylated carotenoids are localized to the thylakoid membrane where the glycosyl moiety serves as a binding motif that enables the proper folding and stacking of the thylakoid membrane (Mohamed *et al.*, 2005). While *Salinispora* lack thylakoid (Maldonado *et al.*, 2005) the role of the glycosyl residue in the cell membrane may be similar in that it extends beyond the membrane and serves as an anchor or attachment site.

Though the *Salinispora* genus is undergoing extensive investigations regarding its chemical potential, biogeography, and evolutionary history (Jensen and Mafnas, 2006; Jensen *et al.*, 2007; Freel *et al.*, 2012), little is known about its growth in nature or its role in their environmental niche. The biological function of sioxanthin, as determined by its chemical structure, may give insight in how the Micromonosporaceae, and bacteria in general, experience and interact with their environment.

Conclusions

Orange pigmentation, characteristic of the genus *Salinispora*, is the result of a novel carotenoid compound, now referred to as “sioxanthin”, the structure of which revealed some surprising features, such as head groups with opposite polarity involving an aromatic hydrocarbon and a glucosylated diol residue. The biosynthesis of this compound is similarly interesting, employing several distant regions of the genome to produce a single compound. This genetic organization diverges from what is commonly believed to be a standard arrangement for microbial carotenoid biosynthetic genes, and indeed secondary metabolites in general. Thus, *Salinispora* provides a special opportunity to apply genetic techniques to the study of this type of biochemical pathway. The exploration of this pathway uncovered previously overlooked genes in carotenoid biosynthesis in actinomycetes and may allow for an improved ability to predict structures in other species.

Experimental procedures

Identification of candidate carotenoid biosynthetic genes in *Salinispora*

Two terpene clusters, named *terp1* and *terp2*, were previously identified and had been predicted to be involved in carotenoid biosynthesis-based gene annotation (Udwary *et al.*, 2007; Penn *et al.*, 2009). Genes in *terp2*, *strop3251*, *strop3248*, and *strop2408* were previously annotated as carotenoid biosynthesis genes based on domain function. BLAST query searches of these genes against the nr database were used to confirm homology. Additional genes were identified via their sequence homology to known carotenoid genes. Protein sequences for known carotenoid biosynthetic genes were gathered from NCBI and used as BLASTx query sequences against the *Salinispora tropica* CNB-440 and *Salinispora arenicola* CNS-205 closed genomes. Query sequences came from actinomycete genomes where possible. Positive gene hits were those which had greater than 35% sequence identity with more than 50% query coverage, except in the case of *terp1* genes in which the top hits from a *Meithermus ruber* DSM 1279 query sequence were used.

Analysis of carotenoid biosynthetic genes and gene synteny

Eighty-eight *Salinispora* and 15 other Micromonosporaceae genomes are available on the Integrated Microbial Genomes database hosted by the Joint Genome Institute (Markowitz *et al.*, 2014). Synteny of the carotenoid gene clusters and neighborhoods was investigated by performing a BLAST sequence homology search of the gene of interest against the genomes. The BLAST hits were then used to explore the gene neighborhoods with the same top COG (Cluster of Orthologous Genes) hit.

Bacterial strains and growth conditions

Wild-type *Salinispora* were grown in A1 liquid (per liter: 10 g starch, 4 g yeast extract, 2 g peptone, and 28 g Instant Ocean Marine Salts) and on solid media (A1 with addition of 18 g agar/liter) at 30°C (Beer and Moore, 2007). Liquid cultures were grown as 50 mL cultures in 250 mL Erlenmeyer flasks containing a spring with shaking at 200 rpm. After five days of growth, 5 mL of the starter culture were transferred under sterile conditions to 2.5 L flasks containing 1 L A1 liquid medium. Cultures were allowed to continue growing for 10–14 days at 30°C with shaking at 220 rpm. Mutant *Salinispora* strains were grown as above, but with the addition of apramycin (50 µg/mL as a selection marker and naladixic acid (100 µg/ml).

Escherichia coli strains were grown in LB (Luria-Bertani) broth and on solid LB-agar media with the following antibiotics added as appropriate: carbenicillin (100 µg/mL), chloramphenicol (12.5 µg/mL), and apramycin (50 µg/mL). *E. coli* strains were grown at 37°C, except for strain BW25113, which was grown at 30°C when necessary to maintain the temperature-sensitive pKD20 plasmid.

Inactivation of putative carotenoid genes in *Salinispora*

Gene inactivation experiments were carried out in *Salinispora tropica* CNB-440 using PCR-directed gene mutagenesis, as previously described (Gust *et al.*, 2003; Eustaquio *et al.*, 2008). Tailed PCR primer sequences are shown in Supplemental Table S3.

Extraction and analysis of mutant and wild-type carotenoids

Cultures were centrifuged in a Beckman-Coulter Avanti J-E centrifuge at 11,000 × g for 30 minutes at 4°C. Media was discarded and cell pellets from 6–10 liters of cell culture were combined for extraction. Pigment extracts were obtained by soaking cell pellets in approximately 100 mL of acetone for several hours with occasional stirring. The colored solvent was passed through a filter to remove cell debris. Acetone extractions were repeated until the solvent was no longer colored after exposure to the cells. Extractions were combined and washed with brine, dried with magnesium sulfate or sodium sulfate, filtered to remove salts, and finally dried *in vacuo*.

Analytical HPLC conditions

Crude extracts were resuspended in acetonitrile and analyzed on an Agilent 1200 series analytical HPLC using reversed-phase conditions, at 450 nm detection. Extracts were analyzed on a Luna 5µm C18 100 × 4.6 mm column, with 98% acetonitrile and 2% water

under isocratic conditions. The HPLC comparing crude extracts of *S. tropica* and *S. arenicola* used an isocratic solvent system of 80% acetonitrile, 15% methanol and 5% isopropanol on a Luna 5 µm C18 150 × 4.6 mm column.

Derivatization of the crude extract

The crude pigment extracts were acetylated to improve stability during the purification steps. The dry extracts were resuspended in 1–3 mL dichloromethane (CH₂Cl₂). For each mL of CH₂Cl₂, 250 µL triethylamine (Et₃N), 250 µL acetic anhydride, and 0.1 mg 4-dimethylaminopyridine were added. The reaction was allowed to sit at room temperature for at least four hours. The pigmented solution was diluted with ethyl acetate (EtOAc) and washed with brine. The EtOAc extract was dried with sodium sulfate. Salts were removed by passing the solution through a glass wool filter, and the resulting solution was dried under nitrogen. The conversion of the reaction was assessed by resuspending the extract in acetonitrile and analyzing the solution on a reversed phase analytical HPLC, as above. The peak corresponding to the major product had a longer retention time, indicating its increased hydrophobicity, and the wild-type compound was completely consumed.

Carotenoid biosynthesis gene cluster analysis in sequenced bacteria

The *S. tropica* CNB-440 CrtY protein sequence (strop_2408) was used as a BLASTp query sequence against the nr database to find bacterial genomes with *crtY*-type lycopene cyclase genes. The genomes were then reanalyzed with BLAST using strop_4441 as a query sequence to find the locations of the phytoene synthase genes. Genes were determined to be clustered if they were in close proximity to each other and were separated by no more than ten genes unrelated to carotenoid biosynthesis genes. Graphs were made in Excel to show the proportion of pathways in each category with a particular genome architecture. The NCBI database provided phylogenetic and environmental information for the bacterial genomes.

Gene phylogenies and ancestral state

The species tree was built by combining *rpoB* sequences of all of the actinomycete genomes which had homologs to the six sioxanthin genes used to study ancestral state (*strop0241*, *strop2408*, *strop3245*, *strop3247*, *strop4440* and *strop4441*). Cyanobacterial sequences were used as an outgroup. All sequence alignment and tree building were done on the phylogeny.fr website (www.phylogeny.fr) (Dereeper *et al.*, 2008). Sequences were aligned using MUSCLE (Edgar, 2004). Alignment curation was done by Gblocks (Castresana, 2000) allowing for smaller final blocks, gap positions within the final blocks, and less strict flanking positions. Maximum-likelihood trees were built using PhyML 3.0 with an SH-like Approximate Likelihood-Ratio Test (Guindon *et al.*, 2010). Trees were visualized and edited in MEGA 5.2 (Tamura *et al.*, 2011) The ancestral node was inferred using the trace character history function implemented in Mesquite v2.75 (Maddison and Maddison, 2011). A character matrix was created for each gene homolog and likelihood calculations were performed using an Mk1 model. Likelihood scores of greater than 50% on the ancestral nodes were used to infer the points of gene acquisition.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

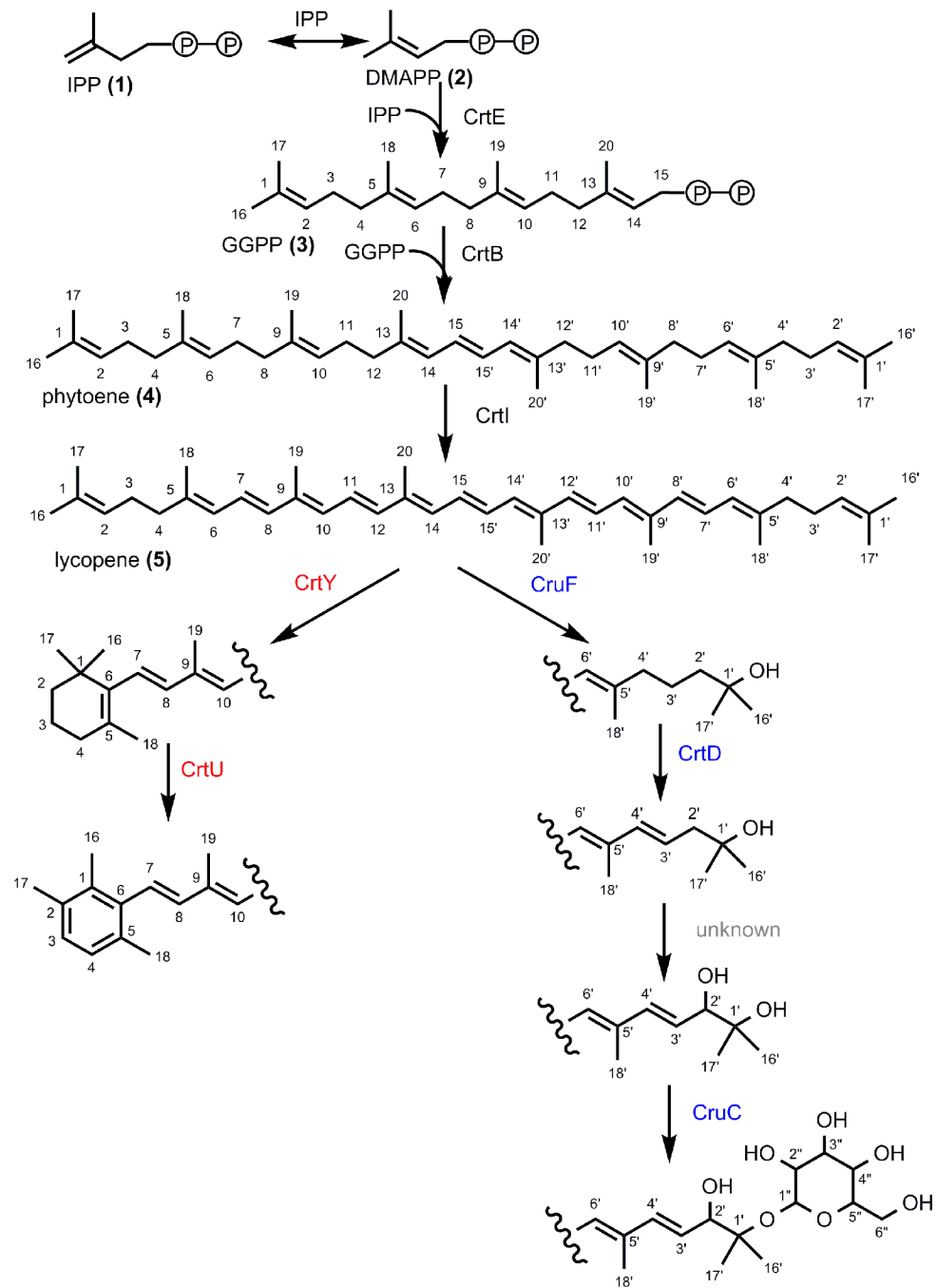
This work was supported by a grant from the NIH (GM085770) to B.S.M. and a California Sea Grant Traineeship to T.K.S.R. We thank Dr. Paul Jensen for providing *Salinispora* strains, Dr. Nadine Ziemert for assistance with phylogenetic techniques, and the UCSD Chemistry and Biochemistry Molecular MS Facility and Brendan Duggan at the UCSD NMR Facility for assistance in analytical techniques.

References

- Ahmed L, Jensen PR, Freel KC, Brown R, Jones AL, Kim B-Y, Goodfellow M. *Salinispora pacifica* sp nov., an actinomycete from marine sediments. *Anton Leeuw Int J G.* 2013; 103:1069–1078.
- Armstrong GA. Eubacteria show their true colors - genetics of carotenoid pigment biosynthesis from microbes to plants. *J Bacteriol.* 1994; 176:4795–4802. [PubMed: 8050991]
- Armstrong, GA. Genetics of eubacterial carotenoid biosynthesis: A colorful tale. In: Ornston, LN., editor. *Annu Rev Microbiol. USA: Annual Reviews Inc.*; 1997. p. 629-659.
- Armstrong GA, Hearst JE. Carotenoids 2: Genetics and molecular biology of carotenoid pigment biosynthesis. *FASEB J.* 1996; 10:228–237. [PubMed: 8641556]
- Ballouz S, Francis AR, Lan R, Tanaka MM. Conditions for the evolution of gene clusters in bacterial genomes. *PLoS Comput Biol.* 2010; 6:e1000672. [PubMed: 20168992]
- Beer LL, Moore BS. Biosynthetic convergence of salinosporamides A and B in the marine actinomycete *Salinispora tropica*. *Org Lett.* 2007; 9:845–848. [PubMed: 17274624]
- Blanco J, Coque JJR, Martin JF. The folate branch of the methionine biosynthesis pathway in *Streptomyces lividans*: Disruption of the 5,10-methylenetetrahydrofolate reductase gene leads to methionine auxotrophy. *J Bacteriol.* 1998; 180:1586–1591. [PubMed: 9515933]
- Britton G. Structure and properties of carotenoids in relation to function. *FASEB J.* 1995; 9:1551–1558. [PubMed: 8529834]
- Britton, GL-JSPH. *Carotenoids handbook.* Basel; Boston: Birkhäuser Verlag; 2004.
- Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol.* 2000; 17:540–552. [PubMed: 10742046]
- Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* 2008; 36:W465–W469. [PubMed: 18424797]
- Dieser M, Greenwood M, Foreman CM. Carotenoid pigmentation in Antarctic heterotrophic bacteria as a strategy to withstand environmental stresses. *Arct Antarct Alp Res.* 2010; 42:396–405.
- Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004; 32:1792–1797. [PubMed: 15034147]
- Eustaquio AS, Pojer F, Noe JP, Moore BS. Discovery and characterization of a marine bacterial SAM-dependent chlorinase. *Nat Chem Biol.* 2008; 4:69–74. [PubMed: 18059261]
- Feling RH, Buchanan GO, Mincer TJ, Kauffman CA, Jensen PR, Fenical W. Salinosporamide A: A highly cytotoxic proteasome inhibitor from a novel microbial source, a marine bacterium of the new genus *Salinispora*. *Angew Chem Int Ed.* 2003; 42 355–+
- Fischbach MA, Walsh CT, Clardy J. The evolution of gene collectives: How natural selection drives chemical innovation. *Proc Natl Acad Sci USA.* 2008; 105:4601–4608. [PubMed: 18216259]
- Fischbach M, Voight CA. Prokaryotic gene clusters: a rich toolbox for synthetic biology. *Biotechnol J.* 2010; 5:1277–1296. [PubMed: 21154668]
- Freel KC, Edlund A, Jensen PR. Microdiversity and evidence for high dispersal rates in the marine actinomycete '*Salinispora pacifica*'. *Environ Microbiol.* 2012; 14:480–493. [PubMed: 22117917]

- Galbis-Martinez M, Padmanabhan S, Murillo FJ, Elias-Arnanz M. CarF mediates signaling by singlet oxygen, generated via photoexcited protoporphyrin IX, in *Myxococcus xanthus* light-induced carotenogenesis. *J Bacteriol.* 2012; 194:1427–1436. [PubMed: 22267513]
- Garcia-Asua G, Lang HP, Cogdell RJ, Hunter CN. Carotenoid diversity: a modular role for the phytoene desaturase step. *Trends Plant Sci.* 1998; 3:445–449.
- Goo K-S, Tsuda M, Ulanova D. *Salinispora arenicola* from temperate marine sediments: new intra-species variations and atypical distribution of secondary metabolic genes. *Anton Leeuw Int J G.* 2014; 105:207–219.
- Graham JE, Bryant DA. The biosynthetic pathway for myxol-2' fucoside (myxoxanthophyll) in the cyanobacterium *Synechococcus* sp strain PCC 7002. *J Bacteriol.* 2009; 191:3292–3300. [PubMed: 19304845]
- Gruszecki, W. Carotenoids in Lipid Membranes. In: Landrum, JT., editor. *Carotenoids Physical, Chemical, and Biological Functions and Properties.* Hoboken: CRC Press; 2009.
- Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst Biol.* 2010; 59:307–321. [PubMed: 20525638]
- Gulder TAM, Moore BS. Salinosporamide Natural Products: Potent 20S Proteasome Inhibitors as Promising Cancer Chemotherapeutics. *Angew Chem Int Ed.* 2010; 49:9346–9367.
- Gust B, Challis GL, Fowler K, Kieser T, Chater KF. PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc Natl Acad Sci U S A.* 2003; 100:1541–1546. [PubMed: 12563033]
- Hertzber S, Jensen SL. Bacterial carotenoids 20. Carotenoids of *Mycobacterium phlei* strain Vera. 2. Structures of the phlei-xanthophylls - two novel tertiary glucosides. *Acta Chem Scand.* 1967; 21:15–41. [PubMed: 4291656]
- Jensen PR, Mafnas C. Biogeography of the marine actinomycete *Salinispora*. *Environ Microbiol.* 2006; 8:1881–1888. [PubMed: 17014488]
- Jensen PR, Williams PG, Oh D-C, Zeigler L, Fenical W. Species-specific secondary metabolite production in marine actinomycetes of the genus *Salinispora*. *Appl Environ Microbiol.* 2007; 73:1146–1152. [PubMed: 17158611]
- Kim SH, Park YH, Schmidt-Dannert C, Lee PC. Redesign, reconstruction, and directed extension of the *Brevibacterium linens* C-40 carotenoid pathway in *Escherichia coli*. *Appl Environ Microbiol.* 2010; 76:5199–5206. [PubMed: 20525861]
- Klassen JL. Phylogenetic and evolutionary patterns in microbial carotenoid biosynthesis are revealed by comparative genomics. *PLoS One.* 2010; 5:e11257. [PubMed: 20582313]
- Krugel H, Krubasik P, Weber K, Saluz HP, Sandmann G. Functional analysis of genes from *Streptomyces griseus* involved in the synthesis of isorenieratene, a carotenoid with aromatic end groups, revealed a novel type of carotenoid desaturase. *BBA-Mol Cell Biol L.* 1999; 1439:57–64.
- Kuzuyama T, Takahashi S, Dairi T, Seto H. Detection of the mevalonate pathway in *Streptomyces* species using the 3-hydroxy-3-methylglutaryl coenzyme A reductase gene. *J Antibiot (Tokyo).* 2002; 55:919–923. [PubMed: 12523826]
- Liu GY, Nizet V. Color me bad: microbial pigments as virulence factors. *Trends Microbiol.* 2009; 17:406–413. [PubMed: 19726196]
- Lutnaes BF, Strand A, Petursdottir SK, Liaaen-Jensen S. Carotenoids of thermophilic bacteria - *Rhodothermus marinus* from submarine Icelandic hot springs. *Biochem Syst Ecol.* 2004; 32:455–468.
- Maddison, WP.; Maddison, DR. Mesquite: a modular system for evolutionary biology analysis. 2011. <http://mesquiteproject.org>
- Maldonado LA, Fenical W, Jensen PR, Kauffman CA, Mincer TJ, Ward AC, et al. *Salinispora arenicola* gen. nov., sp nov and *Salinispora tropica* sp nov., obligate marine actinomycetes belonging to the family Micromonosporaceae. *Int J Syst Evol Micr.* 2005; 55:1759–1766.
- Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, Pillay M, et al. IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Res.* 2014; 42:D560–D567. [PubMed: 24165883]

- Martin MJ, Herrero J, Mateos A, Dopazo J. Comparing bacterial genomes through conservation profiles. *Genome Res.* 2003; 13:991–998. [PubMed: 12695324]
- Marz U. The Global Market for Carotenoids. *The Global Market for Carotenoids.* 2011:189.
- Mohamed HE, van de Meene AML, Roberson RW, Vermaas WFJ. Myxoxanthophyll is required for normal cell wall structure and thylakoid organization in the cyanobacterium, *Synechocystis* sp strain PCC 6803. *J Bacteriol.* 2005; 187:6883–6892. [PubMed: 16199557]
- Moraleda-Munoz A, Perez J, Fontes M, Murillo FJ, Munoz-Dorado J. Copper induction of carotenoid synthesis in the bacterium *Myxococcus xanthus*. *Mol Microbiol.* 2005; 56:1159–1168. [PubMed: 15882411]
- O'Brien J, Wright GD. An ecological perspective of microbial secondary metabolism. *Curr Opin Biotechnol.* 2011; 22:552–558. [PubMed: 21498065]
- Osborn A. Secondary metabolic gene clusters: evolutionary toolkits for chemical innovation. *Trends Genet.* 2010; 26:449–457. [PubMed: 20739089]
- Pal C, Hurst LD. Evidence against the selfish operon theory. *Trends Genet.* 2004; 20:232–234. [PubMed: 15145575]
- Penn K, Jenkins C, Nett M, Udvary DW, Gontang EA, McGlinchey RP, et al. Genomic islands link secondary metabolism to functional adaptation in marine Actinobacteria. *ISME J.* 2009; 3:1193–1203. [PubMed: 19474814]
- Pezzoni M, Costa CS, Pizarro RA, Oppezzo OJ. The relationship between carotenoids and sunlight response in members of the family Micrococcaceae. *J Basic Microbiol.* 2011; 51:325–329. [PubMed: 21298681]
- Rahlert N, Fraser PD, Sandmann G. A crtA-related gene from Flavobacterium P99-3 encodes a novel carotenoid 2-hydroxylase involved in myxol biosynthesis. *FEBS Lett.* 2009; 583:1605–1610. [PubMed: 19383499]
- Takaichi S, Maoka T, Takasaki K, Hanada S. Carotenoids of *Gemmatimonas aurantiaca* (Gemmatimonadetes): identification of a novel carotenoid, deoxyoscollol 2-rhamnoside, and proposed biosynthetic pathway of oscollol 2,2'-dirhamnoside. *Microbiol-Sgm.* 2010; 156:757–763.
- Takaichi S, Maoka T, Akimoto N, Carmona ML, Yamaoka Y. Carotenoids in a Corynebacterineae, *Gordonia terrae* AIST-1: Carotenoid glucosyl mycoloyl esters. *Biosci Biotech Bioch.* 2008; 72:2615–2622.
- Takano H, Obitsu S, Beppu T, Ueda K. Light-induced carotenogenesis in *Streptomyces coelicolor* A3(2): Identification of an extracytoplasmic function sigma factor that directs photodependent transcription of the carotenoid biosynthesis gene cluster. *J Bacteriol.* 2005; 187:1825–1832. [PubMed: 15716454]
- Takano H, Asker D, Beppu T, Ueda K. Genetic control for light-induced carotenoid production in non-phototrophic bacteria. *J Ind Microbiol Biotechnol.* 2006; 33:88–93. [PubMed: 16091943]
- Tamura K, Peterson D, Peterson N, Steche RG, Nei M, Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol.* 2011; 28:2731–2739. [PubMed: 21546353]
- Tian B, Hua Y. Carotenoid biosynthesis in extremophilic Deinococcus-Thermus bacteria. *Trends Microbiol.* 2010; 18:512–520. [PubMed: 20832321]
- Udvary DW, Zeigler L, Asolkar RN, Singan V, Lapidus A, Fenical W, et al. Genome sequencing reveals complex secondary metabolome in the marine actinomycete *Salinispora tropica*. *Proc Natl Acad Sci U S A.* 2007; 104:10376–10381. [PubMed: 17563368]
- Vershinin A. Biological functions of carotenoids - diversity and evolution. *Biofactors.* 1999; 10:99–104. [PubMed: 10609869]
- Yen HC, Marrs B. Map of genes for carotenoid and bacteriochlorophyll biosynthesis in *Rhodospseudomonas capsulata*. *J Bacteriol.* 1976; 126:619–629. [PubMed: 1262313]
- Zheng Y-T, Toyofuku M, Nomura N, Shigeto S. Correlation of carotenoid accumulation with aggregation and biofilm development in *Rhodococcus* sp SD-74. *Anal Chem.* 2013; 85:7295–7301. [PubMed: 23802134]
- Ziemert N, Lechner A, Wietz M, Millan-Aguinaga N, Chavarria KL, Jensen PR. Diversity and evolution of secondary metabolism in the marine actinomycete genus *Salinispora*. *Proc Natl Acad Sci U S A.* 2014; 111:E1130–E1139. [PubMed: 24616526]

**Figure 1.**

Proposed biosynthesis of sioxanthin. Early steps (black enzymes) are common to carotenoid biosynthetic pathways in general. Structural diversity derives from modification of lycopene, shown here for sioxanthin. Colors correspond to genome localization of the genes required for each step where black is *terp2* genes, blue is *terp1* genes, red are non-clustered genes.

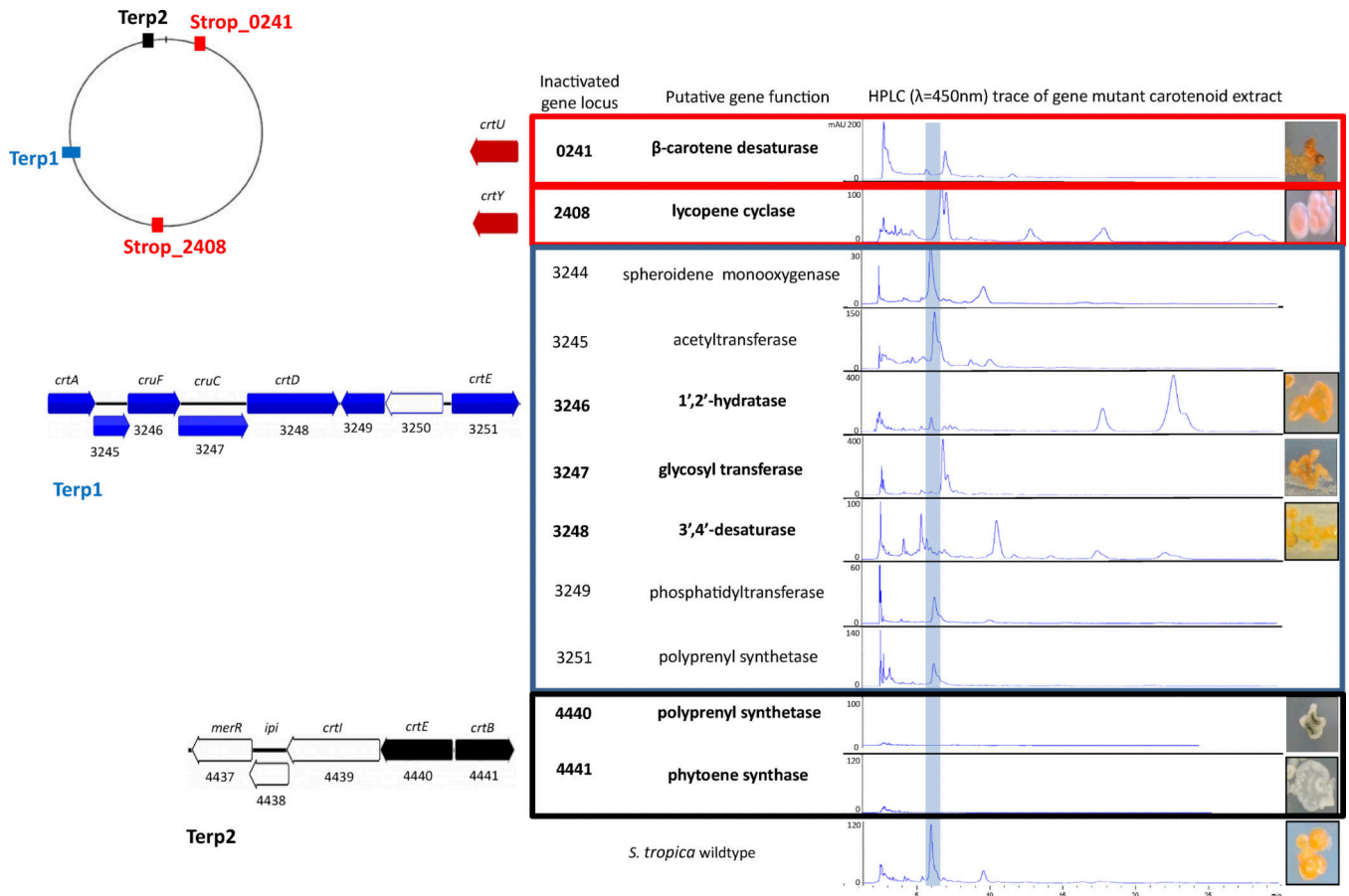
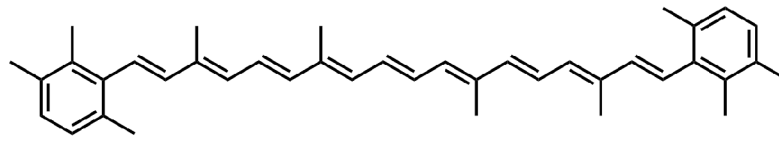
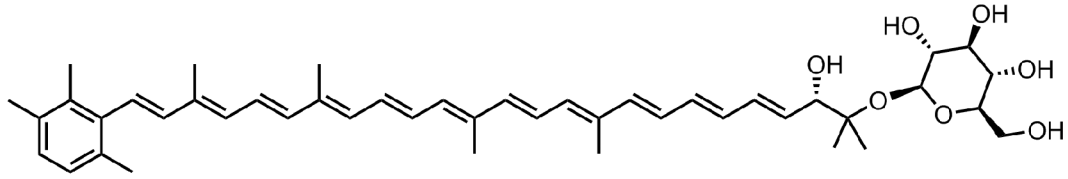


Figure 2.

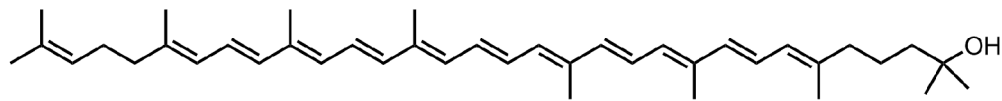
Carotenoid biosynthetic genes in *Salinispora* are subclustered and dispersed throughout the genome. Genome map shows the relative locations of carotenoid biosynthetic genes including two terpene clusters (*terp1* blue and *terp2* black) and two non-clustered genes (*strop_0241* and *strop_2408* red). Numbers refer to the gene locus, gene code is the predicted carotenoid biosynthesis homolog. Table shows results of gene inactivation of carotenoid biosynthesis with the blue line highlighting the sioxanthin biosynthetic peak. Genes determined to be part of the sioxanthin biosynthetic pathway, based on altered retention time, are shown in bold. Photographs show visual phenotypes of wildtype and mutant carotenoid producers. Numbers in parenthesis refer to the structure.



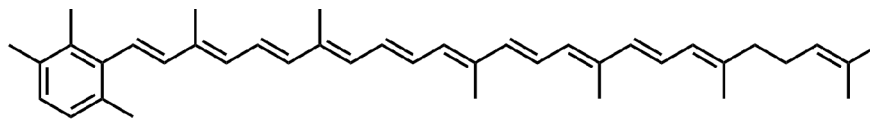
isorenieratene (6)



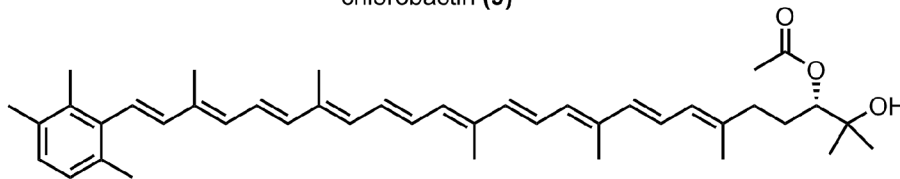
sioxanthin (7)



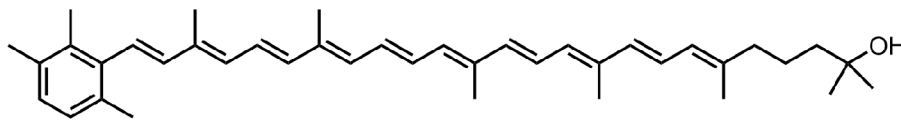
rhodopin (8)



chlorobactin (9)



acetyl sioxanthin aglycone (10)



chlorobactene (11)

Figure 3.
Chemical structures of sioxanthin and related compounds.

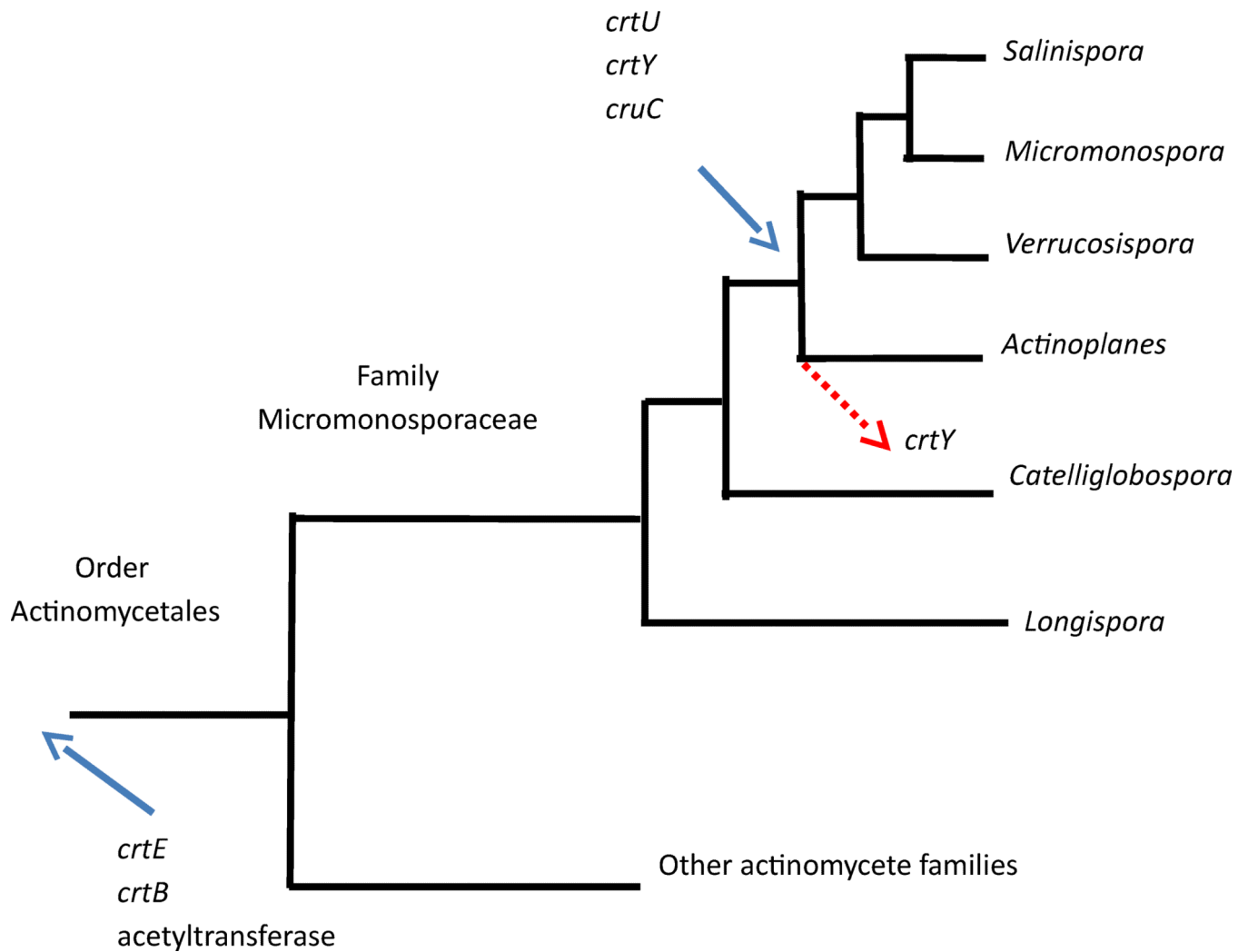


Figure 4. Tree showing relative relationship of the Micromonosporaceae genera and other actinomycetes, highlighting the time of entry of sixxanthin genes in to the genomes. Three genes were present prior to the subdivision of the order actinomycetales. Three others entered later (blue arrow) followed by some gene loss (red arrow).

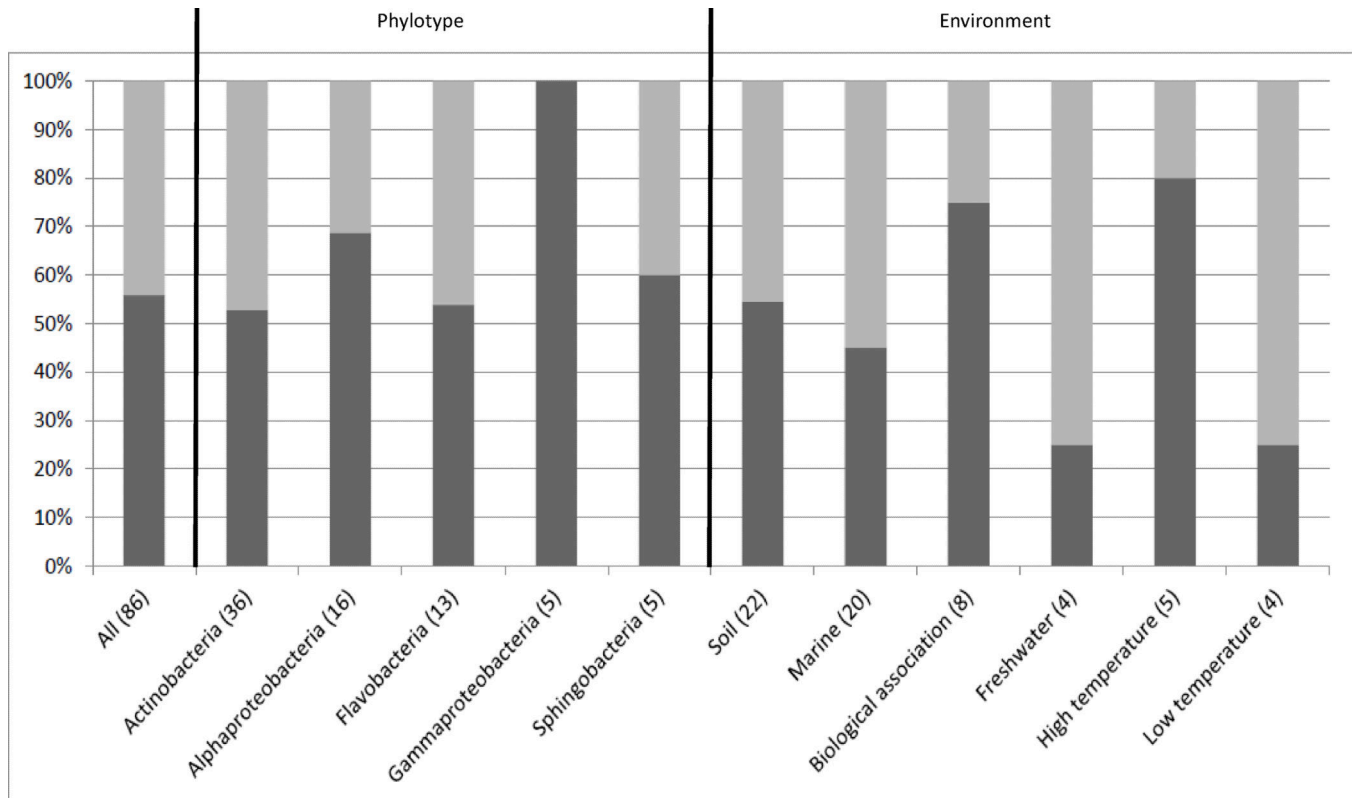


Figure 5.

Chart of percentage of clustered (dark grey) and non-clustered (light grey) carotenoid biosynthetic pathways in selected bacterial genomes, shown by total genomes, individual phylogenetic groupings, and by environment in which the host organism was isolated. The total number of genomes in each category is in parenthesis.