DNA Research 21, 639–647, (2014)
Advance Access publication on 12 August 2014
doi:10.1093/dnares/dsu026

Draft Genome Sequence of *Rhodomicrobium udaipurense* JA643^{T} with Special Reference to Hopanoid Biosynthesis

L. Tushar[1], Ch. Sasikala[2], and Ch. V. Ramana[1,*]

Department of Plant Sciences, School of Life Sciences, University of Hyderabad, P.O. Central University, Hyderabad 500 046, India[1] and Bacterial Discovery Laboratory, Centre for Environment, IST, JNT University, Hyderabad, Kukatpally, Hyderabad 500 085, India[2]

*To whom correspondence should be addressed. Tel. +91 40-23-134-502. Fax. +91 40-23-010-120. Email: cvr449@yahoo.com

Edited by Dr Naotake Ogasawara
(Received 29 January 2014; accepted 4 July 2014)

Abstract

Hopanoids are present in vast amounts as integral components of bacteria and plants with their primary function to strengthen rigidity of the plasma membrane. To establish their roles more precisely, we conducted sequencing of the whole genome of *Rhodomicrobium udaipurense* JA643^{T} isolated from a fresh water stream of Udaipur in Himachal Pradesh, India, by using the Illumina HiSeq pair end chemistry of 2 × 100 bp platform. Determined genome showed a high degree of similarity to the genome of *R. vannielii* ATCC17100^{T} and the 13.7 million reads generated a sequence of 3,649,277 bp possessing 3,611 putative genes. The genomic data were subsequently investigated with respect to genes involved in various features. The machinery required for the degradation of aromatic compounds and resistance to solvents as well as all that required for photosynthesis are present in this organism. Also, through extensive functional annotation, 18 genes involved in the biosynthesis of hopanoids are predicted, namely those responsible for the synthesis of diploptene, diplopterol, adenosylhopane, ribosylhopane, aminobacteriohopanetriol, glycosyl group containing hopanoids and unsaturated hopanoids. The hopanoid biosynthetic pathway was then inferred based on the genes identified and through experimental validation of individual hopanoid molecules. The genome data of *R. udaipurense* JA643^{T} will be useful in understanding the functional features of hopanoids in this bacterium.

Key words: *Rhodomicrobium udaipurense* JA643^{T}; genome sequence; Illumina Hiseq; hopanoid biosynthesis pathway

1. Introduction

Hopanoids are a group of natural pentacyclic triterpenoid lipids widely distributed in plant and bacterial systems, act as cell membrane rigidifiers, analogues to sterols present in eukaryotes which have tetracyclic rings. Apart from their biological occurrence, hopanoids were also observed in geo-sediments (geohopanoids) formed due to diagenetic processes and are considered as 'molecular fossils' for ancient bacteria.[1] Till date, ∼200 structures of hopanoids are identified from bio-/geo-sources.[2] Hopanoids play a key role in conferring membrane integrity under extreme conditions of stress and any loss in hopanoids makes the bacterium sensitive due to weakening of outer membrane integrity.[3,4] Hopanoids are synthesized from six isopentenyl units forming squalene, an immediate precursor in hopanoid synthesis.[5] In a highly complex cyclization reaction which is similar to oxido squalene to sterols conversion, the hopane skeleton is formed from squalene by the squalene hopene cyclase coded by the gene *shc*.[5] Though the occurrence of *shc* gene is widespread among bacteria, there is limited information on the hopanoids identified from different members.[1]

Among the hopanoids of phototrophic bacteria, much of the work was focused on *Rhodopseudomonas palustris* TIE-1, which accumulates substantial amounts of diploptene, diplopterol, aminobacteriohopanetriol and 2β-methyl-bacteriohopanepoyol.[6] A large diversity

© The Author 2014. Published by Oxford University Press on behalf of Kazusa DNA Research Institute.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

of hopanoids occur in a recently described photosynthetic proteobacterium, *Rhodomicrobium udaipurense* JA643^T belonging to the family Hyphomicrobiaceae.[7] There are nine genomes sequenced from the family Hyphomicrobiaceae in the class Alphaproteobacteria (www.ncbi.nlm.nih.gov/genomes/MICROBES/microbial_taxtree.html), and the genus *Rhodomicrobium* has two recognized species names; *R. vannielii* and *R. udaipurense*, which are phototrophic.[7] While *R. vannielii* ATCC17100^T could tolerate solvent stress and degrade aromatic hydrocarbons, *R. udaipurense* JA643^T was psychrotolerant with diverse metabolic capabilities.[7,8] To assess the metabolic potentials with particular reference to the hopanoid biosynthetic genes of *R. udaipurense* JA643^T, we have sequenced the whole genome of this bacterium. The present genome sequence will expand our understanding of hopanoid biosynthesis and provide more insights into the role of hopanoids in *R. udaipurense* JA643^T.

2. Materials and methods

2.1. Organism and DNA preparation

Rhodomicrobium udaipurense JA643^T was grown photoheterotrophically in light with 2,400 lux at 30°C for 48 h maintaining micro-anaerobic conditions in the screw cap test tube. Culture was grown in liquid mineral medium with pyruvate as carbon source.[9] DNA was isolated with the QIAamp minikit (Qiagen) according to the manufacturer's protocol. The authenticity of the genome was confirmed by 16S rRNA gene sequencing, which was done as previously described.[10]

2.2. Genome sequencing and annotations

Genome sequencing was outsourced to the NxGenBio Life Sciences (India). In brief, genomic DNA was fragmented using Covaris system for a mean size range of 300–350 bp. Genomic DNA shotgun library was prepared using standard Illumina TruSeq protocol. Sequencing was carried out using Illumina Hiseq 2X 100 bp paired end chemistry. Two sequencing runs in total were carried out. The quality of raw reads was checked by fastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/). The raw reads were trimmed by fastx toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). The sequence data were *de novo* assembled using Newbler assembly software. Annotations were performed with *de novo* assembled sequence using the RAST (Rapid Annotation using Subsystem Technology) servers.[11] tRNA and rRNA were predicted using RNAmmer.[12] *In silico* DNA–DNA hybridization was carried out using formula 2 of GGDC website server (http://ggdc.dsmz.de/).

Average nucleotide identity (ANI) was calculated using ANI calculator (http://enve-omics.ce.gatech.edu/ani/). SpecI server was used to confirm the species identification.[13] Phylogenetic tree was constructed by MEGA 5 software.[14] The sequence data were also mapped to the reference sequence of *R. vannielii* ATCC 17100^T using Burrows Wheel Aligner (BWA version 0.7.5). The reference genome sequence of *R. vannielli* ATCC17100^T (GCA_000166055.1) was taken from NCBI. The reference genome map was generated using CGView server.[15]

2.3. Extraction and characterization of hopanoid by GC–MS

Cells were grown photoheterotrophically as previously described by Venkata *et al.*[7] with pyruvate (34 mM) as carbon source. Cells were pelleted at 8,000*g* for 10 min at 4°C. Hopanoids were extracted according to Welender *et al.*[4] In brief, cells were sonicated in 10 ml methanol:dichloromethane (DCM):water (10:5:4) and centrifuged at 3,000*g* for 10 min. Supernatant was transferred to a new tube; sonication was repeated one more time and the supernatants were pooled. Ten millilitres of water and DCM were added to the supernatant and centrifuged at 3,000*g* for 10 min. The organic phase was combined and evaporated to dryness. Hopanoids were derivatized by acetic anhydride:pyridine (1:1) and analysed by GC–MS (Agilent 7890). Separation was achieved using DB-1HT column (30 m × 0.25 mm i.d.; 0.1 μm film thickness) with helium as carrier gas. Ramping program was started with 100°C (held for 2 min), then ramped from 100 to 200°C at 10°C min^{-1} and from 200 to 360°C at 6°C min^{-1} (held for 10 min). The system was operated with the following parameters: electron voltage 70 eV, source temperature 200°C, interface temperature 350°C and acquisition delay for 120 s. Mass spectrometer was operated in full-scan mode (m/z 30–1,000).

3. Results and discussion

13.7 million reads of 101 bp long were obtained after sequencing the genome of *R. udaipurense* JA643^T. Raw reads were trimmed and assembled into 256 contigs having 3,649,277 bp (3.64 Mb) in total. The largest contig size was 103,250 bp, and N50 of contig was 26,179 bp. However, the Newbler software predicted that the genome size of *R. udaipurense* JA643^T would be ∼3.8 Mb. Thus, genome sequence has revealed 96% of the whole genome, with an average G + C content of 62.4 mol% and 3,611 open-reading frames (ORFs). The protein coding bases were 3,000,076 bp in total, covering 82.21% of the total bases determined. It contains 68 tRNAs and 3 rRNA (Table 1).

Though *R. udaipurense* JA643^T and *R. vannielii* ATCC17100^T have 100% 16S rRNA gene sequence similarity, based on genome relatedness (46.1% after DNA–DNA hydridization) and phenotypic traits, these

Table 1. General features of *Rhodomicrobium udaipurense* JA643^T draft genome sequence

Total bases	3,649,277
Average gene length	830 bp
Protein coding features	3,611
Protein coding bases	3,000,076 (82.21%)
tRNAs	68
rRNAs	3
G + C percentage	62.4 mol%
DNA–DNA hybridization (*in vitro*) with *Rhodomicrobium vannielii* DSM162^T (=ATCC17100^T)[1]	46.1%
In silico DNA–DNA hybridization with the genome of *R. vannielii* ATCC17100^T [Genome accession number; GCA_000166055.1]	58.60%

were described as two different species.[7] Delineation of the two species was further confirmed by *in silico* regression DDH analysis of the whole genomes, and the determined value of 58.6% is much lower than the gold standard value of 70% for delineation of bacterial species.[16] Further, SpecI analysis using 40 phylogenetic marker genes (pMGs) also indicated an average identity (Supplementary Table S1) of 94.69% which concludes a clear delineation between the two species.[13] Calculated ANI value of 95.26% between the genomes of *R. udaipurense* JA643^T and *R. vannielii* ATCC17100^T strengthens the separation into two species.

Protein-coding genes of *R. udaipurense* JA643^T have an average length of 830 bases, ranging from 70 to 4,796 bases. Out of 3,611 ORFs identified, 2,295 (63.55%) were functionally annotated, 1,238 (34.28%) were hypothetical, 64 were putative and 14 were conserved hypothetical genes. Most of the annotated protein coding sequences [~1,367 genes (62.74%)] were involved in metabolic and biosynthetic processes. In addition, 39 genes were involved in cell signaling and regulations, 26 genes in motility and chemotaxis, 119 genes in RNA metabolism, 74 genes in membrane transport, 168 genes in protein metabolism, 109 genes in DNA metabolism, 46 genes into defence, 88 genes in cell wall and capsule formation and 259 genes in other cellular functions, which were not distributed in subsystem category (Supplementary Figs S1 and S2).

When contig sequences of *R. udaipurense* JA643^T were mapped on *R. vannielii* ATCC17100^T genome sequence, only 2,989 genes got mapped, indicating that the remaining 622 genes were diverged from *R. vannielii* ATCC17100^T counterparts or absent in *R. vannielii* ATCC17100^T genome. *Rhodomicrobium udaipurense* JA643^T showed phototrophic growth, and it contained sets of genes for bacteriochlorophyll biosynthesis and a set of light harvesting complexes as well as *puf* L,M,C,H genes. A set of light harvesting complex gene sequences were observed with one α and a β gene. For N_2 fixation, two types of nitrogenase dimmers, molybdenum-dependent and vanadium-dependent nitrogenase, were present with all related cofactors and accessory genes. Secondary transport system genes included facilitator super family (MFS) transporter, resistance nodulation cell division (RND) pumps, tripartite ATP-dependent periplasmic (TRAP) transporter and organic solvent resistance transporter. The heavy metal efflux transporters and organic solvent resistance transporters present may allow the organism to survive in high concentrations of metals and other solvents.[16] *Rhodomicrobium udaipurense* JA643^T has the genes involved in the degradation of nitrogen-containing compounds like amino acids and heterocyclic aromatic compounds and may dehalogenate with the help of different monooxygenase, dioxygenase, peroxidases and dehalogenase.

Draft genome sequence of *R. udaipurense* JA643^T showed the presence of genes involved in benzoate degradation. Anaerobic benzoate degradation takes place by two known pathways; one occurs through –CoA ligation, while the other by hydroxylation, and *R. udaipurense* JA643^T has genes involved in both the pathways.[17] There are 12 genes related to aromatic hydrocarbon metabolism; of these, 6 genes are involved in benzoate degradation through hydroxylation. However, the RAST subsystem classifies genes involved in benzoate degradation through hydroxylation as nonfunctional in this organism. The presence of aromatic hydrocarbon metabolizing genes and utilization of benzoate by *R. udaipurense* JA643^T and *R. vannielii* ATCC17100^T indicate their probable role in the aromatic hydrocarbon recycling in oxic/anoxic environments.[1,17]

Some of homologous sequences unique in *R. udaipurense* JA643^T may compensate the function; sterol-binding domain-containing protein, hemolysins and related proteins containing CBS domain, putative homolysin, myosin cross-reactive streptococcal antigen, cobalt–zinc–cadmium/protein resistance, sporulation domain protein, sperimidine synthase, antibiotic biosynthesis monooxygenase, stage 0 sporulation protein J and sporulation initiation inhibitor protein. The summary of all these genes is given as Supplementary Table S2.

4. Hopanoid biosynthesis

Information on the genes involved in the biosynthesis of hopanoids is limited, and our study mainly focused on the genomic insight of hopanoid biosynthesis in *R. udaipurense* JA643^T.[3,4,6] Hopanoids are biosynthesized from isopentenyl pyrophosphate (IPP) which is synthesized either through mevalonate pathway or methyl erythritol phosphate pathway (MEP).[18] Genes involved in the MEP pathway are present in *R. udaipurense* JA643^T, and the genes are scattered in the genome (Fig. 1). Only

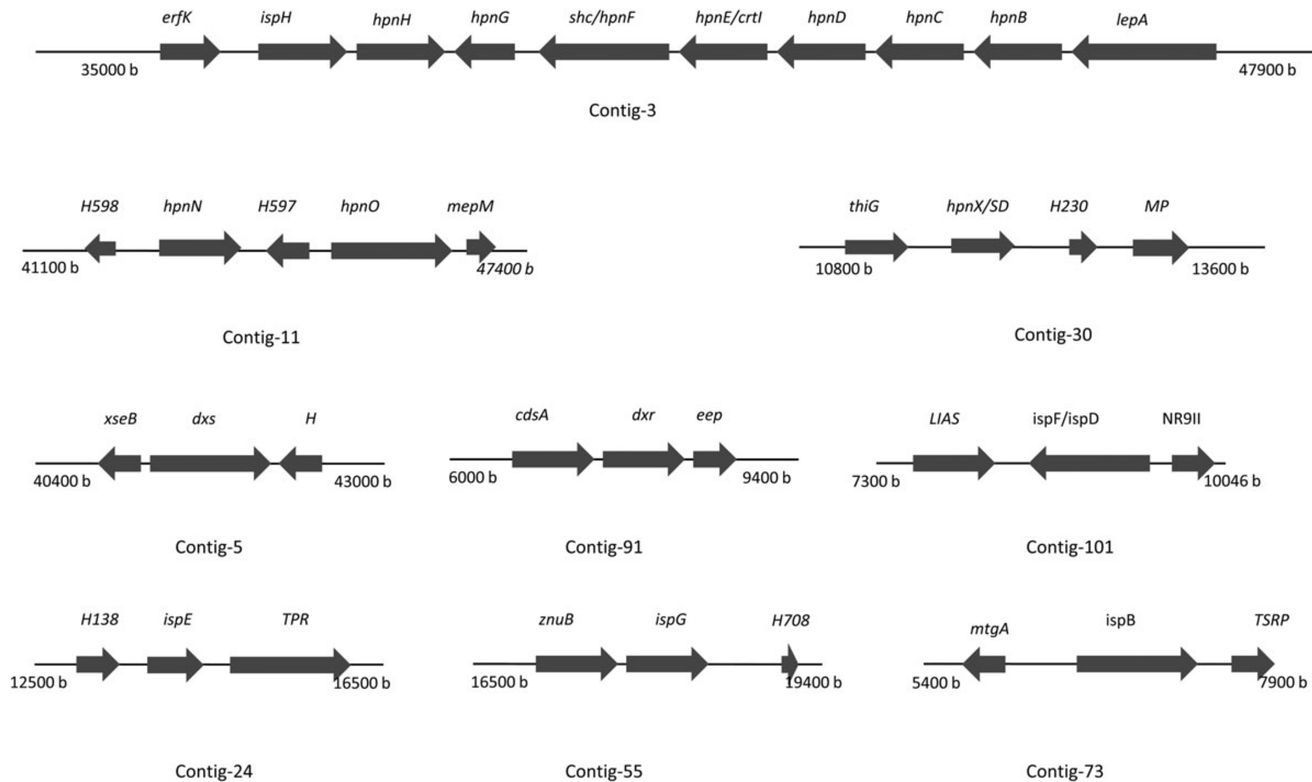

Figure 1. Location of the hopanoid biosynthetic genes in the genome of *Rhodomicrobium udaipurense* JA643^{T}. The representation is not up to the scale. *erfK: ErfK/YbiS/YcfS/YnhG* family gene; *ispH*: 4-hydroxy-3-methylbut-2-enyl diphosphate reductase; *hpnH*: radical SAM protein required for addition of adenosine to hopane skeleton; *hpnG*: adenosylhopane nucleoside/hopanoid-associated phosphorylase; *shc/hpnF*: squalene hopene cyclase; *hpnE/crtI*: squalene-associated, FAD-dependent desaturase/phytoene desaturase; *hpnD*: squalene synthase; *hpnC*: squalene synthase; *hpnI/hpnB*: glycosyl transferase, family 2; *lepA*: translation elongation factor LepA; mepM: murein endopeptidase; *hpnO*: *acetylornithine aminotransferase*; *H597*: hypothetical gene-597; *hpnN*: *hopanoid-associated RND transporter; H598*: hypothetical gene-598; *MP*: membrane protein; *H230*: hypothetical gene-230; *hpnX/SD*: sterol desaturase family protein; *thiG*: thiazole biosynthesis protein; H: hypothetical protein; *dxs*: 1-deoxy-D-xylulose 5-phosphate synthase; *xseB*: exodeoxyribonuclease VII small subunit; *cdsA*: phosphatidate cytidylyltransferase; *dxr*: *1-deoxy-D-xylulose 5-phosphate reductoisomerase; eep*: membrane-associated zinc metalloprotease; *LIAS*: lipoate synthase; *ispF/ispD*: 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; *NR9II*: nitrogen regulation protein NR(II); *TRP*: *TPR domain protein; ispE*: 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase; *H138*: hypothetical gene-138; *znuB*: zinc ABC transporter; *ispG*: 1-hydroxy-2-methyl-2-(*E*)-butenyl 4-diphosphate synthase; *H708*: hypothetical gene-708; *TSRP*: two-component system regulatory protein; *ispB*: eranyltranstransferase (farnesyldiphosphate synthase); *mtgA*: peptidoglycan transglycosylase.

4-hydroxy-3-methylbut-2-enyl diphosphate reductase coding gene (*ispH*) was present in the cluster of seven hopanoid biosynthesis genes. List of seven genes involved in the MEP pathway is given in the Supplementary Table S3. IPP isomerase catalyzes the conversion of relatively unreactive IPP to electrophile dimethylallyl pyrophosphate (DMAPP).[19] Farnesyldiphosphate synthase (*ispA*/*ispB*) catalyzes the formation of (2*Z*, 6*E*)-farnesyl diphosphate from three IPP.[20] Squalene synthase (*hpnC*/*hpnD*) catalyzes stepwise tail-to-tail addition of farnesyl diphosphate (C_{15}) which condenses to form presqualene pyrophosphate, an intermediate product.[5] Presqualene pyrophosphate is converted to squalene (I) by squalene synthase (*hpnC*/*hpnD*). Squalene synthase (*hpnC*/*hpnD*) is homologous to phytoene synthase (involved in carotenoid biosynthesis) and the function of *hpnC*/*hpnD* in the hopanoid biosynthesis is poorly understood.[21,22]

Eight genes involved in the hopanoid biosynthesis were located on contig no. 3 between 36,000 and 45,850 bp including *ispH* which encodes for 4-hydroxy-3-methylbut-2-enyl diphosphate reductase, an enzyme involved in the MEP pathway. These genes encode radical SAM protein (*hpnH*), adenosine nucleosidase (*hpnG*), squalene hopene cyclase (*shc*/*hpnF*), squalene-associated, FAD-dependent desaturase (*hpnE*/*crtI*), squalene synthase (*hpnD* and *hpnC*) and glycosyltransferase family protein (*hpnB*). Ten additional genes involved in hopanoid biosynthesis were scattered in the genome. The region having hopanoid biosynthetic genes and their upstream and downstream regions are shown in Fig. 1. The fragmented organization of the genes indicates the complex regulatory network of hopanoid biosynthesis. Squalene (I), hop-22(29)-ene (II), diplopterol (III), tetrahymanol (IV), 2-methyldiploptetrol (VI), adenosylhopane (VII), bacteriohopanetetrol (XI), aminobacteriohopanetetrol (XII), *N*-tryptophanyl aminobacteriohopanetetrol (XV) were identified through GC–MS analysis (Fig. 2), and their structures are given in Supplementary Fig. S3.

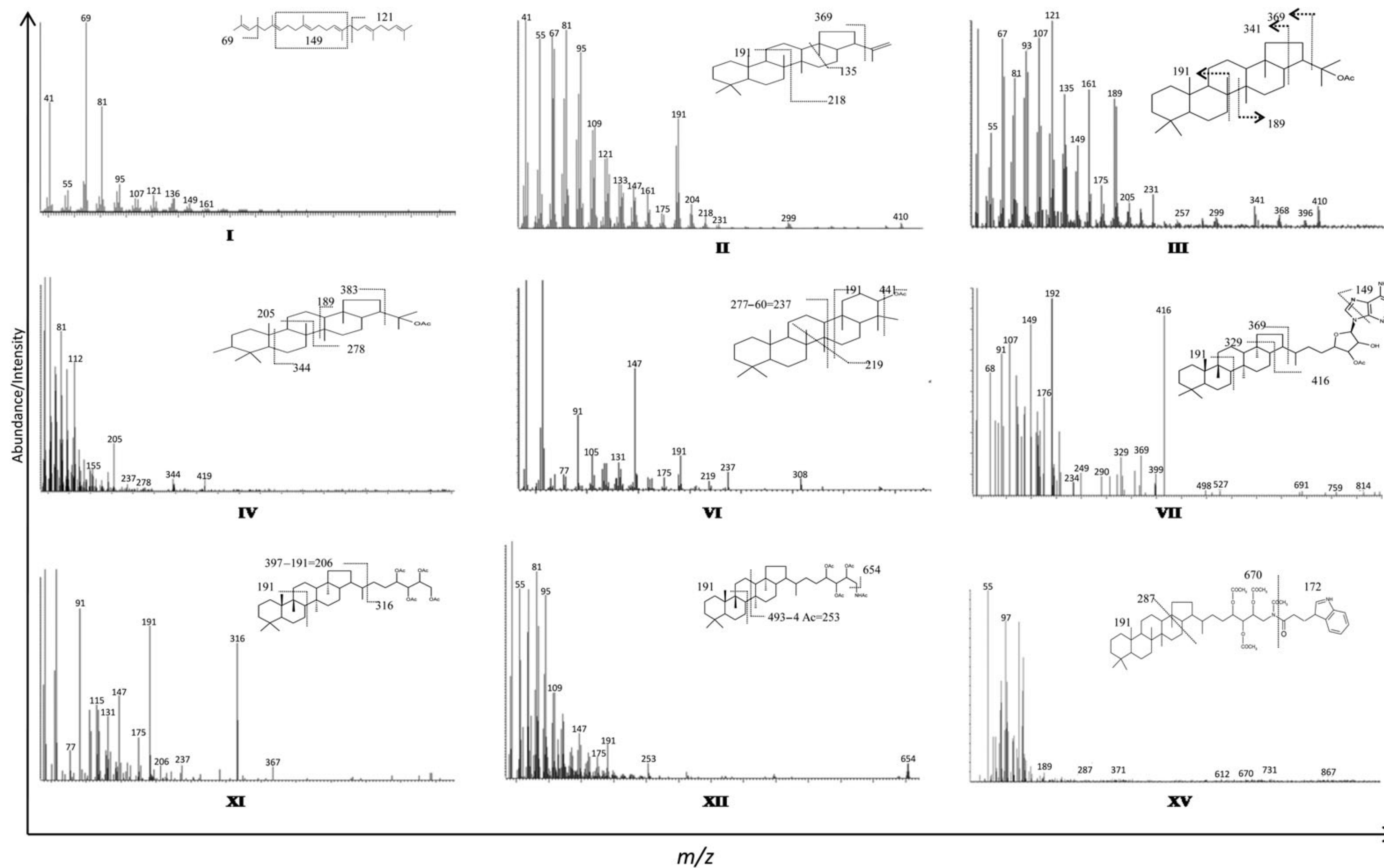

Figure 2. Mass spectra of hopanoids detected after acetylation of total lipid extract from the *Rhodomicrobium udaipurense* JA643^{T}. Identification of hopanoids was done by comparing the previously published data. Squalene (I), hop-22(29)-ene (II), hopan-22-ol (III), tetrahymanol (IV), 2-methyldiplopterol (VI), adenosylhopane (VII), bacteriohopanetetrol (XI), aminobacteriohopanetriol (XII) and *N*-tryptophanyl-35-aminobacteriohopanetriol (XV).

Table 2. Genes involved in the hopanoid biosynthesis pathway

Gene name	Code	EC number	Gene length (bp)	Similarity[a] (%)
Squalene hopene cyclase	*Shc/hpnF*	5.4.99.17	2,039	97
Glycosyl transferase, family 2, hpnB	*hpnI/hpnB*		1,166	99
5′-Methylthioadenosine nucleosidase/*S*-adenosylhomocysteine nucleosidase	*hpnG*	3.2.2.16/3.2.2.9	701	98
Radical SAM protein required for addition of adenosine to hopane skeleton	*hpnH*	–	1,142	99
Acetylornithine aminotransferase	*hpnO/argD*	2.6.1.11	1,199	98
Hopanoid-associated RND transporter	*hpnN*		2,666	98
Phytoene desaturase/squalene-associated FAD-dependent desaturase	*hpnE/crtI*	1.14.99	1,325	99
Phytoene synthase/squalene synthase	*hpnC*	2.5.1.32	857	97
Phytoene synthase/squalene synthase	*hpnD*	2.5.1.32	881	98
Sterol desaturase family protein	*erg3/hpnX*		776	97
Geranyltranstransferase (farnesyldiphosphate synthase)	*ispB/ispA*	2.5.1.10	890	95

[a]Similarity with respective genes of *Rhodomicrobium vannielii* ATCC17100^T.

In the hopanoid biosynthetic pathway, squalene (I) is converted to hop-22(29)-ene/diploptene (II) or 22-hopanol/diplopterol (III) by squalene hopene cyclase (*shc*) or squalene hopanol cyclase. Cyclization of squalene (I) into the pentacyclic tripterpenoid is one of the complex processes catalyzed by squalene hopene cyclase (*shc/hpnF*). A single enzyme catalyzes the stepwise formation of 13 covalent bonds and 9 stereo centres.[5] Accumulation of squalene (I) by *Δshc* mutant could not produce hopanoids including tetrahymanol (IV),[4] suggesting the role of *shc* gene in the biosynthesis of tetrahymanol (IV) as well. Phylogenetic tree using 339 *shc* genes from different microorganisms showed (Supplementary Fig. S4) a distinct clade for *Rhodomicrobium* spp. and has a sequence similarity of 97% between *shc* genes of *R. udaipurense* JA643^T and *R. vannielii* ATCC17100^T (Table 2).

hpnP encoding 2-methyltransferase was not identified in the draft genome sequence of *R. udaipurense* JA643^T. However, identification of a methylated hopanoid (2-methyldiplopterol; Fig 2) indicates the presence of *hpnP* gene in *R. udaipurense* JA643^T which might have not been covered in the draft genome sequence. The hop-22(29)-ene (II) is further converted to adenosylhopane (VII) by the radical SAM protein (*hpnH*) required for the addition of adenosine to hopane skeleton. *hpnH* gene product is a B_{12}-binding radical SAM protein, and mutation in *hpn*H resulted in the production of C_{30} hopanoids (diploptene and diplopterol).[6] Removal of adenine moiety from adenosylhopane (VII) carried out by adenosylhopane nucleoside/hopanoid-associated phosphorylase, encoded by *hpn*G gene (a putative nucleoside hydrolase) lead to the formation of ribosylhopane (VIII). Mutation in *hpn*G gene resulted in loss in the synthesis of bacteriohopanetetrol (XI) and aminobacteriohopanetriol (XII), and consequently an intermediate adenosylhopane (VII) was accumulated.[6]

A lactone hopanoid, ribosylhopane (VIII), was identified as an intermediate in the biosynthesis of C_{35} hopanoids.[23] The lactone ring of this hopanoid could be opened via a lactone hydrolysis, a reaction which is similar to the hydrolysis of the quorum-sensing molecule (acylhomeserine lactone [AHL]) by acylhomoserine lactonase.[24] However, gene *O*-acylhomoserine lactonase was not observed in the draft genome of *R. udaipurense* JA643^T. The lactone ring of hopanoid might be opened via oxidation to form putative ribonylhopane (IX) precursor.[25] The presence of ribonylhopane (IX) in some *R. palustris* mutants suggested that the biosynthesis of the C_{35} hopanoids would require an oxidation of the ribosylhopane (VIII) produced by the *hpnG* to form ribonylhopane (IX).[6] The pathway may proceed with the interconversion of ribonylhopane (IX) to formylhopane (X), because ribonylhopane (IX) is less reactive, whereas formylhopane (X) is more reactive. In the GC–MS analysis, ribosylhopane (VIII), ribonylhopane (IX) and formylhopane were not detected. The gene involved in the interconversion of ribosylhopane to formylhopane is also not yet known.

Formylhopane (X) may undergo reductive amination leading to the formation of an amino group containing hopanoid, aminobacteriohopanetriol (XII).[26] Knockout study of ornithine: oxo-acid aminotransferase (*hpnO)* showed that it was involved in the formation of aminobacteriohopanetriol (XII).[6] *hpnO* bears homology with *N*-acetylornithine aminotransferase (*ArgD*). *ArgD* acts on an aldehyde rather than a hydroxyl group supporting the idea that the formylhopane (X) is more likely to be the correct substrate. In GC–MS analysis, *N*-tryptophanyl-35-aminobacteriohopane-32,33,34-triol was observed, in which amino acid residue is linked to the

amino group of aminobacteriohopanetriol, indicating that there is a modification of aminobacteriohopanetriol in *R. udaipurense* JA643^T. The genes involved in the synthesis of bacteriohopanetetrol and *N*-tryptophanyl-35-aminobacteriohopane-32,33,34-triol are not yet known, but products were identified in GC–MS analysis.

Sterol desaturase (*erg32/hpnX*) gene was present on contig no. 30 in the genome of *R. udaipurense* JA643^T from 959,672 to 958,896 bp. Upstream and downstream regions are shown in Fig. 1. Sterol desaturase family protein known to catalyze formation of a C-5 double bond in the B ring of ergosterol. *erg32* is involved in the biosynthesis of ergosterol which is important for the plasma membrane structure, function and localization of the plasma membrane proteins. However, in *R. udaipurense* JA643^T sterols are not present. As hopanoids are the sterol surrogates and some hopanoids with double bond at C_6 and C_{11} positions were reported previously (XIV),[27] it is hypothesized that sterol desaturase (*hpnX*) might be responsible for the incorporation of double bond in the hopanoids.

hpnI/*hpnB* genes are associated with hopanoid glycosyltransferase, and glycosylated hopanoids were identified with *Geobacter sulfurreducens* and *Geobacter metallireducens.*[28,29] Hopanoids with glycosyl moiety is expected in *R. udaipurense* JA643^T since *hpnI*/*hpnB* genes were observed in the genome of this organism; however, no glycosyl group containing hopanoids were detected in the GC–MS analysis. *hpnC*, *hpnD* and *hpnE* genes are involved in the modification of squalene. *hpnE* is a squalene-associated, FAD-dependent desaturase which has oxidoreductase activity. Perzl *et al.* suggested that it might be involved as an oxidoreductase in hopanoid or terpenoid biosynthesis pathways.[21] However, its role in hopanoid biosynthesis is not understood. Gene *hpnA* encode for NAD-dependent nucleoside diphosphate–sugar epimerase/dehydrates protein, which was not observed in the draft genome of *R. udaipurense* JA643^T. Hopanoid biosynthetic genes between *R. udaipurense* JA643^T and *R. vannielii* ATCC17100^T were compared. Sequence similarity ranges from 95 to 99% between them (Table 2,

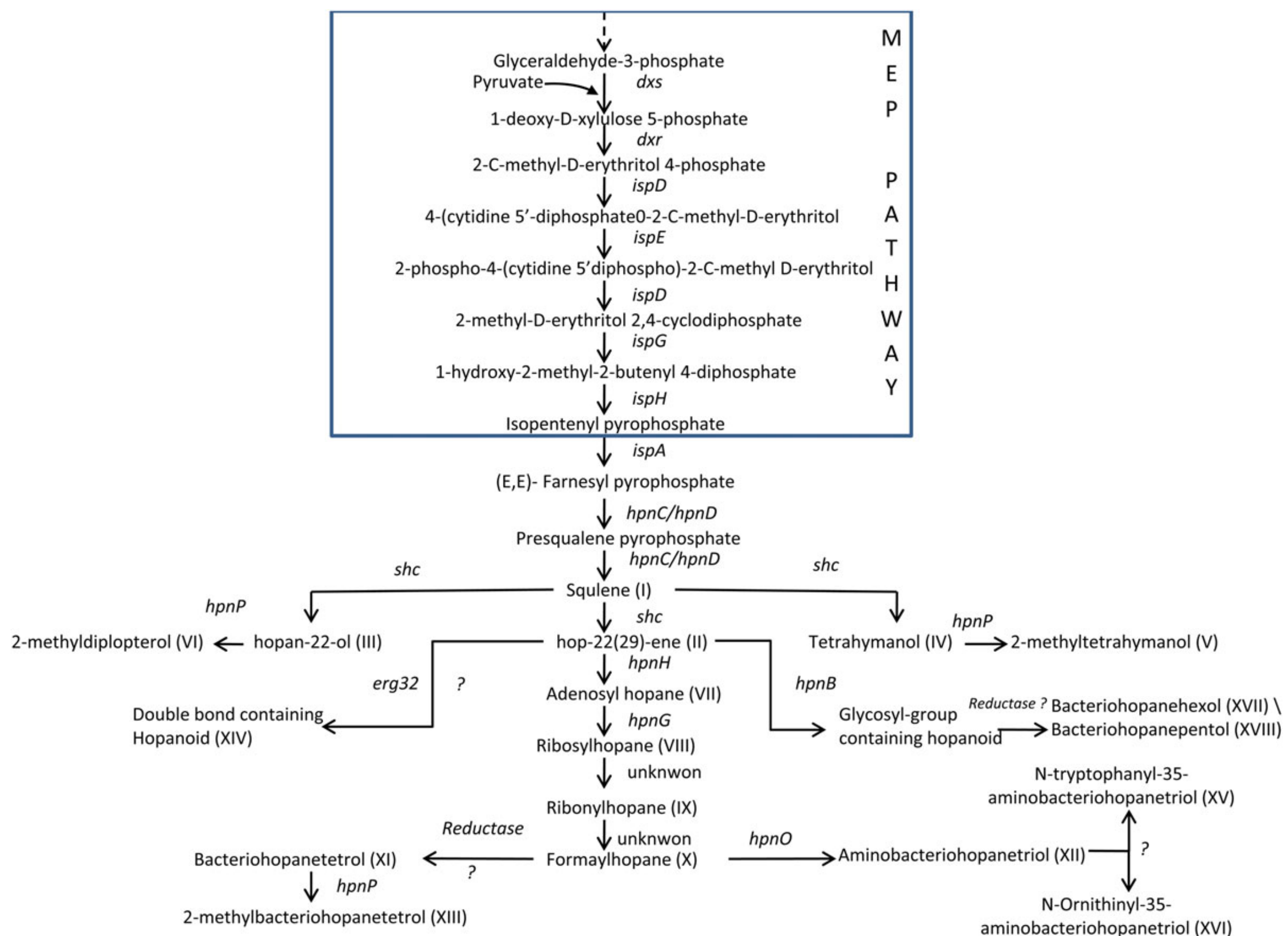

Figure 3. Schematic representation of hopanoid biosynthetic pathway in *Rhodomicrobium udaipurense* JA643^T based on genes identified, GC–MS analysis and previous reports.[3,4,15–26,30]

Supplementary Table S3). The hopanoid biosynthetic pathway was then inferred based on the genes identified, experimental validation of hopanoid molecules and available literature (Fig. 3).[3,4,15–30]

Apart from hopanoid biosynthetic genes, few genes playing a key role in localization of hopanoids, hopanoid regulation and transportation were also observed in *R. udaipurense* JA643^T. One such gene is hopanoid-associated RND transporter (*hpnN*) which encodes an RND-like transporter that helps to localize hopanoids to the outer membrane. *hpnN* and *hpnO* were located close together with a gene coding for hypothetical protein in between. *ΔhpnN* mutant no longer contain any hopanoids in the outer membrane, suggesting its importance in hopanoid transportation.[31] *hpnN* mutant produces more 2-methylhopanoids, possibly indicating some role for *hpnN* in hopanoid homeostasis.[31] Similarly, extracellular sigma factor (ECF) regulates the methylation of hopanoids at C-2 position.[32] Genome sequence of *R. udaipurense* JA643^T showed that it has two RNA polymerase ECF-type sigma factors. Probably these sigma factors may be the regulators for the expression of *hpnP* gene and production of methylated hopanoids.

In summary, genome sequence of *R. udaipurense* JA643^T has a set of genes required for degrading aromatic compounds and resistance to solvents. It has all the machinery required for photosynthesis. There were 18 genes involved in the hopanoid biosynthesis and transport of hopanoids in this organism. The MEP pathway exists in *R. udaipurense* JA643^T which leads to the formation of IPP. Major hopanoids were diploptene, diplopterol, bacteriohopanetetrol, aminobacteriohopanetriol and *N*-tryptophanyl-35-aminobacteriohopanetriol. Though *hpnP* gene was not covered in the genome of *R. udaipurense* JA643^T, the presence of methylated hopanoids (Fig. 2) indicates that this gene might be present in the complete genome of this organism. On the other hand, we could not find *hpnP* gene from the complete genome of *R. vannielii* ATCC17100^T, and this makes the difference in the hopanoids of *R. udaipurense* JA643^T. The genes involved in the formation of amino acid group containing hopanoids need an in-depth investigation.

Hopanoid-deficit mutants can grow at optimized laboratory conditions; however, the mutants act differently under stress conditions. Hopanoids confer membrane integrity and help the bacterium to tolerate bile salts, pH, antibiotics and temperature.[4] Hopanoids are also important for a bacterium to undergo normal cell cycle.[31] However, the role of hopanoids in *R. udaipurense* JA643^T is yet to be revealed, and this part of the work is in progress at our laboratory, while comparisons are also being made with *R. vannielii* ATCC17100^T, whose complete genome sequence is available.

4.1. Data access

This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession JFZJ00000000 and SRR1016881. The version described in this paper is version JFZJ01000000.

Acknowledgements: Our research group acknowledges Mr Ashwin Kumar Mishra for the help in the annotations and other related work. The help of Mr B. Prasanth in the GC–MS data analysis is greatly acknowledged. All the authors thank Prof. G. Padmaja and Dr Sharmistha Banerjee, Associate Professor, University of Hyderabad, for language editing the manuscript.

Supplementary Data: Supplementary data are available at www.dnaresearch.oxfordjournals.org

Funding

This work was funded by Council of Scientific and Industrial Research (CSIR) project. T.L. is supported by research fellowship from the University Grants Commission (UGC). Funding to pay the Open Access publication charges for this article was provided by Council of Scientific and Industrial Research (CSIR) project [Grant ID 38(1351)/13/EMR-II dated on 14-2-2013].

References

1. Welander, P.V., Coleman, M.L., Sessions, A.L., Summons, R.E. and Newman, D.K. 2010, Identification of a methylase required for 2-methylhopanoid production and implications for the interpretation of sedimentary hopanes, *Proc. Natl. Acad. Sci. USA*, **107**, 8537–42.
2. Ourisson, G. and Albrecht, P. 1992, Hopanoids. 1. Geohopanoids: the most abundant natural products on Earth? *Acc. Chem. Res.*, **25**, 398–402.
3. Malott, R.J., Steen-Kinnaird, B.R., Lee, T.D. and Speert, D.P. 2011, Identification of hopanoid biosynthesis genes involved in polymyxin resistance in *Burkholderia multivorans*, *Antimicrob. Agents Chemother.*, **56**, 464–71.
4. Welander, P.V., Hunter, R.C., Zhang, L., Sessions, A.L., Summons, R.E. and Newman, D.K. 2009, Hopanoids play a role in membrane integrity and pH homeostasis in *Rhodopseudomonas palustris* TIE-1, *J. Bacteriol.*, **191**, 6145–56.
5. Siedenburg, G. and Jendrossek, D. 2011, Squalene-hopene cyclases, *Appl. Environ. Microbiol.*, **77**, 3905–15.
6. Welander, P.V., Doughty, D.M., Wu, C.H., Mehay, S., Summons, R.E. and Newman, D.K. 2012, Identification and characterization of *Rhodopseudomonas palustris* TIE-1 hopanoid biosynthesis mutants, *Geobiology*, **10**, 163–77.
7. Venkata, V.R., Shalem Raj, P., Tushar, L., Sasikala, Ch. and Ramana, Ch.V. 2013, *Rhodomicrobium udaipurense* sp. nov., a psychrotolerant, phototrophic alphaproteobacterium isolated from a freshwater stream, *Int. J. Syst. Evol. Microbiol.*, **63**, 2684–9.

8. Wright, G.E. and Madigan, M.T. 1991, Photocatabolism of aromatic compounds by the phototrophic purple bacterium *Rhodomicrobium vannielii*, *Appl. Environ. Microbiol.*, **57**, 2069−73.
9. Lakshmi, K.V.N.S., Sasikala, Ch. and Ramana, Ch. 2009, *Rhodoplanes pokkaliisoli* sp. nov., a phototrophic alphaproteobacterium isolated from a waterlogged brackish paddy soil, *Int. J. Syst. Evol. Microbiol.*, **59**, 2153−7.
10. Subhash, Y., Tushar, L., Sasikala, C. and Ramana, C.V. 2013, *Falsirhodobacter halotolerans* gen. nov., sp. nov., isolated from dry soils of a solar saltern, *Int. J. Syst. Evol. Microbiol.*, **63**, 2132−7.
11. Aziz, R.K., Bartels, D., Best, A.A., et al. 2008, The RAST server: rapid annotations using subsystems technology, *BMC Genomics.*, **9**, 75.
12. Lagesen, K., Hallin, P.F., Rodland, E., Staerfeldt, H.H., Rongnes, T. and Ussery, D.W. 2007, RNAmmer: consistent and rapid annotations of ribosomal RNA genes, *Nucleic Acid Res.*, **35**, 3100−8.
13. Mende, D.R., Sunagawa, S., Zeller, G. and Bork, P. 2013, Accurate and universal delianeation of prokaryotic species, *Nat. Method.*, **10**, 881−4.
14. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. 2011, MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance and maximum parsimony methods, *Mol. Biol. Evol.*, **28**, 2731−9.
15. Grant, J.R. and Stothard, P. 2008, The CGView Server: a comparative genomics tool for circular genomes, *Nucleic Acids Res.*, **36**, 181−4.
16. Wayne, L., Brenner, D.J. and Colwell, R.R., et al. 1987, Report of the Ad Hoc Committee on reconciliation of approaches to bacterial systematics, *Int. J. Syst. Bacteriol.*, **37**, 463−4.
17. Harwood, C.S., Burchhardt, G., Herrmann, H. and Fuchs, G. 1999, Anaerobic metabolism of aromatic compounds via the benzoyl-CoA pathway, *FEMS Microbiol. Rev.*, **22**, 439−58.
18. Rohmer, M. 2008, From molecular fossils of bacterial hopanoids to the formation of isoprene units: discovery and elucidation of the methylerythritol phosphate pathway, *Lipids*, **43**, 1095−107.
19. Street, I.P., Coffman, H.R., Baker, J.A. and Poulter, C.D. 1990, Isopentenyldiphosphate: dimethylallyaldiphosphate isomerase: construction of a high-level heterologous expression system for the gene from *Saccharomyces cerevisiae* and identification of an active-site nucleophile, *Biochemistry*, **29**, 7531−8.
20. Cantera, J.J., Kawasaki, H. and Seki, T. 2002, Farnesyl diphosphate synthase gene of three phototrophic bacteria and its use as a phylogenetic marker, *Int. J. Syst. Evol. Microbiol.*, **52**, 1953−60.
21. Perzl, M., Reipen, I.G., Schmitz, S., et al. 1998, Cloning of conserved genes from *Zymomonas mobilis* and *Bradyrhizobium japonicum* that function in the biosynthesis of hopanoid lipids, *Biochim. Biophys. Acta*, **1393**, 108−18.
22. Kannenberg, E.L. and Poralla, K. 1999, Hopanoid biosynthesis and function in bacteria, *Naturwissenschaften.*, **86**, 168−76.
23. Seeman, M., Bisseret, P., Tritz, J.P., Hooper, A.B. and Rohmer, M. 1999, Novel bacterial triterpenoids of the hopane series from *Nitrosomonas europaea* and their significance for the formation of the C-35 bacteriohopane skeleton., *Tetra. Lett.*, **40**, 1681−4.
24. Wang, W.Z., Morohoshi, T., Ikenoya, M., Someya, N. and Ikeda, T. 2009, AiiM, a novel class of N- acylhomoserine lactonase from the leaf associated bacterium *Microbacterium testaceum*, *Appl. Environ. Microbiol.*, **76**, 2524−30.
25. Duvold, T. and Rohmer, M. 1999, Synthesis of ribosylhopane, the putative biosynthetic precursor of bacterial triterpenoids of the hopane series, *Tetrahedron.*, **55**, 9847−58.
26. Rohmer, M. 1993, The biosynthesis of triterpenoids of the hopane series in the Eubacteria: a mine of new enzyme reactions, *Pure Appl. Chem.*, **65**, 1293−8.
27. Simonin, P., Tindall, B. and Rohmer, M. 1994, Structure elucidation and biosynthesis of 31-methylhopanoids from *Acetobacter europaeus* studies on a new series of bacterial triterpenoids, *Eur. J. Biochem.*, **225**, 765−71.
28. Methe, B.A., Nelson, K.E., Eisen, J.A., et al. 2003, Genome of *Geobacter sulfurreducens*: metal reduction in subsurface environments, *Science*, **302**, 1967−9.
29. Eickhoff, M., Birgel, D., Talbot, H.M., Peckmann, J. and Kappler, A. 2013, Bacteriohopanoid inventory of *Geobacter sulfurreducens* and *Geobacter metallireducens*, *Org. Geochem.*, **58**, 107−14.
30. Meguro, N., Kodama, Y., Gallegos, M.T. and Watanabe, K. 2005, Molecular characterization of resistance nodulation division transporters from solvent and drug resistant bacteria in petroleum contaminated soil, *Appl. Environ. Microbiol.*, **71**, 580−6.
31. Doughty, D.M., Coleman, M.L., Hunter, R.C., Sessions, S.L., Summons, R.E. and Newman, D.K. 2011, The RND-family transporter, HpnN, is required for hopanoid localization to the outer membrane of *Rhodopseudomonas palustris* TIE-1, *Proc. Natl. Acad. Sci. USA*, **108**, E1045−51.
32. Kulkarni, G., Wu, C.H. and Newman, D.K. 2013, The general stress response factor EcfG expression of the C-2 hopanoid methylase HpnP in *Rhodopseudomonas palustris* TIE-1, *J. Bacteriol.*, **195**, 2490−8.