

Supplementary files of

**Horizontal Gene Transfer of Genes Encoding Copper-Containing Membrane-Bound Monooxygenase (CuMMO) and Soluble Di-iron Monooxygenase (SDIMO) in Ethane- and Propane-Oxidizing *Rhodococcus* Bacteria**

Bin Zou<sup>a</sup>, Ying Huang<sup>a</sup>, Pan-Pan Zhang<sup>a</sup>, Xiao-Ming Ding<sup>a</sup>, Huub J.M. Op den Camp<sup>b</sup>, Zhe-Xue Quan<sup>a,c,\*</sup>

<sup>a</sup>Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, Institute of Biodiversity Science, School of Life Sciences, Fudan University, Shanghai, China

<sup>b</sup>Department of Microbiology, Radboud University, Nijmegen, Netherlands

<sup>c</sup>Shanghai Engineering Research Center of Industrial Microorganisms, School of Life Sciences, Fudan University, Shanghai, China

\*Address correspondence to Zhe-Xue Quan, [quanzx@fudan.edu.cn](mailto:quanzx@fudan.edu.cn)

Table S1 Growth of strains on different substrates

Substrate <sup>1</sup>	ZPP	ZPP ( $\Delta$ <i>smoXYB1C1Z-hmoCAB</i> )	<i>R. erythropolis</i> ATCC 25544 (pRZP1)	<i>R. erythropolis</i> ATCC 25544
Methane	- <sup>2</sup> (10 d)	n.a. <sup>3</sup>	n.a.	n.a.
Methanol	- (10 d)	- (10 d)	- (10 d)	- (10 d)
Ethanol	0.71±0.02 (5 d)	0.28±0.05 (5 d)	0.36±0.03 (5 d)	0.14±0.19 (5 d)
1-propanol	1.11±0.01 (5 d)	0.66±0.12 (5 d)	0.59±0.04 (5 d)	0.35±0.03 (5 d)
2-propanol	0.77±0.02 (8 d)	0.14±0.00 (8 d)	- (10 d)	- (10 d)

<sup>1</sup>The consumption of methane was measured by gas chromatography. The consumption of methanol, ethanol, 1-propanol, and 2-propanol was measured through growth of cells at OD<sub>600</sub>.

<sup>2</sup>- means no consumption or growth was measured, compared with the blank controls.

<sup>3</sup>n.a. (not applicable) means no data provided.

Table S2 Representative *IS* elements in pRZP1 and homologous regions of other seven plasmids

Strain	Annotation	Identity	Identity Rate	Query Start	Query End	Subject strand
ZPP plasmid pRZP1 (CP058640)	<i>/SRjo4 Family: IS256</i>	1311/1445	91%	65651	67093	Minus
	<i>/SRjo4 Family: IS256</i>	1311/1445	91%	68606	70048	Plus
	<i>/SRjo1 Family: IS1380</i>	1473/1694	87%	98851	100534	Minus
	<i>/SRjo4 Family: IS256</i>	1311/1445	91%	102764	104206	Minus
	<i>/SRhosp5 Family: IS3</i>	1164/1331	87%	104964	106282	Minus
	<i>/SRhosp1 Family: IS3</i>	1185/1323	90%	109435	110754	Plus
	<i>/SMpa1 Family: IS110</i>	285/349	82%	165650	165998	Plus
	<i>/SMpa1 Family: IS110</i>	358/412	87%	166263	166674	Plus
	<i>/SRhosp5 Family: IS3</i>	1164/1331	87%	172458	173776	Minus
	<i>/IS1164 Family: IS256</i>	1090/1310	83%	191152	192456	Plus
	<i>/SRhosp5 Family: IS3</i>	1164/1332	87%	194048	195367	Minus
	<i>/SRhosp2 Family: IS701</i>	287/319	90%	205051	205367	Minus
	<i>/SRjo4 Family: IS256</i>	1311/1445	91%	207559	209001	Minus
	<i>/SRhosp1 Family: IS3</i>	1209/1321	92%	225623	226940	Minus
	<i>/SRhosp5 Family: IS3</i>	391/422	93%	231550	231970	Plus
<i>/SRhosp5 Family: IS3</i>	565/601	94%	236159	236758	Plus	
B4 plasmid pROB01 (AP011116)	<i>/SRhosp1 Family: IS3</i>	317/326	97%	41325	41650	Minus
	<i>/SRhosp1 Family: IS3</i>	817/841	97%	46753	47590	Minus
	<i>/SRhosp1 Family: IS3</i>	205/213	96%	48717	48929	Minus
	<i>/SRop2 Family: ISAzo13</i>	2171/2171	100%	75197	77367	Minus
RHA1 plasmid pRHL2 (CP000433)	<i>/SRjo2 Family: IS630</i>	1100/1147	96%	75996	77142	Minus
	<i>/SRjo1 Family: IS1380</i>	610/715	85%	95394	96102	Plus
	<i>/SRhosp2 Family: IS701</i>	244/292	84%	114418	114706	Plus
R7 plasmid pPDG3 (CP008950)	<i>/SArsp6 Family: Tn3</i>	202/235	86%	119086	119320	Plus
	<i>/SRhosp5 Family: IS3</i>	1211/1334	91%	126537	127865	Minus
S2-17 plasmid pRB29 (CP021356)	<i>/IS1164 Family: IS256</i>	690/815	85%	212226	213040	Plus
	<i>/SRop2 Family: ISAzo13</i>	2128/2171	98%	213049	215219	Plus
	<i>/IS1164 Family: IS256</i>	288/348	83%	215282	215629	Plus
	<i>/SMav1 Family: IS21</i>	283/332	85%	237792	238123	Minus
NJ-530 plasmid unnamed4 (CP034156)	<i>/STesp1 Family: IS3</i>	301/380	79%	162549	162926	Minus
WAY2 plasmid pRWAY03 (CP046575)	<i>/SMav1 Family: IS21</i>	300/356	84%	335748	336103	Plus

*R. erythropolis* PR4 plasmid pREL1 (AP008931) has no representative *IS* elements in homologous regions.

Fig. S1

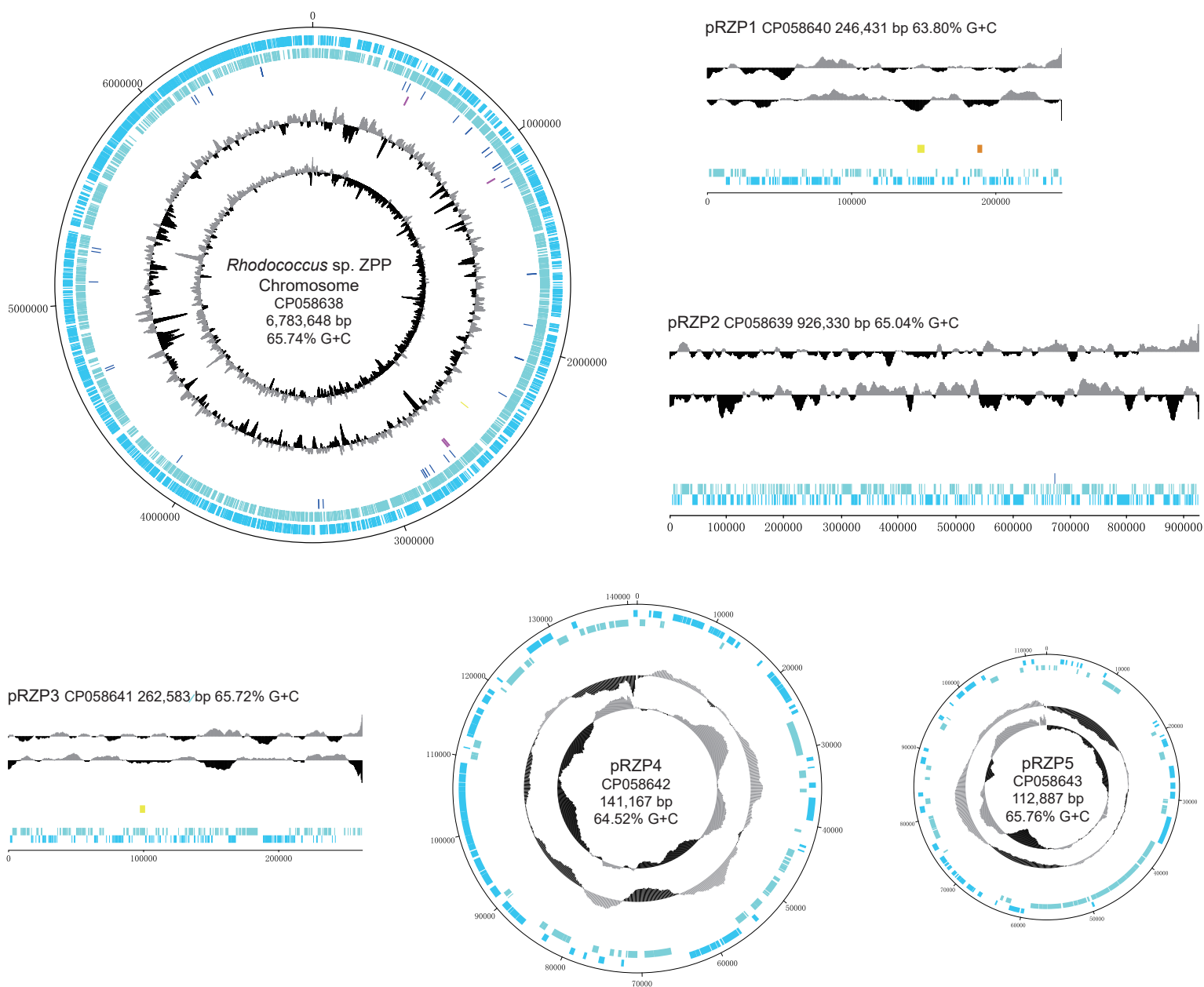


Fig. S1. Genomic map of *Rhodococcus* sp. ZPP replicons. Chromosome, linear mega-plasmids pRZP1, pRZP2, and pRZP3, and circular plasmids pRZP4 and pRZP5 are shown. The outer/bottom two rings/rows represent coding sequences in the forward and reverse strands. The third circle represents tRNA genes. The fourth circle represents rRNA genes. The fifth circle represents the operon of HmoCAB genes (orange on pRZP1), the operon of group 3-like SDIMO (yellow on pRZP1), the operon of group 5 SDIMO (yellow on chromosome), and the operon of group 6 SDIMO (yellow on pRZP3). the sixth circle/row represents deviation in G+C content, and the seventh circle/row represents the G+C skew.



Fig. S3

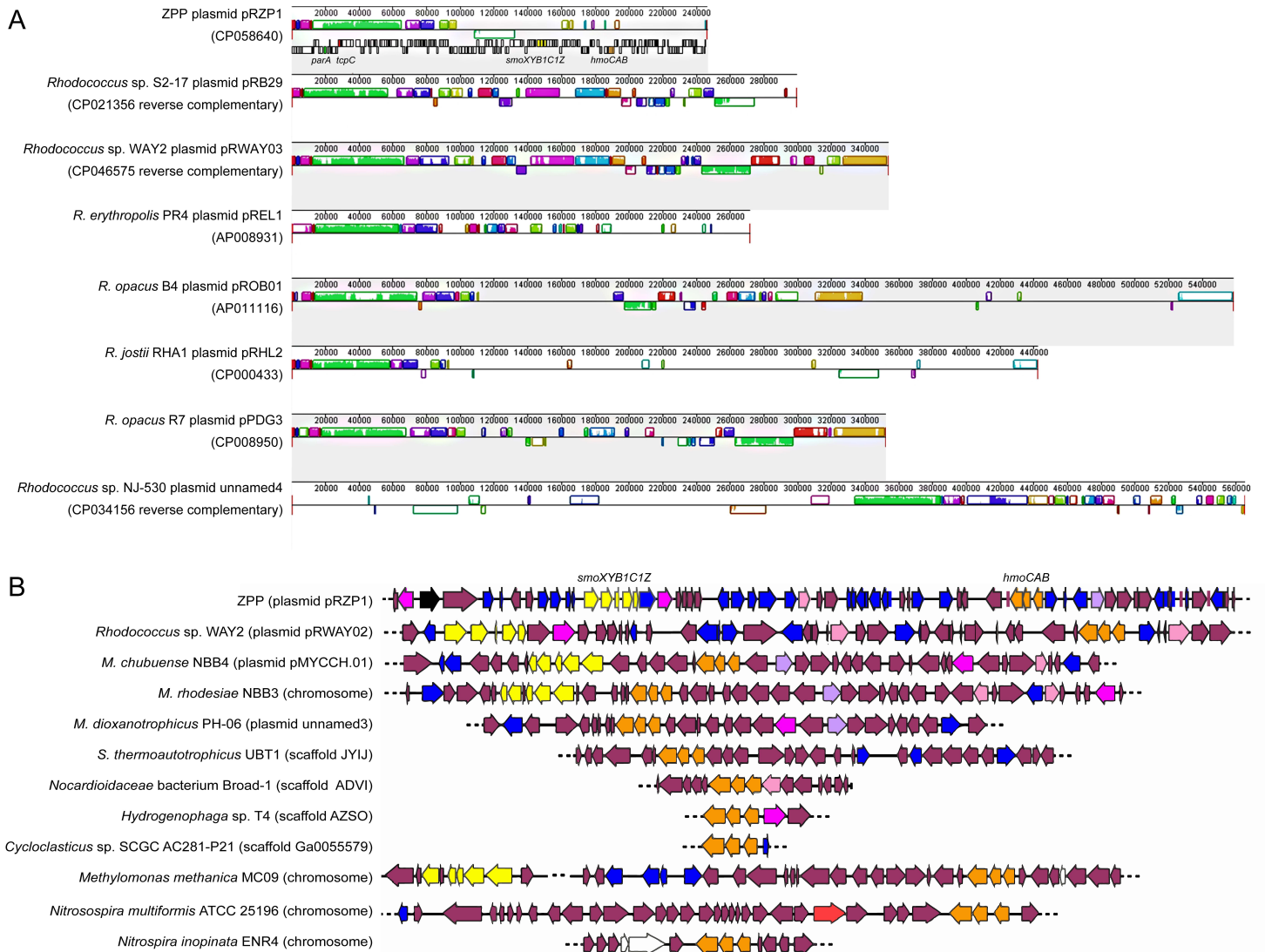


Fig. S3. The genomic alignment of eight plasmids in different strains of *Rhodococcus* (A) and the clusters of CuMMO and SDIMO genes in representative methane-, short-chain alkane-, and ammonia-oxidizing microbes (B). The alignment in (A) was performed using Mauve (1), and the homologous regions are in the same color. The thumbnail of pRZP1 annotation are drawn below the homologous region of pRZP1, and *parA*, *tcpC*, *smoXYB1C1Z* and *hmoCAB* are marked at the corresponding position. In (B), The SDIMO gene (*smoXYB1C1Z*) clusters are marked in yellow, and the CuMMO gene (*hmoCAB*, *amoCAB* and *pmoCAB*) clusters are marked in orange. The alcohol dehydrogenase genes (*adh*) are marked in light pink; the NDMA-dependent methanol dehydrogenase genes (*mdh*) are marked in purple; the aldehyde dehydrogenase genes (*aldh*) are marked in pink; the acetate-CoA ligase gene (*acs*) is marked in black; the periplasmic copper binding protein (*copCD*) is marked in white; the hydroxylamine oxidoreductase gene is marked in red and the mobile elements (including *IS* elements, transposases, and phages) are marked in blue.

Fig. S4

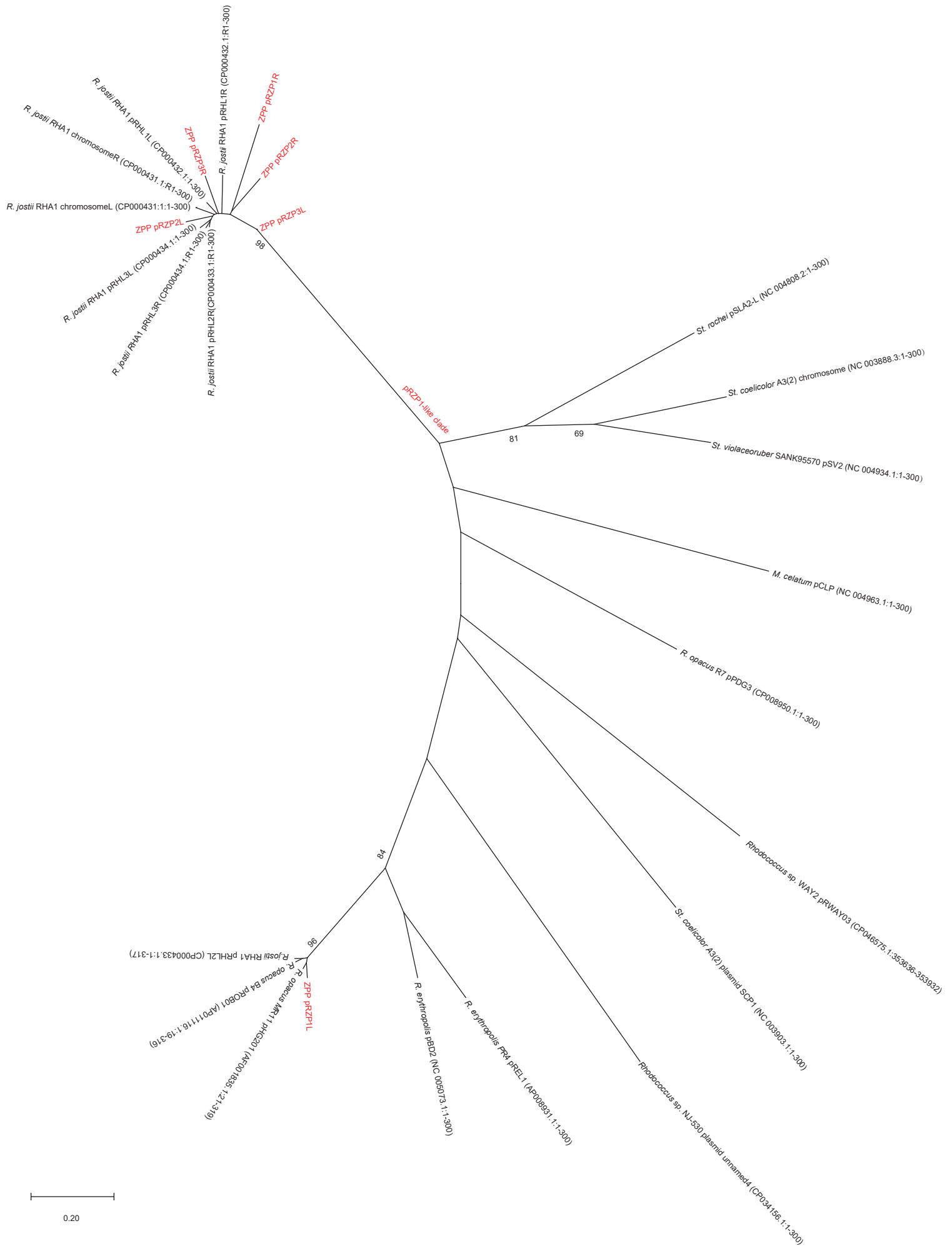


Fig. S4. Phylogenetic analysis of telomeres from linear replicons. The first 300-bp from telomeres were aligned using ClustalW with default parameters and a 500 bootstrapped maximum likelihood tree was constructed using MEGA X. The end of the replicon is designated with a subscript L for left end and R for right. Where the telomeres were identical or not resolved on the scale of this tree, right and left are not indicated. Nodes whose bootstraps  $\geq 60$  are shown. The tree was based on Supporting Figure 4 of a previous study (2).



Fig. S5

1

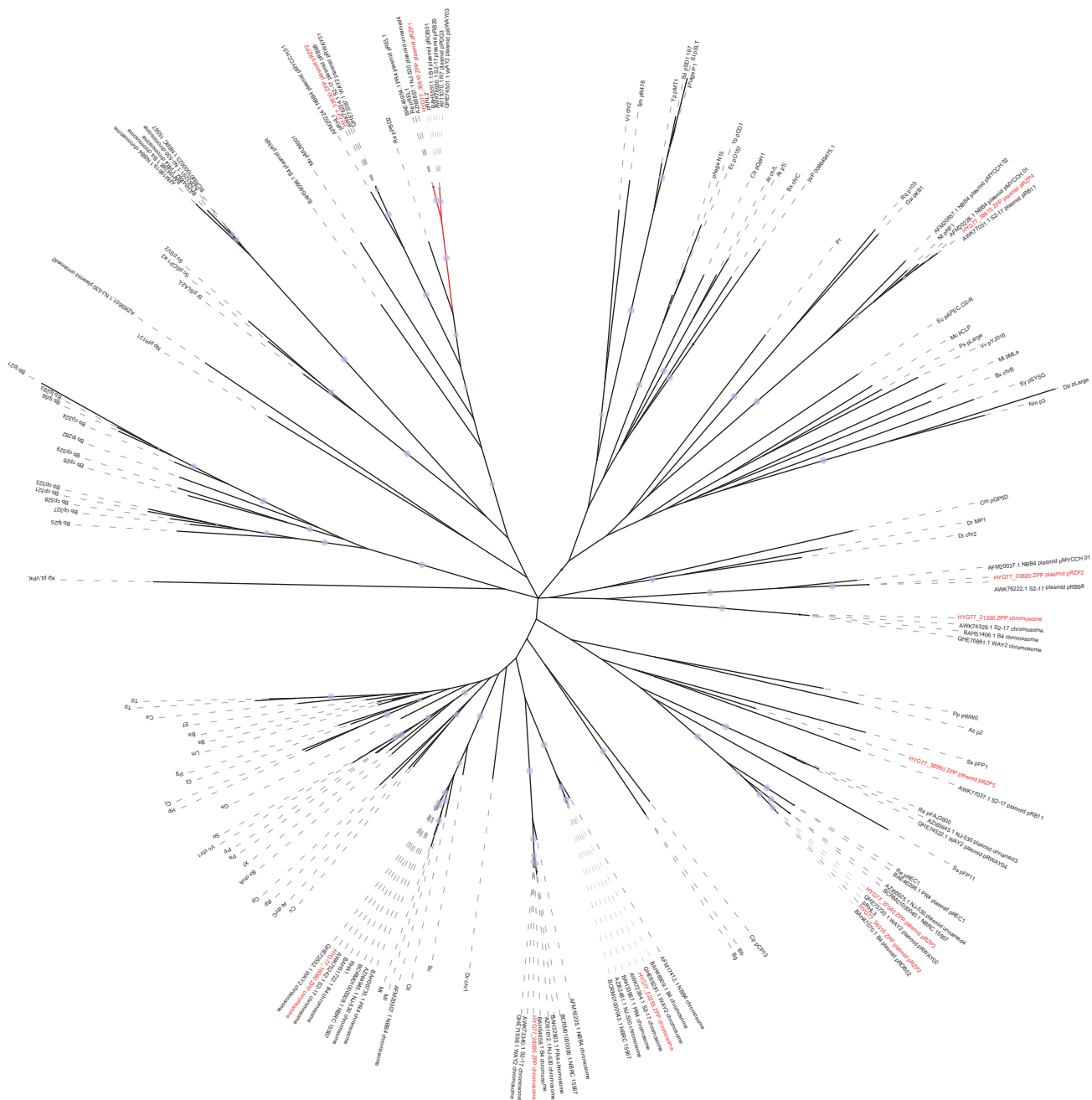


Fig. S5 Phylogenetic tree of ParA proteins. Proteins were aligned using Muscle with default parameters and 100 bootstraps with 95% partial deletion. A maximum likelihood tree was constructed using MEGA X. Nodes whose bootstraps  $\geq 60$  are colored in purple. The branch of pRZP1 is colored in red. Genomes related to this study are labelled with their accession numbers. The tree was based on Supporting Figure 3 of a previous study (2).

Fig. S6

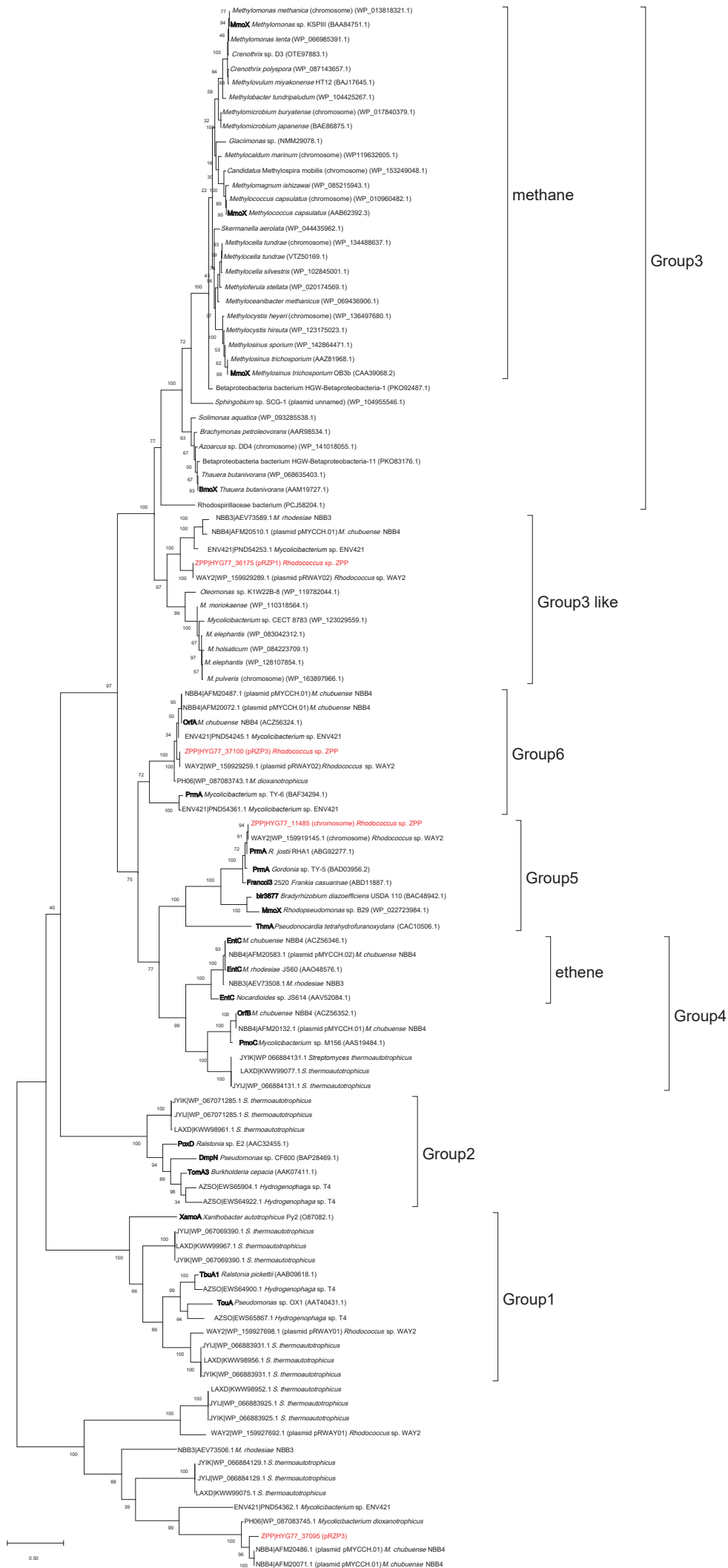


Fig. S6. Phylogenetic tree of *smoXYB1C1Z*-encoding group 3-like soluble di-iron monooxygenases (SDIMO) subunit  $\alpha$ . MEGA X was used to build the maximum likelihood tree. The number of bootstraps was 100, and the partial deletion was set to 95%. The sequence that has a pipe-delimited name is the sequence appearing in the HmoA tree (Fig. 3). The marker (chromosome) indicates that the gene is located on the chromosome, and the marker (plasmid) indicates that it is located on the plasmid. The tree was based on previous work and text in bold indicates that these sequences were extracted from that study (3).

Fig. S7

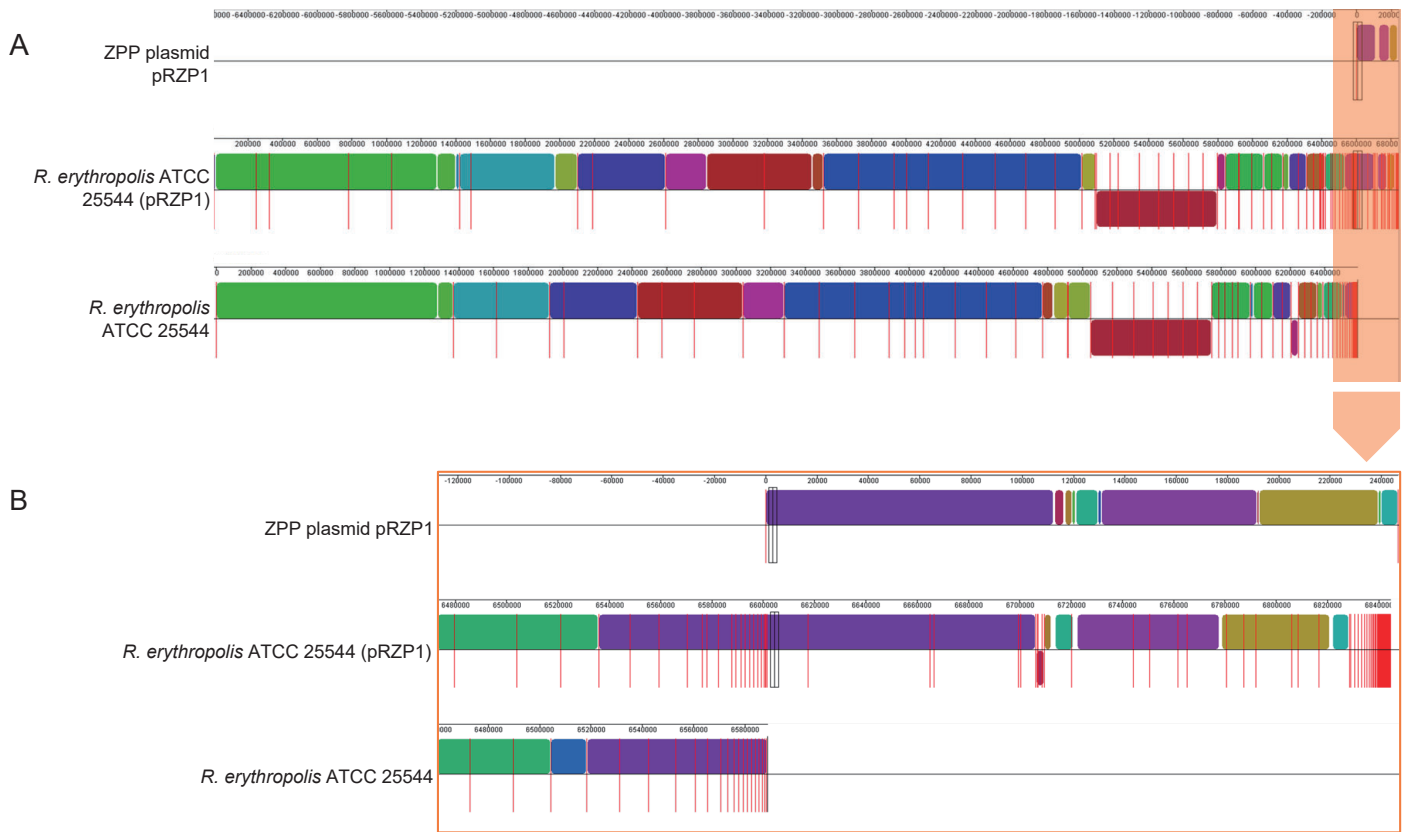


Fig. S7. Genomic comparison of *Rhodococcus* sp. ZPP plasmid pRZP1, ATCC 25544 (pRZP1), and ATCC 25544 (NCBI Assembly ID: GCA\_001552595.1). (A) is the entire genomic alignment and (B) is the region of pRZP1.

Fig. S8

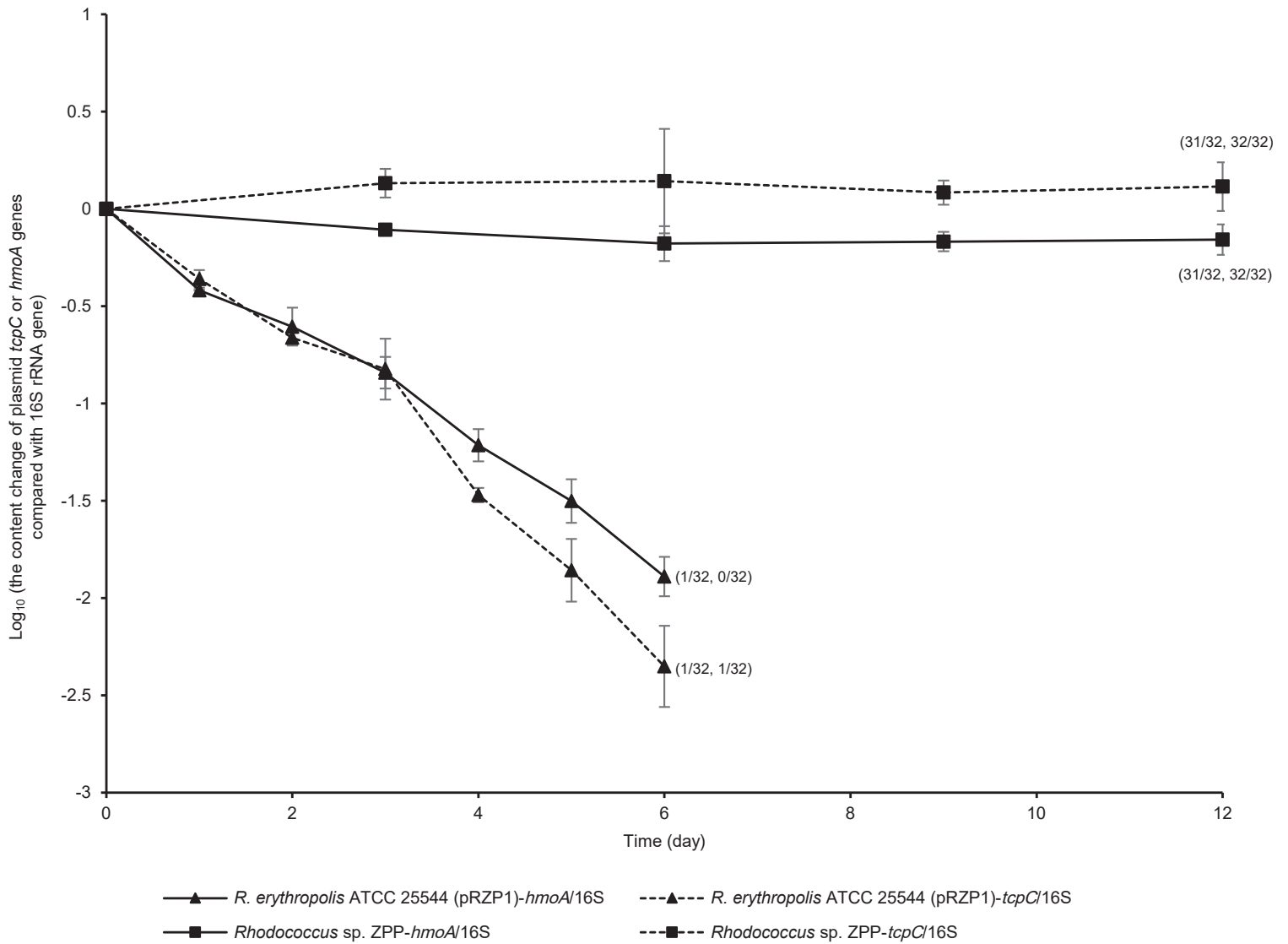


Fig. S8. Plasmid loss of wild-type ZPP and the transconjugant ATCC 25544 (pRZP1) using quantitative PCR (qPCR). The curves show the logarithm of the content change of plasmid *tcpC* or *hmoA* genes compared with 16S rRNA gene over time. Data are the average of two independent experiments, and error bars indicate 1 standard deviation. After culture, dilution plating was performed and 32 single colonies of each culture were selected to verify plasmid loss with the same primers. The number of colonies with positive PCR result for plasmid *tcpC* or *hmoA* genes is shown on the right side of the corresponding curve.

Fig. S9

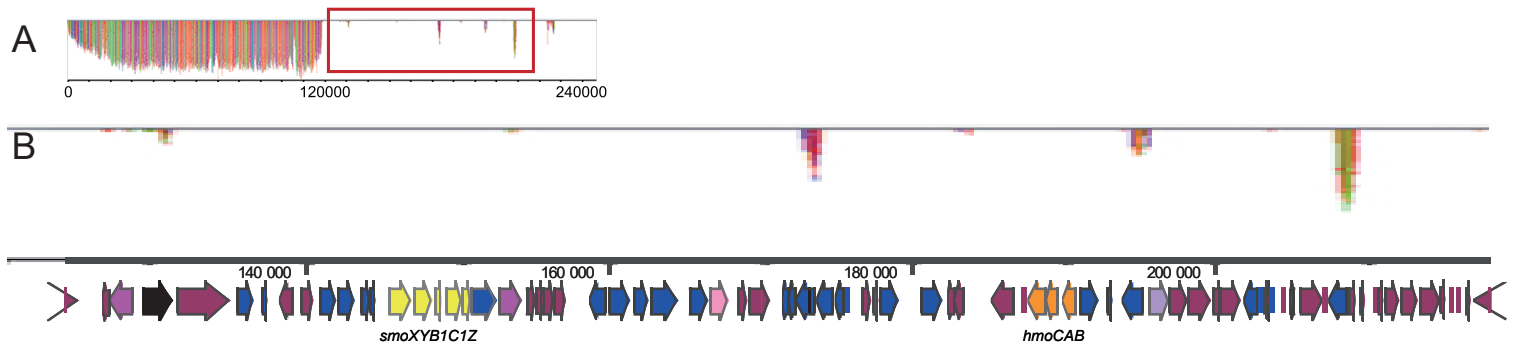


Fig. S9. Burrows-Wheeler Aligner (BWA) mapping result of genomic sequencing of ZPP on plasmid pRZP1  $\Delta(smoXYB1C1Z-hmoCAB)$  (A) and its detailed mapping in *smoXYB1C1Z-hmoCAB* region (B). The bar chart shows the coverage of each position. In (B), the SDIMO gene (*smoXYB1C1Z*) clusters are marked in yellow, and the CuMMO gene (*hmoCAB*, *amoCAB* and *pmoCAB*) clusters are marked in orange. The alcohol dehydrogenase genes (*adh*) are marked in light pink; the NDMA-dependent methanol dehydrogenase genes (*mdh*) are marked in purple; the aldehyde dehydrogenase genes (*aldH*) are marked in pink; the acetate-CoA ligase gene (*acs*) is marked in black and the mobile elements (including *IS* elements, transposases, and phages) are marked in blue.

## References

1. Darling AC, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 14:1394-403.
2. McLeod MP, Warren RL, Hsiao WWL, Araki N, Myhre M, Fernandes C, Miyazawa D, Wong W, Lillquist AL, Wang D, Dosanjh M, Hara H, Petrescu A, Morin RD, Yang G, Stott JM, Schein JE, Shin H, Smailus D, Siddiqui AS, Marra MA, Jones SJM, Holt R, Brinkman FSL, Miyauchi K, Fukuda M, Davies JE, Mohn WW, Eltis LD. 2006. The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. *Proc Natl Acad Sci USA* 103:15582.
3. Holmes AJ. 2009. The diversity of soluble di-iron monooxygenases with bioremediation applications, p 91-102. *In* Singh A, Kuhad RC, Ward OP (ed), *Advances in Applied Bioremediation* doi:10.1007/978-3-540-89621-0\_5. Springer Berlin Heidelberg, Berlin, Heidelberg.