Supporting Information

Eustáquio et al. 10.1073/pnas.0901237106

SI Text

Bacterial Strains, Plasmids, Culture Conditions, and DNA Manipulations. *Salinispora tropica* CNB-476 and CNB-440 (1) were cultured in flasks containing a stainless steel spring or glass beads and A1 sea water-based medium as described in ref. 2. *Escherichia coli* DH5α (Invitrogen) was used for cloning experiments and grown as described in ref. 3. The use of either *E. coli* ET12567/pUZ8002 (4) (DNA methylation-deficient) or *E. coli* S17–1 (5) (DNA methylation-proficient) as donor strains in conjugation experiments yielded comparable numbers of exconjugants, indicating that *S. tropica* has no restriction system for methylated DNA in contrast to the well-studied actinomycete *Streptomyces coelicolor* A3(2) (6, 7).

The REDIRECT technology kit for PCR targeting (8) was obtained from Plant Bioscience Limited. Plasmid pIJ790 containing the λ Red functions is not compatible with pCC1FOS-based fosmids (chloramphenicol resistance gene in both), therefore pKD20 (*bla*) (9) was used instead. pCC1FOS-based (Epicentre) fosmids BHXS3930, BHXS2728, BHXS5918, and BHXS1782 were obtained from the Joint Genome Institute. Cosmid pLBcos23 contains the *sal* gene cluster (spanning from the end of *salU* to ORF2) in pOJ446 (10) and was generated in this study. Plasmid pAEM3 used for insertional inactivation of *salA* was constructed by cloning a 2,622-bp *Pst*I fragment from pLBcos23 into the same site of pKC1132 (10). Apramycin (200 μg/mL for *S. tropica*; 50 μg/mL for *E. coli*), chloramphenicol (5 μg/mL for *S. tropica*; 12–25 μg/mL for *E. coli*), carbenicillin (100 μg/mL), and nalidixic acid (100 μg/mL) were used for selection of recombinant strains. DNA isolation, manipulation, and Southern blot analysis were performed according to standard procedures (3, 6).

Cloning, Sequencing, and Analysis of the *sal* Gene Cluster. The pOJ446 library of *S. tropica* CNB-476 was constructed by using *Mbo*I instead of *Sau*3AI for partial digestion of genomic DNA. Other steps were as described in ref. 11. DNA analysis and assembly was performed with Sequencher. FramePlot (12, 13) was used for sequence analysis and for BLAST homology searches in the National Center for Biotechnology Information database. The *sal* cluster from strain CNB-440 was obtained during genome sequencing (1).

Phylogenetic Analysis. Gene sequences obtained from GenBank (see Table S2) were aligned based on their protein sequence by ClustalW using MEGA 4.0.2 (14). Nexus files were generated with Mesquite (15). Bayesian analysis of DNA sequences was conducted by using MrBayes (16) with the most appropriate model of nucleotide substitution for the sequences estimated with MrModelTest (17). The analysis was run for 2,500,000 generations to achieve standard deviation of split frequencies of 0.0003 ($<$0.01). The Maximum Parsimony tree of aligned protein sequences was obtained from heuristic searches (10,000 replicates, sequences added randomly) using PAUP* (18). Bootstrap values were obtained from 100 replicates. Trees were visualized with FigTree (http://tree.bio.ed.ac.uk/software/figtree/).

Conjugation Protocol for *Salinispora tropica*. One mL mycelium from a 2- to 3-day-old second generation A1 liquid culture (late exponential phase) was resuspended in 1 mL A1 medium. One-half mL of the *E. coli* suspension treated as described in ref. 6 was then mixed with 0.5 mL of the *S. tropica* suspension, and the mixture was spread onto 4 A1 agar plates. After 20-h incubation at 32–34 °C, the plates were overlaid with 1 mg of nalidixic acid and 4 mg of apramycin. Exconjugants were visible after 1 week.

Inactivation of *sal* Genes. All genes besides *salA* (see *Methods*) were inactivated by PCR targeting (8) with some modifications as described in ref. 19. Fosmids targeted in each case were BHXS3930 (*salG*, *salH*, *salM*, *salN*), BHXS2728 (*salQ*, *salS*, *salT*), and BHXS5918 (*Strop_3612*). Primer pairs used to amplify the apramycin resistance [*aac (3)IV*] cassette from pIJ773 (8) by PCR are described below. *salG*, P1-salG (5′-CAG GAG GAG TGC GCC ATG GGC CAG CTG ACC GAC GCC GTG ATT CCG GGG ATC CGT CGA CC-3′) and P2-salG (5′-CGC GCT GGG GCG CGC TCG ATG CCG ACG AAG AGT GGG TTA TGT AGG CTG GAG CTG CTT C-3′). *salH*, salHf (5′-TTC GTC GGC ATC GAG CGC GCC CCA GCG CGC GCC TCC ATG ATT CCG GGG ATC CGT CGA CC-3′) and salHr (5′-CTC CTC GGC CTT ACC TGA CGC TGT CGC CGG GCC CTG GGT TGT AGG CTG GAG CTG CTT C-3′). *salM*, salMf (5′-TGG TGA AGG GCT GCC GAA CAG GTC AGG AGG TGG TTG ATG ATT CCG GGG ATC CGT CGA CC-3′) and salMr (5′-GTC GGA CCA TCT ATG CGA TGG TCG ACG GCG CTC GTG TCA TGT AGG CTG GAG CTG CTT C-3′). *salN*, salNf (5′-CGG CGA GAA AGC CGC CCC CCA GGG AGG TAC ACG ACC GTG ATT CCG GGG ATC CGT CGA CC-3′) and salNr (5′-GAT TGG CCG TGT ACG GCG CTG CCG TCC AAC ATA CGA TCA TGT AGG CTG GAG CTG CTT C-3′). *salQ*, salQf (5′-ATG ATC GGT TCA GAC GAA TTG CCG AGA AGG AGT CGC ATG ATT CCG GGG ATC CGT CGA CC-3′) and salQr (5′-GAA GGA CCA GCC CAG TAA CGA GAG AGT CGG GCT ACC CTA TGT AGG CTG GAG CTG CTT C-3′). *salS*, salSf (5′-GCG CCG ACC GCT ACC CAC TAT CAA CGG AGG AAG TGA ATG ATT CCG GGG ATC CGT CGA CC-3′) and salSr (5′-GTG GCT GCC CGC GAT CGG TCT GCT GGT GGC GAG CGA TCA TGT AGG CTG GAG CTG CTT C-3′). *salT*, salTf (5′-CCC AAG GCA CAT CGG CAT CCG CAG GGG GAA AGA ATG ATT CCG GGG ATC CGT CGA CC-3′) and salTr (5′-GCC ATA ACG GCA CCT CGT CAA CTG GGG AGC AGC CGC TCA TGT AGG CTG GAG CTG CTT C-3′). *Strop_3612*, CCRfnew (5′-CGG CCA CGC GCG GGG GCG GAC GAA CGG GAG GTC GAA GTG ATT CCG GGG ATC CGT CGA CC-3′) and CCRr (5′-CAG CCG CCA CAT TGG TGG CGC CCC TCG CCC GTG TGG TCA TGT AGG CTG GAG CTG CTT C-3′). Underlined letters represent 39-nt homologous extensions to the DNA regions immediately upstream and downstream of the targeted genes, respectively, including the putative start and stop codons. The mutated fosmids were introduced into *S. tropica* CNB-440 by conjugation from *E. coli* S17–1. Gene replacement was confirmed by Southern blot analysis and/or PCR.

RT-PCR Analysis. RNA was isolated from cells collected from a second generation *S. tropica* A1 liquid culture grown for 72 h (late exponential phase) and by using the RiboPure-Bacteria kit (Ambion). Cell lysis was obtained by homogenization using the FastPrep system (BIO 101 Inc.) (5 cycles of 30 s at 5.5 speed). DNase I treatment was carried out for 5 h following the manufacturer's instructions. After confirming the absence of genomic DNA in the sample by PCR using 16S rRNA primers, cDNA synthesis was carried out by using the SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen). The

PCR step was run for 25 cycles using Taq polymerase (New England Biolabs) and 250 ng of cDNA in 25-μL reactions. Primers used were 16S rRNA (forward GAC GGT ACC TGC AGA AGA, reverse GTC TCC CCT ACC GAA CTC TA), salH (GAT CGA GAT CAA CAC CAT CG and ATC GTG CCG GAA TAG ATG AC), salL (TCG TGG ACA TAA CCC ATG AC and AGG ACC TCG TGA CAC TCG AC), salO (AGG TGG GAT GAA GTG CAG AG and CGT ACA CGG CCA ATC GG) and salS (CTC GCT GAA GGG ATG TGT G and TGT TGA GCC CGT AGT TCA CC), giving a PCR product of ≈250 bp.

***salG* Inactivation and Promoter Insertion Before *salH*.** The cassette for promoter insertion before *salH* was generated by PCR as shown in Fig. S4. The apramycin resistance cassette from pIJ773 (8) was fused to the putative *salA* promoter region (P_{salA}, i.e., *salA* and *salK* intergenic region) by a 2-step PCR in which cassette and promoter sequences were amplified in 2 parallel reactions, mixed to equimolar concentrations and subjected to a third PCR using outer primers with extensions (H1, H2 in Fig. S4) necessary for *salG* gene replacement. Primer sequences were: P1 (ATT CCG GGG ATC CGT CGA CC), P2_salAp (GAC GAC GAC CGG CAG GAG CTT GAC ATG TCT C*TG TAG GCT GGA GCT GCT TC*, priming sequence underlined), P3_salAp (GAG ACA TGT CAA GCT CCT GC), P4_salAp (GTG AAT CCC TCA TCG GTG AG), P5 (P1salG, see above), and P6_salAp (CTC GAT GCC GAC GAA GAG TGG GTT AGG CCG GGT CGT CAC *GTG AAT CCC TCA TCG GTG AG*, priming site underlined). Conditions for PCR were as follows. The "PCR 1" reaction was carried out by using the High Fidelity PCR system from Roche as described in ref . 8, with the exception that the cassette from pIJ773 was not gel purified, but rather the PCR was digested with *Dpn*I after amplification. P_{salA} amplification (PCR 2) was achieved by using Pfu Turbo DNA polymerase (Stratagene) and fosmid BHXS1782 as template with denaturation (94 °C), annealing (58 °C) and extension (72 °C) steps of 45 s each in a total of 30 cycles. PCR products were purified with the QIAquick PCR purification kit (Qiagen) and their concentration estimated by gel electrophoresis. "PCR 3" used the High Fidelity PCR system from Roche with 1 cycle of 3-min denaturation at 94 °C and 5-min annealing/elongation at 72 °C, followed by the conditions described in ref. 8, with the exception that extension steps were of 2 min. The correct gene replacement was confirmed by PCR.

Purification of Recombinant SalG. *salG* was amplified by PCR from *S. tropica* genomic DNA using primers 5′-CGTGGTTC*CCATGG*CATGGGCCAGCTGA-3′ (*Nco*I site underlined) and 5′-GCTCGAATTC*AAGCTT*TTAGGCCGGGTC-3′ (*Hin*dIII site underlined) designed for ligation into pHIS8 (20), yielding plasmid pSG1, which was used to transform *E. coli* BL21(DE3). The resulting transformant was inoculated onto 1 L of LB medium supplemented with 50 μg/mL kanamycin, grown at 37 °C to an optical density (OD_{600}) of 0.7, induced with 0.1 mM IPTG, and grown at room temperature for further 16–18-h. SalG was purified by Ni-affinity chromatography (according to standard protocols) and FPLC gel-filtration using a Superdex G200 column (GE). The purified protein fractions were stored in the FPLC buffer (50 mM phosphate, 100 mM NaCl, pH 7.2) at −80 °C after the addition of 10% glycerol.

SalG Activity Assay. Determination of apparent kinetic constants was performed by continuous assay measuring NADPH consumption at 340 nm (see *Methods*). Data were collected in triplicate and nonlinear regression with GraFit 4.012 was used to determine k_{cat} and K_M values. Confirmation of the butyryl-CoA and ethylmalonyl-CoA products from 5 μL SalG catalyzed reactions with crotonyl-CoA, and the corresponding chlorinated products from 4-chlorocrotonyl-CoA was carried out by LC-MS analysis. LC conditions comprised of a water to methanol gradient containing 5 mM ammonium acetate with flow rate of 0.3 mL/min on an Agilent 1200 system using a 2.1 μm Discovery HS C18 reverse phase column (Supelco). Mass spectra were collected on a micrOTOF Q (Bruker Daltonics) mass spectrometer equipped with an electrospray ion source operating in positive mode. ESI-MS *m/z* 870.1 (4-chlorocrotonyl-CoA), *m/z* 872.1 (4-butyryl-CoA), *m/z* 915.8 (chloroethylmalonyl-CoA).

Chemical Synthesis of Compounds Used in Feeding Experiments. All chlorinated substrates were synthesized according to known or slightly modified literature procedures and are summarized in Fig. S5. 5-ClR was obtained in 5 steps from D-ribose by protection, activation of the primary alcohol through tosylation, and subsequent chlorination and deprotection (21–23). The final chemical reaction involving deprotection of 5-chloro-5-deoxy-2,3-O-isopropylidene-β-D-ribose to yield 5-ClR is detailed below. Ribonic lactone was chlorinated using thionyl chloride to obtain 5-ClRL according to the literature (24). Base treatment of 5-ClRL yielded 5-ClRI (25) as detailed below. The hydride reduction of ethyl 4-chloro acetoacetate afforded known ethyl 4-chloro-3-hydroxylbutyrate (26). Chlorocrotonic acid (ClCA) was synthesized by the selective selenocatalytic allylic halogenation of β,γ–unsaturated butyric acid according to the literature (27).

5-Chloro-5-deoxy-D-ribofuranoside (5-ClR). To methyl 5-chloro-5-deoxy-2,3-O-isopropylidene-D-ribofuranoside (250 mg, 1.12 mmol) was added 0.04 N HCl (5 mL), and the emulsion that formed on stirring was heated at reflux for 20 min. On cooling to room temperature, Dowex Monosphere ion exchange resin (550A-OH^- form) was added to the mixture until a pH of 7 was reached. The resin was removed by filtration and washed with water (30 mL). After evaporation of the solvents, the residue was purified by column chromatography to give 5-ClR (72 mg, 38%) as a transparent syrup. At equilibrium in D_2O the product was an approximate 2:1 mixture of α- and β-anomers. α-anomer ^1H NMR (400 MHz, D_2O) δ 3.63–3.79 (m, 2H), 3.96 (dd, J = 4.8, 1.6 Hz, 1H), 4.06–4.10 (m, 1H), 4.21–4.24 (m, 1H), 5.19 (d, J = 1.6 Hz, 1H); ^{13}C NMR (100 MHz, D_2O, reference acetone-d_6 δ 29.92) δ 45.7 (CH_2), 71.9 (CH), 75.6 (CH), 81.7 (CH), 101.5 (CH). β-anomer ^1H NMR (400 MHz, D_2O and $(CD_3)_2CO$) δ 3.63–3.79 (m, 2H), 4.06–4.10 (m, 2H), 4.21–4.24 (m, 1H), 5.34 (d, J = 4.4 Hz, 1H); ^{13}C NMR (100 MHz, D_2O, reference acetone-d_6 δ 29.92) δ 45.0 (CH_2), 71.1 (CH), 71.4 (CH), 82.0(CH), 97.0 (CH).

5-Chlororibonic Acid (5-ClRI). To 5-chloro-5-deoxy-D-ribono-1,4-lactone (25 mg, 0.15 mmol) was added 0.45 N K_2CO_3 (1 mL). After stirring at room temperature for 5 min, the reaction mixture was washed with ethyl acetate (3 × 5 mL), acidified to pH 2–3 with 1.0 M HCl and extracted with EtOAc (8 × 10 mL). The combined organic layer was dried with anhydrous Na_2SO_4. The solvent was evaporated, and the residue was dried in the vacuum to obtain 5-ClRI (6 mg, 21%) as white solid. ^1H NMR (400 MHz, D_2O) δ 3.71 (dd, J = 12, 4.8 Hz, 1H), 3.79 (dd, J = 12, 2.4 Hz, 1H), 3.97–4.05 (m, 2H), 4.42 (d, J = 2.0 Hz, 1H); ^{13}C NMR (100 MHz, D_2O, reference acetone-d_6 δ 29.92) δ 48.0 (CH_2), 69.6 (CH), 72.2 (CH), 73.1(CH), 175.6 (C = O); HRMS calcd for [$C_5H_9ClO_5$ –H]: 183.0055, found: 183.0060.

4-Chlorocrotonyl-CoA. To a solution of 4-chlorocrotonic acid (ClCA, 22 mg, 0.186 mmol, 6 equiv.) in anhydrous THF (8 mL), cooled to 0 °C and under nitrogen, was added triethylamine (26 μL, 0.186 mmol, 6 equiv.) followed by ethyl chloroformate (18 μL, 0.188 mmol, 6 equiv), and stirred at 0 °C for 45 min. The stirring was then stopped to allow the resulting solid to precip-

itate. The clear THF supernatant was added slowly to a solution of hydrated CoA (24 mg, 0.031 mmol, 1 equiv.) and sodium bicarbonate (15 mg, 0.18 mmol) in dd H_2O (3 mL). The reaction mixture was stirred at room temperature for 4 h after which the THF was removed under vacuo. The aqueous solution was acidified to pH 3 using 1 N HCl and then extracted with EtOAc (3 × 3 mL) to remove excess ClCA. The aqueous solution was then lyophilized and the resulting solid was washed with MeOH to recover 4-chlorocrotonyl-CoA as a white solid (18 mg, 65% yield). ^{1}H NMR (500 MHz, CD_3OD) δ 8.67 (s, 1H), 8.41 (s, 1H), 6.93–6.86 (m, 1H), 6.40 (dd, J = 15.5 Hz, 1.5 Hz, 1H), 6.14 (d, J = 6.5 Hz, 1H), 4.74 (s, 1H), 4.5 (s, 1H), 4.31 (m, 2H), 4.24 (d, J = 6.0 Hz, 2H), 4.05 (m, 3H), 3.67 (m, 3H), 3.46 (m, 4H), 3.06 (t, J = 8.6 Hz, 2H), 2.42 (t, J = 8.6 Hz, 2H), 1.01 (s, 3H), 0.89 (s, 3H); MS for $C_{25}H_{39}ClN_7O_{17}P_3S$ (ESI+) m/z 870.1 (M + H).

1. Udwary DW, et al. (2007) Genome sequencing reveals complex secondary metabolome in the marine actinomycete *Salinispora tropica*. *Proc Natl Acad Sci USA* 104:10376–10381.
2. Beer LL, Moore BS (2007) Biosynthetic convergence of salinosporamides A and B in the marine actinomycete *Salinispora tropica*. *Org Lett* 9:845–848.
3. Sambrook J, Russell DW (2001) *Molecular Cloning*: *A Laboratory Manual* (Cold Spring Harbor Lab Press, New York).
4. Paget MS, Chamberlin L, Atrih A, Foster SJ, Buttner MJ (1999) Evidence that the extracytoplasmic function sigma factor sigmaE is required for normal cell wall structure in *Streptomyces coelicolor* A3(2). *J Bacteriol* 181:204–211.
5. Simon R, Priefer U, Pühler A (1983) A broad host range mobilization system for in vivo genetic-engineering—transposon mutagenesis in gram-negative bacteria. *Bio-Technology* 1:784–791.
6. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000) *Practical Streptomyces Genetics* (John Innes Foundation, Norwich, UK).
7. Flett F, Mersinias V, Smith CP (1997) High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting streptomycetes. *FEMS Microbiol Lett* 155:223–229.
8. Gust B, Challis GL, Fowler K, Kieser T, Chater KF (2003) PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc Natl Acad Sci USA* 100:1541–1546.
9. Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* 97:6640–6645.
10. Bierman M, et al. (1992) Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene* 116:43–49.
11. Piel J, et al. (2000) Cloning, sequencing and analysis of the enterocin biosynthesis gene cluster from the marine isolate *Streptomyces maritimus*: Evidence for the derailment of an aromatic polyketide synthase. *Chem Biol* 7:943–955.
12. Ishikawa J, Hotta K (1999) FramePlot: A new implementation of the frame analysis for predicting protein-coding regions in bacterial DNA with a high G + C content. *FEMS Microbiol Lett* 174:251–253.
13. Bibb MJ, Findlay PR, Johnson MW (1984) The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. *Gene* 30:157–166.
14. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599.
15. Maddison WP, Maddison DR (2008) Mesquite: A Modular System for Evolutionary Analysis (http://mesquiteproject.org).
16. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
17. Nylander JAA (2004) Mr. Modeltest V2. (program distributed by the author, Evolutionary Biology Centre, Upsalla University).
18. Swofford DL (2003) *PAUP**, Phylogenetic Analysis Using Parsimony (*and Other Methods) (Sinauer Associates, Sunderland, Massachusetts), Version 4.
19. Eustáquio AS, Pojer F, Noel JP, Moore BS (2008) Discovery and characterization of a marine bacterial SAM-dependent chlorinase. *Nat Chem Biol* 4:69–74.
20. Jez JM, Ferrer JL, Bowman ME, Dixon RA, Noel JP (2000) Dissection of malonyl-coenzyme A decarboxylation from polyketide formation in the reaction mechanism of a plant polyketide synthase. *Biochemistry* 39:890–902.
21. Moon BS, et al. (2005) Synthesis of *O*-(3-[F-18]fluoropropyl)-L-tyrosine (L-[F-18]FPT) and its biological evaluation in 9L tumor bearing rat. *Bull Korean Chem Soc* 26:91–96.
22. Vanes T, Rabelo JJ (1973) Derivatives of 5-deoxy-5-seleno-D-xylose. *Carbohydr Res* 29:252–254.
23. Gudmundsson KS, Drach JC, Wotring LL, Townsend LB (1997) Synthesis and antiviral activity of certain 5′-modified analogs of 2,5,6-trichloro-1-(beta-D-ribofuranosyl)benzimidazole. *J Med Chem* 40:785–793.
24. Bouchez V, Stasik I, Beaupere D, Uzan R (1997) Regioselective halogenation of pentono-1,4-lactones. Efficient synthesis of 5-chloro- and 5-bromo-5-deoxy derivatives. *Carbohydr Res* 300:139–142.
25. Kold H, Lundt I, Pedersen C (1994) Synthesis of L-ribono-lactone and L-lyxono-lactone. *Acta Chem Scand* 48:675–678.
26. Hoff BH, Anthonsen T (1999) Lipase-catalyzed resolution of esters of 4-chloro-3-hydroxybutanoic acid: Effects of the alkoxy group and solvent on the enantiomeric ratio. *Tetrahedron-Asymmetry* 10:1401–1412.
27. Tunge JA, Mellegaard SR (2004) Selective selenocatalytic allylic chlorination. *Org Lett* 6:1205–1207.
28. Hall BG (2005) Comparison of the accuracies of several phylogenetic methods using protein and DNA sequences. *Mol Biol Evol* 22:792–802.
29. Reeves CD, et al. (2001) Alteration of the substrate specificity of a modular polyketide synthase acyltransferase domain through site-specific mutations. *Biochemistry* 40:15464–15470.

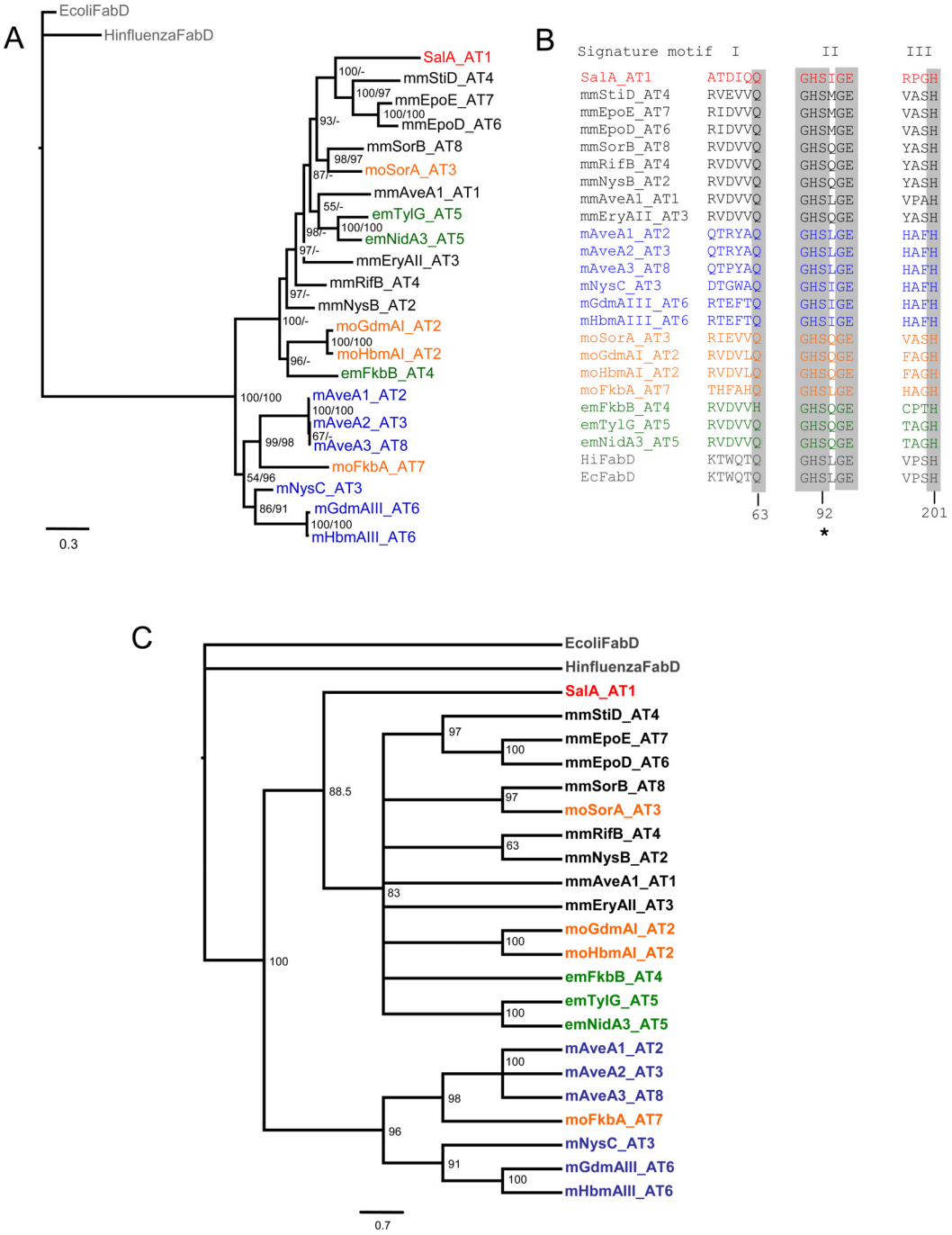

Fig. S1. Phylogenetic analysis of SalA_AT$_1$ domain using Bayesian inference and Maximum Parsimony (28). (*A*) Bayesian tree of codon aligned DNA sequences. The scale bar indicates changes per nucleotide. Values on nodes are estimates of the reliability of a particular grouping, that is, posterior probability obtained from the Bayesian inference and bootstrap scores from the Maximum Parsimony analysis of protein sequences, respectively (in %); a dash indicates that the respective clade was either not present in the Parsimony tree or its bootstrap value was <50%. Methylmalonyl- (mm, black), malonyl- (m, blue), methoxymalonyl- (mo, orange) and ethylmalonyl-specific (em, green) domains are color coded. The malonyl-CoA:ACP transacylase component of fatty acid synthase, FabD, from *E. coli* and *Haemophilus influenzae* (gray) was used as outgroup to root the tree. (*B*) Sequence alignment showing divergence of signature motifs in PKS AT domains with different substrate specificity (29). Numbers below the alignment correspond to *E. coli* FadD, the catalytic serine is indicated by an asterisk. (*C*) Maximum Parsimony tree of aligned protein sequences showing topology only. Strict consensus tree with bootstrap scores (in %) out of 1,000 replicates. Nodes were collapsed if the bootstrap value was <50%. According to Hall (28), Bayesian estimation of trees from codon aligned DNA sequences seems to be the most accurate method for inferring phylogeny. Yet, it appears reasonable to use both Bayesian inference of DNA sequences and Maximum Parsimony analysis of protein sequences, with any topological differences representing real uncertainty (28). See *SI Text* and Table S2 for details about sequences used.

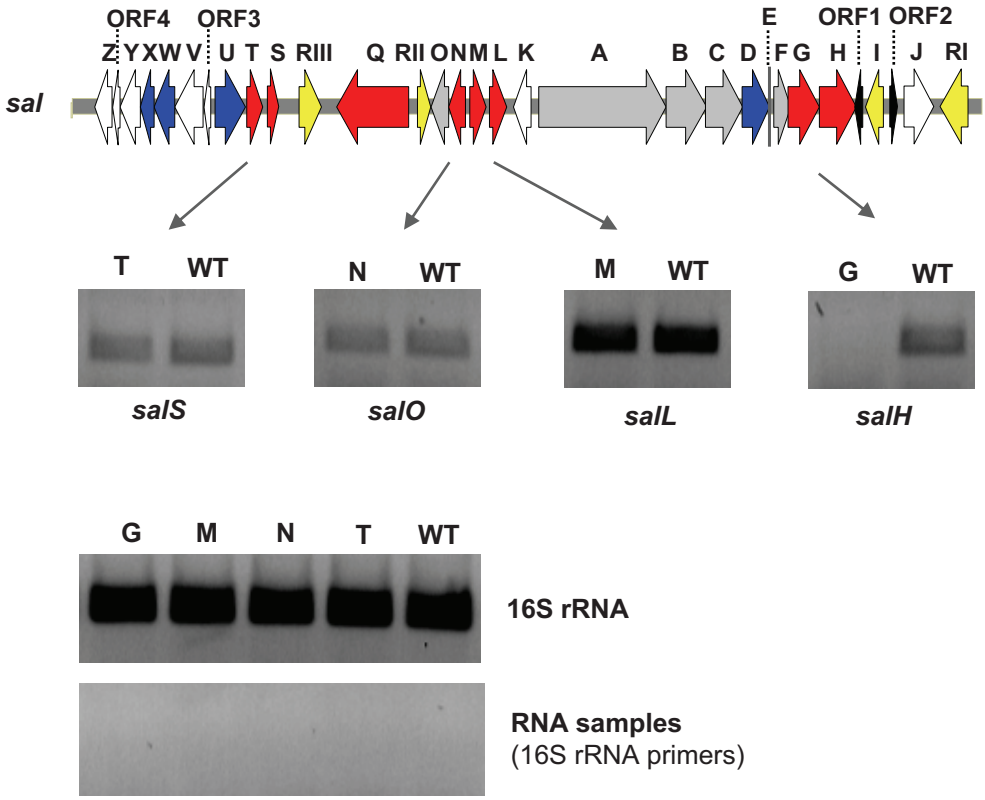

Fig. S2. Does mutation of *sal* genes cause polar effects on downstream genes? Semiquantitative reverse transcription-PCR (RT-PCR) analysis of *salT* (T), *salN* (N), *salM* (M), and *salG* (G) defective mutants compared with the wild type (WT). Transcripts analyzed were *salS*, *salO*, *salL*, and *salH*, respectively. 16S rRNA was used as control of cDNA quality and integrity. The negative control with RNA instead of cDNA samples confirms the absence of genomic DNA contamination. Samples were collected from 3-day-old second generation A1 liquid cultures.

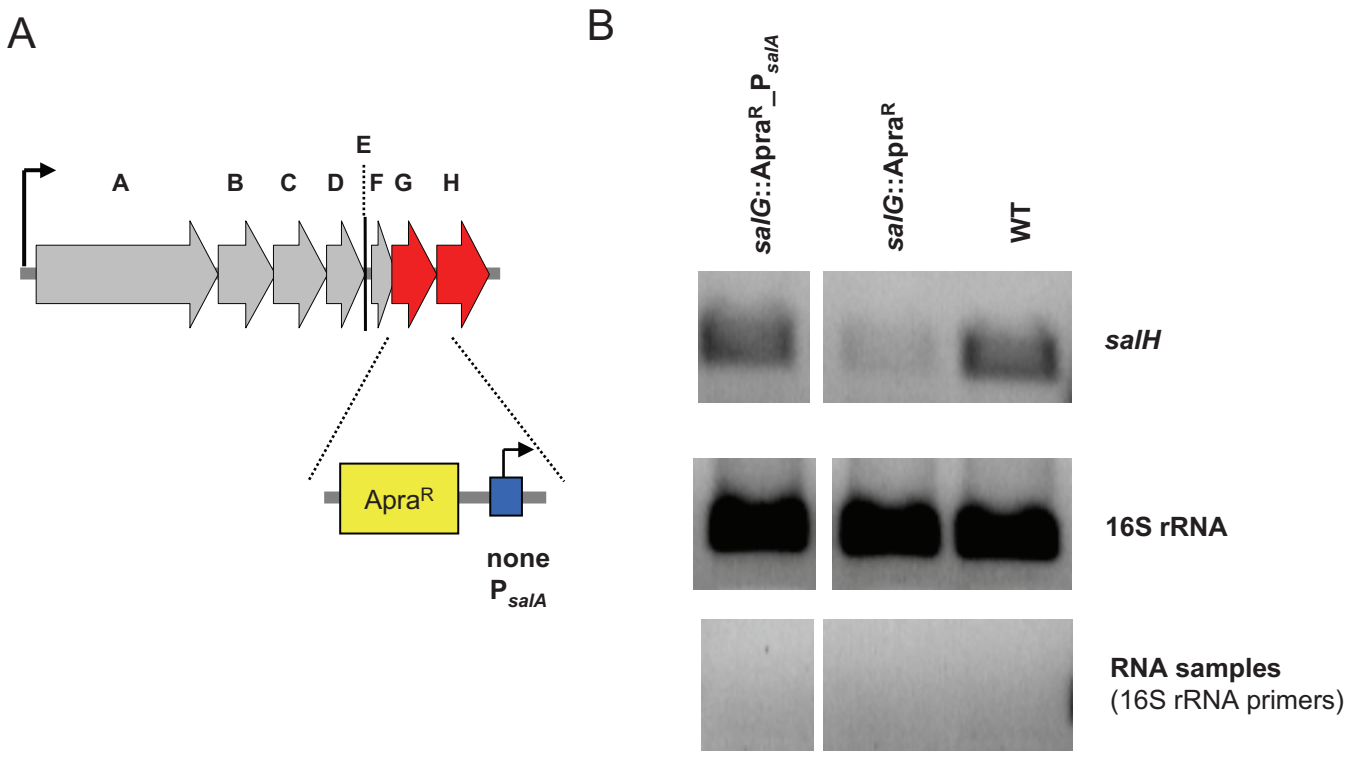

Fig. S3. Construction of an additional *salG* mutant in which the putative *salA* promoter region (P_{salA}) was inserted before *salH* to normalize its transcription levels. (*A*) Schematic representation of the mutation. (*B*) Semiquantitative RT-PCR analysis of the original *salG* mutant (*salG*::Apra^R) and the mutant containing P_{salA} inserted before *salH*.

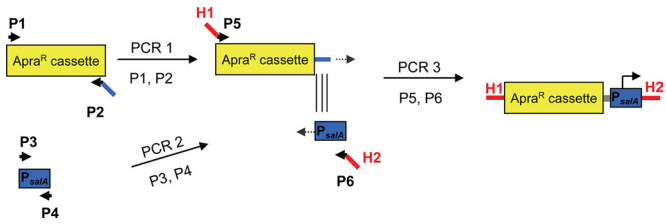

Fig. S4. Schematic representation of the PCR strategy used to fuse the apramycin resistance cassette from pIJ773 to the *salA* promoter region (P_{salA}), generating a new cassette for *salG* inactivation that avoids polar effects on *salH*.

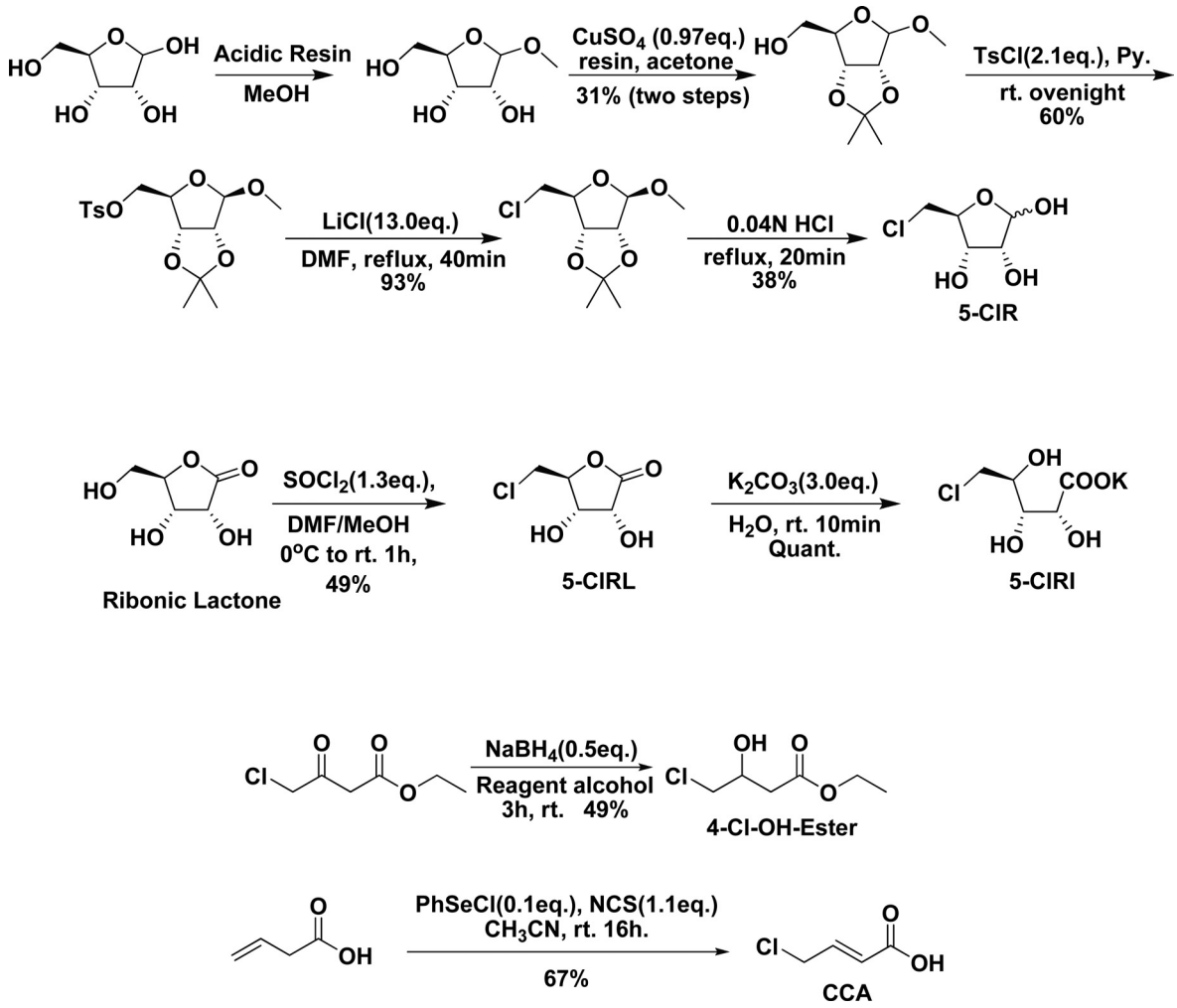

Fig. S5. Summary of synthetic schemes to produce the chlorinated compounds used in this study.

Table S1. Deduced functions of the proteins encoded by *sal* genes

Strop_	Protein	Residues	Proposed function	Sequence identity to (%)	Accession number
1024	SalA	2,053	ACP_L-KS-AT_L-AT_1-ACP_1-C (PKS)	Type I PKS StiA of *Stigmatella aurantiaca* (43%)	CAD19085
1023	SalB	632	A-PCP (NRPS)	NRPS of *Myxococcus xanthus* (43%)	ABF87402
1022	SalC	597	Ketosynthase	Beta-ketoacyl synthase domain of PKS from *Mycobacterium* sp. (39%)	ZP_01278815
1021	SalD	417	Cytochrome P450 hydroxylase	Tiol of *Micromonospora* sp. (51%)	CAJ34365
1020	SalE	71	MbtH-like	CDA-ORFX of *Streptomyces coelicolor* (47%)	AAD18046
1019	SalF	286	Thioesterase	Hydrolase of *Streptomyces coelicolor* (51%)	NP_627445
1018	SalG	460	4-Chlorocrotonyl-CoA reductase/ carboxylase	Crotonyl-CoA reductase of *Streptomyces hygroscopicus* (69%)	AAR32675
1017	SalH	592	Dihydroxyacid dehydratase	Dihydroxyacid dehydratase of *Rubrobacter xylanophilus* (64%)	YP_644102
1016	ORF1	128	Partial transposase	Transposase of *Streptomyces avermitilis* (61%, partial)	NP_821430
1015	SalI	278	Proteasome β-subunit	Proteasome β-type subunit 2 of *Rhodococcus erythropolis* (59%)	AAC45736
1014	ORF2	110	Partial transposase	Transposase of *Corynebacterium glutamicum* (67%, partial)	YP_226786
1013	SalJ	451	Conserved hypothetical protein	Conserved hypothetical protein of *Frankia* sp. (49%)	ABD11397
1012	SalRI	365	Regulatory protein, MerR-family	MerR-family transcriptional regulator of *Streptomyces coelicolor* (76%)	NP_631737
1025	SalK	281	Conserved hypothetical protein	Conserved hypothetical protein of *Frankia alni* (44%)	YP_710332
1026	SalL	283	5′-Chloro-5′-deoxyadenosine synthase	5′-Fluoro-5′-deoxyadenosine synthase of *Streptomyces cattleya* (35%)	CAE46446
1027	SalM	255	Dehydrogenase/Reductase	Short-chain dehydrogenase/reductase of *Mycobacterium* sp. (66%)	YP_638874
1028	SalN	258	Phosphatase	Conserved hypothetical protein of *Geobacillus thermodenitrificans* (35%)	YP_001125271
1029	SalO	277	Cyclase	Putative cyclase of *Silicibacter pomeroyi* (34%)	AAV94733
1030	SalRII	225	Regulatory protein, LuxR-family	Putative DNA binding regulatory protein, LuxR-type, of *Streptomyces cattleya* (37%)	CAJ20008
1031	SalQ	1,149	α-Ketoacid decarboxylase	Putative pyruvate decarboxylase of *Nocardioides* sp. (54%)	ZP_00656820
1032	SalRIII	340	Regulatory protein, LysR-family	Transcriptional regulator, LysR-family, of *Pseudomonas putida* (30%)	NP_744719
1034	SalS	151	Acyl dehydratase	Acyl dehydratase of *Ralstonia eutropha* (49%)	YP_725579
1035	SalT	267	Purine nucleoside phosphorylase	5′-methylthioadenosine phosphorylase of *Streptomyces coelicolor* (59%)	CAB90972
1036	SalU	454	DAHP synthase	DAHP synthase of *Actinosynnema pretiosum* (58%)	AAC13561
1037	ORF3	99	Conserved hypothetical protein	Unknown orf7 of *Streptomyces laurentii* (50%)	ACN80659
1038	SalV	471	Acyl-CoA ligase	Phenylacetyl-CoA ligase of *Photorhabdus luminescens* (39%)	CAE15415
1039	SalW	320	Amino acid aminotransferase	Branched-chain amino acid aminotransferase of *Rubrobacter xylanophilus* (44%)	YP_645847
1040	SalX	217	Prephenate dehydratase	Prephenate dehydratase BacA of *Bacillus subtilis* (34%)	AAM90568
1041	SalY	329	Cyclodeaminase	Ornithine cyclodeaminase of *Silicibacter pomeroyi* (40%)	AAV97046
1042	ORF4	97	Conserved hypothetical protein DUF1330	Hypothetical protein ms17859 of *Mesorhizobium loti* (42%)	NP_108090
1043	SalZ	278	Epimerase	Predicted epimerase, PhzC/PhzF homolog of *Brevibacterium linens* (57%)	ZP_00378489

Table S2. Sequences used in Fig. S1

Protein	Organism	NCBI† accession code	AT no. (position)
AveA1	*Streptomyces avermitilis*	NP_822113	AT1 (25..357); AT2 (2633..2961)
AveA2	*Streptomyces avermitilis*	BAA84475	AT3 (591..925)
AveA3	*Streptomyces avermitilis*	BAA84478	AT8 (2419..2753)
*Ec*FabD	*Escherichia coli*	P0AAI9	—
EpoD	*Sorangium cellulosum*	AAF62883	AT6 (5626..5949)
EpoE	*Sorangium cellulosum*	AAF62884	AT7 (556..879)
EryAII	*Saccharopolyspora erythraea*	YP_001102990	AT3 (556..881)
FkbA	*Streptomyces hygroscopicus* var. *ascomyceticus*	AAF86396	AT7 (516..787)
FkbB	*Streptomyces hygroscopicus* var. *ascomyceticus*	AAF86393	AT4 (6472..6791)
GdmAI	*Streptomyces hygroscopicus*	AAO06916	AT2 (3709..4035)
GdmAIII	*Streptomyces hygroscopicus*	AAO06918	AT6 (576..907)
HbmAI	*Streptomyces hygroscopicus*	AAY28225	AT2 (3704..4030)
HbmAIII	*Streptomyces hygroscopicus*	AAY28227	AT6 (578..917)
*Hi*FabD	Haemophilus influenzae	YP_001290329	—
NidA3	*Streptomyces caelestis*	AAC46026	AT5 (2004..2328)
NysB	*Streptomyces noursei*	AAF71775	AT2 (2110..2437)
NysC	*Streptomyces noursei*	AAF71776	AT3 (539..863)
RifB	*Amycolatopsis mediterranei*	AAC01711	AT4 (557..878)
SalA	*Salinispora tropica*	YP_001157876	AT1 (1064..1386)
SorA	*Sorangium cellulosum*	AAK19883	AT3 (4703..5026)
SorB	*Sorangium cellulosum*	AAA79984	AT8 (7307..7630)
StiD	*Stigmatella aurantiaca*	CAD19088	AT4 (576..899)
TylG	*Streptomyces fradiae*	AAB66506	AT5 (2092..2416)

†NCBI, National Center for Biotechnology Information.