Supplementary Data: McHugh et al., 2022.

Rebecca E McHugh[a,b], John T Munnoch[a], Robyn E Braes[a], Iain J. W. McKean[a,c], Josephine Giard[a], Andrea Taladriz-Sender[c], Frederik Peschke[c], Glenn A Burley[c], Andrew J Roe[b], & Paul A Hoskisson[a#].

Biosynthesis of aurodox, a Type III secretion system inhibitor from *Streptomyces goldiniensis.*

[a] Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, 161 Cathedral Street, Glasgow, G4 0RE, UK

[b] Institute of Infection, Immunity and Inflammation, University of Glasgow, Glasgow, G12 8TA, UK

[c] Department of Pure and Applied Chemistry, University of Strathclyde, 252 Cathedral Street, Glasgow, G1 1BX.

 # Correspondence: paul.hoskisson@strath.ac.uk

Supplementary Tables: McHugh et al., 2022.

Table S1: Summary of biosynthetic gene clusters encoded by *S. goldiniensis* as predicted by antiSMASH.

Region	Type	Region	Most similar	% similarity
1	T2PKS	32,278 - 104,355	Spore pigment	83
3	Lanthipeptide, lassopeptide	486,431 - 514,192	Citrulassin E (ripp)	100
4	Siderophore	2,159,777 - 2,170,184	NA	0
5	Terpene	2,487,747 - 2,508,882	Lipopeptide	4
6	NRPS, tfuA related	2,543,021 - 2,606,769	Diisonitrile antibiotic SF2768	66
7	Bacteriocin	2,705,875 - 2,716,826	NA	0
8	Terpene	2,777,350 - 2,799,499	Geosmin	0
9	NRPS-like	2,818,277 - 2,858,728	S56P1	11
10	NRPS-like, siderophore	3,085,338 - 3,126,119	Paulomycin	13
11	NRPS-like	3,541,716 - 3,582,164	Octacosamicin	12
12	Terpene	3,625,827 - 3,643,980	NA	0
13	Indole	4,158,688 - 4,181,228	Terfestatin	28
14	bacteriocin,bottromycin,NRPS,transAT-PKS,T1PKS,hglE-KS,lassopeptide	4,213,370 - 4,484,508	Kirromycin	40
15	NRPS-like	4,599,349 - 4,657,342	Echosides	17
16	Terpene	4,783,720 - 4,802,511	NA	0
17	Linardin	5,046,397 - 5,067,017	Pentostatine	17
18	Siderophore	5,085,773 - 5,100,346	NA	0
19	NRPS, melanin	5,167,798 - 5,232,264	Scabichelin	100
20	Terpene	5,257,095 - 5,277,644	NA	0
21	NRPS, T1PKS	5,284,180 - 5,346,580	Herboxidiene	4
22	Terpene	5,480,336 - 5,505,711	Isorenieratene	100
23	T2PKS	5,770,343 - 5,856,353	Fluostatins	60
24	Lanthipeptide, bacteriocin	6,153,412 - 6,179,099	Informatipeptin	100
25	NRPS-like, T1PKS	6,410,852 - 6,457,973	Marineosin	45
26	Terpene	6,507,863 - 6,532,462	Hopene	100
27	Ectoine	6,852,064 - 6,861,735	Ectoine	100
28	Ladderane	6,937,157 - 6,978,500	Colabomycin	11
29	T3PKS, Terpene	7,764,486 - 7,805,661	Merochlorin	12
30	Melanin	8,189,939 - 8,200,406	Melanin	80
31	Siderophore	8,344,177 - 8,355,949	Desferrioxamine A	83
32	Ectoine	8,471,171 - 8,481,599	Ectoine	75
33	Terpene	9,170,031 - 9,191,797	NA	0
34	Bacteriocin	9,838,044 - 9,848,331	NA	0
35	Terpene	9,968,421 - 10,005,017	NA	0

Table S2: Description of putative aurodox gene function and their closest BLASTp hit.

Gene	Type	Function	Closest BLAST hit	AA similarity
aurQ	Other Gene	Hypothetical protein	MULTISPECIES *Streptomyces*	94/107 (87%)
aurHII	Other gene	Hypothetical protein	Streptomyces globosus	225/259 (86%)
aurHIII	Other Gene	Hypothetical protein	*Streptomyces sp*. SM11	242/288 (84%)
aurAVII	Core Biosynthetic Gene	Non-ribosomal peptide synthetase	*Streptomyces roseoverticillatus*	887/1023 (86%)
*aurM**	Additional Biosynthetic Gene	SAM dependant- O-Methyltransferase	*Streptomyces roseoverticillatus*	296/333 (88%)
aurCI	Core Biosynthetic Gene	S-malonyltransferase	*Streptomyces roseoverticillatus*	973/1076(90%)
aurOII	Additional Biosynthetic Gene	Cytochrome P450	*Streptomyces sp.* CB00455	344/401 (85%)
aurD	Other Gene	aspartate 1-decarboxylase	*Streptomyces roseoverticillatus*	129/137 (94%)
aurM	Additional Biosynthetic Gene	class I SAM-dependent O- methyltransferase	*Streptomyces klenkii*	301/319 (94%)
aurCII	Core Biosynthetic Gene	Acyltransferase domain-containing protein	*Streptomyces sp. CB00455*	383/440(87%)
aurX	Other Gene	Dieckmann Cyclase	*Streptomyces globosus*	231/259 (89%)
aurHIV	Other gene	Uncharacterized protein, DUF2087 family	*Streptomyces sp. SM11*	149/178 (83%)
aurN	Additional Biosynthetic Gene	crotonyl-CoA carboxylase/reductase	*Streptomyces azureus*	427/443 (96%)
aurT	Transport related gene	Major Facilitator SuperfamilyTransporter	*Streptomyces klenkii*	509/557 (91%)
aurHVI	Other Gene	Magnesium ATP-ase	*Streptomyces sp. SM11*	97/109 (88%)
aurAI	Core Biosynthetic Gene	Type 1 Polyketide Synthase	*Streptomyces roseoverticillatus*	2513/3129 (80%)
aurAII	Additional Biosynthetic Gene	SDR family NAD(P)-dependent oxidoreductase	*Streptomyces roseoverticillatus*	1922/2361 (81%)
aurAIII	Additional Biosynthetic Gene	SDR family NAD(P)-dependent oxidoreductase	*Streptomyces sp. AVP053U2*	5416/6729 (80%)
aurAIV	Core Biosynthetic Gene	non-ribosomal peptide synthetase	*Streptomyces roseoverticillatus*	1350/1551 (87%)

aurAV	Core Biosynthetic Gene	SDR family NAD(P)-dependent oxidoreductase	*Streptomyces klenkii*	3627/4506 (80%)
aurAVI	Additional Biosynthetic Gene	SDR family NAD(P)-dependent oxidoreductase	*Streptomyces klenkii*	1707/2018 (84%)
aurR	Regulatory Gene	TetR/AcrR family transcriptional regulator	*MULTISPECIES Streptomyces*	178/182 (97%)
aurOI	Additional Biosynthetic Gene	Cytochrome p450	*Streptomyces viridosporus*	390/406 (96%)
aurHVI	Other gene	Phytanoyl-coA Dioxygenase	*Streptomyces viridosporus*	267/278 (96%)

Table S3: List of Bacterial Strains used in this study.

Strain	Description	Genotype	Reference
Escherichia coli DH5α	*E. coli* K12 derivative	*huA2* (*argF-lacZ*)U169 *phoA glnV44* 80 (*lacZ*)M15 *gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	(1)
E. coli DH10β Top10	*E. coli* K12 derivative	mcrA, Δ(mrr-hsdRMS-mcrBC), ΦlacZ(del)M15, ΔlacX74, deoR, recA1, araD139, Δ(ara-leu)7697, galU, galK, rpsL(SmR), endA1, nupG	Invitrogen
E. coli BW21153	*E. coli* K12 derivative	lacI+ , rrnBT14, ΔlacZWJI6, hsdR514, ΔaraBADAH33, ΔrhaBADLD78	(2)
E. coli ET12567	*E. coli* K12 derivative	dam13::Tn9, dcm6, hsdM, hsdR, recF143, zij201::Tn10, galK2, galT22, ara14, lacY1, xylS, leuB6, thi-1, tonA31, rpsL136, hisG4, tsx78, mtli, glnV44, F	(3)
E. coli E2348/69	Wild type EPEC O127:H6	NA	(4)
Staphylococcus aureus ATCC43300	Methicillin-resistant *Staph. aureus*	Methicillin and oxacillin resistant *Staphylococcus aureus* strain used in susceptibility testing	ATCC 43300
Streptomyces goldiniensis ATCC 21386	Wild type, Aurodox-producing strain	Wild-type	ATCC 21386
Streptomyces collinus Tü 365	Wild type, Kirromycin-producing strain	Wild-type	Gift from Wolfgang Wolheben
Streptomyces coelicolor M1152	*Streptomyces coelicolor* 'superhost'	*Δact Δred Δcpk Δcda rpoB*(C1298T)	(5)
Streptomyces coelicolor M1152	pESAC-13A-Aur1	Aurodox cluster on PAC - AprR,	This study
Streptomyces coelicolor M1152	pESAC-13A-Aur1ΔAurP1	Aurodox cluster on PAC - AprR,	This study
Streptomyces coelicolor M1152	pESAC-13A	A pESAC-13A AprR,	This study
Streptomyces collinus Tü 365	Wild type, Kirromycin-producing strain + *aurM** from *S. goldiniensis*	Wild-type + pIJ6902_aurM*	This study

Table S4: Summary of Plasmids/BACs used in this study.

Plasmid name	Description	Resistance	Reference
pESAC-13A	Bacterial Artificial Chromosome, ΦC31 integrase, lacZ,	Apramycin, Carbenicillin	(6)
pESAC-13A-Aur1	Bacterial Artificial Chromosome, ΦC31 integrase, triple positive clone for aurodox cluster	Apramycin	This study
pESAC-13A-Aur1ΔAurP1	Bacterial Artificial Chromosome, ΦC31 integrase, triple positive clone for aurodox cluster, ΔaurP1	Apramycin, Hygromycin	This study
PIJ6902	ΦC31 integrase containing vector with Thiostrepton resistance and a MCS containing and upstream thiostrepton inducible promoter	Apramycin, Thiostrepton	(7)
PIJ6902_*aurM**	ΦC31 integrase, AurM* expression vector	Apramycin, Thiostrepton	This study
PIJ10790	λ-RED (gam, bet, exo), cat, araC, rep101ts	Chloramphenicol	(8)
PIJ10700	Re-direct template	Hygromycin	(8)

Table S5: List of Oligonucleotides used in this study.

Primer name	Sequence	Tm (^{O}C)
ClusterCheckA F	GAGCTACCAACTCTTTTTCC	50.9
ClusterCheckA R	CCGCGGATGACCGCGTGGACGTGA	69.2
ClusterCheckB F	CATCGGCAGCCTCTTCTCCGC	63.7
ClusterCheckB R	CGGCTCATGCCGGAAGTCTCCCGA	66.8
ClusterCheckC F	TTCCGCTCCCTGGATCGGCAT	66.7
ClusterCheckC R	AGCACTGAACGCCGGCACGGAAGA	67
MTGibson F	ATTCGAGCTCGGTACGGATGGAAATCCTCAAGCTCGCC	70.9
Methyltransferase_Gibson R	GTCAGAGAAGGGAGCGGACATCAGGAGGCACGCAGGGA	71.8
PIJ6902MT_check F	GAGATCGGCGTCGTACAT	54.3
PIJ6902MT_check R	GAGCTACCAACTCTTTTTCC	50.9
RedirectPKS1 F	TTCCATGCCGCGATTCCGGCCGGACCGTGTGTCCGCATGATTCCGGGGACCGTCGA	76.5
RedirectPKS1 R	GAACGAGCGCGCCGGGGCGCGTGAGGAAGGACCGGGTTATGTAGGCTGGAGCTGCTT	76.2
Hyg_F	GTGACACAAGAATCCCTGTTACTTCTC	62.7
Hyf_R	CTCGAACACCTCGAAGTCCAGGAA	65.9
PKSredirectcheckF	GGCCTGATCATGCGGTTGGGCGCT	69.5
PKSredirectcheckR	TGGAGGACATCCGGCGCT	66

Supplementary Figures: McHugh et al., 2022.

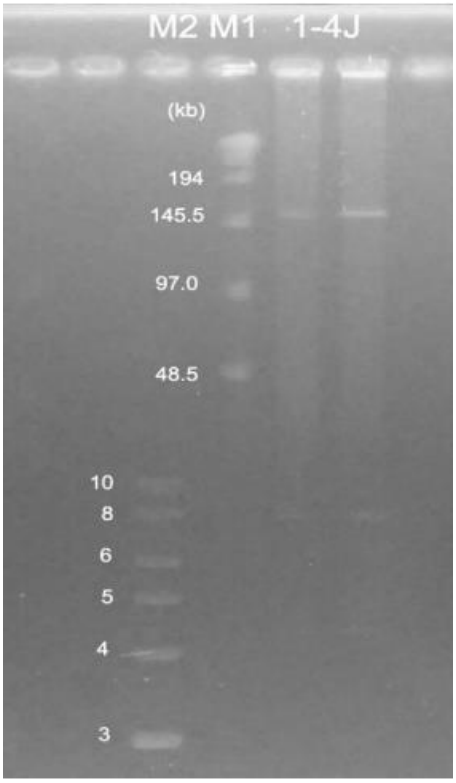

Figure S1: **Pulse field gel electrophoresis for confirmation of PAC insert size**. PAC1 (1-4J) were digested with BamHI and subjected to PFGE. Insert sizes was 160 kb. Image was supplied and PFGE was carried out by Bio S & T.

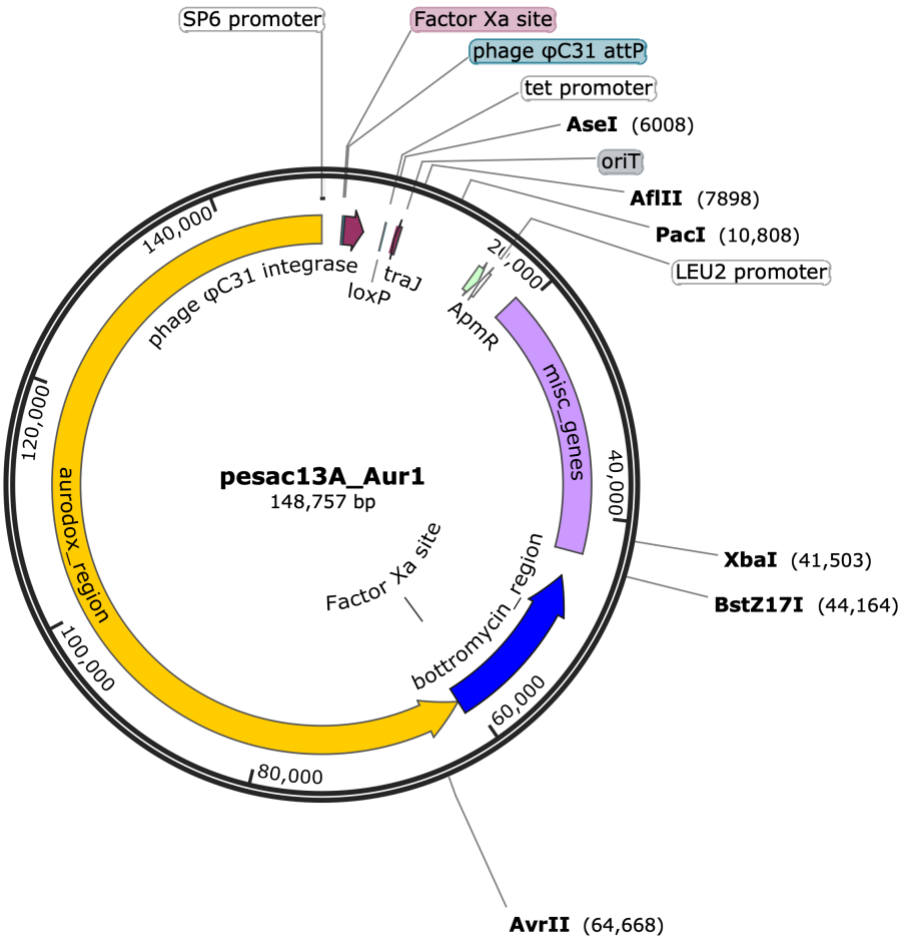

Figure S2: PAC map of pESAC-13A_Aur1 [pAur1] (9). The PAC was completely sequenced via Illumina MiSeq sequencing provided by Novogene (**Supp_File. pAur1)**. Vector encodes ϕC31 integrase, ampicillin resistance gene (replaced by gDNA at BamHI sites during PAC construction) and apramycin resistance gene. Genes encoded in the amber region represent putative aurodox genes which were also confirmed through PCR, genes in the blue genes represent predicted bottromycin genes which flank aurodox-encoding genes in the native producer, *S. goldiniensis.* Genes in the lilac region encode miscellaneous proteins with no obvious association specialised metabolite production.

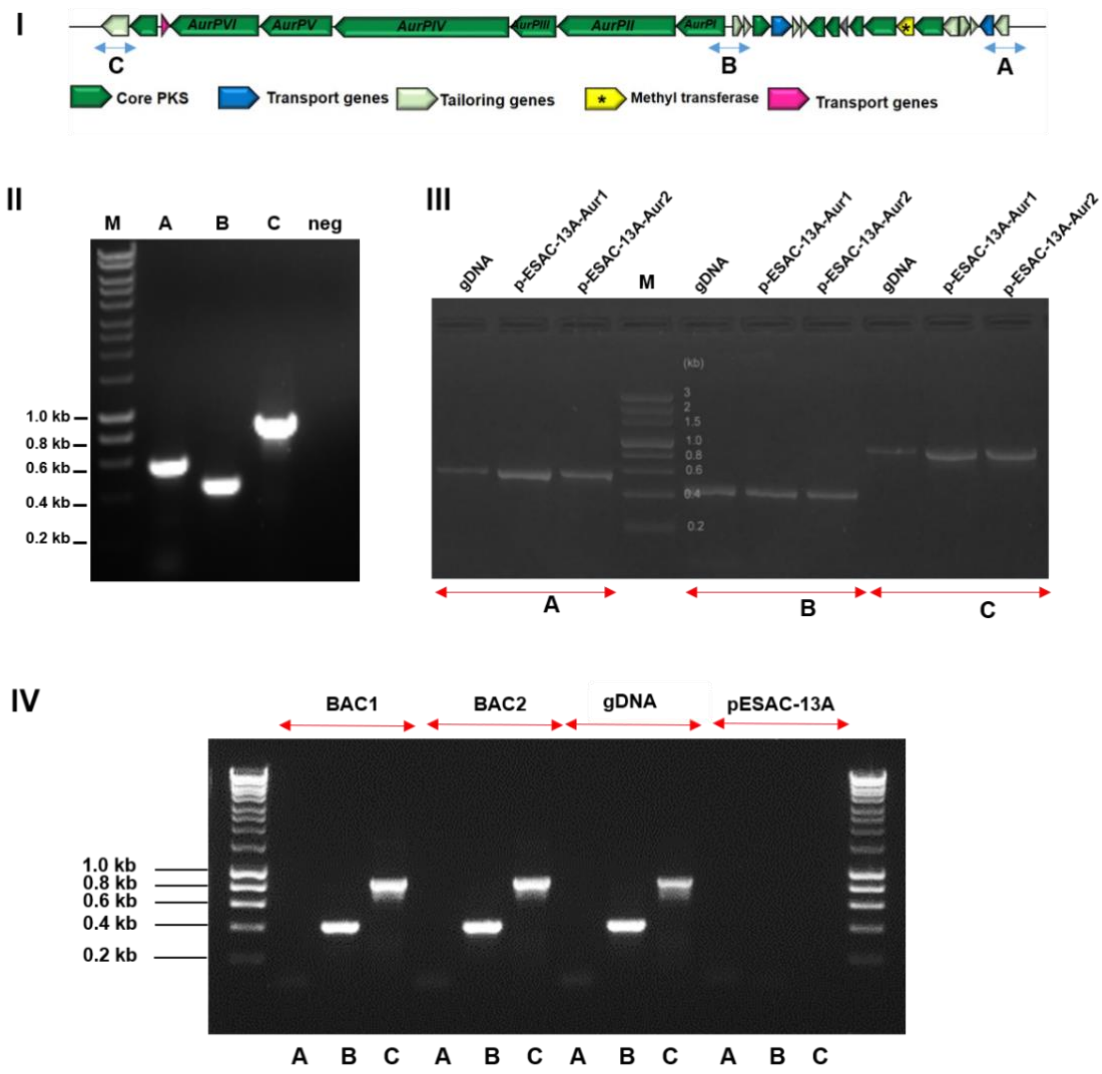

Figure S3: **Screening of pESAC-13A PAC library for triple positive aurodox clones. (I)** Map of aurodox gene cluster indicating target regions for PCR amplification during PAC screening. (Oligonucleotide primers can be found in **Supp Table S5**) (**II)** Agarose gel electrophoresis showing design of primers for Bio S & T screening and validation of their results. PCR products are 580, 512 and 863 bp respectively. Sequencing of amplicons from eurofins was used to confirm PCR products matched the putative aurodox cluster **(III)** Identification of two triple positive clones for putative aurodox encoding cluster, agarose gel electrophoresis carried out by Bio S and T **(IV)** PCR (primers can be found in **Supp Table S5**) and agarose gel electrophoresis confirmation that PACs derived from Bio S & T screening were positive for the aurodox cluster.

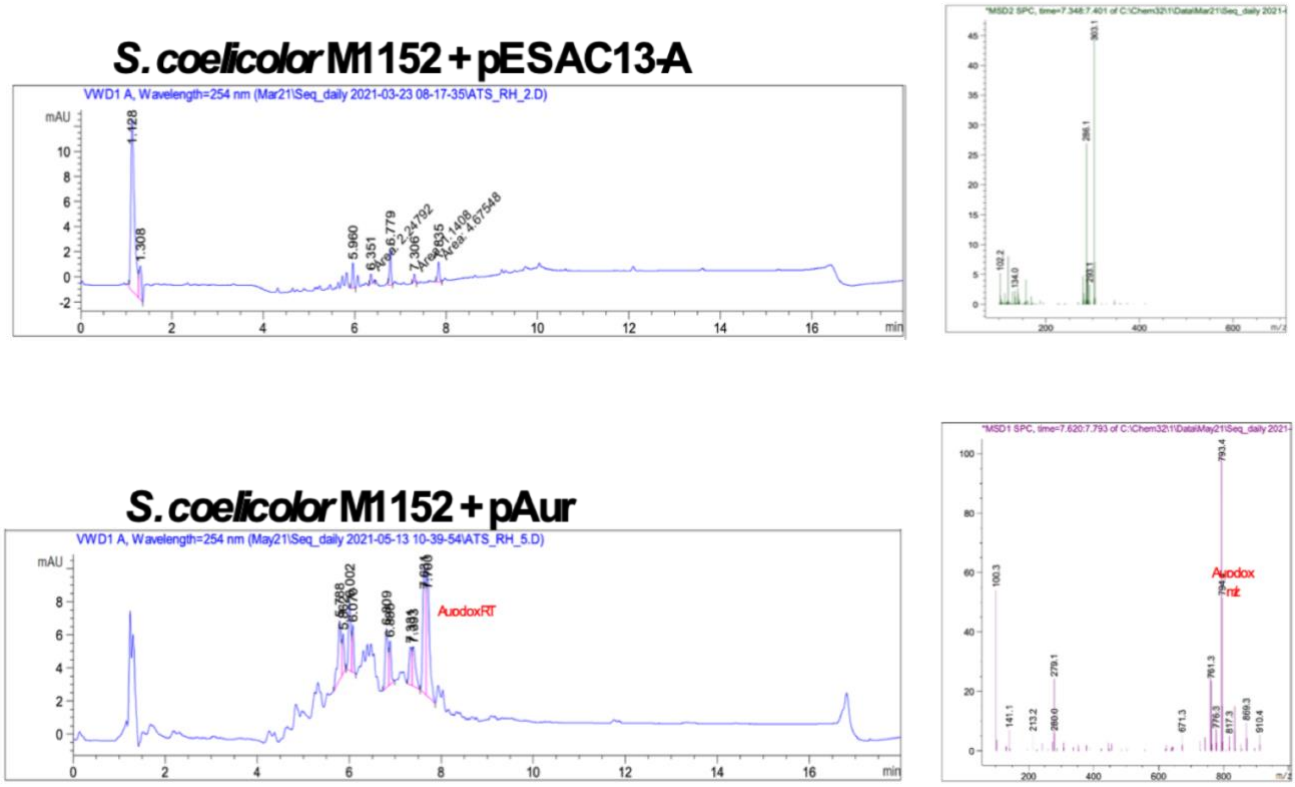

Figure S4: **LC-MS analysis of fermentation extracts demonstrating pAur confers production of aurodox**. *S. coelicolor* M1152 + pESAC-13A (Empty vector control) and *S. coelicolor* M1152 + pAurI. **(A)** Total ion chromatogram from analysis of *S. coelicolor* M1152 + pESAC-13A (Empty vector control) with no significant peaks corresponding to aurodox. **(B)** Total ion chromatogram from analysis of *S. coelicolor* M1152 + pAurI, peak corresponding to aurodox (~ 7.2 minutes) is indicated by red text, negative scan mode MS peaks are shown, with 793 m/z peak corresponding to aurodox anion.

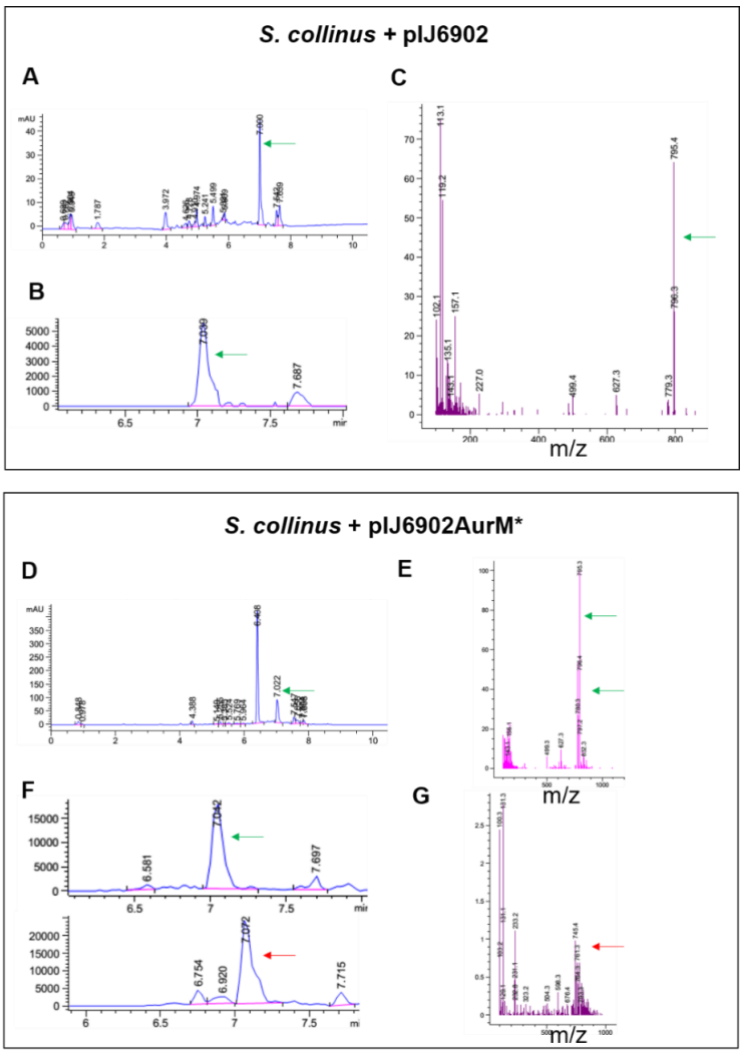

Figure S5: **LCMS analysis of *S. collinus* + pIJ6902AurM* versus empty vector control**. Extracts were generated through standard fermentation and extraction protocol. **(A)** Total Ion Chromatogram of extracts from *S.collinus* + pIJ6902, kirromycin peak (retention time 7.0 minutes) is indicated with a green arrow. **(B)** HPLC trace showing signal intensity from kirromycin associated peak. **(C)** Mass Spectrometry analysis of 7.0 min retention time peak, peak with m/z ratio of 795 (kirromycin) is indicated with a green arrow. **(D)** Total Ion Chromatogram of extract from *S. collinus* + pIJ6902AurM with kirromycin peak indicated with a green arrow. **(E)** Mass Spectrometry analysis of 7.0 min retention time peak, peak with m/z ratio of 795 (kirromycin) is indicated with a green arrow **(F)** HPLC trace showing signal intensity from kirromycin associated peak (green arrow) and aurodox associated peak (red arrow). **(G)** Mass Spectrometry analysis of 7.1 minute retention time aurodox peak signal with m/z ratio of 793 corresponding to aurodox is indicated with a red arrow.

References

1. Grant SG, Jessee J, Bloom FR, Hanahan D. 1990. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. Proc National Acad Sci 87:4645–4649.

2. Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc National Acad Sci 97:6640–6645.

3. MacNeil DJ, Gewain KM, Ruby CL, Dezeny G, Gibbons PH, MacNeil T. 1992. Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. Gene 111:61–68.

4. Campellone KG, Robbins D, Leong JM. 2004. EspFU Is a Translocated EHEC Effector that Interacts with Tir and N-WASP and Promotes Nck-Independent Actin Assembly. Dev Cell 7:217–228.

5. Gomez-Escribano JP, Bibb MJ. 2012. *Streptomyces coelicolor* as an expression host for heterologous gene clusters. Methods Enzymol 517:279–300.

6. Sosio M, Giusino F, Cappellano C, Bossi E, Puglia AM, Donadio S. 2000. Artificial chromosomes for antibiotic-producing actinomycetes. Nat Biotechnol 18:343–345.

7. Huang J, Shi J, Molle V, Sohlberg B, Weaver D, Bibb MJ, Karoonuthaisiri N, Lih C, Kao CM, Buttner MJ, Cohen SN. 2005. Cross-regulation among disparate antibiotic biosynthetic pathways of *Streptomyces coelicolor*. Molecular Microbiology 58:1276–1287.

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. Proceedings of the National Academy of Sciences of the United States of America 100:1541–6.

9. Sosio M, Bossi E, Donadio S. 2001. Assembly of large genomic segments in artificial chromosomes by homologous recombination in *Escherichia coli*. Nucleic Acids Res 29:e37–e37.