

RESEARCH

Open Access



# Characterization and engineering of the biosynthesis gene cluster for antitumor macrolides PM100117 and PM100118 from a marine actinobacteria: generation of a novel improved derivative

Raúl García Salcedo<sup>1</sup>, Carlos Olano<sup>1</sup>, Cristina Gómez<sup>1</sup>, Rogelio Fernández<sup>2</sup>, Alfredo F. Braña<sup>1</sup>, Carmen Méndez<sup>1</sup>, Fernando de la Calle<sup>2</sup> and José A. Salas<sup>1\*</sup>

## Abstract

**Background:** PM100117 and PM100118 are glycosylated polyketides with remarkable antitumor activity, which derive from the marine symbiotic actinobacteria *Streptomyces caniferus* GUA-06-05-006A. Structurally, PM100117 and PM100118 are composed of a macrocyclic lactone, three deoxysugar units and a naphthoquinone (NQ) chromophore that shows a clear structural similarity to menaquinone.

**Results:** Whole-genome sequencing of *S. caniferus* GUA-06-05-006A has enabled the identification of PM100117 and PM100118 biosynthesis gene cluster, which has been characterized on the basis of bioinformatics and genetic engineering data. The product of four genes shows high identity to proteins involved in the biosynthesis of menaquinone via futasoline. Deletion of one of these genes led to a decay in PM100117 and PM100118 production, and to the accumulation of several derivatives lacking NQ. Likewise, five additional genes have been genetically characterized to be involved in the biosynthesis of this moiety. Moreover, the generation of a mutant in a gene coding for a putative cytochrome P450 has led to the production of PM100117 and PM100118 structural analogues showing an enhanced in vitro cytotoxic activity relative to the parental products.

**Conclusions:** Although a number of compounds structurally related to PM100117 and PM100118 has been discovered, this is, to our knowledge, the first insight reported into their biosynthesis. The structural resemblance of the NQ moiety to menaquinone, and the presence in the cluster of four putative menaquinone biosynthetic genes, suggests a connection between the biosynthesis pathways of both compounds. The availability of the PM100117 and PM100118 biosynthetic gene cluster will surely pave a way to the combinatorial engineering of more derivatives.

**Keywords:** Polyketide, Antitumor, Modular polyketide syntase, Deoxysugar, Glycosyltransferase, Structural analogue

## Background

Actinobacteria is an extensive phyla within the domain bacteria with a wide distribution in nature, encompassing both terrestrial and aquatic environments [1, 2].

Numerous actinobacteria species have an outstanding medical value as producers of cancer chemotherapeutic drugs [3], among other biologically active compounds [4, 5]. Terrestrial actinobacteria are the source of the vast majority of natural antitumor agents discovered up to date, many of which are the base of currently available chemotherapeutic treatments or are in advance clinical trials [6]. However, in spite of the enormous potential of soil actinobacteria as producers of antitumor agents, the

\*Correspondence: jasalas@uniovi.es

<sup>1</sup> Departamento de Biología Funcional e Instituto Universitario de Oncología del Principado de Asturias (I.U.O.P.A), Universidad de Oviedo, 33006 Oviedo, Asturias, Spain

Full list of author information is available at the end of the article

rate of discovery of new antitumor drugs and other bioactive compounds from the terrestrial environment has nowadays decayed [7–9]. This, together with the worldwide-rising occurrence of cancer and the appearance of multi-drug resistant tumor cell lines, urges to extend the screening of new and improved chemotherapeutic products to less explored environments.

Ocean is a major component of the biosphere and an example of unexploited habitat with the potential to host the chemical and biological diversity required for the discovery of novel anticancer agents. A staggering diversity of actinobacteria species has been isolated from different marine substrates such as sediments seawater, seaweeds or mangroves [2]. Furthermore, tissues of mollusks and invertebrates are the niche of remarkably diverse symbiotic actinobacteria populations many of which could have lost the ability to grow independently of their host [10]. Some of the symbiotic functions attributable to symbiotic actinobacteria entail the production of bioactive secondary metabolites (e.g., for host defense) with potential pharmaceutical significance. Indeed, many natural products isolated from diverse marine invertebrates have been proven to be produced by their symbiotic microorganisms. That is the case, for example, of the dibenzodiazepinone diazepinomicin [11], the thiodipeptide thiocoraline [12] and the polyketide bryostatin [13], three potent antitumor drug leads; or the tetrahydroisoquinoline ecteinascidin 743 [14], an anticancer drug in current clinical use.

Although the screening of antitumor natural products from yet unexploited environments is a valid plan to alleviate the problem of the apparent chemical exhaustion of terrestrial sources, the use of metabolic engineering and combinatorial biosynthesis strategies intended to generate novel analogues from known natural agents [15–17], can be a no less interesting approach. However, application of these genetic engineering strategies requires to a certain extent detailed knowledge on the genetic and biochemical basis of the biosynthesis of relevant natural metabolites. Hence, the identification and characterization of biosynthetic gene clusters is not only an invaluable tool for the elucidation of the biosynthesis pathway of bioactive natural agents, but also an essential requirement for tackling the combinatorial engineering of novel analogues. In parallel with the development of these genetic engineering approaches, advances in next-generation whole-genome sequencing techniques [18] has favored the availability of a growing number of genomes from actinobacteria producing clinical attractive bioactive compounds. Mining of these genomes has enabled the identification and analysis of an exponentially increasing number of clusters [19], setting the stage for subsequent combinatorial engineering of novel

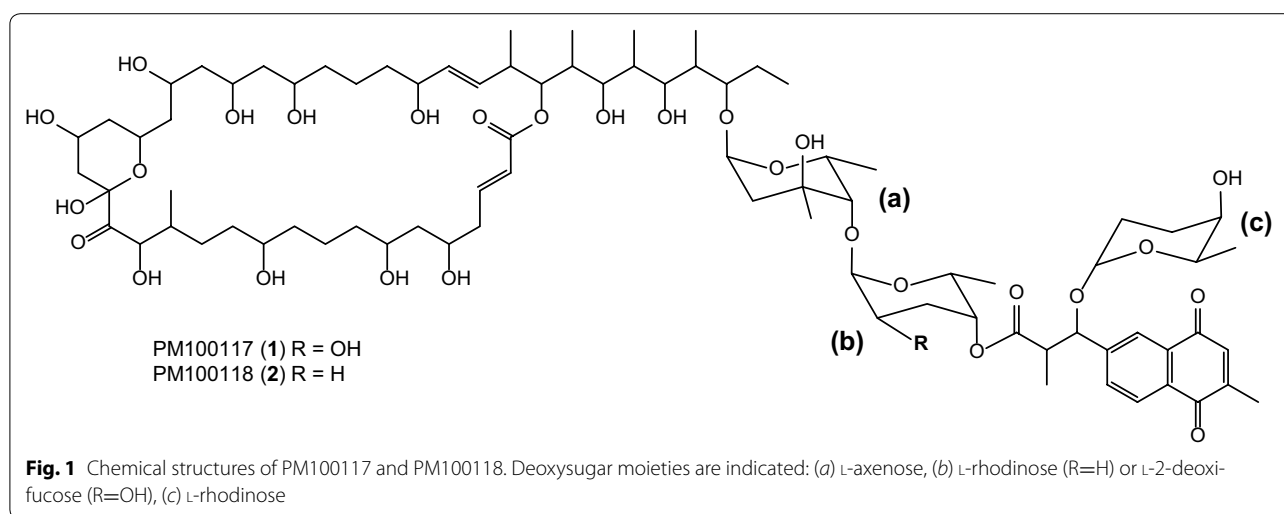
derivatives. Furthermore, genome analysis by enhanced bioinformatics platforms associated to database [20–24] allows for disclosing the chemical diversity potentially hosted in these actinobacteria, contributing to widen the availability of natural products by activation of cryptic clusters [19].

In a recent publication [25], the discovery of PM100117 and PM100118 was reported. These glycosylated polyketide compounds, with remarkable antitumor activity, are produced by *Streptomyces caniferus* GUA-06-05-006A, a symbiotic actinobacteria isolated from the marine sponge *Filigrana* sp. The structures of PM100117 and PM100118, consisting of a macrocyclic lactone, three deoxysugars and a 1, 4-naphthoquinone chromophore (Fig. 1), share a clear similarity with other antitumor polyketide compounds, including langkolide [26] and the promising anticancer drug lead GT35 [27]. Given their attractive as potential anticancer clinical drugs, PM100117 and PM100118 represent interesting targets for the combinatorial engineering of novel derivatives. The goals of this work were the identification and characterization of the PM100117 and PM100118 biosynthesis gene cluster, and engineering the gene cluster to generate novel derivatives with improved biological activity.

## Results

### Sequencing and bioinformatics analysis of *S. caniferus* GUA-06-05-006A genome

The first step to accomplish the identification of the PM100117 and PM100118 biosynthesis gene cluster was sequencing the *S. caniferus* GUA-06-05-006A chromosome. Whole-genome sequencing of this strain generated a total number of 548,579 paired-end reads with an average length of 372.5 nucleotides, producing a total sequence of 204.3 Mb. This represents a 20-fold coverage of the chromosome sequence, which is estimated in 9.8 Mb. De novo assembly of those sequences resulted in 907 contigs. Median (N50) of the contig assembly was 20.9 Kb, and the largest was around 119.3 Kb. Subsequent contig arrangement rinsed 33 scaffolds (mean: 276.2 Kb; N50: 1.5 Mb) generating a draft genome of 9.1 Mb with a G + C content of 70.63 %. *In silico* genome analyses with the antibiotics and Secondary Metabolite Analysis Shell (antiSMASH) algorithm [22] revealed the presence of 8582 open reading frames (ORFs), having at least 5045 proteins an assigned putative function. Sequence analyses with antiSMASH also showed the presence of 32 gene clusters potentially involved in the biosynthesis of secondary metabolites. Seven of these clusters were identified as containing gene sequences belonging to the type I (four clusters), II (two clusters) and III (one cluster) family of polyketide synthases (PKS). Likewise, genome sequence analysis detected nine additional gene clusters



comprising modular enzyme-coding genes such as non-ribosomal peptide synthetase (NRPS, seven clusters) and hybrid PKS-I-NRPS genes (two clusters). Other products from these gene clusters include one nucleoside, five terpenes and three butyrolactones, as well as compounds with a peptidic backbone such as three siderophores, three lantipeptides and one ectoine.

#### Identification of PM100117 and PM100118 biosynthetic gene cluster

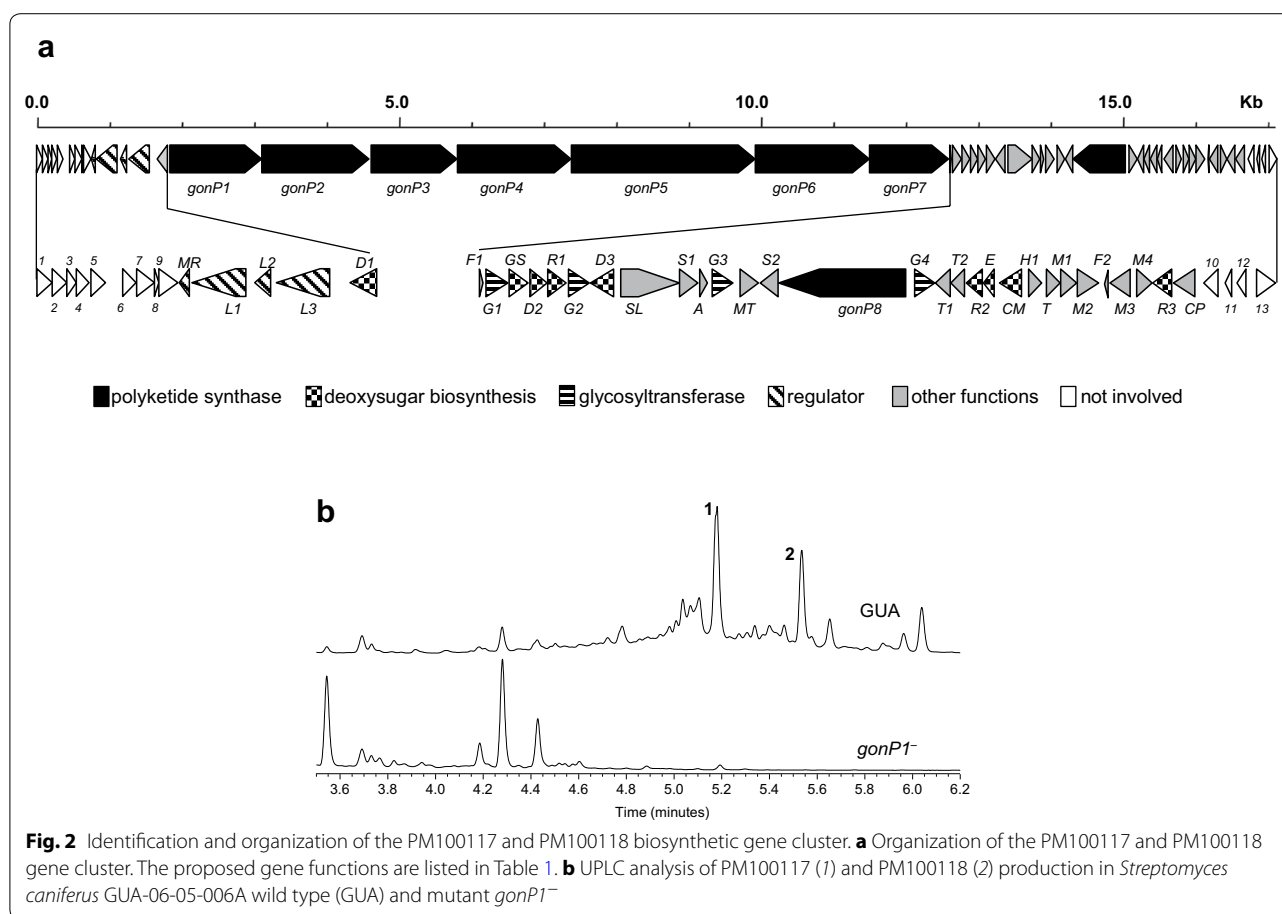
PM100117/18 chemical structures have been previously elucidated by nuclear magnetic resonance (NMR) spectroscopy [25]. They contain a 48 carbons aglycone that is likely biosynthesized by the condensation of 21 ketide moieties. The aglycone forms a 36-membered macrolactone ring and is decorated with a side chain consisting of three 2,6-dideoxy sugars and a 1,4-naphthoquinone chromophore (Fig. 1). The first sugar moiety attached to the aglycone is L-axenose, which can be linked either to L-2-deoxy-fucose (PM100117) or L-rhodinose (PM100118) as second sugar. The second sugar unit is linked to the naphthoquinone moiety, which shares a clear structural similarity with menaquinone (MK). An L-rhodinose moiety attached to the naphthoquinone structure represents the third deoxysugar of the PM100117/18 glycosylation pattern. Based on these structural features, glycosyltransferase and sugar biosynthetic genes, as well as a minimum of 21 PKS modules, are presumably involved in the biosynthesis of PM100117/18. Among the seven PKS gene clusters identified by antiSMASH, one of them contained ORFs designated with such putative functions, representing the most suitable candidate cluster to accomplish PM100117/18 biosynthesis. The putative PM100117/18 cluster covers a 171 kb region and contains 54 ORFs (Fig. 2a) coding for proteins with

the putative functions listed in Table 1. The involvement of this cluster in PM100117/18 biosynthesis was demonstrated by inactivation of the PKS gene *gonP1*. The analysis of the resulting mutant strain showed that *gonP1* inactivation abolished PM100117/18 biosynthesis, thus confirming the implication of this cluster in the production of PM100117/18 (Fig. 2b).

#### *In silico* analysis of PM100117 and PM100118 gene cluster and proposed biosynthesis pathway

##### Polyketide ring and post-PKS modifications

The PM100117/18 cluster encompasses seven contiguous PKS genes (*gonP1–gonP7*) coding for a multimodular PKS that comprises a loading domain (LD) and 20 extension modules (M1–M20), in agreement with the 21 condensation steps required for the biosynthesis of the PM100117/18 macrolide ring. Sequence analysis of GonP1–GonP7 allowed to define ketosynthase (KS or KSQ), acyltransferase (AT), ketoreductase (KR), dehydratase (DH), enoylreductase (ER) and acyl-carrier-protein (ACP) domains. LD contains a ketosynthase (KSQ) domain, in which the essential cysteine of the conserved DTxCSxS sequence at the active site is replaced by glutamine [28]. Sequence alignment of the cluster PKS domains active sites is shown as supplementary data (Additional file 1: Figure S1). AntiSmash analysis also predicted substrate specificity of AT domains within the modular PKS for methylmalonyl-CoA (ATp) in seven modules (LD, M1–M4, M14 and M19) and for malonyl-CoA (ATa) in 14 modules (M5–M13, M15–M18 and M20). The organization of the PKS genes and the proposed biosynthesis pathway of the PM100117/18 macrolide backbone are shown in Fig. 3. Overall, the predicted substrate specificity of AT domains alongside the disposal of the defined KR, DH and ER domains are



consistent with the hypothetical elongation through collinear reactions of the polyketide chain. Only a few discrepancies are found between this biosynthesis model and the chemical structure of the polyketide moiety, suggesting the inactivity/inability of certain PM100117/18 PKS domains or the occurrence of post-PKS modifications. The presence of a DH domain in M6 (GonP3), as well as a DH and an ER domain in M9 (GonP4), apparently disturbs the collinear biosynthesis of the polyketide, given the presence of hydroxyl groups at carbon (C) 31 and C25 of the polyketide skeleton (Fig. 3). Nonetheless, a detailed sequence analysis of the M6 DH domain revealed an arginine in place of an otherwise conserved histidine in the NADPH motif LxxHxxxGxxxxP [29] at the active site (Additional file 1: Figure S1). This, together with a shorter C-terminal sequence in comparison with other DH domains within the cluster PKSs, suggests that M6 DH is most possibly inactive. By contrast, sequence analysis did not show any alteration of the M9 DH and ER domains, which are probably functional. Thus, the unexpected presence of these domains could be explained by a “domain skipping” mechanism preventing dehydrogenation of C25 during polyketide elongation

[30]. Alternatively, the presence of a hydroxyl group at C25 can be the result of a post-PKS modification following polyketide biosynthesis, catalyzed by any of the two putative oxygenases coded by *gonCP* and *orf9*. Moreover, the involvement of the M13 DH domain (GonP5) in the polyketide biosynthesis is unlikely due to the lack of a KR domain in this module catalyzing a previous ketoreduction reaction. One last inconsistency is found between the substrate specificity of the AT domain of M19 (GonP7), which is predicted to utilize methylmalonyl-CoA as extension unit, and the absence of a methyl group at C4 (Fig. 3). However, the presence in this domain of a YASH motif, which specifies utilization of methylmalonyl-CoA [31], can be confirmed as shown in the alignment of cluster AT active sites (Additional file 1: Figure S1). Despite this, no PM100117/18 analogues containing an additional methyl group at C4 have been detected in *S. caniferus* GUA-06-05-006A cultures.

#### Sugars biosynthesis

A distinctive feature of the PM100117/18 biosynthesis gene cluster, among the 32 clusters for secondary metabolites hosted by *S. caniferus* GUA-06-05-006A, is

**Table 1 Deduced functions of ORFs in PM100117 and PM100118 biosynthetic gene cluster**

Gene	Deduced aa. length	Protein homologue (accession no.)	Identity/similarity (%)	Proposed function
<i>orf1</i>	264	WP_030324639.1	84/92	Aspartate dehydrogenase
<i>orf2</i>	265	WP_030324637.1	77/83	Hydrolase
<i>orf3</i>	173	WP_018535062.1	94/98	Cupin domain-containing protein
<i>orf4</i>	228	EDY62802.1	81/86	3-Ketoacyl-(acyl-carrier-protein) reductase
<i>orf5</i>	260	WP_005320771.1	83/89	Short-chain dehydrogenase
<i>orf6</i>	541	WP_033304767.1	82/87	Acetolactate synthase
<i>orf7</i>	313	WP_005320767.1	79/85	Ferredoxin
<i>orf8</i>	75	WP_031101148.1	70/75	Hypothetical protein
<i>orf9</i>	337	WP_033304770.1	88/93	3-Phenylpropionate dioxygenase
<i>gonMR</i>	176	WP_030624728.1	92/95	MarR family transcriptional regulator
<i>gonL1</i>	959	WP_006607813.1	86/90	LuxR family transcriptional regulator
<i>gonL2</i>	241	EJJ02892.1	80/84	LuxR family transcriptional regulator
<i>gonL3</i>	939	WP_006607811.1	80/87	LuxR family transcriptional regulator
<i>gonD1</i>	469	WP_006607808.1	90/94	NDP-hexose 2,3-dehydratase
<i>gonP1</i>	4219	AJE81080.1	60/69	Type I polyketide synthase (loading module and modules 1 and 2)
<i>gonP2</i>	4939	AAZ94388.1	57/68	Type I polyketide synthase (modules 3 to 5)
<i>gonP3</i>	3959	EJJ02484.1	86/89	Type I polyketide synthase (modules 6 and 7)
<i>gonP4</i>	5234	EJJ02491.1	86/90	Type I polyketide synthase (modules 8 to 10)
<i>gonP5</i>	8461	EDX25447.1	55/67	Type I polyketide synthase (modules 11 to 15)
<i>gonP6</i>	5238	EJJ02498.1	87/90	Type I polyketide synthase (modules 16 to 18)
<i>gonP7</i>	3641	EJJ02497.1	87/91	Type I polyketide synthase (modules 19 and 20)
<i>gonF1</i>	67	WP_006608219.1	76/80	Ferredoxin
<i>gonG1</i>	415	ABO28820.1	66/80	Glycosyltransferase
<i>gonG5</i>	358	WP_006608216.1	94/97	Glucose-1-phosphate thymidyltransferase (glucose synthase)
<i>gonD2</i>	333	WP_040902050.1	91/93	dTDP-glucose 4,6-dehydratase
<i>gonR1</i>	356	WP_051007140.1	70/79	NDP-4-keto-6-deoxyhexose reductase
<i>gonG2</i>	408	WP_006608213.1	74/83	Glycosyltransferase
<i>gonD3</i>	433	BAP34744.1	73/84	NDP-hexose-3,4-dehydratase
<i>gonSL</i>	1107	WP_006608212.1	68/78	AMP-dependent synthetase and ligase (acyl-CoA ligase, DH, KR, ACP)
<i>gonS1</i>	347	WP_040902110	82/91	3-oxoacyl-ACP synthase III
<i>gonA</i>	146	WP_040902049	72/84	Cold-Shock DNA-binding domain-containing protein
<i>gonG3</i>	402	WP_006608209	55/68	Glycosyltransferase
<i>gonMT</i>	345	WP_014208548	49/62	Methylase involved in ubiquinone/menaquinone biosynthesis
<i>gonS2</i>	344	WP_040246457	81/90	3-Oxoacyl-ACP synthase III
<i>gonP8</i>	2408	WP_006608212	58/59	Type I polyketide synthase (loading module and extension module 1)
<i>gonG4</i>	368	WP_042158955	89/93	Glycosyltransferase
<i>gonT1</i>	286	WP_040902100.1	83/92	Multidrug ABC transporter permease
<i>gonT2</i>	260	WP_006608193.1	89/94	Multidrug ABC transporter permease
<i>gonR2</i>	319	AAB63047.1	49/63	dTDP-4-keto-2,3,6-trideoxyhexulose reductase
<i>gonE</i>	211	WP_006608191.1	83/89	dTDP-4-deoxyglucose 3,5-epimerase

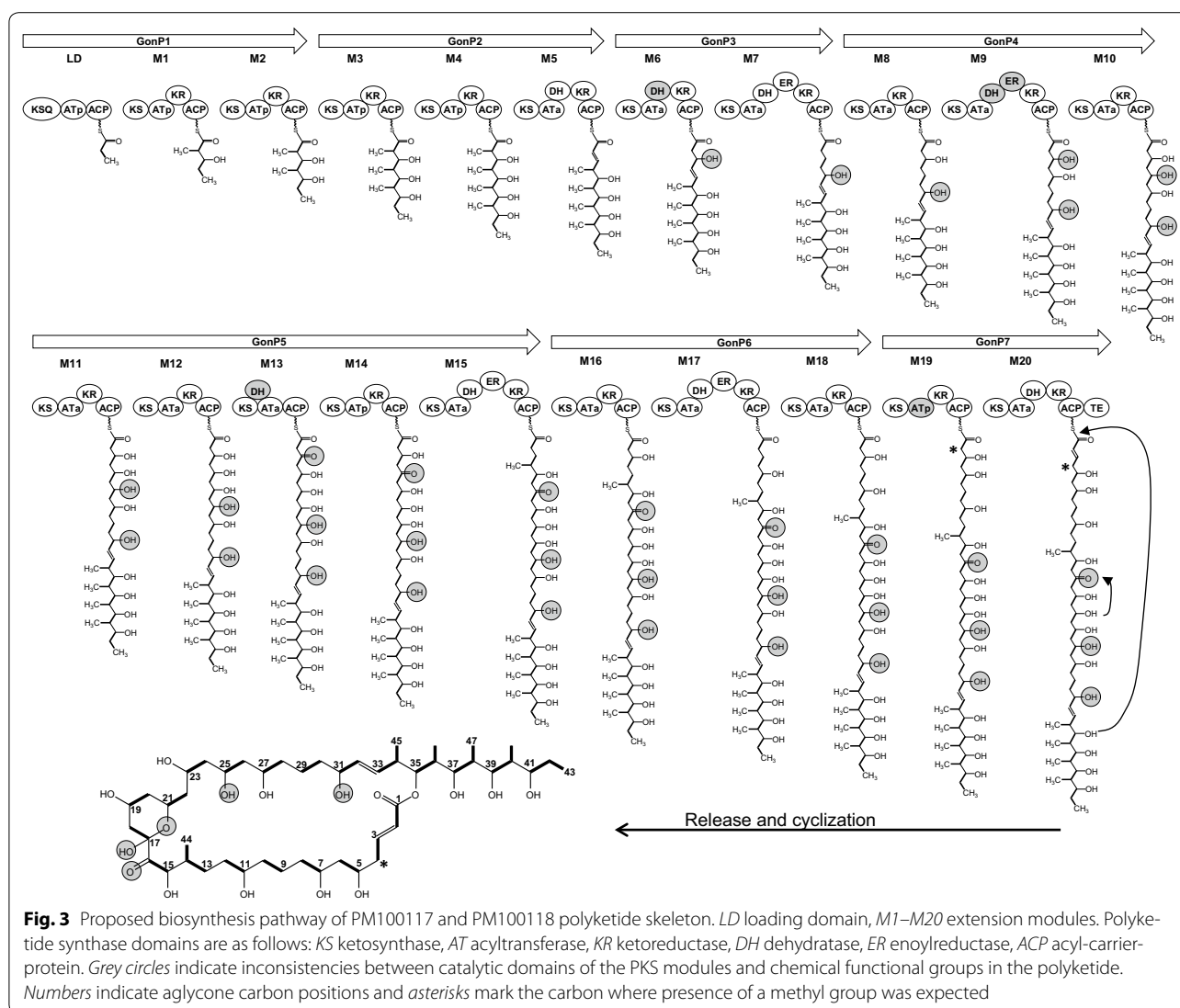
**Table 1 continued**

Gene	Deduced aa. length	Protein homologue (accession no.)	Identity/similarity (%)	Proposed function
<i>gonCM</i>	410	AKT74284.1	79/87	NDP-hexose-3-C-methyltransferase
<i>gonH1</i>	252	WP_006608189.1	79/88	Hypothetical protein
<i>gonT</i>	256	WP_006608180.1	90/94	Thioesterase
<i>gonM1</i>	301	WP_046708157.1	77/86	Menaquinone biosynthesis protein (chorismate dehydratase)
<i>gonM2</i>	407	EJJ02566.1	96/97	Menaquinone biosynthesis protein (dehypoxanthine futasoline cyclase)
<i>gonF2</i>	101	WP_006608177.1	94/96	Ferredoxin
<i>gonM3</i>	387	WP_006608176.1	98/99	Menaquinone biosynthesis protein (aminofutasoline synthase)
<i>gonM4</i>	291	WP_006608174.1	87/91	Menaquinone biosynthesis protein (1,4-dihydroxy-6-naphthoate synthase)
<i>gonR3</i>	364	WP_006608173.1	91/94	NDP-hexose 3-ketoreductase
<i>gonCP</i>	482	WP_006608172.1	91/95	Cytochrome P450
<i>orf10</i>	264	WP_030985702.1	88/92	3-Oxoacyl-ACP reductase
<i>orf11</i>	128	WP_053688399.1	86/91	HxlR family transcriptional regulator
<i>orf12</i>	174	WP_051733431.1	93/96	Hypothetical membrane protein
<i>orf13</i>	355	WP_037651728.1	90/93	NADPH: quinone reductase

the presence of ORFs with assigned putative functions implicated in biosynthesis or transfer of deoxysugars. Based on these putative functions the biosynthesis pathway of L-axenose, L-2-deoxy-fucose and L-rhodinose can be predicted (Fig. 4). Proteins GonGS and GonD2, putative NDP-glucose synthase and NDP-glucose 4,6-dehydratase, respectively, might catalyze the biosynthesis of the key intermediate NDP-4-keto-6-deoxy-D-glucose [32], which should be then transformed into NDP-4-keto-2,6-dideoxy-D-glucose (2,6-DG) by the activity of the putative NDP-hexose 2,3-dehydratase GonD1 and the putative NDP-hexose 3-ketoreductase GonR3. Biosynthesis of L-2-deoxy-fucose from 2,6-DG requires 3,5-epimerization (3,5-EPI) and C4-ketoreduction (C4-KR) reaction steps, which are possibly catalyzed by the putative dTDP-deoxyglucose 3,5-epimerase GonE and one of the putative NDP-4-keto-6-deoxyhexose reductases GonR1 or GonR2, respectively. In addition to 3,5-EPI and C4-KR reactions, biosynthesis of L-axenose from 2,6-DG involves a C3-methylation step, presumably catalyzed by the putative NDP-hexose-3-C-methyltransferase GonCM. Moreover, biosynthesis of L-rhodinose requires the C3-dehydration reaction of 2,6-DG, possibly catalyzed by the putative NDP-hexose-3,4-dehydratase enzyme GonD3, followed by 3,5-EPI and C4-KR. After macrolactone formation, four putative glycosyltransferase-coding genes (*gonG1*, *gonG2*, *gonG3* and *gonG4*) could be involved in the transfer of the three deoxysugar moieties to generate PM100117/18.

#### **Naphthoquinone (NQ) biosynthesis**

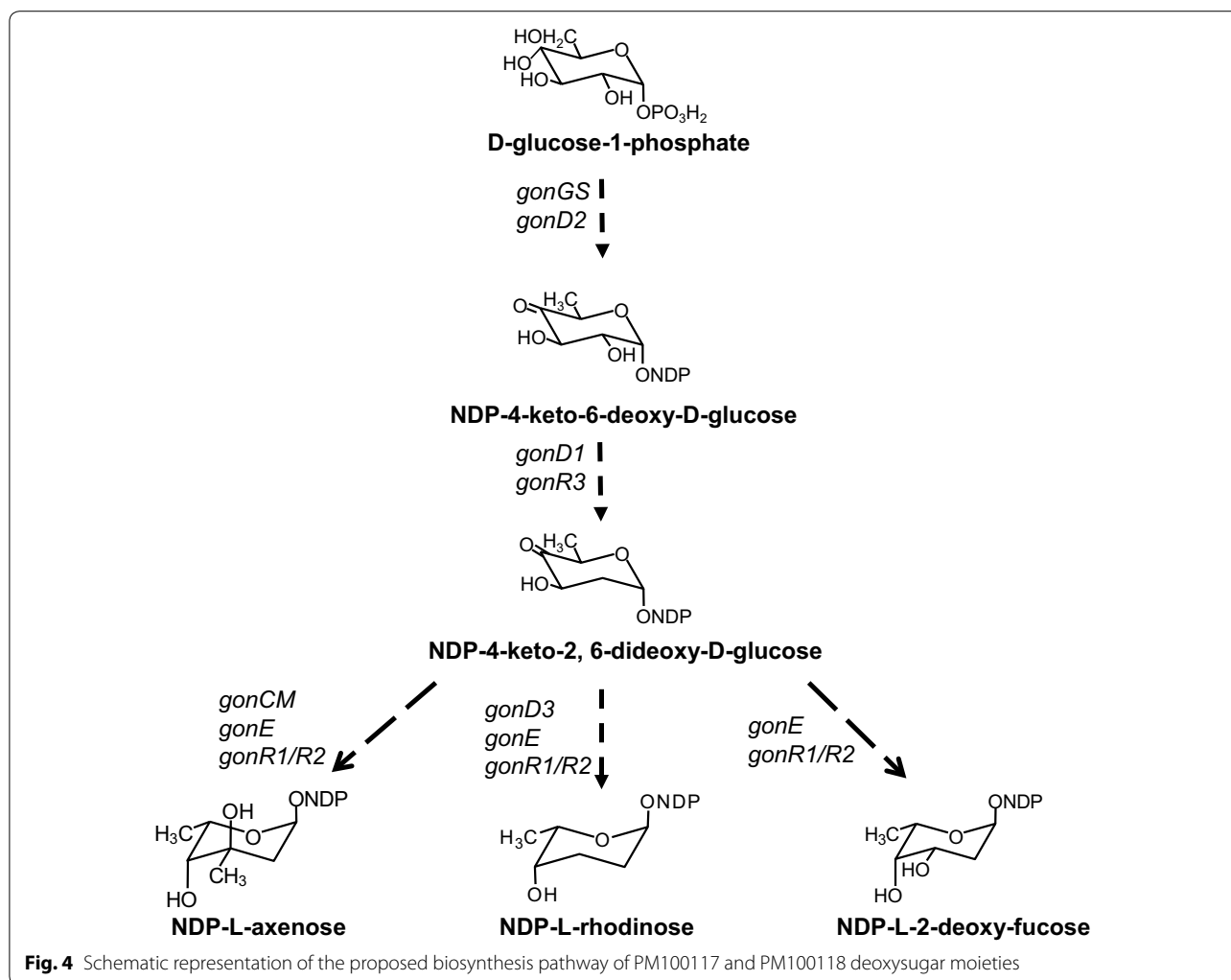
A further peculiarity of the PM100117/18 biosynthesis cluster is the presence of four ORFs (*gonM1*, *gonM2*, *gonM3* and *gonM4*) with high identity to genes previously reported as involved in menaquinone biosynthesis via futasoline. This newly discovered MK pathway was first described in *S. coelicolor* A3(2) [33–35] but bioinformatics analyses suggest its presence also in other bacteria lacking the classic MK biosynthetic pathway via isochorismate [33, 36, 37]. Sequence analysis reveals a high degree of identity (I) and similarity (S) of GonM1, GonM2 and GonM4 to the *S. coelicolor* A3(2) proteins SCO4506/MqnA (I: 60.1 %, S: 74.4 %), SCO4550/MqnC (I: 85 %, S: 91.6 %) and SCO4326/MqnD (I: 69.5 %, S: 77.1 %), respectively. These *S. coelicolor* A3(2) proteins, together with SCO4327/MqnB, have been previously shown to be involved in the biosynthesis of the futasoline MK pathway intermediate 1,4-dihydroxy 6-naphthoic acid (DH6N) [33–35]. Furthermore, GonM3 shares sequence resemblance with SCO4494/MqnE (I: 93 %, S: 96.4), a protein proposed to catalyze the biosynthesis of aminodeoxyfutasoline [38], an alternative substrate for MK biosynthesis. An additional copy of genes coding for proteins with high identity and similarity to *S. coelicolor* A3(2) MqnA (I: 56.6 %, S: 62.2 %), MqnB (I: 68.1 %, S: 72.9 %), MqnC (I: 92.7 %, S: 97.0 %) and MqnD (I: 72.2 %, S: 75.7 %) has been detected in the *S. caniferus* GUA-06-05-006A genome, outside the PM100117/18 gene cluster. The structural similarity of DH6N with the NQ unit led us to suspect a role of genes *gonM1–gonM4* on



the biosynthesis of this compound. Other ORFs potentially involved in the biosynthesis of NQ code for AMP-dependent synthetase and ligase GonSL, 3-oxoacyl-ACP synthase III GonS1 and GonS2, type I PKS GonP8 and methyltransferase GonMT. GonSL contains an apparently active CoA ligase (CAL) and ACP domain, and a KR and DH domain with important amino acid substitutions (Additional file 1: Figure S1). The putative PKS gene *gonP8* consists of a loading module containing a CAL and an ACP domain, and an extension module with a KS, ATp, KR, DH and ACP domain, all of them with no significant active site amino acid replacements (Additional file 1: Figure S1).

The proposed biosynthesis pathway of the PM100117/18 NQ moiety is depicted in Fig. 5. GonM1 (MqnA), GonM2 (MqnC), GonM4 (MqnD) and a futalosine hydrolase (MqnB) enzyme coded by a gene external

to the PM100117/18 cluster could catalyze the biosynthesis of DH6N, which would be then methylated by the putative methyltransferase GonMT to form 3-methyl-DH6N (Fig. 5). Interestingly, GonMT is only 11 % identical to SCO4556, which is the enzyme proposed to catalyze the quinone-ring C2-methylation in the last step of MK biosynthesis [33–35]. There is no available information to deduce the possible mechanism by which 3-methyl-DH6N is elongated with a propionate unit to form the NQ moiety. However, based on the fatty acid activation process catalyzed by fatty acyl-AMP ligases [39], we could speculate that the putative synthetase-ligase GonSL could catalyze the synthesis of a 3-methyl-DH6N-AMP adduct, and the subsequent transfer of 3-methyl-DH6N to the pantetheine group of its own ACP domain. In a next step, 3-methyl-DH6N might be transferred to the LD of PKS GonP8, and then elongated



by the GonP8 extension module. Binding of 3-methyl-DH6N-AMP to the GonSL ACP domain, and its following transfer to the LD of PKS GonP8, may require the participation of any of the putative 3-oxoacyl-ACP synthase III (KSIII) enzymes coded by genes *gonS1* and *gonS2*. The involvement of KSIII proteins in the priming of starter units alternative to malonyl-CoA and methyl-malonyl-CoA has been previously reported in the biosynthesis of a number of compounds [40–42]. Likewise, the mechanism by which NQ is eventually transferred to PM100117/18 may also require intervention of KSIII, GonP8 or both.

#### Pathway regulation

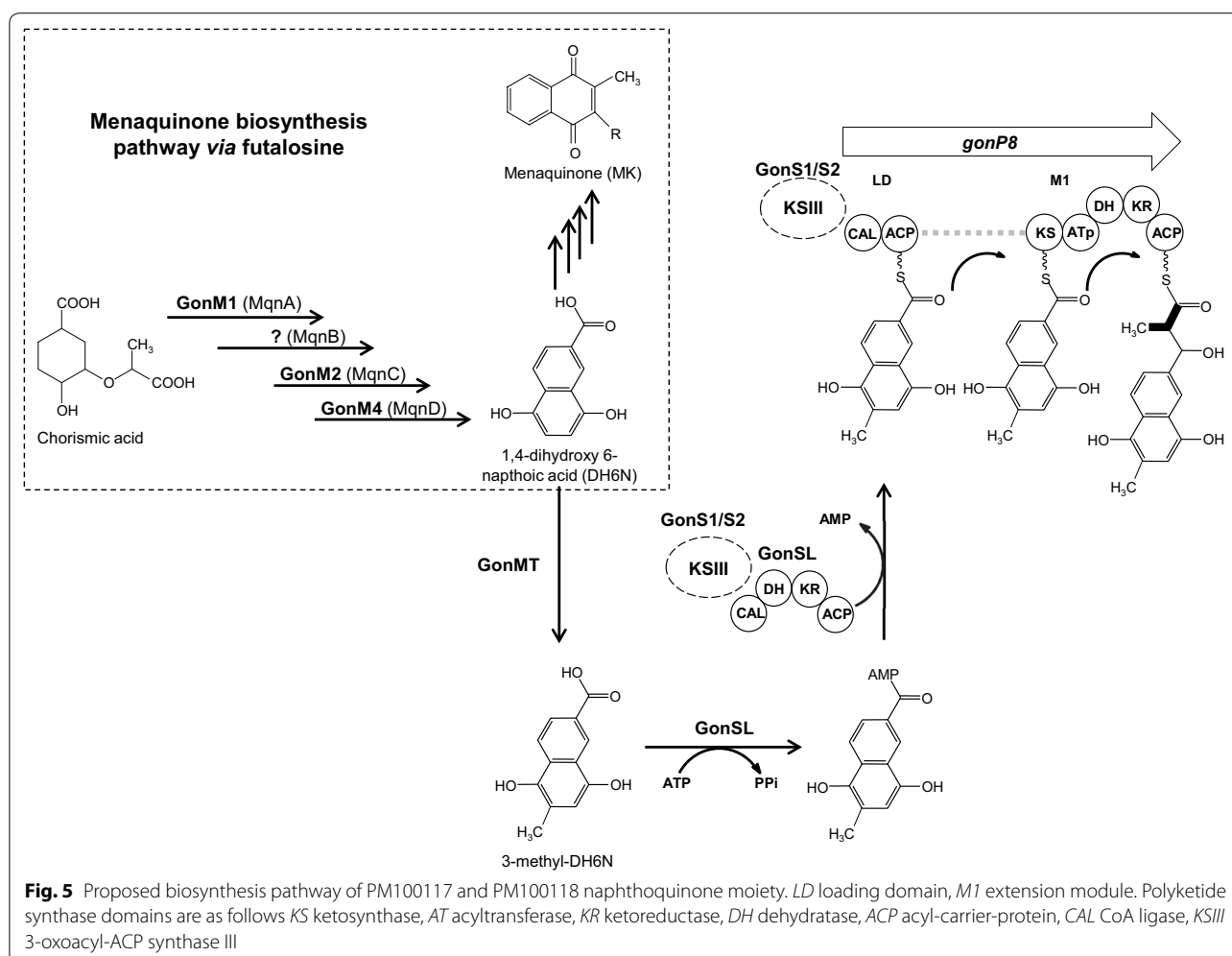
Four ORFs (*gonMR*, *gonL1*, *gonL2* and *gonL3*) could be responsible for PM100117/18 pathway regulation as they code for proteins with high sequence resemblance to transcription regulatory proteins. GonMR contains a helix-turn-helix (HTH) motif (smart00347) highly

conserved among members of the MarR family of protein regulators. GonL1 and GonL3 contain a N-terminal nucleoside triphosphate binding motif (pfam13191) and a HTH domain of the LuxR-type at de C terminus (smart00421), which are distinctive functional features of the LAL family of transcription factors [43, 44]. In addition to a LuxR-like C-terminal HTH motif, GonL2 contains a PAS-sensor binding fold at the N terminus (cd00130). This domain is designated PAS because of its homology to the *Drosophila* period protein (Per), the aryl hydrocarbon receptor nuclear translocator protein (ARNT) and the *Drosophila* single-minded protein (Sim) [45].

#### PM100117 and PM100118 transport

ORFs *gonT1* and *gonT2* code for a putative ATP-binding cassette (ABC) transport system. ABC transport complexes are comprised of a hydrophilic protein containing an ATP-binding domain and a hydrophobic protein with



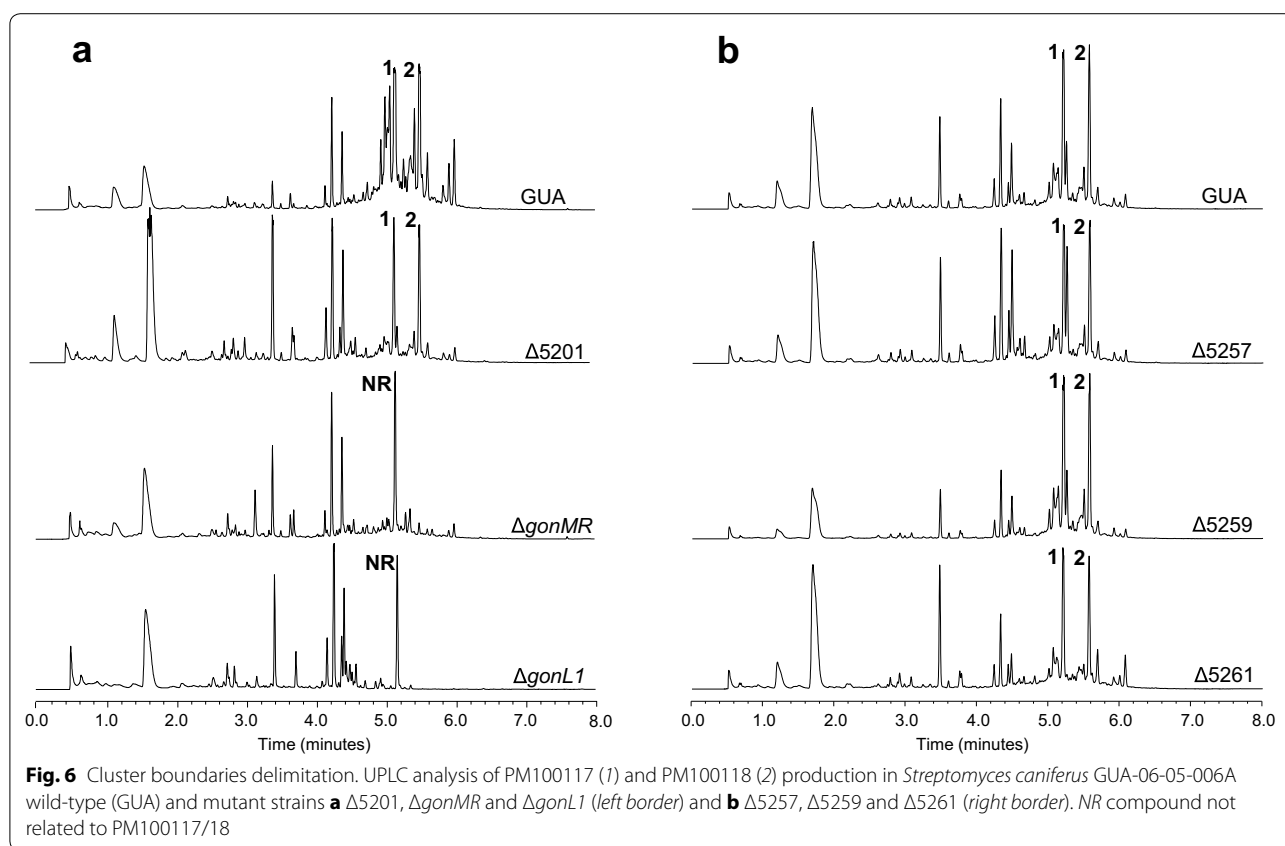


six membrane-spanning domains [46]. GonT1 shares high sequence identity and similarity with ABC-type membrane permeases and its analysis with the TMHMM server predicts the formation in this protein of six hydrophobic transmembrane domains. Conversely, GonT2 displays an ABC-like nucleoside triphosphate hydrolase domain (cl21455), which in combination with the putative permease GonT1 could produce a complex to facilitate PM100117/18 transport across the membrane.

#### Cluster boundaries analysis

On the cluster left side, genes *gonMR*, *gonL1*, *gonL2* and *gonL3* were presumed to code for pathway-specific regulators for PM100117/18 biosynthesis, and *orf9*, which codes for a putative dioxygenase, was a candidate gene to accomplish polyketide post-PKS oxygenations. Thus, to verify this boundary, we performed ultra-performance liquid chromatography (UPLC) analysis of PM100117/18 production in mutant strains  $\Delta\text{gonL1}$ ,  $\Delta\text{gonMR}$  and  $\Delta\text{5201}$  in which *gonL1*, *gonMR* and *orf9*, respectively,

have been deleted. The result of this analysis showed that in  $\Delta\text{5201}$  PM100117/18 production was not altered with respect to *S. caniferus* GUA-06-05-006A wild type. By contrast, in  $\Delta\text{gonL1}$  and  $\Delta\text{gonMR}$ , PM100117/18 biosynthesis was absent or severely decreased (Fig. 6a). Curiously, deletion of these genes induced the production of a new compound (NR), with a maximum absorption wavelength at 260 nm, apparently not related to PM100117/18 biosynthesis. Since PM100117/18 derivatives lacking small functional groups could possess retention times close to those of the parental products,  $\Delta\text{5201}$  extracts were also analyzed by LC-MS, confirming that peaks labeled as **1** and **2** indeed correspond to PM100117 and PM100118, respectively. These results verified the involvement of genes *gonL1* and *gonMR* in PM100117/18 biosynthesis, presumably coding for positive pathway-specific transcriptional regulators. PM100117/18 production was partially recovered in mutant strains  $\Delta\text{gonL1}$  and  $\Delta\text{gonMR}$  when a copy of *gonL1* and *gonMR*, respectively, was re-introduced (Additional file 1: Figure S2).

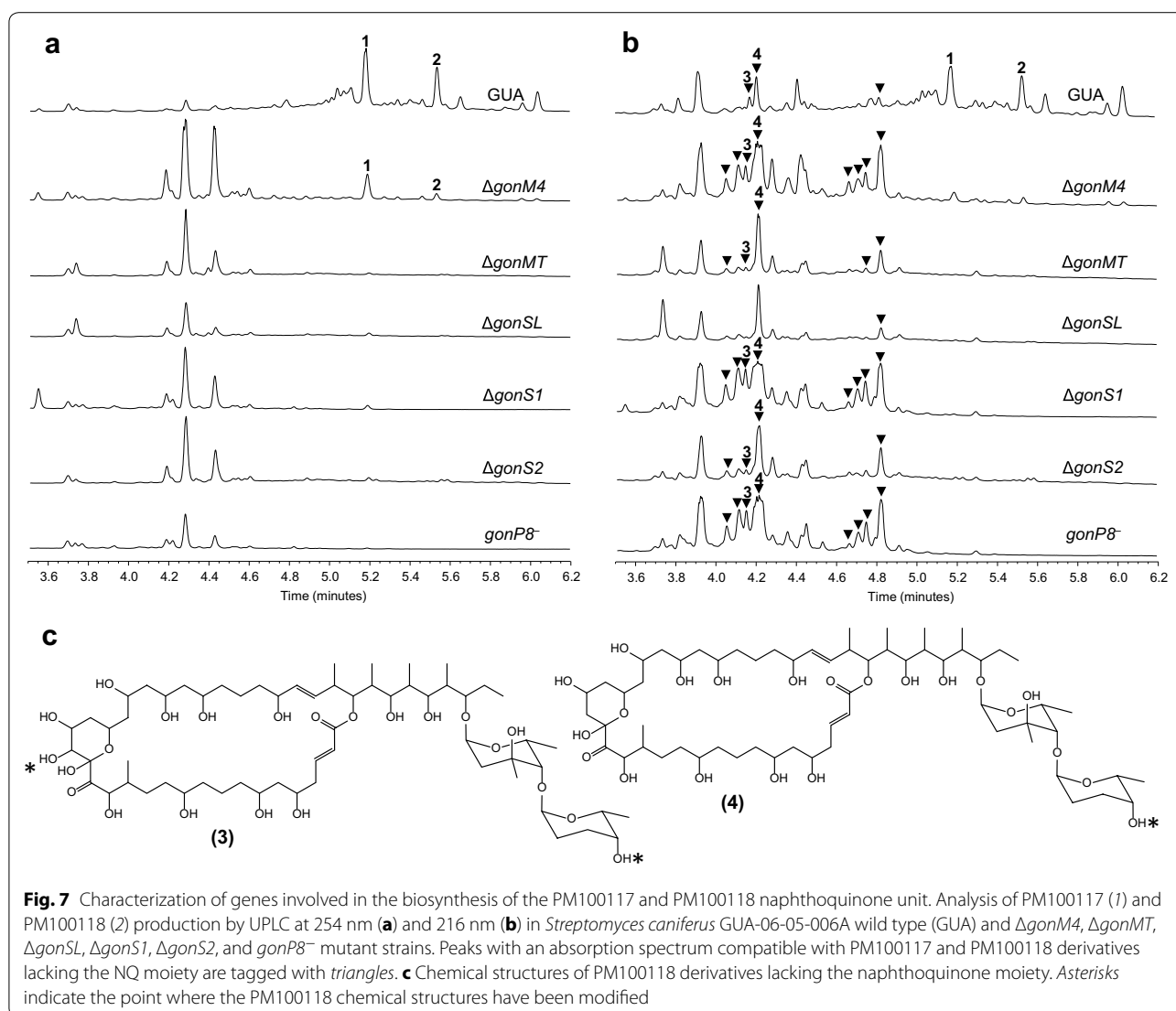


The cluster left border is thus apparently defined by *gonMR*. Upstream to this gene, antiSMASH detected nine ORFs coding for proteins with putative functions (Table 1) likely not pertaining to PM100117/18 biosynthesis. Most of these putative activities are related to oxidation-reduction reactions, such as ferredoxin (*orf7*), short-chain dehydrogenase (*orf5*), 3-ketoacyl-ACP reductase (*orf4*) and aspartate dehydrogenase (*orf1*), as well as a cupin domain-containing protein (*orf3*).

On the cluster right side, gene *gonCP* has been demonstrated by genetic engineering to be involved in the polyketide post-PKS oxygenation. Detailed data on this finding are described in section below. Adjacent to *gonCP*, antiSMASH detected ORFs coding for proteins with putative activities 3-oxoacyl-ACP reductase (*orf10*), HxLR family transcriptional regulator (*orf11*), hypothetical membrane protein (*orf12*) and NADPH-dependent quinone reductase (*orf13*). Hence, in order to verify the cluster right boundary, strains  $\Delta 5257$ ,  $\Delta 5259$  and  $\Delta 5261$ , which lack *orf10*, *orf11* and *orf13*, respectively, were assessed for PM100117/18 production by UPLC. As shown in Fig. 6b, deletion of these genes does not have any effect on PM100117/18 biosynthesis, confirming that they do not belong to the cluster. Therefore, the right cluster boundary seems to be defined by gene *gonCP*.

#### Generation and characterization of novel PM100117/18 derivatives

The main goal of this work was engineering novel PM100117/18 analogues with improved antitumor properties. We first sought to obtain structural analogues lacking the NQ moiety. For that purpose, a series of mutant strains affected in NQ biosynthesis putative genes (Fig. 5) was generated by disruption of *gonP8*, or individual deletion of *gonM4*, *gonMT*, *gonSL*, *gonS1* or *gonS2*. The resulting strains,  $gonP8^-$ ,  $\Delta gonM4$ ,  $\Delta gonMT$ ,  $\Delta gonSL$ ,  $\Delta gonS1$  and  $\Delta gonS2$ , were examined for PM100117/18 production by UPLC at 254 nm (Fig. 7a). These analyses did not detect PM100117/18 biosynthesis in  $gonP8^-$ ,  $\Delta gonMT$ ,  $\Delta gonSL$ ,  $\Delta gonS1$  and  $\Delta gonS2$ . Only in  $\Delta gonM4$  some traces of PM100117/18 were detected (Fig. 7a). To assess whether the reason of these changes in the PM100117/18 production level was the absence or decrease of NQ biosynthesis, we examined the accumulation of PM100117/18 intermediates lacking the NQ moiety in the mutant strains. It is important to note that loss of the NQ unit causes a change in the maximum absorption wavelength with respect to the parental compounds, shifting from 254 to 216 nm. UPLC and LC-MS analysis at 216 nm detected in the six mutant strains several compounds (Fig. 7b, triangles) with the expected absorption



spectra. In addition, two of these products (3 and 4) possessed molecular weights compatible with PM100117 [1, UPLC  $R_t = 5.182$  min,  $m/z$  1601.9 ( $M + Na$ )<sup>+</sup>] and PM100118 [2, UPLC  $R_t = 5.536$  min,  $m/z$  1585.9 ( $M + Na$ )<sup>+</sup>] biosynthetic intermediates lacking the NQ moiety. The chemical structures of compounds 3 [UPLC  $R_t = 4.14$  min,  $m/z$  1231.7 ( $M + Na$ )<sup>+</sup>] and 4 [UPLC  $R_t = 4.21$  min,  $m/z$  1215.7 ( $M + Na$ )<sup>+</sup>] were determined by NMR (Additional file 2: Figure S5), confirming that both products correspond to PM100118 analogues lacking the NQ moiety (Fig. 7c). Interestingly, compound 3 carries an additional hydroxyl group at the aglycone C18, which is not present in any of the parental products, suggesting that this derivative could belong to a PM100118 shunt biosynthetic pathway. Mutant complementation with the corresponding genes partially restored

PM100117/18 production (Additional file 1: Figure S2). These results confirm the involvement of genes *gonM4*, *gonMT*, *gonSL*, *gonS1*, *gonS2* and *gonP8* in the biosynthesis and/or transfer of the NQ unit to PM100117/18.

The *in vitro* antitumor activity of 3 and 4 was determined by measuring their  $GI_{50}$  (50 % inhibition on cell growth), TGI (total growth inhibition) and  $LC_{50}$  (50 % cell death) concentrations [47], against cancer cell lines A549 (human lung carcinoma cells), PSN1 (pancreas carcinoma), MDA-MB-231 (human breast adenocarcinoma) and HT29 (human colorectal carcinoma). The values of these three antitumor indicators for compounds 3 and 4 were remarkably higher than those of PM100117/18, indicating a weaker cytotoxic activity of the derivative compounds in comparison with the natural drugs (Table 2).

**Table 2 In vitro antitumor activity of compounds 1–6**

Compounds	A549 ( $\mu\text{M}$ )	HT29 ( $\mu\text{M}$ )	MDA-MB-231 ( $\mu\text{M}$ )	PSN1 ( $\mu\text{M}$ )
<b>1</b>				
PM100117				
GI <sub>50</sub>	1.52	3.04	2.66	nd
TGI	1.84	3.23	2.79	nd
LC <sub>50</sub>	2.22	3.61	2.97	nd
<b>2</b>				
PM100118				
GI <sub>50</sub>	2.24	1.92	1.73	nd
TGI	3.13	2.81	2.75	nd
LC <sub>50</sub>	4.28	4.09	4.16	nd
<b>3</b>				
from mutant <i>gonP8</i> <sup>-</sup>				
GI <sub>50</sub>	>8.40	>8.40	>8.40	>8.40
TGI	>8.40	>8.40	>8.40	>8.40
LC <sub>50</sub>	>8.40	>8.40	>8.40	>8.40
<b>4</b>				
from mutant <i>gonP8</i> <sup>-</sup>				
GI <sub>50</sub>	>8.40	>8.40	>8.40	>8.40
TGI	>8.40	>8.40	>8.40	>8.40
LC <sub>50</sub>	>8.40	>8.40	>8.40	>8.40
<b>5</b>				
from mutant $\Delta$ <i>gonCP</i>				
GI <sub>50</sub>	0.38	0.13	0.71	0.90
TGI	0.54	1.80	0.83	0.96
LC <sub>50</sub>	0.77	2.50	1.03	1.03
<b>6</b>				
from mutant $\Delta$ <i>gonCP</i>				
GI <sub>50</sub>	1.11	3.39	2.60	2.80
TGI	1.17	4.30	2.80	3.00
LC <sub>50</sub>	1.24	5.48	3.00	3.19

GI<sub>50</sub> compound concentration that produces 50 % inhibition on cell growth as compared to control cells, TGI compound concentration that produces total growth inhibition as compared to control cells, LC<sub>50</sub> compound concentration that produces 50 % cell death as compared to control cells, Nd values not determined

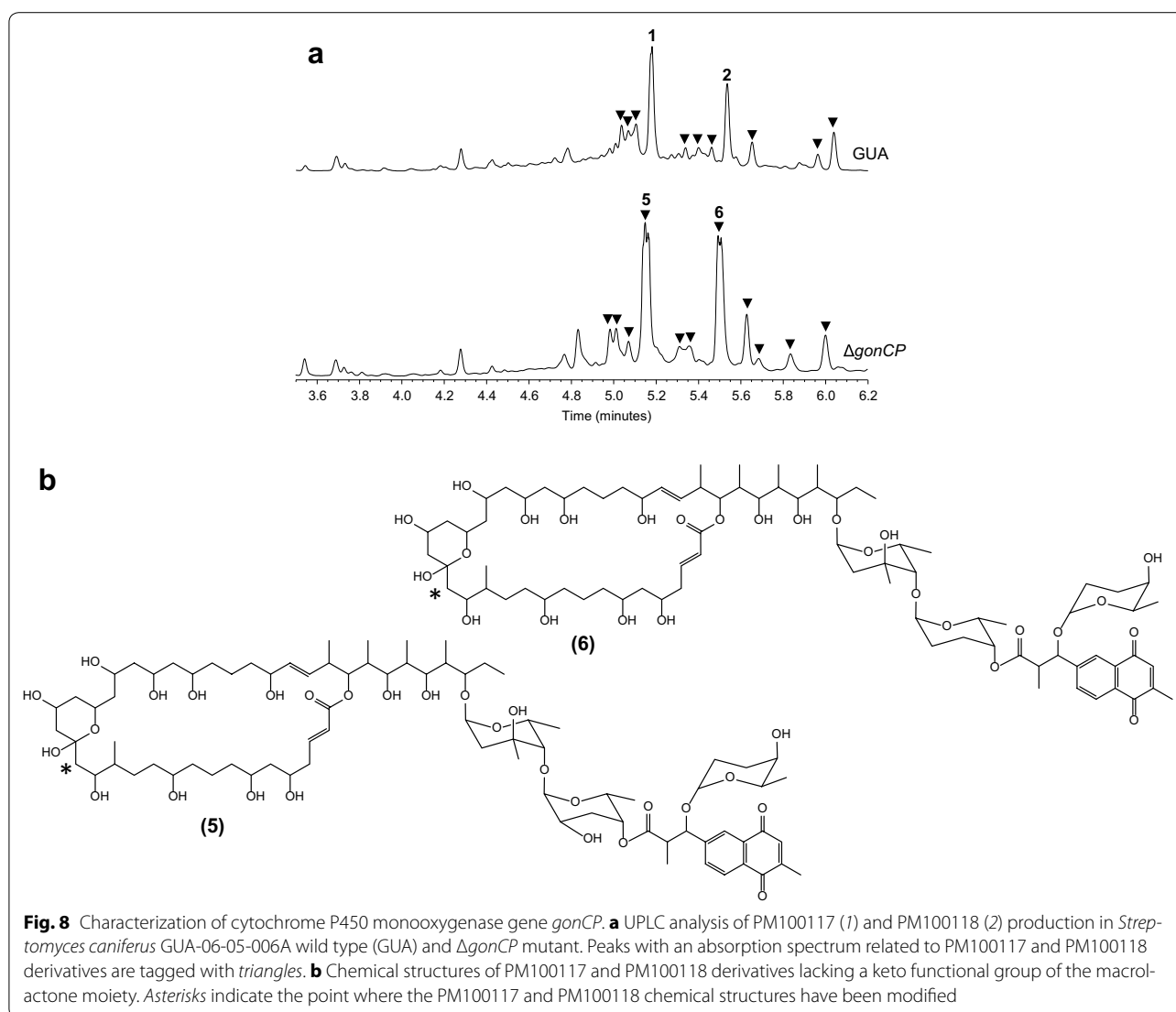
Given the cytotoxicity decay caused by the loss of the NQ moiety, we sought to perform a genetic manipulation leading to minor modifications of the PM100117/18 structure while preserving the NQ unit. With this aim, we deleted *gonCP*, which codes for a putative cytochrome P450 monooxygenase potentially involved in the oxygenation of the aglycone C16 and C17 (Fig. 3). UPLC analysis revealed the ability of the resulting mutant strain,  $\Delta$ *gonCP*, to produce several compounds (triangles, Fig. 8a) with absorption spectra related to PM100117/18. Furthermore, analysis of fermentation extracts from  $\Delta$ *gonCP* by LC-MS showed that two of these compounds, **5** [UPLC R<sub>t</sub> = 5.14 min, m/z 1587.9 (*M* + Na)<sup>+</sup>] and **6** [UPLC R<sub>t</sub> = 5.50 min, m/z 1571.9 (*M* + Na)<sup>+</sup>], possess molecular weights corresponding to PM100117 (**1**) and

PM100118 (**2**) analogues, respectively, lacking a keto group. Compounds **5** and **6** were then purified and analyzed by NMR to determine their chemical structures (Additional file 2: Figure S5), confirming that they derive from PM100117 and PM100118, respectively, by loss of the C16 aglycone keto functional group (Fig. 8b). This confirms the implication of gene *gonCP* in C16 oxygenation, but the question on the enzyme that catalyzes C17 hydroxylation still remains. The production of the natural compounds was restored when *gonCP* was re-introduced in the  $\Delta$ *gonCP* mutant (Additional file 1: Figure S2).

Likewise compounds **3** and **4**, the antitumor activity of derivatives **5** and **6** was examined against various cancer cell lines. Interestingly, compound **5** possesses an in vitro cytotoxicity threefold to fourfold higher than its corresponding parental product (**1**, PM100117) against A549 and MDA-MB-231 cell lines, as indicated by the lower GI<sub>50</sub>, TGI and LC<sub>50</sub> concentration values (Table 2). The GI<sub>50</sub> concentration of this compound also showed an outstanding 23-fold increase of antitumor activity, relative to PM100117, against HT29 tumor cells. However, this result is not concomitant with a similar decrease of the TGI and LC<sub>50</sub> concentrations, which keep values close to that of compound PM100117. Additional analysis of **5** by antibiotic activity assays against *Micrococcus luteus* and *Saccharomyces cerevisiae* further confirms an enhanced bioactivity of this derivative with respect to the parental products (Additional file 1: Figure S3). PM100118 analogue **6** showed a twofold to threefold improvement in the in vitro cytotoxic activity against the A549 cell line, as revealed by its lower GI<sub>50</sub>, TGI and LC<sub>50</sub> values relative to the natural compound (**2**, PM100118).

## Discussion

In this work, we have identified the PM100117 and PM100118 biosynthetic gene cluster, which has been characterized on the basis of bioinformatics analysis and genetic engineering data. The cluster spans a region of 169 kb and contains 41 genes coding for all the putative functions presumably required for PM100117/18 biosynthesis (Fig. 2; Table 1). From these activities, biosynthesis of PM100117/18 can be predicted as follows. Firstly, a type I multimodular PKS containing a loading module and 20 extension modules, catalyzes the 21 condensation reactions required for the biosynthesis of the aglycone polyketide chain (Fig. 3). The macrolatone moiety is then decorated with two 2,6-dideoxysugar moieties. Next, the NQ unit is transferred to the second deoxysugar and then glycosylated with the third 2,6-dideoxysugar. The presence in the cluster of four putative glycosyltransferase genes is consistent with the PM100117/18 glycosylation pattern. One of the four glycosyltransferases might transfer L-axenose to the first position of the glycosylation



profile. Two glycosyltransferases might be responsible for the transfer of L-2-deoxy-fucose (PM100117) or L-rhodinose (PM100118) to the second glycosylation position. Finally, the last glycosyltransferase would transfer L-rhodinose to the NQ moiety. Carbons 16 and 17 of the aglycone contain a keto and a hydroxyl group, respectively. These oxygenations could be introduced at any step of PM100117/18 biosynthesis, after completion of the polyketide biosynthesis, as a tailoring modification. We have shown that the putative cytochrome P450 monooxygenase GonCP is responsible, at least, for C16 oxygenation (Fig. 8). There is not a clear candidate for C17 oxygenation, being the closest *orf9*. Deletion of *orf9*, which codes for a putative dioxygenase and lies outside the PM100117/100118 cluster boundaries, does not show any effect on C17 hydroxylation (Fig. 6). However, the

involvement of *gonCP* acting as a multifunctional oxygenase in C17 modification cannot be excluded. The oxygenation of consecutive carbons at polyketide aglyca by multifunctional cytochrome P450 monooxygenases has been previously reported [48]. Taking this into consideration we cannot discard the possibility of *orf9* product, or any other oxygenase coding gene at *S. caniferus* GUA-06-05-006A chromosome, complementing the putative GonCP C17 oxygenation in  $\Delta gonCP$  mutant strain. Four genes code for putative transcriptional regulators belonging to the MarR (GonMR) and LuxR (GonL1, GonL2 and GonL3) family, indicating that PM100117/18 biosynthesis might be subjected to a tight transcriptional regulation. Genetic engineering data suggest that proteins GonMR and GonL1 act as putative transcriptional activators. This is an interesting finding because MarR transcriptional

activators have been rarely described in the literature [49]. However, other mechanisms can also be considered regarding the mode of action of GonMR, such as a dual role as activator-repressor [50] or co-activator.

The noticeable structural similarity of the NQ unit to MK induced us to envisage a connection between the biosynthesis of both compounds (Fig. 5). This hypothesis is further supported by the presence in the cluster of four genes (*gonM1*, *gonM2*, *gonM3* and *gonM4*), coding for proteins highly identical to those involved in the early steps of MK biosynthesis via futasoline [33–35]. Based on their assigned putative functions, GonM1 (MqnA), GonM2 (MqnC) and GonM4 (MqnD) could catalyze three of the four reaction steps leading to the biosynthesis of DH6N, which, according to our model, defines the branching point towards NQ and MK biosynthesis (Fig. 5). Even though C2-methylation of the NQ quinone ring could occur at any step of its biosynthesis, it is very plausible that this reaction delivers DH6N to the NQ branch. Deletion of *gonM4* diminishes PM100117/18 production, and leads to the accumulation of biosynthetic intermediates lacking the NQ unit (Fig. 7c). This result confirms the involvement of *gonM4* in NQ biosynthesis. However, in contrast to *gonP8*<sup>-</sup>,  $\Delta$ *gonMT*,  $\Delta$ *gonSL*,  $\Delta$ *gonS1* and  $\Delta$ *gonS2* mutant strains, in which PM100117/18 biosynthesis is completely abolished, in  $\Delta$ *gonM4* a low PM100117/18 production level can be detected (Fig. 7a). This indicates that the additional *mqnD* homologue, present outside the PM100117/18 gene cluster, is functional and able to partially complement the loss of *gonM4*. We do not fully understand the role of duplicated MqnA, MqnC and MqnD gene functions in the genome. However, we can speculate that the presence in the cluster of dedicated *mqnA*, *mqnC* and *mqnD* genes leading to DH6N biosynthesis could offer two main advantages. One is to provide sufficient supply of DH6N, which is probably a limiting intermediate as it is common to MK and NQ biosynthesis. The second advantage could be to facilitate a coordinated regulation with other genes involved in PM100117/18 biosynthesis, thus ensuring an optimal PM100117/18 production when these are required. To date, this is the first time that MK biosynthesis is described as a nexus between secondary and primary metabolism. Nonetheless, the existence of duplicated gene homologues in a single genome, one being part of the primary metabolism and the other present in a secondary metabolites gene cluster, has been frequently described for genes (*ccr*) coding for crotonyl-coA carboxylase/reductase. Multiple *ccr* homologues have been identified in gene clusters for the biosynthesis of various polyketide natural products [51–54]. As we have speculated for the presence of MK biosynthesis genes in the PM100117/18 gene cluster, the role of *ccr*

duplications has also been postulated to be the supply of sufficient precursor building blocks for polyketide biosynthesis [51].

The generation of PM100117/18 structural analogues with improved antitumor activity relative to their natural products, represents a substantial achievement of this work. These derivatives were accomplished by a genetic manipulation, *gonCP* deletion, leading to a minor modification of PM100117/18 structure, the loss of a keto functional group. The twofold to fourfold improvement of antitumor activity showed by these derivatives is an interesting finding because the generation of truncated products, or even the removal of minor structural elements of a natural compound, often affects bioactivity negatively [55]. For instance, deletion of a gene coding for a cytochrome P450 monooxygenase of the pimaricin gene cluster in *S. natalensis* conduces to the production of analogue 4,5-deepoxypimaricin, which differs from the natural compound in a single oxygenation and shows a diminished antibacterial activity relative to the parental product [56]. Similar results have been reported on the pikromycin analogues neopikromycin and narbomycin produced by *S. venezuelae*, which lack a single hydroxyl group at different positions of the polyene ring and possess a noticeably reduced antibacterial activity [57]. Instead, most of the reported genetic engineering approaches, leading to the generation of improved bioactive analogues, consist of the addition or replacement of structural components in the parental product [17, 58]. In this regard, cytochrome P450 monooxygenases have been frequently contemplated as promising targets for engineering the biosynthesis of novel therapeutic natural compounds. As an example, replacement of the C16 carboxyl on the nystatin analogue S44HP with a methyl group by mutation of a P450 monooxygenase gene, yielded a twofold more active antifungal analogue [59]. Herein, compounds 5 and 6 has been only assessed for in vitro cytotoxicity against different tumor cell lines. However, aside from an enhanced bioactivity, structural modifications frequently yield derivatives with other additional desirable pharmacological properties, such as lower toxicity or improved solubility [59–61]. In future works some of these properties of compounds 5 and 6 might also be addressed.

On the other hand, the loss of the NQ unit caused in vitro cytotoxicity decay, indicating that this moiety is central to PM100117/18 antitumor activity. Curiously, two compounds structurally related to PM100117/18, langkolide [26] and GT35 [27], which harbor a similar naphthoquinone moiety, also possess cytotoxic activity. Furthermore, the anti-proliferative effect of MKs on tumor cells both in vitro and in vivo has been repeatedly reported [62, 63]. Conversely, other macrolides

structurally resembling PM100117/18 but lacking a naphthoquinone unit, such as liposidolide A [64] and polaramycin [65], exhibit antifungal and antibacterial but not antitumor activity. All together, these observations suggest the idea that PM100117/18 cytotoxic activity could, to large extent, stems from the NQ moiety.

The availability of the PM100117 and PM100118 gene cluster, and the genetic insights into their biosynthesis, will help to understand how similar natural compounds are produced. This information enables the engineering of more derivatives with improved pharmacological properties such as an enhanced biological activity.

## Conclusions

PM100117 and PM100118 are members of a group of glycosylated compounds hallmarked by the presence in their structures of a NQ chromophore structurally resembling MK. Our results show that biosynthesis of the NQ chromophore is a complex process that involves diverse enzymes and which is connected to primary metabolism. The presence in secondary metabolite gene clusters of certain primary metabolism genes may be explained in terms of a sufficient supply of limiting intermediates. This connection of secondary metabolism with MK biosynthesis has never been reported before. A similar situation might be encountered in the future when other gene clusters for natural products structurally related to PM100117/18 will be characterized. On the other hand, the analysis of PM100117/18 analogues has shown interesting insights into the structure-bioactivity relationship of these family of natural products. Removal of the C16 keto group leads to an increased antitumor activity of both the PM100117 and PM100118 derivative. However, the overall cytotoxicity level showed by compound 5 is higher than the observed in compound 6. This suggests that PM100117 could be a more promising target to undertake other structural modifications. In addition, based on the results presented in this work and previously described data on similar compounds, the presence of NQ moieties might be a predictive structural feature of cytotoxic activity. This issue could be taken into account at future screening for novel antitumor natural products.

## Methods

### Bacterial strains, tumor cell lines, media and cultivation conditions

Unless otherwise indicated, media used in this work have been described in Kieser et al. [66]. The PM100117/18 producer strain *S. caniferus* GUA-06-05-006A [25] was routinely grown in MA medium (2.1 % morpholinepropanesulfonic acid, 0.5 % glucose, 0.05 % yeast extract, 0.05 % meat extract, 0.1 % casaminoacids, pH 7.0 adjusted with KOH). The *Escherichia coli* strains used as

host for cloning (DH10B) [67] and for intergeneric conjugation (ET12567/pUB307) [66] were grown in 2× TY medium supplemented, when required, with the appropriate antibiotic for plasmid selection. For metabolite production, a seed culture was grown in 50-ml falcon tubes containing 5 ml TSB medium on a rotary shaker at 250 rpm and 30 °C for 24 h. Then, 1.5 ml of the seed culture was used to inoculate 25 ml of SM medium (1 % glucose, 0.4 % yeast extract, 0.4 % peptone, 0.4 % K<sub>2</sub>HPO<sub>4</sub>, 0.2 % KH<sub>2</sub>PO<sub>4</sub>, 0.05 % MgSO<sub>4</sub>, pH 7.0 adjusted with KOH) supplemented with 10.3 % sucrose in a 250-ml shake flask. The culture was continued at 30 °C for 7 days with constant shaking at 200 rpm. Large-scale fermentations for compound purification were carried out for 9 days in a 2.5 L culture final volume inoculated (5 % v/v) with a seed culture grown for 72 h.

A-549 (ATCC CCL-185), lung carcinoma; HT-29 (ATCC HTB-38), colorectal carcinoma and MDA-MB-231 (ATCC HTB-26), breast adenocarcinoma cell lines were obtained from ATCC (<http://www.lgcstandards-atcc.org/>). Cell lines were maintained in RPMI medium 1640 (Gibco-RBL) supplemented with 10 % fetal calf serum (FCS), 2 mM L-glutamine and 100 U/mL penicillin and streptomycin, at 37 °C and 5 % CO<sub>2</sub>.

### Analysis of metabolite production and compound purification

Samples (3 ml) from *S. caniferus* GUA-06-05-006A whole-cultures (see above) were mixed with an equal volume of ethyl acetate and incubated at room temperature for 2 h. The organic phase was then recovered by centrifugation (3000×g, 10 min) and evaporated *in vacuo*. The residue was dissolved in methanol:DMSO (1:1) to perform UPLC and LC-MS analyses as described elsewhere [68].

For the purification of compounds 3, 4, 5 and 6, mycelia of the corresponding producing strains were separated from the culture by centrifugation and extracted twice with ethyl acetate. The supernatants were filtered and applied to a solid-phase extraction cartridge (Sep-Pak Vac C18, 10 g, Waters) that had been fitted with a perforated stopper pierced by a stainless steel HPLC tubing. The culture broth was applied by means of a peristaltic pump and subsequently the cartridge was connected to a HPLC quaternary pump (model 600E, Waters). The retained material was eluted with a mixture of methanol and 0.05 % trifluoroacetic acid (TFA) in water. A linear gradient from 0 to 100 % methanol in 60 min, at 10 ml/min, was used. Fractions were taken every 5 min, collected on 5 ml of 0.1 M phosphate buffer, pH 7.0 and analyzed by UPLC. Those fractions containing the desired compounds were evaporated *in vacuo* and subsequently re-dissolved in a small volume of a mixture of DMSO and

methanol (50:50). The organic extract of the culture pellets was similarly evaporated and re-dissolved. The compounds of interest were purified by preparative HPLC using a SunFire C18 column (10  $\mu$ m, 10  $\times$  250 mm, Waters). Compounds were chromatographed with mixtures of acetonitrile or methanol and 0.05 % TFA in water in isocratic conditions optimized for each peak, at 7 ml/min, and were always collected on 0.1 M phosphate buffer, pH 7.0. Compound **5** was purified with 55 % acetonitrile in a first step and with 82 % methanol in a second step. Compound **6** was purified with 55 % acetonitrile in a first step and with 85 % methanol in a second step. Compounds **3** and **4** were purified with 32 % acetonitrile in a first step and with 37 % acetonitrile in a second step. After every purification step, the collected compounds were diluted fourfold with water and then applied to a solid-phase extraction cartridge (Sep-Pak C18, Waters). The cartridge was washed with water, the retained compound was eluted with methanol and dried *in vacuo*. Once the purification was finished, the compounds were dissolved in a mixture of tert-butanol and water (1:1) and lyophilized.

#### In vitro cytotoxicity assay

Triplicate cultures were incubated for 72 h in the presence or absence of test compounds (at ten concentrations ranging from 10 to 0.0026 mg/mL). For quantitative estimation of cytotoxicity, the colorimetric sulforhodamine B (SRB) method was used [69]. Briefly, cells were washed twice with PBS, fixed for 15 min in 1 % glutaraldehyde solution, rinsed twice in PBS, and stained in 0.4 % SRB solution for 30 min at room temperature. Cells were then rinsed several times with 1 % acetic acid solution and air-dried. Sulforhodamine B was then extracted in 10 mM trizma base solution and the absorbance measured at 490 nm. Using the mean  $\pm$  SD of triplicate cultures, a dose-response curve was automatically generated using nonlinear regression analysis. Three reference parameters were calculated (NCI algorithm) by automatic interpolation: GI<sub>50</sub> = compound concentration that produces 50 % cell growth inhibition, as compared to control cultures; TGI = total cell growth inhibition (cytostatic effect), as compared to control cultures, and LC<sub>50</sub> = compound concentration that produces 50 % net cell killing (cytotoxic effect).

#### Mass spectra and structural elucidation

(+)-HRESIMS was performed on an Agilent 6230 time of flight LC/MS. NMR spectra were recorded on a Varian "Unity 500" spectrometer at 500/125 MHz (<sup>1</sup>H/<sup>13</sup>C). Chemical shifts were reported in ppm using residual CD<sub>3</sub>OD (d 3.31 for <sup>1</sup>H and 49.0 for <sup>13</sup>C) as internal reference. HMBC experiments were optimized for a <sup>3</sup>J<sub>CH</sub> of 8 Hz. ROESY spectra were measured with a mixing time

of 500 ms. The structures were established by <sup>1</sup>H- and <sup>13</sup>C-NMR and two dimensional NMR experiments correlation spectroscopy (COSY), heteronuclear multiple quantum coherence (HMQC), heteronuclear multiple-bond correlation (HMBC).

#### DNA manipulation and plasmids construction

The isolation and manipulation of DNA were carried out following standard general methods previously described for *E. coli* [70] and *Streptomyces* [66]. PCR amplifications were conducted by using Herculase II Fusion polymerase (Agilent Technologies) with a touchdown PCR procedure. The thermocycler (SureCycler 8800, Agilent Technologies) was programmed as follow: initial denaturation at 99.9 °C for 4 min; 20 cycles of 99.9 °C for 20 s, 65–45 °C touchdown for 20 s and 72 °C for tx (20 s/kb) min followed by 10 cycles of 99.9 °C for 20 s, 60 °C for 20 s and 72 °C for tx (20 s/kb) min. Final extension was performed at 72 °C for 3 min. PCR products of the expected size were gel-purified and sequenced.

A detailed description of the construction of plasmids used in this work can be found in Additional file 3: Methods S1. Plasmids for inactivation of genes *gonP1* and *gonP8* were constructed in the conjugative plasmid pOJ260 [71], which lacks the capacity to replicate in *Streptomyces* and carries the *aac(3)IV* gene marker that confers resistance to apramycin (Apm<sup>R</sup>). To achieve the single deletion of genes *gonM4*, *gonMT*, *gonSL*, *gonS1*, *gonS2*, *gonCP*, *gonMR*, *gonL1*, *orf9*, *orf10*, *orf11* and *orf13*, DNA sequences flanking the target genes were amplified with the primer pairs designated in Additional file 3: Table S1 and cloned at both sides of the *aac(3)IV* gene in plasmid pEFBA-oriT [72]. The hygromycin B resistance (Hyg<sup>R</sup>) gene marker, *hyg*, was then extracted from plasmid pLHyg [73] and introduced in the deletion plasmids (Additional file 3: Methods S1). Gene *hyg* allows recognizing clones in which a complete gene replacement by a double cross-over has taken place (Hyg<sup>S</sup> Apm<sup>R</sup>) from those in which a single cross-over event has integrated the deletion plasmid into the chromosome (Hyg<sup>R</sup> Apm<sup>R</sup>). A suitable plasmid backbone to accomplish the complementation of Apm<sup>R</sup> mutants was constructed as follows. The integrative plasmid pSETec [68], which harbors the constitutive *ermE*\**p* promoter, was digested with NcoI. A 1.6 Kb fragment containing *hyg* was extracted from pLHyg by NheI/SpeI digestion. The linearized pSETec plasmid and the pLHyg NheI-SpeI fragment were then blunt-ended with the Klenow fragment of DNA polymerase I and ligated to afford plasmid pSETHe. Complementation plasmids were generated by inserting the target genes into the XbaI/EcoRV sites of pSETHe, under the transcriptional control of the *ermE*\**p* promoter (Additional file 3: Methods S1).



### Gene mutation and complementation by intergeneric conjugation

Plasmids pOJ-*gonP1* and pOJ-*gonP8*, pD-*gonM4*, pD-*gonMT*, pD-*gonSL*, pD-*gonS1*, pD-*gonS2*, pD-*gonCP*, pD-*gonMR*, pD-*gonL1*, pD-*orf9*, pD-*orf10*, pD-*orf11* and pD-*orf13* (Additional file 3: Methods S1) were transferred into *S. caniferus* GUA-06-05-006A by intergeneric conjugation to generate the mutant strains *gonP1*<sup>-</sup>, *gonP8*<sup>-</sup>,  $\Delta$ *gonM4*,  $\Delta$ *gonMT*,  $\Delta$ *gonSL*,  $\Delta$ *gonS1*,  $\Delta$ *gonS2*,  $\Delta$ *gonCP*,  $\Delta$ *gonMR*,  $\Delta$ *gonL1*,  $\Delta$ 5201,  $\Delta$ 5257,  $\Delta$ 5259 and  $\Delta$ 5261, respectively. A standard conjugation procedure [66] with minor modifications was followed. Briefly, *S. caniferus* GUA-06-05-006A mycelia fragments stored at -20 °C from a SMS culture (5 ml) supplemented with 10 mM MgCl<sub>2</sub> were used as plasmid recipient. *E. coli* donor cells were grown to an absorbance of 0.6 at 600 nm in 2× TY media containing the appropriate antibiotics for plasmid selection and 10 mM MgCl<sub>2</sub>. Donor cells (5 × 10<sup>-8</sup>) were washed twice with 2× TY and mixed with mycelia in 2× TY to a final volume of 1 ml. The mating mixture was spread on mannitol-soya (MS) plates supplemented with 25 mM MgCl<sub>2</sub> and incubated for 20 h at 30 °C. Plates were then overlaid with 3 ml of SNA containing nalidixic acid (0.6 mg) and apramycin (1.2 mg), and further incubated for 7–10 days. Exconjugants potentially harboring a gene deletion were transferred to MA plates with and without hygromycin B (100 µg/ml) and cultivated for 2 days. After several rounds of non-selective growth, Hyg<sup>s</sup> recombinants were selected for gene replacement confirmation by PCR. Likewise, *gonP1*<sup>-</sup> and *gonP8*<sup>-</sup> exconjugants were refreshed in MA plates with nalidixic acid and apramycin and tested for correct gene inactivation by PCR (Additional file 3: Table S1). The resulting mutant strains were complemented with the conjugative plasmids pSEThe, pC-*gonP8*, pC-*gonM4*, pC-*gonMT*, pC-*gonSL*, pC-*gonS1*, pC-*gonS2*, pC-*gonCP*, pC-*gonMR* and pC-*gonL1* (Additional file 3: Methods S1), generating the strains *GUA*-pS, *CPgonP8*, *CPgonM4*, *CPgonMT*, *CPgonSL*, *CPgonS1*, *CPgonS2*, *CPgonCP*, *CPgonMR* and *CPgonL1*, respectively. Selection of exconjugants carrying the corresponding complementation plasmid was performed by overlying conjugation plates with 3 ml of SNA containing nalidixic acid (0.6 mg) and hygromycin (2 mg).

### Sequencing and bioinformatics analysis

The *S. caniferus* GUA-06-05-006A chromosome was sequenced at Lifesequencing Ltd., Valencia, Spain by Roche/454 pyrosequencing [74] on a Genome Sequencing FLX platform. The genome was assembled in the Newbler assembler package [75] version 2.8 by using default parameters. Identification of gene clusters for biosynthesis of secondary metabolites was performed by

the antibiotics and secondary metabolite analysis shell: antiSMASH 3.0.4 [22]. Annotation of ORFs within the PM100117/18 biosynthesis gene cluster was based on database searching of the corresponding proteins carried out by BLAST algorithm [76] at the National Center for Biotechnology Information (NCBI). Additional sequence alignments were conducted by ClustalW2 [77] and EMBOSS needle [78] from the European Molecular Biology Laboratory (EMBL). Prediction of transmembrane domains was performed by TMHMM Server v. 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>).

The nucleotide sequence of the PM100117/18 gene cluster was deposited in GeneBank (accession numbers: LN997801 and LN997802).

### Additional files

**Additional file 1: Figure S1.** Sequence alignment of the cluster PKS domains active sites. LD, loading domain; M1-M20, extension modules. Polyketide synthase domains are as follows: KS, ketosynthase; AT, acyltransferase; KR, ketoreductase; DH, dehydratase; ER, enoylreductase; ACP, acyl carrier protein; CAL, CoA-ligase. **Figure S2.** Genetic complementation of mutant strains. UPLC analysis of PM100117 (1) and PM100118 (2) production in strains *GUA*-pS, *CPgonP8*, *CPgonM4*, *CPgonMT*, *CPgonSL*, *CPgonS1*, *CPgonS2*, *CPgonCP*, *CPgonMR* and *CPgonL1*. **Figure S3.** Antibiotic activity test of compound 5. Diffusion disc assay against *Saccharomyces cerevisiae* and *Micrococcus luteus*. The length (cm) of the inhibition-growth halo is indicated by numbers and yellow lines. **Methods S2.** Antibiotic activity assay. Format: PDF.

**Additional file 2: Figure S4.** <sup>1</sup>H-NMR (CD<sub>3</sub>OD, 500 MHz) spectra of PM100117 and PM100118. **Figure S5.** Mass and NMR spectra of compounds 3, 4, 5 and 6. <sup>1</sup>H-NMR (CD<sub>3</sub>OD, 500 MHz), <sup>1</sup>H and MS spectra of compounds 3, 4, 5 and 6. <sup>13</sup>C, gCOSY, TOCSY, gHSQC, gHMBC and ROESY spectra of compounds 3, 4 and 5. Format: PDF.

**Additional file 3. Methods S1.** Plasmids construction. **Table S1.** Primers used in this work. Format: PDF.

### Abbreviations

2,6-DG: NDP-4-keto-2,6-dideoxy-D-glucose; 3,5-EPI: 3,5-epimerization; ABC: ATP-binding cassette transporters; ACP: acyl-carrier-protein; Aprm<sup>R</sup>: apramycin resistant phenotype; ARNT: aryl hydrocarbon receptor nuclear translocator protein; AT: acyltransferase; ATa: malonyl-CoA acyltransferase; ATP: methylmalonyl-CoA acyltransferase; ATP: adenosine triphosphate; C4-KR: C4-ketoreductase; CAL: CoA ligase; CoA: Coenzyme A; DH: dehydratase; DH6N: 1,4-dihydroxy 6-naphthoic acid; ER: enoylreductase; HPLC: high-performance liquid chromatography; HTH: helix-turn-helix; Hyg<sup>R</sup>: hygromycin B resistant phenotype; Hyg<sup>S</sup>: hygromycin B sensitive phenotype; I: identity; KR: ketoreductase; KS and KSQ: ketosynthase; KSIII: 3-oxoacyl-ACP synthase III; LAL: large ATP binding members of the LuxR family; LC: liquid chromatography; LD: loading domain; M: module; MK: menaquinone; MS: mass spectrometry; NMR: nuclear magnetic resonance; NQ: naphthoquinone; NRPS: non-ribosomal peptide synthetase; ORF: open reading frame; PAS: PER-ARNT-SIM homologous; PER: *Drosophila* period protein; PKS: polyketide synthases; Rt: retention time; S: similarity; SIM: *Drosophila* single-minded protein; TFA: trifluoroacetic acid; UPLC: ultra-performance liquid chromatography.

### Authors' contributions

JAS, FDC, CM and CO conceived and designed the project; RGS and CG conducted experiments and analyzed the data; RGS and CO performed sequence *in silico* analysis; AFB carried out compound purifications; RF performed structural elucidation of novel derivatives; RGS drafted the manuscript and CO, JAS and CM contributed to preparing the final version of the paper. All authors read and approved the final manuscript.

**Author details**

<sup>1</sup> Departamento de Biología Funcional e Instituto Universitario de Oncología del Principado de Asturias (I.U.O.P.A), Universidad de Oviedo, 33006 Oviedo, Asturias, Spain. <sup>2</sup> Drug Discovery Area, PharmaMar SA, Avda. de los Reyes 1, Colmenar Viejo, 28770 Madrid, Spain.

**Acknowledgements**

This work was supported by the Spanish Ministry of Economy and Competitiveness in the INNFACTO programme 2012–2015, with the project number IPT-2011-0752-900000 (Bioketido).

**Competing interests**

The authors declare that they have no competing interests.

Received: 10 November 2015 Accepted: 11 February 2016

Published online: 22 February 2016

**References**

- Ward AC, Bora N. Diversity and biogeography of marine actinobacteria. *Curr Opin Microbiol*. 2006;9:279–86.
- Goodfellow M, Fiedler HP. A guide to successful bioprospecting: informed by actinobacterial systematics. *Antonie Van Leeuwenhoek Int J Gen Mol Microbiol*. 2010;98:119–42.
- Olano C, Méndez C, Salas JA. Antitumor compounds from marine actinomycetes. *Mar Drugs*. 2009;7:210–48.
- Strohl W. Chapter 31: antimicrobials. In: *Microbial diversity and bioprospecting*. 2004. doi: [10.1128/9781555817770.ch31](https://doi.org/10.1128/9781555817770.ch31).
- Ambavane V, Tokdar P, Parab R, Sreekumar ES, Mahajan G, Mishra PD, et al. Caerulomycin A—An Antifungal Compound Isolated from Marine Actinomycetes. *Adv Microbiol*. 2014;04:567–78.
- Newman DJ, Cragg GM. Microbial antitumor drugs: natural products of microbial origin as anticancer agents. *Curr Opin Investig Drugs*. 2009;10:1280–96.
- Bull AT, Ward AC, Goodfellow M. Search and discovery strategies for biotechnology: the paradigm shift. *Microbiol Mol Biol Rev*. 2000;64:573–606.
- Baltz RH. Renaissance in antibacterial discovery from actinomycetes. *Curr Opin Pharmacol*. 2008;8:557–63.
- Donadio S, Maffioli S, Monciardini P, Sosio M, Jabes D. Antibiotic discovery in the twenty-first century: current trends and future perspectives. *J Antibiot (Tokyo)*. 2010;63:423–30.
- König GM, Kehraus S, Seibert SF, Abdel-Lateff A, Müller D. Natural products from marine organisms and their associated microbes. *ChemBioChem*. 2006;7:229–38.
- Charan RD, Schlingmann G, Janso J, Bernan V, Feng X, Carter GT. Diazepinomicin, a new antimicrobial alkaloid from a marine *Micromonospora* sp. *J Nat Prod*. 2004;67:1431–3.
- Perez Baz J, Cañedo LM, Fernández Puentes JL, SilvaElipse MV. Thiocoraline, a novel depsipeptide with antitumor activity produced by a marine *Micromonospora*. II. Physico-chemical properties and structure determination. *J Antibiot (Tokyo)*. 1997;50:738–41.
- Kortmansky J, Schwartz GK. Bryostatin-1: a novel PKC inhibitor in clinical development. *Cancer Invest*. 2003;21:924–36.
- Schofield MM, Jain S, Porat D, Dick GJ, Sherman DH. Identification and analysis of the bacterial endosymbiont specialized for production of the chemotherapeutic natural product ET-743. *Environ Microbiol*. 2015;17:3964–75.
- Floss HG. Combinatorial biosynthesis—potential and problems. *J Biotechnol*. 2006;124:242–57.
- Olano C, Lombó F, Méndez C, Salas JA. Improving production of bioactive secondary metabolites in actinomycetes by metabolic engineering. *Metab Eng*. 2008;10:281–92.
- Olano C, Méndez C, Salas JA. Post-PKS tailoring steps in natural product-producing actinomycetes from the perspective of combinatorial biosynthesis. *Nat Prod Rep*. 2010;27:571–616.
- Shin SC, Ahn DH, Kim SJ, Lee H, Oh T-J, Lee JE, et al. Advantages of single-molecule real-time sequencing in high-GC content genomes. *PLoS One*. 2013;8:e68824.
- Rutledge PJ, Challis GL. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat Rev Microbiol*. 2015;13:509–23.
- Starcevic A, Zucko J, Simunkovic J, Long PF, Cullum J, Hranueli D. ClustScan: an integrated program package for the semi-automatic annotation of modular biosynthetic gene clusters and in silico prediction of novel chemical structures. *Nucleic Acids Res*. 2008;36:6882–92.
- Li MHT, Ung PMU, Zajkowski J, Garneau-Tsodikova S, Sherman DH. Automated genome mining for natural products. *BMC Bioinformatics*. 2009;10:185.
- Weber T, Blin K, Duddela S, Krug D, Kim HU, Brucoleri R, et al. antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res*. 2015;43:W237–43.
- Medema MH, Kottmann R, Yilmaz P, Cummings M, Biggins JB, Blin K, et al. Minimum Information about a Biosynthetic Gene cluster. *Nat Chem Biol*. 2015;11:625–31.
- Medema MH, Fischbach MA. Computational approaches to natural product discovery. *Nat Chem Biol*. 2015;11:639–48.
- Perez M, Schleissner C, Fernandez R, Rodriguez P, Reyes F, Zuniga P, et al. PM100117 and PM100118, new antitumor macrolides produced by a marine *Streptomyces caniferus* GUA-06-05-006A. *J Antibiot*. 2015. (In press).
- Helaly SE, Kulik A, Zinecker H, Ramachandran K, Tan GYA, Imhoff JF, et al. Langkolide, a 32-membered macrolactone antibiotic produced by *Streptomyces* sp. acta 3062. *J Nat Prod*. 2012;75:1018–24.
- Takahashi I, Nishiie Y, Uosaki Y, Ochiai K. New substance GT35 and its production (Patent JPH09100290). 1997.
- Bisang C, Long PF, Cortés J, Westcott J, Crosby J, Matharu AL, et al. A chain initiation factor common to both modular and aromatic polyketide synthases. *Nature*. 1999;401:502–5.
- Bevitt DJ, Cortes J, Haydock SF, Leadlay PF. 6-Deoxyerythronolide-B synthase 2 from *Saccharopolyspora erythraea*. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme. *Eur J Biochem*. 1992;204:39–49.
- Beck BJ, Yoon YJ, Reynolds KA, Sherman DH. The hidden steps of domain skipping: macrolactone ring size determination in the pikromycin modular polyketide synthase. *Chem Biol*. 2002;9:575–83.
- Yadav G, Gokhale RS, Mohanty D. Computational Approach for Prediction of Domain Organization and Substrate Specificity of Modular Polyketide Synthases. *J Mol Biol*. 2003;328:335–63.
- Lombó F, Olano C, Salas JA, Méndez C. Chapter 11. Sugar biosynthesis and modification. *Methods Enzymol*. 2009;458(A):277–307.
- Hiratsuka T, Furihata K, Ishikawa J, Yamashita H, Itoh N, Seto H, et al. An Alternative Menaquinone Biosynthetic Pathway Operating in Microorganisms. *Science*. 2008;321:1670–3.
- Dairi T. An alternative menaquinone biosynthetic pathway operating in microorganisms: an attractive target for drug discovery to pathogenic *Helicobacter* and *Chlamydia* strains. *J Antibiot (Tokyo)*. 2009;62:347–52.
- Dairi T. Menaquinone Biosyntheses in Microorganisms. *Methods Enzymol*. 2012;515:107–22.
- Brooijmans R, Smit B, Santos F, van Riel J, de Vos WM, Hugenholtz J. Heme and menaquinone induced electron transport in lactic acid bacteria. *Microb Cell Fact*. 2009;8:e28.
- Barta ML, Thomas K, Yuan H, Lovell S, Battaile KP, Schramm VL, et al. Structural and biochemical characterization of *Chlamydia trachomatis* hypothetical protein CT263 supports that menaquinone synthesis occurs through the futasoline pathway. *J Biol Chem*. 2014;289:32214–29.
- Mahanta N, Fedoseyenko D, Dairi T, Begley TP. Menaquinone biosynthesis: formation of aminofutasoline requires a unique radical SAM enzyme. *J Am Chem Soc*. 2013;135:15318–21.
- Arora P, Goyal A, Natarajan VT, Rajakumara E, Verma P, Gupta R, Yousuf M, Trivedi OA, Mohanty D, Tyagi A, Sankaranarayanan R, Gokhale RS. Mechanistic and functional insights into fatty acid activation in *Mycobacterium tuberculosis*. *Nat Chem Biol*. 2009;5:166–73.
- Tang Y, Lee TS, Kobayashi S, Khosla C. Ketosynthases in the initiation and elongation modules of aromatic polyketide synthases have orthogonal acyl carrier protein specificity. *Biochemistry*. 2003;42:6588–95.
- Billigin T, Hyun C-G, Williams JS, Czisny AM, Thorson JS. The hedamycin locus implicates a novel aromatic PKS priming mechanism. *Chem Biol*. 2004;11:959–69.

42. Waldman AJ, Balskus EP. Lomaiviticin biosynthesis employs a new strategy for starter unit generation. *Org Lett*. 2014;16:640–3.
43. Guerra SM, Rodríguez-García A, Santos-Aberturas J, Vicente CM, Payero TD, Martín JF, et al. LAL regulators SCO0877 and SCO7173 as pleiotropic modulators of phosphate starvation response and actinorhodin biosynthesis in *Streptomyces coelicolor*. *PLoS One*. 2012;7:e31475.
44. Martín JF, Liras P. Cascades and networks of regulatory genes that control antibiotic biosynthesis. *Subcell Biochem*. 2012;64:115–38.
45. Antón N, Santos-Aberturas J, Mendes MV, Guerra SM, Martín JF, Aparicio JF, PimM, a PAS domain positive regulator of pimarinin biosynthesis in *Streptomyces natalensis*. *Microbiology*. 2007;153:3174–83.
46. Lubelski J, Konings WN, Driessen AJM. Distribution and physiology of ABC-type transporters contributing to multidrug resistance in bacteria. *Microbiol Mol Biol Rev*. 2007;71:463–76.
47. Skehan P, Storeng R, Scudiero D, Monks A, McMahon J, Vistica D, et al. New colorimetric cytotoxicity assay for anticancer-drug screening. *J Natl Cancer Inst*. 1990;82:1107–12.
48. Podust LM, Sherman DH. Diversity of P450 enzymes in the biosynthesis of natural products. *Nat Prod Rep*. 2012;29:1251–66.
49. Zhang Q, Chen Q, Zhuang S, Chen Z, Wen Y, Li J. A MarR Family Transcriptional Regulator, DptR3, Activates Daptomycin Biosynthesis and Morphological Differentiation in *Streptomyces roseosporus*. *Appl Environ Microbiol*. 2015;81:3753–65.
50. Oh S-Y, Shin J-H, Roe J-H. Dual role of OhrR as a repressor and an activator in response to organic hydroperoxides in *Streptomyces coelicolor*. *J Bacteriol*. 2007;189:6284–92.
51. Wilson MC, Moore BS. Beyond ethylmalonyl-CoA: the functional role of crotonyl-CoA carboxylase/reductase homologs in expanding polyketide diversity. *Nat Prod Rep*. 2012;29:72–86.
52. Cundliffe E, Bate N, Butler A, Fish S, Gandecha A, Merson-Davies L. The tylosin-biosynthetic genes of *Streptomyces fradiae*. *Antonie van Leeuwenhoek*. 2001;79:229–34.
53. Quade N, Huo L, Rachid S, Heinz DW, Müller R. Unusual carbon fixation gives rise to diverse polyketide extender units. *Nat Chem Biol*. 2011;8:117–24.
54. Karray F, Darbon E, Oestreicher N, Dominguez H, Tuphile K, Gagnat J, et al. Organization of the biosynthetic gene cluster for the macrolide antibiotic spiramycin in *Streptomyces ambofaciens*. *Microbiology*. 2007;153:4111–22.
55. Schlünzen F, Zarivach R, Harms J, Bashan A, Tocilj A, Albrecht R, et al. Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature*. 2001;413:814–21.
56. Mendes MV, Recio E, Fouces R, Luiten R, Martín JF, Aparicio JF. Engineered biosynthesis of novel polyenes: a pimarinin derivative produced by targeted gene disruption in *Streptomyces natalensis*. *Chem Biol*. 2001;8:635–44.
57. Lee SK, Park JW, Kim JW, Jung WS, Park SR, Choi CY, et al. Neopikromycin and novapikromycin from the pikromycin biosynthetic pathway of *Streptomyces venezuelae*. *J Nat Prod*. 2006;69:847–9.
58. Han AR, Shinde PB, Park JW, Cho J, Lee SR, Ban YH, et al. Engineered biosynthesis of glycosylated derivatives of narbomycin and evaluation of their antibacterial activities. *Appl Microbiol Biotechnol*. 2012;93:1147–56.
59. Brautaset T, Sletta H, Nedal A, Borgos SEF, Degnes KF, Bakke I, et al. Improved antifungal polyene macrolides via engineering of the nystatin biosynthetic genes in *Streptomyces noursei*. *Chem Biol*. 2008;15:1198–206.
60. Seco EM, Fotso S, Laatsch H, Malpartida F. A tailoring activity is responsible for generating polyene amide derivatives in *Streptomyces diastaticus* var. 108. *Chem Biol*. 2005;12:1093–101.
61. Carmody M, Murphy B, Byrne B, Power P, Rai D, Rawlings B, et al. Biosynthesis of amphotericin derivatives lacking exocyclic carboxyl groups. *J Biol Chem*. 2005;280:34420–6.
62. Samyukty A, Shetty AV, Dakshinamoorthy G, Kalyanasundaram R, Zheng G, Chen A, et al. Vitamin k2, a naturally occurring menaquinone, exerts therapeutic effects on both hormone-dependent and hormone-independent prostate cancer cells. *Evid Based Complement Alternat Med*. 2013;2013:287358.
63. Hitomi M, Yokoyama F, Kita Y, Nonomura T, Masaki T, Yoshiji H, et al. Antitumor effects of vitamins K1, K2 and K3 on hepatocellular carcinoma in vitro and in vivo. *Int J Oncol*. 2005;26:713–20.
64. Kihara T, Koshino H, Shin YC, Yamaguchi I, Isono K. Liposidolide A, a new antifungal macrolide antibiotic. *J Antibiot (Tokyo)*. 1995;48:1385–7.
65. Meng W, Jin WZ. [Structure determination of new antifungal antibiotics, polaramycins A and B]. *Yao Xue Xue Bao*. 1997;32:352–6.
66. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA. *Practical Streptomyces genetics*. Norwich: The John Innes Foundation; 2000.
67. Grant SGN, Jessee J, Bloom FR, Hanahan D. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc Natl Acad Sci USA*. 1990;87:4645–9.
68. Cano-Prieto C, García-Salcedo R, Sánchez-Hidalgo M, Braña AF, Fiedler HP, Méndez C, et al. Genome Mining of *Streptomyces* sp. Tü 6176: characterization of the nataxazole biosynthesis pathway. *ChemBioChem*. 2015;16:1461–73.
69. Papazisis KT, Geromichalos GD, Dimitriadis KA, Kortsaris AH. Optimization of the sulforhodamine B colorimetric assay. *J Immunol Methods*. 1997;208:151–8.
70. Sambrook JW, Russell D. *Molecular cloning: a laboratory manual*. 3rd ed. New York: Cold Spring Harbor Laboratory Press; 2001.
71. Bierman M, Logan R, O'Brien K, Seno ET, Rao RN, Schoner BE. Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene*. 1992;116:43–9.
72. Horna DH, Gómez C, Olano C, Palomino-Schätzlein M, Pineda-Lucena A, Carbajo RJ, et al. Biosynthesis of the RNA polymerase inhibitor streptolydigin in *Streptomyces lydicus*: tailoring modification of 3-methyl-aspartate. *J Bacteriol*. 2011;193:2647–51.
73. Olano C, Moss SJ, Braña AF, Sheridan RM, Math V, Weston AJ. Biosynthesis of the angiogenesis inhibitor borrelidin by *Streptomyces parvulus* Tü4055: insights into nitrile formation. *Mol Microbiol*. 2004;52:1745–56.
74. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 2005;437:376–80.
75. Miller JR, Koren S, Sutton G. Assembly algorithms for next-generation sequencing data. *Genomics*. 2010;95:315–27.
76. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25:3389–402.
77. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007;23:2947–8.
78. Li W, Cowley A, Uludag M, Gur T, McWilliam H, Squizzato S, et al. The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res*. 2015;43:W580–4.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
www.biomedcentral.com/submit

