



HAL
open science

Transcriptional regulation of congocidine (netropsin) biosynthesis and resistance

Audrey Vingadassalon, Florence Lorieux, Maud Juguet, Alba Noël, Luisa Santos, Laura Marin Fernandez, Jean-Luc Pernodet, Stéphanie Bury-Moné, Sylvie Lautru

► **To cite this version:**

Audrey Vingadassalon, Florence Lorieux, Maud Juguet, Alba Noël, Luisa Santos, et al.. Transcriptional regulation of congocidine (netropsin) biosynthesis and resistance. *Applied and Environmental Microbiology*, 2021, 87 (24), pp.e01380-21. 10.1128/AEM.01380-21 . hal-03787547

HAL Id: hal-03787547

<https://hal.science/hal-03787547>

Submitted on 25 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Transcriptional regulation of congocidine (netropsin)**
2 **biosynthesis and resistance.**

3

4 **Audrey Vingadassalon^{1,2}, Florence Lorieux¹, Maud Juguet^{1,3}, Alba Noël^{1,4}, Luisa D. F.**
5 **Santos¹, Laura Marin Fernandez^{1,5}, Jean-Luc Pernodet¹, Stéphanie Bury-Moné¹ and Sylvie**
6 **Lautru^{1*}**

7 ¹Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198,
8 Gif-sur-Yvette, France.

9 ² Current address: Université des Antilles, COVACHIM M2E (EA 3592), UFR SEN, Campus de
10 Fouillole, F-97 110 Pointe-à-Pitre, France.

11 ³ Current address: METabolic EXplorer S.A., Biopôle Clermont-Limagne, Saint-Beauzire,
12 France

13 ⁴ Current address: IRHS-UMR1345, Université d'Angers, INRAE, Institut Agro, SFR 4207
14 QuaSaV, 49071, Beaucouzé, France

15 ⁵ Current address: Unité Fonctionnalité et Ingénierie des Protéines (UFIP), Université de
16 Nantes, CNRS, UMR 6286, F-44000, Nantes, France.

17 *** Correspondence:**

18 Sylvie Lautru

19 sylvie.lautru@i2bc.paris-saclay.fr

20

21

22 **Keywords:** congocidine, pyrrolamide, antibiotic, regulation, *Streptomyces*, resistance.

23 **Running title:** Regulation of congocidine biosynthesis and resistance

24

25 **ABSTRACT**

26 The production of specialized metabolites by *Streptomyces* bacteria is usually temporally
27 regulated. This regulation is complex and frequently involves both global and pathway-
28 specific mechanisms. *Streptomyces ambofaciens* ATCC23877 produces several specialized
29 metabolites, including spiramycins, stambomycins, kinamycins and congocidine. The
30 production of the first three molecules has been shown to be controlled by one or several
31 cluster-situated transcriptional regulators. However, nothing is known regarding the
32 regulation of congocidine biosynthesis. Congocidine (netropsin) belongs to the family of
33 pyrrolamide metabolites, which also includes distamycin and anthelvencins. Most
34 pyrrolamides bind into the minor groove of DNA, specifically in A/T-rich regions, which gives
35 them numerous biological activities, such as antimicrobial and antitumoral activities. We
36 previously reported the characterization of the pyrrolamide biosynthetic gene clusters of
37 congocidine (*cgc*) in *S. ambofaciens* ATCC23877, distamycin (*dst*) in *Streptomyces netropsis*
38 DSM40846 and anthelvencins (*ant*) in *Streptomyces venezuelae* ATCC14583. The three gene
39 clusters contain a gene encoding a putative transcriptional regulator, *cgc1*, *dst1* and *ant1*
40 respectively. Cgc1, Dst1 and Ant1 present a high percentage of amino acid sequence
41 similarity. We demonstrate here that Cgc1, an atypical orphan response regulator, activates
42 the transcription of all *cgc* genes in the stationary phase of *S. ambofaciens* growth. We also
43 show that the *cgc* cluster is constituted of eight main transcriptional units. Finally, we show
44 that congocidine induces the expression of the transcriptional regulator Cgc1 and of the
45 operon containing the resistance genes (*cgc20* and *cgc21*, coding for an ABC transporter),
46 and propose a model for the transcriptional regulation of the *cgc* gene cluster.

47

48 **Importance**

49 Understanding the mechanisms of regulation of specialized metabolite production can have
50 important implications both at the level of specialized metabolism study (expression of silent
51 gene clusters) and the biotechnological level (increase of the production of a metabolite of
52 interest). We report here a study on the regulation of the biosynthesis of a metabolite from
53 the pyrrolamide family, congocidine. We show that congocidine biosynthesis and resistance
54 is controlled by Cgc1, a cluster-situated regulator. As the gene clusters directing the
55 biosynthesis of the pyrrolamides distamycin and anthelvencin encode a homolog of Cgc1,
56 our findings may be relevant for the biosynthesis of other pyrrolamides. In addition, our
57 results reveal a new type of feed-forward induction mechanism, in which congocidine
58 induces its own biosynthesis through the induction of the transcription of *cgc1*.

59 INTRODUCTION

60 The control of specialized metabolite production in *Streptomyces* is a complex process that
61 often involves multiple levels of regulation, including transcriptional and translational
62 regulations and precursor availability (1). At the transcriptional level only, a large diversity of
63 regulation mechanisms may be involved. The transcription of a specialized metabolite
64 biosynthetic gene cluster can be controlled by 'global' regulators, sensing and responding to
65 the nature of the nutrient source (for example DasR), or to nutrient limitation (GlnR for
66 nitrogen or PhoP for phosphate limitations) (1). Small signaling molecules (γ -butyrolactones,
67 furans, ppGpp) may also be required, participating to the regulatory cascade leading to the
68 activation or repression of cluster-situated regulators (CSRs, also called pathway-specific
69 regulators), encoded by genes located within biosynthetic gene clusters and generally
70 dedicated to their regulation (1, 2). In some cases, specialized metabolites themselves (*e.g.*
71 actinorhodin or simocyclinone D8) or their late biosynthetic intermediates (simocyclinone
72 C4) can also contribute to the control of the transcription of their biosynthetic genes,
73 through interactions with various transcription factors (2–5). Understanding the regulation
74 mechanisms that govern the biosynthesis of a metabolite have important implications. In
75 addition to participating to a better knowledge of the biology of the producer, it may help to
76 determine conditions favorable to this metabolite production, or give means to increase this
77 production (6). As specialized metabolites remain an important source of drugs, yield
78 improvement can be economically valuable.

79 We previously isolated and characterized the biosynthetic gene clusters of three pyrrolamide
80 antibiotics, congocidine (also called netropsin, Figure 1) in *Streptomyces ambofaciens*
81 ATCC23877, distamycin in *Streptomyces netropsis* DSM40846 and anthelvencins in

82 *Streptomyces venezuelae* ATCC14583 (7–9). Pyrrolamides are characterized by the presence
83 of 4-aminopyrrole-2-carboxylate groups linked by amide bonds. They exhibit a large range of
84 biological activities (antiviral, antimicrobial or anthelmintic activities), mainly due to their
85 ability to bind into the minor groove of the DNA (10). Indeed, congocidine and distamycin
86 have been extensively studied for the specificity of their binding to DNA. Physical studies
87 (e. g. X-ray structure determination and circular dichroism measurements) have shown that
88 they exhibit a marked preference for A/T rich DNA regions (11, 12). Many analogs and
89 hybrids of congocidine and distamycin have been synthesized chemically, in attempts to
90 modify or increase the sequence specificity of the DNA binding for gene expression
91 regulation applications (10), or to create new nontoxic antimicrobial molecules (13).

92

93 The regulation of pyrrolamide production by *Streptomyces* species has not been explored so
94 far. Our previous studies have revealed the presence of one gene, *cgc1* in the *S. ambofaciens*
95 congocidine gene cluster, *dst1* in the *S. netropsis* distamycin/congocidine/ disgocidine gene
96 cluster and *ant1* in the *S. venezuelae* anthelvencin gene cluster, coding for three
97 homologous putative transcriptional response regulators (7–9). This, combined with the
98 strikingly similar genetic organization of the three pyrrolamide gene clusters, suggests that
99 the transcriptional regulation of these three gene clusters could be similar.

100 In this study, we show that Cgc1, an atypical orphan response regulator, activates the
101 transcription of the congocidine biosynthetic and resistance genes during the stationary
102 phase in *S. ambofaciens*. We determine the transcriptional organization of the *cgc* gene
103 cluster. Finally, we show that congocidine induces the expression of *cgc1* encoding the

104 transcriptional regulator Cgc1 and of the operon containing the resistance genes, *cgc20* and
105 *cgc21*; and we propose a model for the transcriptional regulation of the *cgc* gene cluster.

106 **RESULTS**

107 **Cgc1 activates the transcription of the *cgc* genes.**

108 We analyzed the sequence of *S. ambofaciens* Cgc1 using BLAST (14) to search for
109 homologous protein and CD search (15) to look for conserved domains. These analyses
110 suggested that *cgc1* encodes a transcriptional regulator. This led us, in a previous study, to
111 propose that Cgc1 may control the transcription of the *cgc* genes (7). To examine this
112 hypothesis, *cgc1* was deleted in the SPM110 reference strain, yielding the CGCA021 strain. *S.*
113 *ambofaciens* SPM110 is a strain derived from *S. ambofaciens* ATCC23877 that produces
114 congocidine but does not produce spiramycins (16). The production of congocidine and the
115 transcription of the *cgc* genes were then analyzed. Comparative LC analyses of SPM110 and
116 CGCA021 culture supernatants (120 h of culture in MP5 production medium) showed that
117 production of congocidine is abolished in CGCA021 (Figure 2A). RT-PCR analyses were
118 carried out on RNA extracted after 43 h of culture in MP5 medium, a time at which
119 congocidine production has started. These analyses showed that all *cgc* genes are
120 transcribed in the SPM110 reference strain whereas transcription of all of them is abolished
121 in CGCA021 (Figure 2B). This suggested that Cgc1 acts as an activator of the transcription of
122 the *cgc* gene cluster.

123 To confirm this result, we genetically complemented CGCA021 (*cgc1* deletion mutant). We
124 cloned *cgc1* under the control of the strong *ermE** promoter in the pOSV554 integrative
125 vector and introduced the resulting pCGC081 plasmid in CGCA021 (yielding CGCA031 strain).
126 LC analysis showed that congocidine production is restored into CGCA031 (Figure 2A) and

127 RT-PCR analyses at 43 h confirmed that all *cgc* genes are expressed in the complemented
128 strain (Figure 2B). Altogether, these data demonstrate that Cgc1 is a pathway specific
129 activator of the transcription of the *cgc* genes in *S. ambofaciens*.

130 **Cgc1 is an atypical orphan response regulator belonging to the NarL family.**

131 BLAST and CD search analyses (14, 15) of the amino acid sequence of Cgc1 showed that this
132 protein belongs to the family of two-component response regulators. More specifically, Cgc1
133 belongs to the NarL family of transcriptional regulators, constituted of a N-terminal receiver
134 (REC) domain coupled with a C-terminal LuxR-type DNA-binding Helix Turn Helix (HTH)
135 domain (17). In bacterial genomes, genes encoding response regulators are typically found
136 clustered with genes coding for histidine kinases, with which they form the classical two-
137 component signal transduction systems (18). No histidine kinase is encoded within the *cgc*
138 gene cluster, indicating that Cgc1 is an orphan response regulator. Such regulators may rely
139 on kinases encoded elsewhere in the genome for their activation. However, some orphan
140 response regulators, called atypical regulators, do not rely on phosphorylation (and thus on
141 a histidine kinase) for their activation and their receiver domain lacks some of the conserved
142 amino acids of the phosphorylation pocket (19).

143 To determine whether Cgc1 is an atypical orphan response regulator, we aligned Cgc1 REC
144 domain with REC domains from either classical (SCO3008; SCO2281; SCO3818; SCO1654;
145 BldM and Aor1) or atypical (RedZ, Whil, JadR1, SCO3134; SCO4144; PapR6 and VmsT) orphan
146 response regulators from *Streptomyces* species and with the prototypical member of the
147 family, CheY (*Escherichia coli*) (Figure S1 in Supplemental material). REC domains from Dst1,
148 homolog of Cgc1 for the biosynthesis of congocidine, disgocidine and distamycin in *S.*
149 *netropsis* DSM40846 (8), and from Ant1, homolog of Cgc1 for the biosynthesis of

150 anthelvencins in *S. venezuelae* ATCC14583 (9) were also included. The alignment shows that
151 the residues constituting the phosphorylation pocket of REC domains, conserved in typical
152 response regulators, are absent in Cgc1, Dst1 and Ant1, as in other atypical orphan response
153 regulators (Figure S1). This strongly suggests that Cgc1 and its homologs encoded in other
154 pyrrolamide biosynthetic gene clusters are atypical orphan response regulators.

155 **The *cgc* gene cluster is constituted of eight transcriptional units.**

156 Having demonstrated that Cgc1 controls the transcription of the *cgc* genes, we established
157 the transcriptional organization of the *cgc* cluster. The 22 *cgc* genes are clustered on the
158 chromosome in a 29 kb region. Given their respective orientation (Figure 1), the minimal
159 number of transcriptional units is two. We undertook the study of the transcriptional
160 organization of the *cgc* gene cluster. The absence or small size of intergenic regions between
161 several genes of the *cgc* cluster means that the promoter of a gene is located in the coding
162 sequence of the upstream gene. For this reason, we chose the approach consisting in the
163 insertion of a transcriptional terminator as described by Dangel and colleagues (21). To
164 identify the transcriptional units, we replaced the first gene of a putative unit by a cassette
165 containing the apramycin resistance gene flanked with transcription terminators (Ω_{aac}) (20).
166 The transcription of the downstream genes was then studied by RT-PCR. However, this
167 strategy could not be used to determine whether *cgc1* is co-transcribed with *cgc2* and the
168 other downstream genes in the SPM110 strain, as Cgc1 controls the transcription of all *cgc*
169 genes. Similarly, replacing *cgc20* with the apramycin resistance cassette was problematic as
170 (i) *cgc20* codes for a subunit of the ABC transporter conferring congocidine resistance, and
171 (ii) export of congocidine by this ABC transporter is likely the only resistance mechanism of
172 *S. ambofaciens* against congocidine (7). Thus, deleting *cgc20* in SPM110 could result in a

173 congoicidine-sensitive strain still producing congoicidine (if *cgc22* is not co-transcribed with
174 *cgc20* and *cgc21*), a situation that would most likely be lethal. In these two cases, co-
175 transcription was investigated by amplifying various large regions encompassing these
176 adjacent genes by RT-PCR, using RNA extracted from the reference strain SPM110 as a
177 matrix. Total RNA extraction for the different strains was performed after 43 h of cultivation,
178 a time at which all the *cgc* genes are transcribed.

179 The transcriptional analysis of the *cgc* gene cluster is detailed in the Material and Methods
180 section and in Supplemental Figure S2. We failed to amplify the *cgc1-cgc2* intergenic region
181 using primers located in *cgc1* and *cgc2*, suggesting that *cgc1* and *cgc2* are not co-transcribed
182 (data not shown). We observed that the insertion of a transcriptional terminator into *cgc2*,
183 *cgc3*, *cgc6* and *cgc19* had no influence on the transcription of the downstream genes,
184 indicating that they are transcribed alone. On the contrary, insertion of the same terminator
185 into *cgc4* abolished the transcription of *cgc5* and insertion into *cgc7* abolished transcription
186 of all genes from *cgc8* to *cgc18*. We detected a transcript overlapping the three genes *cgc20*,
187 *cgc21* and *cgc22*. Altogether, our results indicate that the cluster is constituted of eight main
188 transcriptional units, as shown in Figure 1.

189 **The *cgc* gene cluster expression is induced during metabolic differentiation**

190 In a previous, more general study, we performed RNA-seq experiments with *S. ambofaciens*
191 grown for different time periods in MP5 medium (GEO accession number GSE162865) (22).
192 Focusing on the *cgc* gene cluster, we noticed that the current NCBI annotation of *S.*
193 *ambofaciens* ATCC 23877 genome (NZ_CP012382.1) incorrectly defines the *cgc1* start codon,
194 the most probable start codon being located 450 bp upstream the defined one. We
195 therefore conducted a new transcriptome analysis on the modified annotation. Comparing

196 the transcriptome of 24 h old cultures, when congocidine is not yet produced, with the
197 transcriptome of 48 h old cultures, at the onset of congocidine production, showed
198 important changes in gene expression (Figure 3A). About two thirds of genes are regulated in
199 a statistically significant manner (30 % being induced and 27 % repressed in stationary
200 phase). Remarkably, the *cgc* genes are among the most expressed and up-regulated genes in
201 stationary phase (Figure 3A). During exponential phase of growth (24 h), all *cgc* genes are
202 poorly transcribed, except *cgc1* (Figure 3B). On the contrary, during stationary phase, the
203 whole cluster is induced in a statistically significant manner. In both conditions, *cgc1* is the
204 most expressed gene of the *cgc* cluster. This emphasizes the importance of Cgc1 as a
205 regulator involved both, in the initial induction and in the maintenance of the *cgc* gene
206 cluster expression.

207 When the reads from the RNA-seq experiment performed after 48h of growth were mapped
208 on the DNA sequence of the *cgc* gene cluster, the results were in agreement with the
209 transcriptional organization determined by genetic approaches (Supplemental Figure S3). In
210 particular, *cgc1* and *cgc2* do not constitute an operon, confirming the results obtained
211 previously. Interestingly, the relative level of gene expression decreased from the beginning
212 to the end of an operon, so that *cgc22* (last gene in the *cgc20-cgc22* operon) and *cgc10* to
213 *cgc18* (at the end of the *cgc7-cgc18* operon) were expressed at a relatively low level
214 compared to the other genes of the congocidine cluster (Figure 3B). This result suggests that
215 the relative abundance of the Cgc proteins may greatly vary even if they are encoded by
216 genes from the same operon.

217 **Congocidine induces the expression of the congocidine resistance genes in *S. ambifaciens*.**

218 The *cgc20* and *cgc21* genes encode the subunits of an ABC transporter that is sufficient to
219 confer congocidine resistance to a *Streptomyces* sensitive strain (*Streptomyces lividans* TK23)
220 (7). Moreover, a *S. ambofaciens* strain deleted for the entire *cgc* gene cluster is as sensitive
221 to congocidine as *S. lividans* TK23, suggesting that no gene outside of the cluster is involved
222 in congocidine resistance (7). We observed in the present study that *cgc1* appears to control
223 the transcription of all *cgc* genes (Figure 2), yet the *cgc1* deletion mutant (CGCA021) is
224 resistant to congocidine at a level comparable to the one of the parental SPM110 strain,
225 while a strain deleted for the complete *cgc* gene cluster is sensitive to congocidine in the
226 same conditions (Figure S4 in supplemental material).

227 To explain the observed resistance of CGCA021 ($\Delta cgc1$) to congocidine, we hypothesized
228 that the resistance genes *cgc20* and *cgc21* might be expressed independently of Cgc1 in the
229 presence of congocidine. To test this hypothesis, we cultivated CGCA021 in the presence of
230 5 $\mu\text{g/ml}$ (11.6 μM) congocidine in MP5 medium. Total RNAs were extracted at 43 h and the
231 expression of *cgc20* and *cgc21*, together with the expression of a few biosynthetic genes
232 (*cgc22*, co-transcribed with *cgc20* and *cgc21*, *cgc2*, *cgc3* and *cgc19*), were determined by RT-
233 PCR (Figure 4).

234 While in the absence of congocidine, no *cgc* gene is transcribed in CGCA021 (Figure 2B), in
235 the presence of exogenous congocidine, transcription of *cgc20*, *cgc21*, *cgc22* (encoding a
236 acyl-CoA synthetase) and, to a lesser extent, of *cgc2* is observed. This indicates that the
237 presence of congocidine triggers the transcription of the *cgc20-cgc22* operon, but not the
238 transcription of other biosynthetic genes (with the possible exception of *cgc2*), by a
239 mechanism independent of Cgc1. As Cgc1 is the only transcriptional regulator encoded with
240 the *cgc* gene cluster, the transcription of *cgc20* and *cgc21* in the absence of *cgc1* and

241 presence of congoicidine could involve a transcriptional regulator gene located elsewhere in
242 the genome. In this case, the expression of the resistance genes may not be induced by
243 congoicidine in a heterologous host.

244 **Congoicidine induces the expression of the congoicidine resistance genes in the**
245 **heterologous host *S. lividans* TK23.**

246 To test whether the induction of the expression of the *cgc20-21* resistance genes would also
247 occur in a heterologous context, we constructed a *S. lividans* strain expressing the *cgc* gene
248 cluster except for *cgc1*. By PCR targeting, we replaced *cgc1* by an *att2aac* cassette in the BAC
249 (pCGC002) that we previously used for the heterologous expression of the *cgc* gene cluster
250 (7). After excision of the *att2aac* cassette, the resulting BAC (pCGC313), lacking *cgc1*, was
251 introduced in *S. lividans* TK23 by intergeneric conjugation and the resulting strain was
252 named CGCL095. The strain CGCL095 was cultivated in the presence or in the absence of
253 11.6 μ M congoicidine in MP5 medium, respectively. Total RNAs were extracted at 43 h and
254 the expression of *cgc20* and *cgc21*, together with the expression of a few biosynthetic genes
255 (*cgc22*, co-transcribed with *cgc20* and *cgc21*, and *cgc2*, *cgc3*, *cgc4*, *cgc7* and *cgc19*), were
256 determined by RT-qPCR (Figure 5).

257 Addition of 11.6 μ M congoicidine in the culture medium resulted in a statistically significant
258 increase of the expression of the genes from the *cgc20-cgc22* operon (between 2.8 and 3.9
259 fold, Figure 5). The expression of the biosynthetic genes tested did not change in a
260 statistically significant manner (*cgc2*, *cgc3* and *cgc19*) or decreased slightly (*cgc4*, *cgc6* and
261 *cgc7*) (Figure 5). Thus, in a congoicidine non-producing host, as already observed in the
262 natural producer, *S. ambifaciens*, the presence of congoicidine induces the expression of the

263 *cgc20-cgc22* operon comprising the resistance genes, suggesting that congocidine may exert
264 a direct effect on the promoter of these genes.

265 **Exogenous congocidine only affects the transcription of a subset of *cgc* genes in *S.***
266 ***ambofaciens*.**

267 We next studied the effect of congocidine on *S. ambofaciens* ATCC23877 in absence of
268 endogenous production. For this purpose, we grew *S. ambofaciens* ATCC23877 in YEME10
269 medium, a medium in which limited production of congocidine is detected in culture
270 supernatants after metabolic differentiation (22). The strain was grown for 24 h in medium
271 supplemented or not with 2.3 μ M (1 μ g/ml) congocidine, respectively. We used this
272 concentration of congocidine as it alters the growth of a *Streptomyces*-sensitive strain while
273 having no impact on the growth of the native producer (22).

274 The addition of congocidine had very little impact on gene expression levels, which were
275 modified in a statistically significant manner in the presence of congocidine for six genes
276 only (Figure 6, Supplemental Dataset 1). Among them, only one gene, SAM23877_RS20590
277 (encoding a hypothetical protein), is not located within the *cgc* cluster. Its expression is only
278 slightly repressed in presence of congocidine (0.49-fold, adjusted *p* value 0.0042). The
279 largest changes in gene expression, however, were observed in the *cgc* gene cluster. The
280 expression of *cgc20*, *cgc21* and *cgc22*, for which an induction in presence of congocidine had
281 already been observed, was strongly induced (50.3-fold, 57.1-fold and 40.5-fold induction,
282 respectively). In addition, an induction of the expression was observed for *cgc1* (9-fold
283 induction) and, at a lower level, for *cgc2* (3.4-fold induction). Thus, at low concentration (less
284 than 1% of the concentration usually reached in culture supernatants), congocidine appears

285 to exert a very specific effect on the *cgc20-cgc1* intergenic region that contains the
286 promoters for the transcription of *cgc1* and of the *cgc20-cgc22* operon.

287 **Congocidine increases transcription driven by the *cgc1* and *cgc20* promoters, in a**
288 **heterologous host.**

289 To further investigate the effect of congocidine on the promoters of *cgc1* and *cgc20*, we
290 cloned the *cgc20-cgc1* intergenic region in both orientations (P_{cgc20} and P_{cgc1}) upstream of
291 *gusA* (see the Materials and Methods section for details). The *gusA* gene codes for a β -
292 glucuronidase and is commonly used as a reporter gene (23). The resulting plasmids
293 (pCGC310 and pCGC311, Table 1), together with a control one not containing any promoter
294 region upstream of the *gusA* reporter gene (pCGC309), were introduced by intergeneric
295 conjugation into the *S. lividans* TK23 heterologous host. The resulting exconjugants were
296 verified by PCR (see the Materials and Methods section for details). The reporter strains
297 were first cultivated at 28°C in 10 ml of TSB medium for 24h, a time at which 0, 1 $\mu\text{g/ml}$
298 (2.3 μM) or 5 $\mu\text{g/ml}$ (11.6 μM) congocidine was added to the cultures. Cultures were
299 continued for five hours and cells were harvested and lysed. GusA activity in the crude
300 soluble protein fraction was measured spectrophotometrically for each strain using p-
301 nitrophenyl- β -D-glucuronide as substrate. The results from four independent experiments are
302 presented in Figure 7. In the absence of congocidine, a low level of GusA activity is observed
303 when *gusA* is under the control of P_{cgc1} whereas no activity is observed when *gusA* is under
304 the control of P_{cgc20} . Addition of 2.3 μM or 11.6 μM congocidine to the culture medium leads
305 to a significant increase of GusA activity (about 11.5-fold when *gusA* expression is under the
306 control of P_{cgc1} , between 12- and 68-fold when *gusA* expression is under the control of
307 P_{cgc20}). Thus, in a congocidine non-producing host, congocidine acts as an inducer of the

308 expression of genes under the control of both the P_{cgc20} and P_{cgc1} promoters. A difference of
309 GusA activity between the strain expressing *gusA* under the control of P_{cgc1} and the one
310 expressing *gusA* under the control of P_{cgc20} is observed. This difference may be due to
311 various factors, including a differential effect of congocidine on the two promoters or
312 different levels of GusA protein expression due to different Ribosome Binding Sites (RBS).
313 Nevertheless, RNA-seq analyses (Figures 3 and 6) also show similar trends in basal levels and
314 expression induction for *cgc1* and *cgc20*. This suggests more likely a difference in the
315 strength of the divergent promoters that control the expression of these genes.

316 The *cgc20/cgc21* genes code for the ABC transporter responsible for the export of and
317 resistance to congocidine. Therefore, we introduced the plasmids pCGC309, pCGC310 and
318 pCGC311 described above in *S. lividans* strains expressing the *cgc20/cgc21* genes under the
319 control of P_{cgc20} . Similar to what was observed in strains not harboring the *cgc20/cgc21*
320 genes, addition of 2.3 μ M and 11.6 μ M congocidine resulted in an increase of GusA activity
321 (Supplemental Figure S5). However, the levels of GusA activity were lower than those
322 obtained from strains not expressing the *cgc20/cgc21* genes. This reduced activity reflects a
323 weaker transcription from the P_{cgc20} and P_{cgc1} promoters, consistent with a reduction in the
324 intracellular concentration of congocidine, under the action of the ABC transporter Cgc20-
325 Cgc21.

326 **The *cgc20-cgc1* intergenic region is rich in strong congocidine binding sites.**

327 Congocidine is a minor groove DNA binder that binds at A/T-rich regions, more precisely a
328 succession of four or more A/T bases (24). Its binding into DNA minor groove is known to
329 perturb DNA topology locally, narrowing the minor groove, bending the DNA or introducing
330 negative supercoiling, in a sequence-dependent manner (25, 26). Thus, we considered the

331 possibility that the transcription of *cgc20-cgc22* and *cgc1* in presence of congocidine could
332 be a direct consequence of congocidine DNA binding in the *cgc20-cgc1* intergenic region.
333 This region (740 nt) contains eleven sequences of four or more consecutive A/T bases (five of
334 four A/T bases, four of five A/T bases, one of six A/T bases and one of nine A/T bases).
335 Studies have shown that the strength of congocidine binding to (A/T)₄ motifs depends on the
336 sequence of the motif. In particular, 5'-AAAA-3', 5'-TTTT-3' and 5'-AATT-3' have been
337 identified as strong congocidine binding motifs (27, 28). On the contrary, motifs that contain
338 the dinucleotide 5'-TA-3' sequence have significantly lower binding constants. There are
339 eight non-overlapping strong congocidine binding motifs in the *cgc20-cgc1* intergenic region
340 (Figure S6 in Supplemental Materials). When we searched for non-overlapping strong
341 congocidine binding motifs in all *S. ambifaciens* ATCC23877 intergenic regions of more than
342 4 bp, we found two intergenic regions, both duplicated as they are located in the terminal
343 inverted repeats, much larger than the *cgc20-cgc1* intergenic region (7300 nt and 2268 nt vs
344 740 nt) that contain more motifs than the *cgc20-cgc1* intergenic region (Figure S7 in
345 Supplemental Materials). The occurrence of multiple strong congocidine binding motifs in
346 the *cgc20-cgc1* bidirectional promoter region may therefore explain the effect of
347 congocidine on the transcription of the *cgc20-cgc22* and *cgc1* genes in the absence of Cgc1,
348 possibly by modifying the topology of the DNA.

349 **The expression of the *cgc20* and *cgc21* resistance genes is also controlled by Cgc1.**

350 The induction of the expression of the *cgc20-cgc22* operon by congocidine raised the
351 question of the control of *cgc20-cgc22* expression by Cgc1. Indeed, in the absence of *cgc1*,
352 the lack of expression of *cgc20-cgc22* could be due an absence of congocidine. However, we
353 noticed that the *E. coli* ET12567/pUZ8002 strain used for intergeneric conjugation became

354 resistant to 2 µg/ml (4.6 µM) congocidine after the introduction of the BAC pCGC002
355 containing the complete *cgc* gene cluster, while not producing congocidine (data not
356 shown). This prompted us to compare the growth of three *E. coli* strains: ET12567/pUZ8002
357 (control strain), ET12567/pUZ8002/pCGC0002 containing the complete *cgc* cluster and
358 ET12567/pUZ8002/pCGC313 containing the *cgc* gene cluster deleted for *cgc1*. Results
359 presented in Figure 8 show that the strain containing the complete *cgc* gene cluster is much
360 more resistant to congocidine than the control strain, showing that *cgc20-cgc21* confer
361 congocidine resistance to *E. coli*. The strain ET12567/pUZ8002/pCGC313 containing the
362 cluster deleted for *cgc1* presents an intermediate level of congocidine resistance. Thus, the
363 induction of the expression of the *cgc20-cgc21* genes by congocidine observed in
364 *Streptomyces* also occurs in *E. coli* and confers congocidine resistance at low concentrations
365 (5 µg/ml, 11.6 µM). However, Cgc1 is required to obtain higher levels of resistance (10
366 µg/ml, 23.2 µM) in *E. coli*. This demonstrates that Cgc1 also controls the expression of the
367 *cgc20-cgc21* resistance genes.

368 **DISCUSSION**

369 In this study, we aimed at characterizing the transcriptional regulation of the genes
370 constituting the congocidine biosynthetic gene cluster. The sequence analysis of Cgc1, the
371 predicted transcriptional regulator encoded within the *cgc* gene cluster, indicated that Cgc1
372 belongs to the NarL family of response regulators (RR). Cgc1 is an orphan response regulator,
373 as *cgc1* is not clustered with a gene encoding a histidine kinase. In addition, the receiver
374 domain of Cgc1 lacks most of the residues constituting the phosphorylation pocket. In
375 particular, the phosphorylatable aspartate residue is replaced by a histidine. Thus, in
376 addition to being an orphan RR, Cgc1 is also an atypical RR. Several of such regulators have

377 been studied and characterized, notably in *Streptomyces* species, for which some of them
378 are involved in specialized metabolism regulation (29). Thus, JadR1, the jadomycin pathway-
379 specific transcriptional activator, and RedZ, the undecylprodiginin pathway-specific
380 transcriptional activator, lack most of the conserved amino acids of the phosphorylation
381 pocket (30, 31). It has notably been reported that JadR1 is not phosphorylated when
382 incubated in presence of small phosphodonor molecule such as phosphoramidate (31).

383 The deletion of *cgc1* in the reference strain *S. ambofaciens* SPM110 (not producing
384 spiramycin) and the genetic complementation of the mutant strain carried out in this study
385 demonstrate that similarly to JadR1 and RedZ, Cgc1 activates the transcription of the *cgc*
386 genes. Determination of the transcriptional organization of the *cgc* gene cluster showed that
387 this cluster is constituted of eight main transcriptional units. Upstream of four of these
388 transcriptional units, the intergenic regions are extremely small or inexistent (upstream of
389 *cgc4*, *cgc6*, *cgc7* and *cgc19*), suggesting that in these situations the promoter and the Cgc1
390 operator sites are located within the preceding coding sequence. Using MEME (32) to search
391 for potential motifs for Cgc1 DNA binding did not allow to identify any convincing motif. It is
392 possible that as observed for JadR1, the Cgc1 recognition sequence is poorly conserved (31).

393 Analysis of the transcriptome of *S. ambofaciens* ATCC23877 carried out in the congocidine
394 production medium MP5 showed that during the exponential growth phase, *cgc1* is the only
395 *cgc* gene expressed, whereas all *cgc* genes are induced in a statistically significant manner in
396 stationary phase. The presence of transcripts of *cgc1* but not of biosynthetic/resistance
397 genes in exponential phase raised the question of the existence of another level of
398 regulation, possibly at the translational level. Indeed, the *cgc1* sequence contains two TTA
399 codons separated by two codons. TTA codons are rarely used in *Streptomyces*, which have

400 GC-rich genomes (around 72%), but they are often found in genes encoding cluster-situated
401 transcriptional activators (33, 34). In several *Streptomyces* species, it has been proposed that
402 the presence of a TTA codon in a gene could constitute a temporal regulation mechanism, as
403 *bldA* (SAM23877_RS14785), the only tRNA that recognizes the UUA codon, is more abundant
404 during the stationary phase than during the growth phase (33, 34). In *S. ambofaciens*,
405 however, the abundance of *bldA* transcript is constantly decreasing during metabolic
406 differentiation (transcriptomes analyzed in MP5 and YEME10 media at 24h, 30h, 36h, 48h
407 and 72h (see Liroy *et al.* (22), Table S5), indicating that *bldA* transcript abundance cannot
408 explain the delay between the transcription of *cgc1* and that of the other *cgc* genes. Yet, the
409 *bldA* tRNA may still be involved in some translational regulation of Cgc1 production as post-
410 transcriptional modifications of tRNAs are often needed for them to be fully functional (35,
411 36).

412 During the course of our study, we noticed that a *cgc1* deletion mutant is resistant to
413 congocidine, while a strain deleted of the entire *cgc* gene cluster is as sensitive to
414 congocidine than a congocidine non-producing *Streptomyces* strain (7). Analysis of the
415 expression of *cgc* genes in the *cgc1* deletion mutant showed that the *cgc20-cgc22* operon is
416 expressed in presence of congocidine, even in the absence of Cgc1. This operon contains the
417 genes coding for the ABC transporter responsible for resistance to congocidine (7).

418 Repeating this experiment in a heterologous host context (*S. lividans* TK23) yielded the same
419 observation, the induction of the expression of the *cgc20-cgc22* operon in presence of
420 congocidine and absence of Cgc1. Furthermore, the analysis of *S. ambofaciens* ATCC23877
421 transcriptome in the presence or absence of low levels of congocidine (1 µg/ml, 2.3 µM) in a
422 condition of limited congocidine production showed a modulation of the expression of only

423 six genes, all but one related to the *cgc* gene cluster (Figure 6). The expression of the *cgc20-*
424 *cgc22* operon and to small extent of *cgc2*, already observed in the *cgc1* deletion mutant
425 (Figure 4) was induced. In addition, the induction of the expression of the transcriptional
426 activator Cgc1 was observed. Altogether, this data suggests that congocidine exert a very
427 specific effect on the *cgc20-cgc1* intergenic region, which contains the *cgc1* and *cgc20*
428 promoters. To further explore this effect, we expressed the *gusA* reporter gene under the
429 control of the P_{*cgc1*} and P_{*cgc20*} promoters, in presence or absence of the resistance genes
430 *cgc20-cgc21*. In both cases, an increase of GusA activity was observed in presence of
431 congocidine, albeit at a lower level when the *cgc20-cgc21* resistance genes were present
432 (Figure 7, Figure S5). This confirmed the capacity of congocidine to induce the expression of
433 genes under the control of the P_{*cgc1*} and P_{*cgc20*} promoters. The lower GusA activity observed
434 in the presence of *cgc20-cgc21* may be due to the partial exclusion of congocidine from the
435 cells, comforting their proposed role in congocidine efflux (7).

436 Induction of resistance mechanisms by the product of a gene cluster is not unprecedented
437 (3, 37–39). For example, in the actinorhodin gene cluster, the TetR family member ActR
438 controls the expression of the actinorhodin export pumps ActA/ActB. The repression of
439 *actA/actB* expression by ActR is relieved by the binding of actinorhodin or some of its late
440 biosynthetic intermediates to ActR (39). Similarly, the repression of the expression of the
441 SimX efflux pump by SimR (TetR family) is lifted by the binding of simocyclinone D8 and its
442 biosynthetic intermediate simocyclinone C4 (3). In these two cases, ActR and SimR
443 specifically control the expression of the resistance genes, while other(s) transcriptional
444 regulator(s) control the transcription of the biosynthetic genes (ActII-orf4 in the actinorhodin
445 cluster (4) and SimReg1 in the simocyclinone D8 cluster (5)). In the *cgc* gene cluster, Cgc1 is

446 the sole transcriptional regulator encoded within the cluster. The experiments presented in
447 this study show that it controls both, the expression of the biosynthetic as well as resistance
448 genes. In contrast to what has been reported for actinorhodin and simocyclinone D8, the
449 induction of the congocidine resistance genes does not seem to rely exclusively on a
450 transcriptional regulator, and may involve a direct binding of congocidine to the P_{cgc20}
451 promoter. Indeed, the induction of the resistance genes by congocidine occurs in the
452 absence of *cgc1* and in heterologous hosts as distantly related as *E. coli*. Furthermore, the
453 *cgc20-cgc1* intergenic region, which contains the promoters for the *cgc20-cgc22* operon and
454 for *cgc1*, is rich in congocidine strong binding sites (Figure S6). Thus, the binding of
455 congocidine into the DNA minor groove could perturb DNA topology locally and favor the
456 transcription of *cgc20-cgc22* and of *cgc1*. In fact, it has been shown that congocidine can
457 activate transcription initiation in *E. coli* from promoters that have been engineered to
458 include a congocidine binding site in the spacer sequence between the -35 and -10 regions
459 (40). Interestingly, to our knowledge, only one other DNA-binding metabolite, daunorubicin
460 (intercalating agent, binding to GC-rich motifs), has been suggested to be directly involved in
461 the regulation of its own biosynthesis. In this case, binding of daunorubicin to DNA is
462 proposed to exert a negative feedback on its production by inhibiting the transcription of
463 two of the three transcriptional activators of the regulatory cascade leading to daunorubicin
464 biosynthesis (41, 42). In addition to the direct effect of congocidine on the *cgc20-cgc1*
465 region, our results show that Cgc1 is functional in *E. coli*, which is phylogenetically distant
466 from *S. ambifaciens*. This suggests that the Cgc1 dependent or independent regulation of
467 the divergent *cgc20-cgc1* promoters do not rely on *cis* or *trans* elements other than the
468 intergenic region of *cgc20-cgc1*, the congocidine and/or Cgc1.

469 Altogether, the data obtained in this study lead us to propose a model for the regulation of
470 the *cgc* gene cluster (Figure 9). During exponential growth, before the onset of congocidine
471 production, the regulatory gene *cgc1* is the only *cgc* gene transcribed at significant but
472 relatively low level. At this stage, the amount of Cgc1 may not be sufficient for the efficient
473 activation of the transcription of the *cgc* genes and/or additional levels of regulation may be
474 required. Later, in response to signal(s) yet to be determined, metabolic differentiation
475 occurs and Cgc1 activates the transcription of the *cgc* genes, leading to congocidine
476 production and efflux. Results from experiments performed with *S. ambofaciens* ATCC23877
477 or with heterologous hosts demonstrate that congocidine induces the expression of the
478 resistance genes *cgc20-cgc21* and of the pathway-specific transcriptional activator gene,
479 *cgc1*. This induction can occur in the absence of Cgc1, probably by the direct binding of
480 congocidine to the promoter region of the *cgc20-cgc22* operon and of *cgc1*, rich in strong
481 congocidine binding sites. In a cell population, in which the onset of congocidine production
482 is not synchronized, such induction will first offer protection against congocidine to cells that
483 are not yet producing it. It will also increase the synthesis of Cgc1, first step toward the
484 biosynthesis of congocidine, thus affording the concerted production of this metabolite
485 among the cell population. Such feed-forward mechanism, in which a metabolite is inducing
486 its own production, has already been observed in Actinobacteria (planosporicin,
487 *Planomonospora alba*)(43) or in Bacilli (nisin, *Lactococcus lactis*; subtilin, *Bacillus subtilis*)
488 (44). It has been proposed that such concerted biosynthesis may be required to reach
489 effective levels of antibiotics. In *B. subtilis* and in *L. lactis*, induction of the gene clusters by
490 subtilin and nisin respectively is mediated through a two-component system. In *P. alba*,
491 planosporicin has been proposed to interact with PspW, an anti- σ factor, releasing the ECF- σ
492 factor that controls the transcription of the biosynthetic genes (43). The mechanism by

493 which congocidine induces its own biosynthesis constitute a new type of feed-forward
494 induction in which the metabolite itself induces the transcription of the pathway-specific
495 transcriptional regulator.

496 **MATERIALS AND METHODS**

497 **Bacterial strains, plasmids and growth conditions.**

498 The strains and plasmids used in this study are listed in Table 1. *Escherichia coli* strains were
499 grown in LB medium with appropriate antibiotics when necessary (ampicillin 50µg/ml;
500 apramycin 50 µg/ml; chloramphenicol 35 µg/ml; hygromycin 100 µg/ml). *Streptomyces*
501 strains were grown at 28°C on solid medium SFM (Soya Flour Mannitol) for genetic
502 manipulations and spore stock preparations (45). They were grown in liquid MP5 medium
503 (46) for the analysis of congocidine production and RNA extraction. For RNA-seq analysis, *S.*
504 *ambofaciens* ATCC23877 strain was grown in MP5 or YEME10 medium 10.3 % saccharose (3
505 g/l yeast extract, 5 g/l bactotryptone, 3 g/l malt extract, 10 g/l glucose, 103 g/l saccharose;
506 pH 7.0-7.2) with or without 2.3 µM (1 µg/ml) congocidine at 30°C in a shaking agitator (220
507 rpm, INFORS HT multitron standard) as previously described in (22).

508 **Preparation and DNA manipulations.**

509 *E. coli* transformations and *E. coli*/*Streptomyces* conjugations were performed under
510 standard conditions (45, 47). High-fidelity DNA Polymerase Phusion (ThermoFischer) was
511 used to amplify DNA cassettes and *cgc1*, and Taq polymerase (Qiagen) was used for PCR
512 verification of gene deletions in BACs or *Streptomyces* strains. DNA fragments and PCR
513 products were purified using the Nucleospin Gel and PCR clean-up kit from Macherey-Nagel.
514 All oligonucleotides used in this work are listed in Table 2.

515 **Construction of the *cgc1* and *cgc2* mutants in *S. ambofaciens* SPM110.**

516 To facilitate the construction of the *cgc1* and *cgc2* deletion mutants, the pCGC001 BAC (7)
517 was engineered to allow conjugational transfer from *E. coli* ET12567/pUZ8002 or S17.1. A 3
518 kb HindIII/XbaI/Klenow fragment from pOSV010 (48), containing the hygromycin resistance
519 gene and the RK2 origin of transfer was cloned into the HpaI-digested pCGC001 BAC, yielding
520 pCGC006.

521 Attempts to replace *cgc1* and *cgc2* by the *att2aac* or *att2 Ω aac* cassettes (20) directly in
522 pCGC006 by PCR-targeting failed. Therefore, a 7.3 kb BstBI/BsiWI/Klenow fragment from
523 pCGC001, containing *cgc1* and *cgc2* was cloned into a SmaI digested pUC19, yielding
524 pCGC041 (Supplemental Figure S8). The *cgc1* and *cgc2* genes were replaced by the *att2aac*
525 (from pOSV232) and the *att2 Ω aac* (from pOSV221) cassettes respectively, by PCR-targeting
526 (50), using the oligonucleotide pairs cmj4F/cmj17 and cmj13F/cmj13R. This resulted in
527 plasmids pCGC043 and pCGC044. The *cgc1* gene was then replaced by the *att2aac* cassette
528 in pCGC006 by λ -RED recombination using the 8.5 kb PvuII fragment from pCGC043, yielding
529 pCGC051 (Supplemental Figure S8). The same protocol was followed to replace *cgc2* with
530 the *att2 Ω aac* cassette (using pCGC044), resulting in pCGC052.

531 The BACs pCGC051 and pCGC052 were introduced into *S. ambofaciens* SPM110 by
532 conjugation from *E. coli* S17.1 and exconjugants were selected for with apramycin.
533 Conjugants were screened for sensitivity to hygromycin, indicating a double-crossover allelic
534 exchange. Replacements of *cgc1* and *cgc2* by the apramycin resistance cassettes in *S.*
535 *ambofaciens* SPM110 were verified by PCR using the primer pairs cmj12F/cmj12R and
536 cmj14F/cmj14R respectively, and the mutant strains were named CGCA013 and CGCA017,
537 respectively. To excise the *att2aac* cassette in CGCA013, pOSV235 (55) was introduced into

538 CGCA013 by conjugation from *E. coli* S17.1 and conjugants were selected for with
539 thiostrepton. Mutant strains for which the *att2aac* cassette had been excised by site specific
540 recombination were screened for apramycin sensitivity. After the loss of the unstable
541 pOSV235 plasmid, in-frame deletion of *cgc1* was verified by PCR using the primer pair
542 cmj12F/cmj12R and sequencing of the PCR product and the strain was named CGCA021.

543 **Genetic complementation of the *cgc1* mutant.**

544 To verify that the abolition of the transcription of the *cgc* genes in CGCA021 was due to the
545 deletion of *cgc1*, a plasmid expressing *cgc1* under the control of the strong constitutive
546 promoter P_{ermE}^* was constructed. The *cgc1* gene was amplified using the primers
547 *cgc(amb)1F* (containing a HindIII site) and *cgc(amb)1R* (containing a Stul site). The PCR
548 product was purified, incubated 10 min at 72°C with the Taq polymerase and cloned into the
549 pGEMT-Easy vector, yielding pCGC071. The insert was verified by DNA sequencing. The
550 0.75kb HindIII/Stul fragment from pCGC071 was then cloned into the conjugative and
551 integrative vector pOSV554 (16) digested by the same enzymes. The resulting plasmid,
552 pCGC081, was introduced by conjugation from *E. coli* S17.1 into the *S. ambifaciens* mutant
553 strain deleted for *cgc1*, CGCA021. Conjugants were selected on SFM containing apramycin
554 and verified by PCR using the primers *cgc(amb)1F/R*. The resulting strain was called
555 CGCA031.

556 **Construction of the *cgc3*, *cgc4*, *cgc6*, *cgc7* and *cgc18* replacement mutants.**

557 To study the transcriptional organization of the *cgc* cluster, we replaced several *cgc* genes by
558 the *att1Ωaac* cassette (containing an apramycin resistance gene flanked by transcription
559 terminators (20)). The pCGC006 BAC contains a hygromycin gene whose promoter is

560 identical to the promoter of the apramycin resistance gene in the *att1Ωaac* cassette.
561 Consequently, the BAC pCGC006 could not be used to construct the mutants by PCR
562 targeting. Thus, the pCGC001 BAC was re-engineered to allow conjugational transfer from *E.*
563 *coli* using pOSV408 (8) as the source of a cassette containing a kanamycin resistance gene
564 and the RK2 origin of transfer. The cassette was amplified by PCR and integrated in pCGC001
565 by PCR targeting (51), yielding the BAC pCGC079.

566 The same strategy was used to replace each of the *cgc3*, *cgc4*, *cgc6*, *cgc7* and *cgc18* genes by
567 the *att1Ωaac* cassette by PCR-targeting in pCGC079 (see Table S1 for the names of the
568 strains and plasmids and Table S2 for the oligonucleotides used). Briefly, for the replacement
569 of *cgc3*, the *att1Ωaac* cassette apramycin was amplified from pOSV230 using the *cgc3*-dis-T-
570 F/R primers and used for the replacement of *cgc3* in pCGC079, yielding pCGC084. Correct
571 replacement of *cgc3* was verified by PCR using the primers *cgc3*-SC-F/R and pCGC084 was
572 subsequently introduced into *S. ambofaciens* SMP110 by conjugal transfer via S17.1.
573 Conjugants were selected for apramycin resistance and then screened for kanamycin
574 sensitivity to isolate double cross-over recombinants. The resulting mutant strain was
575 verified by PCR and named CGCA039. Replacement of *cgc4*, *cgc6*, *cgc7* and *cgc18* by the
576 *att1Ωaac* cassette yielded the strains CGCA044, CGCA040, CGCA041 and CGCA043
577 respectively.

578 **Construction of a *S. lividans* TK23 strain expressing the *cgc* gene cluster deleted for *cgc1*.**

579 To study the expression of the resistance genes in a heterologous host and in a genetic
580 context similar to the one in *S. ambofaciens*, a *S. lividans* TK23 strain expressing the *cgc* gene
581 cluster deleted for *cgc1* was constructed. The pCGC043 plasmid previously constructed was
582 digested by PstI/EcoRI and the 8.5 kb fragment (containing *cgc1* replaced by the *att2aac*

583 cassette) was purified and used for the replacement of *cgc1* by *att2aac* in pCGC002 by PCR
584 targeting. The resulting BAC was verified by PCR and by enzymatic digestions and named
585 pCGC312. The *aac(3)IV* gene was next excised using the pOSInt3 plasmid bearing the pSAM2
586 *int* and *xis* genes (20). The correct excision was verified by PCR and the resulting BAC named
587 pCGC313. This BAC was introduced into *S. lividans* TK23 by intergeneric conjugation using
588 the *E. coli* ET12567/pUZ8002 as the host strain. The resulting strain was named CGCL095.
589

590 **RNA extraction from *Streptomyces* liquid culture and gene expression analysis by RT-PCR**
591 **and RT-qPCR.**

592 RNA was isolated from 2 ml of 43 h-old liquid cultures carried out in MP5 medium
593 complemented with congocidine 5 µg/ml when necessary, using the NucleoSpin RNAII and
594 NucleoSpin RNA/DNA Buffer Set kits (Macherey-Nagel). The RNA samples were treated with
595 RNase-free DNase (Ambion or Qiagen) followed by purification using the NucleoSpin RNA
596 Clean-up kit (Macherey-Nagel). Quality and quantity of RNAs were assessed by UV
597 spectroscopy and agarose gel electrophoresis.

598 For RT-PCR analyses, the purified RNA samples were tested for potential DNA contamination
599 by 27 cycles of PCR using primers for *hrdB* amplification with Taq polymerase (Qiagen). They
600 were then subjected to reverse transcription by ThermoScript™ RT-PCR System (Invitrogen)
601 25°C for 10 min followed by 50°C for 50 min and 85°C for 5 min. 10% of the resulting cDNA
602 was used as template for 27 cycles of PCR amplification with appropriate primers (20- to 22-
603 mers with average melting temperature from 65°C to 70°C) designed to amplify products of
604 about 300 bps (see Table S2). The PCR amplifications were performed using the following
605 program: 6 cycles at 94°C for 1 min, 55°C for 30 sec, and 72°C for 30 sec, followed by 21

606 cycles at 94°C for 1 min, 58°C for 30 sec, and 72°C for 30 sec. 1 The *hrdB* gene (encoding the
607 major sigma factor) was used as a control, as it expressed at constant level (22). PCR
608 products were analyzed by 1.2% agarose gel electrophoresis.

609 For RT-qPCR analyses, 200 ng of total RNA were reverse-transcribed in 20 µl final reaction
610 volume using the ThermoScientific Revertaid Reverse Transcription kit (K1622) following the
611 manufacturer's instructions. For each sample, negative reverse transcription reaction was
612 done to verify the absence of genomic contamination in subsequent q-PCR. Primer
613 sequences (see Supplemental Table S2) were designed using NCBI primer-blast software
614 online. SYBR-Green q-PCR reactions were performed using the Bio-Rad[®] CFX Connect Real-
615 Time PCR detection system in 96 well optical reaction plates. One µl of cDNA (env. 3.3
616 ng/reaction), standard or water (no-template control) were used as template for q-PCR
617 reactions with LightCycler(R)FastStart DNA Master SYBR Green I kit (Roche) and primers at
618 500 nM final concentration. Real-time q-PCR amplifications were carried out (95°C for 10
619 min, followed by 6 cycles of 95°C for 10 sec, 55°C for 30 sec and 72°C for 30 sec, then 39
620 cycles of 95°C for 10 sec, 58°C for 30 sec and 72°C for 30 sec and a final dissociation curve
621 analysis step from 65°C to 95°C). Technical triplicate experiments were performed for each
622 biological sample (3 independent clones). The amplification efficiency of each probe was
623 generated using the slopes of the standard curves obtained by ten-fold dilution series of
624 either genomic DNA or cDNA mix. The efficiency of the q-PCR amplifications for all the genes
625 tested was higher than 90 % (except for *cgc19* for which the efficiency was at 80%).
626 Amplification specificity for each q-PCR reaction was confirmed by the dissociation curve
627 analysis and by gel electrophoresis. Determined Ct values were then exploited for further
628 analysis (56). The values were normalized on four reference genes (*aspS*, *rpoB*, *gyrA*, *gyrB*),

629 the mean value being used to calculate the fold change and performed statistical analyses of
630 the results. Data were analyzed with R software (57) and statistical significance was assessed
631 by means of Welch two sample *t* tests. Asterisks depict statistical significance (**p*<0.05,
632 ***p*<0.01, ****p*<0.001).

633 **Bioinformatic analyses and code availability.**

634 We previously described RNA extraction and sequencing from *S. ambofaciens* ATCC 23877
635 strain grown in MP5 and YEME10 media supplemented or not with congocidine (22).
636 Accordingly we analyzed the data deposited in the NCBI Gene Expression Omnibus (GEO,
637 <https://www.ncbi.nlm.nih.gov/geo/>) under the accession number GSE162865. We manually
638 curated the annotation of the start and end position of some *cgc* genes (Supplemental
639 Dataset 1) to generate a new annotation file. The scripts used for RNA-seq data analyses and
640 for strong congocidine motif counting are available on the following Github link:
641 <https://github.com/Bury-Mone/Congomics> (supplemental methods). The Integrative
642 Genomics Viewer (IGV) tool was used to visualize simultaneously RNAseq data and genomic
643 annotations (58).

644 **Liquid Chromatography analyses.**

645 After five days of culture at 28°C in MP5 medium, supernatants were filtered through Mini-
646 UniPrep syringeless filter devices (0.2 µm, Whatman) and analyzed on an Atlantis T3 column
647 (250 mm x 4.6 mm, 5 µm, column temperature 30°C) using an Agilent 1200 HPLC instrument
648 equipped with a quaternary pump. Samples were eluted with isocratic 0.1% HCOOH in H₂O
649 (solvent A)/ 0.1% HCOOH in CH₃CN (solvent B) (95:5) at 1 ml.min⁻¹ for 7 min, followed by a

650 gradient to 40:60 A/B over 23 min. Congocidine was detected by monitoring absorbance at
651 297 nm.

652 ***E. coli* ET1267/pUZ8002 growth curves.**

653 *E. coli* ET1267/pUZ8002 strain was transformed with pCGC002 (complete *cgc* gene cluster)
654 or pCGC313 (*cgc* gene cluster deleted for *cgc1*) and selected on LB medium with hygromycin
655 (50 µg/ml) and kanamycin (50 µg/ml). Bacteria were cultivated overnight in 2 ml of LB
656 supplemented with kanamycin (50 µg/ml) at 37°C under agitation (180 rpm). The samples
657 were then diluted to an adjusted OD_{600nm} of 0.02 in LB medium supplemented or not with
658 congocidine (11.6 or 23.2 µM). 200 µl of these dilutions were inoculated in triplicates into a
659 96-well plate (655161 GBO), sealed with an adhesive (4Ti – 0516/96). Bacteria were grown at
660 37°C in a CLARIOstar microplate reader (BMG LABTECH), under agitation (400 rpm), set with
661 an automatically repeating protocol of cell density (OD_{600nm}) readings every 20 min for
662 each well during 16 h. The data obtained from the plate reader measurements were
663 analyzed with Excel (Microsoft) and R software.

664 **Study of the *cgc20-cgc1* intergenic region using the reporter gene *gusA* in *S. lividans* TK23**

665 The complete *cgc20-cgc1* intergenic region was amplified by PCR using LS-RP1 and LS-RP2, or
666 LS-RP3 and LS-RP4 primers to introduce NotI and NdeI restriction sites in both orientations.
667 The amplicons were cloned in pCR-blunt (Invitrogen) and sequenced. These plasmids were
668 digested by NdeI and NotI, and the fragments containing the *cgc20-cgc1* intergenic region
669 were cloned upstream of *gusA* in the pOSV751 plasmid linearized by the same enzymes,
670 yielding pCGC306 (*gusA* under the control of the *cgc1* promoter, P_{*cgc1*}) and pCGC307 (*gusA*
671 under the control of the *cgc20* promoter, P_{*cgc20*}). The negative control, pCGC308, was

672 obtained after religation of NdeI/NotI/Klenow-digested pOSV751. The three plasmids were
673 next digested by SphI and PvuII and the fragments encompassing the three reporter
674 cassettes (*gusA* without promoter or under the control of P_{cgc1} or P_{cgc20} promoters)
675 surrounded by terminators were cloned into SphI/PvuII-digested pOVS808 (conjugative and
676 integrative plasmid) (54). The resulting plasmids were named pCGC309 (*gusA* without
677 promoter), pCGC310 (*gusA* under the control of P_{cgc1}), and pCGC311 (*gusA* under the control
678 of P_{cgc20}), respectively and introduced in *S. lividans* TK23 and CGCL0107 (expressing the
679 *cgc20-cgc21* resistance genes) by intergeneric conjugation, yielding the CGCL101 to CGCL106
680 strains. Genomic DNA was extracted from the exconjugants and analyzed by PCR using
681 primers CEA31+LS04 and CEA29+LS63 for the proper integration at the VWB *attB* site and
682 primers LS25, LS26 and Cmj12R for the presence of the appropriate reporter cassette.

683 **Spectrophotometric measurement of β -Glucuronidase activity in cell lysates.**

684 For each strain, 10^5 spores were inoculated into 10 ml of TSB medium at 28°C under
685 agitation (180 rpm, INFORS HT Unitron). After 24 h, the cultures were supplemented or not
686 with congocidine (2.3 or 11.6 μ M) and incubated in the same conditions for five additional
687 hours. The mycelia were harvested by centrifugation, washed once with sodium phosphate
688 buffer (100 mM, pH 7) and resuspended in 1 ml of buffer. Cells were disrupted using a One
689 Shot homogenizer (Constant Systems Ltd.) at 2.5 bar. Lysates were centrifuged at 16 000 g
690 and 4°C for 15 min. The supernatants were used to determine the protein concentration
691 using the Bradford assay (59). GusA enzymatic activity was measured as follows: 100 μ g of
692 proteins were diluted in 950 μ l (final volume) of sodium phosphate buffer (100 mM, at pH 7)
693 and placed in a spectrophotometer (UV-1800 Shimadzu) with temperature control and
694 connected to a computer. Reactions were performed at 37°C and started by addition of 50 μ l

695 of 20 mM p-nitrophenyl- β -D-glucuronide (Megazyme). The optical density at 415 nm was
696 automatically measured all 30 sec during 30 min. The data obtained from
697 spectrophotometer measurements were analyzed with Excel (Microsoft) and R software.

698 **ACKNOWLEDGMENTS**

699 AV, MJ, SBM, SL and JLP designed the experiments, AV, FL, MJ, AN, SBM, LFDS, LMF carried
700 them out. AV, MJ, SBM, AN, LFDS, JLP and SL analyzed them. AV, JLP, SBM, AN, LFDS and SL
701 drafted the manuscript and SBM, JLP and SL reviewed and edited the draft manuscript. This
702 work was supported by fellowships from the French government (to A.V.) and by grants from
703 Region Ile-de-France (to M.J.). A.V would like to thank the Fondation pour la Recherche
704 Médicale (FRM) for financial support.

705 **REFERENCES**

- 706 1. Hoskisson PA, Fernández-Martínez LT. 2018. Regulation of specialised metabolites in
707 Actinobacteria - expanding the paradigms. *Environ Microbiol Rep* 10:231–238.
- 708 2. Niu G, Chater KF, Tian Y, Zhang J, Tan H. 2016. Specialised metabolites regulating
709 antibiotic biosynthesis in *Streptomyces* spp. *FEMS Microbiol Rev* 40:554–573.
- 710 3. Le TBK, Fiedler H-P, den Hengst CD, Ahn SK, Maxwell A, Buttner MJ. 2009. Coupling
711 of the biosynthesis and export of the DNA gyrase inhibitor simocyclinone in
712 *Streptomyces antibioticus*. *Mol Microbiol* 72:1462–1474.
- 713 4. Arias P, Fernández-Moreno MA, Malpartida F. 1999. Characterization of the pathway-
714 specific positive transcriptional regulator for actinorhodin biosynthesis in *Streptomyces*
715 *coelicolor* A3(2) as a DNA-binding protein. *J Bacteriol* 181:6958–6968.

- 716 5. Horbal L, Rebets Y, Rabyk M, Makitrynsky R, Luzhetskyy A, Fedorenko V, Bechthold
717 A. 2012. SimReg1 is a master switch for biosynthesis and export of simocyclinone D8
718 and its precursors. *AMB Express* 2:1.
- 719 6. Lu F, Hou Y, Zhang H, Chu Y, Xia H, Tian Y. 2017. Regulatory genes and their roles
720 for improvement of antibiotic biosynthesis in *Streptomyces*. *3 Biotech* 7:250.
- 721 7. Juguet M, Lautru S, Francou F-X, Nezbedová S, Leblond P, Gondry M, Pernodet J-L.
722 2009. An iterative nonribosomal peptide synthetase assembles the pyrrole-amide
723 antibiotic congocidine in *Streptomyces ambofaciens*. *Chem Biol* 16:421–431.
- 724 8. Vingadassalon A, Lorieux F, Juguet M, Le Goff G, Gerbaud C, Pernodet J-L, Lautru S.
725 2015. Natural combinatorial biosynthesis involving two clusters for the synthesis of
726 three pyrrolamides in *Streptomyces netropsis*. *ACS Chem Biol* 10:601–610.
- 727 9. Aubry C, Clerici P, Gerbaud C, Micouin L, Pernodet J-L, Lautru S. 2020. Revised
728 Structure of Anthelvencin A and Characterization of the Anthelvencin Biosynthetic
729 Gene Cluster. *ACS Chem Biol* 15:945–951.
- 730 10. Seedorf T, Kirschning A, Solga D. 2021. Natural and Synthetic Oligoarylamides:
731 Privileged Structures for Medical Applications. *Chem Weinh Bergstr Ger* 27:7321–
732 7339.
- 733 11. Zimmer C, Marck C, Schneider C, Guschlbauer W. 1979. Influence of nucleotide
734 sequence on dA.dT-specific binding of Netropsin to double stranded DNA. *Nucleic*
735 *Acids Res* 6:2831–2837.

- 736 12. Kopka ML, Yoon C, Goodsell D, Pjura P, Dickerson RE. 1985. The molecular origin of
737 DNA-drug specificity in netropsin and distamycin. *Proc Natl Acad Sci U S A* 82:1376–
738 1380.
- 739 13. Barrett MP, Gemmell CG, Suckling CJ. 2013. Minor groove binders as anti-infective
740 agents. *Pharmacol Ther* 139:12–23.
- 741 14. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment
742 search tool. *J Mol Biol* 215:403–410.
- 743 15. Marchler-Bauer A, Bryant SH. 2004. CD-Search: protein domain annotations on the fly.
744 *Nucleic Acids Res* 32:W327-331.
- 745 16. Nguyen HC, Karray F, Lautru S, Gagnat J, Lebrihi A, Huynh TDH, Pernodet J-L. 2010.
746 Glycosylation steps during spiramycin biosynthesis in *Streptomyces ambofaciens*:
747 involvement of three glycosyltransferases and their interplay with two auxiliary proteins.
748 *Antimicrob Agents Chemother* 54:2830–2839.
- 749 17. Santos CL, Correia-Neves M, Moradas-Ferreira P, Mendes MV. 2012. A walk into the
750 LuxR regulators of Actinobacteria: phylogenomic distribution and functional diversity.
751 *PloS One* 7:e46758.
- 752 18. Capra EJ, Laub MT. 2012. Evolution of two-component signal transduction systems.
753 *Annu Rev Microbiol* 66:325–347.
- 754 19. Martin JF, Sola-Landa A, Rodríguez-García A. 2012. Two-component Systems in
755 *Streptomyces*, p. 315–332. *In* Two-Component Systems in Bacteria. Roy Gross and
756 Dagmar Beier. Caister Academic Press, University of Würzburg, Germany.

- 757 20. Raynal A, Karray F, Tuphile K, Darbon-Rongère E, Pernodet J-L. 2006. Excisable
758 cassettes: new tools for functional analysis of *Streptomyces* genomes. *Appl Environ*
759 *Microbiol* 72:4839–4844.
- 760 21. Dangel V, Härle J, Goerke C, Wolz C, Gust B, Pernodet J-L, Heide L. 2009.
761 Transcriptional regulation of the novobiocin biosynthetic gene cluster. *Microbiol Read*
762 *Engl* 155:4025–4035.
- 763 22. Liroy VS, Lorenzi J-N, Najah S, Poinsignon T, Leh H, Saulnier C, Aigle B, Lautru S,
764 Thibessard A, Lespinet O, Leblond P, Jaszczyszyn Y, Gorrichon K, Varoquaux N,
765 Junier I, Boccard F, Pernodet J-L, Bury-Moné S. 2021. Dynamics of the
766 compartmentalized *Streptomyces* chromosome during metabolic differentiation. *Nat*
767 *Commun* 12:5221.
- 768 23. Myronovskiy M, Welle E, Fedorenko V, Luzhetskyy A. 2011. Beta-glucuronidase as a
769 sensitive and versatile reporter in actinomycetes. *Appl Environ Microbiol* 77:5370–
770 5383.
- 771 24. Neidle S. 2001. DNA minor-groove recognition by small molecules. *Nat Prod Rep*
772 18:291–309.
- 773 25. Triebel H, Bär H, Geuther R, Burckhardt G. 1995. Netropsin-induced changes of DNA
774 supercoiling; sedimentation studies, p. 45–54. *In Analytical Ultracentrifugation* Behlke J.
775 (eds). Steinkopff Verlag.
- 776 26. Tevis DS, Kumar A, Stephens CE, Boykin DW, Wilson WD. 2009. Large, sequence-
777 dependent effects on DNA conformation by minor groove binding compounds. *Nucleic*
778 *Acids Res* 37:5550–5558.

- 779 27. Ward B, Reh fuss R, Goodisman J, Dabrowiak JC. 1988. Determination of netropsin-
780 DNA binding constants from footprinting data. *Biochemistry* 27:1198–1205.
- 781 28. Abu-Day a A, Brown PM, Fox KR. 1995. DNA sequence preferences of several AT-
782 selective minor groove binding ligands. *Nucleic Acids Res* 23:3385–3392.
- 783 29. McLean TC, Lo R, Tschowri N, Hoskisson PA, Al Bassam MM, Hutchings MI, Som
784 NF. 2019. Sensing and responding to diverse extracellular signals: an updated analysis
785 of the sensor kinases and response regulators of *Streptomyces* species. *Microbiol Read*
786 *Engl* 165:929–952.
- 787 30. Guthrie EP, Flaxman CS, White J, Hodgson DA, Bibb MJ, Chater KF. 1998. A
788 response-regulator-like activator of antibiotic synthesis from *Streptomyces coelicolor*
789 A3(2) with an amino-terminal domain that lacks a phosphorylation pocket. *Microbiol*
790 *Read Engl* 144 (Pt 3):727–738.
- 791 31. Wang L, Tian X, Wang J, Yang H, Fan K, Xu G, Yang K, Tan H. 2009. Autoregulation
792 of antibiotic biosynthesis by binding of the end product to an atypical response regulator.
793 *Proc Natl Acad Sci U S A* 106:8617–8622.
- 794 32. Bailey TL, Elkan C. 1994. Fitting a mixture model by expectation maximization to
795 discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* 2:28–36.
- 796 33. Chater KF, Chandra G. 2008. The use of the rare UUA codon to define “expression
797 space” for genes involved in secondary metabolism, development and environmental
798 adaptation in *streptomyces*. *J Microbiol Seoul Korea* 46:1–11.

- 799 34. Hackl S, Bechthold A. 2015. The Gene *bldA*, a regulator of morphological
800 differentiation and antibiotic production in *streptomyces*. Arch Pharm (Weinheim)
801 348:455–462.
- 802 35. Koshla O, Yushchuk O, Ostash I, Dacyuk Y, Myronovskyi M, Jäger G, Süßmuth RD,
803 Luzhetskyy A, Byström A, Kirsebom LA, Ostash B. 2019. Gene *miaA* for post-
804 transcriptional modification of tRNAXXA is important for morphological and metabolic
805 differentiation in *Streptomyces*. Mol Microbiol 112:249–265.
- 806 36. Sehin Y, Koshla O, Dacyuk Y, Zhao R, Ross R, Myronovskyi M, Limbach PA,
807 Luzhetskyy A, Walker S, Fedorenko V, Ostash B. 2019. Gene *ssfg_01967* (*miaB*) for
808 tRNA modification influences morphogenesis and moenomycin biosynthesis in
809 *Streptomyces ghanaensis* ATCC14672. Microbiol Read Engl 165:233–245.
- 810 37. Memili E, Weisblum B. 1997. Essential role of endogenously synthesized tylosin for
811 induction of ermSF in *Streptomyces fradiae*. Antimicrob Agents Chemother 41:1203–
812 1205.
- 813 38. Tahlan K, Ahn SK, Sing A, Bodnaruk TD, Willems AR, Davidson AR, Nodwell JR.
814 2007. Initiation of actinorhodin export in *Streptomyces coelicolor*. Mol Microbiol
815 63:951–961.
- 816 39. Xu Y, Willems A, Au-Yeung C, Tahlan K, Nodwell JR. 2012. A two-step mechanism
817 for the activation of actinorhodin export and resistance in *Streptomyces coelicolor*. mBio
818 3:e00191-00112.
- 819 40. Bruzik JP, Auble DT, deHaseh PL. 1987. Specific activation of transcription initiation
820 by the sequence-specific DNA-binding agents distamycin A and netropsin. Biochemistry
821 26:950–956.

- 822 41. Furuya K, Hutchinson CR. 1996. The DnrN protein of *Streptomyces peucetius*, a
823 pseudo-response regulator, is a DNA-binding protein involved in the regulation of
824 daunorubicin biosynthesis. *J Bacteriol* 178:6310–6318.
- 825 42. Ajithkumar V, Prasad R. 2010. Modulation of *dnrN* expression by intracellular levels of
826 DnrO and daunorubicin in *Streptomyces peucetius*. *FEMS Microbiol Lett* 306:160–167.
- 827 43. Sherwood EJ, Bibb MJ. 2013. The antibiotic planosporicin coordinates its own
828 production in the actinomycete *Planomonospora alba*. *Proc Natl Acad Sci U S A*
829 110:E2500-2509.
- 830 44. Kleerebezem M. 2004. Quorum sensing control of lantibiotic production; nisin and
831 subtilin autoregulate their own biosynthesis. *Peptides* 25:1405–1414.
- 832 45. Kieser T, Bibb MJ, Buttner MJ, Chater KK, Hopwood DA. 2000. *Practical streptomyces*
833 *genetics*. John Innes Foundation.
- 834 46. Pernodet JL, Alegre MT, Blondelet-Rouault MH, Guérineau M. 1993. Resistance to
835 spiramycin in *Streptomyces ambofaciens*, the producer organism, involves at least two
836 different mechanisms. *J Gen Microbiol* 139:1003–1011.
- 837 47. Sambrook J, Russell D. 2000. *Molecular Cloning: A Laboratory Manual* 3rd Revised
838 edition. Cold Spring Harbor Laboratory Press, U.S.
- 839 48. Jeannin P, Pernodet J-L, Guérineau M, Simonet P, Courtois S, Cappellano C, Francou F,
840 Raynal A, Ball M, Sezonov G, Tuphile K, Frostegard A. 2000. Method for obtaining
841 nucleic acids from an environment sample, resulting nucleic acids and use in synthesis
842 of novel compounds. WO 2001/040497 A2.

- 843 49. Simon R, Priefer U, Pühler A. 1983. A broad host range mobilization system for *in vivo*
844 genetic engineering: transposon mutagenesis in gram negative bacteria. *Nat Biotechnol*
845 1:784–791.
- 846 50. Gust B, Chandra G, Jakimowicz D, Yuqing T, Bruton CJ, Chater KF. 2004. Lambda red-
847 mediated genetic manipulation of antibiotic-producing *Streptomyces*. *Adv Appl*
848 *Microbiol* 54:107–128.
- 849 51. Yu DG, Ellis HM, Lee EC, Jenkins NA, Copeland NG, Court DL. 2000. An efficient
850 recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad*
851 *Sci U S A* 97:5978–5983.
- 852 52. Yanisch-Perron C, Vieira J, Messing J. 1985. Improved M13 phage cloning vectors and
853 host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33:103–
854 119.
- 855 53. Raynal A, Tuphile K, Gerbaud C, Luther T, Guérineau M, Pernodet J-L. 1998. Structure
856 of the chromosomal insertion site for pSAM2: functional analysis in *Escherichia coli*.
857 *Mol Microbiol* 28:333–342.
- 858 54. Aubry C, Pernodet J-L, Lautru S. 2019. Modular and Integrative Vectors for Synthetic
859 Biology Applications in *Streptomyces* spp. *Appl Environ Microbiol* 85.
- 860 55. Karray F. 2005. Etude de la biosynthèse de l'antibiotique spiramycine par *Streptomyces*
861 *ambofaciens*. PhD Thesis, Université Paris-Sud, FR.
- 862 56. Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT-
863 PCR. *Nucleic Acids Res* 29:e45.

- 864 57. RDC. R: A language and environment for statistical computing. Foundation for
865 Statistical Computing, Vienna, Austria.
- 866 58. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov
867 JP. 2011. Integrative genomics viewer. *Nat Biotechnol* 29:24–26.
- 868 59. Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram
869 quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*
870 72:248–254.

871 **Figure legends**

872 **Figure 1:** Genetic organization of the congocidine gene cluster in *S. ambofaciens* (A). The
873 gene functions are indicated by various filling patterns. Promoter regions in the *cgc* gene
874 cluster are indicated by arrows. B) Chemical structure of congocidine. Colors indicate the
875 different types of congocidine substructures.

876 **Figure 2: Determination of the function of Cgc1.** A) LC analysis of congocidine production in
877 the SPM110 reference strain (upper panel), the *cgc1* deletion mutant CGCA021 (middle
878 panel) and the genetically complemented CGC031 strain (lower panel). B) RT-PCR analysis of
879 the expression of the *cgc* genes after 43 h of culture in MP5 medium in the SPM110
880 reference strain strain (upper panel), the *cgc1* deletion mutant CGCA021 (middle panel) and
881 the genetically complemented CGC031 strain (lower panel). *hrdB* transcript was used as a
882 control.

883 **Figure 3 : Congocidine gene cluster transcriptome dynamics over growth in MP5 medium.**

884 Bacteria were grown in MP5 liquid medium and harvested in either exponential (24 h) or
885 stationary (48 h) phase. A) Plot of the log ratio of differential expression as a function of
886 gene expression in stationary phase. The differentially expressed genes (adjusted p value <

887 0.05) are colored in red circles and yellow triangles for *cgc* and non-*cgc* genes, respectively.
888 B) Focus on the RNA-seq analysis of *cgc* gene cluster in exponential and stationary phase.
889 Each dot corresponds to the number of reads (normalized by DESeq2 and on gene size, log2)
890 for one gene in the different RNA-seq condition performed in triplicates. A circle (in the color
891 of the condition of interest) has been added to indicate overlapping dots. Of note, the
892 expression of the *cgc* genes was statistically induced in stationary phase (adjusted p value \leq
893 0.01, Supplemental Dataset S1).

894 **Figure 4 : RT-PCR analysis of the expression of *cgc20* and *cgc21* (resistance genes), of *cgc1***
895 **(transcriptional regulator gene) and of *cgc22*, *cgc2*, *cgc3* and *cgc19* (biosynthetic genes) in**
896 the reference strain SPM110 and in the *cgc1* deletion mutant (CGCA021) in presence of
897 11.6 μ M congocidine. The transcript of *hrdB* was used as a control. It should be noted that
898 the primers used do not allow the detection of *cgc1* transcripts for the CGCA021 (Δ *cgc1*)
899 strain as they are complementary to the deleted region.

900 **Figure 5: RT-qPCR analysis of the expression of the *cgc* genes by the *S. lividans* CGCL095**
901 **strain cultivated in presence or absence of 11.6 μ M exogenous congocidine. *S. lividans***
902 CGCL095 strain was cultured 43 h in 50 ml liquid medium supplemented or not with 11.6 μ M
903 congocidine. The results correspond to the mean fold change (\pm standard error) of RNA
904 transcripts quantified in presence of congocidine over the control condition (no congocidine)
905 in three independent clones. The quantifications were performed in triplicates and
906 normalized on the mean of four reference genes (*aspS*, *rpoB*, *gyrA*, *gyrB*). Significant
907 differences between the level of expression in presence or in absence of congocidine are
908 indicated with asterisks (based on a Welch two sample *t* test, * p <0.05, ** p <0.01).

909

910 **Figure 6: Congocidine gene cluster transcriptome dynamics in presence of exogenous**
911 **congocidine.** Bacteria were grown in YEME10 liquid medium in presence or absence of 2.3
912 μM of congocidine and harvested in exponential (24 h) phase. Each dot corresponds to the
913 number of reads (normalized by DESeq2 and on gene size) for one gene in the different RNA-
914 seq condition performed in triplicates. A circle (in the color of the condition of interest) has
915 been added to indicate overlapping dots. Genes whose expression is statistically induced
916 (adjusted p value ≤ 0.01 , supplemental Dataset 1) are framed in dotted lines, and their fold
917 change (considering unlogged numbers of reads) indicated.

918

919 **Figure 7: GusA activity in *S. lividans* strains with *gusA* expression under the control of P_{cgc1}**
920 **and P_{cgc20} promoters in the presence or absence of congocidine.** Results correspond to the
921 mean values (\pm standard errors) of four independent experiments. Statistical differences
922 were determined by using a Welch Two Sample t-tests (one symbol: p value < 0.05 , two
923 symbols: p value < 0.01). Black asterisks correspond to statistical differences in the GusA
924 activity of reporter strains compared to the control strain grown in the same condition, blue
925 asterisks indicate statistical differences in the GusA activity of a given strain after growth in
926 the absence versus in the presence of congocidine.

927 **Figure 8: Growth of *E. coli* ET1267/pUZ8002 strains harboring the complete or *cgc1*-deleted**
928 **congocidine biosynthetic gene cluster in presence or absence of congocidine**
929 Black triangles: ET12567/pUZ8002/pCGC002 (complete *cgc* gene cluster); dark grey squares:
930 ET12567/pUZ8002/pCGC313 (*cgc* gene cluster deleted for *cgc1*); light grey disks:
931 ET12567/pUZ8002 (control).

932 Results correspond to the mean values (\pm standard errors) of at least 4 independent
 933 experiments performed in triplicates.

934 **Figure 9: Proposed model for the transcriptional regulation of the *cgc* gene cluster.**

935 **Table 1:** Strains and plasmids used in this study.

Name	Description	Reference
<i>E. coli</i> strains		
S17 .1	Host strain for conjugation from <i>E. coli</i> to <i>Streptomyces</i>	(49)
ET12567/ pUZ8002	Host strain for conjugation from <i>E. coli</i> to <i>Streptomyces</i>	(50)
DH5 α	General cloning host	Promega
DY330	Strain used for λ RED-mediated recombination	(51)
BW25113/ pIJ790	Strain used for λ RED-mediated recombination	(50)
<i>Streptomyces</i> strains		
<i>S. ambofaciens</i> ATCC 23877	Wild-type <i>S. ambofaciens</i> strain (type strain)	ATCC
SPM110	Derived from <i>S. ambofaciens</i> OSC2, spiramycin non producer	(16)
CGCA013	SPM110- <i>cgc1</i> :: <i>att2aac</i>	This work
CGCA017	SPM110- <i>cgc2</i> :: <i>att2Ωaac</i>	This work
CGCA019	SPM110- <i>cgc22</i> /19:: <i>att3</i>	(7)
CGCA021	SPM110- <i>cgc1</i> :: <i>att2</i>	This work
CGCA031	CGCA021 containing pCGC081	This work
CGCA039	SPM110- <i>cgc3</i> :: <i>att1Ωaac</i>	This work
CGCA040	SPM110- <i>cgc6</i> :: <i>att1Ωaac</i>	This work
CGCA041	SPM110- <i>cgc7</i> :: <i>att1Ωaac</i>	This work
CGCA043	SPM110- <i>cgc18</i> :: <i>att1Ωaac</i>	This work
CGCA044	SPM110- <i>cgc4</i> :: <i>att1Ωaac</i>	This work
CGCL095	<i>S. lividans</i> TK23 harboring pCGC313 (<i>cgc</i> gene cluster deleted for <i>cgc1</i>)	This work
CGCL101	<i>S. lividans</i> TK23 harboring pCGC309 (<i>gusA</i> reporter cassette, no promoter)	This work
CGCL102	<i>S. lividans</i> TK23 harboring pCGC310 (<i>gusA</i> under the control of P _{<i>cgc1</i>})	This work
CGCL103	<i>S. lividans</i> TK23 harboring pCGC311 (<i>gusA</i> under the control of P _{<i>cgc20</i>})	This work
CGCL104	<i>S. lividans</i> CGCL107 harboring pCGC309 (<i>gusA</i> reporter cassette, no promoter)	This work
CGCL105	<i>S. lividans</i> CGCL107 harboring pCGC310 (<i>gusA</i> under the control of P _{<i>cgc1</i>})	This work
CGCL106	<i>S. lividans</i> CGCL107 harboring pCGC311 (<i>gusA</i> under the control of P _{<i>cgc20</i>})	This work

Name	Description	Reference
CGCL107	<i>S. lividans</i> TK23 harboring pCGC090 (expressing <i>cgc20</i> and <i>cgc21</i> under their own promoter, P_{cgc20})	This work
Plasmids		
pGEMT-easy	<i>E. coli</i> cloning vector	Promega
pOSV010	source of the $\Omega_{hyg-oriT}$ fragment	(48)
pOSV221	source of the <i>att2Ωaac</i> cassette	(16)
pOSV230	source of the <i>att1Ωaac</i> cassette	(20)
pOSV232	source of the <i>att2aac</i> cassette	(20)
pOSV235	same as pOSV236 (16) with the <i>pac-oriT</i> insert in the opposite orientation	(20)
pOSV408	<i>E. coli</i> / <i>Streptomyces</i> shuttle vector integrative in <i>Streptomyces</i>	(8)
pOSV554	<i>Streptomyces</i> expression vector	(16)
pUC19	<i>E. coli</i> cloning vector	(52)
pOSInt3	Plasmid bearing the pSAM2 <i>xis/int</i> genes, used to excise the <i>att2aac</i> cassette	(53)
pGUS	pSET152 derivative containing omega <i>addA</i> cassette and <i>gusA</i>	(23)
pOSV751	pGUS derivative with a NotI- <i>cat</i> -NdeI gene cassette cloned upstream of <i>gusA</i> and a T_{fd} terminator downstream of <i>gusA</i> .	This work
pOSV808	Conjugative vector containing hygromycin resistance, VWB integrase and <i>amilCP</i>	(54)
pCGC001	pBeloBAC11 containing a 43.4 kb <i>S. ambofaciens</i> DNA fragment carrying the entire <i>cgc</i> gene cluster	(7)
pCGC002	Integrative and conjugative BAC carrying the complete <i>cgc</i> gene cluster	(4)
pCGC006	$\Omega_{hyg-oriT}$ HindIII/XbaI/Klenow fragment from pOSV010 cloned in pCGC001/HpaI	This work
pCGC041	pUC19 containing the 7.3 kb (<i>Bst</i> BI/ <i>Bsi</i> BW) fragment from pCGC001	This work
pCGC043	pCGC041- <i>cgc1</i> :: <i>att2aac</i>	This work
pCGC044	pCGC041- <i>cgc2</i> :: <i>att2Ωaac</i>	This work
pCGC051	pCGC006- <i>cgc1</i> :: <i>att2aac</i>	This work
pCGC052	pCGC006- <i>cgc2</i> :: <i>att2Ωaac</i>	This work
pCGC071	<i>cgc1</i> cloned into pGEM-T Easy	This work
pCGC079	pCGC001 containing <i>neo-oriT</i> from pOSV408	This work
pCGC081	<i>cgc1</i> under the control of <i>ermE</i> *p cloned in pOSV554	This work
pCGC084	pCGC079- <i>cgc3</i> :: <i>att1Ωaac</i>	This work
pCGC085	pCGC079- <i>cgc6</i> :: <i>att1Ωaac</i>	This work
pCGC086	pCGC079- <i>cgc7</i> :: <i>att1Ωaac</i>	This work
pCGC088	pCGC079- <i>cgc18</i> :: <i>att1Ωaac</i>	This work
pCGC089	pCGC079- <i>cgc4</i> :: <i>att1Ωaac</i>	This work
pCGC090	pSET152 derivative with <i>cgc20</i> and <i>cgc21</i> under the control of their native promoter P_{cgc20}	This work
pCGC306	pOSV751 derivative with a deletion of the <i>cat</i> cassette	This work
pCGC307	pOSV751 derivative with <i>gusA</i> under the control of P_{cgc1}	This work

Name	Description	Reference
pCGC308	pOSV751 derivative with <i>gusA</i> under the control of P _{<i>cgc20</i>}	This work
pCGC309	pOSV808 derivative, <i>amilCP</i> replaced by <i>gusA</i> without promoter sequence	This work
pCGC310	pOSV808 derivative, <i>amilCP</i> replaced by P _{<i>cgc1_gusA</i>}	This work
pCGC311	pOSV808 derivative, <i>amilCP</i> replaced by P _{<i>cgc20_gusA</i>}	This work
pCGC312	pCGC002- <i>cgc1</i> :: <i>att2aac</i>	This work
pCGC313	pCGC002- <i>cgc1</i> :: <i>att2</i>	This work

936 **Table 2:** Oligonucleotides used in this study.

Name	Sequence	Description
cgc(amb)1F	<u>AAGCTTGAGAAGGGAGCGGACATATGAGTCAGAC</u> GCTGTGCCA	Amplification <i>cgc1</i> , HindIII site underlined
cgc(amb)1R	<u>AGGCCTTACGCGGCCGCGTGGTGGTGGTGGTGG</u> TGCGCGGCCGCTCAGCAGGTGATGAGACCGT	Amplification <i>cgc1</i> , StuI site underlined
cmj4F	GGGCGCACGGACGGGGACGTCCCCGGCCGCTGA GGCCCCGATCTACCTCTTCGTCCCGAAGCAACT	<i>cgc1</i> in frame deletion or replacement
cmj17	GGGGGGCCAGAGGCGCAAGGCGGCAGTGAGGAG GAGTACGATCGGCGCGCTTCGTTCGGGACGAA	
cmj12F	CGAGACGGACGGCACGGTGAC	Verification <i>cgc1</i> deletion
cmj12R	TGGCAACTCGGCCGCACTCG	
cmj13F	GTGCCCTCGGCCGTGTCAGCCCACCCGGGGACGA GGCCCCGATCTACCTCTTCGTCCCGAAGCAACT	<i>cgc2</i> in frame deletion or replacement
cmj13R	TACGGTGAGCGGGCGTGACAGCCCCGGAGAGG AGCGAGCATCGGCGCGCTTCGTTCGGGACGAAG	
cmj14F	TCGGAGCCTTCGGCCACAG	Verification <i>cgc2</i> deletion
cmj14R	CGGGCAGGGCCATATGTACC	
cgc3-dis-T-F	CGAGTTCGTCTCGACCGCGTGCCGACGGAACG GGGTGATCTACCTCTTCGTCCCGAAG	Replacement of <i>cgc3</i>
cgc3-dis-T-R	CGCCGGCCACCGTGGTGTCCCTCATGACATCTCCC GATCATCGCGCGCTTCGTTCGGGAC	
cgc3-SC-FOR	CGGCGTCGATGATGTCCC	Verification of <i>cgc3</i> replacement
cgc3-SC-REV	ACCGTGACCGTCCAC	
cgc4-dis-T-F	GCCGAGTTCGACGTGGGTGCCGTGGTGCATC AGCTATCTACCTCTTCGTCCCGAAG	Replacement of <i>cgc4</i>
cgc4-dis-T-R	GGACGGAGGGCCTCGACATGACCGGAGTCAGTGT CTTCATCGCGCGCTTCGTTCGGGAC	
cgc4-SC-FOR	GGCGTGACGACGACGATGGG	Verification of <i>cgc4</i> replacement
cgc4-SC-REV	GAGGTCGTGGGTTCAAATCC	
cgc6-dis-T-F	GCTCCGCGCGCACGACCTCGGGGCGCAGGTACAT CATCAATCTACCTCTTCGTCCCGAAG	Replacement of <i>cgc6</i>

Name	Sequence	Description
cgc6-dis-T-R	GACCTGGTCGGACGCAGTGTCAAGAGGTCATGAG GTGAGATCGCGCGCTTCGTTCCGGGAC	
cgc6-SC-FOR	GACGGGATGCCGGTAGACAC	Verification of <i>cgc6</i> replacement
cgc6-SC-REV	CACTGCACCGGGAGATGCAC	
cgc7-dis-T-F2	CTCGGGAAAGCGCATGGTCATCGTGCAGTAGAGC GAGGCATCTACCTCTTCGTCCCGAAG	Replacement of <i>cgc7</i>
cgc7-dis-T-R2	ACTCGTGCACCGCGCGCCTTCGACCAGAGGAGAA CCCATGATCGCGCGCTTCGTTCCGGGA	
cgc7-SC-FOR2	CCGGTGGGATCCTTCAGGAC	Verification of <i>cgc7</i> replacement
cgc7-SC-REV	GGGAGCGCCTGTTCTGGCAC	
cgc18-dis-T-F	CCGCGTCCTCTCGAGCATCGCGGTGGCGTCGAAC TTGGATCTACCTCTTCGTCCCGAAG	Replacement of <i>cgc18</i>
cgc18-dis-T-R	GCGGCCGGCAAGGTGGTGTGCTGATCCCGTGAGCAT CTCCAATCGCGCGCTTCGTTCCGGGAC	
cgc18-SC-FOR	CTCGAGGACACCCGAGAAGC	Verification of <i>cgc18</i> replacement
cgc18-SC-REV	CGTTCGGTGTTCGACTCC	
cmj31F	CCGACGGCCCAAGTTACTGC	RT-PCR <i>cgc1</i> , 307 bp
cmj31R	CCGGGCGATCTGTTTGTTC	
cmj32F	TCCGTCGCCGGTTGAAGTAG	RT-PCR <i>cgc20</i> , 302 bp
cmj32R	GCCGTACATCGCCCTCCATC	
cgc22F	TCGTGTGCCGCGACGAACAG	RT-PCR <i>cgc22</i> , 311 bp
cgc22R	CTCGACGACCTGCTCGACAG	
cmj34F	GTTTCGCACCGTGTTCGGTTC	RT-PCR <i>cgc2</i> , 326 bp
cmj34R	TGCAGCACGCTCTCCTTGTC	
cmj35F	GATGGGCGACACCTTCAGGG	RT-PCR <i>cgc3</i> , 317 bp
cmj35R	GGTGCACGTAGACGCGCTTG	
cmj36F	GACGAACCATGCGGACAACC	RT-PCR <i>cgc19</i> , 303 bp
cmj36R	GGAGTGGCGTTCTCGAAGAGAC	
cmj40F	GACGACGGTGTGCTGGAGTTC	RT-PCR <i>cgc18</i> , 302 bp
cmj40R	TCGACCTTGCCGTTGGGATTC	
cmj43F	GCGCATGGTCATCGTGCAGTAG	RT-PCR <i>cgc7</i> , 333 bp
cmj43R	TTCCGCCGGTTCGAGAACTTCC	
cmj62F	GAAGGAGGCGATGCTGGTGAG	RT-PCR <i>cgc21</i> , 312 bp

Name	Sequence	Description
cmj62R	GCCGAGGTCAACGGCTATGTG	
cmj63F	CTCCGCTCATGTCGCCGAGAAG	RT-PCR <i>cgc4</i> , 292 bp
cmj63R	GGCCAGGTCTGAAGACCGTCAC	
cmj64F	CACCGGCATGCTCAACAGCAC	RT-PCR <i>cgc5</i> , 296 bp
cmj64R	CGGCCAGCTTGACGCTGAAC	
cmj65F	CGTGATCAACGGGCACTACGAG	RT-PCR <i>cgc6</i> , 316 bp
cmj65R	CACGCCCTGGTTGGAGATCTTC	
cmj66F	GACGCCCGGATCCTGCTCTC	RT-PCR <i>cgc8</i> , 313 bp
cmj66R	GGACCCGCCAGGTGTCGTAG	
cmj67F	CCACCTCCTCGACTGGCTCTC	RT-PCR <i>cgc9</i> , 301 bp
cmj67R	CTCGACGAACTGCGGGATCAC	
cmj68F	GTGAAGGTCCAGCCGTTCCC	RT-PCR <i>cgc10</i> , 296 bp
cmj68R	GGTCCCTGGCCGATGATGTG	
cmj69F	CCTGTGGTCCCACCACAAGAAG	RT-PCR <i>cgc11</i> , 285 bp
cmj69R	CAGTCGCCCTCGATGACGTAG	
cmj70F	TGGCCCTGATCGAGGACTGC	RT-PCR <i>cgc12</i> , 306 bp
cmj70R	CGAGCTGGACACGTCCGATG	
cmj55F	CGTCTTCTGGGCCGACTTTG	RT-PCR <i>cgc13</i> , 300 bp
cmj55R	GAGTCCGCGTGGATGATCTC	
cmj71F	GGGACTGAGCGGACTGAAGAC	RT-PCR <i>cgc14</i> , 312 bp
cmj71R	GGCTGGTACGAGCCGAAGATG	
cgc15F	ACGTCGCCGTCCTCAGCAAC	RT-PCR <i>cgc15</i> , 308 bp
cgc15R	CGACTTGACGCGGGAGAATC	RT-PCR <i>cgc15</i> , 308 bp
cgc16F	GGGAACTGGAACGGCTCTAC	RT-PCR <i>cgc16</i> , 319 bp
ccg16R	GCGAGCAGCAGCATGAACAC	
cmj74F	CGTGGTCCTGCTCCCGATCC	RT-PCR <i>cgc17</i> , 298 bp
cmj74R	GGTCAGGTCGGCGATCTCCC	
cAV08	AGGCCACGACTTCATCACG	RT-PCR <i>cgc22-21-20</i> , 2412 bp

Name	Sequence	Description
cAV09	CGCACCCCTTGC GTGCAGTTC	
cAV010	GGCGGACGCAATCTGTCCAC	RT-PCR <i>cgc22-cgc21</i> , 1391 bp
cAV11	GCCTCCGTGCAGCCGTAGAC	
cgc1_dis_F	TCCCAGTCACGACGTTGTAAAACGACGGCCAGTGC CAAGGCGAGGCAAGCTTATCGATG	Replacement of <i>cgc1</i>
cgc1_dis_R	GCAAGGCGGCAGTGAGGAGGAGTACGAGTGAGT CAGACGCACGAGGCCCTTTCGTCTTC	
cgc1_SC_F	AGAAGCTGGGCGAACAAACG	Verification of <i>cgc1</i> replacement
cgc1_SC_R	GCGGATGTTGCGATTACTTC	
SBM21	CCAAGAACAGCATCGACGAG	RT-qPCR, <i>rpoB</i> , 234 bp
SBM22	TGTTCGATGGAGGAGTCGAAG	
AspS_RT_F1	CTGCTGATGATCTCGGGCTT	RT-qPCR, <i>aspS</i> , 104 bp
AspS_RT_R1	TCGAGCTGGTAGAACTCGCC	
GyrA_RT_F1	GGCGACTCCTCCATCTACGA	RT-qPCR, <i>gyrA</i> , 166 bp
GyrA_RT_R1	GGACCATCTCCATCGACAGC	
GyrB_RT_F1	TCGAGACCACCGACTACTCCTT	RT-qPCR, <i>gyrB</i> , 163 bp
GyrB_RT_R1	CTCTTGACCTCGTGCTTCTCG	
cgc2RT-F3	ACGGCACCCCTCGTCCAA	RT-qPCR, <i>cgc2</i> , 66 bp
cgc2RT-R3	GACCGATGGCCCACTTC	
cgc3RT-F3	GATCGCGCCCTCCAACACTAC	RT-qPCR, <i>cgc3</i> , 125 bp
cgc3RT-R3	GAAGGTGTCGCCATCATCA	
cgc4RT-F1	GCCTCGACATGACCGGAGT	RT-qPCR, <i>cgc4</i> , 125 bp
cgc4RT-R1	GCGACATGAGCGGAGAAGAC	
cgc6RT-F2	TGATCAACGGGCACTACGAGA	RT-qPCR, <i>cgc6</i> , 115 bp
cgc6RT-R2	ACCAGGCTCCACCAGCTGAA	
cgc7_RT F1	TCCCCATGAACGCGATGAGT	RT-qPCR, <i>cgc7</i> , 113 bp
cgc7_RT R1	CGAACGACCCGTCCTGAAT	
cgc19RT-F2	CGCGGTGACTTCTTCGAACT	RT-qPCR, <i>cgc19</i> , 77 bp
cgc19RT-R2	ATCCTGCCTATGGCTTGACG	
cgc20RT-F4	CAGGTCTTCGCCAGTCTGCTC	RT-qPCR, <i>cgc20</i> , 110 bp

Name	Sequence	Description
cgc20RT-R4	ACCATCAGGTTTCAGGGACTCG	
cgc21_RT F1	GCCTGGCTGTTCACCTACCTC	RT-qPCR, <i>cgc21</i> , 122 bp
cgc21_RT R1	GTGTCGTGGAACGTCATGTCC	
cgc22RT-F1	TACGGGGAAGTGCACGCAAG	RT-qPCR, <i>cgc22</i> , 245 bp
cgc22RT-R2	CGGAGTCTCGCCTCGTGTG	
LS-RP1	<u>CATATGTCGTACTCCTCCTCACTGC</u>	Amplification of the <i>cgc20-cgc1</i> intergenic region, NdeI site underlined
LS-RP2	<u>GCGGCCGCGGCACTCCTTCTCT</u>	Amplification of the <i>cgc20-cgc1</i> intergenic region, NotI site underlined
LS-RP3	<u>CATATGCGGCACTCCTTCTCTGG</u>	Amplification of the <i>cgc1-cgc20</i> intergenic region, NdeI site underlined
LS-RP4	<u>GCGGCCGCTCGTACTCCTCCTCA</u>	Amplification of the <i>cgc1-cgc20</i> intergenic region, NotI site underlined
LS04	TGCCACAGGGATGCCACAAC	Verification of the CGC101-CGCL106 strains, used with CEA31
LS25	GATGTTACCCGAGAGCTTGG	Verification of the CGC101-CGCL106 strains, used with LS26 or CmJ12R
LS26	GCTGGTCGATACCGCAGTTC	Verification of the CGC101-CGCL106 strains, used with LS25 or CmJ12R
LS63	CATCCGGGTACGACAACGGC	Verification of the CGC101-CGCL106 strains, used with CEA29
CEA29	AAACGAGAACGGGCTCCACTG	Verification of the CGC101-CGCL106 strains, used with LS04
CEA31	ACGCTTGACAGCACGAGAC	Verification of the CGC101-CGCL106 strains, used with LS63

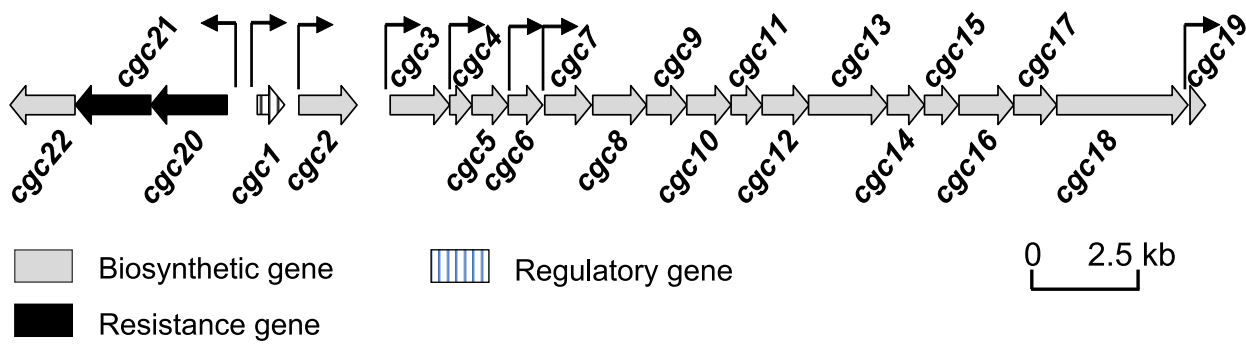
937

938

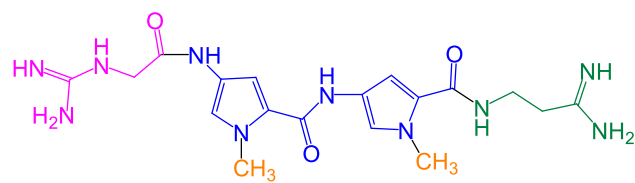
939

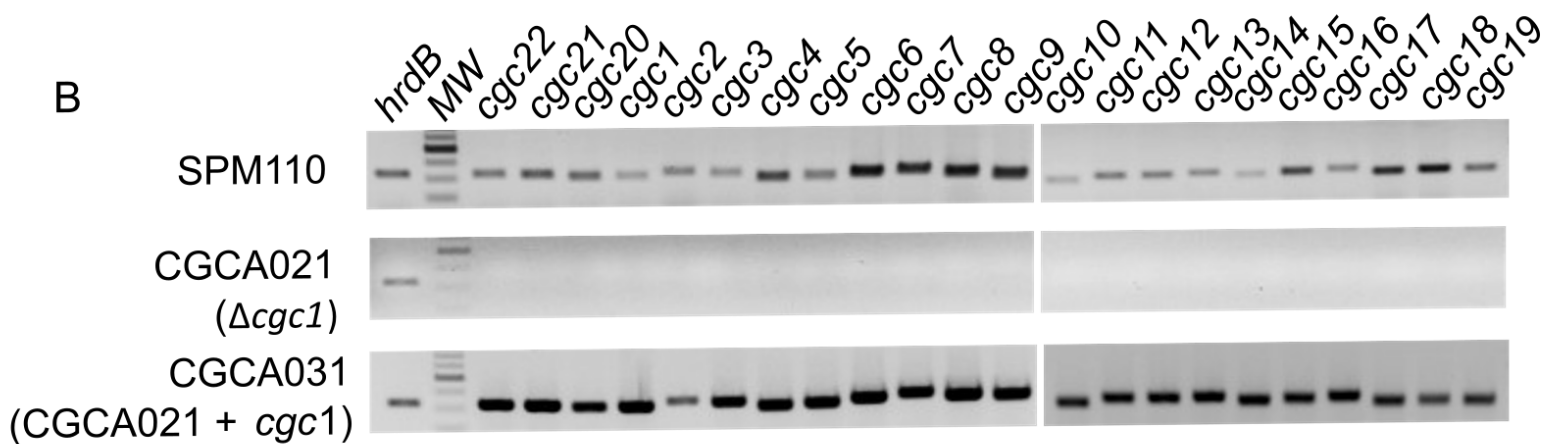
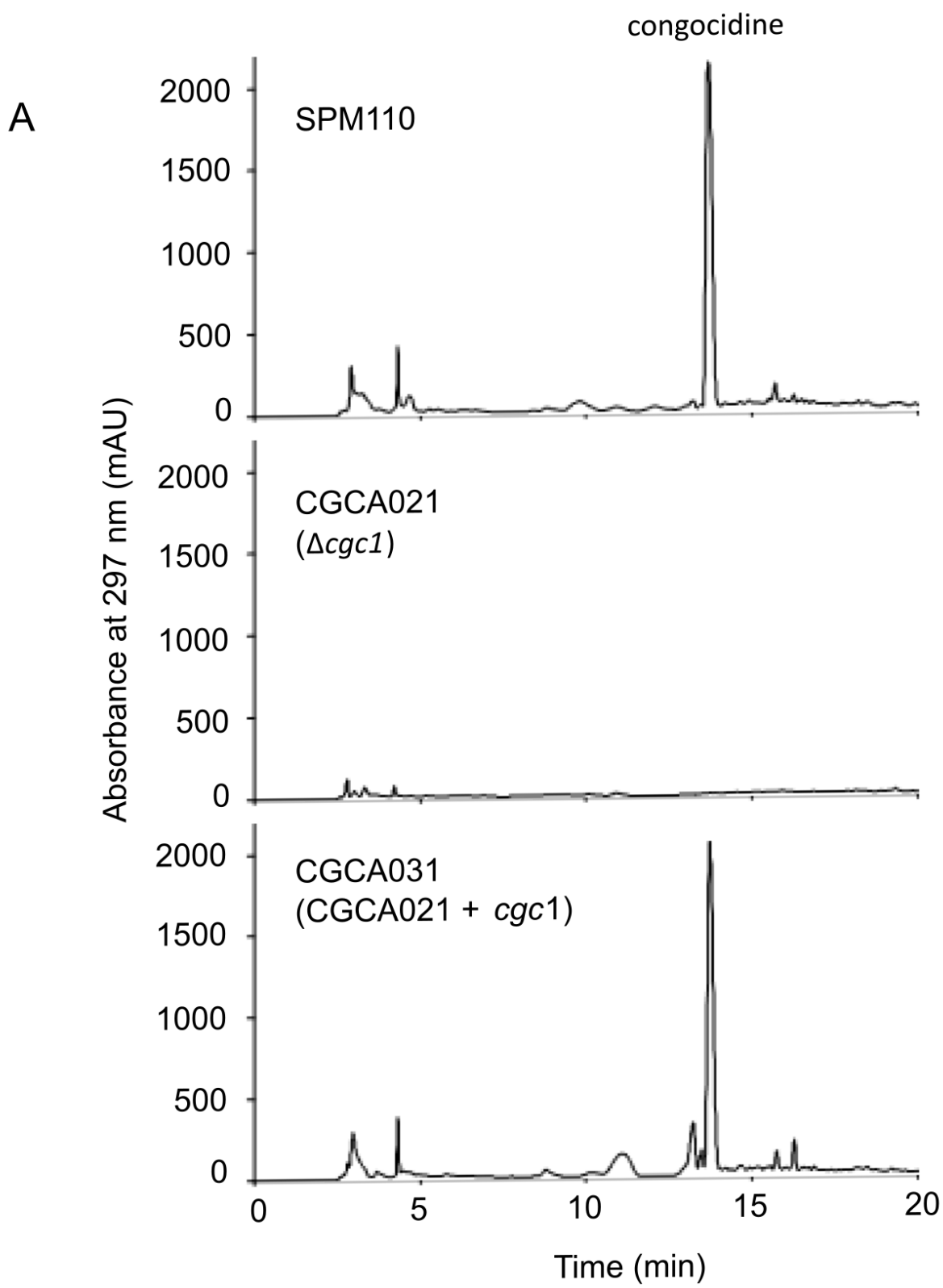
A

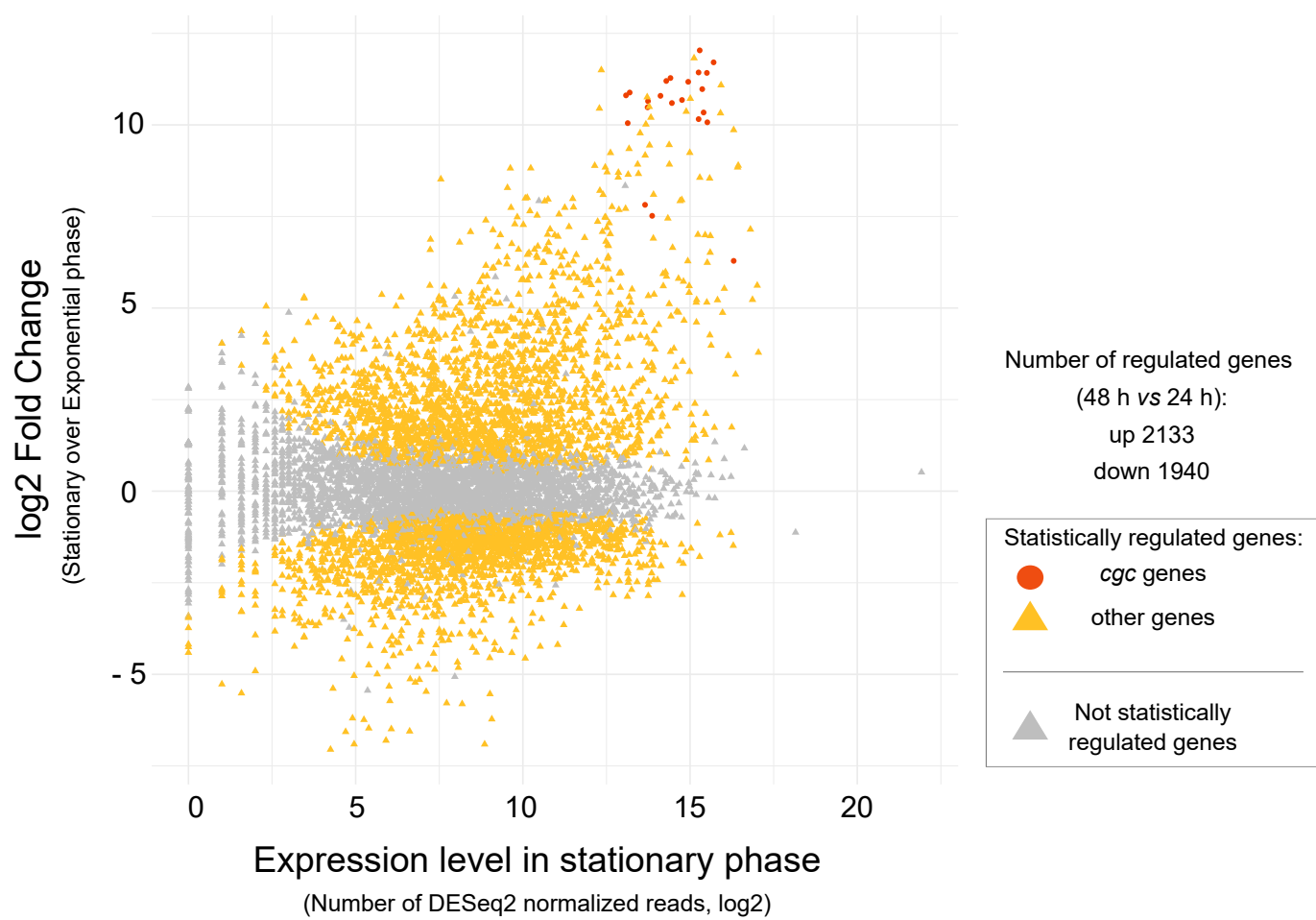
S. ambofaciens



B





A.**B.**