University of Palermo
Department of Biological Chemical and Pharmaceutical Sciences and Technology
(**STEBICEF**)

Streptomyces coelicolor: DNA cytosine methylation and differentiation

Doctoral Thesis of

Annalisa Pisciotta

PhD in Cellular Biology

XXVI cycle

Scientific area code **BIO 19**

Supervisor:
Dr. Rosa Alduina

PhD coordinator:
Prof. Maria Carmela Roccheri

Acknowledgment

I would like to express my sincere gratitude to my supervisor Dr. Rosa Alduina for the continuous support of my Ph.D study and related research, for her patience, motivation, and knowledge. Her guidance helped me in all the time of research and writing of this thesis.

My sincere thank also goes to Prof. Anna Maria Puglia, for funding the BS sequencing of MG medium.

My honest thank also goes to Dr. Angel Manteca, who provided me an opportunity to join his team, who gave access to the laboratory and research facilities and for funding part of this research project.

I thank my lab mates in Oviedo, Dr. Maria Teresa Lopez, Dr Paula Yagüe, Beatriz Gutierrez Beatriz Rioseras, Nathaly Gonzalez and my lab mates in Palermo, Dr. Letizia Lo Grasso for the stimulating discussions and for the help in the progress of this research, and Federica Attardo for the passion with which contributed to development of this project.

Abstract

DNA methylation is an epigenetic modification regulating many aspects of biological processes. DNA cytosine methylation plays mainly a regulatory role in chromatin organization, genome maintenance and gene expression in eukaryotes, while its role in prokaryotes has been less studied.

Streptomyces coelicolor is a mycelial soil microorganism, producer of several antibiotics, with a complex life cycle that includes three different cell types: unigenomic spores, a compartmentalized mycelium (MI) and a multinucleated mycelium (substrate, aerial and sporulating, MII). This life cycle is finely regulated through several mechanisms: two events of programmed cell death, PCD, and three biochemical pathways (*bld* cascade, *sky* pathway and *whi* cascade).

The importance of DNA methylation has been already described in Streptomycetes, but its biological role remains unknown.

The aim of this project is to investigate the relationship between DNA cytosine methylation and morphological and physiological differentiation in *S. coelicolor*.

Dot blot analysis of genomic DNAs extracted from *S. coelicolor* grown in different media (two liquid, MG and R5A, and a solid one, GYM) revealed that the global level of methylated cytosines changes along the growth during development (MI, MII and spores) both on solid and in liquid culture.

To describe the extent of cytosine methylation, sequencing of bisulfite-treated genomic DNA, extracted from *S. coelicolor* along the growth in all the three media, was performed and the corresponding cytosine methylomes were defined.

Bioinformatic analysis of DNA cytosine methylome in liquid MG revealed that 30% of *S. coelicolor* genes contain two cytosine methylation consensus sequences ($GCC^{m}CG$ and $GGC^{m}CGG$) in their upstream regions. Among these, genes involved in morphological and

physiological differentiation, genes encoding putative transcriptional regulators and sigma factors, and genes related to DNA replication and chromatin condensation were identified.

DNA cytosine methylome in R5A contains 28% of *S. coelicolor* genes methylated. These genes contain three cytosine methylation consensus sequences (GCC^{m}CG, GGC^{m}CGG and C^{m}GGGC) in their regulation region; among these genes, 58.8% was methylated during the MI phase, such as genes involved in antibiotic production and morphological differentiation.

DNA cytosine methylome in GYM is constituted by 28.8% of *S. coelicolor* genes methylated. These genes contain the same three cytosine methylation consensus sequences (GCC^{m}CG, GGC^{m}CGG and C^{m}GGGC) in their regulation region as those found in R5A; out of these genes, 63.9% was methylated during the MI phase, such as genes involved in antibiotic production and morphological differentiation.

To study the effect of DNA cytosine methylation, liquid and solid cultures of *S. coelicolor* were treated with 5-aza-2'-deoxycytidine (aza-dC, a cytidine analogous that inhibits DNA-methyltransferase activity). This treatment influenced *S. coelicolor* germination and antibiotic production both in liquid and on solid culture and sporulation on solid culture.

Altogether these results demonstrate a strong relationship between DNA cytosine methylation and *S. coelicolor* growth and differentiation.

Combining methylome results and transcriptomic analysis, a correlation between the cytosine methylation consensus sequences and gene expression was found in R5A and GYM media. Indeed, the GGC^{m}CGG consensus sequence was found to be methylated in MI phase in a set of genes (25%), whose transcription was repressed in MI phase, while this consensus sequence was found twice methylated in all the phases in a set of genes (7.5%), whose transcription was constitutive. Regarding the GCC^{m}CG consensus sequence, it was found to be methylated in MII phase in a set of genes (13.5%) that were more transcribed in MII phase, differently C^{m}GGGC consensus sequence was found to be methylated in MII phase in a set of genes (20%) whose transcription was repressed in the same phase.

The search for DNA methyltransferase genes into *S. coelicolor* genome revealed that it contains 17 genes coding for putative DNA (5-cytosine)-methyltransferases; among them, *SCO1731* is mainly expressed in MI phase, both on solid and in liquid culture. Thus, an independent mutant of *SCO1731* gene was generated to start the study of cytosine demethylation. Phenotypic analysis of the mutant showed the same results obtained after aza-dC treatment, indeed, the morphological and physiological differentiation was delayed on solid GYM and in liquid R5A.

The results of this study demonstrate that DNA cytosine methylation is related to morphological and physiological differentiation in *S. coelicolor*.

Introduction

1. DNA methylation

The term epigenetic is defined as “not genetic, but heritable through DNA replication” and is used to distinguish among three modes of DNA methylation: (I) genetic methylation, for example, in the biosynthesis of dTMP from dUMP, with subsequent incorporation into DNA by the replication machinery; (II) epigenetic methylation, as in C^5-methylcytosine (m5C), N^4-methylcytosine (m4C) and N^6-methyladenine (m6A) which are inherited by maintenance methylation after DNA replication; and (III) non-genetic and non-epigenetic methylation, e.g. O^6-methylguanine. The non-epigenetic and non-genetic DNA methylation in O^6-methylguanine is known to trigger cell death (1). Another class of DNA modification involves the use of a base other than A, T, G and C in DNA. For example, some bacteriophage genomes carry hydroxymethylcytosine instead of cytosine, and dUMP is often incorporated in place of dTMP into some bacteriophage genomes.

DNA methylation is the best characterized epigenetic mechanism that regulates multiple processes in prokaryotes and eukaryotes (2).

In eukaryotes, epigenetic DNA methylation plays roles in chromatin organization, gene expression, and genome maintenance, and its disturbance is related to human diseases (3). In prokaryotes, it is crucial for processes, including cell-cycle regulation, transcriptional regulation, and host-pathogen interaction. It is involved in silencing selfish genetic elements and other aspects of intra-genomic conflicts in eukaryotes and prokaryotes.

1.1 DNA methylation in eukaryotes

In eukaryotes, DNA methylation occurs mainly in cytosines, in particular in CpG islands that are in the promoter region of genes; m5C is typically associated with the regulation of gene expression, whereas gene body DNA cytosine methylation is often correlated with active gene transcription.

In the case of CpG islands containing promoters, the lack of methylation is usually associated with the chromatin pattern of actively transcribed genes, as characterized by an opened nucleosome configuration, reduces amount of histone H1 and presence of acetylated histones (4). The ability of methylation to silence genes with CpG islands was studied on inactivated genes on X-chromosome. Transfection studies showed that this silencing is mainly a result of chromatin condensation which makes DNA less accessible for transcription factors (5). The role of single methyl groups, preventing binding of specific factors, appears to be less important in this case.

In contrast to this, genes without CpG islands are dependent on the methylation of single sites within their promoter regions. This observation can be explained by the property of some transcription factors whose binding to DNA is methylation dependent, i.e. the protein binds to its binding site only in the case of unmethylated DNA.

DNA cytosine methylation is conserved in most major eukaryotic groups, including plants, many fungi and animals. In plants, DNA cytosine methylation occurs in C^5-methylcytosine (m5C) in the contexts of CG, CH[1]G and CHH (6); in mammals, it is restricted to the symmetric CG context (as describe above), although non-CG methylation is prevalent in embryonic stem cells (7).

The rate of DNA cytosine methylation differs strongly between species: 14% of cytosines are methylated in *Arabidopsis thaliana*, 4% in *Mus musculus*, 0.03% in *Drosophila*, and virtually none (< 0.0002%) in yeast species (8).

On the contrary, DNA adenine methylation is low in eukaryotes, and its function remains unknown (9), probably it is involved in the regulation of gene expression and replication of mitochondrial DNA (10).

1.2 DNA methylation in prokaryotes

In prokaryotes, DNA methylation occurs mainly in N^4-methylcytosine and N^6-methyladenine.

[1] H=A,C or T

N^4-methylcytosine has been associated with the restriction–modification (RM) systems (11) and N^6-methyladenine has been associated with (i) gene expression, (ii) DNA repair and (iii) coordination of replication and cell division (12).

1.2.1 Restriction-modification system

In prokaryotes, many epigenetic DNA methyltransferases are paired with a restriction enzyme. Restriction enzymes are DNA endonucleases that recognize specific DNA sequences and introduce a double-strand break. This activity restricts establishment of invading DNAs that lack proper DNA methylation, such as bacteriophage DNA genomes, plasmids, and DNA fragments delivered through natural transformation machinery. The potentially lethal cleavage of cellular DNA in cells that harbour a restriction enzyme is prevented by epigenetic DNA methylation by the cognate DNA methyltransferase that recognizes the same sequence as the restriction enzyme. Genes encoding the restriction enzyme and the methyltransferase are often located next to each other and form a unit called a restriction-modification system. Restriction-modification systems are classified into four types, Type I, II, III, and IV, based on their genetic and biochemical characteristics (1):

- Type I restriction enzymes are composed of three subunits, S, M, and R. The S subunit recognizes a specific DNA sequence. A complex of M and S subunits exhibits methyltransferase activity at the recognition site. The joining of the R subunit to this complex is essential for endonuclease activity. After binding to an unmodified recognition sequence, the restriction enzyme complex translocates DNA towards itself from both directions in a reaction coupled to ATP hydrolysis. Type I restriction modification enzymes have two modes of action that are controlled by the methylation state of their recognition sequence. If the sequence is fully methylated, the enzyme complex does not bind. When the sequence is hemi-methylated, the methyltransferase complex catalyzes an efficient methyltransfer reaction to the other strand. When the

sequence is unmethylated, the restriction enzyme complex is formed and translocation begins, leading to cleavage.

- Type II restriction enzymes bind to a recognition sequence and cleave DNA in their vicinity and are frequently used in DNA engineering. For example, *EcoRI*, *BamHI*, and *PvuII* are Type II enzymes. Many variants are classified into subtypes within this type, based on biochemical characteristics. In many subtypes, restriction activity is present in one enzyme molecule, whereas modification activity is present on the other. Restriction enzymes in this class are divergent in amino acid sequence and three-dimensional structure and can be also classified based on these features.
- Type III restriction enzymes are composed of two subunits: Mod (for modification) and Res (for restriction). The Mod subunit has DNA methyltransferase activity, and the Mod–Res complex has restriction activity. When the restriction enzyme complex binds to an unmethylated site, it cleaves DNA through interaction with another restriction enzyme complex on the same DNA. This process is dependent on ATP hydrolysis. The cleavage mechanism is not clear yet (1), although diverse and sometimes mutually contradictory models have been proposed.
- Type IV systems contain a class of enzymes that cleave DNA only when the recognition site is methylated. In *Escherichia coli*, McrA, McrBC, and Mrr are enzymes of this class that show different restriction spectra. McrBC is a methyl-specific DNase, *in vitro* it recognizes and cleaves the DNA. Although McrA and Mrr are believed to be endonucleases, their DNA cleavage activities have not been observed *in vitro*.

  The type IV, or methylation-dependent restriction, enzymes have a behavior that is opposite to that of the other types, in that they cleave only when bases within the recognition sites are methylated. The bacteria encoding these enzymes do not have a methyltransferase associated with the restriction enzyme.

1.2.2 Gene expression

Computational analysis on the *E. coli* genome revealed that GATC sequences accumulate more frequently in promoter regions, than in coding regions (13). Whole genome analysis of gene expression in *E. coli* strains showed that many genes are mis-regulated in the absence of Dam (DNA adenine methyltransferase). GATC sequences were found in the promoter regions of most genes, suggesting that adenine methylation directly regulates gene transcription.

One of the gene regulated by DNA adenine methylation is *dnaA*, which codes for an essential factor for the initiation of replication from the chromosomal origin. The *dnaA* promoter region contains six GATC sites in *E. coli* and two GATNC sites in *Caulobacter crescentus.* In both bacteria species, the *dnaA* promoter is more active when it is in a fully methylated state, than when it is in a hemimethylated state. Since the *dnaA* gene is located next to the origin of replication, in *E. coli* and *C. crescentus*, it is immediately replicated and is hemi-methylated after the initiation of DNA replication. Using this mechanism, the transcription of *dnaA* is switched down immediately after DnaA has initiated DNA replication, with the DNA replication cycle. In *C. crescentus*, the *dnaA* promoter is maintained hemimethylated until CcrM re-accumulates in pre-divisional cells. In *E. coli*, the *dnaA* promoter region is protected from Dam, mediated methylation that directly binds to the hemimethylated *dnaA* promoter region. The reason why the hemimethylated *dnaA* promoter region is less active for transcription than fully methylated region is not clear yet.

Other two genes in *C. crescentus* are regulated by DNA methylation: *ctrA* and *ccrM*.

Like DnaA, the CtrA protein is a dual function protein that binds to the origin of the replication and other promoter of *C. crescentus*. CtrA inhibits the initiation of DNA replication and the cellular levels of CtrA are cell cycle regulated. The transcription of *ctrA* is temporally controlled by two promoters, one of which, the *ctrAP1*, contains one GANTC sequence and is regulated by DNA methylation. Indeed, the activity of the *ctrAP1* is stimulated by hemimethylation after the

passage of the replication fork through *ctrAP1*. The *ctrAP1* is also repressed by CtrA, suggesting that binding of CtrA could be affected by methylation.

DnaA and CtrA have opposite functions in regulating the initiation of DNA replication and their transcriptional regulation are mirror images: *dnaA* is activated by methylation, while *ctrA* is repressed by methylation.

Like *ctrA*, transcription from the *ccrM* promoter is repressed by CcrM. The *ccrM* promoter is also activated by CtrA, and CtrA-binding is located between two pairs of GANTC sites. One of these pairs repressed *ccrM* transcription (2).

Differential methylation of DNA is part of the basis for immune-stimulation of macrophages by bacterial DNA but not by DNA from vertebrates. The effect of methylation on gene regulation in bacterial pathogens, like *Haemophilus influenzae*, results in increased or decreased levels of gene expression depending on the site of methylation (14).

1.2.3 DNA repair

During DNA replication, mismatched base repair inevitably occur, and their repair requires discrimination between the template strand and the newly synthesized, error-prone strand. This information is provided by transient lack of adenine methylation in the newly synthesized DNA strand. After recognition of a mismatched base repair by MutS protein, a complex involving MutS, MutL and MutH is formed (MMR complex) (Figure 1).

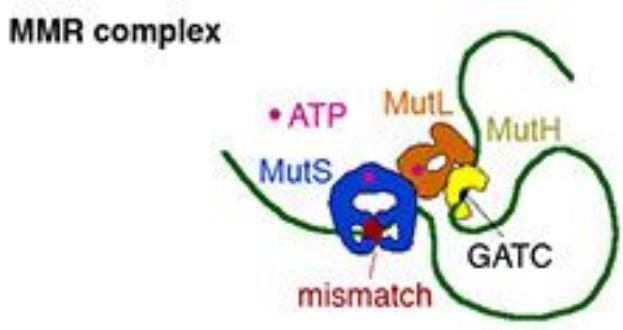

Figure 1 Assembly of the MMR complex.

When part of the ternary complex is assembled at a DNA mismatch, MutH, a GATC-specific endonuclease, cleaves the 5'-P to the G nucleotide in the non-methylated, newly synthesized DNA strand. Transient GATC hemi-methylation in the newly synthesized strand thus provides the signal for strand discrimination by MutH. The resulting 3'-end is a substrate for a helicase and for exonucleases which degrade the daughter strand that erroneously incorporated nucleotide. Re-synthesis of the gap by DNA polymerase III and phosphodiester bond formation by ligase then follow.

1.2.4 Coordination of replication and cell division

The role of adenine methylation in cell cycle of *Caulobacter crescentus* was deeply studied (Figure 2).

The life cycle involves two cell types, a stalked cell, that has a tubular stalk structure, and a swarmer cell, that has a single flagellum that provides swimming mobility. Chromosome replication occurs once and only once per cell cycle. The cyclic developmental program begins with a stalked cell, in which the chromosome is fully methylated, at the replication origin (*oriC*) containing GANTC sites recognized by DnaA. As the cell grows and replicates its DNA, it becomes a pre-divisional cell with hemi-methylated DNA molecules. In the late pre-divisional stage, a flagellum is formed at the swarmer cell pole. The production of CcrM, a methylase adenine specific for the GANTC sequence, is followed by methylation of the daughter chromosome. During this time the swarmer cell becomes incompetent for DNA replication, in fact the DNA is fully methylated and CcrM protein is degraded, in addition CtrA binding to the chromosomal origin of replication inhibits DNA replication initiation, as DnaA can not bind the replication origin (15). Cell separation leads to two different cell types: one cell is a stalked cell which reenters the cyclic developmental program, it contains hemi-methylated DNA (16), the other one is a swarmer cell that differentiates into a stalked cell. This differentiation comprises the noncyclic developmental program.

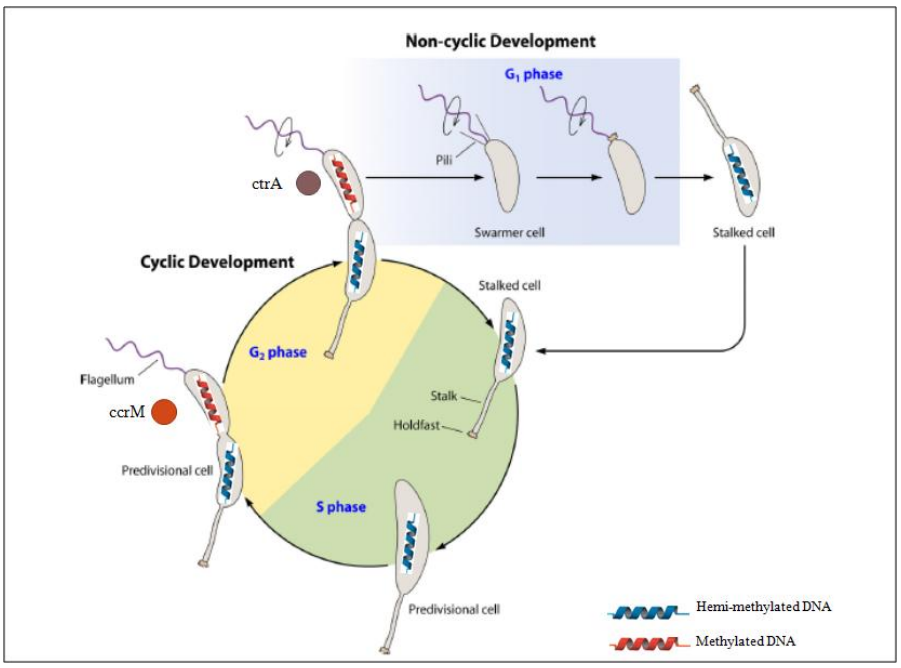

Figure 2 In *C. crescentus*, asymmetric cell division produces a non-replicating swarmer cell and a replicating stalked cell. DNA replication, in cyclic development, is only initiated on the methylated *Cori* of the stalked cell. Modified from Kodzon *et al.* 2013.

1.2.5 Cytosine methylation

Relatively recent studies suggest that C^5-methylcytosine may have physiological roles including regulation of gene expression. In *Helicobacter pylori*, lack of an orphan m5C-methyltransferase alters the expression of genes involved in motility, adhesion and virulence (17).

In *E. coli*, some years ago, it was performed high-throughput sequencing of bisulfite treated genomic DNA (18) to describe, for the first time, the DNA cytosine methylation in *E coli* represent 2.3% of genome. Whereas most target sites (C^mCW^2GG) were fully methylated in stationary phase cells, many sites (CC^mCWGG) were only partially methylated in exponentially growing cells.

Moreover, in *E. coli* the loss of *dcm* (DNA cytosine methyltransferase) causes an increase in gene expression of several categories of genes. Thus, it was demonstrated that DNA cytosine methylation is a regulator of transcription, especially during the stationary phase (19-20).

[2]W=A or T

Experimental alteration of epigenetic DNA methylation system in prokaryotes can cause a variety of changes, for example, an aberrant cytosine methylation pattern leads to death. In *E. coli* an induced ipermethylation level of cytosines was reported to trigger cell death through chromosomal cleavage; when a cytosine methylation system enters the cells (or becomes activated) and begins to methylate cytosines in the host genome, McrBC, a methyl-specific DNA endonuclease (type IV restriction enzyme) senses these epigenetic changes and triggers cell death through chromosomal cleavage (Figure 3). Cell death through cleavage of the chromosome at the level of methylated cytosines occurs upon entry or induction of a methyltransferase gene and aborts its establishment or activation.

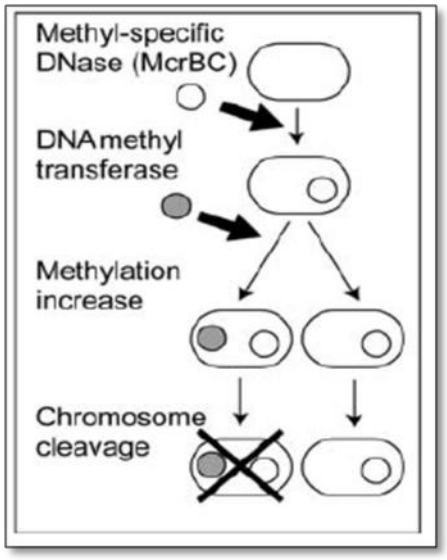

Figure 3 When a DNA cytosine methylation system enters a cell and begins to methylate chromosomal recognition sites, McrBC senses the change and triggers cell death by chromosomal cleavage. From Fukuda *et al.*, 2008.

1.3 Bacterial methylomes

To study the epigenetic modification it is necessary to use a method that reveals the complete methylation pattern. Detection of C^5-methylcytosine by bisulfite genomic sequencing is a method used in eukaryotes to characterize the cytosine methylome; in the last few years, it has been extended to study the bacterial cytosine methylome. Treatment of DNA with bisulfite

converts cytosine residues to uracil, but leaves C^5-methylcytosine residues unaffected. Thus, bisulfite treatment introduces specific changes in the DNA sequence that depends on the methylation status of individual cytosine residues. The limit of this technique is that the sequencing can be carried out only for known bacterial genomes to compare the DNA of the sequenced sample to its reference sequence.

Another technique to study the methylation state of nucleotides N^6-methyladenine, C^5-methylcytosine and N^4-methylcytosine is the single-molecule real-time (SMRT) sequencing (19) that monitors in real time the activity of single DNA polymerase molecules that use fluorescent nucleotides to synthesize DNA complementary to a template. This technique can be carried out also for *de novo* genome sequencing.

An example of identification of bacterial methylomes by SMRT was performed by Murray *et. al* (20), the methylome of six bacteria (*Geobacter metallireducens* GS-15, *Chromohalobacter salexigens*, *Vibrio breoganii* 1C-10, *Bacillus cereus* ATCC 10987, and two species of *Campylobacter jejuni subsp*.) was characterized. In *G. metallireducens* GS-15 the sequences 5'-G^{m6}ATCC-3' and 5'-TCC^{m6}AGG-3' were found and correspond to 99% of all consensus sequences. In *C. salexigens*, 5'-CC^{m6}AC(N[3])$_6$CTC-3' and 5'-R[4]G^{m6}ATCY-3' were found and correspond to 98% and 76% of all consensus sequences, respectively. In *V. breoganii* 1C-10, 5'-AGH[5]^{m6}A(N)$_7$TGAC-3', 5'-CT^{m6}AG(N)6RTAA-3' and 5'-G^{m6}ATC-3' were found and correspond to 97% of all motifs.

In *C. jejuni* subsp. Jejuni 81-176, 5'-RA^{m6}ATTY-3', 5'-GCA^{m6}AGG-3', 5'-GGRC^{m6}A-3', 5'-CAm6AYN6ACT-3' and 5'-TA^{m6}AYN5TGC-3' were found and correspond to 97% of all consensus sequences. In *C. jejuni* NCTC 11168, 5'-TA^{m6}AYN$_5$TGC-3', 5'-RA^{m6}ATTY-3', 5'-G^{m6}AGN$_5$GT-3' and 5'-GKA^{m6}AYG-3' were found and correspond to 98% of all consensus

[3] N: any base
[4] R: A or G
[5] H: A, C or T

sequences. In *B. cereus,* 5'-TA^{m6}AGN7TGG-3', 5'-A^{m4}CGGC-3' and 5'-G^{m5}CWGC-3' were found and correspond to 93%, 34% and 93% of all consensus sequences, respectively.

Combined with a genome-wide analysis of transcription in the wild type and in a DNA methyltransferase mutant, identification of unmethylated DNA targets can spot transcriptional units under putative DNA methylation control.

2. Streptomycetes

Streptomycetes are Gram positive soil bacteria with CG rich genomes (70%), known as the most prolific producers of naturally occurring bioactive molecules, i.e. most antibiotics in current therapeutic use (21).

Streptomyces coelicolor A(3)2 strain M145 is the best-known species of genus streptomycetes at the genetic and molecular level (22) and has been considered the model streptomycete for studying physiological (antibiotic production) and morphological differentiation. It produces four types of antibiotics: undecylprodigiosin (red pigment, Red), actinorhodin (blue pigment, Act), the calcium-dependent lipopeptide antibiotic (CDA) and the cryptic polyketide with antibacterial activity (yellow pigment, Cpk).

The classical *S. coelicolor* developmental model, growth in liquid medium MG, is characterized by two rapid growth phases, a transition and a stationary phase (Figure 4). During the first rapid growth phase (RG1) the cells divide quickly. The transition phase (T) is characterized by a growth arrest, transitional arrest of macromolecular biosynthesis and differentiation, turnover of ribosomal proteins and beginning of secondary metabolite synthesis, like antibiotics. During the second rapid growth phase (RG2) an increased secondary metabolite production occurs; in the stationary phase (S) cellular growth is stopped but the antibiotic production continues (23).

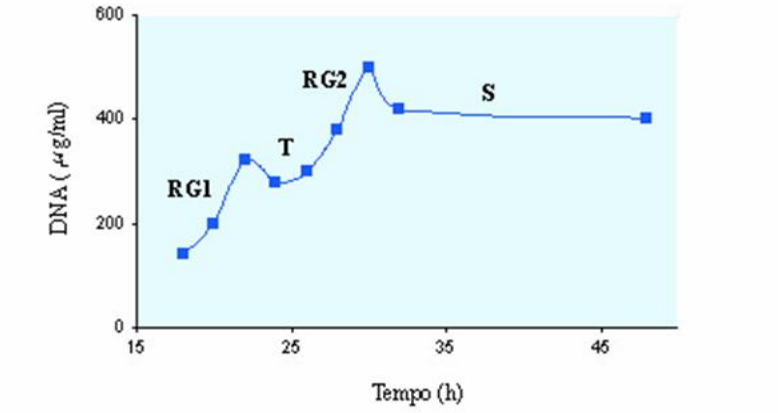

Figure 4 Growth curve of *S. coelicolor* in liquid medium MG. From Puglia *et al.* 1995.

The classical developmental model for confluent growth, on rich solid R2YED, assumed that differentiation started with germination and then completely viable vegetative mycelia (substrate) grow on the surface and inside agar until they undergo a programmed cell death (PCD), followed by hyphae differentiation into a reproductive (aerial) mycelium characterized by the presence of hydrophobic covers. Substrate and aerial mycelia are multinucleated, but at the end of the cycle, aerial hyphae form septa and spore chains (24).

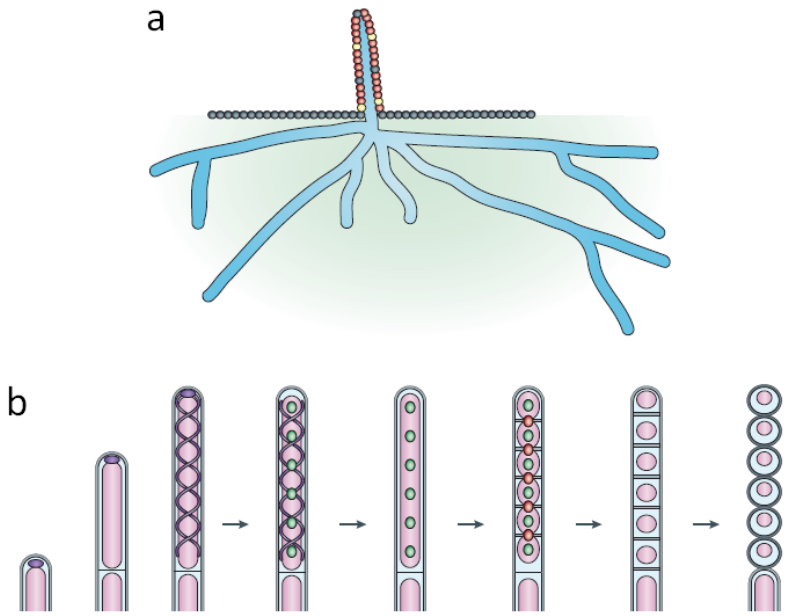

Figure 5 a) Hydrophobic cover formation during growth in solid cultures; b) Chromosome segregation from aerial hyphae to spore chains. Modified from Flärdh and Buttner, 2009.

On solid culture (i.e. GYM) viable vegetative mycelium is produced by the germination of a spore, it grows on the surface and inside agar forming MI (first compartmentalized mycelium); MI undergoes a highly ordered PCD and the remaining viable segments of these hyphae begin to enlarge in the form of MII (second multinucleated mycelium), in this stage it is possible to distinguish MII substrate grown inside agar, and MII aerial characterized by the presence of

hydrophobic layers which precedes the stage of MII sporulating, that undergoes a second round of PCD followed by spore formation (25).

The traditionally denominated "substrate mycelium" corresponded to MII coated with these layers.

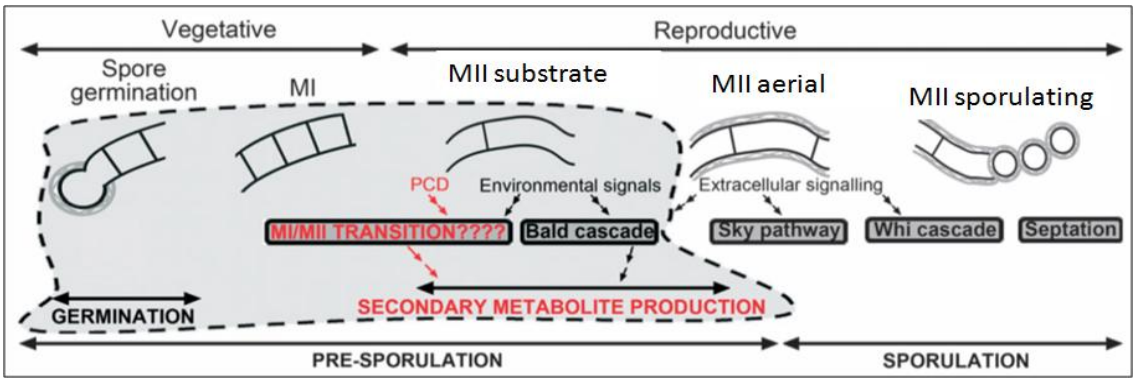

Figure 6 Biochemical pathways regulating *Streptomyces* differentiation. Pathways involved in hydrophobic covers formation ('bald', 'sky') and sporulation ('why', 'septation') are illustrated. Development stages (MI/MII) and presporulation pathways (MI/MII transition) switching on secondary metabolite production are indicated in red. Modified from Yagüe *et al.* 2013.

Bacterial programmed cell death can be defined as any type of genetically controlled cell dismantling involving the activation of specific cell death transducers, regulators and effectors (26). PCD was described in bacteria from different taxa, such as *Bacillus, E. coli*, *Caulobacter* (27), *Streptococcus* (28) and *Streptomyces*, it is activated in response to cellular stress or damage or in response to developmental signals. With a few exceptions, the biochemical pathways controlling bacterial PCD are poorly understood and there is no general biochemical model applicable to all bacterial PCD.

Life cycle of Streptomyces is regulated by different pathways and mutant strains defective in different stages of development were used: (i) the ''bald'' (*bld*) mutants (considered defective in aerial growth) are affected in genes that regulate the so-called ''sky-pathway'' and activate the expression of genes encoding proteins forming hydrophobic covers (*rdl, chp, ram*); (ii) the

‘‘white’’ (*whi*) mutants (defective in the formation of mature grey spores on the tips of the white, fluffy aerial mycelium) affected in genes that activate (whose mutants are hyphae septation and sporulation. In contrast to aerial mycelium and sporulation, *Streptomyces* pre-sporulation stages (germination and MI/MII transition) had been poorly studied. Spore germination was proven to include a succession of distinctive steps. Hardisson (29) organized these steps nicely into three stages: darkening, swelling and germ tube emergence. Darkening merely requires the presence of exogenous divalent cations (Ca^{2+}, Mg^{2+} or Fe^{2+}) with energy being obtained from spore reserves; swelling needs an exogenous carbon source, and germ tube emergence requires additional carbon and nitrogen sources.

A similar morphological differentiation occurs also in some liquid media (i.e. R5A) (30), in which the MI and the MII are present (Figure 7). The MII emergence is preceded by a transient growth arrest, which is the consequence of MI PCD. MII is the antibiotic-producing mycelium (25). The lifespan of MI in liquid cultures is longer than on solid media (around 48h vs 18h).

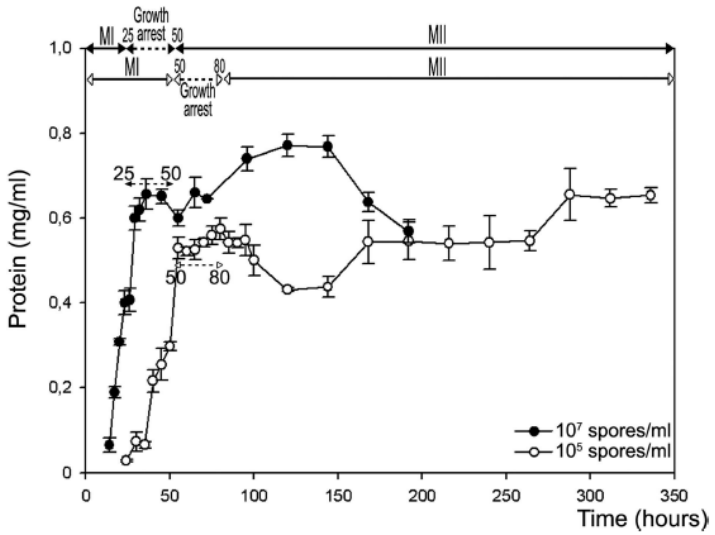

Figure 7 Growth curve of *S. coelicolor* in liquid medium R5A. From Manteca *et al.* 2008.

2.1 *Streptomyces coelicolor* transcriptome on solid GYM and in liquid R5A

Transcriptomic analysis was performed by RNA extracted from *S. coelicolor* growth on solid GYM and in liquid R5A, demonstrated that most genes had similar expression patterns in all the MII stages analyzed (MII24h ‘‘substrate’’; MII48h ‘‘aerial’’; MII72h ‘‘sporulating’’), the same genes were up- or down-regulated in MII with respect to MI (31).

The genes involved in primary metabolism, such as genes involved in oxidative phosphorylation, in glycolysis and gluconeogenesis, coding for ribosomal proteins and sporulation regulators, genes related to Cpk production and conjugation, recombination, or mutagenesis were up-regulated in MI (Table 1).

The genes involved in hyphal differentiation, such as activators of aerial mycelium differentiation, genes involved in the formation of hydrophobic covers, sporulation regulatory genes, genes involved in Act, Red and CDA cluster were up-regulated in MII.

Another secondary metabolite produced by *S. coelicolor* is the odorant geosmin. The expression of one of the genes involved in its biosynthesis (*geoA*, SCO6073) was seen to be up-regulated in MII with respect to MI.

Interestingly, all the genes encoding transposons and insertion sequences were revealed to be up-regulated in MII. The opposite occurred with genes involved in conjugation, recombination, or mutagenesis, which were up-regulated in the MI phase (Table 1).

Table 1 Genes up-regulated in MI and MII phase grouped in functional categories.

Genes up-regulated in:	
MI	**MII**
Oxidative phosphorylation (SCO0924; SCO3945; SCO3946)	Aerial mycelium differentiation (*bldN*, SCO3323; *bldC*, SCO4091; *bldM*, SCO4768; *bldkB*, SCO5113; *bldB*, SCO5723)
Glycolysis and gluconeogenesis (SCO1947 and SCO7511)	Formation of hydrophobic covers (*sapA*, SCO0409; *chpC*, SCO1674; *chpH*, SCO1675; *chpE*, SCO1800; *chpF*, SCO2705; *chpD*, SCO2717; *rdlA*, SCO2718; *ramC*, SCO6681; *ramS*, SCO6682; *ramA*, SCO6683);
Sporulation regulatory genes (*whiJ*, SCO4543)	Sporulation regulatory genes (*wblA*, SCO3579; *wblC*, SCO5190; *wblE*, SCO5240; *whiE*, SCO5316; *whiE protein II*, SCO5319; *whiG*, SCO5621; *whiH*, SCO5819)
Cpk production (SCO6273, SCO6274, SCO6275, SCO6276, SCO6277, SCO6278, SCO6279, SCO6280, SCO6281, SCO6286, SCO6288);	Act and Red production (*actV*, SCO5077; *actII-4*, SCO5085, *redF*, SCO5898)
Ribosomal proteins (*rplI*, SCO3909)	Geosmin production (*geoA*, SCO6073)
Conjugation, recombination, or mutagenesis	Encoding transposons and insertion sequences

3. Methylation in Streptomyceteae

S. coelicolor has a type IV restriction-modification system with enzymes that cleave methylated DNA, but the endonucleases responsible for this activity have not been found yet (32). Up to now, nothing is known regarding the level of DNA methylation in *S. coelicolor*.

The role of DNA methylation was studied in Streptomyceteae but its function remains unknown. Preliminary studies on DNA methylation were carried out in *Streptomyces antibioticus* (33).

Two different types of drugs to study the effect of DNA demethylation were used. *S. antibioticus* solid culture, treated with 1mM of sinefungin, slight delayed growth and blocked sporulation when the drug was added at the beginning of the incubation. When sinefungin was added to cultures after young mycelium was formed did not inhibit sporulation, suggesting that the compound might enter the cells only at the stage of germination.

In addition, *S. antibioticus* solid culture was treated with 1mM of 5-azacytidine. The drug added at the beginning of incubation did not affect growth, but blocked sporulation (34).

Surprisingly, rhodomycin, which was not produced in control cultures, was synthesized at early stages when 5-azacytidine was present. Some stimulatory effect of 5-azacytidine was also observed on the production of actinorhodin by *S. coelicolor* grown under analogous conditions (35).

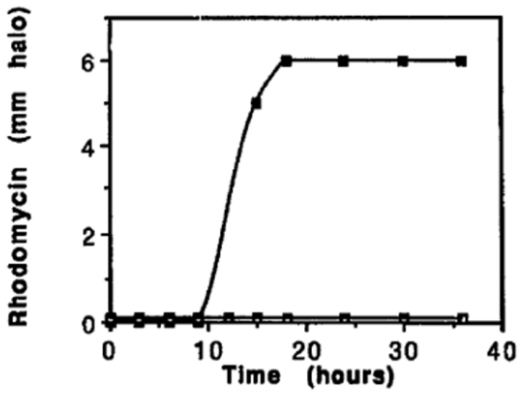

Figure 8 Effect of 5-azacytidine on rhodomycin production during growth of *S. antibioticus*. From Novella *et.al* 1995.

5-azacytidine can be transported into both bacterial and mammalian cells; after its conversion to a nucleotide and it is incorporated into RNA and DNA of treated cells. After incorporation, the drug inhibits the corresponding cytosine methyltransferases through the formation of a covalent complex with the DNA which leads to enzyme inactivation and as a consequence to demethylation.

These works demonstrated a specific effect of 5-azacytidine on physiological and morphological differentiation of *S. antibioticus,* but they did not explain the relationship between methylation and differentiation.

Aim

The aim of this PhD project was to study the DNA cytosine methylation of *S. coelicolor* and to understand whether it can be considered as an epigenetic modification related to its physiological and morphological differentiation.

The DNA cytosine methylation level was investigated in different media

- the defined liquid medium MG (used for a preliminary analysis of DNA cytosine methylation during growth);
- the rich liquid medium R5A;
- the rich solid medium GYM.

Liquid R5A and solid GYM media were chosen since the developmental cycle of *S. coelicolor* was deeply investigated and results of the transcriptomic analysis are available. Thus, the correlation between morphological/physiological differentiation, gene expression and cytosine DNA methylations is expected to be easier and straightforward.

For all of them, the same approach was used:

1. the global level of methylated cytosines was investigated by Dot Blot analysis along the growth and by sequencing of bisulfite-treated genomic DNA extracted from specific growth phases. Both analyses revealed that cytosine methylation is modulated along the growth and allowed to define the cytosine methylomes for each medium and for each growth phase.
2. the effect of DNA cytosine demethylation was studied by treating *S. coelicolor* cultures with 5-aza-2'-deoxycytidine (aza-dC, a cytidine analogous that inhibits DNA-methyltransferase activity); this analysis revealed that cytosine demethylation of DNA impairs morphological and physiological differentiation both in liquid and on solid cultures.

In addition, the role of a putative DNA-methyltransferase *SCO1731* was studied by generating an independent mutant of this gene. *SCO1731* is mainly expressed in MI phase, both on solid

and in liquid culture. Phenotypic analysis revealed that the morphological and physiological differentiation was delayed on solid GYM and in liquid R5A.

Finally, to investigate whether the DNA cytosine methylation is conserved in the Streptomycetaceae family, dot blot analysis of several *Streptomyces* strains (*S. lividans, S. griseus* and *S. avermitilis*) was performed.

Results

1. DNA cytosine and adenine methylation in the defined liquid MG

At the beginning of this project, the methylation effect on the *S. coelicolor* growth was preliminarly investigated in the defined liquid medium MG.

The growth curve showed a first rapid growth phase (RG1) till 18h, a transition phase (T) till 20h and a second rapid growth phase (RG2) from 20h to 22h, followed by a stationary phase (S) till 26h and a decline phase characterized by a considerable decrease in biomass till 42h (Figure 10). In this medium, cells are in MII phase (Figure 9).

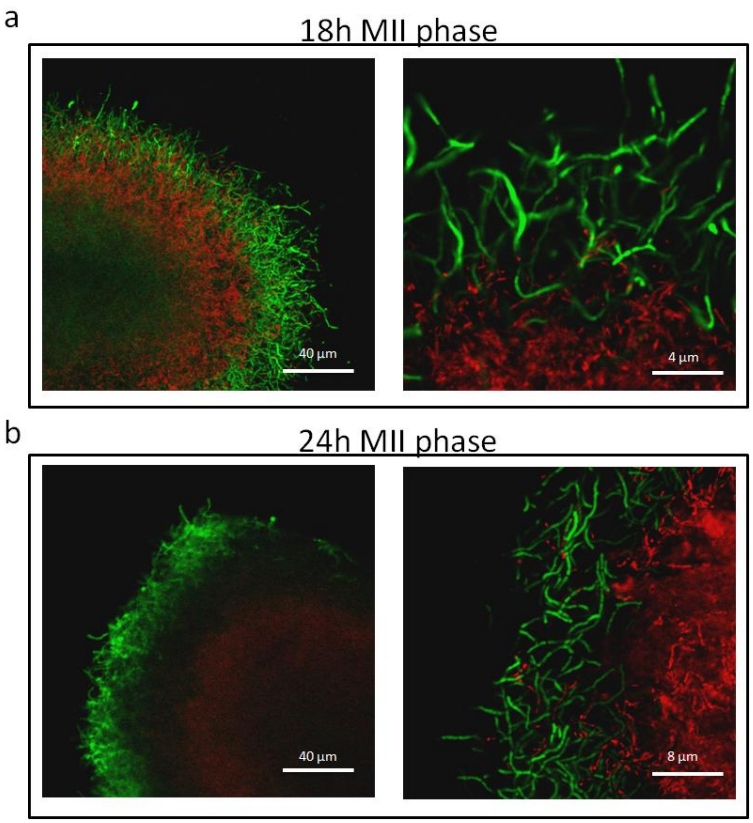

Figure 9 CLSM analysis (a,b) of MII phase of *S. coelicolor* in liquid MG after 18h and 24h of growth. Images correspond to culture preparations stained with SYTO 9 and PI. Culture time points (hours) and the growth phase are indicated.

Dot blot analysis demonstrated that the global level of methylated adenines and cytosines changes during growth in MG. In particular, the level of methylated adenines was higher at the times 18h and 22h, when the cells were in rapid growth phase (Figure 10). On the contrary, the

level of methylated cytosines was higher at the times (18h and 24h-26h) preceding the transition and decline phase, respectively (Figure 10).

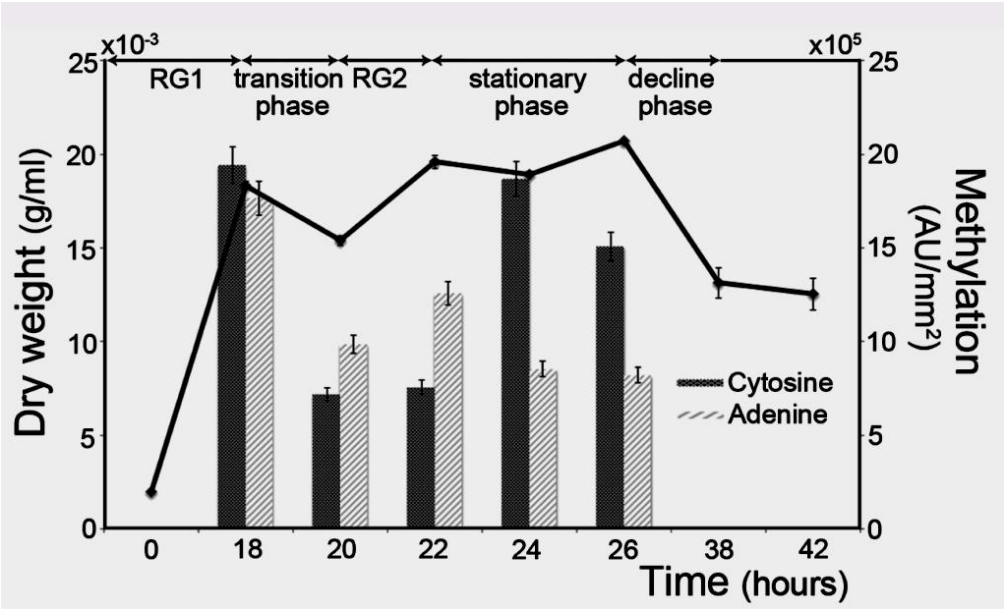

Figure 10 Growth curve of *S. coelicolor* in liquid MG and DNA adenine (grey light bars) and cytosine methylation pattern (grey dark bars) during growth. The level of adenine and cytosine methylation was quantified by Molecular Imager ChemiDoc XRS System Biorad.

1.1 Characterization of *S. coelicolor* cytosine methylome

After this preliminary result, the attention was focused exclusively on cytosine methylation pattern for two reasons:

1. the changes of DNA cytosine methylation patter during *S. coelicolor*'s growth might suggest a role in differentiation;
2. its role is underestimated in bacteria.

Methylome analysis, by Bisulfite sequencing, was carried out on genomic DNA extracted after 18h and 24h of growth in MG (Figure 10), these times preceding the transition and the decline phase, or rather, the "death phases".

A total of 3360 (0.04%) cytosines of the genome (8.867 Mb) are methylated and the most frequent methylation motif is C^mG (2000, 0.025%), followed by the C^mHG (1200, 0.015%) and C^mHH[6] (160, 0.002%) (Figure 11).

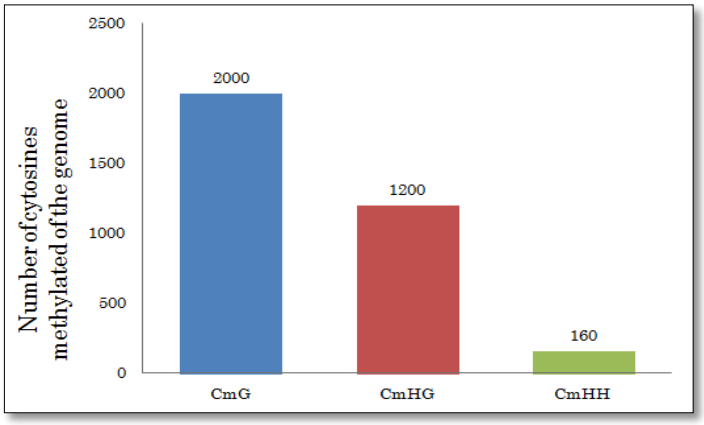

Figure 11 Number of cytosine methylation motifs of the genome of *S. coelicolor*. The X axis indicates the three motifs and the Y axis the the number of methylated cytosines. H stands for A, T or C.

Thus, the search for genes containing methylated cytosines in the upstream region was carried out, considering a region of 400 bp upstream the translation start site of each gene; this will enable us to correlate the presence of cytosine methylated in the upstream region with the regulation of gene expression.

This analysis showed a total of 325 (4.15%) genes, 138 genes containing the C^mG motif, 158 the C^mHG motif and 29 the C^mHH motif (Figure 12).

[6] H stands for A,T or C.

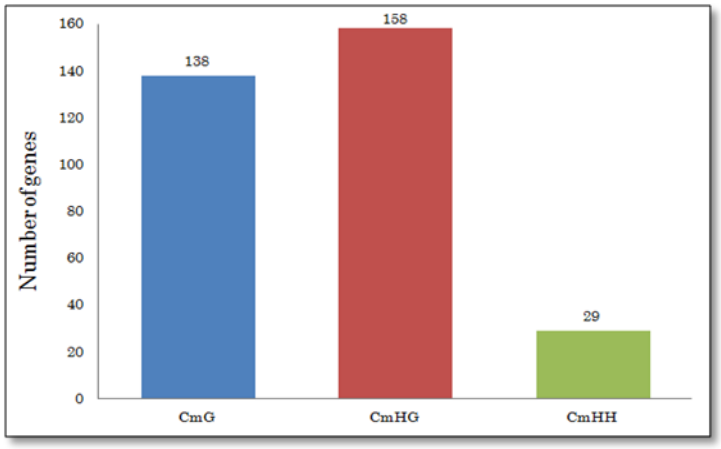

Figure 12 Number of genes containing C^{m}G, C^{m}HG and C^{m}HH motifs in their upstream region. Number of genes contain the cytosine methylation motifs of the genome of *S. coelicolor*. The X axis indicates the three motifs and the Y axis the number of genes with each motif. H stands for A, T or C.

In particular, regarding C^{m}G motif 88 genes are in common between 18h and 24h samples, 26 and 24 genes are specific to 18h and 24h sample; regarding C^{m}HG motif 139 genes are in common between 18h and 24h samples, 8 and 11 genes are specific to 18h and 24h sample; finally, regarding C^{m}HH motif 27 and 2 genes are specific to 18h and 24h sample (Figure 13).

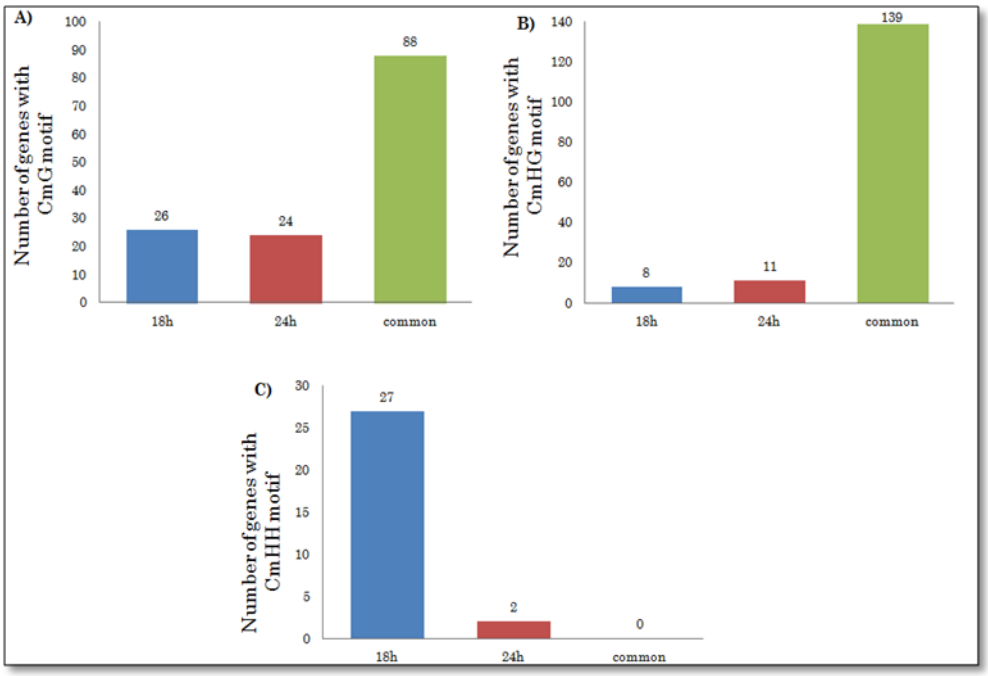

Figure 13 Number of genes containing in their upstream region the methylation motifs **A)** C^{m}G, **B)** C^{m}HG and **C)** C^{m}HH, at 18h, at 24h and both 18h and 24h (indicated like 'common'). The X axis indicates the time of methylation and the Y axis the number of genes containing the three motifs. H stands for A, T or C.

Moreover, the genes were classified on the basis of the number of methylated cytosines present in their upstream region. This analysis revealed that C^mG motif is present once in the upstream region of 109 genes at 18h and of 108 genes at 24h, twice in 4 genes both at 18h and 24h and three times in 1 gene at 18h; C^mHG motif is present once in the upstream region of 140 genes at 18h and of 147 genes at 24h, twice in 2 genes both at 18h and 24h and three times in 1 gene both at 18h and 24h; C^mHH motif is present once in the upstream region of 27 genes at 18h and in 2 genes at 24h (Figure 14).

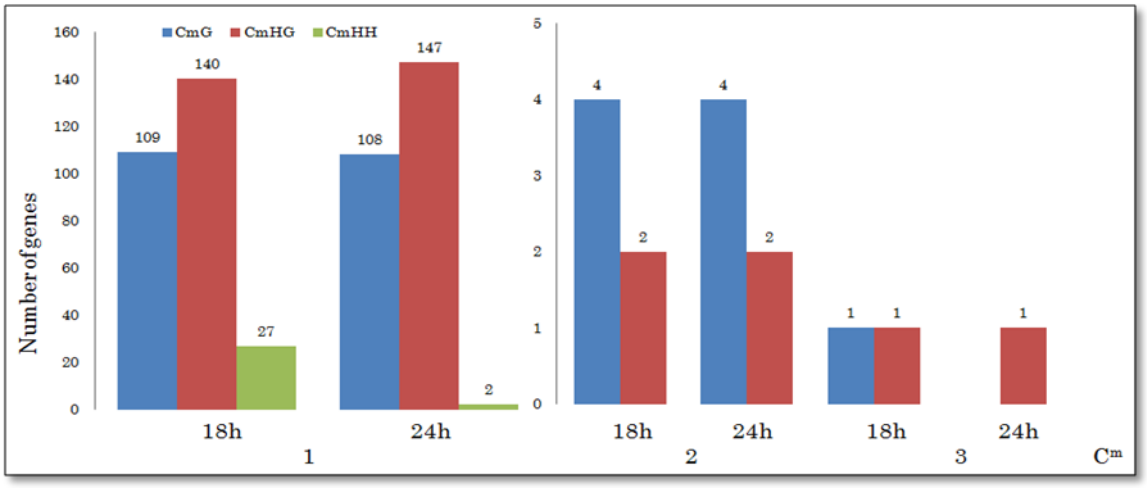

Figure 14 Number of genes containing 1, 2 or 3 methylated cytosines in the upstream region. The X axis indicates the time of methylatilation and the Y axis the number of genes. H stands for A, T or C.

These genes were grouped on the basis of their function: 97 genes are involved in primary metabolism, 39 genes in DNA/RNA metabolism, 35 genes encode putative transcriptional regulators and sigma factors, 52 genes encode membrane proteins and 102 are hypothetical proteins (Table 11).

In particular, 83 genes (Table 12), that are known to be involved in the morphological and physiological differentiation and in DNA replication and chromatin condensation, i. e. *SCO2964 (stgR), SCO3925 (ssgR), SCO5316 (whiE), SCO2077 (divIVA), SCO2716 (chpA), SCO6685*

(ramR), SCO4034 (sigN), SCO3011 (dnaB), SCO3896 (PAP), involved in degradation of nucleic acids), *SCO3109 (MFD),* devoted to DNA repair. and *SCO4662 (tuf),* were identified.

In order to find the most recurring consensus sequence of methylation in our dataset, a background file for the MEME search (36) was generated. MEME analysis revealed two most recurring cytosine methylation consensus sequences, GGC^{m}CGG and GCC^{m}CG (Figure 15).

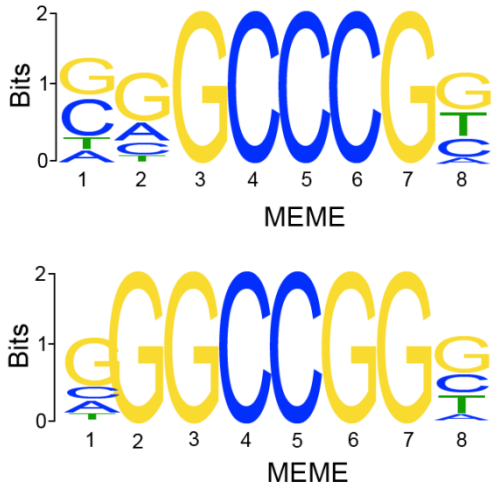

Figure 15 The consensus sequences surrounding the cytosines methylated at 18h and 24h.

S. coelicolor genome was analyzed for the presence of these two sequences. Interestingly, ~ 2150 genes contain these consensus sequences; some of them were not revealed by BS sequencing thus suggesting that they are not methylated at least in the analyzed time points . Among these genes, *SCO4034 (sigN), SCO3925 (ssgR)* and *SCO5316 (whiE),* (genes involved in morphological differentiation) and *SCO5881 (redZ), SCO5878 (redX), SCO5880 (redY)* (genes involved in undecylprodigiosin production) and *SCO3217 (cdaR)* (transcriptional regulator of CDA production) were found.

1.2 Optimization of the use of 5-aza-2'-deoxycytidine (aza-dC)

Since the level of cytosine methylation changed during development and since the number of DNA methyltransferase genes into *S. coelicolor* genome is high (about 30 genes), we tried to inhibit all DNA methyltransferases. We used 5-aza-2-deoxycytidine (aza-dC). Aza-dC is reported to hypomethylate DNA by inhibiting DNA methyltransferases (37).

The amount of aza-dC to add to the medium was chosen after checking the effect of increasing concentrations of the drug on the cells on solid medium GYM (Figure 16). aza-dC is reported to have a half-life of 20h-24h under conditions of physiological temperature and neutral pH (37), so the treatment was repeated every 24h, from 0 to 96h. A control experiment was done in parallel using DMSO (the solvent of aza-dC).

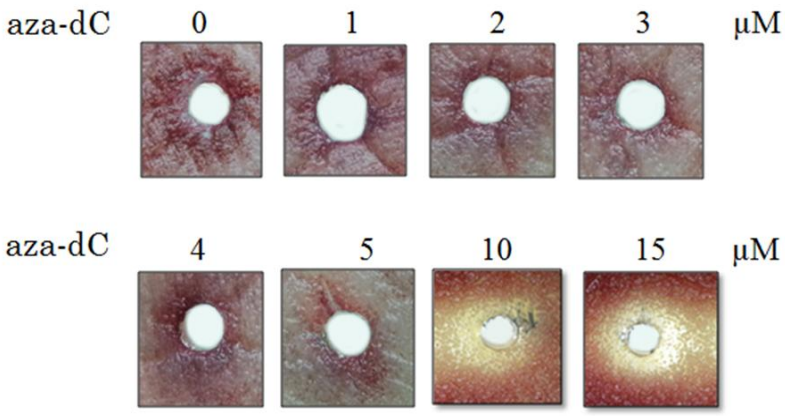

Figure 16 Effect of aza-dC on the cells in the presence of 1, 2, 3, 4, 5, 10 and 15 µM of aza-dC added every 24h, 0= control with DMSO. The pictures were taken after 50h of growth.

5 µM aza-dC was used for further treatments because it was the highest concentration in which the cells were still viable, while 10 and 15 were lethal for the cells, indeed a halo of growth inhibition was present.

The efficiency of DNA cytosine demethylation of the treatment carried out with aza-dC every 24h and every 12h, was evaluated. The efficiency was evaluated, after 24h of growth of *S. coelicolor* in the presence of 5 µM of aza-dC added every 24h and 12h, by dot blot analysis. The

results, showed in Figure 17 revealed that the efficiency of the treatment carried out every 24h was 72%, while every 12h was 99,5%.

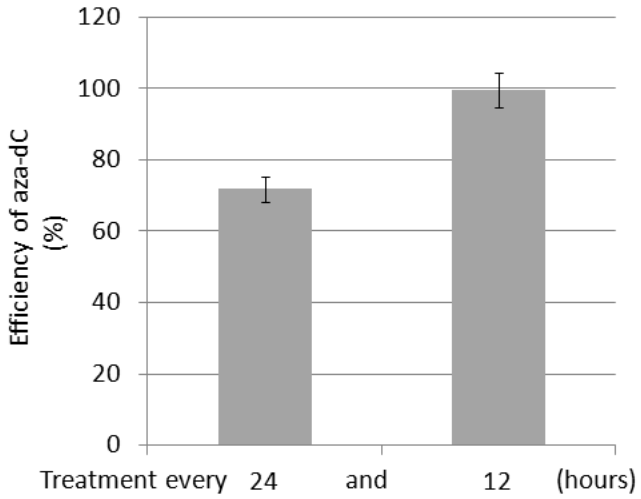

Figure 17 Efficiency percentage of demethylation of aza-dC adding every 12h and 24h.

In order to investigate the effect of aza-dC on growth and physiological differentiation both treatments (every 24h and every12h) were carried out in the liquid medium MG.

1.3 Effect of cytosine demethylation on growth (treatment every 24h)

5 µM aza-dC was added to *S. coelicolor* culture at the time of inoculation of the cells in the liquid medium MG and after 24 hours of growth. A control curve was done in parallel using DMSO.

Growth curve (Figure 18, dashed curve) of *S. coelicolor* treated culture showed a first rapid growth phase (RG1) till 18h, a transition phase (T) from 18h to 20h and a second rapid growth phase (RG2) till 24h, followed by a stationary phase, after the addition of aza-dC, no decline phase was detected.

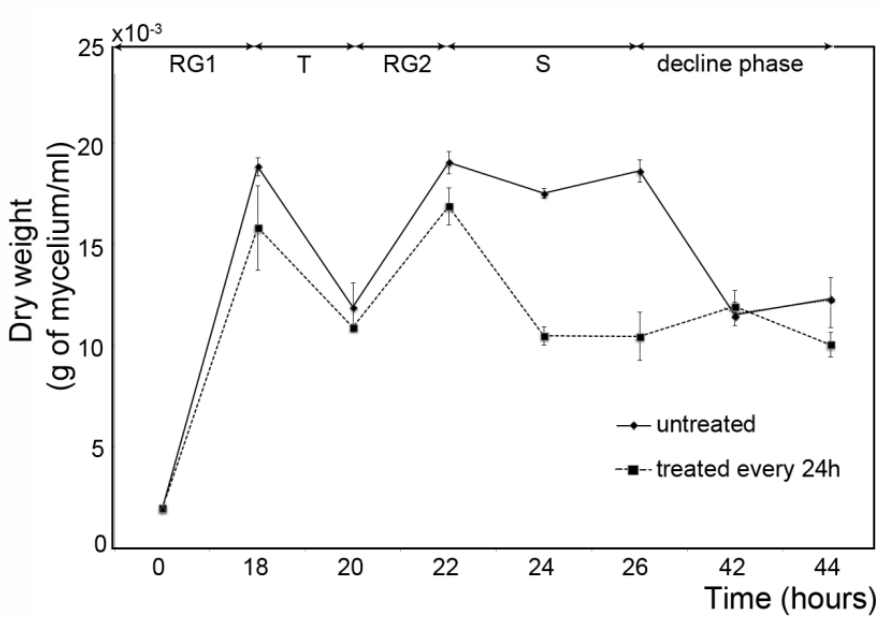

Figure 18 Growth curves of *S. coelicolor* in liquid medium MG, treated with aza-dC every 24h (dashed line) and untreated (continuous line).

1.3.1 Effect of cytosine demethylation on physiological differentiation (treatment every 24h)

To determine the effects of DNA cytosine methylation on physiological differentiation of *S. coelicolor*, production of two antibiotics undecylprodigiosin and actinorhodin was evaluated in the treated and the untreated culture. In the treated culture, undecylprodigiosin production was decreased in respect to the untreated culture, while actinorhodin production seems not to be influenced (Figure 19).

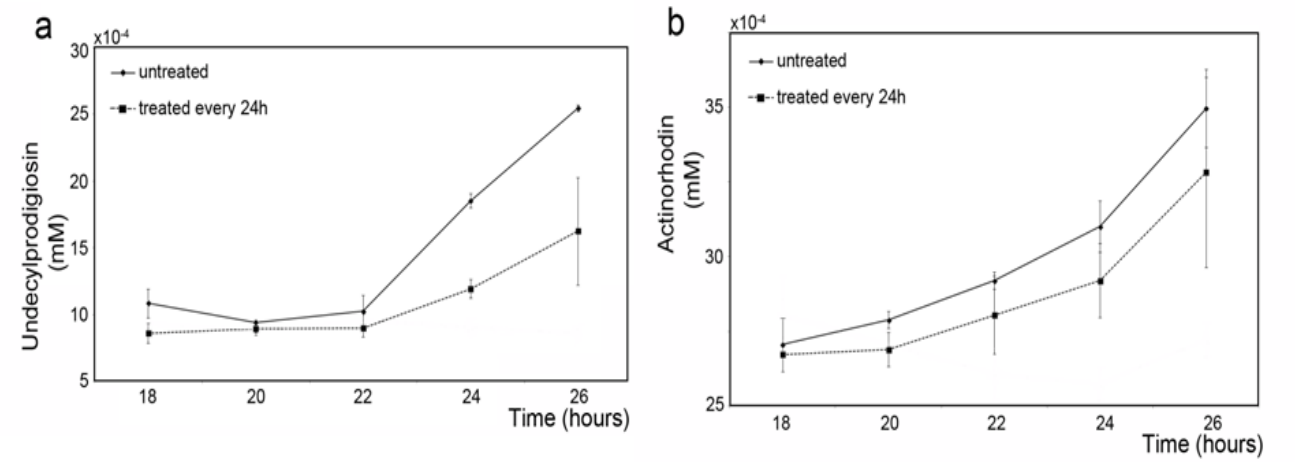

Figure 19 Quantitative analysis of undecylprodigiosin and actinorhodin of *S. coelicolor* in liquid MG, continuous line indicate the antibiotic production in the untreated culture; dashed line in culture treated every 24h. The treated sample corresponds to aza-dC.

1.4 Effect of cytosine demethylation on growth (treatment every 12h)

Since the treatment with aza-dC every 24h did not show strong effects, the treatment was performed every 12h during growth (Figure 20). A control curve was done in parallel using DMSO.

Growth curve (Figure 20, dashed curve) of *S. coelicolor* 12h-treated culture showed a slow growth phase (RG1) till 24, and decline phase was detected from 24 to 42h. The treatment with aza-dC carried out every 12h displayed a stronger effect on the cell growth and the cells were still viable.

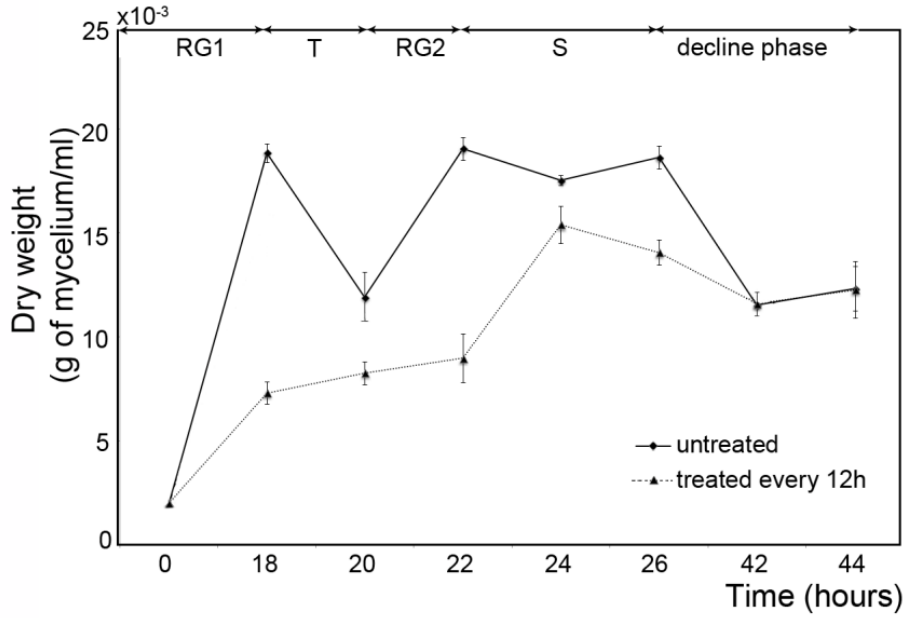

Figure 20 Growth curves of *S. coelicolor* in liquid medium MG, treated with aza-dC every 12h (dashed line) and untreated (continuous line).

1.4.1 Effect of cytosine demethylation on physiological differentiation (treatment every 12h)

To determine the effects of DNA cytosine methylation on physiological differentiation of *S. coelicolor*, undecylprodigiosin and actinorhodin production was evaluated in the 12h-treated and the untreated culture.

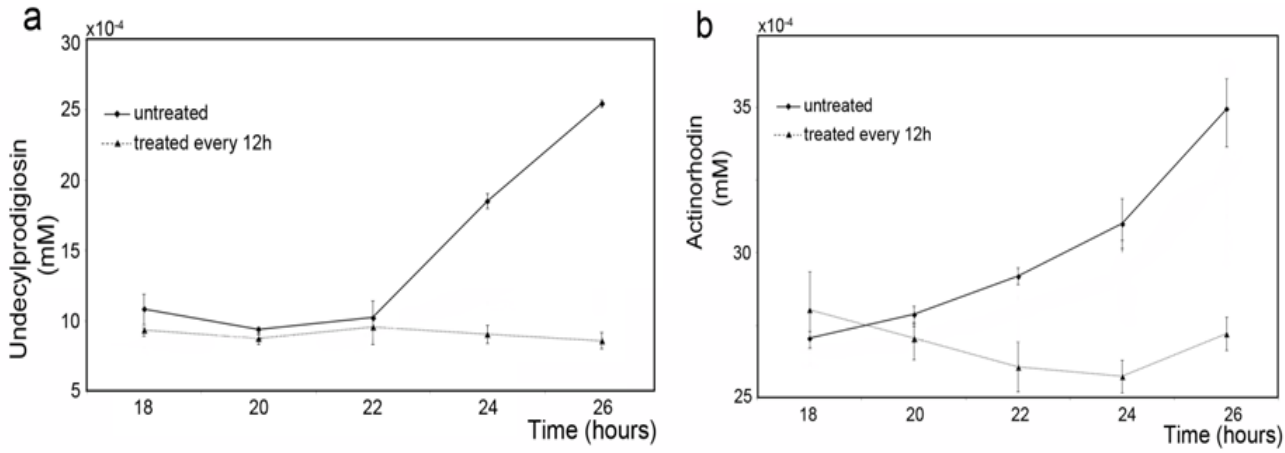

Figure 21 Quantitative analysis of undecylprodigiosin and actinorhodin of *S. coelicolor* in liquid MG. Continuous line indicate the antibiotic production in the untreated culture; dashed line in culture treated every 12h. The treated sample corresponds to aza-dC.

In the 12h-treated culture, undecylprodigiosin and actinorhodin production was decreased in respect to the untreated culture (Figure 21).

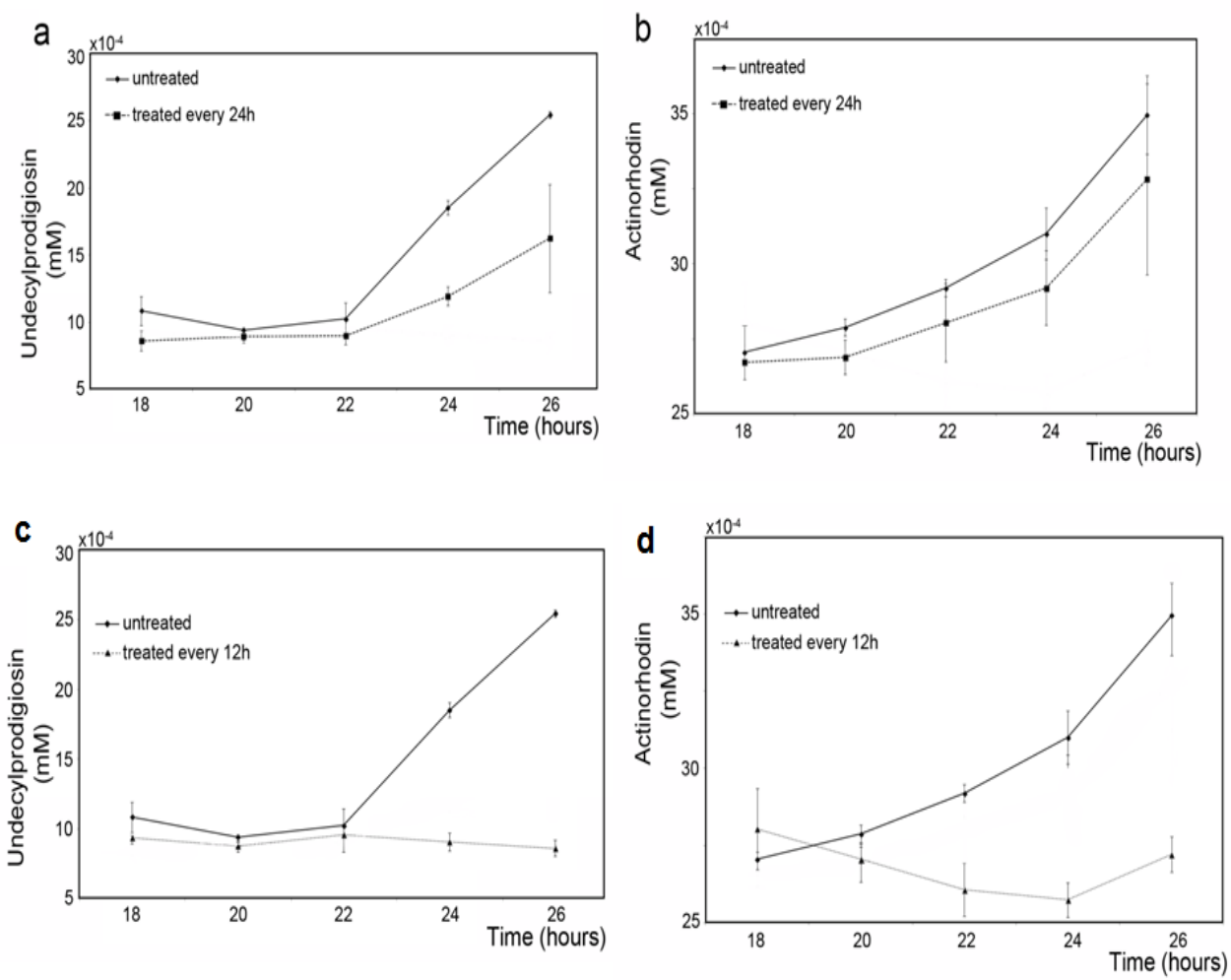

Figure 22 Quantitative analysis of undecylprodigiosin and actinorhodin of *S. coelicolor* cultures (liquid MG) untreated and treated with aza-dC every 24h and every 12h. Continuous lines indicate the antibiotic production in the untreated culture; dashed lines in the treated culture. The treated sample corresponds to aza-dC.

The treatment carried out every 12h (Figure 22 c and d) had a greater effect on the antibiotic production than the treatment every 24h (Figure 22 a and b).

1.5 DNA cytosine methylation and gene expression

To correlate DNA cytosine methylation and gene expression the transcriptional analysis of some genes containing the methylated upstream region was performed.

Two different sets of genes were selected among those listed in the Table 12 List of 83 genes containing the methylated upstream region in MG and involved in physiological and morphological differentiation. In particular, the transcriptional analysis was performed on genes involved in protein and amino acidic metabolism and sigma factor: *SCO6164* (dnaK-like), *SCO2571* (*leuS*), *SCO5820* (*hrdB*) and on the gene *SCO6685 (ramR)* coding for the regulator ramR, involved in cover hydrophobic formation .

These genes show a methylation pattern as indicated in the following table:

Table 2 Genes analyzed by qRT-PCR.

Gene	Main metabolism pathways	Cellular function	Time of methylation	Consensus sequence 1	Consensus sequence 2
SCO2571	Aminoacidic metabolism	leucyl-tRNA synthetase	18-24		GCC^{m}CG
SCO6164	DNA metabolism	molecular chaperone DnaK-like	24		GCC^{m}CG
SCO6685	Regulator	two-component system response regulator (ramR)	24	GGC^{m}CGG	
SCO5820	RNA metabolism	RNA polymerase sigma factor (HrdB)	18-24	GGC^{m}CGG GGC^{m}CGG	

Transcriptional levels of these genes were evaluated at time point 18h and 24h of *S coelicolor* grown in liquid MG untreated and treated with aza-dC every 12h and 24h.

Total RNA was extracted and analyzed for gene expression by qRT-PCR using 16S rRNA gene as endogenous control (Table 10).

The gene *SCO2571*, coding for leucyl-tRNA synthetase and involved in amino acid metabolism, contains the GCC^{m}CG consensus sequence in its upstream region, that it was found to be methylated both at 18h and 24h of growth. The gene *SCO6164*, coding for a molecular chaperone DnaK-like, contains the GCC^{m}CG consensus sequence in its upstream region, that it was found to be methylated at 24h of growth.

The gene *SCO5820,* coding for the RNA polymerase sigma factor *hrdB*, contains two GGC^{m}CGG consensus sequences in its upstream region, that it was found to be methylated both at 18h and 24h of growth. The gene *SCO6685*, coding for the two-component system response

regulator (*ramR*), contains the GGC^{m}CGG consensus sequence in its upstream region, that it was found to be methylated at 24h of growth.

Regarding the treatment every 24h, the transcription levels of these genes at 18h and 24h of growth in 24h-treated samples showed no difference in respect to the untreated samples for the genes *SCO2571* and *SCO6164*, and a little difference in the transcription levels for the genes *SCO5820* and *SCO6685* (Figure 23).

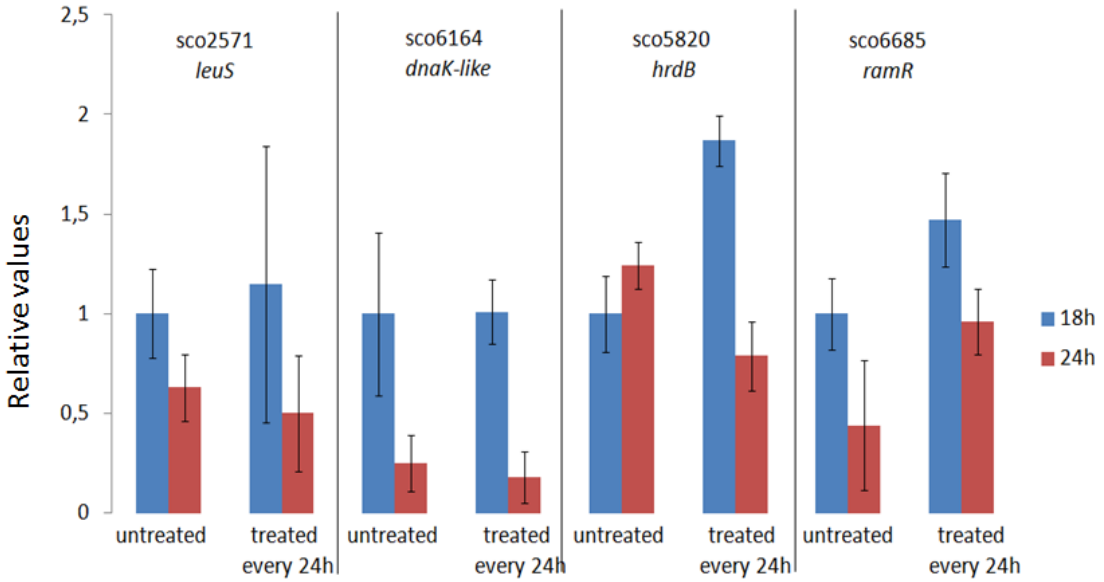

Figure 23 qRT-PCR analysis of *SCO2571*, *SCO6164*, *SCO5820* and *SCO6685* after 18h and 24h of growth and under conditions of untreated and treated with aza-dC. mRNA levels are expressed relative to 16S rRNA transcripts, with the ratio values for the 18h sample arbitrarily set to 1. The standard deviations (indicated by error bars) were calculated from three independent determinations of mRNA abundance in each.

This result suggests that the 24h-treatment did not influence the methylation level of the upstream of these four genes or that the methylation of the upstream did not influence the transcription at 24h.

Regarding the treatment every 12h, the transcription levels of these genes at 18h and 24h of growth in 12h-treated samples showed a big difference in the transcription level in respect to the

untreated samples (Figure 24). For the genes *SCO2571* and *SCO5820*, the 12h-treatment strongly inhibited the transcriptional levels both at 18h and 24h; on the contrary, for the genes *SCO6164* and *SCO6685* the transcriptional levels the 12h-treatment did not influence the transcriptional levels at 18h but induced the transcription at 24h.

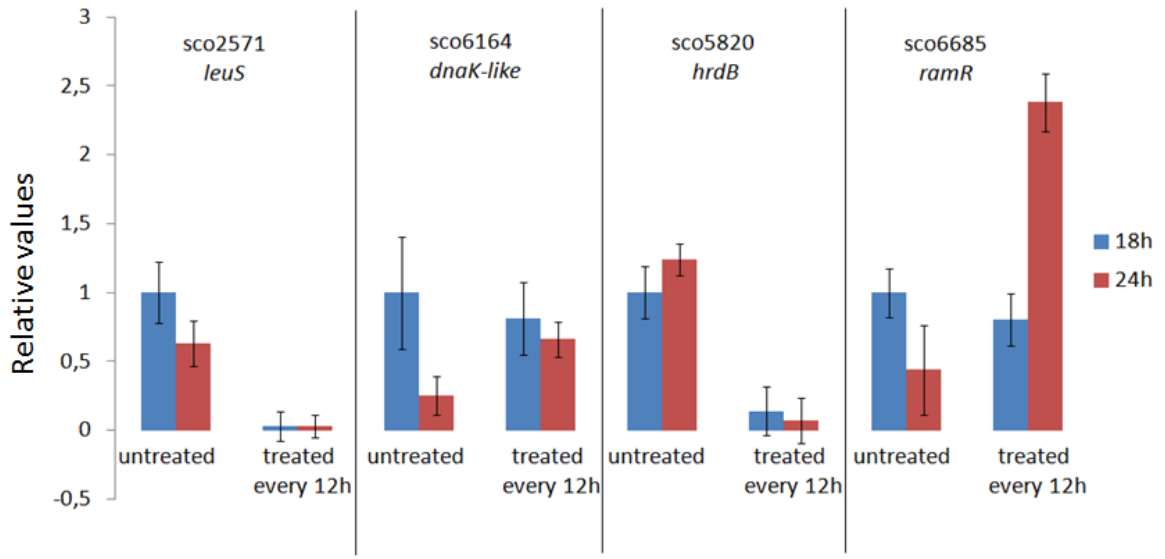

Figure 24 qRT-PCR analysis of *SCO2571*, *SCO6164*, *SCO5820* and *SCO6685* after 18h and 24h of growth and under conditions of untreated and treated with aza-dC. mRNA levels are expressed relative to 16S rRNA transcripts, with the ratio values for the 18h sample arbitrarily set to 1. The standard deviations (indicated by error bars) were calculated from three independent determinations of mRNA abundance in each.

This result suggests that the 12h-treatment had an effect on transcriptional levels of genes containing methylated cytosines in their upstream region and that the 12h-treatment inhibited the transcription of the genes *SCO2571* and *SCO5820* and increased the transcription of the genes *SCO6164* and *SCO6685*.

On the basis of the obtained results about the effect of the treatment carried out with aza-dC every 24h and every 12h in liquid MG, aza-dC was added every 12h for the treatment of *S. coelicolor* both on solid and in liquid culture for further experiments.

2. DNA cytosine methylation in the rich liquid R5A

2.1 DNA cytosine methylation during development

After having demonstrated that cytosine methylation plays a role in growth and physiological differentiation in the liquid medium MG, the role of cytosine methylation was investigated in another liquid medium R5A.

R5A was chosen since in this medium the cell cycle was deeply investigated (28) and at each growth phase (MI and MII) transcriptomic and proteomic analysis was already carried out (34) (35). Thus, combining methylome and gene expression analyses the comprehension of the role of cytosine methylation in *S. coelicolor* could be accelerated.

In R5A, *S. coelicolor* showed a partial life cycle: the MI phase lasts after spore germination until 30h, then there is a mix of both MI and MII until 40-45h and after 45h the cells are in MII phase (Figure 25).

To investigate the levels of cytosine methylation during growth in this medium, genomic DNAs of *S. coelicolor* was extracted from MI (20h) and MII phase (55h) and analyzed by dot blot.

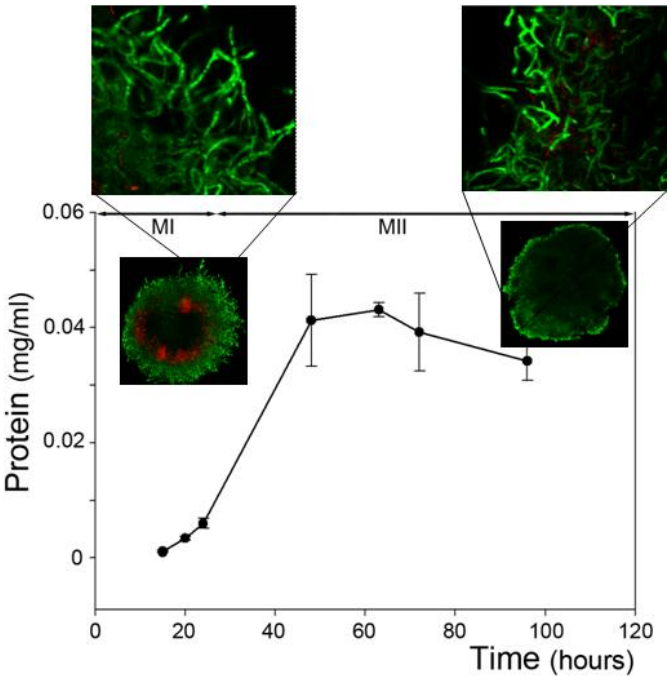

Figure 25 Growth curves of *S. coelicolor* in liquid R5A. Pictures of MI and MII are shown.

The level of cytosine methylation changes during growth in liquid R5A. Specifically, DNA cytosine methylation was higher at the MI stage than in the MII stage (Figure 26).

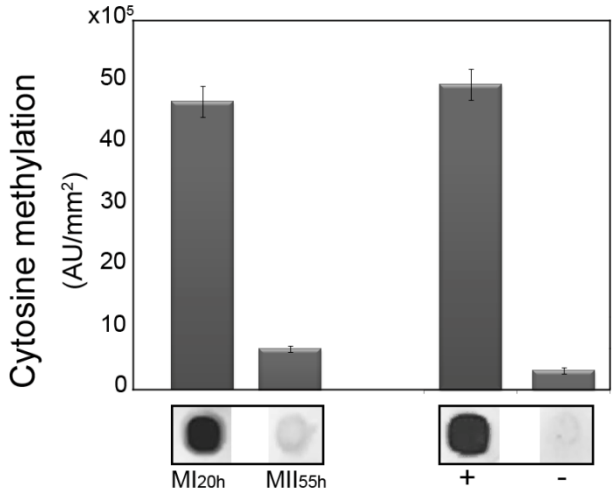

Figure 26 DNA cytosine methylation pattern during growth in liquid R5A. The level of cytosine methylation was quantified by Molecular Imager ChemiDoc XRS System Biorad. +: positive control (genomic DNA of *E.coli* dcm+); -: negative control (genomic DNA of *E.coli* dcm-). Dot blot is shown at the bottom of the graphs.

2.2 Characterization of *S. coelicolor* cytosine methylome

Bisulfite (BS) sequencing of DNA samples extracted of genomic DNAs extracted from MI (20h) and MII (55h) of *S. coelicolor* grown in R5A (Figure 25) was carried out.

Bioinformatic analysis of these results of BS revealed that 2190 genes (28%) were methylated during the development.

The number of methylated genes was compared between MI and MII phase. Among the 2190 genes, 1275 (58.5%) genes were methylated during the MI phase (20h), 665 (30%) were methylated during the MII phase (55h) and 250 (11.5%) were methylated both in MI and MII phase, as depicted by a Venn diagram (Figure 27). The methylated cytosines comprised in the region between -400 and +100 bp of genes were considered.

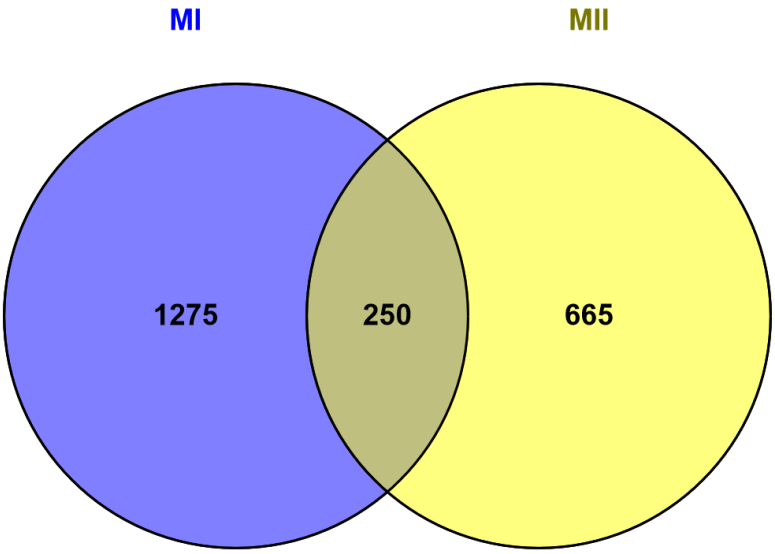

Figure 27 Venn diagram showing the comparison between genes identified in MI and MII phase of liquid R5A.

MEME analysis revealed that the two most recurring cytosine methylation sequences were the same found for the analysis of BS sequencing performed in liquid medium MG (GGC^{m}CGG and GCC^{m}CG). Interestingly, another cytosine methylation consensus sequence C^{m}GGGC was identified (Figure 28). In particular, the consensus sequence GGC^{m}CGG was exclusively found in genes methylated in MI phase, on the contrary, GCC^{m}CG and C^{m}GGGC consensus sequences were common in genes methylated in MII phase.

Figure 28 The third methylation consensus sequence found in R5A.

Next, the genes identified by BS sequencing were grouped in functional categories (Figure 29): 342 genes involved in primary metabolism (DNA/RNA replication, aerobic and anaerobic production, glycolysis and gluconeogenesis, pentose phosphate pathway, amino acid

metabolism, nucleotide metabolism, translation, protein folding, RNA/protein processing, nucleases/RM methylases); 58 genes implicated in secondary metabolism; 243 genes coding for regulatory proteins (transcriptional regulators, kinases, other regulatory proteins, sigma factors); 38 genes correlated to differentiation (TTA BldA targets, Bld and Whi proteins); 175 genes coding for transporters and secreted (ABC transporters, transporters and secreted proteins); 61 genes involved in catabolism and degradation; 26 genes related to lipid metabolism; 26 genes coding for stress and defense proteins and 1221 genes with unknown function.

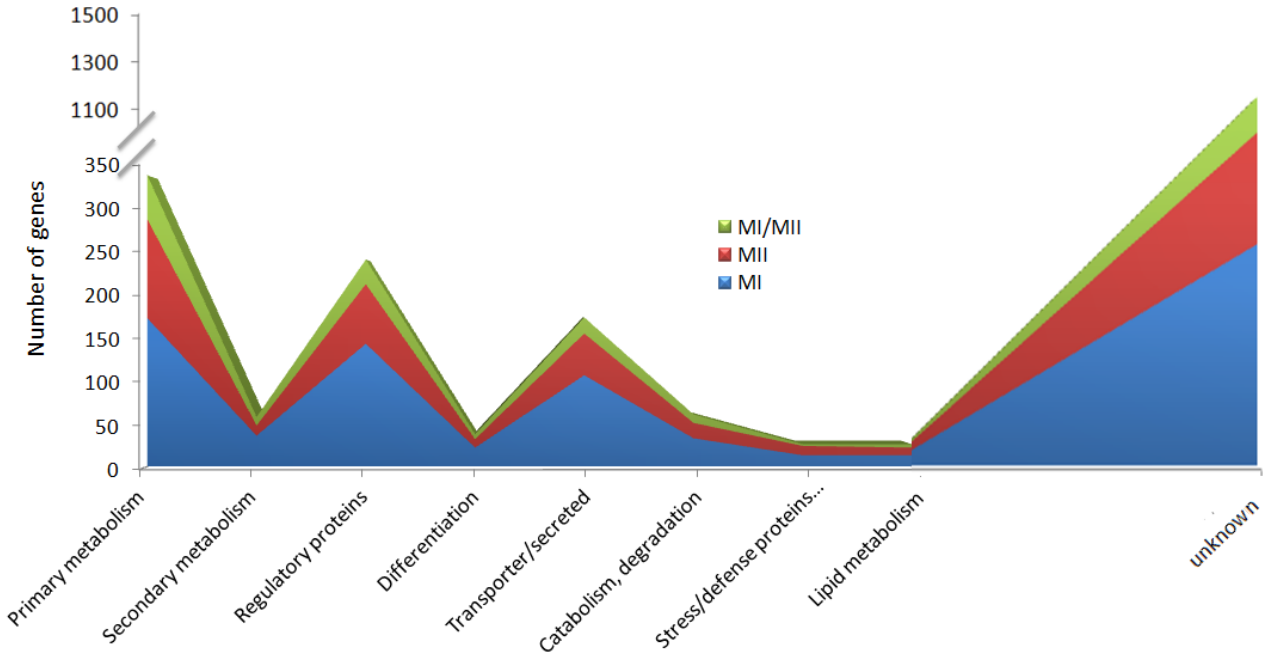

Figure 29 Genes identified by BS sequencing were grouped in functional categories: primary metabolites, secondary metabolites, regulatory proteins, differentiation, transporters and secreted proteins, catabolism and degradation, stress and defense proteins, lipid metabolism, unknown. Blue line represents the genes methylated in MI phase, red the MII phase and green the genes methylated both in MI and MII phase

Next, the MI and MII cytosine methylomes were compared and it was found that genes involved in secondary metabolism, genes coding for regulatory proteins and genes involved in differentiation were methylated during the MI phase in respect to MII phase. In particular, the genes found were those involved in undecylprodigiosin production: *SCO5878 (redX), SCO5881*

(redZ), SCO5882 (redV), SCO5887 (redP), SCO5892 (redL), SCO5896 (redH), SCO5879 (redW);

genes related to actinorhodin production: *SCO5080 (actV4), SCO5088 (actI-ORFII), SCO5084 (actII-3), SCO5092 (actVB), SCO5091 (actIV), SCO5090 (actV-ORF2);*

genes involved in CDA production: *SCO3217 (cdaR), SCO3231 (cdaPSII), SCO3232 (cdaPSIII);*

genes correlated with cell division and septation: *SCO5587 (ftsH)*, and with geosmin production: *SCO6073 (geoA);*

genes associated with differentiation: *SCO2716 (chpA), SCO1800 (chpE), SCO5315 (whiE-ORFIV), SCO5316 (whiE-ORFV), SCO5114 (bldKC), SCO6685 (ramR);*

genes coding for regulatory proteins: *SCO4035 (sigF), SCO4908 (sigQ), SCO4116 (afsR-like), SCO2964 (stgR).*

In MII phase some genes involved in undecylprodigiosin and actinorhodin production *SCO5897 (redG), SCO5076 (actVA1), SCO5078 (actVA2), SCO5085 (actII-ORF4)* were found;

genes correlated with differentiation: *SCO5113 (bldKB). SCO1489 (bldD), SCO2077 (divIVA);*

a gene associated with primary metabolism: *SCO2571 (leuS)*;

genes coding for regulatory proteins: *SCO3068 (sigI), SCO3892 (sigT)*, regulator of *SCO0599 (sigB)* and *SCO0895 (hrdC).*

The genes that contained methylated cytosines in their regulation region both in MI and MII phase were *SCO5820 (hrdB)*, coding for a vegetative sigma factor, genes involved in antibiotic production *SCO5898 (redF), SCO5078 (actVA3)* and *SCO3230 (cdaPSI);*

a gene involved in RM system: *SCO6635 (plgY)* and two genes related to differentiation: *SCO5112 (bldKA)* and *SCO5115 (bldKD).*

2.3 Effect of cytosine demethylation on growth

S. coelicolor was grown in liquid R5A and aza-dC was added every 12h from the time of inoculation till 96h. The growth curve of (Figure 30) shows that the addition of aza-dC decreased the rate of growth. The germination phase was delayed also in liquid R5A, in fact CLSM analysis revealed the presence of ungerminated spores at 20h in the treated culture (data not shown).

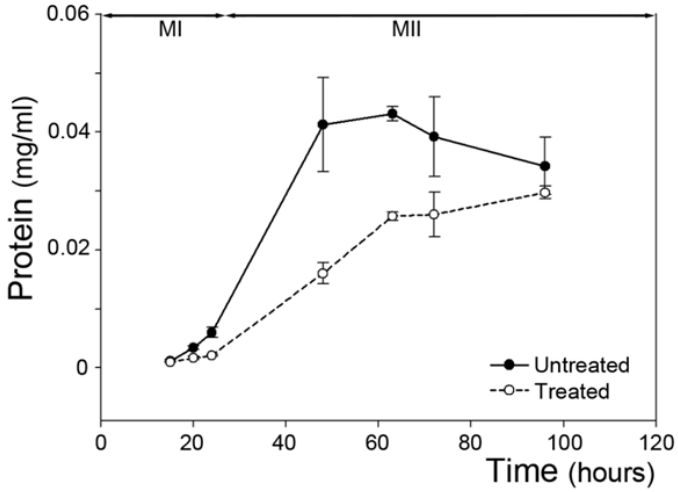

Figure 30 Growth curves of *S. coelicolor* in liquid R5A; treated with aza-dC (5 µM) (dashed line) and untreated (continuous line).

2.4 Effect of cytosine demethylation on physiological differentiation

To determine the effects of DNA cytosine methylation on physiological differentiation of *S. coelicolor*, production of the antibiotics undecylprodigiosin and actinorhodin was evaluated in the untreated and the treated culture.

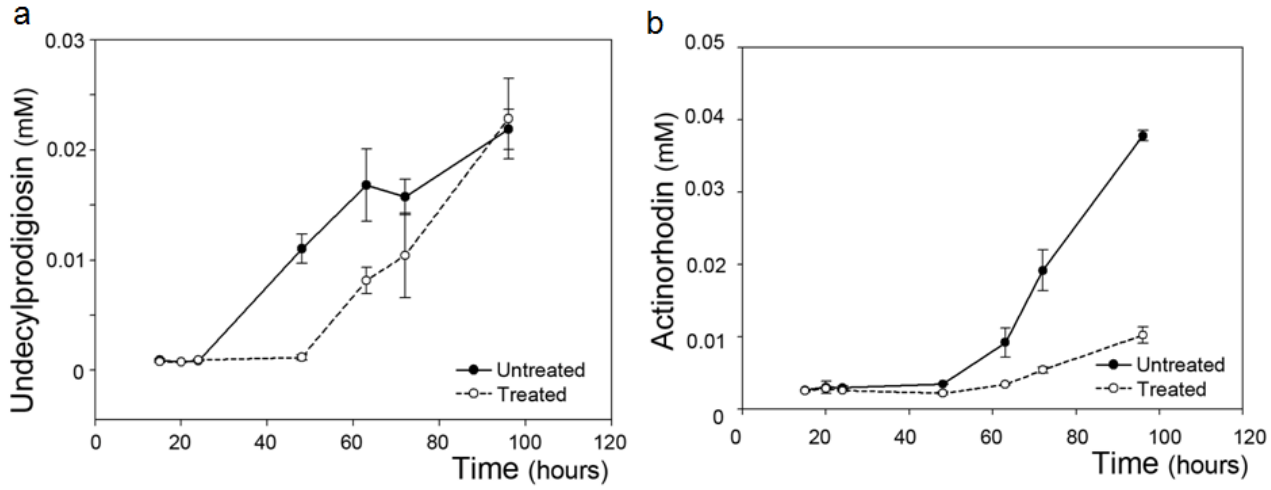

Figure 31 Quantitative analysis of undecylprodigiosin and actinorhodin production of *S. coelicolor* in liquid R5A. Continuous lines indicate the antibiotic production in the untreated culture; dashed lines in the treated culture. The treated sample corresponds to aza-dC.

In the treated culture, undecylprodigiosin and actinorhodin production was decreased and delayed in respect to the untreated culture (Figure 31a-b).

2.5 DNA cytosine methylation and gene expression

The list of the genes containig methylated cytosines, in liquid R5A, was compared to the list of differentially transcribed genes previously obtained in the same liquid medium (31).

2.5.1 GGC^{m}CGG consensus sequence

Interestingly, it was found that ~500 genes (Table 13), containing the GGC^{m}CGG cytosine methylation consensus sequence methylated in MI phase, were more transcribed in MII phase. Among them, genes involved in secondary metabolism and differentiation, such as genes *SCO5881 (redZ), SCO3217 (cdaR), SCO6073 (geoA), SCO6685 (ramR)* and *SCO4035 (sigF),* were found (Table 3). Differently, it was found that ~130 genes containing the GGC^{m}CGG cytosine methylation consensus sequence repeated twice in MI phase, in the regulation region, were constitutively transcribed the same in MI and MII, like *SCO5820 (hrdB)* and *SCO5898 (redF).*

Table 3 Example of genes containing the cytosine methylation consensus sequence GGC^{m}CGG correlated with their gene expression. *the same methylated consensus sequence was found twice. (The entire list of genes is listed in Table 13).

Gene	Description	Function	Consensus sequence	Methylation phase	Expression Phase
SCO5881	*redZ*	secondary metabolism	GGCCGG	MI	MII
SCO3217	*cdaR*	secondary metabolism	GGCCGG	MI	MII
SCO6073	*geoA*	secondary metabolism	GGCCGG	MI	MII
SCO6685	*ramR*	transcription regulator	GGCCGG	MI	MII
SCO4035	*sigF*	transcription regulator	GGCCGG	MI	MII
SCO5820	*hrdB*	transcription regulator	GGCCGG*	MI	Unchanged
SCO5898	*redF*	secondary metabolism	GGCCGG*	MI	Unchanged

To correlate DNA cytosine methylation and gene expression the transcriptional analysis of gene *SCO5881 (redZ)* was performed. The gene *SCO5881 (redZ),* coding for response regulator involved in undecyprodigiosin production cascade, contains the GGC^{m}CGG consensus sequence in its upstream region, that it was found to be methylated in MI phase and expressed in MII phase, thus probably the methylation of this sequence inhibits the gene expression in MI.

Transcriptional levels was evaluated at time point 20h and 55h of *S. coelicolor* grown in R5A with or without aza-dC, corresponding to MI and MII phase, respectively.

The transcription levels of this gene in MII treated sample showed no difference in respect to the untreated sample; differently the transcription level in MI treated sample was higher (1.5 fold) in respect to the MI untreated sample (Figure 32).

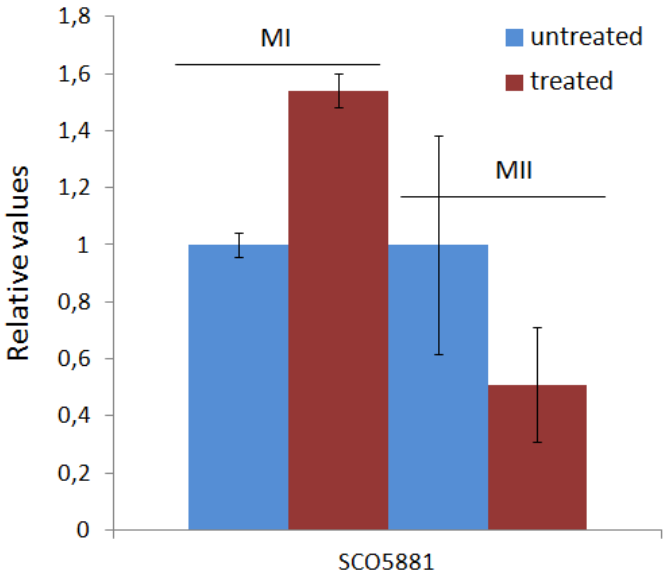

Figure 32 qRT-PCR analysis of *SCO5881* in MI and MII phase under conditions of untreated (blue bars) and aza-dC-treated (red bars) cultures. mRNA levels are expressed as relative to 16S rRNA (primers listed in Table 10) transcripts, with the ratio values for the untreated samples arbitrarily set to 1. The standard deviations (indicated by error bars) were calculated from three independent determinations of mRNA abundance in each.

This result suggests that the methylation of GGC^{m}CGG had an inhibitory effect on transcriptional levels of this gene in MI. In fact, the absence of the methylation, in the treated MI sample, increased the transcription of the gene *SCO5881*.

2.5.2 GCC^{m}CG consensus sequence

Moreover, it was found that ~200 (Table 14) genes containing the GCC^{m}CG cytosine methylation consensus sequence methylated in MII phase, were more transcribed in MII phase, such as *SCO5897 (redG), SCO5085 (actII-ORF4), SCO1489 (bldD), SCO2077 (divIVA)*.

Table 4 Examples of genes containing the cytosine methylation consensus sequence GCC^{m}CG correlated with their gene expression. (The entire list of genes is listed in Table 14).

Gene	Description	Function	Consensus sequence	Methylation phase	Expression Phase
SCO5897	*redG*	transcription regulator	GCCCG	MII	MII
SCO5085	*actII-ORF4*	secondary metabolism	GCCCG	MII	MII
SCO2077	*divIVA*	Differentiation	GCCCG	MII	MII
SCO1489	*bldD*	Differentiation	GCCCG	MII	MII

To correlate DNA cytosine methylation and gene expression the transcriptional analysis of the genes *SCO5085 (actII-ORF4)*, *SCO2077 (divIVA)* and *SCO1489 (bldD)* was performed.

The genes *SCO5085 (actII-ORF4)*, coding for the response regulator involved in actinorhodin production cascade, *SCO2077 (divIVA)*, coding for a protein involved in polar growth, and *SCO1489 (bldD)*, coding for a DNA-binding protein involved in morphological differentiation, contain the GCC^{m}CG consensus sequence in their upstream region; the consensus sequence was found methylated in MII phase for all the three genes and they were expressed in MII phase, thus probably the methylation of this consensus sequence activates the gene expression.

Transcriptional levels of these genes were evaluated at time point 20h and 55h of *S. coelicolor* grown in R5A with or without aza-dC, corresponding at MI and MII phase, respectively.

The transcription levels of these genes in MI treated samples showed no difference in respect to the untreated samples; differently, the transcription levels in MII treated samples were lower (25, 11 and 3 fold, respectively) in respect to the MII untreated samples (Figure 33).

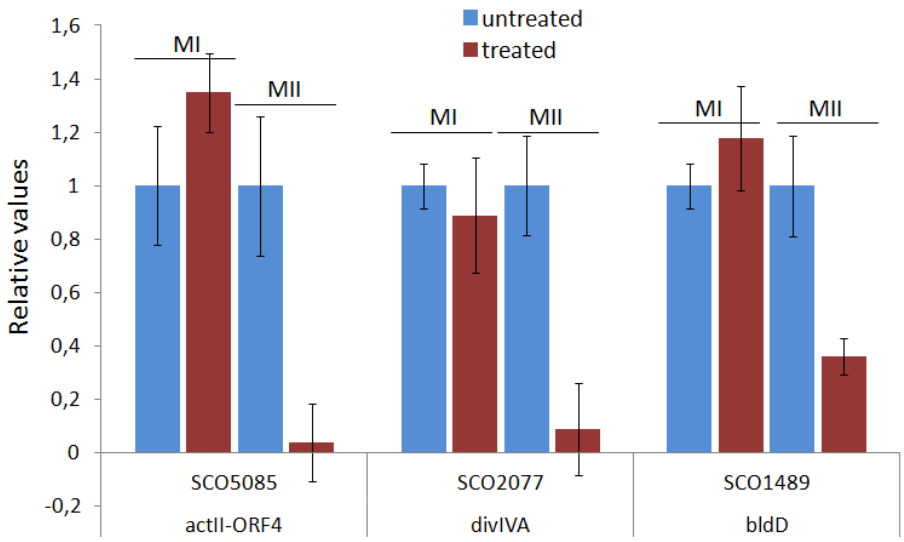

Figure 33 qRT-PCR analysis of *SCO5085*, *SCO2077* and *SCO1489* in MI and MII phase under conditions of untreated (blue bars) and treated with aza-dC (red bars). mRNA levels are expressed relative to 16S rRNA transcripts, with the ratio values for the untreated samples arbitrarily set to 1. The standard deviations (indicated by error bars) were calculated from three independent determinations of mRNA abundance in each.

This result suggests that the presence of GCC^{m}CG methylated consensus sequence had a positive effect on transcriptional levels. In fact, the demethylation of this consensus sequence, in MII treated sample, decreased the transcription of all the three genes analyzed.

2.5.3 C^{m}GGGC consensus sequence

On the contrary, it was found that ~200 (Table 15) genes containing the C^{m}GGGC consensus sequence, that was found to be methylated in MII phase, were more transcribed in MI phase (Table 5), like *SCO0130*, *SCO0191*, *SCO3911 (dnaB)* and *SCO2571 (leuS)*.

Table 5 Example of genes containing the cytosine methylation consensus sequence C^{m}GGGC correlated with their gene expression. (The entire list of genes is listed in Table 15).

Gene	Description	Function	Consensus sequence	Methylation phase	Expression Phase
SCO0130	beta-lactamase	secondary metabolism	CGGGC	MII	MI
SCO0191	licopene cyclase	secondary metabolism	CGGGC	MII	MI
SCO3911	*dnaB*	primary metabolism	CGGGC	MII	MI
SCO2571	*leuS*	primary metabolism	CGGGC	MII	MI

To correlate DNA cytosine methylation and gene expression the transcriptional analysis of *SCO3911 (dnaB)* and *SCO2571 (leuS)* genes was performed. Transcriptional levels of these genes were evaluated at time point 20h and 55h of *S. coelicolor* grown in R5A with or without aza-dC, corresponding at MI and MII phase, respectively.

The transcription levels of these genes in MI and MII treated samples showed no difference in respect to the untreated samples (Figure 34).

This result suggests that the presence of C^{m}GGGC methylated consensus sequence had no effect on transcriptional levels of the analyzed genes.

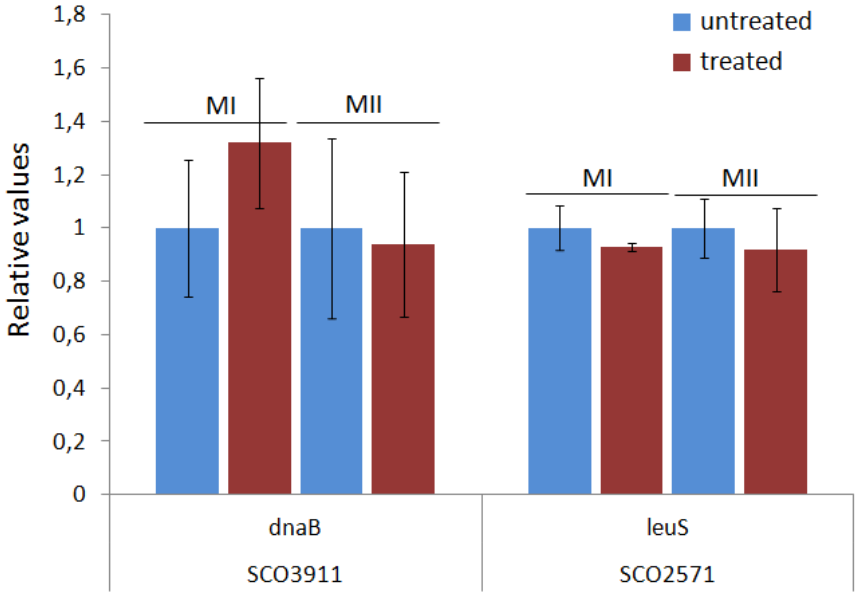

Figure 34 qRT-PCR analysis of *SCO3911* and *SC2571* in MI and MII phase under conditions of untreated (blue bars) and treated with aza-dC (red bars). mRNA levels are expressed relative to 16S rRNA transcripts, with the ratio values for the untreated samples arbitrarily set to 1. The standard deviations (indicated by error bars) were calculated from three independent determinations of mRNA abundance in each.

On the basis of the obtained results about the effect of the treatment carried out with aza-dC it was shown that methylation of the GGC^{m}CGG sequence inhibits gene expression of genes at the MI phase; and methylation of the GCC^{m}CG sequence activates gene expression of genes at the MII phase, while, for the analyzed genes (*SCO3911* and *SCO2571*) no regulatory function in gene expression was found for the C^{m}GGGC sequence.

3. DNA cytosine methylation on solid rich GYM

3.1 DNA cytosine methylation during development

After having demonstrated that cytosine methylation plays a role in growth and physiological differentiation in the MG and R5A liquid media, the role of cytosine methylation was investigated on solid medium GYM.

GYM was chosen since in this medium the cell cycle was deeply investigated (13) and at each growth phase (MI, MII substrate, MI aerial and MI sporulating) transcriptomic and proteomic analysis was already carried out (30-31). Thus, combining methylome and gene expression analyses the comprehension of the role of cytosine methylation in *S. coelicolor* could be accelerate.

On solid GYM, *S. coelicolor* shows a complete life cycle: the MI phase begins after spore germination, subsequent to this vegetative phase the MII substrate is formed(from 24h to 40h) followed by MII aerial formation (from 45h to 63h), the MII sporulating (from 65h to 90) formation anticipate spore formation (from 96h) (Figure 35).

To investigate the levels of cytosine methylation during growth in this medium, genomic DNAs of *S. coelicolor* was extracted from MI (15h), MII substrate (30h), MII aerial (48h), MII sporulating (72h) and spore and analyzed by dot blot.

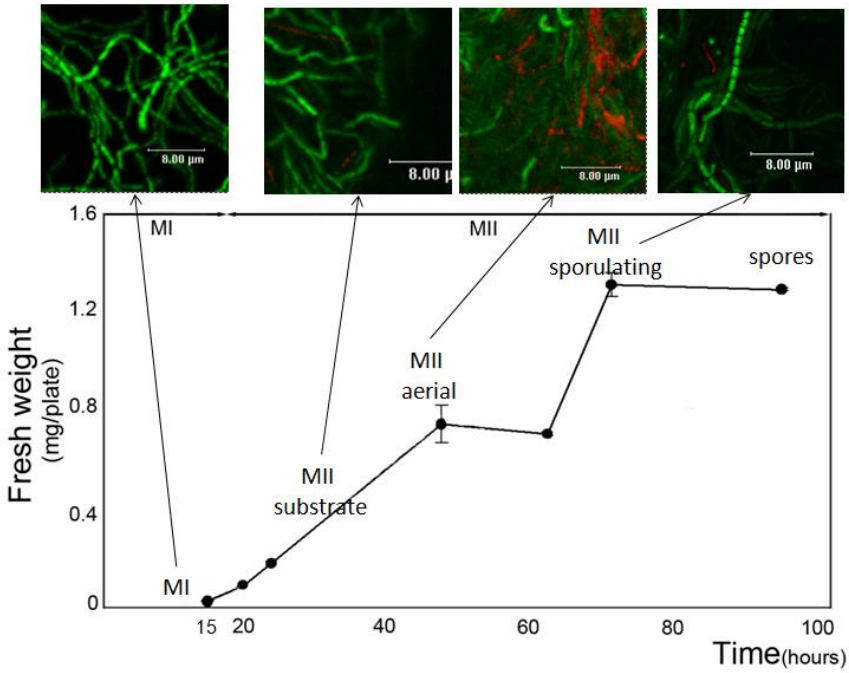

Figure 35 Growth curve of *S. coelicolor* on solid GYM. Pictures of MI, MII substrate (30h), MII aerial (48h) and MII sporulating are shown.

The global level of methylated cytosines changes during growth on solid medium GYM (Figure 36); in particular, the level of methylated cytosines was higher at the MI (15h) stage than in the MII substrate (30h), aerial (48h) and sporulating (72h) and spores.

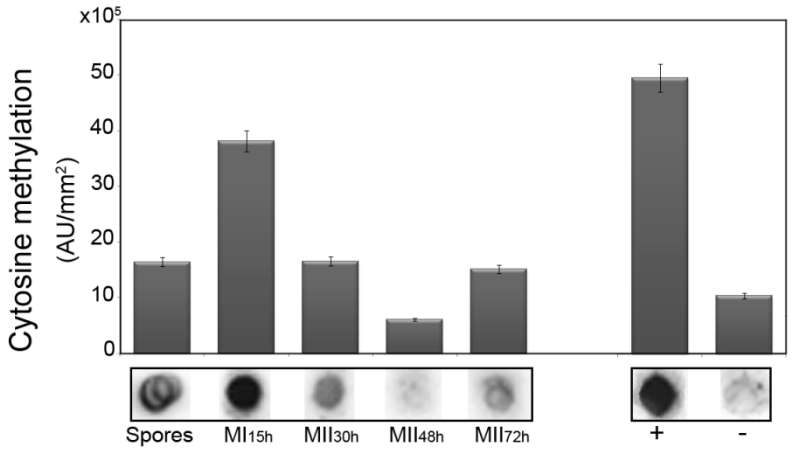

Figure 36 DNA cytosine methylation pattern during growth on solid GYM. The level of cytosine methylation was quantified by Molecular Imager ChemiDoc XRS System Biorad. +: positive control (genomic DNA of *E.coli* dcm+); -: negative control (genomic DNA of *E.coli* dcm-). Dot blot is shown at the bottom of the graphs.

These results confirmed that the chromosomal DNA is subjected to a differential cytosine methylation also during development on a solid medium GYM, as already described in liquid media MG and R5A.

3.2 Characterization of *S. coelicolor* cytosine methylome

Bisulfite (BS) sequencing of DNA samples extracted from the MI (15h), MII substrate (30h), MII aerial (48h), MII sporulating (72h) and spores (96h) of *S. coelicolor* grown in GYM (Figure 35) was carried out.

The number of genes identified, by BS sequencing and bioinformatic analysis, in MI and MII phase is reported in Figure 37. A total of 2251 (28.8%) genes were methylated during the development in their upstream region, out of these 1437 (63.9%) upstream region were methylated during the MI phase (15h), 658 (29.2%) were methylated during the MII substrate phase (30h), 113 (5%) were methylated during the MII aerial phase (48h), 12 (0.5%) were methylated during the MII sporulating phase (72h) and 31 (1.4%) was methylated during the spore phase.

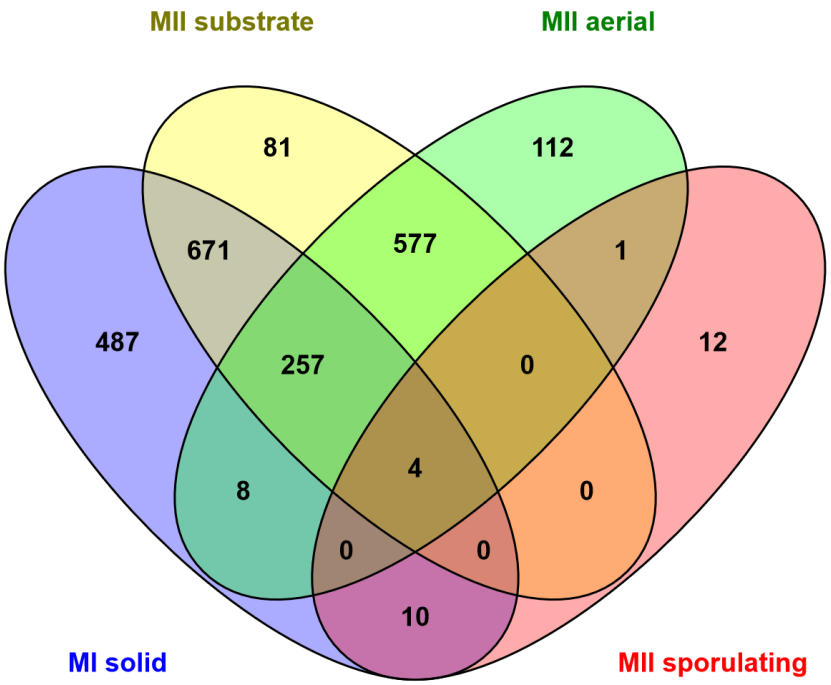

Figure 37 Venn diagram showing the comparison between genes identified in MI and MII phase of solid GYM.

MEME analysis revealed that the three most recurring cytosine methylation sequences were the same to those found for the growth in the liquid MG (GGC^{m}CGG and GCC^{m}CG) and R5A medium (GGC^{m}CGG, GCC^{m}CG and C^{m}GGGC). In particular, the consensus sequence GGC^{m}CGG was exclusively found in genes methylated in MI phase, on the contrary, GCC^{m}CG and C^{m}GGGC consensus sequences were found in genes methylated in MII phase.

Next, the genes identified by BS sequencing were grouped in functional categories (Figure 38): 363 genes involved in primary metabolism (DNA/RNA replication, aerobic and anaerobic production, glycolysis and gluconeogenesis, pentose phosphate pathway, amino acid metabolism, nucleotide metabolism, translation, protein folding, RNA/protein processing, nucleases/RM methylases); 55 genes implicated in secondary metabolism; 258 genes coding for regulatory proteins (transcriptional regulators, kinases, other regulatory proteins, sigma factors); 40 genes correlated to differentiation (TTA BldA targets, Bld and Whi proteins); 179 genes coding for transporters and secreted proteins (ABC transporters, transporters and secreted proteins); 62 genes involved in catabolism and degradation; 25 genes related to lipid metabolism;

24 genes coding for stress and defense proteins and 1245 genes with unknown function (Figure 38).

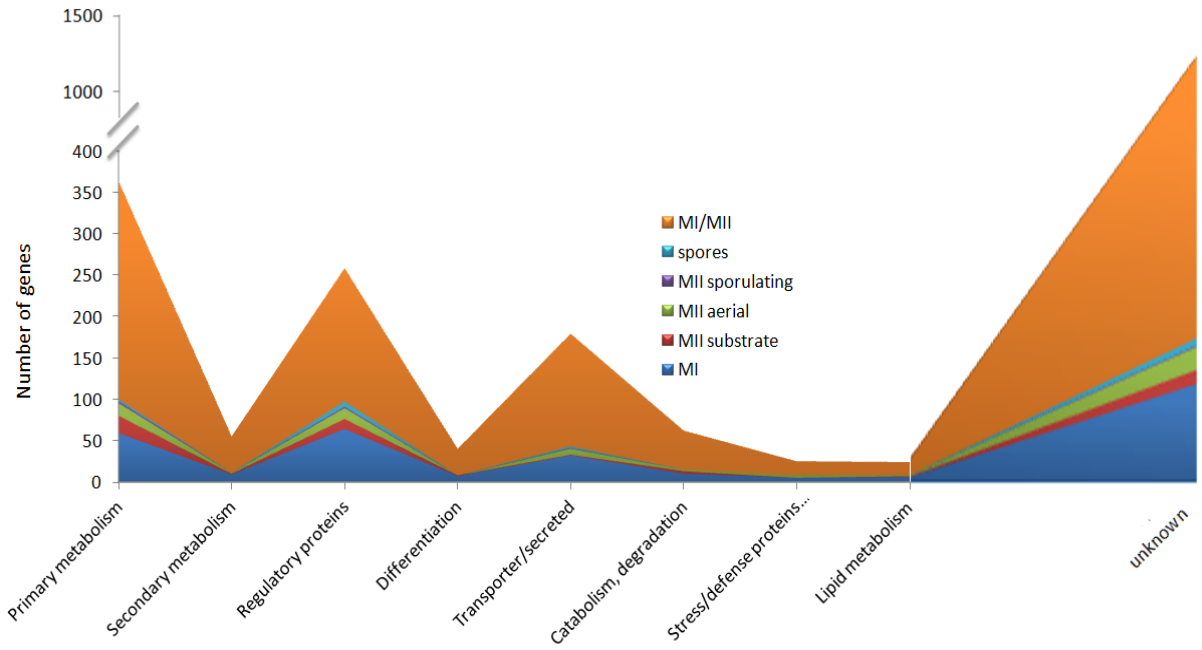

Figure 38 Genes identified by BS sequencing were grouped in functional categories: primary metabolism, secondary metabolism, regulatory proteins, differentiation, transporters and secreted proteins, catabolism and degradation, stress and defense proteins, lipid metabolism, unknown. Blue line represents the genes contain the methylated upstream region in MI phase, red the MII substrate phase, green the MII aerial phase, light blue the spores and orange the MI/MII.

Next, the MI and MII cytosine methylomes were compared and it was found that genes involved in secondary metabolism, genes coding for regulatory proteins and genes involved in differentiation were methylated during the MI phase in respect to MII (substrate, aerial, sporulating) phase. In particular, the genes found were those involved in undecylprodigiosin production: *SCO5882 (redV), SCO5890 (redN);*

genes associated with differentiation: *SCO1800 (chpE), SCO5315 (whiE-ORFIV), SCO5316 (whiE-ORFV);*

genes coding for regulatory proteins: *SCO4035 (sigF).*

The genes that contained methylated cytosines in their regulation region both in MI and MII substrate phase were involved in undecylprodigiosin production: *SCO5887 (redP), SCO5892 (redL), SCO5896 (redH), SCO5879 (redW), SCO5881 (redZ);*

genes related to actinorhodin production: *SCO5080 (actV4), SCO5085 (actII-ORF4), SCO5088 (actI-ORFII), SCO5092 (actVB), SCO5091 (actIV), SCO5090 (actV-ORF2);*

genes involved in CDA production: *SCO3217 (cdaR), SCO3231 (cdaPSII), SCO3232 (cdaPSIII);*

genes correlated with geosmin production: *SCO6073 (geoA)*;

genes associated with differentiation: *SCO2716 (chpA), SCO5114 (bldKC), SCO6681 (ramC), SCO6685 (ramR);*

the gene coding for regulatory protein: *SCO2964 (stgR).*

The genes that contained methylated cytosines in their regulation region both in MII substrate and MII aerial phase were involved in actinorhodin production: *SCO5076 (actVA1), SCO5078 (actVA2);*

genes correlated with differentiation: *SCO5113 (bldKB), SCO1489 (bldD), SCO2077 (divIVA);*

a gene associated with primary metabolism: *SCO2571 (leuS);*

genes coding for regulatory proteins: *SCO3068 (sigI), SCO3892 (sigT)* and *SCO0895 (hrdC).*

The genes that contained methylated cytosines in their regulation region in MI, MII substrate and MII aerial were *SCO5820 (hrdB),* coding for a vegetative sigma factor, genes involved in antibiotic production *SCO5898 (redF), SCO5897 (redG)* and *SCO5079 (actVA4);*

a gene involved in RM system: *SCO6635 (plgY)* and three genes related to differentiation: *SCO5112 (bldKA), SCO5115 (bldKD)* and *SCO4423 (afsK).*

Regarding the spore phase, only 31 genes presented methylated cytosine in their regulation region, out of these 21 have an unknown function.

3.3 Effect of cytosine demethylation on growth

S. coelicolor was grown on solid GYM and aza-dC was added every 12h from the time of inoculation till 96h.

The growth curve (Figure 39, continuous curve) of the untreated culture showed a first rapid growth phase from 15h (MI) to 48h (MII aerial), a transition phase form 48h to 63h, in which the antibiotic production begins, and a second rapid growth phase from 63 to 72h (MII sporulating), followed by a stationary phase till 96h. Spore formation indicated the end of the life cycle.

Growth curve (Figure 39, dashed curve) of the treated culture showed that demethylation did not affect the growth till 63h, indeed the curves were similar. After 63h the treated culture grew very slowly and remained in the stationary phase without sporulation phase. The block of the cells in the stationary phase was previously shown in MG medium.

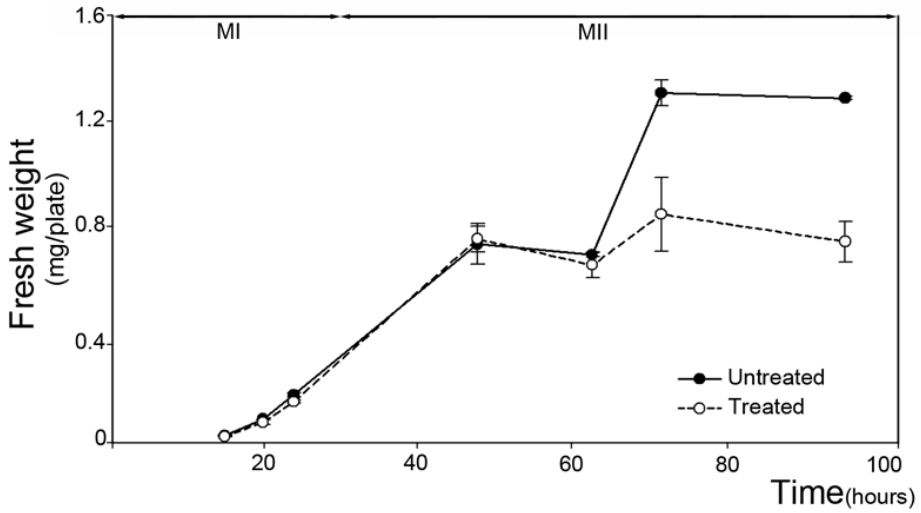

Figure 39 Growth curves of *S. coelicolor* on solid GYM; untreated (continuous line) and treated cultures with aza-dC (5 µM) (dashed line).

3.4 Effect of cytosine demethylation on morphological differentiation

The development processes of *S. coelicolor* cultures, untreated and treated with aza-dC, were followed under CLSM after SYTO 9 and PI staining.

Evaluation of spore germination showed that after 9h of growth in the untreated culture about 95% of the spores germinated, while, in the treated culture, about 65% of spores germinated (Figure 40), after 12h the spores of the treated culture were completely germinated. The analysis of the germination phase revealed that spore germination was delayed after the treatment with aza-dC.

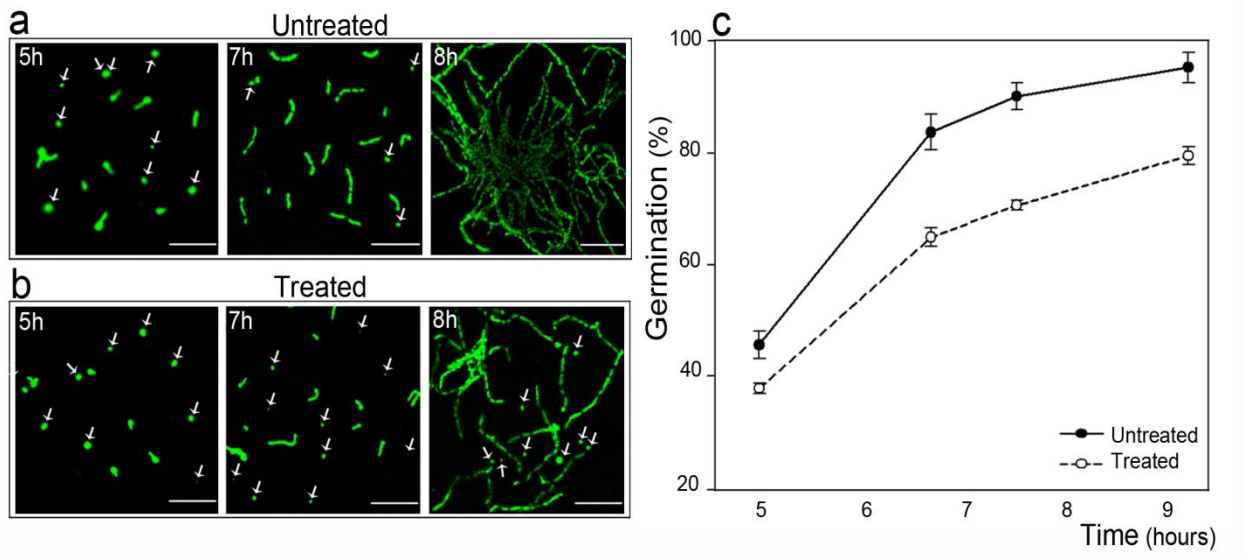

Figure 40 CLSM analysis (a,b) of germination phase of untreated and treated *S. coelicolor* on solid GYM Images correspond to culture preparations stained with SYTO 9 and PI. The arrows indicate ungerinated spores. c) Percentage of spore germination after 5h, 7h, 8h and 9h of growth of untreated (continuous line) and treated (dashed line) *S. coelicolor* with aza-dC (5 µM).

Regarding sporulation, the untreated samples at 72h and 96h, grown on solid GYM analyzed by CLSM showed respectively viable spore chains and single spores (Figure 41); on the contrary, the treated samples showed the multinucleated secondary mycelium (MII) characterized by the non-septate branching hyphae (Figure 41b). Thus, the sporulation phase was blocked in the treated culture.

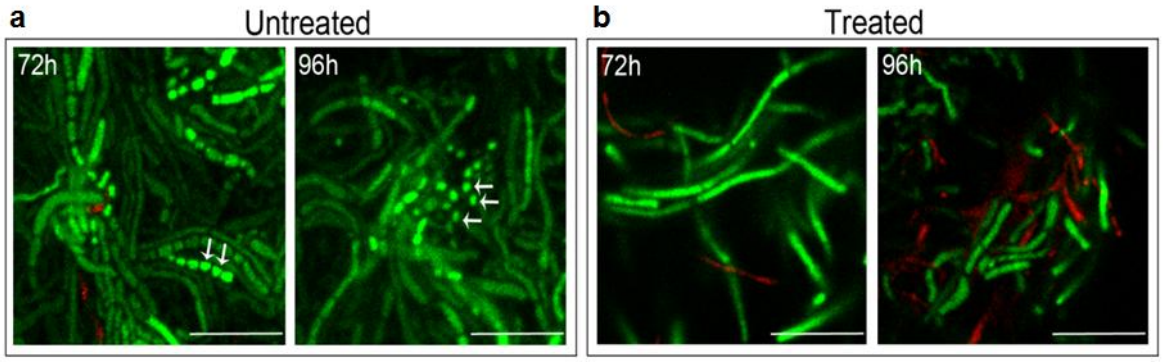

Figure 41 CLSM analysis of *S. coelicolor* grown for 72 and 96h on solid medium GYM (a,b), untreated and treated with aza-dC (5 µM). Images correspond to culture preparations stained with SYTO 9 and PI. Culture time points (hours) are indicated.

To check if the effect on morphological differentiation was a side effect due to the reduced and delayed germination, aza-dC was added only after 48h (MII aerial phase) of growth on solid GYM. The untreated samples at 72h and 96h analyzed by CLSM showed respectively viable spore chains (Figure 42a); on the contrary, the treated samples showed the multinucleate secondary mycelium (MII) characterized by the non-septate branching hyphae (Figure 42b). Thus, the sporulation phase was blocked in the treated culture regardless of delay germination. Thus, this result demonstrated that cytosine methylation exerts a role in the beginning of the sporulation phase.

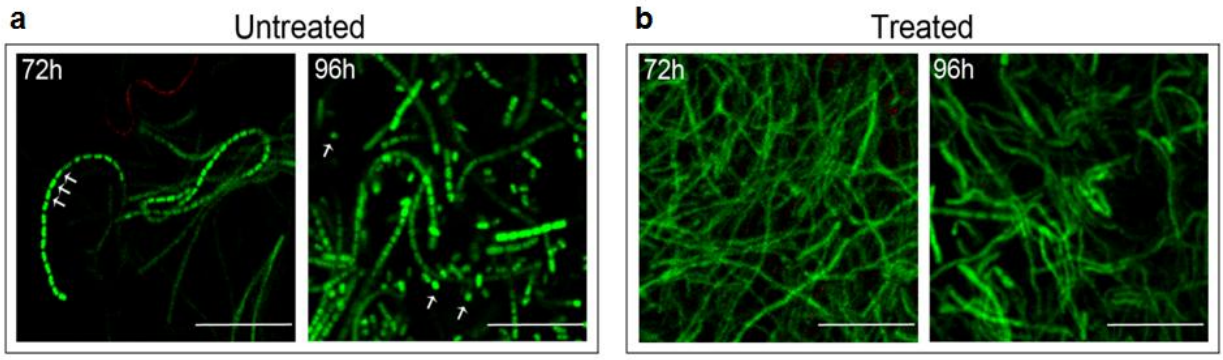

Figure 42 CLSM analysis of *S. coelicolor* grown for 72 and 96h (b,c), untreated and treated with aza-dC (5 µM) after 48h of growth on solid medium GYM. Images correspond to culture preparations stained with SYTO 9 and PI. Culture time points (hours) phase are indicated.

3.5 Effect of demethylation on physiological differentiation

To determine the effects of DNA cytosine methylation on physiological differentiation of *S. coelicolor*, production of the antibiotics undecylprodigiosin and actinorhodin was evaluated in the untreated and the treated culture using GYM.

In the treated culture, the undecylprodigiosin production, characterized by the red pigment, and actinorhodin production, characterized by the blue pigment, was decreased in respect to the untreated culture (Figure 43a-b).

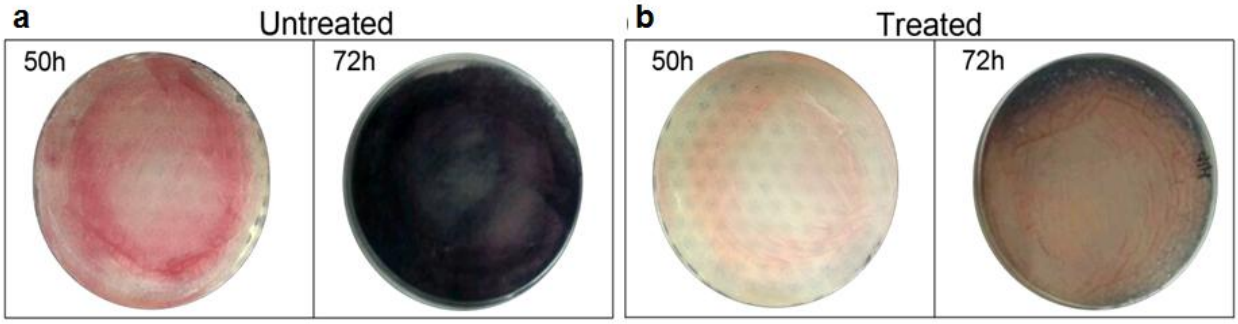

Figure 43 Effect of aza-dC on undecylprodigiosin and actinorhodin production on solid GYM.

3.6 DNA cytosine methylation and gene expression

The list of the genes containig methylated cytosines, on solid GYM, was compared to the list of differentially transcribed genes previously obtained in the same solid medium (34).

3.6.1 GGC^{m}CGG consensus sequence

Interestingly, it was found that ~700 (Table 16) genes, containing the GGC^{m}CGG consensus sequence methylated in MI and MII substrate phases, were more transcribed in MII aerial and sporulating phases. Among them, *SCO5587 (ftsH-like), SCO2716 (chpA), SCO5080 (actVA5), SCO5090 (actV-ORFII), SCO5091 (actIV), SCO5092 (actVB), SCO5879 (redW), SCO5896 (redH), SCO5881 (redZ), SCO6681 (ramC), SCO6685 (ramR)* were identified (Table 6). Differently, it was found that ~200 genes containing the GGC^{m}CGG cytosine methylation

consensus sequence repeated twice in their regulation region both in MI and in MII substrate and aerial phases, were constitutively transcribed, like *SCO6635 (pglY), SCO5897 (redG), SCO5898 (redF), SCO4423 (afsK)* and *SCO5820 (hrdB).*

Table 6 Example of genes containing the cytosine methylation consensus sequence GGC^{m}CGG correlated with their gene expression. *the same methylated consensus sequence was found twice. (The entire list of genes is listed in Table 16).

Gene	Description	Function	Consensus sequence	Methylation phase	Expression phase
SCO5587	*ftsH-like*	cell division/septation	GGCCGG	MI/MII substrate	MII aerial/sporulating
SCO2716	*chpA*	Differentiation	GGCCGG	MI/MII substrate	MII aerial/sporulating
SCO5881	*redZ*	secondary metabolism	GGCCGG	MI/MII substrate	MII aerial/sporulating
SCO5080	*actVA5*	secondary metabolism	GGCCGG	MI/MII substrate	MII aerial/sporulating
SCO5088	*actIORF2*	secondary metabolism	GGCCGG	MI/MII substrate	MII aerial/sporulating
SCO5090	*actVORFII*	secondary metabolism	GGCCGG	MI/MII substrate	MII aerial/sporulating
SCO5091	*actIV*	secondary metabolism	GGCCGG	MI/MII substrate	MII aerial/sporulating
SCO5092	*actVB*	secondary metabolism	GGCCGG	MI/MII substrate	MII aerial/sporulating
SCO5879	*redW*	secondary metabolism	GGCCGG	MI/MII substrate	MII aerial/sporulating
SCO5892	*redL*	secondary metabolism	GGCCGG	MI/MII substrate	MII aerial/sporulating
SCO5896	*redH*	secondary metabolism	GGCCGG	MI/MII substrate	MII aerial/sporulating
SCO6681	*ramC*	Differentiation	GGCCGG	MI/MII substrate	MII aerial/sporulating
SCO6685	*ramR*	Differentiation	GGCCGG	MI/MII substrate	MII aerial/sporulating
SCO6635	*pglY*	nucleases/RM methylases	GGCCGG*	MI/MIIsubstrate/MIIaerial	Unchanged
SCO5897	*redG*	secondary metabolism	GGCCGG*	MI/MIIsubstrate/MIIaerial	Unchanged
SCO5898	*redF*	secondary metabolism	GGCCGG*	MI/MIIsubstrate/MIIaerial	Unchanged
SCO4423	*afsK*	transcriptional regulators	GGCCGG*	MI/MIIsubstrate/MIIaerial	Unchanged
SCO5820	*hrdB*	transcriptional regulators	GGCCGG*	MI/MIIsubstrate/MIIaerial	Unchanged

To confirm these findings, the transcriptional analysis of genes *SCO5881 (redZ)* and *SCO6685 (ramR)* was performed. The genes *SCO5881 (redZ)*, coding for the response regulator involved in undecylprodigiosin production cascade, and *SCO6685 (ramR)*, coding for the response regulator involved in morphological differentiation, contain the GGC^{m}CGG consensus sequence

in their upstream region, that it was found to be methylated in MI and MII substrate phase and expressed in MII aerial and sporulating phase, thus probably the methylation of this consensus sequence has a inhibitory effect on the gene expression in MI and MII substrate.

Transcriptional levels of these genes were evaluated at time point 15h, 30h, 48h and 72h of *S. coelicolor* grown in GYM with or without aza-dC, corresponding to MI, MII substrate, MII aerial and MII sporulating phase, respectively.

The transcription levels of these genes in MII aerial and sporulating treated samples showed no difference in respect to the untreated sample; differently the transcription levels in MI and MII substrate treated samples were higher (for *redZ*, 8.7 and 2.7 fold, for *ramR*, 3.5 and 2 fold, respectively) in respect to the MI untreated sample (Figure 44).

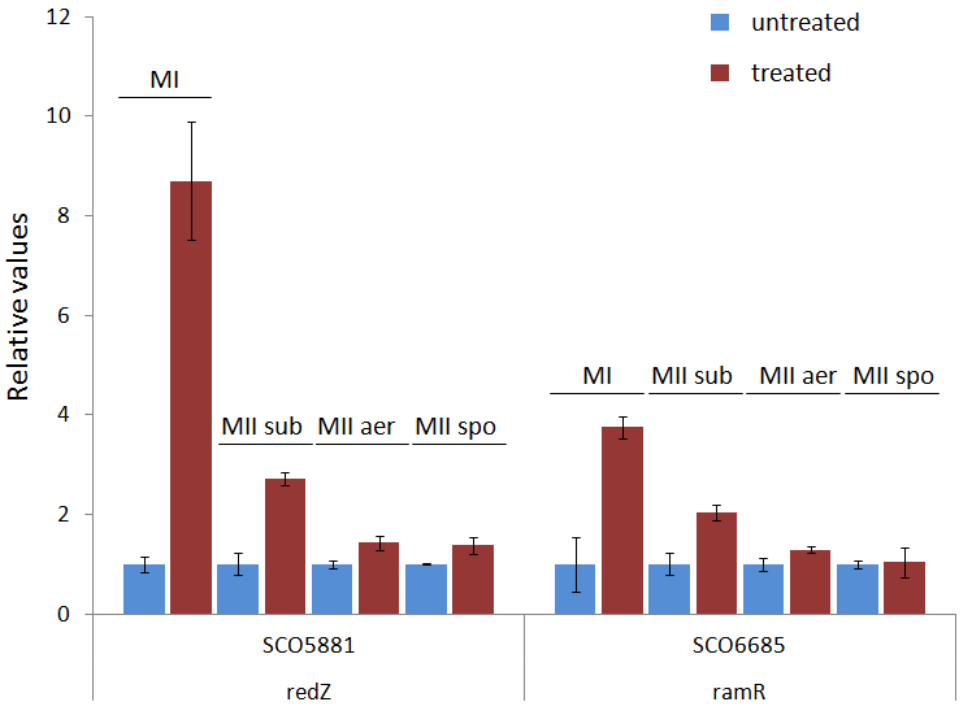

Figure 44 qRT-PCR analysis of *SCO5881* and *SCO6685* in MI, MII sub (substrate), MII aer (aerial) and MII (spo) sporulating phase in the untreated (blue bars) and aza-dC-treated (red bars). mRNA levels are expressed as relative to 16S rRNA transcripts, with the ratio values for the untreated samples arbitrarily set to 1. The standard deviations (indicated by error bars) were calculated from three independent determinations of mRNA abundance in each.

These results suggest that the methylation of GGC^{m}CGG had an inhibitory effect on transcriptional levels of both genes in MI and MII substrate. In fact, the induced de-methylation, in MI and MII substrate treated samples, resulted in an increased transcription level of the genes *SCO5881* and *SCO6685*.

3.6.2 GCC^{m}CG consensus sequence

Moreover, it was found that ~250 (Table 17) genes, containing the GCC^{m}CG cytosine methylation consensus sequence methylated in MII substrate and aerial phase, were more transcribed in one or two of the three MII phases (substrate, aerial and sporulating), such as *SCO2077 (divIVA), SCO5078 (actVA3), SCO5085 (actII-ORF4), SCO1489 (bldD)* and *SCO5113 (bldK)* (Table 7), but a direct link was not straightforward to understand.

Table 7 Example of gGenes containing the cytosine methylation consensus sequence GCC^{m}CG correlated with their gene expression. (The entire list of genesis listed in Table 17).

Gene	Description	Function	Consensus sequence	Methylation phase	Expression phase
SCO2077	*divIVA*	cell division/septation	GCCCG	MIIsubstrate/MIIaerial	MII substrate/aerial
SCO5078	*actVA3*	secondary metabolism	GCCCG	MIIsubstrate/MIIaerial	MII aerial/sporulating
SCO5085	*actII-ORF4*	secondary metabolism	GCCCG	MIIsubstrate/MIIaerial	MII aerial/sporulating
SCO1489	*bldD*	Differentiation	GCCCG	MIIsubstrate/MIIaerial	MII substrate
SCO5113	*bldKB*	Differentiation	GCCCG	MIIsubstrate/MIIaerial	MII substrate

To find a link, the transcriptional analysis of genes *SCO5085 (actII-ORF4), SCO2077 (divIVA)* and *SCO1489 (bldD)* was performed.

The genes *SCO5085 (actII-ORF4),* coding for the response regulator involved in actinorhodin production cascade, *SCO2077 (divIVA),* coding for the protein involved in polar growth, and *SCO1489 (bldD),* coding for a DNA-binding protein involved in morphological differentiation, contain the GCC^{m}CG consensus sequence in their upstream region; it was found to be

methylated in MII substrate and aerial phase and expressed in MII substrate and aerial phase, thus probably the methylation of this consensus sequence activates the gene expression.

The transcription levels of these genes in treated MI and MII sporulating samples showed no difference in respect to the untreated samples; differently, the transcription levels in the treated MII substrate and aerial samples were lower (for *actII*-ORF4, 1.5 fold, for *divIVA*, 1.5 fold and for *bldD*, 2 fold) in respect to the untreated samples (Figure 45).

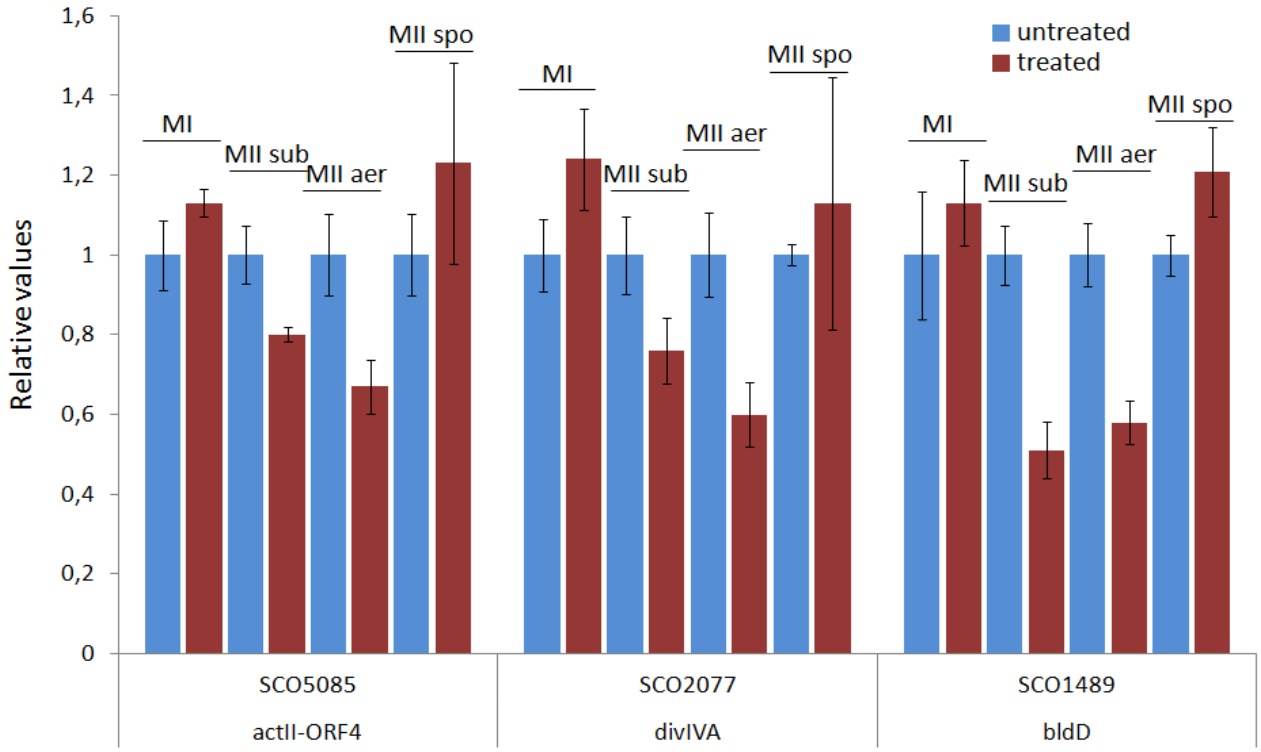

Figure 45 qRT-PCR analysis of *SCO5085*, *SCO2077* and *SCO1489* in MI, MII sub (substrate), MII aer (aerial) and MII (spo) sporulating phase under conditions of untreated (blue bars) and treated with aza-dC (red bars). mRNA levels are expressed relative to 16S rRNA transcripts, with the ratio values for the untreated samples arbitrarily set to 1. The standard deviations (indicated by error bars) were calculated from three independent determinations of mRNA abundance in each.

This result suggests that the methylation of the GCC^{m}CG consensus sequence had a stimulating effect on transcriptional levels in substrate and aerial MII. In fact, the demethylation of this consensus sequence, in MII substrate and aerial treated samples, decreased the transcription of all the three genes analyzed.

3.6.3 C^{m}GGGC consensus sequence

On the contrary, it was found that ~300 genes (Table 18), containing the C^{m}GGGC cytosine methylation consensus sequence methylated in MII substrate and aerial phase, were more transcribed in MI phase, like *SCO2571 (leuS), SCO3911 (dnaB), SCO0895 (hrdC), SCO3068 (sigI)* and *SCO3892 (sigT)* (Table 8).

Table 8 Example of genes containing the cytosine methylation consensus sequence C^{m}GGGC correlated with their gene expression. (The entire list of genes is listed in Table 18).

Gene	Description	Function	Consensus sequence	Methylation phase	Expression phase
SCO2571	*leuS*	primary metabolism	CGGGC	MIIsubstrate/MIIaerial	MI
SCO3911	*dnaB*	primary metabolism	CGGGC	MIIsubstrate/MIIaerial	MI
SCO0895	*hrdC*	transcriptional regulators	CGGGC	MIIsubstrate/MIIaerial	MI
SCO3068	*sigI*	transcriptional regulators	CGGGC	MIIsubstrate/MIIaerial	MI
SCO3892	*sigT*	transcriptional regulators	CGGGC	MIIsubstrate/MIIaerial	MI

To confirm these findings, the transcriptional analysis of genes *SCO3911 (dnaB)* and *SCO2571 (leuS)* was performed.

The genes *SCO3911 (dnaB)*, coding for a replicative DNA helicase, and *SCO2571 (leuS)*, coding for a leucyl-tRNA synthetase, contain the C^{m}GGGC consensus sequence in their upstream region; it was found to be methylated in substrate and aerial MII and expressed in MI phase, thus probably the methylation of this consensus sequence inhibits the gene expression.

The transcription levels of these genes in each treated sample showed no difference in respect to the untreated samples (Figure 46).

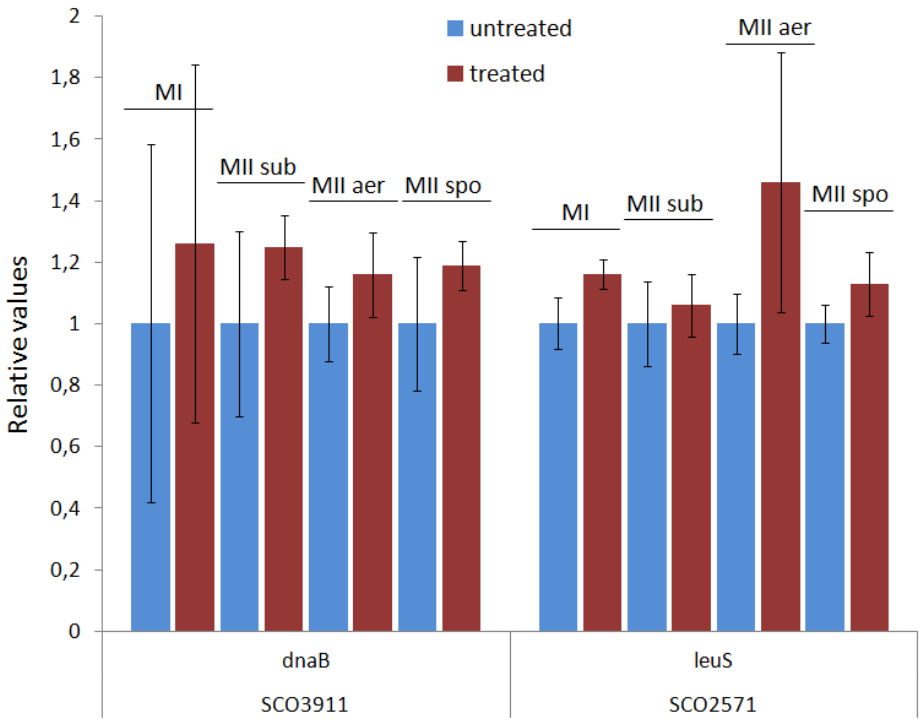

Figure 46 qRT-PCR analysis of *SCO3911* and *SCO2571* in MI, MII sub (substrate), MII aer (aerial) and MII (spo) sporulating phase under conditions of untreated (blue bars) and treated with aza-dC (red bars). mRNA levels are expressed relative to 16S rRNA transcripts, with the ratio values for the untreated samples arbitrarily set to 1. The standard deviations (indicated by error bars) were calculated from three independent determinations of mRNA abundance in each.

On the basis of the obtained results about the effect of the treatment carried out with aza-dC on solid medium, it was shown that methylation of the GGC^{m}CGG consensus sequence inhibits gene expression; and methylation of the GCC^{m}CG consensus sequence activates gene expression, while, for the analyzed genes (*SCO3911* and *SCO2571*) no regulatory function was found for methylation of the C^{m}GGGC. The same result was obtained with the liquid R5A, indicating that the methylation level of these genes do not depend upon the medium, by upon the growth phase.

4. Analysis of *S. coelicolor* cytosine methylome

4.1 Similarities and differences between cytosine methylome in liquid and solid medium

The cytosine methylome analysis revealed that 1525 genes were methylated during the MI phase in R5A and 1437 during the MI phase in GYM

The number of genes, containing methylated cytosines in the regulation region in MI, was compared between solid GYM and liquid MG.

A total of 1303 genes were methylated in the region comprised between -400 and +100 bp of genes both in MI solid and MI liquid phase, as depicted by a Venn diagram (Figure 47). Only 134 genes were methylated on solid medium and 222 in liquid medium

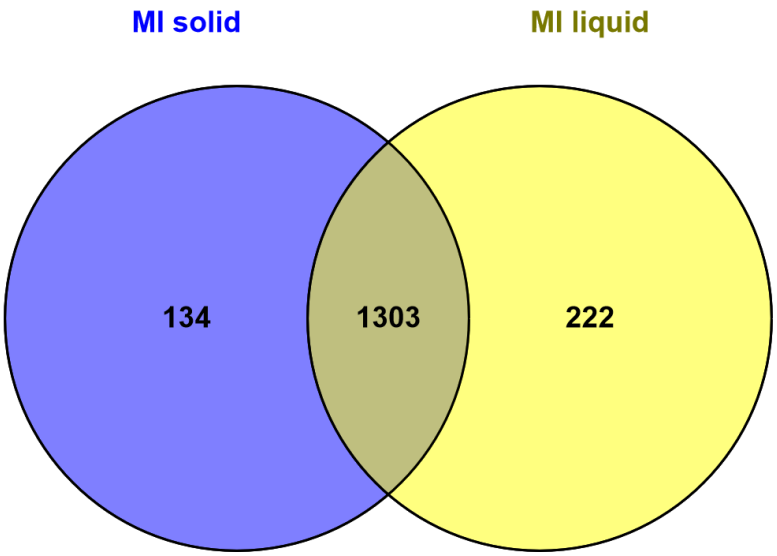

Figure 47 Venn diagram showing the comparison between genes identified in MI solid GYM and MI liquid R5A phase.

This result shows that the MI phases between solid and liquid medium are similar.

Moreover, 1590 genes were methylated during the MII substrate phase in solid GYM (the substrate phase is more similar at MII phase of liquid R5A), and 907 during the MII phase in liquid R5A.

The number of genes was compared between MII solid GYM and MII liquid R5A. A total of 769 genes were methylated both in MII solid and MII liquid phase in the region comprised between -400 and +100 bp of genes, as shown by a Venn diagram (Figure 48). 821 genes were methylated only on solid medium and 138 only in liquid.

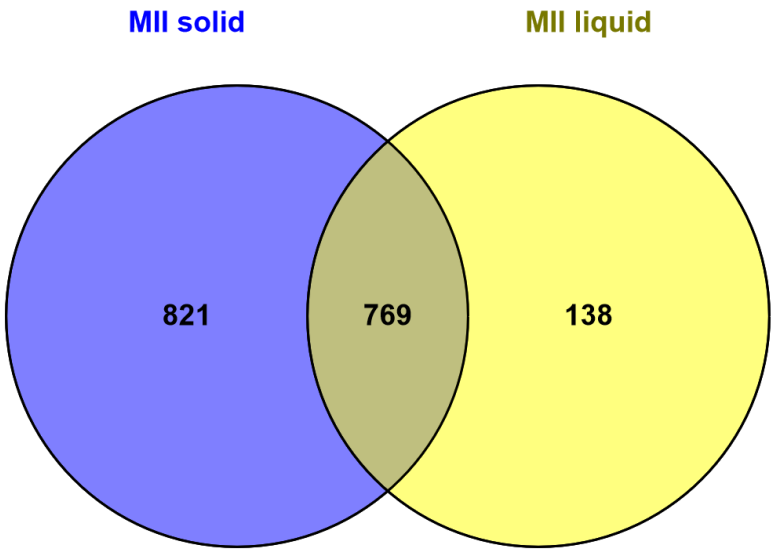

Figure 48 Venn diagram showing the comparison between genes identified in MII solid substrate GYM and MII liquid R5A phase.

This result showed that the MII phase between solid and liquid medium are similar, as obtained for the MI phase.

Finally, 730 genes methylated in the MII phase in liquid R5A were compared to the 325 genes found methylated in the MII phase in MG.

A total of 127 genes was methylated both in MII liquid R5A and MI liquid MG in the region comprised between -400 and +100 bp of genes. as shown by a Venn diagram (Figure 49).

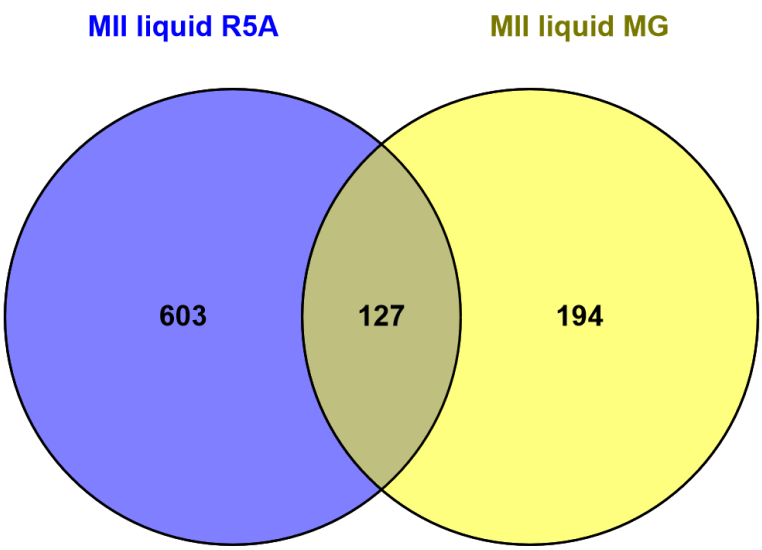

Figure 49 Venn diagram showing the comparison between genes identified in MII liquid R5A and MII liquid MG.

This result revealed that the MII phase in liquid medium R5A and MG do not share many genes.

5. Role of the DNA methyltransferase *SCO1731*

5.1 Construction of a knock out mutant in *SCO1731* gene

A bioinformatics search for 30 genes coding for putative DNA methyltransferases into *S. coelicolor* genome revealed that it contains 17 genes coding for putative DNA (5-cytosine)-methyltransferases (Table 9). Among the putative DNA (5-cytosine)-methyltransferases, the search for methyltransferase genes differentially transcribed in R5A and GYM media was carried out using the transcriptomic results previously obtained (31).

Table 9 List of putative cytosine-methyltransferases in *S. coelicolor.* Underlined the methylatransferases for which mutants were generated.

Putative 5-cytosine-methyltransferase	Expression phase
1. SCO1731	MI
2. SCO5589	MI
3. SCO3215	MI
4. SCO7445	MI
5. SCO7055	MI
6. SCO3545	MI
7. SCO5972	MI
8. SCO0760	MI
9. SCO0995	MIIsubstrate/aerial
10. SCO5257	MIIsubstrate
11. SCO0826	MIIsubstrate/aerial/sporulating
12. SCO0929	MIIsubstrate
13. SCO6541	MIIsubstrate/aerial/sporulating
14. SCO6549	MIIsubstrate/aerial/sporulating
15. SCO5895	MIIaerial/sporulating
16. SCO3744	MIIsubstrate/sporulating
17. SCO6844	Unchanged

This search revealed that the putative DNA(5-cytosine)-methyltransferase *SCO1731* is mainly expressed in MI phase, both on solid and in liquid culture, and *SCO0995*, is mainly expressed in MII phase, both on solid and in liquid culture. Thus, to evaluate if these genes are important for DNA methylation, two independent knock-out mutants were generated.

To generate the mutant in *SCO1731* gene, a cosmid containing the gene replaced by the transposon Tn*5* (Figure 50) was used to perform deletion (21). It contains the apramycin and kanamycin resistance cassettes in Tn*5* and in the cosmid, respectively,

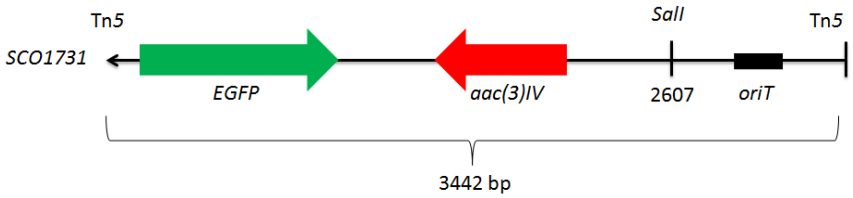

Figure 50 Map of the trasposon Tn*5* used for the generation of mutant containing *EGFP* (green fluorescent protein), *aac(3)IV* (apramycin resistance cassette) and *oriT* (plasmid transfer origin).

8 colonies of putative mutant *ΔSCO1731* were obtained after the growth in SFM (sporulating medium) with apramycin. After 3 passages of these colonies on GYM with apramycin, genomic, DNA was extracted from 4 mutants and analyzed by PCR for the presence of apramycin (~1370 bp) and the absence of kanamycin (902 bp) resistance cassette.

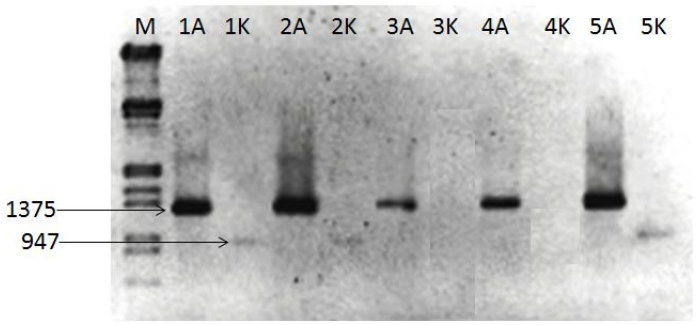

Figure 51 PCR products derived from A (apramycin 1370bp) and K (Kanamycin 902bp) primer sets were separated using a 0,8% agarose gel in 1X TBE buffer. Lane numbers correspond to M: Lambda DNA/*EcoRI+HindIII;* genomic DNA from 4 clones 1A-K: Δ*SCO1731* -1; 2A-K: *SCO1731* -2; 3A-K: Δ*SCO1731* -3; 4A-K: Δ*SCO1731* -4; 5A: apramycin positive control; 5K: kanamycin positive control.

Only two samples (Δ*SCO1731*-3 and -4) had the expected profile and they were analyzed by Southern Blot using genomic DNA digested with *SalI* and Tn*5* as a probe.

The expected restriction profile of the knock-out mutant is shown in the Figure 52. Southern blot analysis revealed that both putative mutants Δ*SCO1731* had the expected restriction profile.

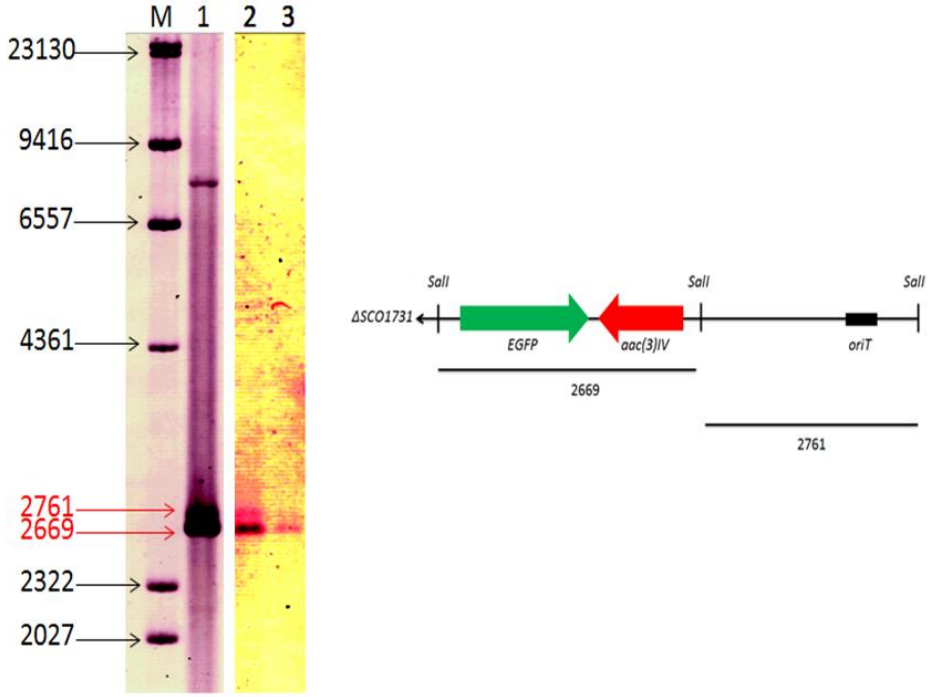

Figure 52 Southern blot analysis and restriction profile of 2 putative mutants of *ΔSCO1731*. M: DNA molecular weight Marker II (Roche); 1: *SalI*-digested cosmid DNA containing resistance cassette of apramycin in *SCO1731*; 2: *SalI*-digested genomic DNA Δ*SCO1731* -3; 3: *SalI*-digested genomic DNA Δ*SCO1731* -4. pQM5062 was used as a probe.

5.2 Effect of the inactivation of *SCO1731* gene

In order to determine the importance of the methyltransferase gene *SCO1731*, growth, antibiotic production and morphological differentiation were analyzed on solid GYM and in liquid R5A.

In liquid R5A, disruption of *SCO1731* did not significantly alter the growth kinetics of the mutant (Figure 53), indicating that this gene is not critical for bacterial growth under the used conditions.

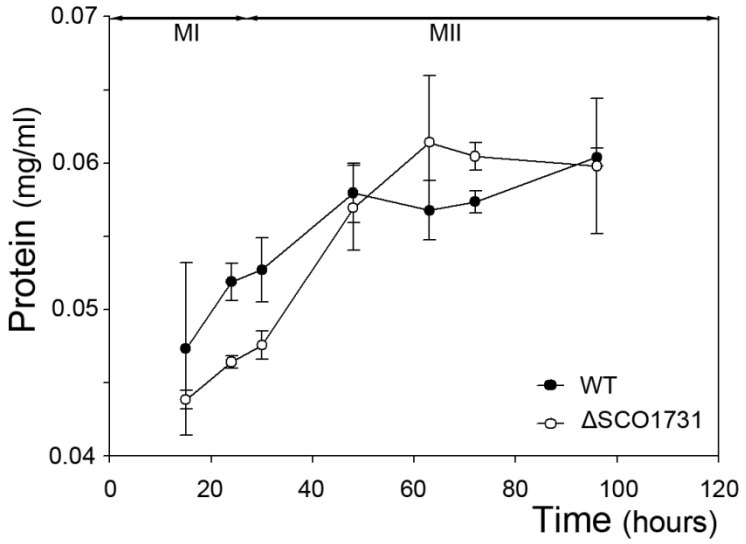

Figure 53 Growth curves in liquid R5A of *S. coelicolor* (WT) indicated by black circles and mutant Δ*SCO1731* indicated by white circles.

In contrast, the undecylprodigiosin and actinorhodin production of Δ*SCO1731* strain in liquid R5A occurs later than the wild type, after 72h and 130h of growth rather than 48h and 72h, respectively (Figure 54).

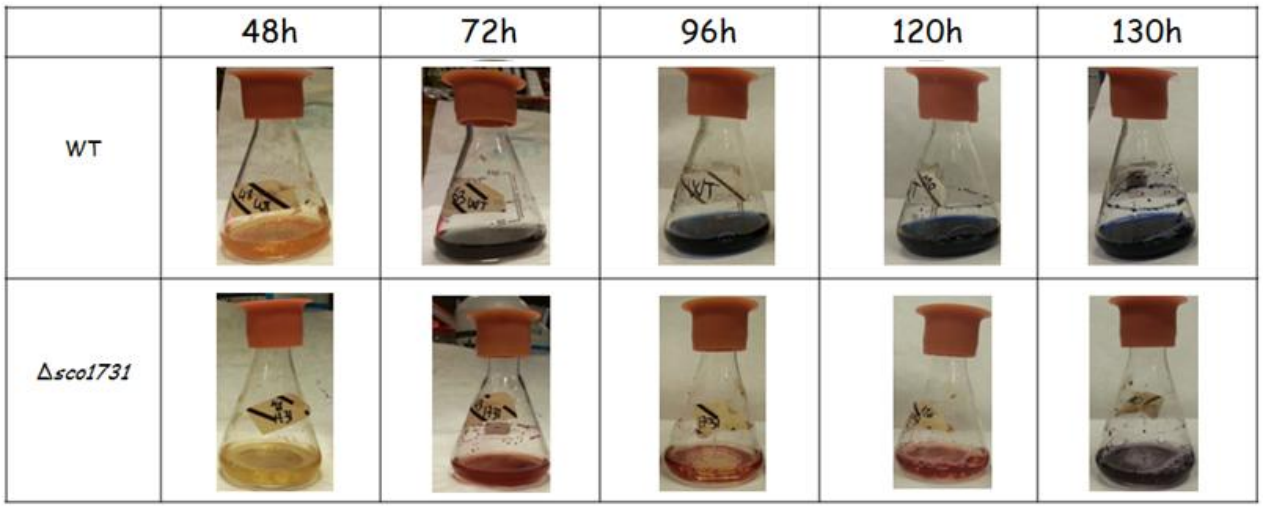

Figure 54 Antibiotic production of *S. coelicolor* (WT) and the mutant Δ*SCO1731* in liquid R5A.

On solid GYM, the inactivation of *SCO1731* caused a delay in physiological and morphological differentiation (Figure 55), in fact, the aerial mycelium formation and the actinorhodin

production start in the mutant Δ*SCO1731* at 96h rather than at 48h; at the end, spore formation occurs after 10 days of growth in solid SFM in the mutant, while the wild type forms spores after 5 days (data not shown).

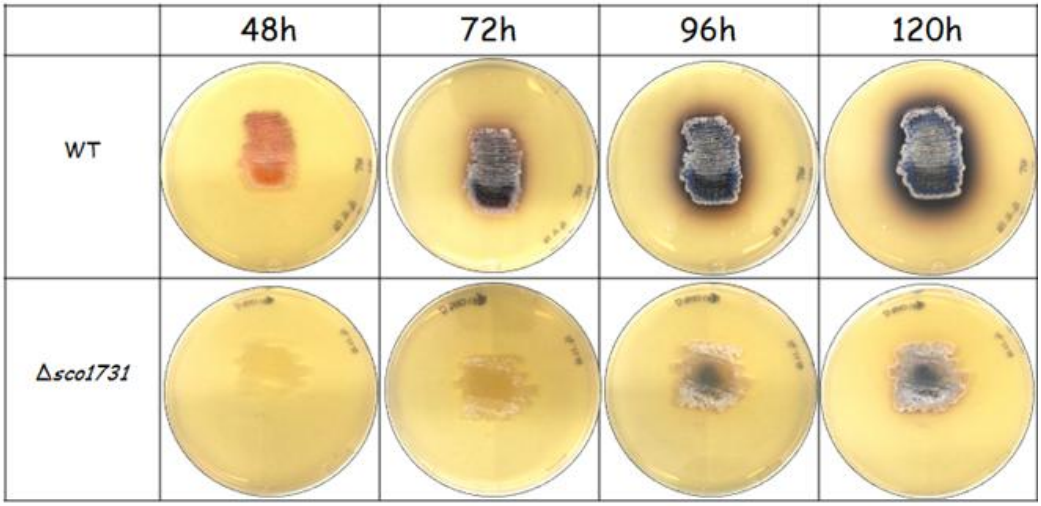

Figure 55 Phenotypic analysis and antibiotic production of *S. coelicolor* (WT) and mutant Δ*SCO1731* on solid GYM.

These results indicate that SCO1731 is involved in the regulation of antibiotic production and morphological differentiation.

5.3 Complementation of mutant Δ*SCO1731* and construction of a strain expressing two copies of the *SCO1731* gene

To complement the mutation, pNG3 integrating vector (37) with a copy of *SCO1731*, placed under the control of its promoter, was used. The plasmid was used to transform the mutant Δ*SCO1731* strain by interspecific conjugation generating the Δ*SCO1731*_compl strain.

Thus, 32 colonies of putative complemented *ΔSCO1731* were obtained after the growth in SFM with hygromycin. After 3 passages of these colonies on GYM with hygromycin, genomic DNA was extracted by 10 putative complemented strains and analyzed by PCR using the primers

SCO4848_F/R. These primers amplified a fragment of 617 bp only if pNG3 is integrated at the *attB* site of *SCO4848* (Figure 56).

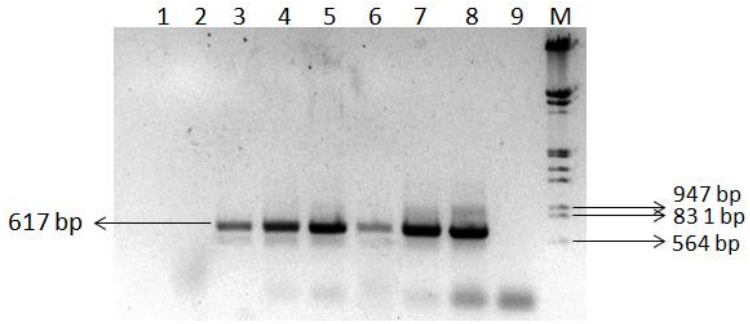

Figure 56 PCR products of 617 bp, derived from Δ*SCO1731*_compl strains, were separated using a 0,8% agarose gel in 1X TBE buffer. Lane numbers correspond to genomic DNA from 1: Δ*SCO1731*_compl-1; 2: *SCO1731*_compl-2; 3: Δ*SCO1731*_compl-3; 4: Δ*SCO1731*_compl-4; 5: Δ*SCO1731*_compl-5; 6: Δ*SCO1731*_compl-6; 7: Δ*SCO1731*_compl-7; 8: Δ*SCO1731*_compl-8; 9: negative control; M: Lambda DNA/*EcoRI+HindIII*

Six of eight samples had the expected profile and only the Δ*SCO1731*_compl-10 was used for further experiment.

In order to investigate if one copy extra of *SCO1731* in *S. coelicolor* caused an opposite physiological effect, the same vector with a copy of *SCO1731* was used to transform *S. coelicolor* (WT) (1copy_*SCO1731* strain).

Thus, 24 colonies of putative complemented *ΔSCO1731* were obtained after the growth in SFM with hygromycin. After 3 passages of these colonies on GYM with hygromycin, genomic DNA was extracted by 12 putative complemented strains and analyzed by PCR using the primers SCO4848_F/R. These primers amplified a fragment of 617 bp only if pNG3 is integrated at the *attB* site of *SCO4848* (Figure 57).

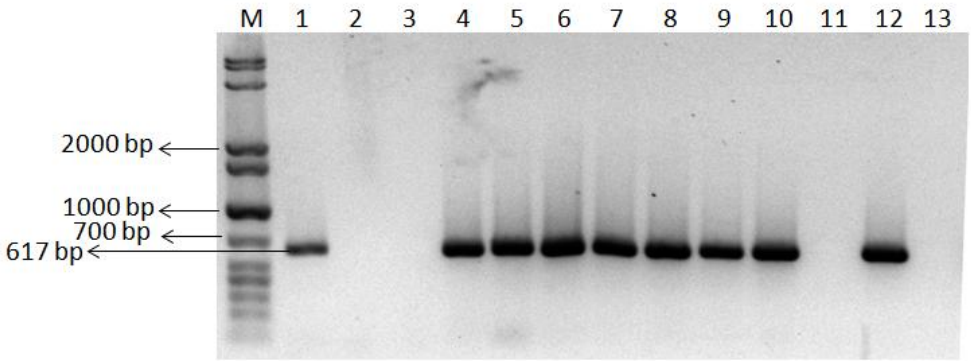

Figure 57 PCR products of 617 bp derived from 1copy_*SCO1731* strains were separated using a 0,8% agarose gel in 1X TBE buffer. Lane numbers correspond to genomic DNA from 1: 1copy_*SCO1731*-1; 2: 1copy_*SCO1731*-2; 3: 1copy_*SCO1731*-3; 4: 1copy_*SCO1731*-4; 5: 1copy_*SCO1731*-5; 6: 1copy_*SCO1731*-6; : 1copy_*SCO1731*-7; 8: 1copy_*SCO1731*-8; 9: 1copy_*SCO1731*-9; 10: 1copy_*SCO1731*-10; 11: 1copy_*SCO1731*-11; 12: 1copy_*SCO1731*-12; 13:negative control; M: 1 kb (PerfectSize DNA Molecular Weight Ladder, 5-Prime).

Nine of twelve samples had the expected profile and only the 1copy_*SCO1731*-6 was used for further experiment.

In agreement with the results obtained with the *SCO1731* mutant, no differences were observed between growth of the parental, Δ*SCO1731*_compl and 1copy_*SCO1731* strains in liquid R5A medium (Figure 58).

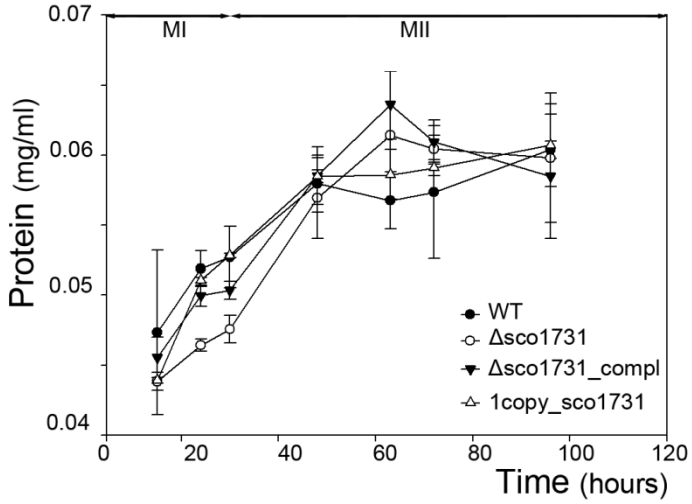

Figure 58 Growth curves in liquid R5A of *S. coelicolor* (WT) indicated by black circles, mutant Δ*SCO1731* indicated by white circles, the complemented mutant Δ*SCO1731*_compl indicated by black triangles and 1copy_*SCO1731* indicated by white triangles.

The undecylprodigiosin production (Figure 59A) is not affected in Δ*SCO1731* strain.

The actinorhodin production is completely absent in Δ*SCO1731* strain. The strain Δ*SCO1731*_compl restored, even if not completely, the actinorhodin production (Figure 59B).

The control strain, containing a copy of the empty vector (WT+pNG3), showed a production profile similar to that of the parental strain, excluding an effect of pNG3 integration on antibiotic production (data not shown).

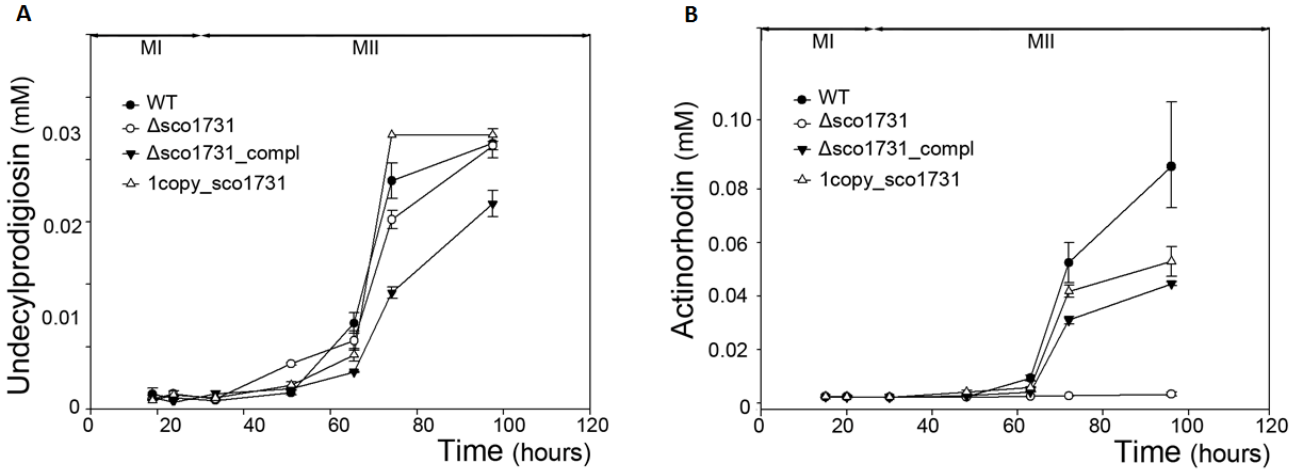

Figure 59 Quantitative analysis of undecylprodigiosin **A)** and actinorhodin **B)**production of *S. coelicolor* WT (black circles), *ΔSCO1731* (white circles), *ΔSCO1731*_compl (black triangles) and 1copy_*SCO1731* (white triangles) strains in liquid R5A. 10^8 spores per each strain were inoculated.

Macroscopic view and CLSM analysis performed on solid GYM showed that the Δ*SCO1731*_compl restored the correct morphological development with aerial and spore production (Figure 60 shows, in the picture of CLSM, spore chains at 72h).

The morphological development of 1copy_*SCO1731* was the same as *S. coelicolor* WT.

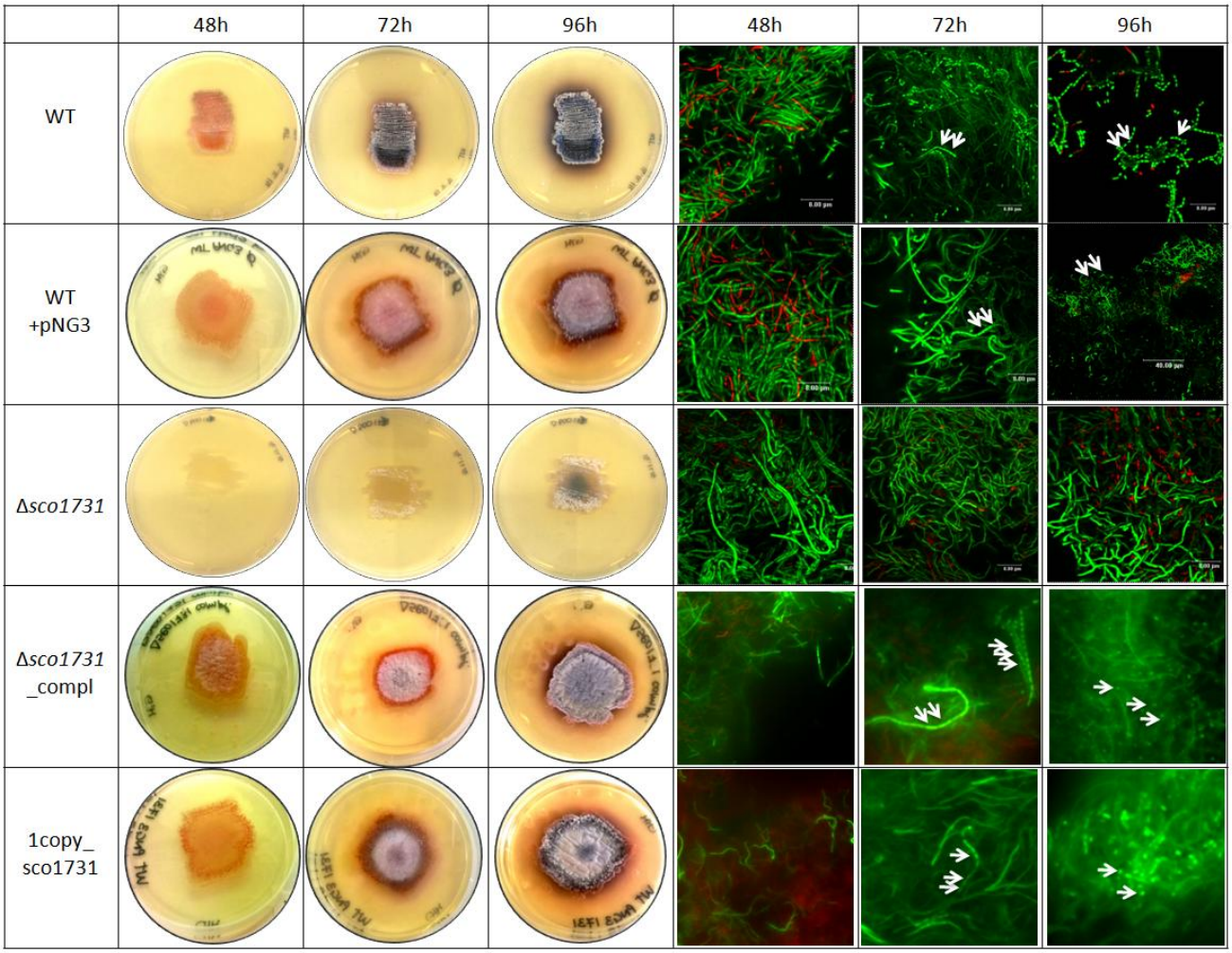

Figure 60 Macroscopic and CLSM analysis view of *S. coelicolor* WT, WT+pNG3*, ΔSCO1731*, *ΔSCO1731*_compl and 1copy_*SCO1731* on solid GYM. White arrows indicate spore chains (73h) and spore (96h). Images in CLSM correspond to culture preparations stained with SYTO 9 and PI.

It would be interesting to carry out the analysis of the mutant methylome to compare it with the wild type methylome and to know which genes are methylated by SCO1731.

6. Role of the DNA methyltransferase *SCO0995*

6.1 Construction of a knock out mutant in *sc0995* gene

Among the putative DNA (5-cytosine)-methyltransferases (Table 9), *SCO0995* is mainly expressed in MII phase, both on solid and in liquid culture. Thus, to evaluate if this gene is important for DNA methylation, a knock-out mutant was generated. A cosmid containing the gene replaced by the transposon Tn*5* (Figure 61) was used to perform deletion (21). It contains the apramycin and kanamycin resistance cassettes in Tn*5* and in the cosmid, respectively. This strategy was used since the *SCO0995* should be the last gene of a polycistronic operon going from *SCO0991* to *SCO0995* on the basis of the intergenic sequence.

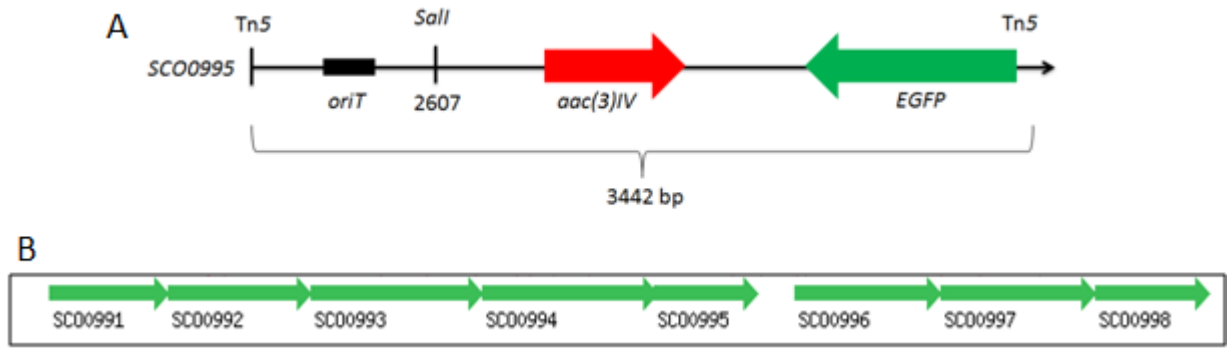

Figure 61 A) Map of the trasposon Tn*5* containing *EGFP* (green fluorescent protein), *aac(3)IV* (apramycin resistance cassette) and *oriT* (plasmid transfer origin). **B)** Map of the genes surrounding SCO0995, as described in StrepDB database.

7 colonies of putative mutant *Δsc0995* were obtained after the growth in SFM (sporulating medium) with apramycin. After 3 passages of these colonies on GYM with apramycin, genomic DNA was extracted from 3 mutants and analyzed by PCR for the presence of apramycin (~1370 bp) and the absence of kanamycin (902 bp) resistance cassette.

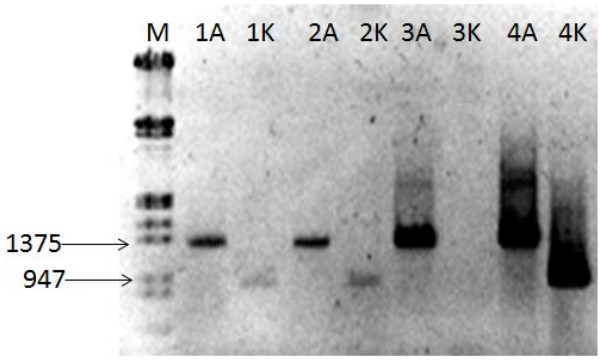

Figure 62 PCR products derived from A (apramycin 1370bp) and K (Kanamycin 902bp) primer sets were separated using a 0,8% agarose gel in 1X TBE buffer. Lane numbers correspond to M: Lambda DNA/*EcoRI+HindIII;* genomic DNA from 3 clones 1A-K: Δ*SCO0995* -1; 2A-K: *SCO0995* -2; 3A-K: Δ*SCO0995* -3; 4A apramycin positive control; 4K: kanamycin positive control.

Only one samples (Δ*SCO0995*-3) had the expected profile and it was further analyzed by Southern Blot using genomic DNA digested with *SalI* and Tn*5* as a probe.

The expected restriction profile of the knock-out mutant is shown in the Figure 63. Southern blot analysis revealed that the putative mutant Δ*SCO0995* had the expected restriction profile.

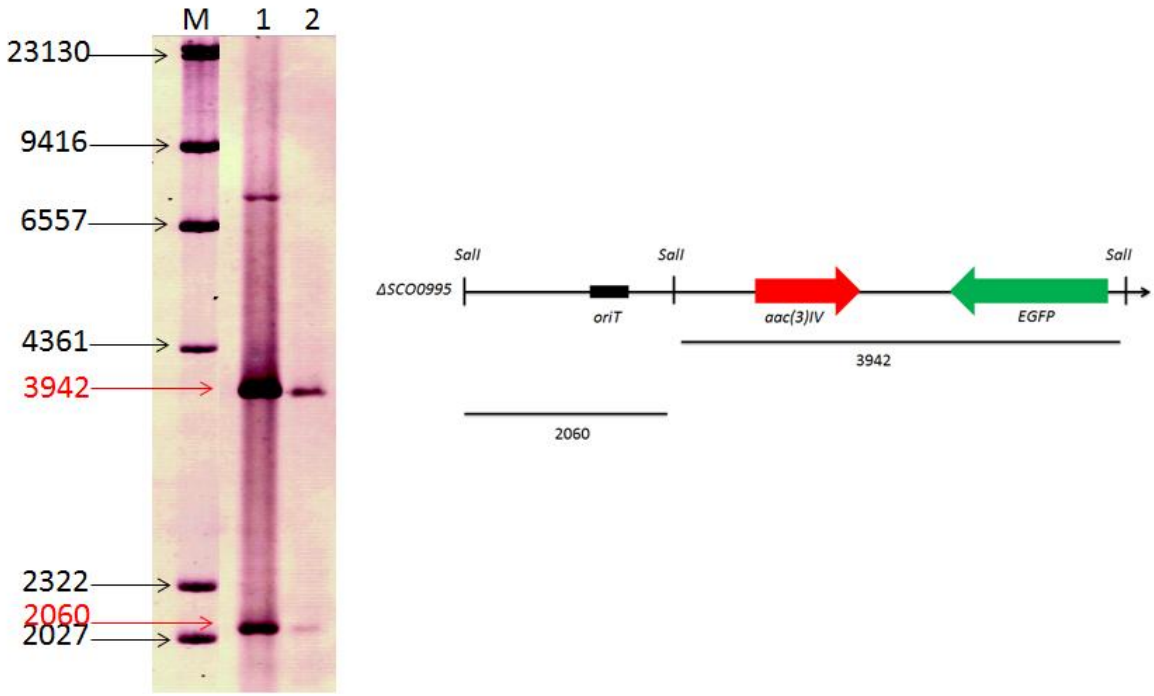

Figure 63 Southern blot analysis and restriction profile of the putative mutant of *ΔSCO0995*. M: DNA molecular weight Marker II (Roche); 1: *SalI*-digested cosmid DNA containing resistance cassette of apramycin in *SCO0995*; 2: *SalI*-digested genomic DNA of clone Δ*SCO0995* -3. pQM5062 was used as a probe.

6.2 Effect of the inactivation of *SCO0995* gene

In order to determine the importance of the methyltransferase gene *SCO0995*, growth, antibiotic production and morphological differentiation were analyzed on solid GYM.

The inactivation of *SCO0995* caused a delay in physiological and morphological differentiation (Figure 64), in fact, the aerial mycelium formation and antibiotic production started in the mutant Δ*SCO0995* at 72h rather than at 48h; surprisingly. at the end, spore formation occurred after 5 days of growth in solid SFM, like the wild type strain (data not shown).

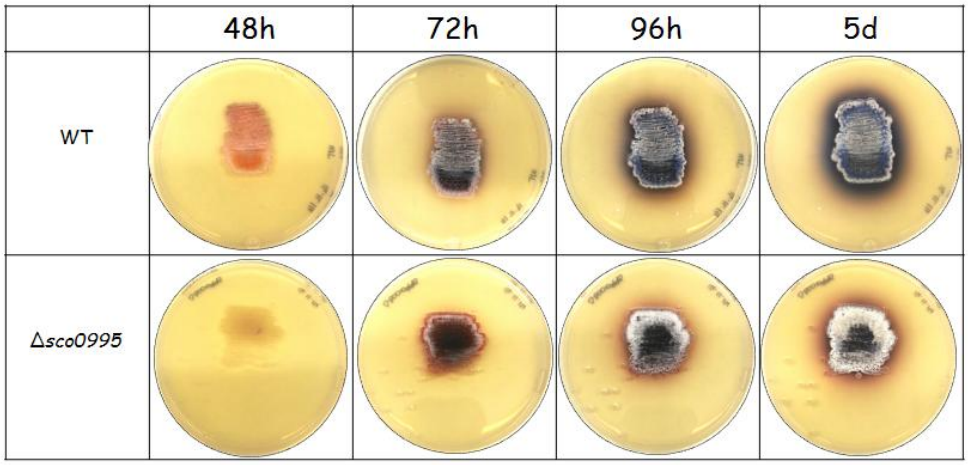

Figure 64 Phenotypic analysis and antibiotic production of *S. coelicolor* (WT) and mutant Δ*SCO1731* on solid GYM.

These results indicate that SCO0995 is involved in the regulation of antibiotic production and morphological differentiation.

6.3 Complementation of mutant Δ*SCO0995*

To complement the mutation, pNG2 integrating vector (37) with a copy of *SCO0995*, placed under the control of erythromycin promoter, was used. The plasmid was used to transform the mutant Δ*SCO0995* strain by interspecific conjugation generating the Δ*SCO0995*_compl strain.

Thus, 28 colonies of putative complemented *ΔSCO0995* were obtained after the growth in SFM with hygromycin. After 3 passages of these colonies on GYM with hygromycin, genomic DNA

was extracted by 10 putative complemented strains and analyzed by PCR using the primers SCO4848_F/R. These primers amplified a fragment of 617 bp only if pNG2 is integrated at the *attB* site of *SCO4848* (Figure 65).

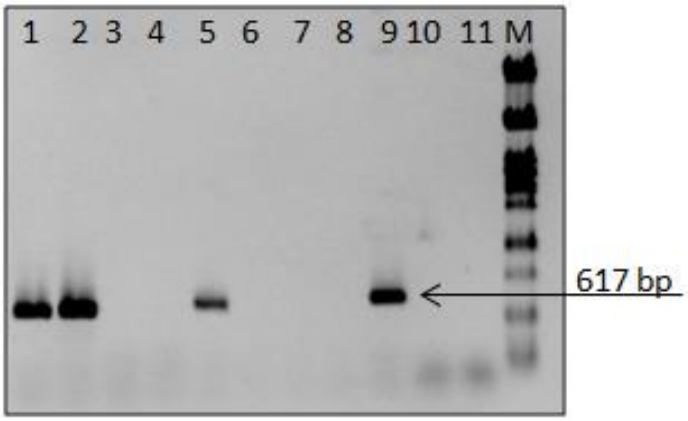

Figure 65 PCR products of 617 bp, derived from Δ*SCO0995*_compl strains, were separated using a 0,8% agarose gel in 1X TBE buffer. Lane numbers correspond to genomic DNA from 1: Δ*SCO0995*_compl-1; 2: *SCO0995*_compl-2; 3: Δ*SCO0995*_compl-3; 4: Δ*SCO0995*_compl-4; 5: Δ*SCO0995*_compl-5; 6: Δ*SCO0995*_compl-6; 7: Δ*SCO0995*_compl-7; 8: Δ*SCO0995*_compl-8; 9: Δ*SCO0995*_compl-9; 10: Δ*SCO0995*_compl-10; 11: negative control; M: Lambda DNA/*EcoRI+HindIII*

Four of ten samples had the expected profile and only the Δ*SCO0995*_compl-2 was used for further experiments.

Macroscopic view and CLSM analysis performed on solid GYM showed that the Δ*SCO0995*_compl did not restore the correct morphological development with aerial and spore production (data not shown).

Thus, a transcriptional analysis was performed to investigate if *SCO0995* was cotranscribed with *SCO0996.*

RT-PCR analysis, using the primers 0995-96F/R, revealed that *SCO0995* is cotranscribed with *SCO0996*. Thus, the phenotype is not due to *SCO0995*, but it could be due to a gene located downstream of *SCO0995* (*SCO0996-0997-0998*).

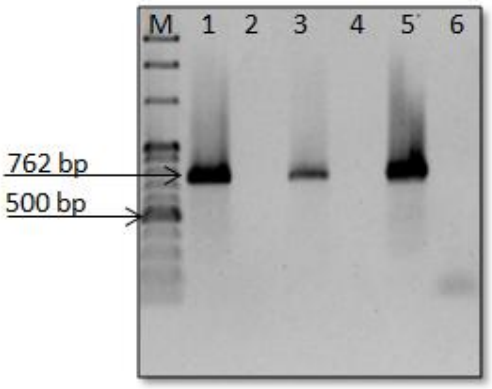

Figure 66 RT-PCR products of 762 bp, derived from *S. coelicolor* wt strains, were separated using a 0,8% agarose gel in 1X TBE buffer. Lane numbers correspond to genomic DNA from : M: Gene Ruler 100 bp plus DNA ladder; 1:RNA 24h with RT; 2: RNA 24h with taq; 3: RNA 48h with RT; 4: RNA 48h with taq; 5: positive control; 6: negative control.

All these genes code for putative membrane proteins, in particular *SCO0997-98* are involved in the Fe uptake system permease. Thus, no further studies on this mutant were carried out.

7. DNA cytosine methylation in several streptomycetes

To evaluate whether the DNA cytosine methylation is shared among *streptomycetes* and to demonstrate that chromosomal DNA cytosine methylation is independent from Restriction Modification (RM) system, dot blot analysis of *Sreptomyces lividans* strain 1326, *Streptomyces griseus* strain NBRC 13356 and *Streptomyces avermitilis* strain ATCC 31267 was performed.

S. griseus and *S. avermitilis*, like *S. coelicolor*, do not accept methylated DNA, while *S. lividans* accepts methylated DNA.

Despite *S. lividans* and *S. coelicolor* are different in accepting methylated DNA they have a very similar genome and the same development (32).

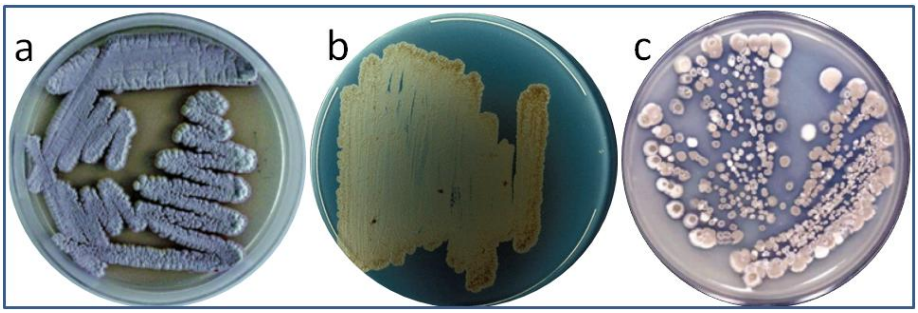

Figure 67 Streptomycetes grown on solid media (a) *S. lividans*, (b) *S. griseus* and (c) *S. avermitilis*.

Dot blot analysis of genomic DNAs extracted from *S. griseus*, *S. avermitilis* and *S. lividans* grown on solid GYM and in liquid R5A, revealed that DNA cytosine methylation changes during growth of these *Streptomycetes*.

In particular, the global levels of DNA cytosine methylation of all strains were higher at the stage of MI in liquid medium than in MII phase (Figure 68).

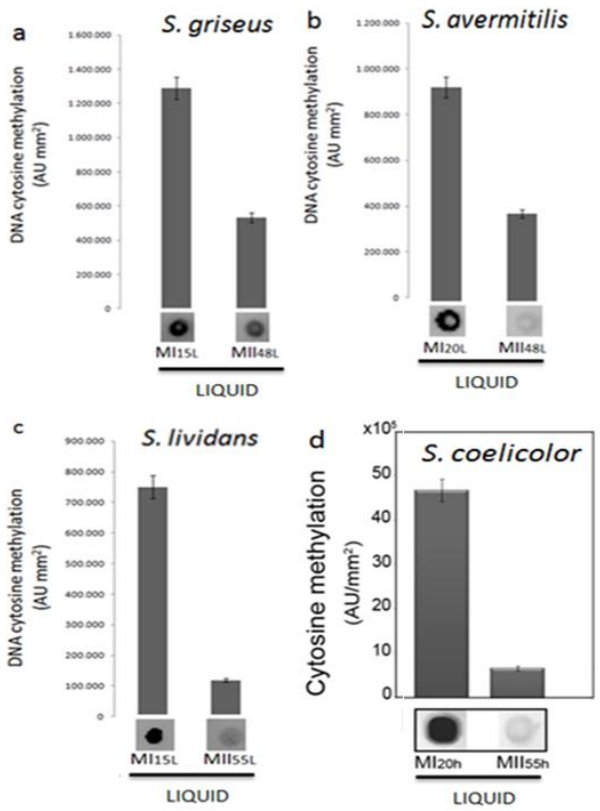

Figure 68 Levels of DNA cytosine methylation of *S. griseus* (a), *S. avermitilis* (b), *S. lividans* (c) and *S. coelicolor* (d) in liquid medium. Dot blot is shown at the bottom of the graphs.

On solid medium, the global levels of DNA cytosine methylation of all strains were higher at the MI phase than MII; only *S. avermitilis* spores showed a high level of methylation (Figure 69).

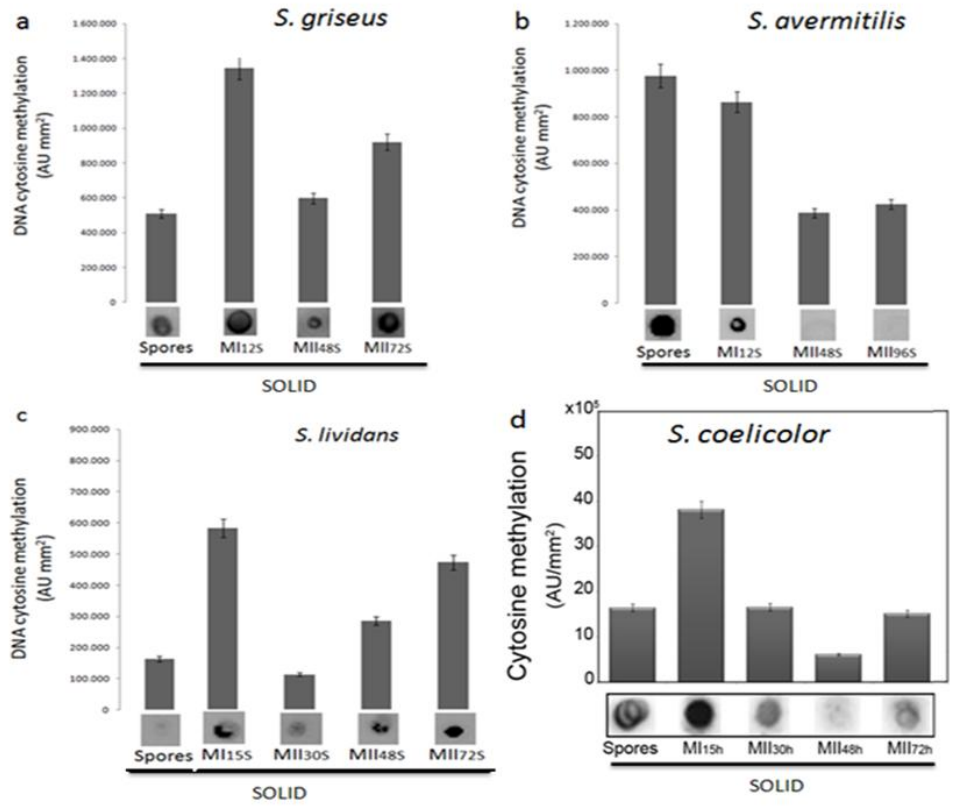

Figure 69 Levels of DNA cytosine methylation of *S. griseus* (a), *S. avermitilis* (b), *S. lividans* (c) and *S. coelicolor* (d) in solid (blue panel) medium. Dot blot is shown at the bottom of the graphs.

These results demonstrate that the cytosine methylation is conserved in the Streptomycetaceae family and that the chromosomal DNA methylation is independent from RM system.

Discussion

This PhD project described for the first time how the DNA cytosine methylation is involved in morphological and physiological differentiation in *S. coelicolor*.

DNA cytosine methylation was studied in three different media: defined liquid medium MG, rich liquid medium R5A and solid rich medium GYM. After preliminary experiments in MG, R5A and GYM media were chosen since the developmental cycle was better defined and transcriptomic analysis results are published and available (31).

The life cycle of *S. coelicolor* is complete on the solid medium GYM: a spore germinates and differentiates in MI mycelium. MI undergoes a highly ordered PCD and the remaining viable segments of these hyphae begin to enlarge in the form of a MII mycelium, that differentiates in substrate, aerial and sporulating (Figure 6). In the liquid medium R5A, the life cycle is almost complete, a spore germinates and differentiates in MI mycelium, a event of PCD occurs to allow the differentiation in MII mycelium that remains undifferentiated. The growth phases on solid GYM and in liquid R5A are very similar, the only difference between these media is the lack of differentiation of MII mycelium in liquid culture.

The growth curve of *S. coelicolor* in liquid MG is characterized by two rapid growth phases, a transition and a decline phase (Figure 10). During the first rapid growth phase (RG1) cells divide quickly. The transition phase (T) is characterized by a growth arrest, transitional arrest of macromolecular biosynthesis, turnover of ribosomal proteins and beginning of secondary metabolites synthesis, like antibiotics. During the second rapid growth phase (RG2) an increased secondary metabolites production occurs. However, along this cycle the cells are in MII phase (Figure 9).

1. DNA cytosine methylation along the growth

DNA cytosine methylation levels were found to change during development in all the media tested. Specifically, DNA cytosine methylation is higher at the MI stage than in the MII or spores in the tested media (liquid R5A, Figure 26 and solid GYM, Figure 35). Regarding the

liquid medium MG, the levels of cytosine methylation were higher at the times preceding the transition and the decline phase (Figure 10).

Interestingly, the phase in which higher levels of cytosine methylation were found corresponds to the phases when programmed cell death (PCD) events occur. In liquid R5A and on solid GYM the first round of PCD occurs at MI phase; in liquid MG cell death events occur at the transition and the decline phase. Some bacteria undergo PCD events (31) in response to cellular stress or damage (*E. coli, C. crescentus, Streptococcus mutans*), or in response to developmental signals (*Bacillus subtilis, S. coelicolor*). Anyway, there is no general biochemical model to explain bacterial PCD. Moreover, it was shown that an induced hypermethylation of the cytosines leads to cell death in *E. coli*, in fact a methyl-specific DNA endonuclease would sense these epigenetic changes and trigger cell death through chromosome cleavage (40). In according with these results, cytosine methylation could have a role in the regulation of PCD event or in the regulation of genes involved in this pathway.

It was found that cytosine methylation is conserved in the streptomycetes family; in fact, *S. lividans*, *S. griseus* and *S. avermitilis* genomic DNAs are methylated along the growth with the highest level of cytosine methylation in MI phase both on solid GYM and in liquid R5A medium (Figure 68 and Figure 69), as in *S. coelicolor*.

2. Cytosine methylomes

In this project the methylomes of *S. coelicolor* along the growth in liquid MG and R5A and on solid GYM were determined (Figure 10, Figure 25 and Figure 35). Firstly, the methylome of MG was studied revealing that only 325 genes (4.15%) of *S. coelicolor* contain methylated cytosines in their upstream regions (-400 at ATG). Secondly, the analysis was expanded using R5A and GYM because in both these media the life cycle of *S. coelicolor* is complete (GYM) and almost complete (R5A).

Methylome of R5A is constituted by 2190 genes (28%) containing methylated cytosines, in the region comprised between -400 and +100 bp of genes. Among these, 58.5% were methylated during the MI phase (20h), 30% during the MII phase (55h) and 11.5% were methylated in both MI and MII phase.

Methylome of GYM contains 2251 genes (28.8%) that were methylated during the development. Among these 63.9% were methylated during the MI phase (15h), 29.2% during the MII substrate phase (30h), 5% during the MII aerial phase (48h), 0.5% during the MII sporulating phase (72h) and 1.4% in the spore.

The genes of the three methylomes were grouped in functional categories (Figure 29 and Figure 38): genes involved in primary and secondary metabolism, genes coding for regulatory proteins, genes correlated to differentiation, genes coding for transporters and secreted proteins, genes involved in catabolism and degradation, genes related to lipid metabolism, genes coding for stress and defense proteins and genes with unknown function.

Interestingly, three most recurring cytosine methylation sequences were found; the GGC^{m}CGG and GCC^{m}CG (Figure 15) consensus sequences are common in all the media, C^{m}GGGC consensus sequence is exclusively present in liquid R5A and on solid GYM (Figure 28), in which both MI and MII phases are present.

In the last years, the methylome of *E. coli* and of other six bacteria was investigated. In *E. coli* the consensus sequence C^{m}CW[7]GG was found to be fully methylated in stationary phase cells and the extended CC^{m}CWGG partially methylated in exponentially growing cells (19), and it is correlated with the regulation of gene expression. In *E.coli* 2.3% of genome contains methylated cytosines.

For the other six bacteria, *Geobacter metallireducens* GS-15, *Chromohalobacter salexigens*, *Vibrio breoganii* 1C-10, *Bacillus cereus* ATCC 10987 and two species of *Campylobacter jejuni subsp*., different consensus sequences were found and they are correlated with the restriction-

[7] W= A or T

modification system (20); about 98% of genome of these bacteria present methylated adenines (m6A) and cytosines (m5C and m4C).

3. Correlation between cytosine methylome and transcriptomic analysis

The cytosine methylomes were compared to the transcriptomic data of the corresponding growth phases (34).

There is a strong relationship between two cytosine methylation consensus sequences and gene expression. In particular, GGC^{m}CGG consensus sequence is methylated in MI phase in genes whose transcription, like *SCO5587 (ftsH), SCO2716 (chpA), SCO5080 (actVA5), SCO5090 (actV-ORFII), SCO5091 (actIV), SCO5092 (actVB), SCO5881 (redZ), SCO5879 (redW), SCO5896 (redH), SCO6681 (ramC)* and *SCO6685 (ramR),* is inhibited in MI, both on solid and in liquid culture. If this consensus sequence is present twice there is a contrary effect, the transcription is stimulated for the genes *SCO6635 (pglY), SCO4423 (afsK)* and *SCO5820 (hrdB)* (Table 3 and Table 6).

Regarding the GCC^{m}CG consensus sequence a correlation was also found, it is methylated in MII phase in genes, i.e. *SCO2077* (*divIVA), SCO5078 (actVA3), SCO1489 (bldD), SCO5085 (actII-ORF4),* and *SCO5113 (bldKB),* whose transcription is activated in MII phase. The same effect was detected in R5A and GYM media (Table 4 and Table 7).

Most of these genes are involved in physiological and morphological differentiation. On the basis of these results, the hypothesis could be that the methylation of these two motifs, GGC^{m}CGG and GCC^{m}CG, regulates transcription of different genes and that at least two different methyl-transferases could be responsible for the progression along the complex life cycle of *S. coelicolor* together with the already known biochemical pathways *bld, sky* and *whi*.

About the presence of C^{m}GGGC consensus sequence, in liquid R5A and on solid GYM, no regulatory effect was detected for the two analyzed genes *SCO2571 (leuS)* and *SCO3911 (dnaB)*.

Methylome of spores showed that only 31 genes, 21 of them with unknown function, contain methylated cytosines in their regulation region. The germination phase has been less studied and the transcriptomic and proteomic analysis (38) showed that most genes related to germination code for hypothetical proteins with unknown function, but a direct correlation between methylation and gene expression was not found.

4. Effect of cytosine demethylation by adding aza-dC

The treatment with 5-aza-2'-deoxycytidine (aza-dC), a cytidine analogous that inhibits DNA-methyltransferase activity, revealed that cytosine methylation affects growth in all the media tested: the treatment delayed the rate of growth in liquid MG (Figure 20) and R5A (Figure 30); and influenced the rate of growth after 63h of growth on solid GYM (Figure 39).

The treatment impairs physiological differentiation. Indeed, in the treated culture (in all media) undecylprodigiosin and actinorhodin production was strongly decreased in respect to the untreated culture in MG, R5A and GYM; only in MG actinorhodin production was weakly decreased in respect to the untreated culture. This result is in contrast with that one obtained by Fernandez *et al*., (33) that demonstrated that 5-azacytidine increased the production of rhodomycin in *S. antibioticus* and actinorhodin in *S. coelicolor*; it is noteworthy that different media and different drug concentrations were used and that the two drugs, 5-azacytidine and 5-aza-2'-deoxycytidine, are differently incorporated with the first being incorporated in both DNA and RNA and the second one only in DNA.

The upstream of the regulatory genes for antibiotic production (*SCO4423, afsK),* undecylprodigiosin *(SCO5897-redG, SCO5898-redF* and *SCO5881-redZ*) and actinorhodin *(SCO5085-actII-ORF4*) biosynthesis are methylated and the genes expressed in MII phase, both on solid and in liquid media. Quantitative transcriptional analysis showed that demethylation of the GCC^{m}CG methylation consensus sequence upstream the gene *SCO5085 (actII-ORF4)* caused a strong decrease in the the transcription level of this gene both in R5A and GYM (Figure 33 and

Figure 45); thus, justifying the decrease of the actinorhodin production due to the aza-dC treatment.

Differently, quantitative transcriptional analysis showed that the gene *SCO5881 (redZ)*, containing the GGC^{m}CGG methylation consensus sequence is more transcribed after demethylation treatment both in R5A and in GYM (Figure 32 and Figure 44). This result does not explain the decrease of the undecylprodigiosin production during the aza-dC treatment, suggesting that probably other proteins control this pathway .

Concerning the morphological differentiation, DNA cytosine demethylation influenced *S. coelicolor* germination, both on solid GYM and in liquid R5A, and sporulation on solid GYM. The same result was obtained in *S. antibioticus* (34).

The genes involved in the differentiation, like *SCO5316 (whiE), SCO2077 (divIVA)* and *SCO1489 (bldD),* are methylated and expressed in MII phase both on solid and in liquid media. qRT-PCR results of the genes *SCO2077 (divIVA)* and *SCO1489 (bldD)*, containing the GCC^{m}CG methylation consensus sequence, in the untreated and treated with aza-dC samples revealed that in the treated samples the transcription level of these genes was lower (11 fold and 1.5 fold, in R5A, and 25 and 2 fold, in GYM) in respect to the untreated sample (Figure 33 and Figure 45). These results explain the block of the morphological differentiation during the aza-dC treatment.

The effect of aza-dC on sporulation occurs both when the drug is added at the beginning of the growth and when added in MII phase (Figure 42) suggesting the involvement of a specific methyltransferase in regulating genes for sporulation.

5. Role of a putative DNA(5-cytosine)-methyltransferase

Among 17 putative DNA(5-cytosine)-methyltransferases 8 are more transcribed at MI phase and 8 at MII phase, only one shows an unchanged expression profile in all the phases, *SCO6844*, it is the only one with a similarity of 39% with DNA(cytosine-5)-methyltransferase 1 (DNMT1), that maintain patterns of methylated cytosine residues in the mammalian genome, and it is considered

the enzyme responsible for most of the maintenance methylation activities occurring in the somatic cells of vertebrates (39). On the basis of this homology, a hypothesis could be that SCO6844 methytransferase is the enzyme for maintenance of methylation levels and patterns of the entire genome in *S. coelicolor*, but further experiments are necessary.

The gene *SCO1731* was more transcribed at MI phase, in which DNA cytosine methylation is higher in respect to MII phases; thus its role was investigated by generating a knockout mutant.

Phenotypic analysis of Δ*SCO1731* mutant strain revealed that the *SCO1731* did not alter the growth kinetics of the mutant in liquid R5A, indicating that this gene is not critical for bacterial growth under the used conditions. Interestingly, the actinorhodin production of Δ*SCO1731* strain occurred later than the wild type, after 130h of growth rather than 72h, (Figure 54). The Δ*SCO1731* strain showed a delayed morphological development (aerial mycelium formation and sporulation) on solid GYM and actinorhodin production both on solid and in liquid R5A and the complemented mutant strain was generated and the wild type phenotype was partially restored (Figure 59 and Figure 60). These results indicate that SCO1731 is involved in the regulation of actinorhodin production and morphological differentiation.

In according with the phenotype of the mutant, putative targets of this methyltransferase could be genes involved in actinorhodin production and aerial mycelium formation and sporulation.

Thus, one hypothesis could be that SCO1731, more transcribed in MI, is responsible for the methylation of genes found methylated in MI. Among them, genes containing the $GGC^{m}CGG$ methylation consensus sequence were found (Table 3 and Table 6), and genes, involved in morphological differentiation, like *SCO5587 (ftsH-like), SCO2716 (chpA), SCO6681 (ramC), SCO6685 (ramR),* and antibiotic production, like *SCO5881 (redZ), SCO5879 (redW), SCO5892 (redL), SCO5896 (redH), SCO5080 (actVA5), SCO5088 (actIORF2), SCO5091 (actIV), SCO5092 (actVB)* could be the putative targets of SCO1731.

BS sequencing and RNA-sequencing will reveal which sequence is methylated and which genes are regulated by SCO1731.

Through the knowledge of cytosine methylomes important progress will be made towards understanding the processes that underlie growth and morphogenesis in *Streptomyces* species.

Material and Methods

1. Strains and media

Stains	Genotype	Origin or reference
Escherichia coli DH10B	F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 Δ(*ara, leu*)7697 *ga*/U *ga*/K λ- *rps*L *nup*G	Woodcock, *et al*. 1989
Escherichia coli ET12567/pUZ8002	F-*dam*-13::Tn9 *dcm*6 *hsd*M *hsd*R *rec*F143 zjj201::Tn10 *gal*K2 *gal*T22 *ara*14 *lac*Y1 *xyl*-5 *leu*B6	Kieser *et al*, 2000
***Streptomyces coelicolor A(3)2* strain M145**	$SCP1^-$, $SCP2^-$	Kieser *et al*, 2000
***Sreptomyces lividans* strain 1326**	$SCP1^-$, $SCP2^-$	Kieser *et al*, 2000
***Streptomyces griseus* strain NBRC 13356**		
***Streptomyces avermitilis* strain ATCC 312**		
ΔSCO1731	*S. coelicolor SCO1731::Tn5062* mutant	This study
wt+pNG3	*S. coelicolor* with pNG3	This study
Δ*SCO1731*_compl	*S. coelicolor SCO1731::Tn5062* mutant with pNG3-SCO1731	This study
1copy_*SCO1731*	*S. coelicolor* with pNG3-SCO1731	This study
ΔSCO0995	*S. coelicolor SCO0995::Tn5062* mutant	This study
Δ*SCO0995*_compl	*S. coelicolor SCO0995::Tn5062* mutant with pNG2-SCO0995	This study

Liquid R5A: (0.25g/l K_2SO_4, 10.12g/l $MgCl_2$ $6H_2O$, 10g/l glucose, 0.1g/l casaminoacids, 2 ml/l trace element solution[8], 5g/l yeast extract, 21g/l MOPS, pH 6,8 with KOH);

liquid MG [(50g/l Maltose, 21g/l MOPS, 0.2g/l $MgSO_4$ $7H_2O$, 0.09g/l $FeSO_4$ $7H_2O$, 0.001g/l $CaCl_2$ $2H_2O$, 0.001g/l NaCl, 40.5g Glutamate, 4.5ml Trace elements, 150ml PO_4 buffer pH6.5 (SolA, 342.5 ml/l,+ SolB, 157.5 ml/l); SolA: 10,88 g KH_2PO_4 (V_f 400ml) SolB: 13,94g KH_2PO_4 (V_f 400ml)];

liquid JM: (100g/l sucrose, 30g/l tryptone, soya broth, 10g/l yeast extract and 10g/l MgCl $6H_2O$) was used for pre-inoculation.

Flasks with liquid R5A and MG media were used for the fermentation of 10^8 viable spores and fresh mycelium (grown in liquid JM), respectively. They were incubated at 30°C and shaken at 200 rpm.

[8] $ZnCl_2$ 40mg/l, $FeCl_3$ $6H_2O$ 200mg/l, $CuCl_2$ $2H_2O$ 10mg/l, $MnCl_2$ $4H_2O$ 10mg/l, $Na_2B_4O_7$ $10H_2O$ 10mg/l, $(NH_4)_6Mo_7O_{24}$ $4\ H_2O$ 10mg/l (39).

Solid cultures were performed on Petri dishes (8.5 cm) with 25 ml of GYM solid medium (5g/l glucose, 4g/l yeast extract, 5g/l malt-extract, 0,5g/l $MgSO_4$ $7H_2O$, 18 g/l agar, after the sterilization add 0,5g/l K_2HPO_4) that were inoculated with $1X10^8$ viable spores, and incubated at 30 °C.

For the sporulation of *S. coelicolor* the solid medium SFM (20g/l mannitol, 20g/l soya flour and 20g/l agar) was used.

2. Plasmids and cosmids

Plasmid	Description	Origin or reference
pCR™-Blunt II-TOPO®	Zero Blunt® TOPO® PCR Cloning Kit, Km^R	Invitrogen
pQM5062	Plasmid containing *eGFP* Tn5062	Bishop *et al*., 2004
pNG3	*bla* cloned into pNG1/*HindIII*/*AvrII* Hyg^R, Amp^R	Gonzalez-Quiñonez *et al.*, 2015 accession KR131850
pNG3-1731compl	pNG3 plasmid containing the ORF of SCO1731	This study
pNG2	*PermE**+RBS+MCS+*fd-ter* cloned into pNG1'/*EcoRV/SpeI* Hyg^R	Gonzalez-*Quiñonez et al.*, 2015; accession KR131849
pNG2-0995compl	pNG2 plasmid containing the ORF of SCO0995	This study
Cosmid		
I11.2.G06	I11 cosmid carrying I11.2.G06 transposant	Fernàndez-Martinez *et al*, 2011 (40)
2StG2.1.G12_2ScG2-1	2StG2 cosmid carrying 2StG2.1.G12_2ScG2-1 trasposant	Fernàndez-Martinez *et al*, 2011

3. Treatment of *S. coelicolor* with aza-dC

For liquid cultures, 5 µM aza-dC was added to liquid medium at the time of inoculation and every 12/24h as indicated in the text. A control experiment was done in parallel with DMSO (it is indicated as untreated in the text).

Solid medium: 5 µM aza-dC was added at the time of inoculation on solid medium and every 12h. A control experiment was done in parallel with DMSO (it is indicated as untreated in the text).

4. Dot Blot assay

Genomic DNA (30 ng/µl) was denatured at 95°C for 10 min, incubated on ice for 5 min, spotted on nitrocellulose filter and fixed by UV (2 cycles at 700 joules). Blocking of non-specific binding was achieved by incubating the membrane in TBS-T (20 mM Tris-HCl pH 8, 150 mM NaCl; pH7.5, 0.05% Tween) containing 1% Bovine serum albumin (Sigma). After 15 min blocking solution was replaced with a solution of TBS-T containing 0.5% BSA and a dilution 1:10000 of the primary antibody anti-m5C (Anti-5-Methylcytosine Mouse mAB, Calbiochem) and incubated ON at RT. After 3 washes for 15 min with TBS-T, the membrane was incubated with a dilution 1:10000 of the secondary antibody anti-mouse (Goat Anti-Mouse IgG-H&L Chain Specific Peroxidase Conjugate, Calbiochem) for 2 hours. After 3 washes for 15 min with TBS-T, Chemidoc (Chemi Hi sensitivity) was used to capture the image with SuperSignal® West Femto Maximum Sensitivity Substrate (Life Technologies). For quantification, spots of the same area were manually labelled. The software measured the number of pixels per each spot and calculated the mean value among three replicates of the same sample. A SCOre for methylation level was assigned to each sample on the basis of the intensity of the signal. Percentage of methylation level was reported as arbitrary units per mm^2 (AU/mm^2). Genomic DNAs, extracted from *Escherichia coli dam*$^+$ *dcm*$^+$ and *Escherichia coli dam*$^-$ *dcm*$^-$ strains, were used as positive and negative control, respectively. The experiments were performed at least twice and in triplicate.

5. Spore germination

Germination was quantified was quantified in solid media with cellophane discs. At different developmental time points, pieces of cellophane discs were cut and processed for Confocal microscopy as described in the previous paragraph. Three biological replicates of the cultures were analysed at different developmental time points. The percentage of germination was assessed from at least 100 spores at each time point. Spores were considered to be germinating when the germ 1 tubes were visible under the confocal microscope.

6. Antibiotic quantification

To measure actinorhodin (intracellular and extracellular), cells were broken in their culture medium by adding KOH 0.1 N. Cellular debris was discarded by centrifugation, and actinorhodin was quantified spectrophotometrically with UV/visible spectrophotometer, applying the linear Beer-Lambert relationship to estimate concentration (ε_{640}=25,320) (41).

Undecylprodigiosin was measured after vacuum drying of the mycelium, followed by extraction with methanol, acidification with HCl (to 0.5 N), and spectrophotometic assay at 530 nm, again using the Beer-Lambert relationship to estimate concentration (ε_{530}=100,500).

Reproducibility has been corroborated by at least three independent cultures at various developmental time points.

7. Protein quantification (for growth curve)

Samples (half milliliter) were collected at different developmental time points, and stored at -20ºC until they were analyzed. Half milliliter of NaOH 1M was added to the half milliliter samples (0.5M final concentration of NaOH), and boiled for 10 min. Cellular debris was removed by centrifugation (at 7740 ×g for 15 min at 4 ºC), in order to obtain the intracellular samples. Protein was quantified using the Bradford method (Bradford, 1976) with bovine serum albumin (Sigma) as the standard.

8. Confocal laser scanning microscopy analysis (CLSM)

Culture samples were obtained and processed for microscopy at different incubation time points. Cells were stained with a non-cell-permeating nucleic acid stain (propidium iodide, PI) in order to detect the dead cell population and with SYTO 9 green fluorescent nucleic acid stain (LIVE/DEAD Bac-Light bacterial viability kit, L-13152; Invitrogen) to detect viable cells. The SYTO 9 green fluorescent stain labels all the cells, i.e., those with intact membranes, as well as those with damaged ones. In contrast, PI penetrates only in bacteria with damaged membranes, decreasing SYTO 9 stain fluorescence when both dyes are present. The samples were observed under a Leica TCS-SP2-AOBS confocal laser-scanning microscope at a wavelength of 488 nm and 568 nm excitation and 530 nm (green) or 640 nm (red) emissions. A significant number of images was analyzed in a minimum of three independent culture analyses.

9. Bisulfite sequencing

Genomic DNA was extracted by salting out procedure (42) and purified by GenElute Bacterial Genomic DNA kit (Sigma). Quantity and quality of DNA samples were measured by NanoDrop ND1000 Spectophotometer and dried samples were sent to BGI Honk Hong CO.

DNA sequencing was carried out by Illumina's HiSeq Technology. The obtained genomic sequences were compared to the annotated *S. coelicolor* M145 genome sequence (link http://www.ncbi.nlm.nih.gov/nuccore/NC_003888.3).

10. Analysis on methylation counts

Barcoded reads obtained from Illumina Genome Analyzer IIx were first assigned to their respective samples. Subsequently, methylation calls were made using Bismark software36, which makes use of Bowtie57 for read-to-genome mapping.

Methylated bases were extracted from the count files and mapped to the *S. coelicolor A3(2)* reference genome. Methylated bases were also mapped to gene's upstream sequences if they

were found in the -400 +100 region around the first translated codon of each gene. Operon information was derived from the MicrobesOnline operon prediction database. The following software libraries have been used in this analsys: *ipython 1.0.0* (43), *Biopython 1.62* (44), *Matplotlib 1.2.1* (45), *NumPy 1.7.1* (46) and SciPy 0.12.0. The MEME software (version 4.10.0) was used to search for consensus sequences around the methylated sites, using 337 sequences and the following parameters: consensus sequence width 21, "zoops" search strategy, consensus sequence "NNNNNNNNNNCNNNNNNNNNN", and minimum number of sites 30. All statistical tests were carried out using R (http://r-project.org).

11. Real-Time Quantitative Reverse Transcription PCR (qRT-PCR)

RNA was extracted after 18h and 24h of growth in liquid medium MG. The cells were broken by using 1 mg of lysozyme/ml in P-buffer (42), and total RNA was isolated by using the RNeasy mini-kit (QIAGEN). DNase I (Roche) treatment was performed at 37°C for 1 h, and ethanol precipitation in the presence of 0.1 V 3 M sodium acetate allowed recovery of the DNase-treated total RNA. After a washing step with 70% ethanol and air drying, the RNA pellet was resuspended in water. Reverse transcription-PCR (RT-PCR) was performed by using a Superscript One-Step RT-PCR kit (Invitrogen) with about 0.2 µg of total RNA as a template, primer pairs internal to genes indicated in Table 10 and the conditions indicated by the supplier, routinely using 40 PCR cycles. For each reaction, a negative control with Taq polymerase and without reverse transcriptase was included.

Expression was analyzed quantitatively by real-time RT-PCR using the Applied Biosystems 7300 real-time PCR system (Applied Biosystems). A high-capacity cDNA archive kit (Applied Biosystems) was used, according to the manufacturer's instructions, to retrotranscribe 5 µg of total RNA, extracted after 18h and 24h of growth, in liquid medium MG, from untreated and treated with aza-dC cultures, in a final volume of 100 µl of water. Then, 3 µl of the cDNA was mixed with 10 µl SYBR green PCR master mix (Applied Biosystem) and 10 pmol of each

primer in a final volume of 20 µl. The PCR was performed under the following conditions: 2 min at 50°C and 10 min at 95°C, followed by 40 cycles of 15 s at 95°C and m1 min at 68°C. Eventually, a dissociation reaction was performed with the following conditions: a 1-min step with a temperature gradient increase of 1°C per step from 55 to 99°C. This last reaction allowed the melting curve of the PCR products and, consequently, their specificity to be determined. A negative control (distilled water) was included in all real-time PCR assays, and each experiment was performed in triplicate. The 16S gene (the vegetative sigma factor encoding gene) was used as an internal control to quantify the relative expression of target genes.

12. Distruption of *SCO1731* and *SCO0995*

Cosmids I11.2.G06 and 2StG2.1.G12 (40) were used for constructing the *SCO1731::Tn5062* mutant (*ΔSCO1731*) and *SCO0995::Tn5062* mutant (*ΔSCO0995)* strains. Gene disruption was made by obtaining double cross-over via conjugation using *E. coli* ET12567/pUZ8002 as a donor strain, and following the protocol described in Kieser *et al*. (2000) (42). *ET12567* [pUZ8002] is methylation-deficient and therefore a convenient donor to use it for conjugation with *S. coelicolor*. The plasmid pUZ8002 is required because of the *tra* gene, which encodes the transfer protein Tra. *E. coli ET12567/pUZ8002* was grown at 37ºC in LB, (Tryptone 10g/l, Yeast Extract 5g/l, NaCl 10g/l), containing kanamycin (25 µg/ml) and chloramphenicol (25 µg/ml).
Mutant strains were confirmed by PCR analysis and southern blotting, using chromosomal DNA digested with *SalI.* Southern hybridisation was carried out by established procedures using the digoxigenin-labelled 3442-bp Tn5062 *PvuII* fragment from plasmid pQM5062 as a probe. No hybridization signal was detected for the parental strain.

13. Complementation and overexpression of *SCO1731*

To construct the plasmid to complement *ΔSCO1731* and to insert one copy of *SCO1731* in *S. coelicolor* (1copy_*SCO1731* strain), a PCR fragment prepared using the primers SCO1731-

speI_F and SCO1731compl_R (Table 10) was cloned into pCR™-Blunt II-TOPO® (Invitrogen), digested with *SpeI* (Invitrogen) and *EcoRV* (Invitrogen), and then inserted between the *SpeI* (Invitrogen) and *EcoRV* (Invitrogen) sites of pNG3. Cloning was confirmed by sequencing the insert. The plasmid was transferred into *S. coelicolor ΔSCO1731* strain and wild type by conjugation from *E. coli* ET12567 (pUZ8002) as described above, generating strains called *ΔSCO1731*_compl and 1copy_*SCO1731*, respectively. As a control, pNG3 was inserted into the *S. coelicolor* chromosome. PCR analysis using the primers SCO4848_F/R was performed to confirm the presence of the integrative vector into the chromosome. These primers were designed to hybridize at the gene *SCO4848* 5' end (SCO4848_F), in *S. coelicolor* chromosome, and the DNA region flanking *attP* in pNG3 (Figure 70) (SCO4848_R) (Gonzalez-Quiñonez *et al.*, 2015). These primers amplify a fragment of 617 bp only if pNG3 is integrated at the *attB* site of *SCO4848*.

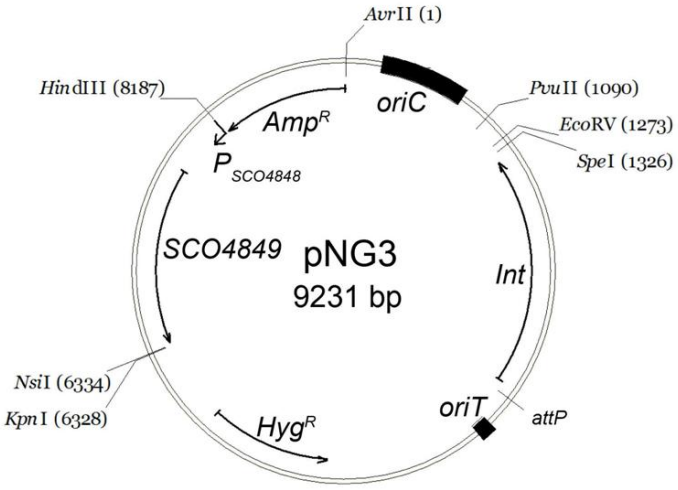

Figure 70 pNG3 vector map.

14. Complementation and overexpression of *SCO0995*

To construct the plasmid to complemen *ΔSCO0995* and to overexpress *SCO0995* in *S. coelicolor* a PCR fragment prepared using the primers over0995_F and over0995_R (Table 10) was cloned

into pCR™-Blunt II-TOPO® (Invitrogen), digested with *SpeI* (Invitrogen) and *NdeI* (Invitrogen), and then inserted between the *SpeI* (Invitrogen) and *NdeI* (Invitrogen) sites of pNG2. Cloning was confirmed by sequencing the insert. The plasmid was transferred into *S. coelicolor ΔSCO0995* strain and wild type by conjugation from *E. coli* ET12567 (pUZ8002) as described above, generating strains called *ΔSCO0995*_compl and *SCO0995*_over. As a control, pNG2 was inserted into the *S. coelicolor* chromosome. PCR analysis using the primers SCO4848_F/R was performed to confirm the presence of the integrative vector into the chromosome. These primers were designed to hybridize at the gene *SCO4848* 5' end (SCO4848_F), in *S. coelicolor* chromosome, and the DNA region flanking *attP* in pNG2 (Figure 71) (SCO4848_R) (Gonzalez-Quiñonez *et al.*, 2015). These primers amplify a fragment of 617 bp only if pNG2 is integrated at the *attB* site of *SCO4848*.

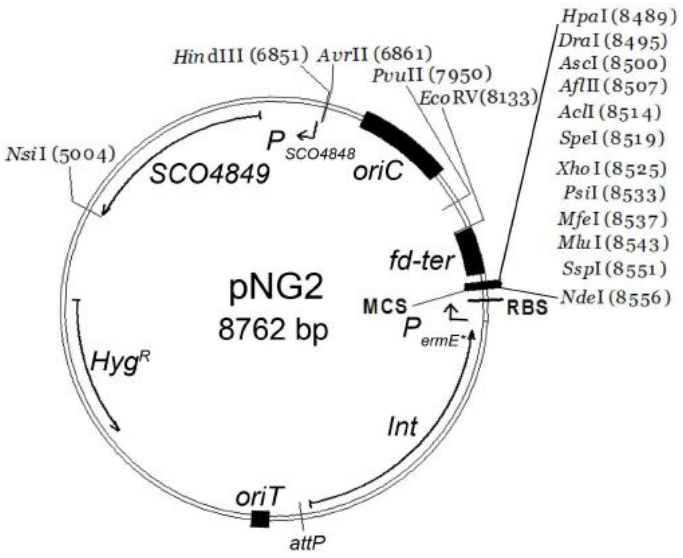

Figure 71 pNG2 vector map.

15. PCR analysis of mutants *ΔSCO1731* and *ΔSCO0995*

250 ng of genomic DNA was used in a PCR with Kana_F/R (amplicon size: 902 bp) and Apra_F/R (amplificon size: 1372 bp) primers to detected the presence of kanamycin and apramycin resistance cassettes, respectively.

	Master mix (Final conc)	Amplification program
Buffer 10X	1X	1: 95°C 5 min
Primer F 10 µM	0.2 µM	2: 95°C 30 sec
Primer R 10 µM	0.2 µM	3: 60°C 30 sec
$MgCl_2$ 50 mM	33 mM	4: 72°C 30 sec
dNTP 2 mM	0.3 mM	Repeat 2-3-4 for 35 cycles
Taq DNA Polymerase 5U/µl (INVITROGEN)	0.5 µl	72°C 10 min
H_2O	up to 20 µl	
Template	2 µl	

References

1. Ishikawa K, Fukuda E, Kobayashi I. Conflicts targeting epigenetic systems and their resolution by cell death: novel concepts for methyl-specific and other restriction systems. DNA Res. 2010 Dec e 7(6):325-42.

2. Collier J., Epigenetic regulation of the bacterial cell cycle. Curr Opin Microbiol. 2009 Dec e 12(6):722-9.

3. Iyer LM, Abhiman S, Aravind L., Natural history of eukaryotic DNA methylation systems. Prog Mol Biol Transl Sci. 2011 e 101:25-104.

4. Tazi, J. and Bird, A. Alternative chromatin structure at CpG islands. Cell 60, 909-20. (1990).

5. Kass, S. U., Landsberger, N., and Wolffe, A. P. DNA methylation directs a time-dependent repression of transcription initiation. Curr Biol (1997) 7, 157-65.

6. He XJ, Chen T, Zhu JK. Regulation and function of DNA methylation in plants and animals.Cell Res. 2011 Mar e Review., 21(3):442-65.

7. Zemach A., McDaniel I. E., Silva P., Zilberman D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. 14 MAY 2010 VOL 328 SCIENCE.

8. Capuano F, Mülleder M, Kok R, Blom HJ, Ralser M. Cytosine DNA methylation is found in Drosophila melanogaster but absent in Saccharomyces cerevisiae, Schizosaccharomyces pombe, and other yeast species. Anal Chem. 2014 Apr 15.

9. Hattman S. DNA-[adenine] methylation in lower eukaryotes. Biochemistry. May, 2005 e 70(5):550-8.

10. Baniushin BF., Methylation of adenine residues in DNA of eukaryotes. Mol Biol (Mosk). 2005 Jul-Aug e 39(4):557-66.

11. Roberts, R.J., Vincze,T., Posfai,J. and Macelis,D. REBASE—a database for DNA restriction and modification:enzymes, genes and genomes. Nucleic Acids Res.2010, 3, D234-D236.

12. Casadesús J, Low D., Epigenetic gene regulation in the bacterial world. Microbiol Mol Biol Rev. 2006 Sep e 70(3):830-56.

13. Oshima T, Wada C, Kaawagoe Y, Ara T, Maeda M, Masuda Y, Hiraga S, Mori H. Genome-wide analysis of deoxyadenosine methyltransferase-mediated control of gene expression in Escherichia coli. Mol Microbiol 2002, 100:4672-4677.

14. Srikhanta YN, Maguire TL, Stacey KJ, Grimmond SM, Jennings MP. The phase varion: a genetic system controlling coordinated, random switching of expression of multiple genes. Proc Natl Acad Sci U S A. 2005 Apr 12 e 31., 102(15):5547-51. Epub 2005 Mar.

15. Kozdon JB, Melfi MD, Luong K, Clark TA, Boitano M, Wang S, Zhou B, Gonzalez D, C.Global methylation state at base-pair resolution of the Caulobacter genome throughout the cell cycle.Proc Natl Acad Sci U S A. 2013 Nov 26 e 11., 110(48):E4658-67. Epub 2013 Nov.

16. Curtis PD, Brun YV. Getting in the loop: regulation of development in Caulobacter crescentus. Mol Cell. 2010 Aug 13 e molcel.2010.07.034., 39(3):319-20.

17. Kumar R, Mukhopadhyay AK, Ghosh P, Rao DN.Comparative transcriptomics of H. pylori strains AM5, SS1 and their hpyAVIBM deletion mutants: possible roles of cytosine methylation. PLoS One. 2012 e 3, 7(8):e42303. doi: 10.1371/journal.pone.0042303. Epub 2012 Aug.

18. Kahramanoglou C, Prieto AI, Khedkar S, Haase B, Gupta A, Benes V, Fraser GM, Luscombe NM, Seshasayee AS. Genomics of DNA cytosine methylation in Escherichia coli reveals its role in stationary phase transcription. Nat Commun. 2012 Jun 6 e 3:886.

19. Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC,Clark TA, Korlach J, Turner SW: Direct detection of DNA methylation during single-molecule, real-time sequencing. Nat Methods 2010, 7:461-465.

20. Murray IA, Clark TA, Morgan RD, Boitano M, Anton BP, Luong K, Fomenkov A, Turner SW, Korlach J, Roberts RJ. The methylomes of six bacteria. Nucleic Acids Res. 2012 Dec e 2., 40(22):11450-62. Epub 2012 Oct.

21. Challis GL, Hopwood DA, Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by Streptomyces species. Proc Natl Acad Sci U S A. 2003 Nov 25 e 2:14555-61., 100 Suppl.

22. Bentley SD. Chater KF, Cerdeño-Tárraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, Bateman A, Brown S,Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). Nature. 2002 May 9 e 417(6885):141-7.

23. Puglia AM, Vohradsky J, Thompson CJ., Developmental control of the heat-shock stress regulon in Streptomyces coelicolor. Mol Microbiol. 1995 Aug e 737-46., 17(4):.

24. Flärdh K, Buttner MJ. Streptomyces morphogenetics: dissecting differentiation in a filamentous bacterium. Nat Rev Microbiol. 2009 Jan e 7(1):36-49.

25. Manteca A, Alvarez R, Salazar N, Yagüe P, Sanchez J. Mycelium differentiation and antibiotic production in submerged cultures of Streptomyces coelicolor. Appl Environ Microbiol. 2008 Jun e 25., 74(12):3877-86. Epub 2008 Apr.

26. Hanna Engelberg-Kulka, Shahar Amitai, Ilana Kolodkin-Gal, Romem Hazan. Bacterial programmed cell death and multicellular behaviour in bacteria. PLoS Genet. Oct 2006 e e135., 2(10):.

27. Julia Bos, Anastasiya A. Yakhnina, Zemer Gitai. BapE DNA endonuclease induces an apoptotic-like response to DNA damage in Caulobacter. Proc Natl Acad Sci U S A. 2012 October 30 e 18096–18101., 109(44):.

28. Sébastien Guiral, Tim J. Mitchell, Bernard Martin, Jean-Pierre Claverys.Competence-programmed predation of noncompetent cells in the human pathogen Streptococcus pneumoniae: Genetic requirements. Proc Natl Acad Sci U S A. 2005 June 14 e 8710–8715., 102(24):.

29. Hardisson C, Manzanal MB, Salas JA, Sua´rez JE. Fine structure, physiology and biochemistry of arthrospore germination in Streptomyces antibioticus. J Gen Microbiol (1978) 105: 203-214.

30. Manteca A, Sanchez J, Jung HR, Schwämmle V, Jensen ON. Quantitative proteomics analysis of Streptomyces coelicolor development demonstrates that onset of secondary metabolism coincides with hypha differentiation. Mol Cell Proteomics. 2010 Jul e 9(7):1423-36.

31. Yagüe P, et. al. Transcriptomic analysis of Streptomyces coelicolor differentiation in solid sporulating cultures: first compartmentalized and second multinucleated mycelia have different and distinctive transcriptomes. PLoS One. 2013 e 10., 8(3):e60665. doi:.

32. Jayapal KP, Lian W, Glod F, Sherman DH, Hu WS. Comparative genomic hybridizations reveal absence of large Streptomyces coelicolor genomic islands in Streptomyces lividans.BMC Genomics. 2007 Jul 10 e 8:229.

33. Fernandez M, Olek A, Walter J, Sanchez J. Analysis of DNA methylation processes related to the inhibition of DNA synthesis by 5-azacytidine in Streptomyces antibioticus ETH 7451. Biol Chem. 1998 Apr-May e 379(4-5):559-62.

34. Novella IS, Sánchez J. Effects of 5-azacytidine on physiological differentiation of Streptomyces antibioticus. Res Microbiol. 1995 Nov-Dec e 146(9):721-8.

35. Fernandez M1, Soliveri J, Novella IS, Yebra MJ, Barbés C, Sánchez J. Effect of 5-azacytidine and sinefungin on Streptomyces development. Gene. 1995 May 19 e 157(1-2):221-3.

36. Timothy L. Bailey and Charles Elkan, Fitting a mixture model by expectation maximization to discover motifs in biopolymers, Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology, pp. 28-36, AAAI Press, Menlo Park,.

37. Rogstad D.K., Herring J.L., Theruvathu J. A., Budzy A., Perry C.C., Neidigh J. W. and Sowers L.C. The chemical decomposition of 5'-aza-2'-deocytidine (Decitabine): kinetic analyses and identification of products by NMR, HPLC, and mass spectrometry. 2009.

38. Gonzalez-Quiñonez N., López-García MT., Yagüe P., Rioseras B., Pisciotta A, Alduina R., Manteca A. New ϕBT1 site-specific integrative vectors with neutral p 1 henotype in Streptomyces. Accepted in Applied Microbiology and Biotechnology, 2015.

39. Strakova E, Bobek J, Zikova A, Vohradsky J. Global features of gene expression on the proteome and transcriptome levels in S. coelicolor during germination. PLoS One. 2013 Sep 9 e 2013., 8(9):e72842. doi: 10.1371/journal.pone.0072842. eCollection.

40. Yen RW, Vertino PM, Nelkin BD, Yu JJ, el-Deiry W, Cumaraswamy A, Lennon GG, Trask BJ, Celano P, Baylin SB. Isolation and characterization of the cDNA encoding human DNA methyltransferase. Nucleic Acids Res June 1992 20 (9): 2287–91.

41. Fernández-Martínez LT, Del Sol R, Evans MC, Fielding S, Herron PR, Chandra G, Dyson PJ. A transposon insertion single-gene knockout library and new ordered cosmid library for the model organism Streptomyces coelicolor A3(2). Antonie Van Leeuwenhoek. 2011.

42. Rioseras B, MT López-García, P Yagüe, J Sánchez, Á Manteca. Mycelium differentiation and development of Streptomyces coelicolor in lab-scale bioreactors: programmed cell death, differentiation and lysis are closely linked to red and act production. 2014.

43. Kieser T, Bibb MJ, Buttner M, Chater K, Hopwood D. Pratical Streptomyces genetics" Norwich, The John Innes Foundation. 2000.

44. http://dx.doi.org/10.1109/MCSE.2007.53. [Online]

45. doi: 10.1093/bioinformatics/btp163. [Online]

46. http://doi.ieeecomputersociety.org/10.1109/MCSE.2007.55. [Online]

47. http://dx.doi.org/10.1109/MCSE.2011.37. [Online]

Supplemental Material

Table 10 List of primers.

Name	Sequence
SCO2571_F	TCAACTCCTGGTACGACGAC
SCO2571_R	CTTGAAGACGGGGAAGTTGC
SCO6164_F	CGCGTCACCCAGCTTCAG
SCO6164_R	CGCGTGTAGGGGAGGATC
ramR_F	GAACGCTTCCTGGACGAGA
ramR_R	CATGTAGTTGCGGACCGTC
hrdB_F	CGTCGAGGGTCTTCGGCTG
hrdB_R	CGCGAGCCCATCTCGCTGC
SCO3911_F	CTCTACATCGACGACTCCCC
SCO3911_R	CGCGGTGACTCCTTCTCATA
redZ_F	AAACGTCGGTCGAAGAACTG
redZ_R	GTCTTGCCCTGGGTCAGTAA
bldD_F	GACTTCCTCGACTCCACAGG
bldD_R	GAGGTGTGAATGGGGAACAC
SCO5085_F	TAATTTCGCATCCGCTGAAC
SCO5085_R	CTACACGAGCACCTTCTCACC
divIVA_F	GTTCCGAGGCCAACAAGATC
divIVA_R	ACTCCAGGTACGACTTCAGC
16Sstrepto_F	TGCCAGCAGCCGCGGTAATA
16Sstrepto_R	GACTGCAGACCCGGGGTTAA
Kana_F	GATGGCTTTCTTGCCGCC
Kana_R	TCGGTCATTTCGAACCCC
Apra_F	CGGGGTACCCTCACGGTAACTGATGCC
Apra_R	ATTTTAATGCGGATGTTGCG
SCO4848_F	CGTCGATCCCCTCGGTTG
SCO4848_R	GAGCCGGGAAAGCTCATTCA
SCO1731-speI_F	GGACTAGTTGGCTGCCTCCTTACGGAT
SCO1731compl_R	AAGATATCGTCTGGACGAGGACGAGTTC
SCO1731over_F	AAGATATCGTCTGGACGAGGACGAGTTC
over0995_F	CATATGGCACGAACGGGCGGGTT
over0995_R	ACTAGTATGACCCGCTGATACTACGG
0995-96_F	CCGCCGAAGGTGGTCCAGC
0995-96_R	GGCTTGTCGTACGTCGTCTT

Table 11 List of genes containing a methylation motif grouped on the basis of their function: primary metabolism, DNA/RNA metabolism, regulators, membrane proteins and hypothetical proteins. CG, CHG and CHH represent the methylation sequences indicated at 18h and 24h of growth. 0 means: no methylation, 1, 2 and 3 indicate 1, 2 and 3 methylated sites.

		18h			24h		
Gene	Function	CG	CHG	CHH	CG	CHG	CHH
SCO0138	Primary metabolism	0	1	0	0	1	0
SCO0191	Primary metabolism	0	1	0	0	1	0
SCO0216	Primary metabolism	0	1	0	0	1	0
SCO0240	Primary metabolism	0	1	0	0	1	0
SCO0449	Primary metabolism	0	0	0	0	1	0
SCO0588	Primary metabolism	0	1	0	0	1	0
SCO0638	Primary metabolism	1	0	0	1	0	0
SCO0812	Primary metabolism	0	0	1	0	0	0
SCO0815	Primary metabolism	0	1	0	0	1	0
SCO0829	Primary metabolism	0	1	0	0	1	0
SCO0917	Primary metabolism	1	0	0	1	0	0
SCO1055	Primary metabolism	0	0	1	0	0	0
SCO1056	Primary metabolism	0	0	1	0	0	0
SCO1084	Primary metabolism	0	1	0	0	1	0
SCO1085	Primary metabolism	0	1	0	0	1	0
SCO1169	Primary metabolism	0	1	0	0	1	0
SCO1170	Primary metabolism	0	1	0	0	1	0
SCO1228	Primary metabolism	0	0	0	0	1	0
SCO1349	Primary metabolism	1	0	0	1	0	0
SCO1350	Primary metabolism	1	0	0	1	0	0
SCO1371	Primary metabolism	0	1	0	0	1	0
SCO1652	Primary metabolism	1	0	0	0	0	0
SCO1847	Primary metabolism	0	1	0	0	1	0
SCO1875	Primary metabolism	0	0	0	0	1	0
SCO1934	Primary metabolism	1	0	0	0	0	0
SCO1935	Primary metabolism	1	0	0	0	0	0
SCO1955	Primary metabolism	1	0	0	0	0	0
SCO2068	Primary metabolism	1	0	0	1	0	0
SCO2093	Primary metabolism	1	0	0	1	0	0
SCO2113	Primary metabolism	0	1	0	0	1	0
SCO2183	Primary metabolism	0	1	0	0	1	0
SCO2187	Primary metabolism	1	1	0	1	1	0
SCO2231	Primary metabolism	0	1	0	0	1	0
SCO2256	Primary metabolism	0	0	1	0	0	0
SCO2267	Primary metabolism	0	1	0	0	1	0
SCO2427	Primary metabolism	1	0	0	1	0	0
SCO2428	Primary metabolism	1	0	0	1	0	0
SCO2553	Primary metabolism	0	1	0	0	1	0
SCO2587	Primary metabolism	3	0	0	1	0	0
SCO2640	Primary metabolism	0	1	0	0	1	0
SCO2710	Primary metabolism	0	0	0	1	0	0
SCO2914	Primary metabolism	1	0	0	1	0	0
SCO3026	Primary metabolism	1	0	0	1	0	0
SCO3314	Primary metabolism	0	1	0	0	1	0
SCO3336	Primary metabolism	0	0	1	0	0	0
SCO3474	Primary metabolism	0	1	0	0	1	0
SCO3540	Primary metabolism	0	1	0	0	1	0
SCO3801	Primary metabolism	0	1	0	0	1	0

SCO3856	Primary metabolism	0	0	1	0	0	0
SCO4006	Primary metabolism	1	0	0	0	0	0
SCO4068	Primary metabolism	0	1	0	0	1	0
SCO4085	Primary metabolism	1	0	0	1	0	0
SCO4109	Primary metabolism	0	0	0	1	0	0
SCO4280	Primary metabolism	0	1	0	0	1	0
SCO4292	Primary metabolism	0	0	0	1	0	0
SCO4293	Primary metabolism	0	0	0	1	0	0
SCO4444	Primary metabolism	1	0	0	1	0	0
SCO4475	Primary metabolism	1	0	0	1	0	0
SCO4651	Primary metabolism	0	1	0	0	1	0
SCO4746	Primary metabolism	1	0	0	1	0	0
SCO4773	Primary metabolism	0	1	0	0	1	0
SCO4774	Primary metabolism	0	1	0	0	1	0
SCO4780	Primary metabolism	0	1	0	0	1	0
SCO4837	Primary metabolism	0	1	0	0	1	0
SCO5251	Primary metabolism	1	0	0	1	0	0
SCO5286	Primary metabolism	0	0	0	1	0	0
SCO5316	Primary metabolism	2	0	0	2	0	0
SCO5365	Primary metabolism	0	1	0	0	1	0
SCO5419	Primary metabolism	0	1	0	0	0	0
SCO5696	Primary metabolism	0	2	0	0	2	0
SCO5744	Primary metabolism	1	0	0	1	0	0
SCO5897	Primary metabolism	1	0	0	1	0	0
SCO6163	Primary metabolism	0	0	0	0	1	0
SCO6178	Primary metabolism	0	1	0	0	1	0
SCO6179	Primary metabolism	0	1	0	0	1	0
SCO6234	Primary metabolism	0	1	0	0	1	0
SCO6347	Primary metabolism	1	0	0	1	0	0
SCO6411	Primary metabolism	0	1	0	0	1	0
SCO6473	Primary metabolism	0	0	1	0	0	0
SCO6522	Primary metabolism	1	0	0	1	0	0
SCO6712	Primary metabolism	1	0	0	1	0	0
SCO6723	Primary metabolism	0	1	0	0	1	0
SCO6799	Primary metabolism	0	1	0	0	1	0
SCO6989	Primary metabolism	1	0	0	0	0	0
SCO7225	Primary metabolism	0	0	1	0	0	0
SCO7225	Primary metabolism	0	0	1	0	0	0
SCO7444	Primary metabolism	1	0	0	1	0	0
SCO7470	Primary metabolism	1	0	0	1	0	0
SCO7471	Primary metabolism	1	0	0	1	0	0
SCO7556	Primary metabolism	1	0	0	1	0	0
SCO7606	Primary metabolism	1	0	0	1	0	0
SCO7607	Primary metabolism	1	0	0	0	0	0
SCO7608	Primary metabolism	1	0	0	0	0	0
SCO7637	Primary metabolism	1	0	0	1	0	0
SCO7659	Primary metabolism	1	0	0	1	0	0
SCO7714	Primary metabolism	0	1	0	0	1	0
SCO7784	Primary metabolism	0	1	0	0	1	0
SCO0178	DNA/RNA metabolism	0	1	0	0	1	0
SCO0325	DNA/RNA metabolism	0	0	0	0	1	0
SCO0895	DNA/RNA metabolism	0	1	0	0	1	0
SCO1301	DNA/RNA metabolism	0	1	0	0	1	0
SCO1421	DNA/RNA metabolism	0	1	0	0	1	0
SCO1973	DNA/RNA metabolism	0	0	1	0	0	0
SCO2092	DNA/RNA metabolism	1	0	0	1	0	0

SCO2170	DNA/RNA metabolism	0	1	0	0	1	0
SCO2571	DNA/RNA metabolism	0	1	0	0	1	0
SCO2639	DNA/RNA metabolism	0	1	0	0	1	0
SCO2950	DNA/RNA metabolism	0	1	0	0	1	0
SCO3109	DNA/RNA metabolism	1	0	0	1	0	0
SCO3149	DNA/RNA metabolism	1	0	0	0	0	0
SCO3365	DNA/RNA metabolism	1	0	0	1	0	0
SCO3543	DNA/RNA metabolism	0	0	0	0	1	0
SCO3702	DNA/RNA metabolism	0	1	0	0	1	0
SCO3732	DNA/RNA metabolism	1	0	0	1	0	0
SCO3798	DNA/RNA metabolism	0	1	0	0	1	0
SCO3896	DNA/RNA metabolism	0	1	0	0	1	0
SCO3911	DNA/RNA metabolism	0	1	0	0	1	0
SCO3974	DNA/RNA metabolism	1	0	0	0	0	0
SCO4027	DNA/RNA metabolism	0	1	0	0	1	0
SCO4034	DNA/RNA metabolism	0	0	0	0	0	1
SCO4096	DNA/RNA metabolism	1	0	0	1	0	0
SCO4146	DNA/RNA metabolism	1	0	0	0	0	0
SCO4176	DNA/RNA metabolism	0	1	0	0	1	0
SCO4236	DNA/RNA metabolism	0	0	1	0	0	0
SCO4441	DNA/RNA metabolism	0	1	0	0	1	0
SCO4662	DNA/RNA metabolism	0	1	0	0	1	0
SCO4960	DNA/RNA metabolism	0	3	0	0	3	0
SCO5802	DNA/RNA metabolism	1	0	0	0	0	0
SCO5820	DNA/RNA metabolism	2	0	0	2	0	0
SCO6164	DNA/RNA metabolism	0	0	0	0	1	0
SCO6380	DNA/RNA metabolism	1	0	0	1	0	0
SCO6606	DNA/RNA metabolism	0	0	0	1	0	0
SCO7213	DNA/RNA metabolism	0	1	0	0	1	0
SCO7445	DNA/RNA metabolism	1	0	0	1	0	0
SCO7740	DNA/RNA metabolism	1	0	0	1	0	0
SCO7773	DNA/RNA metabolism	1	0	0	1	0	0
SCO0322	membran protein/protein	1	0	0	0	0	0
SCO0368	membran protein/protein	1	0	0	1	0	0
SCO0375	membran protein/protein	0	1	0	0	1	0
SCO0752	membran protein/protein	0	1	0	0	1	0
SCO0811	membran protein/protein	0	0	1	0	0	0
SCO1009	membran protein/protein	1	0	0	1	0	0
SCO1010	membran protein/protein	2	0	0	2	0	0
SCO1067	membran protein/protein	0	1	0	1	1	0
SCO1150	membran protein/protein	0	0	1	0	0	0
SCO1516	membran protein/protein	0	1	0	0	1	0
SCO2063	membran protein/protein	0	1	0	0	1	0
SCO2188	membran protein/protein	1	1	0	1	1	0
SCO2328	membran protein/protein	0	0	0	1	0	0
SCO2372	membran protein/protein	1	0	0	1	0	0
SCO2920	membran protein/protein	0	1	0	0	1	0
SCO3024	membran protein/protein	1	0	0	1	0	0
SCO3404	membran protein/protein	0	1	0	0	1	0
SCO3490	membran protein/protein	0	1	0	0	1	0
SCO3564	membran protein/protein	1	0	0	1	0	0
SCO3577	membran protein/protein	1	0	0	0	0	0
SCO4053	membran protein/protein	1	0	0	1	0	0
SCO4157	membran protein/protein	0	1	0	0	1	0
SCO4561	membran protein/protein	0	1	0	0	1	0
SCO4579	membran protein/protein	0	0	1	0	0	0

SCO5157	membran protein/protein	1	0	0	1	0	0
SCO5289	membran protein/protein	1	0	0	1	0	0
SCO5447	membran protein/protein	1	0	0	1	0	0
SCO5629	membran protein/protein	0	1	0	0	1	0
SCO5716	membran protein/protein	0	1	0	0	1	0
SCO5767	membran protein/protein	0	0	0	1	0	0
SCO6257	membran protein/protein	0	0	1	0	0	0
SCO6390	membran protein/protein	0	1	0	0	1	0
SCO6607	membran protein/protein	0	0	0	1	0	0
SCO6614	membran protein/protein	0	1	0	0	0	0
SCO6622	membran protein/protein	0	1	0	0	1	0
SCO6822	membran protein/protein	2	0	0	2	0	0
SCO7197	membran protein/protein	1	1	0	1	1	0
SCO7348	membran protein/protein	0	0	0	0	1	0
SCO7444	membran protein/protein	1	0	0	1	0	0
SCO7605	membran protein/protein	1	0	0	1	0	0
SCO7605	membran protein/protein	1	0	0	1	0	0
SCO7660	membran protein/protein	1	0	0	1	0	0
SCO7746	membran protein/protein	1	0	0	1	0	0
SCO0241	Regulator	0	1	0	0	1	0
SCO0376	Regulator	0	1	0	0	1	0
SCO1774	Regulator	0	1	0	0	1	0
SCO1956	Regulator	1	0	0	0	0	0
SCO2077	Regulator	0	1	0	0	1	0
SCO2200	Regulator	0	0	1	0	0	0
SCO2201	Regulator	0	0	1	0	0	0
SCO2232	Regulator	0	1	0	0	1	0
SCO2489	Regulator	0	1	0	0	1	0
SCO2620	Regulator	1	0	0	1	0	0
SCO2716	Regulator	1	0	0	1	0	0
SCO2745	Regulator	0	0	0	1	0	0
SCO2964	Regulator	1	0	0	1	0	0
SCO2965	Regulator	1	0	0	1	0	0
SCO3158	Regulator	0	1	0	0	1	0
SCO3315	Regulator	0	1	0	0	1	0
SCO3335	Regulator	0	0	1	0	0	0
SCO3413	Regulator	0	1	0	0	1	0
SCO3925	Regulator	1	0	0	0	0	0
SCO4596	Regulator	1	0	0	0	0	0
SCO4628	Regulator	0	1	0	0	1	0
SCO4755	Regulator	0	1	0	0	1	0
SCO4756	Regulator	0	1	0	0	1	0
SCO5025	Regulator	1	0	0	1	0	0
SCO5287	Regulator	0	0	0	1	0	0
SCO5418	Regulator	0	1	0	0	0	0
SCO5614	Regulator	0	1	0	0	1	0
SCO6685	Regulator	0	0	0	1	0	0
SCO6711	Regulator	1	0	0	1	0	0
SCO6722	Regulator	0	1	0	0	1	0
SCO6781	Regulator	1	0	0	0	0	0
SCO7075	Regulator	0	1	0	0	1	0
SCO7093	Regulator	0	1	0	0	1	0
SCO7295	Regulator	0	1	0	0	1	0
SCO7310	Regulator	1	0	0	1	0	0
SCO0014	hypothetical protein	1	0	0	1	0	0
SCO0050	hypothetical protein	0	0	1	0	0	0

SCO0056	hypothetical protein	0	1	0	0	1	0
SCO0133	hypothetical protein	0	0	0	0	0	1
SCO0196	hypothetical protein	0	1	0	0	1	0
SCO0215	hypothetical protein	0	1	0	0	1	0
SCO0244	hypothetical protein	0	1	0	0	1	0
SCO0324	hypothetical protein	0	0	0	0	1	0
SCO0326	hypothetical protein	1	0	0	1	0	0
SCO0753	hypothetical protein	0	2	0	0	2	0
SCO0863	hypothetical protein	0	0	0	1	0	0
SCO0916	hypothetical protein	1	0	0	1	0	0
SCO0989	hypothetical protein	1	0	0	1	0	0
SCO1024	hypothetical protein	0	1	0	0	1	0
SCO1030	hypothetical protein	1	0	0	1	0	0
SCO1107	hypothetical protein	0	0	1	0	0	0
SCO1154	hypothetical protein	1	0	0	1	0	0
SCO1327	hypothetical protein	1	0	0	1	0	0
SCO1328	hypothetical protein	0	0	0	0	1	0
SCO1375	hypothetical protein	0	0	0	1	0	0
SCO1392	hypothetical protein	0	0	0	1	0	0
SCO1536	hypothetical protein	0	0	0	1	0	0
SCO1704	hypothetical protein	1	0	0	0	0	0
SCO1985	hypothetical protein	0	1	0	0	1	0
SCO2006	hypothetical protein	0	1	0	0	1	0
SCO2062	hypothetical protein	0	1	0	0	1	0
SCO2112	hypothetical protein	0	1	0	0	1	0
SCO2255	hypothetical protein	0	0	1	0	0	0
SCO2327	hypothetical protein	0	0	0	1	0	0
SCO2396	hypothetical protein	0	1	0	0	1	0
SCO2454	hypothetical protein	1	0	0	1	0	0
SCO2572	hypothetical protein	0	1	0	0	1	0
SCO2588	hypothetical protein	0	0	0	1	0	0
SCO2621	hypothetical protein	0	1	0	0	1	0
SCO2622	hypothetical protein	0	1	0	0	1	0
SCO2744	hypothetical protein	0	0	0	1	0	0
SCO2842	hypothetical protein	1	0	0	0	0	0
SCO2867	hypothetical protein	0	1	0	0	1	0
SCO2878	hypothetical protein	0	1	0	0	0	0
SCO2921	hypothetical protein	0	1	0	0	1	0
SCO2971	hypothetical protein	0	1	0	0	0	0
SCO3003	hypothetical protein	1	0	0	1	0	0
SCO3004	hypothetical protein	1	0	0	1	0	0
SCO3010	hypothetical protein	0	0	0	1	0	0
SCO3145	hypothetical protein	1	0	0	1	0	0
SCO3152	hypothetical protein	1	0	0	1	0	0
SCO3159	hypothetical protein	0	1	0	0	1	0
SCO3512	hypothetical protein	0	1	0	0	1	0
SCO3576	hypothetical protein	1	0	0	0	0	0
SCO3598	hypothetical protein	0	0	0	1	0	0
SCO3713	hypothetical protein	0	1	0	0	1	0
SCO3760	hypothetical protein	0	1	0	0	1	0
SCO3822	hypothetical protein	1	0	0	0	0	0
SCO3840	hypothetical protein	1	0	0	0	0	0
SCO3888	hypothetical protein	0	1	0	0	1	0
SCO3895	hypothetical protein	0	1	0	0	1	0
SCO3910	hypothetical protein	0	1	0	0	1	0
SCO3922	hypothetical protein	0	0	1	0	0	0

SCO3923	hypothetical protein	0	0	1	0	0	0
SCO4028	hypothetical protein	0	1	0	0	1	0
SCO4064	hypothetical protein	0	1	0	0	1	0
SCO4165	hypothetical protein	0	1	0	0	1	0
SCO4173	hypothetical protein	0	1	0	0	0	0
SCO4175	hypothetical protein	0	1	0	0	1	0
SCO4248	hypothetical protein	0	1	0	0	1	0
SCO4249	hypothetical protein	0	1	0	0	1	0
SCO4281	hypothetical protein	0	1	0	0	1	0
SCO4627	hypothetical protein	0	1	0	0	1	0
SCO4789	hypothetical protein	1	0	0	1	0	0
SCO4806	hypothetical protein	0	1	0	0	1	0
SCO4818	hypothetical protein	1	0	0	0	0	0
SCO4935	hypothetical protein	0	1	0	0	0	0
SCO4941	hypothetical protein	1	0	0	1	0	0
SCO5000	hypothetical protein	0	1	0	0	1	0
SCO5027	hypothetical protein	0	1	0	0	1	0
SCO5158	hypothetical protein	1	0	0	1	0	0
SCO5163	hypothetical protein	0	1	0	0	1	0
SCO5164	hypothetical protein	0	1	0	0	1	0
SCO5168	hypothetical protein	1	0	0	1	0	0
SCO5195	hypothetical protein	0	1	0	0	1	0
SCO5290	hypothetical protein	0	1	0	0	1	0
SCO5294	hypothetical protein	0	1	0	0	1	0
SCO5310	hypothetical protein	0	0	0	1	0	0
SCO5362	hypothetical protein	0	1	0	0	0	0
SCO5439	hypothetical protein	1	0	0	1	0	0
SCO5555	hypothetical protein	1	0	0	0	0	0
SCO5610	hypothetical protein	0	1	0	0	1	0
SCO5619	hypothetical protein	1	0	0	1	0	0
SCO5726	hypothetical protein	0	1	0	0	1	0
SCO6197	hypothetical protein	0	1	0	0	1	0
SCO6389	hypothetical protein	0	1	0	0	1	0
SCO6609	hypothetical protein	0	0	1	0	0	0
SCO6828	hypothetical protein	0	0	1	0	0	0
SCO6846	hypothetical protein	1	0	0	1	0	0
SCO6855	hypothetical protein	1	0	0	1	0	0
SCO6909	hypothetical protein	0	1	0	0	1	0
SCO7224	hypothetical protein	0	0	1	0	0	0
SCO7628	hypothetical protein	1	0	0	1	0	0
SCO7747	hypothetical protein	1	0	0	1	0	0
SCO7772	hypothetical protein	0	1	0	0	1	0
SCO7804	hypothetical protein	0	0	0	1	0	0
SCO7833	hypothetical protein	1	0	0	1	0	0

Table 12 List of 83 genes containing the methylated upstream region in MG and involved in physiological and morphological differentiation.

Name	Description
SCO0014	target bldA
SCO0178	UspA (universal stress protein)
SCO0241	transcriptional regulator
SCO0325	translation initiation factor IF-2 60%
SCO0376	transcriptional regulator
SCO0895	hrdC
SCO0916	SET domain nucleotidase
SCO1024	sporulation protein 97%
SCO1301	Exonuclease
SCO1421	rbpA
SCO1536	Tat pathway signal sequence domain 89%
SCO1774	regulatory protein
SCO1956	LacI family transcription regulator
SCO1973	Methyltransferase
SCO2077	divIVA
SCO2092	S-adenosyl-methyltransferase MraW
SCO2170	Methyltransferase
SCO2200	NmrA family transcriptional regulator
SCO2201	HxlR family transcriptional regulator
SCO2232	malR
SCO2489	TetR family transcriptional regulator
SCO2571	leucyl-tRNA synthetase
SCO2620	trigger factor
SCO2639	σ factor
SCO2716	chpA
SCO2745	LacI family transcriptional regulator
SCO2950	hup (DNA BP)
SCO2964	stgR
SCO3026	PMI
SCO3109	MFD
SCO3149	dimethyladenosine transferase
SCO3158	ssgE
SCO3315	transcriptional regulator
SCO3335	AraC family transcription regulator
SCO3365	DNA binding protein/XRE transcriptional factor S. lividans
SCO3413	transcriptional regulator
SCO3543	topoisomerase I
SCO3702	DNA BP
SCO3732	DEAD/DEAH box helicase
SCO3798	chromosome condensation protein
SCO3896	PAP
SCO3911	dnaB
SCO3925	ssgR
SCO3974	preprotein translocase SecA

SCO4027	Antiσ
SCO4034	sigN
SCO4096	Helicase
SCO4146	ECF family RNA polymerase sigma factor
SCO4176	DNA binding protein
SCO4236	tRNA/rRNA methyltransferase
SCO4441	DNA BP
SCO4596	two-component system response regulator
SCO4628	Regulator
SCO4662	Tuf
SCO4755	transcriptional regulator
SCO4756	transcriptional regulator
SCO4960	σ factor
SCO5025	transcriptional regulator
SCO5287	MarR family transcriptional regulator
SCO5290	cvnB5
SCO5316	whiE
SCO5418	transcriptional regulator
SCO5439	superfamily I DNA and RNA helicase 99% S. lividans
SCO5555	hypothetical protein
SCO5614	transcriptional regulator
SCO5802	Helicase
SCO5820	hrdB
SCO6164	molecular chaperon DnaK
SCO6606	MT
SCO6606	Methyltransferase
SCO6607	Methyltransferase
SCO6685	ramR
SCO6711	transcriptional regulator 99% S. lividans
SCO6722	ssgD
SCO6781	regulatory protein
SCO7075	two component response regulator protein
SCO7093	transcriptional regulator
SCO7213	Methyltransferase
SCO7295	LuxR family transcriptional regulator
SCO7310	regulatory protein
SCO7445	Methyltransferase
SCO7773	histone acetyltrasferase HPA2 97% S. lividans
SCO7833	target bldA

Table 13 List of genes methylated in MI phase in *S. coelicolor* grown in liquid R5A grouped for the methylation GGC^{m}CGG consensus sequence.

Name	Description
SCO0065	extracellular binding protein
SCO0127	beta keto-acyl synthase
SCO0162	hypothetical protein
SCO0163	hypothetical protein
SCO0202	hypothetical protein
SCO0212	hypothetical protein
SCO0217	nitrate reductase subunit beta NarH2
SCO0247	hypothetical protein
SCO0270	hypothetical protein
SCO0293	beta-xylosidase
SCO0305	hypothetical protein
SCO0308	hypothetical protein
SCO0314	hypothetical protein
SCO0321	carboxylesterase
SCO0326	hypothetical protein
SCO0335	hypothetical protein
SCO0347	racemase
SCO0361	hypothetical protein
SCO0363	hypothetical protein
SCO0381	glycosyl transferase
SCO0382	UDP-glucose/GDP-mannose dehydrogenase
SCO0385	hypothetical protein
SCO0387	bi-domain-containing oxidoreductase
SCO0388	hypothetical protein
SCO0391	transferase
SCO0396	hypothetical protein
SCO0397	hypothetical protein
SCO0398	glycosyl transferase
SCO0414	RNA polymerase sigma factor
SCO0416	hypothetical protein
SCO0419	zinc-binding oxidoreductase
SCO0488	hydrolase
SCO0490	esterase
SCO0499	formyltransferase
SCO0511	dehydrogenase
SCO0542	regulatory protein
SCO0571	hypothetical protein
SCO0595	hypothetical protein
SCO0599	regulator of sigB
SCO0606	hypothetical protein
SCO0607	lipoprotein
SCO0609	hypothetical protein
SCO0648	methyltransferase

SCO0676	integral membrane sensor protein
SCO0678	hypothetical protein
SCO0679	hypothetical protein
SCO0686	hypothetical protein
SCO0701	hypothetical protein
SCO0708	branched-chain amino acid ABC transporter
SCO0732	protease
SCO0735	oxidoreductase
SCO0767	hypothetical protein
SCO0777	hypothetical protein
SCO0803	RNA polymerase sigma factor
SCO0815	hypothetical protein
SCO0817	hypothetical protein
SCO0860	cation-transporting ATPase
SCO0863	hypothetical protein
SCO0866	ECF family RNA polymerase sigma factor
SCO0883	peptide deformylase
SCO0916	hypothetical protein
SCO0917	oxygenase
SCO0931	proline-rich protein
SCO0932	hypothetical protein
SCO0935	hypothetical protein
SCO0944	hypothetical protein
SCO0983	malate synthase
SCO0989	hypothetical protein
SCO0993	hypothetical protein
SCO0995	methyltransferase
SCO1009	transposase
SCO1010	integral membrane transport protein
SCO1019	hypothetical protein
SCO1024	hypothetical protein
SCO1030	hypothetical protein
SCO1043	transcriptional regulator
SCO1054	aminotransferase
SCO1068	hypothetical protein
SCO1071	two component system sensor kinase
SCO1072	hypothetical protein
SCO1083	flavin-dependent reductase
SCO1112	oxidoreductase
SCO1113	hypothetical protein
SCO1154	hypothetical protein
SCO1156	hypothetical protein
SCO1170	xylulose kinase
SCO1180	DNA polymerase III subunit beta
SCO1181	hypothetical protein
SCO1184	hypothetical protein
SCO1186	LacI family transcriptional regulator
SCO1192	hypothetical protein

SCO1222	hypothetical protein
SCO1223	ornithine aminotransferase
SCO1266	3-oxoacyl-ACP synthase
SCO1272	acyl carrier protein
SCO1275	halogenase
SCO1281	oxidoreductase
SCO1289	GntR family transcriptional regulator
SCO1290	alkaline phosphatase
SCO1292	hypothetical protein
SCO1308	4-hydroxybenzoate 3-monooxygenase
SCO1331	multidomain-containing protein family
SCO1348	hypothetical protein
SCO1350	hypothetical protein
SCO1359	hypothetical protein
SCO1361	hypothetical protein
SCO1374	hypothetical protein
SCO1375	hypothetical protein
SCO1379	hypothetical protein
SCO1391	phosphoenolpyruvate-protein phosphotransferase
SCO1394	glycosyl hydrolase
SCO1419	phosphodiesterase
SCO1434	CbxX/CfqX family protein
SCO1452	hypothetical protein
SCO1456	hypothetical protein
SCO1490	transcription antitermination protein NusB
SCO1504	regulator
SCO1532	hypothetical protein
SCO1533	hypothetical protein
SCO1534	DNA polymerase III subunit epsilon
SCO1541	regulator
SCO1548	hypothetical protein
SCO1551	protein kinase
SCO1565	glycerophosphoryl diester phosphodiesterase
SCO1571	hypothetical protein
SCO1572	hypothetical protein
SCO1573	oxidoreductase membrane protein
SCO1609	hypothetical protein
SCO1611	short chain dehydrogenase
SCO1612	aldehyde dehydrogenase
SCO1614	transcriptional regulator
SCO1615	hypothetical protein
SCO1625	ribosomal pseudouridine synthase
SCO1643	20S proteasome alpha-subunit
SCO1646	hypothetical protein
SCO1685	hypothetical protein
SCO1686	NTP pyrophosphohydrolase
SCO1703	transcriptional regulator
SCO1704	hypothetical protein

SCO1709	integral membrane transport protein
SCO1712	TetR family transcriptional regulator
SCO1713	hypothetical protein
SCO1720	ABC-transporter transmembrane protein
SCO1744	two-component system sensor kinase
SCO1774	regulatory protein
SCO1781	inorganic polyphosphate/ATP-NAD kinase
SCO1800	chpE
SCO1813	GntR family transcriptional regulator
SCO1836	stress-like protein
SCO1837	hypothetical protein
SCO1859	hypothetical protein
SCO1898	substrate binding protein
SCO1903	transport associated protein
SCO1910	alanine-rich protein
SCO1912	dihydrodipicolinate synthase
SCO1935	transketolase
SCO1936	transaldolase
SCO1956	LacI family transcription regulator
SCO1964	export associated protein
SCO1969	DNA-methyltransferase
SCO1972	sugar kinase
SCO1977	oxidoreductase
SCO1994	hypothetical protein
SCO1996	dephospho-CoA kinase
SCO1997	hypothetical protein
SCO2007	hypothetical protein
SCO2008	branched-chain amino acid ABC transporter substrate-binding protein
SCO2010	branched-chain amino acid ABC transporter permease
SCO2011	branched-chain amino acid ABC transporter ATP-binding protein
SCO2028	hypothetical protein
SCO2035	hypothetical protein
SCO2036	tryptophan synthase subunit alpha
SCO2055	hypothetical protein
SCO2068	alkaline phosphatase
SCO2079	hypothetical protein
SCO2101	carotenoid dehydrogenase
SCO2110	Ser/Thr protein kinase
SCO2124	hypothetical protein
SCO2133	hypothetical protein
SCO2160	large membrane protein
SCO2187	hypothetical protein
SCO2188	peptidase
SCO2210	glutamine synthetase
SCO2211	hypothetical protein
SCO2220	hypothetical protein
SCO2263	hypothetical protein
SCO2267	heme oxygenase

SCO2286	alkaline phosphatase
SCO2302	hypothetical protein
SCO2372	small hydrophobic protein
SCO2373	tetracenomycin C efflux protein
SCO2376	hypothetical protein
SCO2377	aldo/keto reductase
SCO2407	aldose 1-epimerase
SCO2408	aminotransferase
SCO2426	regulatory protein
SCO2428	phosphate binding protein
SCO2448	hypothetical protein
SCO2450	Ser/Thr protein kinase (regulator)
SCO2462	sugar kinase
SCO2472	hypothetical protein
SCO2475	LysR family transcriptional regulator
SCO2477	short chain dehydrogenase
SCO2495	hypothetical protein
SCO2516	hypothetical protein
SCO2529	metalloprotease
SCO2537	DNA-binding protein
SCO2540	carbohydrate kinase
SCO2542	glucarate dehydratase
SCO2544	IclR family transcriptional regulator
SCO2549	protease
SCO2551	hypothetical protein
SCO2556	hypothetical protein
SCO2564	DNA-binding protein
SCO2573	oxidoreductase
SCO2588	hypothetical protein
SCO2596	50S ribosomal protein L27
SCO2623	hypothetical protein
SCO2637	serine protease
SCO2647	MarR family regulatory protein
SCO2657	transcriptional regulator
SCO2663	hypothetical protein
SCO2664	sugar-binding protein
SCO2669	hypothetical protein
SCO2672	hypothetical protein
SCO2684	ATP-binding membrane protein
SCO2696	2-hydroxyacid dehydrogenase
SCO2764	lipoprotein
SCO2765	hypothetical protein
SCO2776	acetyl/propionyl CoA carboxylase subunit beta
SCO2777	acetyl/propionyl CoA carboxylase subunit alpha
SCO2778	hydroxymethylglutaryl-CoA lyase
SCO2782	pyridoxal-dependent decarboxylase
SCO2783	monooxygenase
SCO2800	two component system histidine kinase

SCO2809	hypothetical protein
SCO2816	hypothetical protein
SCO2823	decarboxylase
SCO2824	NAD-binding protein
SCO2832	IclR family transcriptional regulator
SCO2842	hypothetical protein
SCO2875	MerR family transcriptional regulator
SCO2882	ATP/GTP-binding protein
SCO2883	cytochrome P450
SCO2885	hypothetical protein
SCO2908	hypothetical protein
SCO2909	hypothetical protein
SCO2918	nicotinamidase
SCO2926	hypothetical protein
SCO2965	transporter
SCO2977	hypothetical protein
SCO2981	glycosyl transferase
SCO2983	glycosyl transferase
SCO2985	hypothetical protein
SCO3004	hypothetical protein
SCO3023	S-adenosyl-L-homocysteine hydrolase
SCO3061	hypothetical protein
SCO3105	hypothetical protein
SCO3107	lipoprotein
SCO3108	hypothetical protein
SCO3111	ABC transporter ATP-binding protein
SCO3156	penicillin-binding protein
SCO3193	hypothetical protein
SCO3207	TetR family transcriptional regulator
SCO3211	indoleglycerol phosphate synthase
SCO3215	hypothetical protein
SCO3217	transcriptional regulator cdaR
SCO3220	hypothetical protein
SCO3228	glycolate oxidase
SCO3230	CDA peptide synthetase I cdaPSI
SCO3231	CDA peptide synthetase II cdaPSII
SCO3232	CDA peptide synthetase III cdaPSIII
SCO3233	hydrolase
SCO3248	3-oxoacyl-ACP synthase
SCO3287	serine/arginine rich protein
SCO3288	hypothetical protein
SCO3289	large membrane protein
SCO3290	hypothetical protein
SCO3291	regulatory protein
SCO3318	porphobilinogen deaminase
SCO3357	hypothetical protein
SCO3365	hypothetical protein
SCO3371	hypothetical protein

SCO3378	small membrane protein
SCO3407	hypothetical protein
SCO3434	DNA polymerase I
SCO3457	transmembrane protein
SCO3547	membrane-bound proton-translocating pyrophosphatase
SCO3552	hypothetical protein
SCO3557	septum site determining protein
SCO3558	morphological differentiation-associated protein
SCO3577	ion-transporting ATPase
SCO3578	ion-transporting ATPase
SCO3592	hypothetical protein
SCO3658	aminotransferase
SCO3681	hypothetical protein
SCO3701	hypothetical protein
SCO3712	hydrolase
SCO3723	regulatory protein
SCO3850	hypothetical protein
SCO3934	ftsK/SpoIIIE family protein
SCO3959	hypothetical protein
SCO3967	hypothetical protein
SCO3975	regulator
SCO4003	hypothetical protein
SCO4005	RNA polymerase sigma factor
SCO4014	hypothetical protein
SCO4015	hypothetical protein
SCO4025	TetR family transcriptional regulator
SCO4035	RNA polymerase sigma factor sigF
SCO4053	transport integral membrane protein
SCO4069	hypothetical protein
SCO4085	lipoprotein
SCO4087	phosphoribosylaminoimidazole synthetase
SCO4094	hypothetical protein
SCO4108	peptidase
SCO4115	hypothetical protein
SCO4124	two-component sensor kinase
SCO4132	transglycosylase
SCO4148	ABC transporter ATP-binding protein
SCO4199	hypothetical protein
SCO4218	small hydrophilic protein
SCO4224	hypothetical protein
SCO4237	hypothetical protein
SCO4279	acetyltransferase
SCO4290	trehalose-phosphate synthase
SCO4294	hypothetical protein
SCO4296	chaperonin GroEL
SCO4314	hypothetical protein
SCO4317	hypothetical protein
SCO4319	hypothetical protein

SCO4322	hypothetical protein
SCO4332	integral membrane ATPase
SCO4338	hypothetical protein
SCO4367	oxidoreductase
SCO4368	lipase
SCO4369	hypothetical protein
SCO4378	hypothetical protein
SCO4390	hypothetical protein
SCO4394	iron repressor
SCO4399	hypothetical protein
SCO4401	lipoprotein
SCO4402	hypothetical protein
SCO4411	calcium binding protein
SCO4444	glutathione peroxidase
SCO4507	Ser/Thr protein kinase
SCO4565	NADH dehydrogenase subunit D
SCO4566	NADH dehydrogenase subunit E
SCO4567	NADH dehydrogenase subunit NuoF
SCO4569	NADH dehydrogenase subunit H
SCO4570	NADH dehydrogenase subunit I
SCO4574	NADH dehydrogenase subunit M
SCO4575	NADH dehydrogenase subunit N
SCO4577	helicase
SCO4581	hypothetical protein
SCO4584	hypothetical protein
SCO4592	hypothetical protein
SCO4593	hypothetical protein
SCO4596	two-component system response regulator
SCO4632	ATP/GTP binding protein
SCO4650	lipoprotein
SCO4667	two-component system sensor kinase
SCO4743	hypothetical protein
SCO4746	lipase
SCO4762	chaperonin GroEL
SCO4774	glycerol phosphate dehydrogenase
SCO4775	Ser/Thr protein kinase
SCO4779	Ser/Thr protein kinase
SCO4780	succinic semialdehyde dehydrogenase
SCO4783	hypothetical protein
SCO4789	hypothetical protein
SCO4792	two-component system DNA-binding response regulator
SCO4801	hypothetical protein
SCO4803	hypothetical protein
SCO4805	hypothetical protein
SCO4809	succinyl-CoA synthetase subunit alpha
SCO4818	hypothetical protein
SCO4820	Ser/Thr protein kinase
SCO4842	oxidoreductase

SCO4843	hypothetical protein
SCO4861	hypothetical protein
SCO4862	hypothetical protein
SCO4880	transferase
SCO4882	hypothetical protein
SCO4883	peptidase
SCO4892	regulatory protein
SCO4894	hypothetical protein
SCO4903	hypothetical protein
SCO4907	transcriptional regulator
SCO4908	RNA polymerase sigma factor sigQ
SCO4910	hypothetical protein
SCO4941	hypothetical protein
SCO4947	nitrate reductase subunit alpha NarG3
SCO4950	nitrate reductase subunit gamma NarI3
SCO4972	dehydrogenase
SCO4973	hypothetical protein
SCO5006	septum site-determining protein
SCO5007	septum site-determining protein
SCO5010	hypothetical protein
SCO5011	hypothetical protein
SCO5017	AraC family transcription regulator
SCO5024	oxidoreductase
SCO5053	hypothetical protein
SCO5079	hypothetical protein actVA3
SCO5080	hydrolase actVA4
SCO5084	hypothetical protein actII-3
SCO5088	actinorhodin polyketide beta-ketoacyl synthase subunit beta actIORF2
SCO5090	actinorhodin polyketide synthase bifunctional cyclase/dehydratase actVORFII
SCO5091	cyclase actIV
SCO5092	actinorhodin polyketide dimerase actVB
SCO5103	regulatory protein
SCO5109	hypothetical protein
SCO5122	peptidase
SCO5127	hypothetical protein
SCO5149	protease
SCO5151	hypothetical protein
SCO5168	hypothetical protein
SCO5176	reductase
SCO5177	hypothetical protein
SCO5188	ATP-dependent DNA helicase
SCO5206	hydrogen peroxide sensitive repressor
SCO5208	monophosphatase
SCO5209	TetR family transcriptional regulator
SCO5218	hypothetical protein
SCO5219	lipoprotein
SCO5228	acetyltransferase

SCO5247	deaminase
SCO5251	acetyltransferase
SCO5253	hypothetical protein
SCO5254	superoxide dismutase
SCO5255	signal peptidase
SCO5257	methyltransferase
SCO5280	ATP-binding protein
SCO5292	ATP/GTP-binding protein
SCO5306	protein phosphatase
SCO5307	hypothetical protein
SCO5310	hypothetical protein
SCO5315	whiE, ORFIV
SCO5316	whiE, ORFV
SCO5329	hypothetical protein
SCO5349	integrase
SCO5356	homoserine kinase
SCO5375	hypothetical protein
SCO5387	hypothetical protein
SCO5391	ATP/GTP-binding protein
SCO5392	ABC transporter
SCO5444	glycogen phosphorylase
SCO5447	neutral zinc metalloprotease
SCO5463	MerR family transcriptional regulator
SCO5465	hypothetical protein
SCO5473	ATP/GTP binding protein
SCO5474	hypothetical protein
SCO5493	hypothetical protein
SCO5494	NAD-dependent DNA ligase LigA
SCO5512	acetolactate synthase 1 catalytic subunit
SCO5515	D-3-phosphoglycerate dehydrogenase
SCO5517	transcriptional regulator
SCO5519	hypothetical protein
SCO5520	delta-1-pyrroline-5-carboxylate dehydrogenase
SCO5522	3-isopropylmalate dehydrogenase
SCO5523	branched-chain amino acid aminotransferase
SCO5541	ATP-GTP binding protein
SCO5577	chromosome associated protein
SCO5582	regulator
SCO5583	ammonium transporter
SCO5584	nitrogen regulatory protein P-II
SCO5585	PII uridylyl-transferase
SCO5652	hypothetical protein
SCO5658	polyamine-binding lipoprotein
SCO5671	oxidoreductase
SCO5677	ATP/GTP binding protein
SCO5684	two-component system response regulator
SCO5708	ribosome-binding factor A
SCO5710	large Pro/Ala/Gly-rich protein

SCO5748	sensory histidine kinase
SCO5750	ftsK-like protein
SCO5759	hypothetical protein
SCO5761	ATP-dependent DNA helicase
SCO5762	AraC family transcription regulator
SCO5765	hypothetical protein
SCO5781	hypothetical protein
SCO5783	hypothetical protein
SCO5784	two-component sensor
SCO5794	kinase/phosphohydrolase
SCO5795	zinc metalloprotease membrane protein
SCO5797	hypothetical protein
SCO5799	aminotransferase
SCO5800	hypothetical protein
SCO5801	hypothetical protein
SCO5802	ATP-dependent helicase
SCO5807	hypothetical protein
SCO5808	hypothetical protein
SCO5816	hypothetical protein
SCO5820	RNA polymerase sigma factor HrdB
SCO5824	two-component sensor
SCO5827	transmembrane transport protein
SCO5828	two-component transcriptional regulator
SCO5849	AgaS protein
SCO5863	two-component sensor (kinase)
SCO5865	hypothetical protein
SCO5868	deoxyuridine 5'-triphosphate nucleotidohydrolase
SCO5874	hypothetical protein
SCO5878	polyketide synthase RedX
SCO5879	acyl-coa dehydrogenase RedW
SCO5881	response regulator redZ
SCO5882	RedV protein
SCO5887	acyl carrier protein redQ
SCO5888	3-oxoacyl-ACP synthase redP
SCO5890	8-amino-7-oxononanoate synthase redN
SCO5892	polyketide synthase redL
SCO5896	phosphoenolpyruvate-utilizing enzyme redH
SCO5898	hypothetical protein redF
SCO5899	hypothetical protein
SCO5912	protease
SCO5955	hypothetical protein
SCO5980	salicylyl-CoA 5-hydroxylase
SCO5985	hypothetical protein
SCO5986	oxidoreductase
SCO5994	integral membrane cytochrome biogenesis protein
SCO5995	hypothetical protein
SCO6007	transmembrane transport protein
SCO6019	hypothetical protein

SCO6037	transmembrane transport protein
SCO6039	flavoprotein oxidoreductase
SCO6041	protoporphyrinogen oxidase
SCO6047	ABC transporter ATP-binding protein
SCO6073	cyclase geoA
SCO6074	hypothetical protein
SCO6077	transferase
SCO6083	hypothetical protein
SCO6109	hydrolase
SCO6127	carboxylesterase
SCO6145	hypothetical protein
SCO6146	hypothetical protein
SCO6150	ADA-like regulatory protein
SCO6151	methylated-DNA-protein-cysteine methyltransferase
SCO6159	GntR family transcriptional regulator
SCO6171	oxidoreductase
SCO6172	oxidoreductase
SCO6173	permease SC6C509
SCO6175	hypothetical protein
SCO6182	dehydratase
SCO6187	bifunctional synthase/transferase
SCO6195	acetyl-coenzyme A synthetase
SCO6201	glyoxylate carboligase
SCO6247	allantoinase
SCO6248	allantoicase
SCO6262	helicase
SCO6263	hypothetical protein
SCO6272	FAD-binding protein
SCO6276	hypothetical protein
SCO6278	integral membrane transport protein
SCO6280	regulatory protein
SCO6349	transcriptional regulator
SCO6358	hypothetical protein
SCO6380	hypothetical protein
SCO6393	transposase
SCO6407	gamma-glutamyltranspeptidase
SCO6417	integral membrane transporter
SCO6431	peptide synthase
SCO6432	peptide synthase
SCO6433	hypothetical protein
SCO6434	oxidoreductase
SCO6435	hypothetical protein
SCO6437	hypothetical protein
SCO6438	diaminopimelate decarboxylase
SCO6459	regulatory protein
SCO6483	efflux protein
SCO6484	hypothetical protein
SCO6506	gas vesicle protein

SCO6507	gas vesicle synthesis protein
SCO6517	uvrA-like protein
SCO6520	RNA polymerase sigma factor
SCO6522	hypothetical protein
SCO6527	hypothetical protein
SCO6531	ATP/GTP binding protein
SCO6539	hypothetical protein
SCO6540	pterin-4-alpha-carbinolamine dehydratase
SCO6541	hypothetical protein
SCO6550	oxidoreductase
SCO6569	solute binding protein
SCO6570	oxidoreductase
SCO6597	beta-glucosidase
SCO6598	transcriptional regulator
SCO6601	sugar binding protein
SCO6609	hypothetical protein
SCO6621	hypothetical protein
SCO6632	hypothetical protein
SCO6651	glycosyl transferase
SCO6652	hypothetical protein
SCO6653	hypothetical protein
SCO6681	Ser/Thr protein kinase
SCO6685	two-component system response regulator
SCO6691	phospholipase C
SCO6694	transcriptional regulator
SCO6696	regulatory protein
SCO6697	3-oxoadipate enol-lactone hydrolase/4-carboxymuconolactone decarboxylase
SCO6712	copper oxidase
SCO6713	transcriptional regulator
SCO6715	transcriptional regulator
SCO6721	hypothetical protein
SCO6723	oxidoreductase
SCO6731	acetyl-CoA acetyltransferase
SCO6735	hypothetical protein
SCO6746	transcriptional regulator
SCO6759	phytoene synthase
SCO6760	phytoene synthase
SCO6762	phytoene dehydrogenase
SCO6764	squalene-hopene cyclase
SCO6776	hypothetical protein
SCO6799	L-threonine 3-dehydrogenase
SCO6805	integral membrane efflux protein
SCO6822	integral membrane efflux protein
SCO6828	hypothetical protein
SCO6846	hypothetical protein
SCO6886	hypothetical protein
SCO6887	hypothetical protein

SCO6889	hypothetical protein
SCO6907	DNA ligase
SCO6967	beta-ketoadipyl-CoA thiolase
SCO6971	precorrin 6A synthase
SCO6979	solute-binding lipoprotein
SCO6989	hypothetical protein
SCO7004	carbohydrate kinase
SCO7021	hypothetical protein
SCO7058	hypothetical protein
SCO7072	hypothetical protein
SCO7092	hypothetical protein
SCO7093	transcriptional regulator
SCO7104	RNA polymerase sigma factor
SCO7122	acetyltransferase
SCO7134	transcriptional regulator
SCO7142	hypothetical protein
SCO7143	transcriptional regulator
SCO7182	branched-chain amino acid ABC transporter ATP-binding protein
SCO7202	hypothetical protein
SCO7227	hypothetical protein
SCO7232	hypothetical protein
SCO7233	hypothetical protein
SCO7240	Ser/Thr protein kinase
SCO7241	oxidoreductase
SCO7252	regulatory protein
SCO7293	hypothetical protein
SCO7298	thioredoxin reductase
SCO7299	stress-inducible protein
SCO7326	hypothetical protein
SCO7337	hypothetical protein
SCO7338	glycogen debranching protein
SCO7361	DNA-binding protein
SCO7366	hypothetical protein
SCO7367	membrane efflux protein
SCO7368	hypothetical protein
SCO7403	hypothetical protein
SCO7433	small neutral protease regulatory protein
SCO7434	lipoprotein sspA (tarfget of sigF)
SCO7443	phosphoglucomutase
SCO7444	cytochrome P450 (fragment)
SCO7460	lipoprotein
SCO7470	phenylacetic acid degradation protein PaaI
SCO7471	phenylacetate-CoA oxygenase subunit PaaA
SCO7539	TetR family transcriptional regulator
SCO7553	oxidoreductase
SCO7568	regulatory protein
SCO7569	hypothetical protein
SCO7580	hypothetical protein

SCO7600	alanyl tRNA synthetase
SCO7602	TetR family transcriptional regulator
SCO7605	metallopeptidase
SCO7606	amino acid binding protein
SCO7608	hypothetical protein
SCO7610	transcriptional regulator
SCO7620	hypothetical protein
SCO7627	hypothetical protein
SCO7628	hypothetical protein
SCO7657	hypothetical protein
SCO7659	oxidoreductase
SCO7660	voltage-gated potassium channel
SCO7672	hypothetical protein
SCO7677	solute-binding protein
SCO7680	ABC transporter ATP-binding protein
SCO7684	hypothetical protein
SCO7688	hypothetical protein
SCO7697	hydrolase
SCO7715	hypothetical protein
SCO7721	hypothetical protein
SCO7732	hypothetical protein
SCO7779	oxidoreductase
SCO7780	transcriptional regulator
SCO7798	transposase
SCO7800	hypothetical protein

Table 14 List of genes methylated in MII phase in *S. coelicolor* grown in liquid R5A grouped for the methylation GCC^{m}CG consensus sequence.

Name	Description
SCO0178	hypothetical protein
SCO0196	hypothetical protein
SCO0211	hypothetical protein
SCO0324	hypothetical protein
SCO0325	hypothetical protein
SCO0352	solute-binding protein
SCO0375	integral membrane transport protein
SCO0389	lipoprotein
SCO0544	hypothetical protein
SCO0608	regulatory protein
SCO0752	protease
SCO0888	hypothetical protein
SCO0973	hypothetical protein
SCO1084	thioredoxin
SCO1085	acyltransferase
SCO1104	TetR family transcriptional regulator
SCO1121	hypothetical protein
SCO1169	xylose isomerase
SCO1174	aldehyde dehydrogenase
SCO1182	hypothetical protein
SCO1371	oxidoreductase
SCO1421	hypothetical protein
SCO1489	DNA-binding protein
SCO1529	hypothetical protein
SCO1564	RNA polymerase sigma factor
SCO1636	hypothetical protein
SCO1668	hypothetical protein
SCO1671	hypothetical protein
SCO1793	hypothetical protein
SCO1797	hypothetical protein
SCO1968	hydrolase
SCO1981	hypothetical protein
SCO1991	hypothetical protein
SCO2077	hypothetical protein
SCO2094	regulatory protein
SCO2113	bacterioferritin
SCO2231	maltose-binding protein
SCO2251	hypothetical protein
SCO2256	3-methyl-2-oxobutanoate hydroxymethyltransferase
SCO2313	hypothetical protein
SCO2385	hypothetical protein
SCO2404	sugar-binding receptor
SCO2430	beta-galactosidase

SCO2487	nitrite reductase large subunit NirB
SCO2488	nitrite reductase small subunit NirC
SCO2492	hypothetical protein
SCO2517	two-component system response regulator
SCO2621	hypothetical protein
SCO2733	hypothetical protein
SCO2734	LysR family transcriptional regulator
SCO2779	acyl-CoA dehydrogenase
SCO2784	aceytltranferase
SCO2967	carboxy-terminal processing protease
SCO3019	lipoprotein
SCO3036	2-phospho-L-lactate transferase
SCO3106	lipoprotein
SCO3404	cell division protein FtsH-like protein
SCO3490	transposase
SCO3531	hypothetical protein
SCO3562	integral membrane transport protein
SCO3713	hypothetical protein
SCO3767	hypothetical protein
SCO3768	translocase
SCO3835	dehydrogenase
SCO3884	hypothetical protein
SCO3944	aminotransferase
SCO3977	protease
SCO4166	hypothetical protein
SCO4175	hypothetical protein
SCO4176	hypothetical protein
SCO4193	ATP/GTP-binding membrane protein
SCO4280	reductase
SCO4281	hypothetical protein
SCO4297	oxidoreductase
SCO4428	hypothetical protein
SCO4441	DNA-binding protein
SCO4568	NADH dehydrogenase subunit G
SCO4627	hypothetical protein
SCO4633	hypothetical protein
SCO4674	hypothetical protein
SCO4773	nucleotide-sugar dehydrogenase
SCO4788	ntegral membrane protein
SCO4804	hypothetical protein
SCO4826	hypothetical protein
SCO4895	RNA polymerase factor sigma-70
SCO4975	hypothetical protein
SCO5018	hypothetical protein
SCO5074	dehydratase
SCO5075	oxidoreductase
SCO5078	hypothetical protein
SCO5085	actII-4 actinorhodin cluster activator protein

SCO5097	short-chain oxidoreductase
SCO5100	GntR family transcriptional regulator
SCO5104	hypothetical protein
SCO5150	sec-independent translocase
SCO5160	hypothetical protein
SCO5194	hypothetical protein
SCO5202	hypothetical protein
SCO5332	hypothetical protein
SCO5417	zinc-binding oxidoreductase
SCO5489	hypothetical protein
SCO5491	hypothetical protein
SCO5557	hypothetical protein
SCO5565	hypothetical protein
SCO5566	ATP-dependent DNA helicase RecG
SCO5580	docking protein
SCO5604	hypothetical protein
SCO5605	hypothetical protein
SCO5610	hypothetical protein
SCO5614	transcriptional regulator
SCO5629	ATP /GTP-binding protein
SCO5639	hypothetical protein
SCO5645	ribosomal RNA large subunit methyltransferase N
SCO5693	acyl CoA dehydrogenase
SCO5696	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase
SCO5751	hypothetical protein
SCO5788	hypothetical protein
SCO5796	hypothetical protein
SCO5815	ATP-dependent DNA helicase
SCO5852	tagatose-bisphosphate aldolase
SCO5897	oxidase
SCO5917	hypothetical protein
SCO6045	hypothetical protein
SCO6055	carbonic anhydrase
SCO6108	esterase
SCO6188	transferase
SCO6204	catalase
SCO6209	OHCU decarboxylase
SCO6233	transcriptional regulator
SCO6251	reductase
SCO6351	hypothetical protein
SCO6360	ABC transporter ATP-binding protein
SCO6384	integral membrane lysyl-tRNA synthetase
SCO6390	hypothetical protein
SCO6394	IS element ATP binding protein
SCO6411	hydrolase
SCO6439	DNA-binding protein
SCO6461	ADA-like regulatory protein
SCO6493	hypothetical protein

SCO6581	transmembrane transport protein
SCO6593	hypothetical protein
SCO6608	hypothetical protein
SCO6720	ABC transporter
SCO6773	peptidase
SCO6788	acetyl-CoA acetyltransferase
SCO6800	2-amino-3-ketobutyrate coenzyme A ligase
SCO6905	hypothetical protein
SCO7008	ABC transporter ATP-binding protein
SCO7023	hypothetical protein
SCO7059	oxidoreductase
SCO7075	two component response regulator protein
SCO7248	hypothetical protein
SCO7278	RNA polymerase sigma factor
SCO7279	DNA-binding protein
SCO7432	extracellular small neutral protease
SCO7455	isochorismatase
SCO7472	phenylacetate-CoA oxygenase subunit PaaB
SCO7473	phenylacetic acid degradation protein PaaC
SCO7536	hypothetical protein
SCO7560	MutT-family protein
SCO7577	hydrolase
SCO7623	NAD(P) transhydrogenase subunit alpha
SCO7645	TetR family transcriptional regulator
SCO7652	acetyltransferase
SCO7702	GntR family transcriptional regulator
SCO7722	hypothetical protein
SCO7733	transcriptional regulator
SCO7734	hypothetical protein
SCO7784	oxidoreductase

Table 15 List of genes methylated in MII phase in *S. coelicolor* grown in liquid R5A grouped for the methylation $C^{m}GGGC$ consensus sequence.

Name	Description
SCO0069	hypothetical protein
SCO0122	flavin-containing monooxygenase
SCO0130	beta-lactamase
SCO0135	hypothetical protein
SCO0191	lycopene cyclase
SCO0240	oxidoreductase
SCO0241	transcriptional regulator
SCO0311	ligase
SCO0318	oxidoreductase
SCO0467	hypothetical protein
SCO0502	hypothetical protein
SCO0503	hypothetical protein
SCO0530	transcriptional regulator
SCO0531	sugar transporter sugar binding protein
SCO0538	sugar transporter sugar binding lipoprotein
SCO0654	hypothetical protein
SCO0655	gas vesicle synthesis protein
SCO0664	hypothetical protein
SCO0692	hypothetical protein
SCO0693	hypothetical protein
SCO0694	hypothetical protein
SCO0723	fructose transport system kinase
SCO0829	serine protease
SCO0894	hypothetical protein
SCO0895	RNA polymerase principal sigma factor HrdC
SCO0951	transport system permease
SCO0967	reductase
SCO0986	hypothetical protein
SCO1007	oxidoreductase
SCO1037	hypothetical protein
SCO1038	hypothetical protein
SCO1109	oxidoreductase
SCO1354	hypothetical protein
SCO1382	hypothetical protein
SCO1395	mutT-like protein
SCO1418	hypothetical protein
SCO1516	preprotein translocase subunit SecD
SCO1519	Holliday junction DNA helicase RuvA
SCO1579	bifunctional ornithine acetyltransferase/N-acetylglutamate synthase
SCO1867	hydroxylase
SCO1921	aminotransferase
SCO1922	ABC transporter ATP-binding protein
SCO1944	preprotein translocase subunit SecG

SCO2004	formate dehydrogenase
SCO2024	chitosanase
SCO2062	hypothetical protein
SCO2063	small hydrophilic protein
SCO2087	phospho-N-acetylmuramoyl-pentapeptide- transferase
SCO2198	glutamine synthetase
SCO2215	two-component system sensor kinase
SCO2261	hypothetical protein
SCO2276	hypothetical protein
SCO2277	hypothetical protein
SCO2316	hypothetical protein
SCO2361	hypothetical protein
SCO2398	MarR family transcriptonal regulator
SCO2433	sugar transporter membrane protein
SCO2451	rod shape-determining protein MreB
SCO2463	ABC transporter
SCO2553	oxidoreductase
SCO2571	leucyl-tRNA synthetase
SCO2591	hypothetical protein
SCO2639	RNA polymerase sigma factor
SCO2640	aspartate-semialdehyde dehydrogenase
SCO2747	bifunctional carbohydrate binding and transport protein
SCO2815	TetR family transcriptional regulator
SCO2829	amino acid ABC transporter transmembrane protein
SCO2867	hypothetical protein
SCO2868	hypothetical protein
SCO2907	PTS transmembrane protein
SCO2920	protease
SCO2959	nitrate extrusion protein
SCO3006	acetyltransferase
SCO3021	hypothetical protein
SCO3059	phosphoribosylaminoimidazole carboxylase catalytic subunit PurE
SCO3067	anti anti sigma factor
SCO3068	RNA polymerase sigma factor
SCO3140	hypothetical protein
SCO3319	glutamyl-tRNA reductase
SCO3391	hypothetical protein
SCO3439	hypothetical protein
SCO3440	hypothetical protein
SCO3474	sugar kinase
SCO3506	LacI family transcriptional regulator
SCO3509	hypothetical protein
SCO3511	lipoprotein
SCO3512	hypothetical protein
SCO3513	hypothetical protein
SCO3550	helicase
SCO3650	orotate phosphoribosyltransferase

SCO3692	anti-sigma factor antagonist
SCO3727	stress response protein
SCO3743	hypothetical protein
SCO3827	molybdenum cofactor biosynthesis protein
SCO3891	hypothetical protein
SCO3895	hypothetical protein
SCO3896	RNA nucleotidyltransferase
SCO3910	hypothetical protein
SCO3911	replicative DNA helicase
SCO3915	transmembrane efflux protein
SCO3922	hypothetical protein
SCO3923	hypothetical protein
SCO3981	GntR family transcriptional regulator
SCO3982	hypothetical protein
SCO3983	hypothetical protein
SCO3993	hypothetical protein
SCO3994	hypothetical protein
SCO3995	partial replication initiator protein
SCO4016	hypothetical protein
SCO4028	hypothetical protein
SCO4158	LacI-family regulatory protein
SCO4243	hypothetical protein
SCO4248	hypothetical protein
SCO4249	hypothetical protein
SCO4250	hypothetical protein
SCO4252	hypothetical protein
SCO4341	hypothetical protein
SCO4345	hypothetical protein
SCO4366	phosphoserine aminotransferase
SCO4482	LysR family transcriptional regulator
SCO4484	hypothetical protein
SCO4487	Ser/Thr protein kinase
SCO4506	hypothetical protein
SCO4514	hypothetical protein
SCO4520	hypothetical protein
SCO4606	NADH dehydrogenase subunit NuoL2
SCO4692	hypothetical protein
SCO4693	hypothetical protein
SCO4733	hypothetical protein
SCO4734	50S ribosomal protein L13
SCO4756	hypothetical protein
SCO4836	GntR family transcriptional regulator
SCO4837	serine hydroxymethyltransferase
SCO4940	TetR family transcriptional regulator
SCO4964	integral membrane transport protein
SCO4989	IclR family transcriptional regulator
SCO5000	hypothetical protein
SCO5232	sugar transporter sugar binding protein

SCO5365	transferase
SCO5366	ATP synthase I
SCO5367	F0F1 ATP synthase subunit A
SCO5476	oligopeptide transport integral membrane protein
SCO5488	tRNA-specific 2-thiouridylase MnmA
SCO5492	short chain dehydrogenase
SCO5526	urease subunit alpha
SCO5615	hypothetical protein
SCO5672	hypothetical protein
SCO5673	chitinase
SCO5689	beta-galactosidase
SCO5709	tRNA pseudouridine synthase B
SCO5724	hypothetical protein
SCO5726	hypothetical protein
SCO5737	polynucleotide phosphorylase/polyadenylase
SCO5745	hypothetical protein
SCO5818	ABC transporter
SCO5833	hypothetical protein
SCO5848	tagatose 6-phosphate kinase
SCO6033	hypothetical protein
SCO6095	ABC transporter ATP-binding protein
SCO6376	hypothetical protein
SCO6429	hypothetical protein
SCO6585	succinyl-CoA synthetase subunit beta
SCO6604	beta-glucosidase
SCO6607	hypothetical protein
SCO6614	hypothetical protein
SCO6660	hypothetical protein
SCO6732	fatty acid oxidative multifunctional enzyme
SCO6790	long chain fatty acid CoA ligase
SCO6837	arsenic resistance membrane transport protein
SCO6841	hypothetical protein
SCO6849	hypothetical protein
SCO6861	protein kinase-like protein
SCO6878	ATP-binding protein
SCO6883	hypothetical protein
SCO6898	carboxylase
SCO6954	monovalent cation/H+ antiporter subunit A
SCO6957	monovalent cation/H+ antiporter subunit E
SCO7018	hypothetical protein
SCO7047	UDP pyrophosphate phosphatase
SCO7101	dehydrogenase
SCO7159	hypothetical protein
SCO7318	hypothetical protein
SCO7374	nitrate reductase NarB (fragment)
SCO7384	transmembrane transport protein
SCO7452	O-methyltransferase
SCO7504	integral membrane binding-protein-dependent transport protein

SCO7519	sugar acetyltransferase
SCO7576	hydrolase
SCO7674	metal-binding protein
SCO7679	transport system integral membrane protein
SCO7686	cytochrome P450
SCO7687	thioesterase
SCO7703	integral membrane transport protein
SCO7704	TetR family transcriptional regulator
SCO7751	regulatory protein
SCO7762	hypothetical protein
SCO7772	hypothetical protein
SCO7805	hypothetical protein
SCO7811	hypothetical protein

Table 16 List of genes methylated in MI phase in *S. coelicolor* grown on solid GYM grouped for the methylation GGC^{m}CGG consensus sequence.

Name	Product
SCO0026	hypothetical protein
SCO0042	hypothetical protein
SCO0062	LacI family transcriptional regulator
SCO0089	transcriptional regulator
SCO0103	flavohemoprotein
SCO0118	xylosidase/arabinosidase
SCO0147	transmembrane transport protein
SCO0148	transcriptional regulator
SCO0150	hypothetical protein
SCO0162	hypothetical protein
SCO0163	hypothetical protein
SCO0184	hypothetical protein
SCO0202	hypothetical protein
SCO0212	hypothetical protein
SCO0217	nitrate reductase subunit beta NarH2
SCO0219	nitrate reductase subunit delta NarI2
SCO0247	hypothetical protein
SCO0285	sodium/proton antiporter
SCO0294	acetyltransferase
SCO0299	oxidoreductase
SCO0305	hypothetical protein
SCO0308	hypothetical protein
SCO0314	hypothetical protein
SCO0329	esterase
SCO0331	short chain oxidoreductase
SCO0335	hypothetical protein
SCO0347	racemase
SCO0356	oxidoreductase
SCO0361	hypothetical protein
SCO0362	racemase
SCO0381	glycosyl transferase
SCO0382	UDP-glucose/GDP-mannose dehydrogenase
SCO0385	hypothetical protein
SCO0387	bi-domain-containing oxidoreductase
SCO0388	hypothetical protein
SCO0391	transferase
SCO0396	hypothetical protein
SCO0397	hypothetical protein
SCO0398	glycosyl transferase
SCO0406	hypothetical protein
SCO0407	transcriptional regulator
SCO0414	RNA polymerase sigma factor
SCO0419	zinc-binding oxidoreductase

SCO0438	pyrrolidone-carboxylate peptidase
SCO0441	LamB/YcsF family protein
SCO0457	hypothetical protein
SCO0470	hydrolase
SCO0476	ABC transporter ATP-binding protein
SCO0488	hydrolase
SCO0490	esterase
SCO0499	formyltransferase
SCO0541	alpha-galactosidase
SCO0544	hypothetical protein
SCO0545	hypothetical protein
SCO0557	MerR family transcriptional regulator
SCO0571	hypothetical protein
SCO0577	hypothetical protein
SCO0583	cytochrome P450
SCO0599	regulator of sig8
SCO0607	lipoprotein
SCO0638	lipoprotein
SCO0676	integral membrane sensor protein
SCO0678	hypothetical protein
SCO0686	hypothetical protein
SCO0687	oxidoreductase
SCO0701	hypothetical protein
SCO0705	hypothetical protein
SCO0707	branched-chain amino acid ABC transporter permease
SCO0732	protease
SCO0735	oxidoreductase
SCO0740	hydrolase
SCO0777	hypothetical protein
SCO0778	hypothetical protein
SCO0787	hydrolase
SCO0817	hypothetical protein
SCO0842	deoxyribodipyrimidine photolyase
SCO0851	PfkB-family carbohydrate kinase
SCO0853	hypothetical protein
SCO0859	hypothetical protein
SCO0860	cation-transporting ATPase
SCO0863	hypothetical protein
SCO0866	ECF family RNA polymerase sigma factor
SCO0880	hypothetical protein
SCO0883	peptide deformylase
SCO0916	hypothetical protein
SCO0935	hypothetical protein
SCO0938	amino acid transporter
SCO0975	6-phosphogluconate dehydrogenase
SCO0976	hypothetical protein
SCO0982	isocitrate lyase
SCO0983	malate synthase

SCO0989	hypothetical protein
SCO0993	hypothetical protein
SCO0995	methyltransferase
SCO0996	lipoprotein
SCO1010	integral membrane transport protein
SCO1024	hypothetical protein
SCO1030	hypothetical protein
SCO1054	aminotransferase
SCO1068	hypothetical protein
SCO1072	hypothetical protein
SCO1083	flavin-dependent reductase
SCO1112	oxidoreductase
SCO1116	hypothetical protein
SCO1154	hypothetical protein
SCO1156	hypothetical protein
SCO1179	hypothetical protein
SCO1188	cellulose binding protein
SCO1192	hypothetical protein
SCO1206	polyketide synthase
SCO1207	cytochrome P450
SCO1208	hypothetical protein
SCO1212	ligase
SCO1214	6-phosphofructokinase
SCO1253	hypothetical protein
SCO1254	adenylosuccinate lyase
SCO1266	3-oxoacyl-ACP synthase
SCO1272	acyl carrier protein
SCO1289	GntR family transcriptional regulator
SCO1308	4-hydroxybenzoate 3-monooxygenase
SCO1319	hypothetical protein
SCO1327	hypothetical protein
SCO1331	multidomain-containing protein family
SCO1345	3-ketoacyl-ACP reductase
SCO1347	hypothetical protein
SCO1361	hypothetical protein
SCO1369	two component system histidine kinase
SCO1379	hypothetical protein
SCO1394	glycosyl hydrolase
SCO1397	hypothetical protein
SCO1412	hypothetical protein
SCO1419	phosphodiesterase
SCO1420	hypothetical protein
SCO1444	chitinase
SCO1452	hypothetical protein
SCO1456	hypothetical protein
SCO1504	regulator
SCO1510	peptidyl-prolyl cis-trans isomerase
SCO1532	hypothetical protein

SCO1534	DNA polymerase III subunit epsilon
SCO1541	regulator
SCO1553	uroporphyrin-III methyltransferase
SCO1565	glycerophosphoryl diester phosphodiesterase
SCO1566	acyltransferase
SCO1571	hypothetical protein
SCO1609	hypothetical protein
SCO1611	short chain dehydrogenase
SCO1612	aldehyde dehydrogenase
SCO1615	hypothetical protein
SCO1625	ribosomal pseudouridine synthase
SCO1643	20S proteasome alpha-subunit
SCO1646	hypothetical protein
SCO1663	cysteinyl-tRNA synthetase
SCO1664	hypothetical protein
SCO1684	hypothetical protein
SCO1685	hypothetical protein
SCO1689	phosphotransferase
SCO1703	transcriptional regulator
SCO1709	integral membrane transport protein
SCO1720	ABC-transporter transmembrane protein
SCO1724	Ser/Thr protein kinase
SCO1744	two-component system sensor kinase
SCO1745	two-component system response regulator
SCO1774	regulatory protein
SCO1779	hypothetical protein
SCO1781	inorganic polyphosphate/ATP-NAD kinase
SCO1800	hypothetical protein
SCO1824	subtilisin-like protease
SCO1845	low-affinity phosphate transport protein
SCO1853	precorrin-2 C20-methyltransferase
SCO1916	transferase
SCO1927	aminoglycoside acetyltransferase
SCO1929	integral membrane transport protein
SCO1930	integral membrane transport protein
SCO1976	hypothetical protein
SCO1994	hypothetical protein
SCO1996	dephospho-CoA kinase
SCO2008	branched-chain amino acid ABC transporter substrate-binding protein
SCO2010	branched-chain amino acid ABC transporter permease
SCO2011	branched-chain amino acid ABC transporter ATP-binding protein
SCO2022	hypothetical protein
SCO2023	hypothetical protein
SCO2035	hypothetical protein
SCO2046	integral membrane efflux protein
SCO2068	alkaline phosphatase
SCO2076	isoleucyl-tRNA synthetase

SCO2079	hypothetical protein
SCO2086	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase
SCO2093	hypothetical protein
SCO2116	hypothetical protein
SCO2124	hypothetical protein
SCO2133	hypothetical protein
SCO2139	hypothetical protein
SCO2179	leucyl aminopeptidase
SCO2180	dihydrolipoamide dehydrogenase
SCO2181	dihydrolipoamide succinyltransferase
SCO2183	2-oxoacid dehydrogenase subunit E1
SCO2188	peptidase
SCO2190	hypothetical protein
SCO2210	glutamine synthetase
SCO2220	hypothetical protein
SCO2232	maltose operon transcriptional repressor
SCO2263	hypothetical protein
SCO2267	heme oxygenase
SCO2273	FecCD-family membrane transport protein
SCO2287	oxidoreductase
SCO2304	dehydrogenase
SCO2305	ABC transporter ATP-binding protein
SCO2306	ABC transporter
SCO2321	hypothetical protein
SCO2322	hypothetical protein
SCO2323	hypothetical protein
SCO2324	ABC transporter ATP-binding protein
SCO2345	peptidodoglycan-binding membrane protein
SCO2357	hypothetical protein
SCO2372	small hydrophobic protein
SCO2375	hypothetical protein
SCO2376	hypothetical protein
SCO2383	hypothetical protein
SCO2387	ACP S-malonyltransferase
SCO2388	3-oxoacyl-ACP synthase
SCO2407	aldose 1-epimerase
SCO2426	regulatory protein
SCO2448	hypothetical protein
SCO2450	Ser/Thr protein kinase (regulator)
SCO2456	hypothetical protein
SCO2473	nitrate reductase
SCO2486	nitrite reductase NirB
SCO2487	nitrite reductase large subunit NirB
SCO2516	hypothetical protein
SCO2529	metalloprotease
SCO2540	carbohydrate kinase
SCO2545	transmembrane efflux protein
SCO2549	protease

SCO2551	hypothetical protein
SCO2557	hypothetical protein
SCO2573	oxidoreductase
SCO2588	hypothetical protein
SCO2590	glycosyltransferase
SCO2596	50S ribosomal protein L27
SCO2620	trigger factor
SCO2637	serine protease
SCO2663	hypothetical protein
SCO2664	sugar-binding protein
SCO2677	ABC transporter ATP-binding protein
SCO2678	hypothetical protein
SCO2696	2-hydroxyacid dehydrogenase
SCO2703	hypothetical protein
SCO2710	polysaccharide deacetylase
SCO2716	hypothetical protein
SCO2745	LacI family transcriptional regulator
SCO2746	ABC transporter ATP-binding protein
SCO2754	hypothetical protein
SCO2764	lipoprotein
SCO2765	hypothetical protein
SCO2776	acetyl/propionyl CoA carboxylase subunit beta
SCO2782	pyridoxal-dependent decarboxylase
SCO2783	monooxygenase
SCO2794	LacI family transcriptional regulator
SCO2800	two component system histidine kinase
SCO2824	NAD-binding protein
SCO2832	IclR family transcriptional regulator
SCO2836	glycosyl transferase
SCO2840	LysR family transcriptional regulator
SCO2855	hypothetical protein
SCO2856	hypothetical protein
SCO2882	ATP/GTP-binding protein
SCO2883	cytochrome P450
SCO2885	hypothetical protein
SCO2914	amino acid permease
SCO2959	nitrate extrusion protein
SCO3014	translation initiation factor
SCO3015	hypothetical protein
SCO3017	hypothetical protein
SCO3024	transporter
SCO3025	mannose-6-phosphate isomerase
SCO3026	hypothetical protein
SCO3027	hypothetical protein
SCO3041	hypothetical protein
SCO3046	hypothetical protein
SCO3109	transcriptional-repair coupling factor
SCO3125	peptidyl-tRNA hydrolase

SCO3139	sodium:solute symporter
SCO3152	hypothetical protein
SCO3154	hypothetical protein
SCO3155	hypothetical protein
SCO3159	hypothetical protein
SCO3160	integral membrane transport protein
SCO3183	hypothetical protein
SCO3193	hypothetical protein
SCO3207	TetR family transcriptional regulator
SCO3217	transcriptional regulator
SCO3276	integral membrane transporter
SCO3288	hypothetical protein
SCO3289	large membrane protein
SCO3290	hypothetical protein
SCO3291	regulatory protein
SCO3298	oxidoreductase
SCO3306	aminotransferase
SCO3318	porphobilinogen deaminase
SCO3321	redoxin
SCO3336	hydrolase
SCO3347	hypothetical protein
SCO3355	adenine glycosylase
SCO3357	hypothetical protein
SCO3361	transcription regulator AsnC
SCO3365	hypothetical protein
SCO3371	hypothetical protein
SCO3397	integral membrane lysyl-tRNA synthetase
SCO3407	hypothetical protein
SCO3425	30S ribosomal protein S18
SCO3433	hypothetical protein
SCO3444	hypothetical protein
SCO3456	ABC transporter substrate-binding protein
SCO3518	hypothetical protein
SCO3570	hypothetical protein
SCO3578	ion-transporting ATPase
SCO3580	transpeptidase
SCO3592	hypothetical protein
SCO3605	hypothetical protein
SCO3630	hypothetical protein
SCO3658	aminotransferase
SCO3744	hypothetical protein
SCO3747	hypothetical protein
SCO3780	sugar hydrolase
SCO3791	hypothetical protein
SCO3801	aminopeptidase 2
SCO3848	Ser/Thr protein kinase
SCO3850	hypothetical protein
SCO3869	WD-40 repeat-containing protein

SCO3901	penicillin-binding protein
SCO3966	hypothetical protein
SCO3967	hypothetical protein
SCO3973	hypothetical protein
SCO4014	hypothetical protein
SCO4035	RNA polymerase sigma factor
SCO4068	phosphoribosylamine--glycine ligase
SCO4071	phosphoribosylaminoimidazole-succinocarboxamide synthase
SCO4079	phosphoribosylformylglycinamidine synthase II
SCO4085	lipoprotein
SCO4087	phosphoribosylaminoimidazole synthetase
SCO4108	peptidase
SCO4115	hypothetical protein
SCO4118	TetR family transcriptional regulator
SCO4120	hypothetical protein
SCO4160	hydrolase
SCO4199	hypothetical protein
SCO4200	hypothetical protein
SCO4230	response regulator
SCO4237	hypothetical protein
SCO4242	hypothetical protein
SCO4246	hypothetical protein
SCO4250	hypothetical protein
SCO4254	hypothetical protein
SCO4275	histidine autokinase
SCO4276	response regulatory protein
SCO4279	acetyltransferase
SCO4291	hypothetical protein
SCO4294	hypothetical protein
SCO4314	hypothetical protein
SCO4355	nucleotidyltransferase
SCO4362	two component system sensor kinase
SCO4384	enoyl-CoA hydratase
SCO4392	esterase
SCO4399	hypothetical protein
SCO4400	hypothetical protein
SCO4401	lipoprotein
SCO4402	hypothetical protein
SCO4411	calcium binding protein
SCO4444	glutathione peroxidase
SCO4447	hypothetical protein
SCO4450	TetR family transcriptional regulator
SCO4451	export protein
SCO4475	cytochrome biogenesis-like protein
SCO4511	hypothetical protein
SCO4535	hypothetical protein
SCO4536	hypothetical protein
SCO4537	hypothetical protein

SCO4554	bifunctional transferase/deacetylase
SCO4567	NADH dehydrogenase subunit NuoF
SCO4569	NADH dehydrogenase subunit H
SCO4581	hypothetical protein
SCO4592	hypothetical protein
SCO4601	dehydrogenase
SCO4632	ATP/GTP binding protein
SCO4636	hypothetical protein
SCO4637	hypothetical protein
SCO4654	DNA-directed RNA polymerase subunit beta
SCO4655	DNA-directed RNA polymerase subunit beta'
SCO4659	30S ribosomal protein S12
SCO4662	elongation factor Tu
SCO4666	ABC transporter ATP-binding protein
SCO4669	hypothetical protein
SCO4695	hypothetical protein
SCO4729	DNA-directed RNA polymerase subunit alpha
SCO4750	hypothetical protein
SCO4760	hypothetical protein
SCO4766	transcriptional regulator
SCO4774	glycerol phosphate dehydrogenase
SCO4788	ntegral membrane protein
SCO4789	hypothetical protein
SCO4792	two-component system DNA-binding response regulator
SCO4809	succinyl-CoA synthetase subunit alpha
SCO4820	Ser/Thr protein kinase
SCO4824	bifunctional 5,10-methylene-tetrahydrofolate dehydrogenase/ 5,10-methylene-tetrahydrofolate cyclohydrolase
SCO4827	malate dehydrogenase
SCO4843	hypothetical protein
SCO4861	hypothetical protein
SCO4862	hypothetical protein
SCO4882	hypothetical protein
SCO4883	peptidase
SCO4884	lipoprotein
SCO4887	sugar ABC transporter integral membrane protein
SCO4888	sugar ABC transporter integral membrane protein
SCO4890	thymidine phosphorylase
SCO4892	regulatory protein
SCO4909	ATP-binding protein
SCO4910	hypothetical protein
SCO4911	hypothetical protein
SCO4927	ligase
SCO4943	oxidoreductase
SCO4947	nitrate reductase subunit alpha NarG3
SCO4950	nitrate reductase subunit gamma NarI3
SCO4994	hypothetical protein
SCO5006	septum site-determining protein

SCO5007	septum site-determining protein
SCO5010	hypothetical protein
SCO5011	hypothetical protein
SCO5016	hypothetical protein
SCO5039	penicillin-binding protein
SCO5052	hypothetical protein
SCO5064	hypothetical protein
SCO5090	actinorhodin polyketide synthase bifunctional cyclase/dehydratase
SCO5099	hypothetical protein
SCO5109	hypothetical protein
SCO5110	lipoprotein
SCO5112	ABC transporter integral membrane protein BldKA
SCO5114	ABC transporter integral membrane protein BldKC
SCO5115	ABC transporter intracellular ATPase subunit BldKD
SCO5117	peptide transport system peptide-binding protein
SCO5122	peptidase
SCO5123	small membrane protein
SCO5128	hypothetical protein
SCO5149	protease
SCO5151	hypothetical protein
SCO5157	metal-transport protein
SCO5168	hypothetical protein
SCO5188	ATP-dependent DNA helicase
SCO5206	hydrogen peroxide sensitive repressor
SCO5218	hypothetical protein
SCO5219	lipoprotein
SCO5247	deaminase
SCO5251	acetyltransferase
SCO5252	hypothetical protein
SCO5257	methyltransferase
SCO5262	dehydrogenase
SCO5278	magnesium chelatase
SCO5280	ATP-binding protein
SCO5289	two component sensor kinase
SCO5290	hypothetical protein
SCO5291	hypothetical protein
SCO5292	ATP/GTP-binding protein
SCO5293	oxygenase subunit
SCO5315	polyketide cyclase
SCO5316	acyl carrier protein
SCO5336	mutT-like protein
SCO5347	replication initiation protein
SCO5352	arginyl tRNA synthetase
SCO5353	diaminopimelate decarboxylase
SCO5356	homoserine kinase
SCO5387	hypothetical protein
SCO5391	ATP/GTP-binding protein
SCO5392	ABC transporter

SCO5422	hypothetical protein
SCO5444	glycogen phosphorylase
SCO5447	neutral zinc metalloprotease
SCO5455	two-component system response regulator
SCO5465	hypothetical protein
SCO5479	oligopeptide ABC transporter ATP-binding protein
SCO5510	dehydrogenase
SCO5515	D-3-phosphoglycerate dehydrogenase
SCO5522	3-isopropylmalate dehydrogenase
SCO5523	branched-chain amino acid aminotransferase
SCO5529	alpha-isopropylmalate/homocitrate synthase family transferase
SCO5541	ATP-GTP binding protein
SCO5558	acyltransferase
SCO5559	NAD(P)H-dependent glycerol-3-phosphate dehydrogenase
SCO5563	phosphomethylpyrimidine kinase
SCO5582	regulator
SCO5583	ammonium transporter
SCO5584	nitrogen regulatory protein P-II
SCO5585	PII uridylyl-transferase
SCO5594	tRNA (guanine-N(1)-)-methyltransferase
SCO5677	ATP/GTP binding protein
SCO5684	two-component system response regulator
SCO5695	metalloprotease
SCO5699	prolyl-tRNA synthetase
SCO5710	large Pro/Ala/Gly-rich protein
SCO5722	serine protease
SCO5743	FAD-dependent thymidylate synthase
SCO5748	sensory histidine kinase
SCO5759	hypothetical protein
SCO5762	AraC family transcription regulator
SCO5794	kinase/phosphohydrolase
SCO5801	hypothetical protein
SCO5805	vitamin B12-dependent ribonucleotide reductase
SCO5807	hypothetical protein
SCO5808	hypothetical protein
SCO5820	RNA polymerase sigma factor
SCO5822	DNA topoisomerase IV subunit B
SCO5823	hypothetical protein
SCO5827	transmembrane transport protein
SCO5828	two-component transcriptional regulator
SCO5849	AgaS protein
SCO5863	two-component sensor (kinase)
SCO5865	hypothetical protein
SCO5867	hypothetical protein
SCO5868	deoxyuridine 5'-triphosphate nucleotidohydrolase
SCO5879	acyl-coa dehydrogenase RedW
SCO5881	redZ
SCO5897	oxidase

SCO5898	hypothetical protein
SCO5899	hypothetical protein
SCO5913	protease
SCO5936	hypothetical protein
SCO5954	chitinase
SCO5969	hypothetical protein
SCO5980	salicylyl-CoA 5-hydroxylase
SCO5986	oxidoreductase
SCO6007	transmembrane transport protein
SCO6034	hypothetical protein
SCO6035	lysine/ornithine decarboxylase
SCO6037	transmembrane transport protein
SCO6039	flavoprotein oxidoreductase
SCO6041	protoporphyrinogen oxidase
SCO6066	ATP/GTP-binding protein
SCO6073	cyclase
SCO6074	hypothetical protein
SCO6083	hypothetical protein
SCO6109	hydrolase
SCO6116	hypothetical protein
SCO6131	carboxypeptidase
SCO6150	ADA-like regulatory protein
SCO6151	methylated-DNA-protein-cysteine methyltransferase
SCO6159	GntR family transcriptional regulator
SCO6173	permease SC6C509
SCO6180	transferase
SCO6183	transferase
SCO6195	acetyl-coenzyme A synthetase
SCO6201	glyoxylate carboligase
SCO6209	OHCU decarboxylase
SCO6210	hypothetical protein
SCO6247	allantoinase
SCO6248	allantoicase
SCO6254	two-component system response regulator
SCO6296	hypothetical protein
SCO6393	transposase
SCO6394	IS element ATP binding protein
SCO6407	gamma-glutamyltranspeptidase
SCO6416	oxidoreductase
SCO6417	integral membrane transporter
SCO6427	integral membrane transport protein
SCO6436	tRNA synthetase
SCO6438	diaminopimelate decarboxylase
SCO6457	beta-galactosidase
SCO6483	efflux protein
SCO6484	hypothetical protein
SCO6486	transport associated protein
SCO6488	acyl-peptide hydrolase

SCO6520	RNA polymerase sigma factor
SCO6522	hypothetical protein
SCO6531	ATP/GTP binding protein
SCO6540	pterin-4-alpha-carbinolamine dehydratase
SCO6541	hypothetical protein
SCO6542	glycosyl hydrolase
SCO6550	oxidoreductase
SCO6564	3-oxoacyl-ACP synthase
SCO6598	transcriptional regulator
SCO6601	sugar binding protein
SCO6602	transmembrane sugar transport protein
SCO6603	transmembrane sugar transport protein
SCO6619	dehydrogenase
SCO6621	hypothetical protein
SCO6626	protein kinase
SCO6627	hypothetical protein
SCO6632	hypothetical protein
SCO6635	bacteriophage resistance gene pglY
SCO6644	solute-binding lipoprotein
SCO6645	transport system permease
SCO6646	transport system permease
SCO6651	glycosyl transferase
SCO6652	hypothetical protein
SCO6653	hypothetical protein
SCO6681	Ser/Thr protein kinase
SCO6685	ramR
SCO6691	phospholipase C
SCO6696	regulatory protein
SCO6697	3-oxoadipate enol-lactone hydrolase/4-carboxymuconolactone decarboxylase
SCO6711	hypothetical protein
SCO6712	copper oxidase
SCO6713	transcriptional regulator
SCO6715	transcriptional regulator
SCO6721	hypothetical protein
SCO6723	oxidoreductase
SCO6731	acetyl-CoA acetyltransferase
SCO6734	L-asparagine permease
SCO6739	BCCT family transporter
SCO6742	ABC transporter ATP-binding protein
SCO6754	glycerol dehydrogenase
SCO6759	phytoene synthase
SCO6760	phytoene synthase
SCO6761	hypothetical protein
SCO6762	phytoene dehydrogenase
SCO6764	squalene-hopene cyclase
SCO6776	hypothetical protein
SCO6780	hypothetical protein

SCO6783	hypothetical protein
SCO6786	oxidoreductase
SCO6795	hypothetical protein
SCO6796	hypothetical protein
SCO6828	hypothetical protein
SCO6838	monooxygenase
SCO6855	hypothetical protein
SCO6904	hypothetical protein
SCO6905	hypothetical protein
SCO6907	DNA ligase
SCO6915	hypothetical protein
SCO6971	precorrin 6A synthase
SCO6987	hypothetical protein
SCO7004	carbohydrate kinase
SCO7020	alpha-amylase
SCO7021	hypothetical protein
SCO7022	hypothetical protein
SCO7030	binding-protein-dependent transport protein
SCO7031	beta-D-xylosidase
SCO7072	hypothetical protein
SCO7076	two-component histidine kinase
SCO7086	MerR family transcriptional regulator
SCO7093	transcriptional regulator
SCO7104	RNA polymerase sigma factor
SCO7142	hypothetical protein
SCO7143	transcriptional regulator
SCO7203	hypothetical protein
SCO7232	hypothetical protein
SCO7233	hypothetical protein
SCO7239	hypothetical protein
SCO7252	regulatory protein
SCO7293	hypothetical protein
SCO7298	thioredoxin reductase
SCO7299	stress-inducible protein
SCO7304	FAD-binding dehydrogenase
SCO7309	membrane transport protein
SCO7310	regulatory protein
SCO7326	hypothetical protein
SCO7337	hypothetical protein
SCO7338	glycogen debranching protein
SCO7360	hypothetical protein
SCO7368	hypothetical protein
SCO7377	hypothetical protein
SCO7413	short chain dehydrogenase
SCO7417	cytochrome P450-family protein
SCO7443	phosphoglucomutase
SCO7444	cytochrome P450 (fragment)
SCO7448	hypothetical protein

SCO7467	hypothetical protein
SCO7471	phenylacetate-CoA oxygenase subunit PaaA
SCO7506	hydrolase
SCO7539	TetR family transcriptional regulator
SCO7547	sulfatase
SCO7558	beta-glucosidase
SCO7563	ABC transporter solute binding lipoprotein
SCO7568	regulatory protein
SCO7580	hypothetical protein
SCO7582	hypothetical protein
SCO7605	metallopeptidase
SCO7610	transcriptional regulator
SCO7627	hypothetical protein
SCO7628	hypothetical protein
SCO7641	dehydrogenase
SCO7657	hypothetical protein
SCO7659	oxidoreductase
SCO7660	voltage-gated potassium channel
SCO7677	solute-binding protein
SCO7678	metal transport integral membrane protein
SCO7682	non-ribosomal peptide synthase
SCO7683	non-ribosomal peptide synthase
SCO7684	hypothetical protein
SCO7685	hypothetical protein
SCO7688	hypothetical protein
SCO7689	ABC transporter ATP-binding protein
SCO7697	hydrolase
SCO7702	GntR family transcriptional regulator
SCO7703	integral membrane transport protein
SCO7711	two component system sensor kinase
SCO7715	hypothetical protein
SCO7721	hypothetical protein
SCO7730	hypothetical protein
SCO7740	hypothetical protein
SCO7773	hypothetical protein
SCO7778	transcriptional regulator
SCO7798	transposase
SCO7800	hypothetical protein
SCO7801	hypothetical protein

Table 17 List of genes methylated in MII phase in *S. coelicolor* grown on solid GYM grouped for the methylation GCC^{m}CG consensus sequence.

Name	Product
SCO0026	hypothetical protein
SCO0055	membrane-associated oxidoreductase
SCO0096	noncomposite transposon transposase
SCO0178	hypothetical protein
SCO0196	hypothetical protein
SCO0211	hypothetical protein
SCO0217	nitrate reductase subunit beta NarH2
SCO0235	short chain dehydrogenase
SCO0244	hypothetical protein
SCO0267	hydrolase
SCO0305	hypothetical protein
SCO0335	hypothetical protein
SCO0352	solute-binding protein
SCO0375	integral membrane transport protein
SCO0376	transcriptional regulator
SCO0381	glycosyl transferase
SCO0382	UDP-glucose/GDP-mannose dehydrogenase
SCO0389	lipoprotein
SCO0449	solute-binding lipoprotein
SCO0544	hypothetical protein
SCO0545	hypothetical protein
SCO0608	regulatory protein
SCO0676	integral membrane sensor protein
SCO0718	hypothetical protein
SCO0753	hypothetical protein
SCO0790	hydrolase
SCO0888	hypothetical protein
SCO1058	sugar transport integral membrane protein
SCO1084	thioredoxin
SCO1085	acyltransferase
SCO1112	oxidoreductase
SCO1114	uracil-DNA glycosylase
SCO1115	hypothetical protein
SCO1120	hydrolase
SCO1121	hypothetical protein
SCO1122	regulatory protein
SCO1123	hypothetical protein
SCO1174	aldehyde dehydrogenase
SCO1233	urease accessroy protein UreF
SCO1244	biotin synthase
SCO1254	adenylosuccinate lyase
SCO1421	hypothetical protein
SCO1485	hypothetical protein

SCO1489	bldD
SCO1504	regulator
SCO1529	hypothetical protein
SCO1774	regulatory protein
SCO1793	hypothetical protein
SCO1794	hypothetical protein
SCO1797	hypothetical protein
SCO1817	hypothetical protein
SCO1945	triosephosphate isomerase
SCO1968	hydrolase
SCO1980	hypothetical protein
SCO1981	hypothetical protein
SCO2015	nucleotidase
SCO2077	hypothetical protein
SCO2083	sporulation protein
SCO2116	hypothetical protein
SCO2232	maltose operon transcriptional repressor
SCO2234	bifunctional glutamine-synthetase adenylyltransferase/deadenyltransferase
SCO2235	hypothetical protein
SCO2236	hypothetical protein
SCO2241	glutamine synthetase
SCO2251	hypothetical protein
SCO2255	hypothetical protein
SCO2266	methionine aminopeptidase
SCO2275	lipoprotein
SCO2276	hypothetical protein
SCO2277	hypothetical protein
SCO2313	hypothetical protein
SCO2385	hypothetical protein
SCO2386	hypothetical protein
SCO2387	ACP S-malonyltransferase
SCO2396	hypothetical protein
SCO2404	sugar-binding receptor
SCO2405	sugar-transport ATP binding protein
SCO2487	nitrite reductase large subunit NirB
SCO2488	nitrite reductase small subunit NirC
SCO2491	oxidoreductase
SCO2492	hypothetical protein
SCO2503	chitinase
SCO2516	hypothetical protein
SCO2621	hypothetical protein
SCO2622	hypothetical protein
SCO2627	ribose-5-phosphate isomerase B
SCO2628	amino acid permease
SCO2661	sugar hydrolase
SCO2681	ATP /GTP-binding protein
SCO2733	hypothetical protein

SCO2734	LysR family transcriptional regulator
SCO2780	hypothetical protein
SCO2783	monooxygenase
SCO2784	aceytltranferase
SCO2865	regulatory protein
SCO2934	hypothetical protein
SCO2967	carboxy-terminal processing protease
SCO3015	hypothetical protein
SCO3043	hypothetical protein
SCO3045	hypothetical protein
SCO3046	hypothetical protein
SCO3065	hypothetical protein
SCO3093	hydrolase
SCO3203	phosphinothricin acetyltransferase
SCO3289	large membrane protein
SCO3355	adenine glycosylase
SCO3404	cell division protein FtsH-like protein
SCO3490	transposase
SCO3543	DNA topoisomerase I
SCO3580	transpeptidase
SCO3702	DNA-binding protein
SCO3747	hypothetical protein
SCO3791	hypothetical protein
SCO3801	aminopeptidase 2
SCO3835	dehydrogenase
SCO3869	WD-40 repeat-containing protein
SCO3896	RNA nucleotidyltransferase
SCO3906	30S ribosomal protein S6
SCO4064	hypothetical protein
SCO4068	phosphoribosylamine--glycine ligase
SCO4085	lipoprotein
SCO4109	oxidoreductase
SCO4175	hypothetical protein
SCO4176	hypothetical protein
SCO4193	ATP/GTP-binding membrane protein
SCO4199	hypothetical protein
SCO4200	hypothetical protein
SCO4254	hypothetical protein
SCO4257	hydrolytic protein
SCO4258	hydrolytic protein
SCO4280	reductase
SCO4441	DNA-binding protein
SCO4551	hypothetical protein
SCO4561	hypothetical protein
SCO4568	NADH dehydrogenase subunit G
SCO4592	hypothetical protein
SCO4655	DNA-directed RNA polymerase subunit beta'
SCO4693	hypothetical protein

SCO4716	30S ribosomal protein S8
SCO4717	50S ribosomal protein L6
SCO4724	methionine aminopeptidase
SCO4734	50S ribosomal protein L13
SCO4756	hypothetical protein
SCO4773	nucleotide-sugar dehydrogenase
SCO4774	glycerol phosphate dehydrogenase
SCO4890	thymidine phosphorylase
SCO4901	adenosine deaminase
SCO4902	hypothetical protein
SCO4934	lipoprotein
SCO5018	hypothetical protein
SCO5074	dehydratase
SCO5075	oxidoreductase
SCO5078	hypothetical protein
SCO5085	actii-ORF4
SCO5100	GntR family transcriptional regulator
SCO5110	lipoprotein
SCO5111	GTP-binding protein
SCO5112	ABC transporter integral membrane protein BldKA
SCO5113	ABC transporter lipoprotein BldKB
SCO5115	ABC transporter intracellular ATPase subunit BldKD
SCO5117	peptide transport system peptide-binding protein
SCO5128	hypothetical protein
SCO5129	ABC transporter ATP-binding protein
SCO5130	ABC transporter
SCO5149	protease
SCO5150	sec-independent translocase
SCO5151	hypothetical protein
SCO5232	sugar transporter sugar binding protein
SCO5290	hypothetical protein
SCO5491	hypothetical protein
SCO5498	aspartyl/glutamyl-tRNA amidotransferase subunit C
SCO5559	NAD(P)H-dependent glycerol-3-phosphate dehydrogenase
SCO5560	D-alanyl-alanine synthetase A
SCO5580	docking protein
SCO5614	transcriptional regulator
SCO5624	30S ribosomal protein S2
SCO5625	elongation factor Ts
SCO5626	uridylate kinase
SCO5629	ATP /GTP-binding protein
SCO5645	ribosomal RNA large subunit methyltransferase N
SCO5693	acyl CoA dehydrogenase
SCO5696	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase
SCO5718	hypothetical protein
SCO5750	ftsK-like protein
SCO5751	hypothetical protein
SCO5787	(dimethylallyl)adenosine tRNA methylthiotransferase

SCO5847	transcriptional regulator
SCO5897	oxidase
SCO5898	hypothetical protein
SCO5951	transcriptional regulator
SCO5971	hypothetical protein
SCO6033	hypothetical protein
SCO6034	hypothetical protein
SCO6074	hypothetical protein
SCO6109	hydrolase
SCO6178	deacetylase
SCO6179	nucleotide-sugar dehydratase
SCO6195	acetyl-coenzyme A synthetase
SCO6199	esterase
SCO6204	catalase
SCO6360	ABC transporter ATP-binding protein
SCO6384	integral membrane lysyl-tRNA synthetase
SCO6390	hypothetical protein
SCO6393	transposase
SCO6394	IS element ATP binding protein
SCO6407	gamma-glutamyltranspeptidase
SCO6411	hydrolase
SCO6416	oxidoreductase
SCO6439	DNA-binding protein
SCO6440	hypothetical protein
SCO6451	substrate binding protein
SCO6498	hypothetical protein
SCO6601	sugar binding protein
SCO6626	protein kinase
SCO6635	bacteriophage resistance gene pglY
SCO6640	ATP-dependent helicase
SCO6644	solute-binding lipoprotein
SCO6691	phospholipase C
SCO6731	acetyl-CoA acetyltransferase
SCO6732	fatty acid oxidative multifunctional enzyme
SCO6773	peptidase
SCO6788	acetyl-CoA acetyltransferase
SCO6903	hypothetical protein
SCO6904	hypothetical protein
SCO7008	ABC transporter ATP-binding protein
SCO7022	hypothetical protein
SCO7023	hypothetical protein
SCO7059	oxidoreductase
SCO7093	transcriptional regulator
SCO7131	lipase
SCO7159	hypothetical protein
SCO7252	regulatory protein
SCO7279	DNA-binding protein
SCO7348	hypothetical protein

SCO7472	phenylacetate-CoA oxygenase subunit PaaB
SCO7473	phenylacetic acid degradation protein PaaC
SCO7506	hydrolase
SCO7657	hypothetical protein
SCO7674	metal-binding protein
SCO7683	non-ribosomal peptide synthase
SCO7684	hypothetical protein
SCO7722	hypothetical protein
SCO7733	transcriptional regulator
SCO7734	hypothetical protein

Table 18 List of genes methylated in MII phase in *S. coelicolor* grown on solid GYM grouped for the methylation $C^{m}GGGC$ consensus sequence.

Name	Product
SCO0069	hypothetical protein
SCO0122	flavin-containing monooxygenase
SCO0138	short chain dehydrogenase
SCO0191	lycopene cyclase
SCO0240	oxidoreductase
SCO0366	hypothetical protein
SCO0403	hypothetical protein
SCO0404	2-haloalkanoic acid dehalogenase
SCO0502	hypothetical protein
SCO0530	transcriptional regulator
SCO0531	sugar transporter sugar binding protein
SCO0537	hypothetical protein
SCO0538	sugar transporter sugar binding lipoprotein
SCO0654	hypothetical protein
SCO0655	gas vesicle synthesis protein
SCO0692	hypothetical protein
SCO0693	hypothetical protein
SCO0723	fructose transport system kinase
SCO0829	serine protease
SCO0894	hypothetical protein
SCO0895	RNA polymerase principal sigma factor HrdC
SCO0951	transport system permease
SCO0986	hypothetical protein
SCO0987	hypothetical protein
SCO1007	oxidoreductase
SCO1036	phosphotriesterase-family protein
SCO1037	hypothetical protein
SCO1051	hypothetical protein
SCO1170	xylulose kinase
SCO1194	export protein
SCO1223	ornithine aminotransferase
SCO1227	DNA-binding protein
SCO1228	acyltransferase
SCO1250	acetyltransferase
SCO1291	hypothetical protein
SCO1446	hypothetical protein
SCO1448	transporter
SCO1463	transcriptional regulator
SCO1519	Holliday junction DNA helicase RuvA
SCO1579	bifunctional ornithine acetyltransferase/N-acetylglutamate synthase
SCO1617	hypothetical protein
SCO1654	two-component response regulator
SCO1686	NTP pyrophosphohydrolase

SCO1849	cobaltochelatase subunit CobN
SCO1867	hydroxylase
SCO1875	penicillin binding protein
SCO1910	alanine-rich protein
SCO1937	glucose-6-phosphate 1-dehydrogenase
SCO2004	formate dehydrogenase
SCO2024	chitosanase
SCO2028	hypothetical protein
SCO2062	hypothetical protein
SCO2110	Ser/Thr protein kinase
SCO2111	endonuclease
SCO2170	methyltransferase
SCO2187	hypothetical protein
SCO2194	lipoyl synthase
SCO2198	glutamine synthetase
SCO2215	two-component system sensor kinase
SCO2261	hypothetical protein
SCO2264	hypothetical protein
SCO2300	hypothetical protein
SCO2316	hypothetical protein
SCO2336	integral membrane transport protein
SCO2398	MarR family transcriptonal regulator
SCO2428	phosphate binding protein
SCO2433	sugar transporter membrane protein
SCO2451	rod shape-determining protein MreB
SCO2452	two-component sensor histidine kinase
SCO2477	short chain dehydrogenase
SCO2495	hypothetical protein
SCO2553	oxidoreductase
SCO2571	leucyl-tRNA synthetase
SCO2591	hypothetical protein
SCO2636	hypothetical protein
SCO2639	RNA polymerase sigma factor
SCO2640	aspartate-semialdehyde dehydrogenase
SCO2730	regulator
SCO2731	cation-transporting P-type ATPase
SCO2779	acyl-CoA dehydrogenase
SCO2815	TetR family transcriptional regulator
SCO2816	hypothetical protein
SCO2829	amino acid ABC transporter transmembrane protein
SCO2849	hypothetical protein
SCO2876	acetyltransferase
SCO3006	acetyltransferase
SCO3007	hypothetical protein
SCO3022	hypothetical protein
SCO3036	2-phospho-L-lactate transferase
SCO3061	hypothetical protein
SCO3067	anti anti sigma factor

SCO3068	RNA polymerase sigma factor
SCO3140	hypothetical protein
SCO3142	hypothetical protein
SCO3158	hypothetical protein
SCO3168	protease
SCO3230	CDA peptide synthetase I
SCO3236	oxygenase
SCO3319	glutamyl-tRNA reductase
SCO3344	Ser/Thr protein kinase
SCO3378	small membrane protein
SCO3383	pantoate--beta-alanine ligase
SCO3391	hypothetical protein
SCO3439	hypothetical protein
SCO3440	hypothetical protein
SCO3509	hypothetical protein
SCO3511	lipoprotein
SCO3512	hypothetical protein
SCO3513	hypothetical protein
SCO3540	proteinase
SCO3562	integral membrane transport protein
SCO3650	orotate phosphoribosyltransferase
SCO3692	anti-sigma factor antagonist
SCO3723	regulatory protein
SCO3743	hypothetical protein
SCO3774	beta-lactamase
SCO3792	methionyl-tRNA synthetase
SCO3888	hypothetical protein
SCO3892	RNA polymerase sigma factor SigM
SCO3895	hypothetical protein
SCO3910	hypothetical protein
SCO3911	replicative DNA helicase
SCO3931	hypothetical protein
SCO3947	ABC transporter
SCO3981	GntR family transcriptional regulator
SCO3983	hypothetical protein
SCO4005	RNA polymerase sigma factor
SCO4015	hypothetical protein
SCO4016	hypothetical protein
SCO4027	anti sigma factor antagonist
SCO4028	hypothetical protein
SCO4052	glucose 1-dehydrogenase
SCO4158	LacI-family regulatory protein
SCO4249	hypothetical protein
SCO4296	chaperonin GroEL
SCO4345	hypothetical protein
SCO4381	acetyl/propionyl CoA carboxylase subunit alpha
SCO4423	Ser/Thr protein kinase
SCO4520	hypothetical protein

SCO4531	septum determining protein
SCO4532	hypothetical protein
SCO4566	NADH dehydrogenase subunit E
SCO4606	NADH dehydrogenase subunit NuoL2
SCO4607	NADH dehydrogenase subunit NuoM2
SCO4643	UDP-N-acetylenolpyruvoylglucosamine reductase
SCO4705	50S ribosomal protein L2
SCO4733	hypothetical protein
SCO4804	hypothetical protein
SCO4805	hypothetical protein
SCO4848	hypothetical protein
SCO4849	hypothetical protein
SCO4935	hypothetical protein
SCO4940	TetR family transcriptional regulator
SCO4964	integral membrane transport protein
SCO4972	dehydrogenase
SCO5000	hypothetical protein
SCO5097	short-chain oxidoreductase
SCO5194	hypothetical protein
SCO5321	polyketide hydroxylase
SCO5372	F0F1 ATP synthase subunit gamma
SCO5418	transcriptional regulator
SCO5506	regulatory protein
SCO5526	urease subunit alpha
SCO5527	hypothetical protein
SCO5557	hypothetical protein
SCO5672	hypothetical protein
SCO5673	chitinase
SCO5709	tRNA pseudouridine synthase B
SCO5760	DNA glycosylase
SCO5797	hypothetical protein
SCO5817	DNA hydrolase
SCO5832	citrate synthase
SCO5848	tagatose 6-phosphate kinase
SCO6095	ABC transporter ATP-binding protein
SCO6120	hypothetical protein
SCO6188	transferase
SCO6251	reductase
SCO6262	helicase
SCO6376	hypothetical protein
SCO6432	peptide synthase
SCO6570	oxidoreductase
SCO6572	glycosyl hydrolase
SCO6581	transmembrane transport protein
SCO6585	succinyl-CoA synthetase subunit beta
SCO6660	hypothetical protein
SCO6735	hypothetical protein
SCO6790	long chain fatty acid CoA ligase

SCO6799	L-threonine 3-dehydrogenase
SCO6837	arsenic resistance membrane transport protein
SCO6883	hypothetical protein
SCO6956	monovalent cation/H+ antiporter subunit D
SCO6957	monovalent cation/H+ antiporter subunit E
SCO6969	hypothetical protein
SCO7018	hypothetical protein
SCO7064	hypothetical protein
SCO7101	dehydrogenase
SCO7183	branched amino acid transport system permease
SCO7197	amino acid ABC transporter permease
SCO7202	hypothetical protein
SCO7318	hypothetical protein
SCO7384	transmembrane transport protein
SCO7504	integral membrane binding-protein-dependent transport protein
SCO7536	hypothetical protein
SCO7679	transport system integral membrane protein
SCO7680	ABC transporter ATP-binding protein
SCO7756	hypothetical protein
SCO7761	hypothetical protein
SCO7762	hypothetical protein
SCO7805	hypothetical protein
SCO7811	hypothetical protein

Figures and Tables

List of Abbreviations

Act	actinorhodin
aza-dC	5-aza-2’-deoxycytidine
bp	base pairs
BS	Bisulfite
CDA	calcium-dependet lipopetide
Chp	chaplin
CLSM	confocal laser scanning microSCOpy
CpK	cryptic polyketide
DMSO	dimethyl sulfoxide
DNMT1	DNA-(cytosine-5)-methylatrasferase 1
EGFP	green fluorescent protein
m4C	N^4-methylcytosine
m5C	C^5-methylcytosine
m6A	N^6-methyladenine
MI	first compartimetalyzed mycelium
MII	second multinucleated mycelium
PCD	programmed cell death
PCR	polymerase chain reaction
Rdl	rodlin
Red	undecylprodigiosin
RG1	first rapid growth
RG2	second rapid growth
RM	Restriction-Modification system
SMRT	single-molecule-real-time
Taq	DNA polymerase
WT	wild type