



# Time-Resolved Transcriptomics and Constraint-Based Modeling Identify System-Level Metabolic Features and Overexpression Targets to Increase Spiramycin Production in *Streptomyces ambofaciens*

## OPEN ACCESS

### Edited by:

Dimitris G. Hatzinikolaou,  
National and Kapodistrian University  
of Athens, Greece

### Reviewed by:

Georgios Skretas,  
National Hellenic Research  
Foundation, Greece  
Rosa Alduina,  
Università di Palermo, Italy

### \*Correspondence:

Clelia Peano  
clelia.peano@itb.cnr.it  
Pietro Alifano  
pietro.alifano@unisalento.it

† These authors have contributed  
equally to this publication and should  
be considered co-first authors.

### Specialty section:

This article was submitted to  
Systems Microbiology,  
a section of the journal  
Frontiers in Microbiology

Received: 18 November 2016

Accepted: 24 April 2017

Published: 12 May 2017

### Citation:

Fondi M, PinateL E, Talà A,  
Damiano F, Consolandi C,  
Matorre B, Fico D, Testini M,  
De Benedetto GE, Siculella L,  
De Bellis G, Alifano P and Peano C  
(2017) Time-Resolved  
Transcriptomics  
and Constraint-Based Modeling  
Identify System-Level Metabolic  
Features and Overexpression Targets  
to Increase Spiramycin Production  
in *Streptomyces ambofaciens*.  
*Front. Microbiol.* 8:835.  
doi: 10.3389/fmicb.2017.00835

Marco Fondi<sup>1†</sup>, Eva PinateL<sup>2†</sup>, Adelfia Talà<sup>3†</sup>, Fabrizio Damiano<sup>3</sup>, Clarissa Consolandi<sup>2</sup>,  
Benedetta Matorre<sup>1</sup>, Daniela Fico<sup>4</sup>, Mariangela Testini<sup>3</sup>, Giuseppe E. De Benedetto<sup>4</sup>,  
Luisa Siculella<sup>3</sup>, Gianluca De Bellis<sup>2</sup>, Pietro Alifano<sup>3\*</sup> and Clelia Peano<sup>2\*</sup>

<sup>1</sup> Department of Biology, University of Florence, Florence, Italy, <sup>2</sup> Institute of Biomedical Technologies, National Research Council, Segrate, Italy, <sup>3</sup> Department of Biological and Environmental Sciences and Technologies, University of Salento, Lecce, Italy, <sup>4</sup> Laboratory of Analytical and Isotopic Mass Spectrometry, Department of Cultural Heritage, University of Salento, Lecce, Italy

In this study we have applied an integrated system biology approach to characterize the metabolic landscape of *Streptomyces ambofaciens* and to identify a list of potential metabolic engineering targets for the overproduction of the secondary metabolites in this microorganism. We focused on an often overlooked growth period (i.e., post-first rapid growth phase) and, by integrating constraint-based metabolic modeling with time resolved RNA-seq data, we depicted the main effects of changes in gene expression on the overall metabolic reprogramming occurring in *S. ambofaciens*. Moreover, through metabolic modeling, we unraveled a set of candidate overexpression gene targets hypothetically leading to spiramycin overproduction. Model predictions were experimentally validated by genetic manipulation of the recently described ethylmalonyl-CoA metabolic node, providing evidence that spiramycin productivity may be increased by enhancing the carbon flow through this pathway. The goal was achieved by over-expressing the *ccr* paralog *sm4* in an *ad hoc* engineered plasmid. This work embeds the first metabolic reconstruction of *S. ambofaciens* and the successful experimental validation of model predictions and demonstrates the validity and the importance of *in silico* modeling tools for the overproduction of molecules with a biotechnological interest. Finally, the proposed metabolic reconstruction, which includes manually refined pathways for several secondary metabolites with antimicrobial activity, represents a solid platform for the future exploitation of *S. ambofaciens* biotechnological potential.

**Keywords:** *Streptomyces ambofaciens*, antibiotic production, strain improvement, metabolic modeling, transcriptomics, systems biology

## INTRODUCTION

Constraint-based modeling is a widely adopted technique to study microbial metabolic features at the system level. Indeed, it has proven useful for addressing fundamental issues about the metabolic landscape of a variety of microorganisms, including the elucidation of non-trivial metabolic engineering strategies (King et al., 2015; Zhang and Hua, 2016). Genome-scale metabolic networks are formulated and reconstructed using genomic information of the species to be modeled. Next, computational methods (e.g., Flux Balance Analysis, FBA) can be adopted to infer the flux distributions within the cell (Orth et al., 2010), and, even more importantly, to study the metabolic reprogramming following environmental perturbations and/or changes in any of the possible cellular information layers (e.g., gene expression). Many fields can benefit from this technology including clinical microbiology (Sigurdsson et al., 2012), metabolic engineering (Licon-Cassani et al., 2012), environmental microbiology (Fondi et al., 2015), and microbial ecology (Stolyar et al., 2007). A relevant (recent) achievement in the field of constraint-based modeling is the possibility to integrate gene expression data into the metabolic framework. Indeed, by constraining the admissible flux across each reaction in the model on the basis of the corresponding genes' expression level, metabolic flux predictions can be improved or context-specific models created. Despite the fact that a recent work showed that no method performs universally well, the integration of transcriptomics data and constraint-based metabolic modeling might still provide cues to guide the determination of the correct phenotype among the space of solutions (Machado and Herrgard, 2014).

*Streptomyces* representatives' metabolism has been extensively studied (Baltz, 2016), in many cases using genome-scale metabolic reconstruction and constraint-based modeling (Borodina et al., 2005; D'Huys et al., 2012; Huang et al., 2012, 2013). In particular, this computational approach has disclosed potential metabolic engineering strategies (e.g., overexpression and knock-out targets) for daptomycin and FK506 overproduction in *S. roseosporus* and *S. tsukubaensis*, respectively. Furthermore, a comprehensive and system-level understanding of the metabolic landscape of *S. lividans* and *S. coelicolor* has been depicted using this methodology, to predict flux changes that occur when the cell switches from biomass to antibiotic production (Alam et al., 2010). Interestingly, the potential of integrating gene expression data and metabolic modeling for exploring *Streptomyces* biology (and biotechnological potential) has been assessed in a recent work focused on *S. coelicolor* (Kim et al., 2016). Besides benchmarking currently available tools for transcriptomic data integration into metabolic reconstructions and identifying the outperforming tool among them (iMAT, Shlomi et al., 2008), this study has identified a list of potential metabolic engineering targets for the overproduction of the actinorhodin secondary metabolite in this and other streptomycetes.

Our study offers a system biology approach to explore, from a global point of view, the metabolic landscape of *Streptomyces ambofaciens*, a prolific producer of bioactive compounds. The

reference strain *S. ambofaciens* ATCC 23877 (Pinnert-Sindico, 1954), whose genome sequence has been recently delivered (Thibessard et al., 2015), has been known for the last 60 years due to its ability to produce a wide range of secondary metabolites including spiramycin, a macrolide used in human medicine as antibacterial and antiparasitic agents (active against *Toxoplasma* spp.) (Poulet et al., 2005; Chew et al., 2012) and congocidine (netropsin), a pyrrolamide with a broad range of biological activities but no medical applications (Cosar et al., 1952). More recently, a genome mining-guided approach has revealed the ability of this strain to produce, in addition to spiramycin and congocidine, the antibiotic kinamycin, the siderophores coelichelin and desferioxamines, the antifungals antimycin and stambomycins, and novel polyketides with antibacterial and antiproliferative activities (Thibessard et al., 2015).

Here we have applied an approach that combines physiological, transcriptomic and *in silico* modeling with the aim to provide a system-level understanding of *S. ambofaciens* metabolic features, and to identify a list of potential metabolic engineering targets for the overproduction of the secondary metabolites in this microorganism. Besides including the first metabolic reconstruction of *S. ambofaciens*, the experimental validation of model predictions presented herein underlines the power of *in silico* modeling in the context of metabolic engineering.

## MATERIALS AND METHODS

### Bacterial Strains, Media and Growth Conditions

*Streptomyces ambofaciens* ATCC 23877 was obtained from the American Type Culture Collection (ATCC). The identity of this strain was verified by sequence analysis using the 16S rRNA genes specific 'universal' primers 5'-CAGCAGCCGCGTAATAC-3' and 5'-CCGTCAATTCCTTTGAGTTT-3' (Olsen et al., 1986). The strain was stored in 1 ml cryotubes at  $-80^{\circ}\text{C}$  as frozen mycelium in yeast starch (YS) medium containing 15% glycerol at a biomass concentration of approximately 0.25 g dry cell weight (DCW)/ml, or at  $-20^{\circ}\text{C}$  as spores in 20% glycerol at a titre of approximately  $5 \times 10^8$ /ml. The composition (per liter) of the media used in this study for *S. ambofaciens* growth and manipulation is here reported. YS broth: 2 g yeast extract, 10 g soluble starch, (18 g agar in YS agar) (pH 7.0); SMII (medium for conjugation experiments): 15 g dextrose, 10 g soybean flour, 0.5 g  $\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$ , 5 g  $\text{CaCO}_3$ , 15 g agar; 2XYT (medium for spore resuspension): 16 g tryptone, 10 g yeast extract, 5 g NaCl.

*Escherichia coli* strain DH5 $\alpha$  [ $\text{F}^-$   $\Phi$ 80d *lacZ* $\Delta$ M15 *endA1* *recA1* *hsdR17* *supE44* *thi-1*  $\lambda^-$  *gyrA96*  $\lambda$  (*lacZYA-argF*) U169] was used in cloning procedures. This strain was cultivated in Luria-Bertani broth supplemented with 50  $\mu\text{g}/\text{ml}$  ampicillin when required. Conjugation experiments were performed with *E. coli* strain GM2929/pUB307::Tn7. Lennox Broth (LeB) was used to grow, wash and resuspend this strain. The composition (per liter) of the LeB is as follows: 10 g tryptone, 5 g yeast extract, 5 g NaCl, 1 g glucose.

## DNA Procedures and Construction of Recombinant *S. ambofaciens* Strains

*Streptomyces ambofaciens* total genomic DNA was obtained as described by (Kieser et al., 2000). Manipulation of *E. coli* DNA was performed as indicated by Sambrook and Russell (2001). PCR amplification conditions were 94°C for 5 min followed by 34 cycles of 94°C for 1 min, 65°C for 1 min and 72°C for 1 min, and finally 72°C for 10 min.

Overexpression of *S. ambofaciens ccr* (*srm4*) (SAM23877\_RS26790 in *S. ambofaciens* ATCC 23877 genome) coding for crotonyl-CoA carboxylase/reductase was achieved by cloning the PCR-amplified coding region into an *ad hoc* engineered expression vector, pNGEM/OriT/P, under the control of *groES* promoter. To construct pNGEM/OriT/P we utilized the *E. coli*-*Streptomyces* shuttle vector pN702GEM3. pN702GEM3 (Fernandez-Abalos et al., 2003) contains the pIJ702 origin of replication (for Streptomyces), the bifunctional *Neo/Kan* resistance marker from Tn5 (for both *E. coli* and Streptomyces), the pUC origin of replication (for *E. coli*) and a polylinker derived from pGEM3Zf(+) (Promega). This vector is replicative but not conjugative in Streptomyces. To transform it into conjugative vector pNGEM/OriT a DNA fragment containing the *oriT* determinant was amplified by PCR from plasmid pTYM18 (Onaka et al., 2003) using the primer pair OriTHindIII-for and OriTHindIII-rev whose sequence is reported below (*HindIII* restriction sites are underlined):

OriTHindIII-for: 5'-CCGACCAAGCTTCGCCCAACCTGCC  
ATCAGGAGATTTC-3'  
OriTHindIII-rev: 5'-GAGCTGAAAGCTTCAGAAGCCACT  
GGAGCACCTC-3'

The amplified fragment was treated with *HindIII*, and the resulting 930 bp DNA fragment was cloned into pN702GEM3 using the *HindIII* restriction site of the polylinker.

To construct the pNGEM/OriT/P plasmid, the *groES* gene promoter (*P<sub>groES</sub>*) region was amplified by PCR using *S. ambofaciens* ATCC 23877 genomic DNA as template, and the primer pair *groES*-for and *groES*-rev (*XbaI* and *NdeI* restriction sites are underlined):

*groES*-for: 5'-GGCCGTTCTAGACCGTCCGGCGTTTCGAG  
GACGAGG-3'  
*groES*-rev: 5'-GCTGGTTCTAGACATATGCCGACCTCCCC  
CTTCGGAGATCTCACG-3'

The amplified DNA fragment was digested with *XbaI*, and the resulting 325 bp DNA fragment was inserted into the polylinker of pNGEM/OriT, generating pNGEM/OriT/P.

The crotonyl-CoA carboxylase/reductase (CRR) encoding gene *srm4* (locus tag: SAM23877\_RS26790) was amplified from *S. ambofaciens* ATCC 23877 genomic DNA, using the primer pair CCR-for and CCR-rev (*NdeI* and *EcoRI* restriction sites are underlined):

CCR-for: 5'-GACCGTCATATGCCCGAAAGCCATGCGCA  
GAGCGCG-3'

CCR-rev: 5'-CCGCACGGAATTCCGAGCACCTGGTGCC  
GTCACCGGC-3'

The amplified 1475 bp DNA fragment was digested with *NdeI* and *EcoRI* and then ligated to *NdeI* and *EcoRI* -cleaved pNGEM/OriT/P, obtaining the resulting plasmid pNGEM/OriT/P/CCR.

Plasmids pNGEM/OriT/P and pNGEM/OriT/P/CCR were introduced into the *S. ambofaciens* strain by conjugation with *E. coli* GM2929/pUB307::Tn7, as described previously Kieser et al. (2000). To allow plasmid selection, conjugation medium SMII was supplemented with kanamycin (25 µg/ml).

## Overproduction and Purification of CCR (Srm4)

To overproduce and purify His6-tagged *S. ambofaciens* CCR (*Srm4*) in *E. coli*, the entire *ccr* (*srm4*) gene was amplified by PCR using the pNGEM/OriT/P/CCR construct as template, and using the primer pair CCR-for (above reported) and pETCCR-rev (*HindIII* restriction site is underlined):

pETCCR-rev: 5'-GAATTCAAGCTTGCGGAACCGGTTGA  
TCGCGTCCA-3'

The amplified DNA fragment was digested with *NdeI* and *HindIII* and then ligated to *NdeI* and *HindIII* -cleaved pET21b, obtaining the resulting plasmid pET/Srm4. pET/Srm4 was then introduced into *E. coli* BL21(DE3) competent cells.

*Escherichia coli* BL21(DE3) harboring pET/Srm4 was grown in 1 l of LB medium supplemented with ampicillin (10 µg ml<sup>-1</sup>) to A600 = 0.6 at 37°C and 250 r.p.m. Culture was induced by the addition of 0.4 mM IPTG and incubated for 16 h at 20°C and 180 r.p.m. Cells were harvested by centrifugation at 5,500 RCF for 10 min, re-suspended in lysis buffer (20 mM Tris-HCl pH 8.0, 200 mM KCl, 25 mM imidazole, 10% glycerol and 0.1 mg ml<sup>-1</sup> DNase I). Cells were lysed by passing the suspension through a French Press twice at 1000 psi, and the cell debris were removed by centrifugation at 10,000 × g for 20 min at 4°C. The clarified cell lysate was bound to 1 ml of Ni-IMAC resin (Bio-Rad Laboratories) for 1 h at 4°C. The resulting slurry was transferred to a gravity-flow column, and washed twice with 1 ml of lysis buffer. The recombinant protein His6-Srm4 was then eluted with 0.2 ml of elution buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 500 mM imidazole and 10% glycerol), and frozen for storage at -80°C.

## Spiramycin Production Assay

Spiramycin production by *S. ambofaciens* cultures grown in YS broth was assessed by HPLC. At different time intervals, supernatants were filtered through Phenex-RC membrane (0.45 µm; Phenomenex). Four hundred microliter filtrated samples were lyophilized and resuspended in a mixture water-acetonitrile 70:30. The concentration of spiramycins was determined by HPLC (Beckman System Gold Programmable Solvent Module125) equipped with a UV detector (232 nm). The column was a reverse phase LiChrospher® RP-8 HPLC Column (Supelco) and the mobile phase was a mixture of 100 mM phosphate buffer (pH 2.3) containing 50 mM sodium perchlorate

and acetonitrile 70:30 v/v with a flow rate of 0.8 ml/min. Purified spiramycin (Sigma–Aldrich) freshly dissolved in phosphate buffer (pH 7) was used as standard. Two-tailed paired Student's *t*-test was used to assess statistical significance of antibiotic titre differences between bacterial strains during the time course. Statistical significance was declared at *p*-value < 0.05.

## Determination of Crotonyl-CoA Carboxylase/Reductase (CCR) Activity

*Streptomyces ambofaciens* cultures (100 ml) were harvested by centrifugation at 10,000 × *g* at 4°C for 20 min. Supernatants were discarded and mycelial pellets were washed twice in homogenization buffer (50 mM potassium phosphate pH 7.5, 1 mM EDTA, 1 mM dithioerythritol (DTE), 10% glycerol, 0.1 mM phenylmethylsulfonyl fluoride). After resuspension in homogenization buffer, mycelial extracts were prepared by mycelial disruption at 2000 psi using a French Press and subsequent centrifugation of the homogenates at 25,000 × *g* for 5 min at 4°C to remove debris. The supernatant was collected and used for the enzyme activity measurement. Protein concentrations were determined by using the colorimetric Bio-Rad colorimetric assays for protein determination with BSA as a standard. Protein extracts were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) according to standard procedures.

The crotonyl-CoA carboxylase/reductase (CCR) activity was determined as described (Erb et al., 2007) in a reaction mixture containing 50 mM phosphate buffer (pH 7.5), 1 mM EDTA, 1 mM DTE, 1 mM crotonyl-CoA, 30 mM NaHCO<sub>3</sub> and 0.1–0.3 mg of mycelial extract proteins, or 0.10–0.2 μg of purified CCR protein. Upon equilibration at 30°C for 5 min, 6 mM NADPH was added to initiate the reaction. CCR activity was measured spectrophotometrically by following the decrease in absorbance at 340 nm ( $\epsilon_{\text{NADPH}} = 6.22 \text{ mM}^{-1} \text{ cm}^{-1}$ ), which corresponds to the oxidation of the NADPH. Specifically, the decrease in the absorbance due to CCR-independent NADPH oxidation was excluded from the calculation of CCR activity, by performing an analogous reaction in the absence of crotonyl-CoA. A unit (U) of enzyme activity is the amount of enzyme catalyzing the oxidation of 1 μmol NADPH min<sup>-1</sup>.

## RNA Extraction, Real-Time qPCR and RNAseq Experiments

Total bacterial RNA was extracted from *S. ambofaciens* growing in YS medium at 28°C on a rotary shaker at 200 rpm, by using the RNeasy Mini kit according to the manufacturer's instructions (QIAGEN). RNase-free DNase I was used to eliminate traces of DNA in the samples, in accordance with the instructions of the manufacturer (Promega, Madison, WI, USA).

Semi-quantitative analyses of the *ccr*-specific transcripts, normalized to 16S rRNA, were carried out by Real-Time qPCR. Total RNAs (1 μg) were reverse-transcribed by using random hexamer (2.5 μM) with Superscript RT (Invitrogen). About 0.1–1% of each RT reaction was used to run real-time PCR on a SmartCycler System (Cepheid) with SYBR<sup>®</sup> Green JumpStart Taq ReadyMix (Sigma–Aldrich) and the oligonucleotide

primer pairs 16Suniv-1/16S-univ-2 (specific for 16S rRNA gene) and *ccr*-rtF/*ccr*-rtR (specific for *ccr*). The primer sequences (and names) were: 5'-CAGCAGCCGCGTAATAC-3' (16Suniv-1); 5'-CCGTCAATTCCTTTGAGTTT-3' (16Suniv-2); 5'-CGATCGTCACCTGCGCCTCCAC-3' (*ccr*-rtF); 5'-CGAAG TGGGTGCCGACGATAC-3' (*ccr*-rtR). Primers were synthesized as a service by MWG-Biotech AG Oligo Production. RT-PCR product lengths were 185 bp for 16Suniv-1/16S-univ-2 and 96 bp for *ccr*-rtF/*ccr*-rtR. Real-Time qPCR samples were run in triplicate. The Real-Time qPCR conditions were: 30 s at 94°C, 30 s at 60°C, 30 s at 72°C for 35 cycles; detection of PCR products was performed at 83°C. The relative transcript levels were calculated using the 2<sup>-ΔΔCt</sup> method (Livak and Schmittgen, 2001).

For RNAseq experiments ribosomal RNAs were depleted starting from 1 μg of total RNA from each of the time points and biological replicates by using the RiboZero Gram positive kit (Epicentre, Illumina), strand specific RNA-seq libraries were prepared by using the ScriptSeq<sup>™</sup> v2 RNAseq library preparation kit (Epicentre, Illumina) starting from 50 ng of previously rRNA depleted RNA from each biological replicate and for all the time points analyzed. Then each library was sequenced on a MiSeq Illumina sequencer and 76 bp reads were produced.

## RNAseq Data Analysis

Bowtie 2 (v2.2.6) (Langmead and Salzberg, 2012) was used to align raw reads to *S. ambofaciens* ATCC 23877 genome (GCF\_001267885.1). Multi-mapping reads alignment was forced using non-deterministic option to obtain a single mapping locus for the reads located on the terminal arm region (nt 1-202694/8101246-8303940), which is known to be duplicated. High quality reads were selected imposing the following criteria: mapping quality for uniquely mapping reads MAPQ greater than 30 and alignment score greater than -15; for multi-mapping reads the alignment score was set equal or greater than -15. These criteria define high quality reads only those showing a maximum of three variations (mismatches or short indels): the sum of penalty scores assigned to mismatches and short indels should not be greater than -15.

rRNA depletion, strand specificity and gene coverage were evaluated using BEDTools (v2.20.1) (Quinlan and Hall, 2010) and SAMtools (v0.1.19) (Li et al., 2009) to verify the library preparation and sequencing performances (see Supplementary Data Sheet 7 and Table S1). To obtain the raw read counts of protein coding genes on unique chromosomal regions and on the pSAM1 plasmid, strand specific reads overlapping for at least 50% of their length to the genes present in *S. ambofaciens* ATCC 23877 RefSeq annotation were considered in each sample. For protein coding genes located in duplicated regions it was impossible to determine the exact chromosomal origin of the reads. To overcome this problem we reported only one of the two homologs in the final list of genes and the relative raw count is the sum of all the reads aligning to both loci (strand specificity was not considered). The R (R Development Core Team, 2012) package DESeq2 (v1.6.1) (Love et al., 2014) was then used to normalize the counts basing on the median-of-ratio method to estimate the scaling factors.

Time course analysis was performed with maSigPro (Conesa et al., 2006; Nueda et al., 2014). To find the genes whose expression was significantly altered during the time-course, we considered only those expressed in at least five over eight samples. The regression parameters were left to default values ( $\alpha = 0.05$ ;  $Q = 0.05$ ) in both the two steps of the analysis, in particular a forward step regression model was adopted in the second step. Significant genes lists were obtained by imposing R-square  $> 0.7$ , in order to select genes correctly fitting the model as suggested by Conesa et al. (2006). Finally hierarchical clustering was applied to the expression values of the significant genes to individuate 9 time-dependent patterns of gene expression.

The enrichments of functional categories were calculated using Fisher test and BH correction for multiple testing, the categories with an adjusted  $p$ -value  $\leq 0.05$  were considered significantly enriched (see next paragraph for functional categories definition).

Raw data are publicly available at Sequence Reads Archive under accession number BioProject PRJNA342588.

## Metabolic Reconstruction and Modeling

The Refseq *S. ambofaciens* ATCC 23877 genome annotation was manually refined by using RAST (Overbeek et al., 2014) and antiSMASH (Weber et al., 2015): functional categories were defined according to RAST classification, and the genes encoding secondary metabolites, annotated by antiSMASH, were manually added to the category of secondary metabolism. This version of the genome annotation is reported in Supplementary Data Sheet 1. A draft metabolic reconstruction of *S. ambofaciens* ATCC 23877 was initially assembled using RAST annotation system with default parameters and then downloaded from the ModelSEED database (Henry et al., 2010). Afterward, this reconstruction was extensively curated by manual inspection following the main steps listed in Thiele and Palsson (2010). Several additional information sources were interrogated, including KEGG (Kanehisa and Goto, 2000), BRENDA (Scheer et al., 2011) and MetaCyc (Caspi et al., 2006). In addition, models of more/less related organisms (*S. coelicolor* and *M. tuberculosis*) were used as reference in a comparative genomics workflow for identifying potentially missing reactions. BLAST (Altschul et al., 1990) searches (adopting the Bidirectional Best Hit criterion) on the *S. ambofaciens* ATCC 23877 genome were carried out to confirm/reject the inclusion of further genes/reactions from related organisms to the *S. ambofaciens* model. The list of *S. ambofaciens* cellular transporters was refined probing the Transporter Classification Data Base (TCDB, Saier et al., 2014). Biosynthetic pathways leading to specific antibiotics known to be produced by *S. ambofaciens* were not initially included in the draft reconstruction. For this reason, extensive bibliographical data mining was carried out to include spiramycin (Richardson et al., 1990; Karray et al., 2007; Nguyen et al., 2010), stambomycin, antimycin, and congoic acid into the reconstruction.

Overall, at the end of the manual refinement process, 112 reactions, 80 metabolites and 81 genes were added to the initial reconstruction.

To date, no detailed information on the biomass composition of *S. ambofaciens* ATCC23877 is available in scientific literature.

Accordingly, the biomass forming reaction from the *S. coelicolor* A3(2) metabolic model (Borodina et al., 2005) was used to include a reaction accounting for biomass assembly in *S. ambofaciens* ATCC23877 reconstruction.

The FBA method was employed to simulate flux distribution in different conditions. Briefly, FBA is a constraint-based method relying on the representation of the biochemical system under investigation in the form of a stoichiometric matrix  $S$  ( $m \times n$ ), where  $m$  is the number of metabolites and  $n$  the number of reactions. FBA is based on the assumption of the cellular pseudo-steady state, according to which the net sum of all the production and consumption rates of each internal metabolite within a cell is considered to be zero. Under this assumption, the system can be described by the set of linear equations:

$$\frac{dX_i}{dt} = \sum_{j=1}^M S_{ij}v_j = 0, \quad \forall i \in N$$

in which  $X_i$  is the concentration of metabolite  $i$ ,  $S_{ij}$  is the stoichiometric coefficient of the  $i^{\text{th}}$  metabolite in the  $j^{\text{th}}$  reaction,  $v_j$  is the flux of the  $j^{\text{th}}$  reaction,  $N$  the entire set of metabolites and  $M$  the entire set of reactions.

Upper and lower bounds of flux through each reaction act as further constraints and are expressed as:

$$l_b < v_j < u_b$$

where  $l_b$  and  $u_b$  are the lower and upper limits for reaction  $j$ , respectively. Finally, FBA exploits linear programming to determine a feasible steady state flux vector that optimizes a given objective function (e.g., biomass production).

The reconstructed model was analyzed using COBRAToolbox-2.0 (Hyduke et al., 2011; Schellenberger et al., 2011) in MATLAB® R2009b (Mathworks Inc.). Gurobi 5.6<sup>1</sup> and GLPK 4.32<sup>2</sup> solvers were used for computational simulations presented herein.

## Pareto Optimality Calculation

We calculated the growth dependencies between biomass and antibiotics biosynthesis using Pareto optimality as described in Heinken et al. (2013). We built Pareto optimum curves by (i) determining the minimal and maximal flux through biomass assembly reaction, (ii) fixing the flux through this reaction at different steps spanning minimal to maximal flux and (iii) maximizing the flux through each antibiotic synthesis reaction for each step. The procedure was repeated with biomass and antibiotics assembly reactions exchanged.

## In Silico Identification of Overexpression Targets for Spiramycin Overproduction

The FSEOF algorithm for flux scanning based on enforced objective flux was implemented for the selection of gene targets to be amplified as described in Choi et al. (2010).

<sup>1</sup><http://www.gurobi.com>

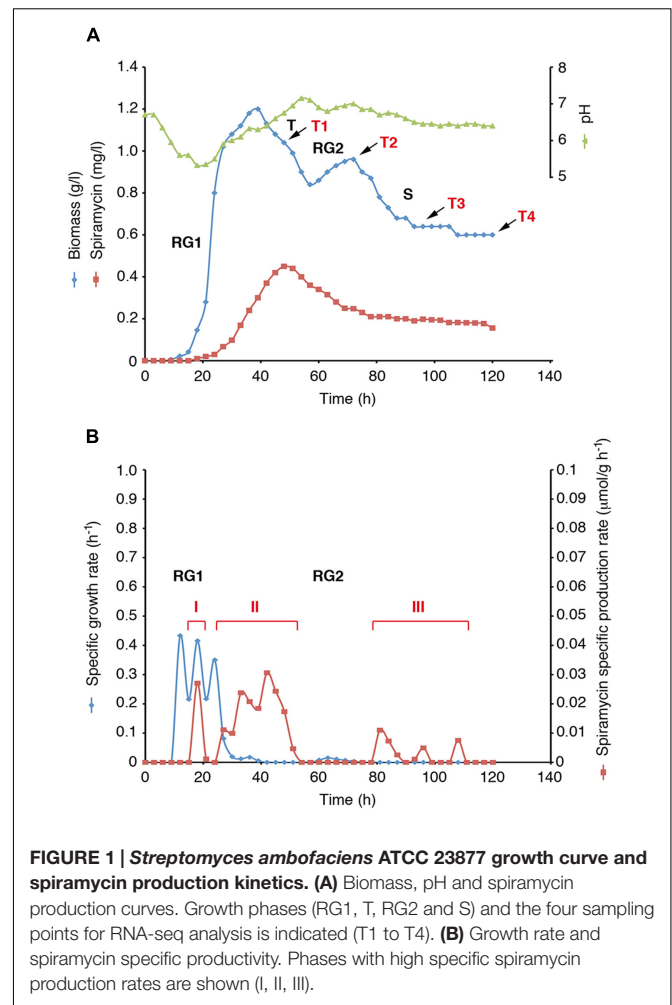
<sup>2</sup><http://www.gnu.org/software/glpk/>

Briefly, initial fluxes were calculated by constraint-based flux analysis using the objective function of maximizing biomass formation and setting (i) upper and lower boundaries to exchange reactions as to simulate growth in complex medium, (see below) and (ii) the spiramycin production rate to the experimentally measured value of  $0.0017 \text{ mmol} \cdot \text{h}^{-1} \cdot \text{g}^{-1}$ . Next, the objective function was set to spiramycin production and the theoretical maximum product formation rate was calculated. Afterward, FSEOF was performed under the objective function of maximizing cell growth while the spiramycin production rate is gradually enforced from the flux value used in the first simulation to the theoretical maximum production rate obtained with the second simulation (i.e., performing ten different model optimizations). Candidate target genes were then collected identifying those fluxes that increased upon the application of the enforced objective flux without changing the reaction's direction. More specifically, we selected those reactions (and corresponding coding genes) showing an increased flux in at least one of the 10 simulations, without changing the reaction's direction. Please refer to Choi et al. (2010) for a detailed and formal explanation of the FSEOF algorithm.

## RESULTS AND DISCUSSION

### Growth of *S. ambofaciens* ATCC 23877 and Spiramycin Production Kinetics

Mycelial growth and spiramycin production were measured during *S. ambofaciens* ATCC 23877 growth in shake flask batch cultivation using complex YS medium (Figure 1 and Supplementary Data Sheet 2). Biomass data displayed the typical two-stage growth curve of actinomycetes consisting of an initial stage of rapid growth (RG1) (12–40 h), a transition phase of stalled growth (T) (40–60 h), a secondary growth phase (RG2) (60–72 h) and a stationary (S) phase (72–120 h) (Figures 1A,B). The analysis of growth kinetics also showed that during RG1 specific growth rates ( $\mu_{\max}$ ) were not constant over time but exhibited three distinct peaks at 12, 18, and 24 h, respectively (Figure 1B). Spiramycin production started at 18 h and reached the highest titre in the medium at 48 h (Figure 1A). Then its concentration decreased progressively till 78 h and afterward it remained nearly constant till 117 h (Figure 1A). The evaluation of specific production rates ( $q_p$ ) led us to distinguish three distinct phases of spiramycin production (Figure 1B and Supplementary Data Sheet 2): a brief pulse of antibiotic production culminating at 18 h (phase I) during RG1, a robust and prolonged phase of production (between 27 and 51 h) (phase II) starting during RG1 and lasting over almost the entire T phase, and a third phase of production (between 81 and 108 h) (phase III) starting after RG2 and lasting over almost the entire S phase. Notably, production phases II and III seem to be preceded by growth rate peaks and are characterized by either growth rate decline (phase II) or no growth (phase III) demonstrating a strict interconnection between growth pulses and activation of spiramycin production (Supplementary Data Sheet 2).

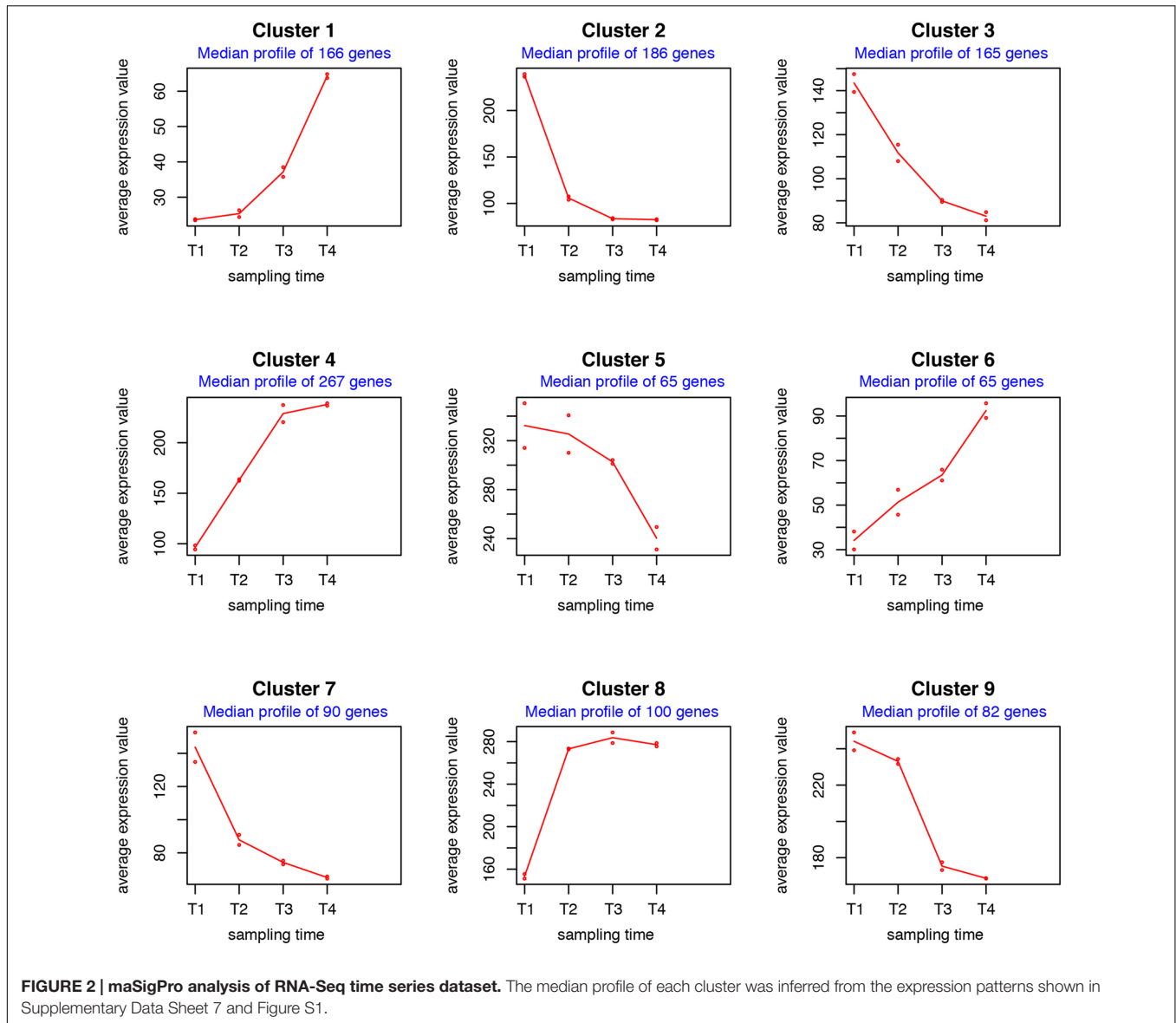


**FIGURE 1 | *Streptomyces ambofaciens* ATCC 23877 growth curve and spiramycin production kinetics. (A)** Biomass, pH and spiramycin production curves. Growth phases (RG1, T, RG2 and S) and the four sampling points for RNA-seq analysis is indicated (T1 to T4). **(B)** Growth rate and spiramycin specific productivity. Phases with high specific spiramycin production rates are shown (I, II, III).

### Time-Resolved Transcriptomics of *S. ambofaciens* ATCC 23877

To gain a better insight into the physiological and genetic mechanisms governing mycelial growth and antibiotic production, after RG1, global transcriptional analysis was carried out by RNA-Seq technology. For RNAseq experiments, samples were collected from shake flask batch cultures in YS medium at four different time points (T1 = 48 h, T2 = 72 h, T3 = 96 h, and T4 = 120 h) for RNA extraction. Two biological replicates were analyzed for each time point.

The maSigPro bioconductor package (Conesa et al., 2006) was adopted to analyze the RNA-Seq time series dataset starting from the normalized gene counts obtained after DESeq2 (Love et al., 2014) standard normalization (see Material and Methods). maSigPro was initially developed for the analysis of single and multi-series time course microarray experiments, and was recently adapted to deal also with Next Generation-Sequencing (NGS) series of data (Nueda et al., 2014). This package follows a two-step regression strategy to find genes with significant temporal expression changes and/or significant differences between experimental groups. Following the first step of regression the software individuated 3551 significant



genes in our dataset and then, in a second step, it refined this number to a total of 1186 variable genes, 598 upregulated and 588 downregulated during the time course. Among them we found 30 upregulated genes coding for biosynthetic enzymes for secondary metabolites, most of which involved in butyrolactone, stambomycin, and alpomycin/kinamycin production; while among the downregulated genes 37 participate in spiramycin production. Moreover 34 genes were positioned on the duplicated arm region where the cluster of genes coding for alpomycin and kinamycin production is located, while 32 genes were expressed by the extrachromosomal plasmid pSAM1 (Supplementary Data Sheet 3).

The 1186 variable genes were finally subdivided into nine clusters according to their gene expression profiles (Supplementary Data Sheets 3, 7 and Figure S1) and median profile was then inferred for each cluster (Figure 2). Clusters

1, 4, 6, and 8 consist of genes whose expression levels increase during the time course, although with distinct patterns and basal expression levels. In particular the expression levels of genes in cluster 8 has high basal expression at T1 (48 h), rapidly increased from T1 (48 h) to T2 (72 h) and then remained roughly constant throughout T3 (96 h) and T4 (120 h). Expression levels of genes in cluster 4 also exhibited high basal expression at T1 (48 h), progressively increased from T1 to T3 and remained high at T4. These two clusters are significantly enriched in functional categories “cell division and cell cycle” (cluster 4), “cell wall and capsule” (clusters 4 and 8), “DNA metabolism” (clusters 4 and 8), and “sulfur metabolism” (cluster 8) (Supplementary Data Sheet 3). The enrichment in these categories is consistent with growth kinetics indicating a secondary growth phase (RG2) between 60 and 72 h (Figure 1).

Genes in clusters **1** and **6** showed low expression levels at T1 (48 h) and gradually reached their highest expression level during T4, although with different kinetics. Cluster **1** is enriched in the functional category “secondary metabolism” (Supplementary Data Sheet 3). In particular, it contains five type I polyketide synthase (PKS)-encoding genes involved in stambomycin biosynthesis, and several regulatory genes (belonging to TetR and SARP families) of the alpomycin/kinamycin gene cluster. Genes involved in butyrolactone biosynthesis also belong to this cluster (Supplementary Data Sheet 3). Cluster **6** is enriched in the functional category “fatty acid, lipid, isoprenoids,” but it also contains four type I PKS-encoding genes of the stambomycin gene cluster (Supplementary Data Sheet 3). This result indicates that all nine PKSs-encoding genes involved in the biosynthesis of the large (51-membered) glycosylated macrolide stambomycin (Laureti et al., 2011) achieve the highest expression levels in the late stationary phase (S). This finding is of interest for management of spiramycin production. In fact, the biosynthetic pathways leading to spiramycin and stambomycin compete with each other for common metabolic precursors (i.e., malonyl-CoA and methylmalonyl-CoA) during the assembly of the macrolide scaffold.

It is also noteworthy that the up-regulated clusters **1**, **4**, **6**, and **8** contain more than 95% of the genes modulated on pSAM1 indicating a global gene expression increase of this extrachromosomal element during late growth.

Such a competition is possibly alleviated by differential temporal expression patterns of the genes involved in spiramycin and stambomycin biosynthesis (Supplementary Data Sheet 3) as suggested by maSigPro analysis that assigned 4 spiramycin biosynthetic genes to cluster **2**, 16 genes to cluster **3**, 5 genes to cluster **7** and 2 genes to cluster **9**. Indeed, although with different kinetics, the spiramycin genes belonging to the clusters **2**, **3**, **5**, **7**, and **9** exhibited a general trend of down-regulation during the time course while the stambomycin genes increased their expression level during the time course (Figure 2 and Supplementary Data Sheet 7 and Figure S1). Consistently with growth kinetics (Figure 1), clusters **2** and **3** contains many genes coding for ribosomal proteins and translation factors, and genes involved in central carbon (and energy) metabolism. Indeed, all these genes are expected to be strongly down-regulated beyond the RG1. Cluster **2** is enriched in the functional categories “membrane transport,” “virulence, disease and defense,” while cluster **3** is enriched in genes involved in “secondary metabolism” (essentially, spiramycin biosynthesis). Cluster **7** exhibits an enrichment in the category “fatty acid, lipid, isoprenoids,” while clusters **5** and **9** exhibit no significant enrichment in any considered functional categories.

## The Metabolic Reconstruction of *S. ambofaciens* Is in Agreement With Experimental Growth Data on Defined Media

A *S. ambofaciens* ATCC 23877 draft reconstruction was obtained from ModelSeed as described in “Material and Methods.” As it is usually expected for automatically reconstructed metabolic

networks, this draft model failed in predicting *S. ambofaciens* growth, both on known defined media and on an arbitrarily complex one. Thus, the model underwent a manual refinement process which led to the inclusion of 112 reactions, 80 metabolites and 81 genes, the re-formulation of the biomass assembly reaction (following the one available for *S. coelicolor*) and the inclusion of gap-filling reactions (i.e., not gene-encoded) to reconcile *S. ambofaciens* growth on defined minimum media (see below). The obtained reconstruction was named iMF1244 according to the standard nomenclature and its main features are reported in Table 1. The SBML-formatted version of iMF1244 is available as Supplementary Data Sheet 4.

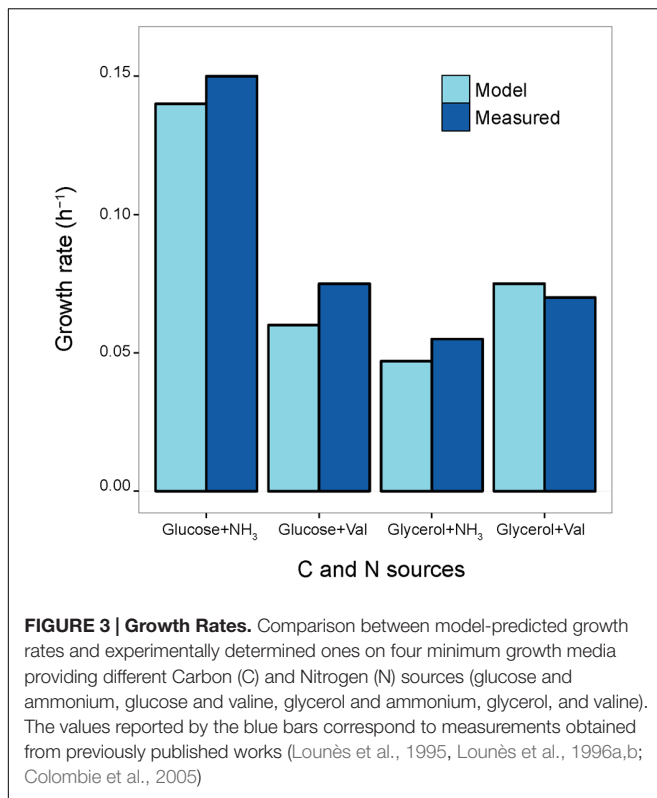
The reliability of the reconstructed *S. ambofaciens* model was inferred by comparing the predicted growth rates on a set of different minimal media against a set of experimentally determined ones available from previous studies (Lounès et al., 1995, 1996a,b; Colombie et al., 2005). Accordingly, *in silico* minimal growth media were defined using exchange reactions present in the model, and biomass optimization was selected as the model objective function (O.F.). More in detail, lower bounds of exchange reactions accounting for all the salts present in the experimental medium were set to  $-10 \text{ mmol g}^{-1} \times \text{h}^{-1}$ , in order to mimic non-limiting conditions. The carbon source of each available minimum growth medium was then chosen as the unique carbon source of this *in silico* medium and its uptake flux was set to the corresponding experimentally determined value. The compositions of all the simulated *in silico* media are reported in Supplementary Data Sheet 7 and Table S2. The growth rates that were inferred by the *in silico* simulation were then compared to those experimentally determined, revealing an overall good agreement (Pearson’s product-moment correlation 0.97, *p*-value 0.02, Figure 3). This indicates that the model should be able to reproduce a realistic picture of the *S. ambofaciens* ATCC 23877 metabolic landscape.

Afterward, we tested the response of the model in a simulated complex medium. The assimilation of yeast extract (YE) amino acids and its effect on the production of spiramycin in culture of *S. ambofaciens* has been extensively analyzed by Benslimane et al. (1995). In this work the authors showed that in a YE medium a specific growth rate ( $\mu_{\text{max}}$ ) of  $0.15 \text{ h}^{-1}$  was soon established after setting up the culture. The specific growth rate then decreased sharply to reach zero value at 24 h and rose up again when a second growth phase with a  $\mu_{\text{max}}$  of  $0.012 \text{ h}^{-1}$

**TABLE 1 | Main features of the *Streptomyces ambofaciens* ATCC23877 metabolic reconstruction.**

<i>S. ambofaciens</i> ATCC23877 genome	
Genome size (bp)	8099129
N. of protein encoding genes	7208
<i>iMF1244</i> model	
N. of genes (% of coding genes)	1244 (17%)
N. of reactions	1473
Gene-associated	1210
Non gene-associated (Exchange reactions)	263 (144)
N. of metabolites	1283





was observed. During this second growth phase, spiramycin biosynthesis was observed, with a specific production rate ( $q_p$ ) of  $1.46 \text{ mg}^{-1} \text{ h}^{-1} \text{ g}^{-1}$  (corresponding to  $0.0017 \text{ mmol} \cdot \text{h}^{-1} \text{ g}^{-1}$ ). The growth on YE of *S. ambofaciens* was characterized by the consumption of all the amino acids available in the medium (Benslimane et al., 1995). Accordingly, to account for such behavior, we added to the model all the 20 amino acids exchange and transport reactions. Since no detailed information is available concerning the single amino acids uptake rates in the second growth phase on YE, each uptake flux was arbitrary set to the value of  $0.015 \text{ mmol} \cdot \text{h}^{-1} \text{ g}^{-1}$ . In such conditions, setting the growth rate exactly to the measured value of  $0.012 \text{ h}^{-1}$ , the model predicts a spiramycin production rate of  $0.0016 \text{ mmol} \cdot \text{h}^{-1} \text{ g}^{-1}$ , which resembles the one experimentally determined ( $0.0017 \text{ mmol} \cdot \text{h}^{-1} \text{ g}^{-1}$ ).

In a typical scenario, the metabolic pathways for the production of secondary products (such as antibiotics) should compete for the pool of available metabolic intermediates with all the other cellular biosynthetic processes (e.g., biomass formation and/or homeostasis maintenance). This is typically accounted for by the so-called Pareto front between the two objectives (Marler and Arora, 2004). Pareto front refers to the set of resources allocation that is Pareto efficient, i.e., the situation in which it is impossible to make any one objective better off without making at least one individually worse. In other words, a pair of objective functions (such as a biomass and antibiotic assembly reactions) can be balanced to find the set of optimal solutions for which one objective can only be improved at the expense of the other. Recently, it has been shown that cells' metabolism operates

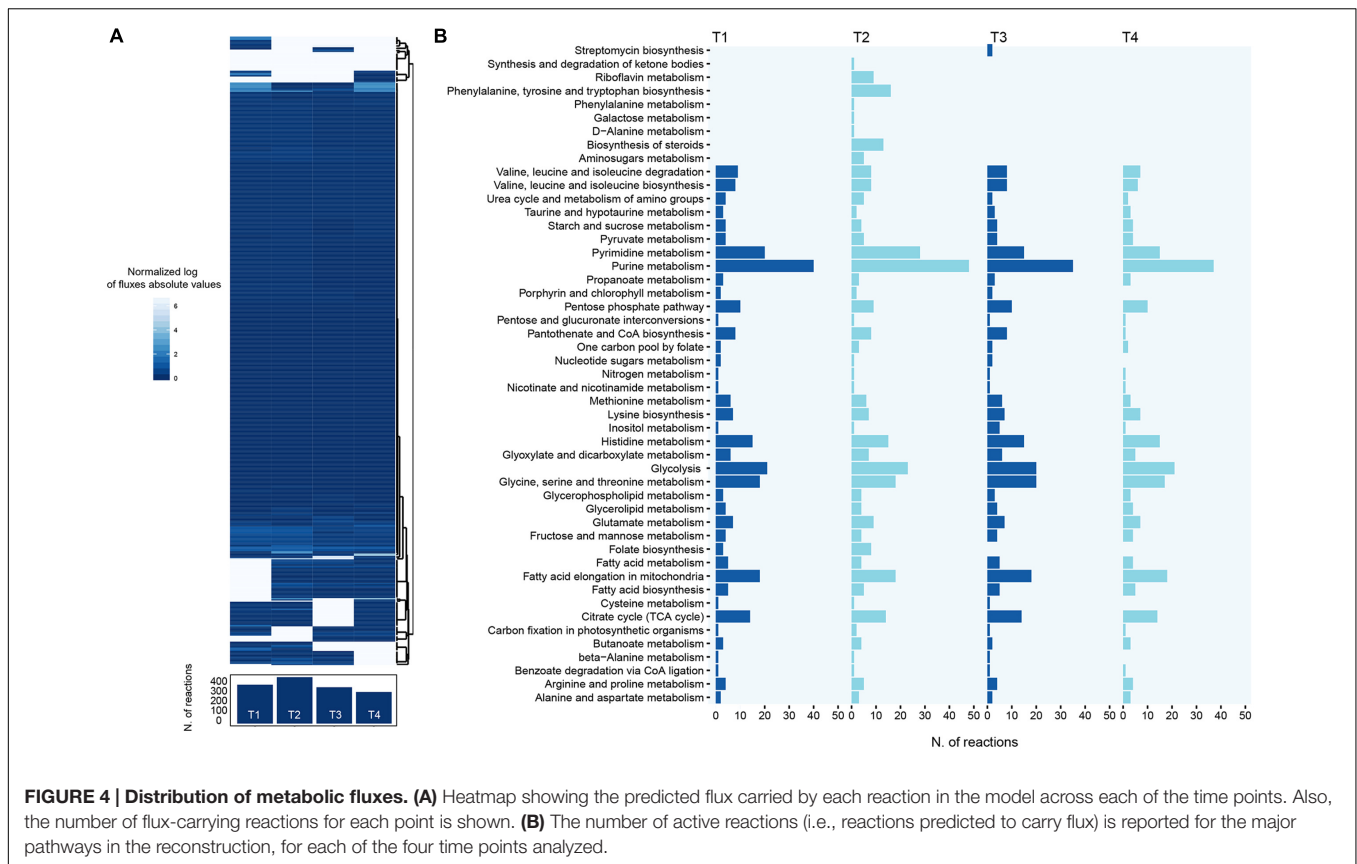
close to the Pareto-optimal (Schuetz et al., 2012) and that the calculation of Pareto fronts represents a technique that is more effective than weighting the objectives (Gevorgyan et al., 2011), especially when genetic algorithms are used to approximate the front (Boghigian et al., 2010; Angione and Lió, 2015).

Accordingly, we computed (see Materials and Methods for details) Pareto optimality for each of the four antibiotics that were included in the reconstruction. As shown in Supplementary Data Sheet 7 and Figure S2 the maximum theoretical yields were observed in the condition of  $\mu = 0$  and were  $0.0052$ ,  $0.0105$ ,  $0.00012$  and  $0.00086 \text{ mmol} \cdot \text{h}^{-1} \text{ g}^{-1}$  for spiramycin, antimycin, stambomycin and congocidine, respectively. Afterward, the model (correctly) predicts a decrease in the production rates of each antibiotic as a consequence of an increase of the growth rate. These results, together with the consistence of growth rates prediction, support the iMF1244 as being a reliable reconstruction of the metabolism of this *S. ambofaciens* ATCC 23877.

## Expression Data Integration With Metabolic Modeling Identifies System-Level Metabolic Trends

Here we used our reconstruction to provide a global description of *S. ambofaciens* metabolism in the analyzed time points, by integrating gene expression data with constraint-based metabolic modeling. In general, many methods to achieve this task have been developed to date (Machado and Herrgard, 2014). However, a recent work (Kim et al., 2016) showed that, among them, iMAT (Shlomi et al., 2008) led to more accurate predictions in the analysis of *Streptomyces coelicolor* metabolic model. Briefly, iMAT uses gene expression values to divide reactions into two groups: highly and lowly expressed. The algorithm next seeks to find the flux distribution that maximizes the consistency with this classification. Accordingly, we first computed the first and third quartiles for RPKM values in each time point to define down- and up-regulated genes, respectively (obtained values are reported in Supplementary Data Sheet 7 and Table S3). Next, we performed four different FBA optimizations (one for each of time points analyzed) using iMAT to constrain fluxes across reactions according to gene expression values. For each optimization the flux through biomass assembly and spiramycin production reactions was set to the values experimentally determined (see Figure 1 and Supplementary Data Sheet 7 and Table S4) and the resulting fluxes distribution analyzed. To account for the cellular maintenance cost, an arbitrary cellular growth rate of  $1e^{-10} \text{ h}^{-1}$  was set in those cases in which no growth had been experimentally observed. Results obtained, however, were not affected when this arbitrary cellular growth rate was replaced by a simulated ATP maintenance cost. In this case a distribution of overlapping fluxes was observed. Figure 4A shows the flux across each reaction for every time step and the number of flux-carrying reactions.

Our model predicts that, on average, 392 reactions are active across all the time points (standard deviation,  $SD = 63.4$ ). Accordingly, around 1000 reactions are predicted to be inactive in each time point (73% of the model reactions, on average).



This is in agreement with the fact that the bacterium is growing in a complex medium and thus can find in the surrounding environment most of its cellular building blocks. Interestingly, the second time point is the one displaying the highest number of flux-carrying reactions (475), in agreement with the fact that, unlike the other time points, *S. ambofaciens* is actively producing biomass in this moment of the growth phase. After this time point, *S. ambofaciens* experiences a decreasing trend in the number of flux-carrying reactions (Figure 4) reflecting with both the reduced growth rate and lower spiramycin productivity observed *in vivo* in last phases of the growth.

From a functional viewpoint, our modeling predicts that the set of active pathways is largely shared by the four time points (Figure 4B). Indeed, a similar scheme of reactions is apparently carrying flux across all the time points, including the main central routes of *S. ambofaciens* metabolism (e.g., purine and pyrimidine metabolism, glycolysis, TCA cycle). This suggests the absence of a massive metabolic reprogramming in the four time points analyzed, i.e., the final part of *S. ambofaciens* growth. The most evident exception to this general trend is represented by the second time point, the one also showing the highest number of active reactions. Our functional annotation of reactions predicted to be carrying flux suggests that this difference is due to the activation of four main pathways i.e., amino-sugar metabolism, steroid biosynthesis, phenylalanine, tyrosine and tryptophan biosynthesis, riboflavin metabolism (Figure 4B). Thus, according to our simulations, these pathways (or part

thereof) play a role when the cell is trying to achieve two objective functions simultaneously, i.e., biomass and antibiotic synthesis. The activation of riboflavin metabolism and a consequent increase in FAD cofactor availability, for example, is consistent with an increased cellular demand of acetyl-CoA (a substrate for the synthesis of many spiramycin intermediates) that, in turn, can be obtained through the FAD-dependent degradation of fatty acids. Similarly, the use of amino sugar metabolism might be propaedeutic to the recycling of nucleotide hexoses (e.g., glucose-1P), valuable precursors for the sugar moieties of the spiramycin molecule.

All these data show how changes in gene expression may redirect the cellular metabolic fluxes within the cell, leading to activate specific routes that, in turn, may represent key nodes in the utilization of intermediates that are common to multiple cellular objectives (biomass formation and spiramycin synthesis in this case). The hypothetical scenario depicted here indicates both FAD production and nucleotide sugars recycling as key factors in allowing the cell to achieve the imposed double optimization.

## Metabolic Modeling Identifies Potential Overexpression Targets for Spiramycin Overproduction

We next exploited the iMF1244 metabolic reconstruction for the genome-wide identification of gene amplification

targets potentially allowing spiramycin overproduction in *S. ambofaciens*. To this purpose we used the FSEOF approach described in Choi et al. (2010), as reported in Section “Materials and Methods.” From the set of identified potential targets, we excluded from further consideration those reactions that (i) were added to the model for gap-filling, (ii) were involved in the import (or export) of nutrients into the system for modeling purposes (i.e., exchange reactions) and (iii) were either encoded by multiple genes (enzymatic complexes) or, at least, two alternative genes (paralogs) in the genome. In all these cases, no sufficiently detailed clues on the gene(s) to be amplified for spiramycin overproduction could be derived. Following these selection criteria, our approach led to the identification of 65 potential reactions (corresponding to 54 genes) whose increased flux may lead to an increased flow through spiramycin synthesis pathway. The complete list of the reactions identified is reported in Supplementary Data Sheet 5.

A large fraction of this set of reactions is involved in the synthesis of precursors of the macrolactone backbone of spiramycin. These include, for example, glycerone phosphate (produced by sn-glycerol-3-phosphate: NAD<sup>+</sup> 2-oxidoreductase) involved in the biosynthesis of methoxymalonyl-ACP, crotonoyl-CoA (the product of butanoyl-CoA: oxygen 2-oxidoreductase) precursor of ethylmalonyl-CoA and acetyl-CoA, the substrate of 3-oxopropionate: NAD<sup>+</sup> oxidoreductase for the generation of malonyl-CoA. Reactions leading to the production of intermediates necessary for the synthesis of the deoxyhexose sugars embedded in the spiramycin molecule (e.g., alpha-D-glucose 1-phosphate 1,6-phosphomutase responsible for the conversion of D-glucose 1-phosphate into D-glucose 6-phosphate) were also identified.

Among the proposed targets, we selected the reaction carried out by crotonyl-CoA carboxylase/reductase and presumably involved in the synthesis of ethylmalonyl-CoA. The spiramycin macrolactone backbone is synthesized by a PKS that, apart from the most common polyketide precursors, malonyl-CoA and methylmalonyl-CoA, was predicted to incorporate the less common precursor, ethylmalonyl-CoA, and the functionalized extender unit, methoxymalonyl-CoA (Kuhstoss et al., 1996; Karray et al., 2007). In particular, the ethylmalonyl-CoA is used as building block to assemble only a few known polyketides, including the polyether antibiotic monensin A, whose biosynthetic pathway was the first established pathway to require ethylmalonyl-CoA (Liu and Reynolds, 1999).

Although there is more than one pathway to ethylmalonyl-CoA, the pathway involving the crotonyl-CoA carboxylase/reductase (CCR) (Erb et al., 2007) (commonly referred to as “the ethylmalonyl-CoA pathway”) appears to be the dominant source of this precursor for polyketide biosynthesis (Wilson and Moore, 2012). This pathway was originally described in purple non-sulfur bacterium *Rhodobacter sphaeroides* and then discovered also in methylotrophic bacteria such as *Methylobacterium extorquens* and in *Streptomyces* spp. In the ethylmalonyl-CoA pathway a C4-compound, acetoacetyl-CoA, derived from two molecules of acetyl-CoA, is converted to the C5-compound mesaconyl-CoA (2-methylfumaryl-CoA) that is then transformed to (2R,3S)-β-methylmalyl-CoA by hydration

and is finally cleaved to glyoxylate and propionyl-CoA. The key enzyme of the pathway is CCR, which simultaneously carboxylates and reduces the C4-compound crotonyl-CoA, forming the C5-compound (2S)-ethylmalonyl-CoA (herein indicated as “ethylmalonyl-CoA”). As illustrated in Supplementary Data Sheet 7 and Figure S3, the ethylmalonyl-CoA and its direct metabolic precursor crotonyl-CoA lies at the crossroad between key metabolic pathways devoted to energy metabolism and energy storage, lipid and amino acid metabolism and secondary metabolism.

In *S. ambofaciens* the ethylmalonyl-CoA is used as a building block to synthesize the polyketide backbone of spiramycin and antimycin, and the full set of genes necessary for a functioning ethylmalonyl-CoA pathway is clustered together in the *S. ambofaciens* chromosome not very distant from the spiramycin biosynthetic cluster. The cluster comprises the ORFs: SAM23877\_RS28800 coding for methylsuccinyl-CoA dehydrogenase (*msd* gene); SAM23877\_RS28805 coding for mesaconyl-CoA hydratase (*mcd* gene); SAM23877\_RS28810 coding for L-malyl-CoA/beta-methylmalyl-CoA lyase (*mclA* gene); SAM23877\_RS28815 encoding ethylmalonyl-CoA mutase (*ecm* gene); SAM23877\_RS28820 encoding the CCR (*ccr* gene); SAM23877\_RS28825 coding for a TetR family transcriptional regulator; SAM23877\_RS28830 coding for 3-hydroxybutyryl-CoA dehydrogenase (Supplementary Data Sheet 7 and Figure S3).

As well as in other polyketides gene clusters, a secondary copy of the CCR-encoding gene is located in both spiramycin (SAM23877\_RS26790, *srm4*) and antimycin (SAM23877\_RS01780) biosynthetic gene clusters. RNAseq data showed that under tested experimental conditions all three CCR-encoding genes were poorly expressed, the gene RS26790 has an expression level higher than the gene RS28820 which is more expressed than the gene RS01780 (RS26790 > RS28820 > RS01780), among them only the gene SAM23877\_RS26790 (*srm4*) shows an expression level above the baseline threshold (RPKM > 10) in all the time points analyzed, thus allowing us to consider it effectively expressed along the growth curve (Table 2). Although the CCR enzymatic activity of the *ccr* paralog gene products has not been proven so far, their role is likely to synchronize ethylmalonyl-CoA synthesis from acetoacetyl-CoA with polyketide assembly, and/or to increase intracellular ethylmalonyl-CoA levels in order to ensure adequate levels of the precursor during polyketide biosynthesis. Therefore, the carbon flow through the ethylmalonyl-CoA pathway was consistently associated with spiramycin productivity by our metabolic model.

**TABLE 2 | Expression of CCR-encoding genes as deduced by RPKM values from RNAseq analysis.**

Time point	Time (hours)	SAM23877_RS26790 (RPKM value)	SAM23877_RS28820 (RPKM value)	SAM23877_RS01780 (RPKM value)
T1	48	13.5	8.5	3.5
T2	72	16	7	6
T3	96	16	10.5	6.5
T4	120	32	26.5	18.5

## Manipulation of the Ethylmalonyl-CoA Metabolic Node Leads to Spiramycin Overproduction

To validate this prediction, we planned to over-express the *ccr* paralog *srn4* of the spiramycin gene cluster in *S. ambofaciens*. However, we first verified that *srn4* actually codes for an enzyme with CCR activity. To this purpose, the His6-tagged *S. ambofaciens* CCR (*Srn4*) protein was overproduced in *E. coli* BL21(DE3) strain and purified (Supplementary Data Sheet 7 and Figure S4). Then its CCR enzymatic activity was determined. The results demonstrated an enzymatic activity of  $449 \pm 11$  U mg<sup>-1</sup>. To over-express *ccr* (*srn4*), the gene was placed under the control of the *PgroES* promoter (Luo et al., 2015) in pN702GEM3, a multicopy *E. coli*-*Streptomyces* shuttle plasmid vector (Supplementary Data Sheet 7 and Figure S5). The pNGEM/OriT/P/CCR or the control plasmid pNGEM/OriT/P were then introduced into *S. ambofaciens* ATCC 23877 by intergeneric conjugation with *E. coli*, and mycelial growth (Figure 5A), spiramycin production (Figure 5B) *ccr* (*srn4*) mRNA levels (Figure 5C), and CCR enzymatic activity in crude cell extracts (Figure 5D) were analyzed in the trans-conjugant strains grown in YS medium. With respect to growth kinetics, compared to control strain (harboring pNGEM/OriT/P) the *ccr*-over-expressing strain exhibited major differences during the RG1 with about 14.6% lower final biomass values but similar growth rates (0.0697 vs. 0.0622 between 24 and 48 h) (Figure 5A and Supplementary Data Sheet 6b).

Strain manipulation did not alter significantly either soluble protein content (expressed as mg protein/g DCW) or electrophoretic pattern as determined, respectively, by Bio-Rad colorimetric assays and SDS-PAGE (Supplementary Data Sheet 7 and Figure S6). In each time point the *ccr* (*srn4*) mRNA expression levels of the strain over-expressing *ccr* (pNGEM/OriT/P/CCR), were compared at the same time both with the expression levels of the wild type and of the strain with control plasmid (pNGEM/OriT/P). As shown in Figure 5C the strain over-expressing *ccr* (*srn4*) exhibited an average maximum increase in the mRNA levels of 2–3-fold when compared respectively with wild type or with the control plasmid strain, thus indicating a moderate gene overexpression at each time point (Figure 5C). It is noteworthy that during the time course, in all the three strains *ccr* (*srn4*) mRNA levels reached the highest values of expression at 48 h. So in the light of these results we can infer that the percentage of the overexpressed *Srn4* protein in the total soluble *S. ambofaciens* proteome is therefore almost negligible.

In the *ccr*-over-expressing strain spiramycin productivity was enhanced during the RG1 and T phases. Compared to control strain spiramycin specific production rates ( $q_p$ ) were increased by a factor of 3.63 at 24 h (due to early triggering of antibiotic biosynthesis) and 1.52 at 48 h (Supplementary Data Sheet 6b), but antibiotic titres were only modestly increased by a factor of 1.503 at 48 h (i.e., an average increase of 50.3% with a standard deviation of 36.5%,  $p$ -value = 0.01078), when spiramycin titres reached the highest values during the time course. Antibiotic titres were instead significantly increased

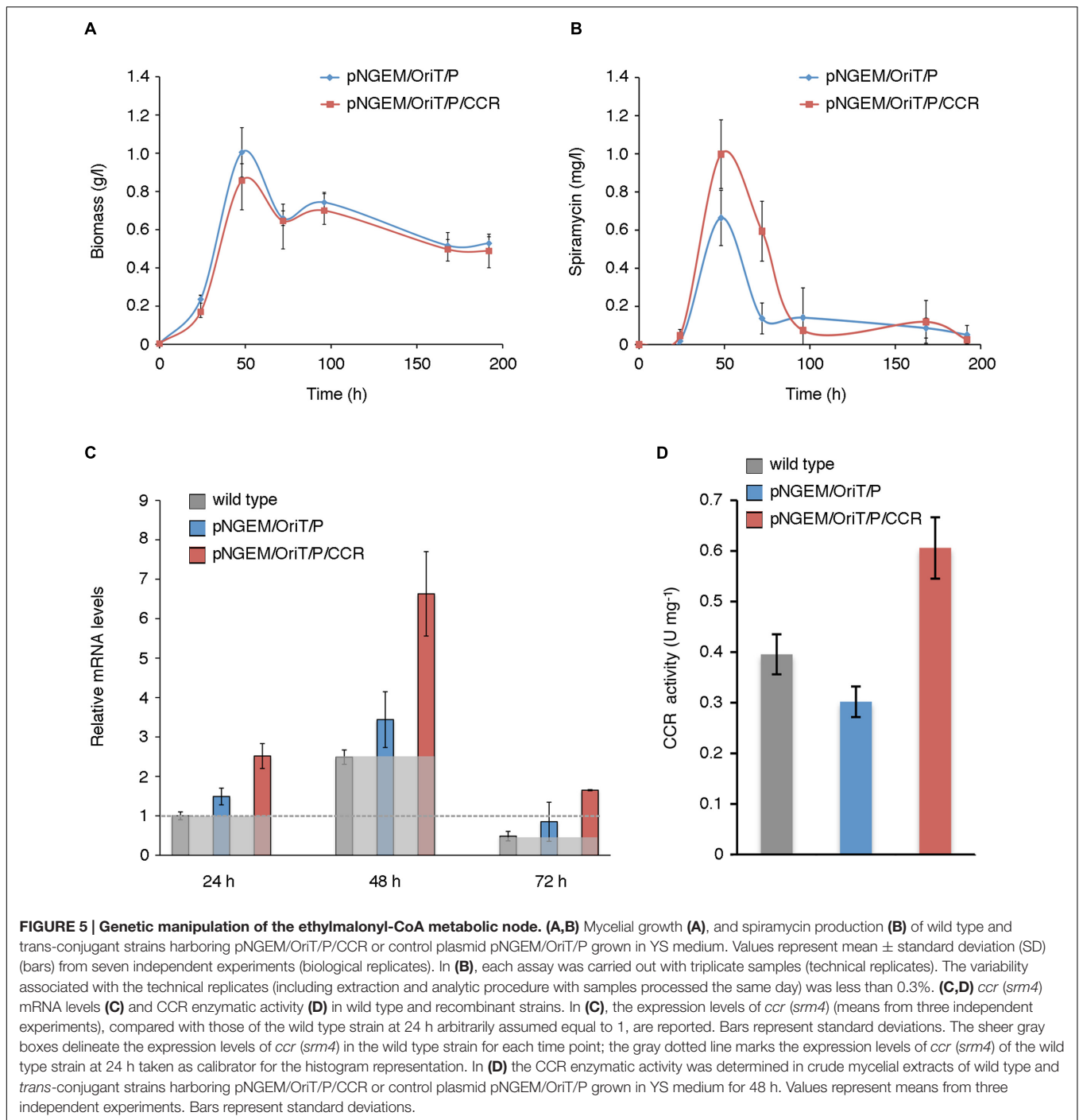
by a factor of 4.336 at 72 h (i.e., an average increase of 333.6% with a standard deviation of 235.6%,  $p$ -value = 0.00024) (Figure 5B and Supplementary Data Sheet 6a). In contrast, no statistically significant differences were observed during the late phase of antibiotic production (during RG2 and S phases).

The achieved increase was rather limited when compared to that obtained with similar metabolic engineering strategy that resulted in about 2- and 1.8-fold increase in actinorhodin production by *S. coelicolor*A3(2) strains overexpressing, respectively, the ribulose 5-phosphate 3-epimerase and the NADP-dependent malic enzyme (Kim et al., 2016). However, the increase was comparable to those obtained in other streptomycetes by genetic modification of single targets (Huang et al., 2012, 2013). For instance, in *Streptomyces tsukubaensis* 35, 40, or 50% improvement in FK506 (a 23-membered polyketide macrolide also known as tacrolimus) production could be achieved, respectively, by overexpression of *zwf2* (coding for glucose-6-phosphate dehydrogenase), *accA2* (coding for acetyl-CoA carboxylase) or *dahp* (coding for 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase). However, the combined effect of *gdhA*-deletion and *dahp*-, *accA2*-, *zwf2*-overexpression enhanced the concentration of FK506 up to 1.47-fold in fed-batch fermentations (Huang et al., 2013). In another streptomycete, such as for example *Streptomyces roseosporus*, co-overexpression of *zwf2*, *dptI*, and *dptJ* genes resulted in a 34.4% higher daptomycin concentration compared with the parental strain (Huang et al., 2012). These comparisons indicate that the achieved increase may not be irrelevant, and that spiramycin production may be further enhanced by the combined effect of *ccr*-overexpression, and genetic manipulation of other identified targets and/or improvement of fermentation conditions. Preliminary data indicate very long-chain acyl-CoA dehydrogenases, and several enzymes involved in glyoxylate cycle and glycine, serine and threonine metabolism as further good targets for strain improvement.

The results of biochemical assay with crude mycelial extracts demonstrated that the increase in spiramycin productivity paralleled with an increase in CCR activity (Figure 5D). In particular, at 48 h of growth, the enzyme activity in the *ccr*-over-expressing strain was about two-fold higher than in control strain (harboring pNGEM/OriT/P). It should be noted also that the order of magnitude of CCR activity that was determined in the protein extracts from *S. ambofaciens* is similar to those assessed in protein extracts from *Rhodobacter sphaeroides* (0.7 U mg<sup>-1</sup>), *Methylobacterium extorquens* (0.8 U mg<sup>-1</sup>) and *Streptomyces coelicolor* A3(2) (0.4 U mg<sup>-1</sup>) (Erb et al., 2007). Altogether these results validate the metabolic model providing evidence that spiramycin production may be increased by genetic manipulation of the ethylmalonyl-CoA metabolic node.

## CONCLUSION

In this study we have applied an integrated approach to explore the metabolic landscape of *Streptomyces ambofaciens* with the aim to provide a system-level understanding of its metabolic



features, and to identify a list of potential metabolic engineering targets for the overproduction of the secondary metabolites in this microorganism (**Figure 6**).

Our work culminated in the construction of a metabolic model that enabled us to identify key molecular targets for strain improvement. In particular, we proposed a set of genes/reactions that might be preferential targets for metabolic engineering experiments aimed at the overproduction of spiramycin. Notably, we selected one of these targets and we validated the model

predictions by providing evidence that spiramycin productivity may be considerably increased by enhancing the carbon flow through the recently described ethylmalonyl-CoA metabolic pathway. This goal was achieved by over-expressing the *ccr* paralog *srm4* in an *ad hoc* engineered plasmid. In the future this strategy will be valuable in reconstructing the genetic and metabolic profiles of other streptomycetes, and of great benefit to the development of industrial overproducers by genetic engineering.

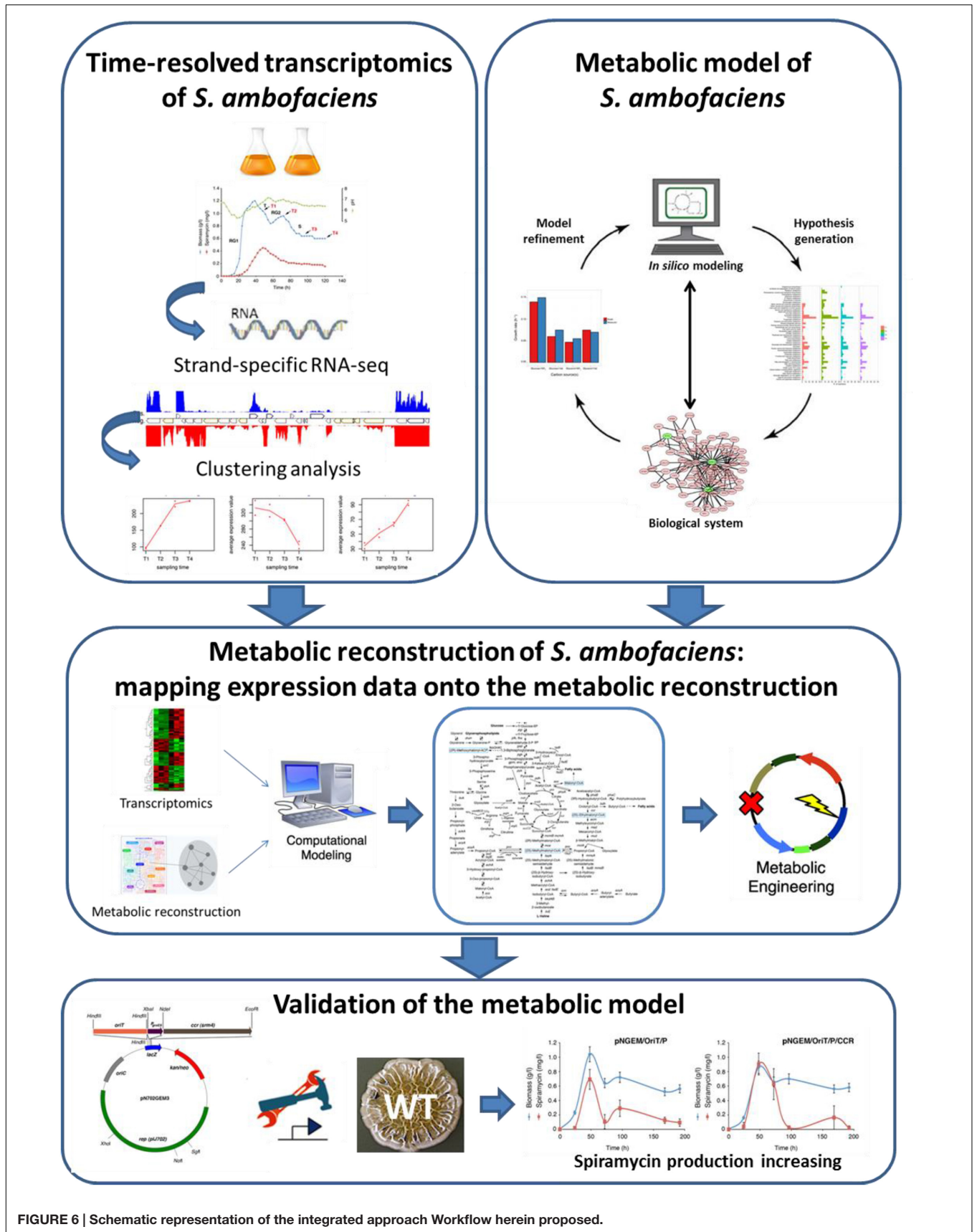


FIGURE 6 | Schematic representation of the integrated approach Workflow herein proposed.

## AUTHOR CONTRIBUTIONS

PA, GDB, MF, and CP conceived the experimental design; CC, CP, and EP produced RNA-seq data and performed the relative analysis; MF and BM constructed the metabolic model; MF performed all the simulations with the model; PA, FD, DF, GEDB, AT, and MT performed the experimental validation and provided the material for RNA-seq library preparation. PA, MF, CP, EP, and AT wrote the draft of the work. All the authors critically revised and finally approved the manuscript.

## FUNDING

This work was supported by the Italian Ministry for Education, Universities and Research [Grant numbers PON01\_02093 and PRIN 2012WJSX8K to PA].

## REFERENCES

- Alam, M. T., Merlo, M. E., Consortium, S., Hodgson, D. A., Wellington, E. M., Takano, E., et al. (2010). Metabolic modeling and analysis of the metabolic switch in *Streptomyces coelicolor*. *BMC Genomics* 11:202. doi: 10.1186/1471-2164-11-202
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.
- Angione, C., and Lió, P. (2015). Predictive analytics of environmental adaptability in multi-omic network models. *Sci. Rep.* 5:15147. doi: 10.1038/srep15147
- Baltz, R. H. (2016). Genetic manipulation of secondary metabolite biosynthesis for improved production in *Streptomyces* and other actinomycetes. *J. Ind. Microbiol. Biotechnol.* 43, 343–370. doi: 10.1007/s10295-015-1682-x
- Benslimane, C., Lebrhi, A., Lounes, A., Lefebvre, G., and Germain, P. (1995). Influence of dextrans on the assimilation of yeast extract amino-acids in culture of *Streptomyces ambofaciens* producer of spiramycin. *Enzyme Microb. Technol.* 17, 1003–1013. doi: 10.1016/0141-0229(95)00029-1
- Boghigian, B. A., Shi, H., Lee, K., and Pfeifer, B. A. (2010). Utilizing elementary mode analysis, pathway thermodynamics, and a genetic algorithm for metabolic flux determination and optimal metabolic network design. *BMC Syst. Biol.* 4:49. doi: 10.1186/1752-0509-4-49
- Borodina, I., Krabben, P., and Nielsen, J. (2005). Genome-scale analysis of *Streptomyces coelicolor* A3(2) metabolism. *Genome Res.* 15, 820–829. doi: 10.1101/gr.3364705
- Caspi, R., Foerster, H., Fulcher, C. A., Hopkinson, R., Ingraham, J., Kaipa, P., et al. (2006). MetaCyc: a multiorganism database of metabolic pathways and enzymes. *Nucleic Acids Res.* 34, D511–D516. doi: 10.1093/nar/gkj128
- Chew, W. K., Segarra, I., Ambu, S., and Mak, J. W. (2012). Significant reduction of brain cysts caused by *Toxoplasma gondii* after treatment with spiramycin coadministered with metronidazole in a mouse model of chronic toxoplasmosis. *Antimicrob. Agents Chemother.* 56, 1762–1768. doi: 10.1128/AAC.05183-11
- Choi, H. S., Lee, S. Y., Kim, T. Y., and Woo, H. M. (2010). In silico identification of gene amplification targets for improvement of lycopene production. *Appl. Environ. Microbiol.* 76, 3097–3105. doi: 10.1128/AEM.00115-10
- Colombie, V., Bideaux, C., Goma, G., and Uribealrrea, J. L. (2005). Effects of glucose limitation on biomass and spiramycin production by *Streptomyces ambofaciens*. *Bioprocess Biosyst. Eng.* 28, 55–61. doi: 10.1007/s00449-005-0015-2
- Conesa, A., Nueda, M. J., Ferrer, A., and Talón, M. (2006). maSigPro: a method to identify significantly differential expression profiles in time-course microarray experiments. *Bioinformatics* 22, 1096–1102. doi: 10.1093/bioinformatics/btl056
- Cosar, C., Ninet, L., Pinnert-Sindico, S., and Preud'homme, J. (1952). Trypanocidal action of an antibiotic produced by a *Streptomyces*. *C. R. Hebd. Seances Acad. Sci.* 243, 1498–1499.
- D'Huys, P. J., Lule, I., Verammen, D., Anne, J., Van Impe, J. F., and Bernaerts, K. (2012). Genome-scale metabolic flux analysis of *Streptomyces lividans* growing

## ACKNOWLEDGMENTS

We would like to thank Giada Caredda and Maria Vurchio (Institute of Biomedical Technologies, National Research Council, Milan) for technical and administrative support. We thank Dr. B. E. Alber (Department of Microbiology, Ohio State University) for providing us with the technical suggestions for CCR purification.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2017.00835/full#supplementary-material>

- on a complex medium. *J. Biotechnol.* 161, 1–13. doi: 10.1016/j.jbiotec.2012.04.010
- Erb, T. J., Berg, I. A., Brecht, V., Müller, M., Fuchs, G., and Alber, B. E. (2007). Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway. *Proc. Natl. Acad. Sci. U.S.A.* 104, 10631–10636. doi: 10.1073/pnas.0702791104
- Fernandez-Abalos, J. M., Reviejo, V., Diaz, M., Rodriguez, S., Leal, F., and Santamaria, R. I. (2003). Posttranslational processing of the xylanase Xys1L from *Streptomyces halstedii* JM8 is carried out by secreted serine proteases. *Microbiology* 149, 1623–1632. doi: 10.1099/mic.0.26113-0
- Fondi, M., Maida, I., Perrin, E., Mellera, A., Mocali, S., Parrilli, E., et al. (2015). Genome-scale metabolic reconstruction and constraint-based modelling of the antarctic bacterium *Pseudoalteromonas haloplanktis* TAC125. *Environ. Microbiol.* 17, 751–766. doi: 10.1111/1462-2920.12513
- Gevorgyan, A., Bushell, M. E., Avignone-Rossa, C., and Kierzek, A. M. (2011). SurreyFBA: a command line tool and graphics user interface for constraint-based modeling of genome-scale metabolic reaction networks. *Bioinformatics* 27, 433–434. doi: 10.1093/bioinformatics/btq679
- Heinken, A., Sahoo, S., Fleming, R. M., and Thiele, I. (2013). Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. *Gut Microbes* 4, 28–40. doi: 10.4161/gmic.22370
- Henry, C. S., DeJongh, M., Best, A. A., Frybarger, P. M., Lindsay, B., and Stevens, R. L. (2010). High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol.* 28, 977–982. doi: 10.1038/nbt.1672
- Huang, D., Li, S., Xia, M., Wen, J., and Jia, X. (2013). Genome-scale metabolic network guided engineering of *Streptomyces tsukubaensis* for FK506 production improvement. *Microb. Cell Fact.* 12:52. doi: 10.1186/1475-2859-12-52
- Huang, D., Wen, J., Wang, G., Yu, G., Jia, X., and Chen, Y. (2012). In silico aided metabolic engineering of *Streptomyces roseosporus* for daptomycin yield improvement. *Appl. Microbiol. Biotechnol.* 94, 637–649. doi: 10.1007/s00253-011-3773-6
- Hyduke, D., Hyduke, D., Schellenberger, J., Que, R., Fleming, R., Thiele, I., et al. (2011). COBRA Toolbox 2.0. *Protoc. Exch.* 1–35. doi: 10.1038/protex.2011.234
- Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30.
- Karray, F., Darbon, E., Oestreicher, N., Dominguez, H., Tiphile, K., Gagnat, J., et al. (2007). Organization of the biosynthetic gene cluster for the macrolide antibiotic spiramycin in *Streptomyces ambofaciens*. *Microbiology* 153, 4111–4122. doi: 10.1099/mic.0.2007/009746-0
- Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., and Hopwood, D. A. (2000). *Practical Streptomyces Genetics*. Norwich: The John Innes Foundation, doi: 10.4016/28481.01
- Kim, M., Yi, J. S., Lakshmanan, M., Lee, D. Y., and Kim, B. G. (2016). Transcriptomics-based strain optimization tool for designing secondary metabolite overproducing strains of *Streptomyces coelicolor*. *Biotechnol. Bioeng.* 113, 651–660. doi: 10.1002/bit.25830

- King, Z. A., Lloyd, C. J., Feist, A. M., and Palsson, B. O. (2015). Next-generation genome-scale models for metabolic engineering. *Curr. Opin. Biotechnol.* 35, 23–29. doi: 10.1016/j.copbio.2014.12.016
- Kuhstoss, S., Huber, M., Turner, J. R., Paschal, J. W., and Rao, R. N. (1996). Production of a novel polyketide through the construction of a hybrid polyketide synthase. *Gene* 183, 231–236. doi: 10.1016/S0378-1119(96)00565-3
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923
- Lauerti, L., Song, L., Huang, S., Corre, C., Leblond, P., Challis, G. L., et al. (2011). Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in *Streptomyces ambifaciens*. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6258–6263. doi: 10.1073/pnas.1019077108
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Licona-Cassani, C., Marcellin, E., Quek, L. E., Jacob, S., and Nielsen, L. K. (2012). Reconstruction of the *Saccharopolyspora erythraea* genome-scale model and its use for enhancing erythromycin production. *Antonie van Leeuwenhoek* 102, 493–502. doi: 10.1007/s10482-012-9783-2
- Liu, H., and Reynolds, K. A. (1999). Role of crotonyl coenzyme A reductase in determining the ratio of polyketides monensin A and monensin B produced by *Streptomyces cinnamonensis*. *J. Bacteriol.* 181, 6806–6813.
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262
- Lounès, A., Lebrühi, A., Benslimane, C., Lefebvre, G., and Germain, P. (1995). Glycerol effect on spiramycin production and valine catabolism in *Streptomyces ambifaciens*. *Curr. Microbiol.* 31, 304–311.
- Lounès, A., Lebrühi, A., Benslimane, C., Lefebvre, G., and Germain, P. (1996a). Effect of nitrogen/carbon ratio on the specific production rate of spiramycin by *Streptomyces ambifaciens*. *Process Biochem.* 31, 13–20.
- Lounès, A., Lebrühi, A., Benslimane, C., Lefebvre, G., and Germain, P. (1996b). Regulation of spiramycin synthesis in *Streptomyces ambifaciens*: effects of glucose and inorganic phosphate. *Appl. Microbiol. Biotechnol.* 45, 204–211.
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8
- Luo, Y., Zhang, L., Barton, K. W., and Zhao, H. (2015). Systematic identification of a panel of strong constitutive promoters from *Streptomyces albus*. *ACS Synth. Biol.* 4, 1001–1010. doi: 10.1021/acssynbio.5b00016
- Machado, D., and Herrgard, M. (2014). Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism. *PLoS Comput. Biol.* 10:e1003580. doi: 10.1371/journal.pcbi.1003580
- Marler, R. T., and Arora, J. S. (2004). Survey of multi-objective optimization methods for engineering. *Struct. Multidiscip. Optim.* 26, 369–395.
- Nguyen, H. C., Karray, F., Lautru, S., Gagnat, J., Lebrühi, A., Huynh, T. D., et al. (2010). Glycosylation steps during spiramycin biosynthesis in *Streptomyces ambifaciens*: involvement of three glycosyltransferases and their interplay with two auxiliary proteins. *Antimicrob. Agents Chemother.* 54, 2830–2839. doi: 10.1128/AAC.01602-09
- Nueda, M. J., Tarazona, S., and Conesa, A. (2014). Next maSigPro: updating maSigPro bioconductor package for RNA-seq time series. *Bioinformatics* 30, 2598–2602. doi: 10.1093/bioinformatics/btu333
- Olsen, G. J., Lane, D. J., Giovannoni, S. J., Pace, N. R., and Stahl, D. A. (1986). Microbial ecology and evolution: a ribosomal RNA approach. *Annu. Rev. Microbiol.* 40, 337–365. doi: 10.1146/annurev.mi.40.100186.002005
- Onaka, H., Taniguchi, S., Ikeda, H., Igarashi, Y., and Furumai, T. (2003). pTOYAMAcos, pTYM18, and pTYM19, actinomycete-*Escherichia coli* integrating vectors for heterologous gene expression. *J. Antibiot.* 56, 950–956.
- Orth, J. D., Thiele, I., and Palsson, B. O. (2010). What is flux balance analysis? *Nat. Biotechnol.* 28, 245–248. doi: 10.1038/nbt.1614
- Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., et al. (2014). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* 42, D206–D214. doi: 10.1093/nar/gkt1226
- Pinnert-Sindico, S. (1954). A new species of *Streptomyces* producing antibiotics *Streptomyces ambifaciens* n. sp., cultural characteristics. *Ann. Inst. Pasteur* 87, 702–707.
- Poulet, P. P., Duffaut, D., Barhet, P., and Brumpt, I. (2005). Concentrations and in vivo antibacterial activity of spiramycin and metronidazole in patients with periodontitis treated with high-dose metronidazole and the spiramycin/metronidazole combination. *J. Antimicrob. Chemother.* 55, 347–351. doi: 10.1093/jac/dki013
- Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033
- R Development Core Team (2012). *R: A Language and Environment for Statistical Computing*. Available at: <http://www.r-project.org/>
- Richardson, M. A., Kuhstoss, S., Huber, M. L., Ford, L., Godfrey, O., Turner, J. R., et al. (1990). Cloning of spiramycin biosynthetic genes and their use in constructing *Streptomyces ambifaciens* mutants defective in spiramycin biosynthesis. *J. Bacteriol.* 172, 3790–3798.
- Saier, M. H., Reddy, V. S., Tamang, D. G., and Västermark, Å. (2014). The transporter classification database. *Nucleic Acids Res.* 42, D251–D258. doi: 10.1093/nar/gkt1097
- Sambrook, J., and Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*, 3rd Edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, doi: 10.1002/humu.1186.abs
- Scheer, M., Grote, A., Chang, A., Schomburg, I., Munaretto, C., Rother, M., et al. (2011). BRENDA, the enzyme information system in 2011. *Nucleic Acids Res.* 39, D670–D676. doi: 10.1093/nar/gkq1089
- Schellenberger, J., Que, R., Fleming, R. M. T., Thiele, I., Orth, J. D., Feist, A. M., et al. (2011). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat. Protoc.* 6, 1290–1307. doi: 10.1038/nprot.2011.308
- Schuetz, R., Zamboni, N., Zampieri, M., Heinemann, M., and Sauer, U. (2012). Multidimensional optimality of microbial metabolism. *Science* 336, 601–604. doi: 10.1126/science.1216882
- Shlomi, T., Cabili, M. N., Herrgard, M. J., Palsson, B. O., and Ruppian, E. (2008). Network-based prediction of human tissue-specific metabolism. *Nat. Biotechnol.* 26, 1003–1010. doi: 10.1038/nbt.1487
- Sigurðsson, G., Fleming, R. M. T., Heinken, A., and Thiele, I. (2012). A systems biology approach to drug targets in *Pseudomonas aeruginosa* biofilm. *PLoS ONE* 7:e34337. doi: 10.1371/journal.pone.0034337
- Stolyar, S., Van Dien, S., Hillesland, K. L., Pintel, N., Lie, T. J., Leigh, J. A., et al. (2007). Metabolic modeling of a mutualistic microbial community. *Mol. Syst. Biol.* 3, 92. doi: 10.1038/msb4100131
- Thibessard, A., Haas, D., Gerbaud, C., Aigle, B., Lautru, S., Pernodet, J. L., et al. (2015). Complete genome sequence of *Streptomyces ambifaciens* ATCC 23877, the spiramycin producer. *J. Biotechnol.* 214, 117–118. doi: 10.1016/j.jbiotec.2015.09.020
- Thiele, I., and Palsson, B. O. (2010). A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* 5, 93–121. doi: 10.1038/nprot.2009.203
- Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Brucoleri, R., et al. (2015). antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* 43, W237–W243. doi: 10.1093/nar/gkv437
- Wilson, M. C., and Moore, B. S. (2012). Beyond ethylmalonyl-CoA: the functional role of crotonyl-CoA carboxylase/reductase homologs in expanding polyketide diversity. *Nat. Prod. Rep.* 29, 72–86. doi: 10.1039/c1np00082a
- Zhang, C., and Hua, Q. (2016). Applications of genome-scale metabolic models in biotechnology and systems medicine. *Front. Physiol.* 6:413. doi: 10.3389/fphys.2015.00413

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Fondi, Pinatel, Talà, Damiano, Consolandi, Mattorre, Fico, Testini, De Benedetto, Siculella, De Bellis, Alifano and Peano. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.